Proceedings of

The 10th International Conference on Frontier Computing (FC 2020)

July 13 – 19, 2020

Singapore

The Institution of Engineering and Technology

The 10th International Conference on Frontier Computing

FC 2020
Proceeding

Singapore
July 13 - 16, 2020
Alternative Presentation September 7 - 9,2020

Organized by

Frontier Computing Conference Group

Message from Organizing Committees

The International Conference on Frontier Computing – Theory, Technologies, and Applications (FC) was first proposed in early 2010 on an IET executive meeting. This conference series aims at providing an open forum to reach a comprehensive understanding to the recent advances and emergence in information technology, science, and engineering, with the themes in the scope of Communication Technology and Applications, Business Intelligence and Knowledge Management, Artificial Intelligence, and any related fields that prompt the development of information technology. This will be the ninth event of the series, in which fruitful results can be found in the digital library or conference proceedings of FC 2010 (Taichung, Taiwan), FC 2012 (Xining, China), FC 2013 (Gwangju, Korea), FC2015 (Bangkok, Thailand), FC2016 (Tokyo, Japan), and FC2017 (Osaka, Japan), FC2018 (Kuala Lumpur, Malaysia), FC ABH2019 (Taichung, Taiwan), FC2019 (Kitakyushu Japan). Each event brings together the researchers worldwide to have excited and fruitful discussions as well as the future collaborations.

This year FC2020 is the 10th-anniversary event of FC conference series, it was planned to be hold in Singapore, however, due to the CODIV-19 pandemic, our conference has to change to the online form and each presentation is planned as a video stream in the website. The papers accepted for inclusion in the conference proceeding primarily cover the topics of current frontier computing areas. The FC2020 is organized together with The 4th International Conference on Advanced Information Technology with Sensor or Sensor Network, International Conference for Convergence Information Technology & Applications and Fi Award Competition 2020, and there are server special sessions that are cooperated within FC2020. These events present the current developments of frontier computing.

We send our sincere appreciations to the authors for their valuable contributions and the other participants of this conference. The conference would not have been possible without their support. Appreciates are also due to the many experts who contributed to making the event a success. We hope we can have a next FC event, FC2021, in onsite form, and it is the next milestone of FC conference toward to the next future ten years.

FC 2020 Organizing Committees
FC Conference Group
September 2020

Organizing Committees

Honorary Chairs

Han-Chieh Chao, National Dong Hwa University, Taiwan
Hideyuki Takagi, Kyushu University, Japan

Steering Chairs

Jason C. Hung, National Taichung University of Science and Technology, Taiwan
Neil Y. Yen, University of Aizu, Japan

General Chair

Jen-Shiun Chiang, Tamkang University, Taiwan

General Co-Chair

Mario Koeppen, Kyushu Institute of Technology, Japan

Vice General Chair

Qingguo Zhou, Lanzhou University, China

Program Chairs

Yan Pei, University of Aizu, Japan
Jia Wei Chang, National Taichung University of Science and Technology, Taiwan
Hai Jiang, Arkansas State University, USA
Pedro, Peris López, Carlos III University of Madrid, Spain
Zhou Rui, Lanzhou University, China
Dmitry Novikov, Institute of Control Sciences V. A. Trapeznikov, Academy of Sciences, Russia
Daniel Shapiro, Clockrr Inc., Canada
Mahdi Zamani, Yale Univesity, USA
Chao-Tung Yang, Tunghai University, Taiwan

Workshop Chairs

Carmen Camara, Technical University of Madrid, Spain
Shih-Nung Chen, Asia University, Taiwan
Young-Ae Jung, Sun Moon University, Korea
Sujata Pandey, Amity University Uttar Pradesh, India
Jun Shen, University of Wollongong, Australia
Gerald Schaefer, Loughborough University, U.K.
Wei-Chen Wu, Feng Chia University, Taiwan

Chih-Chuan Yeh, Overseas Chinese University, Taiwan

Special Session Chairs

Kuan-Chou Lai, National Taichung University of Education, Taiwan
Jenn-Wei Lin, Fu-Jen University, Taiwan
Xinghua Sun, Hebei North University, China
Chengjiu Yin, Kobe University, Japan
Xiaokang Zhou, Shiga University, Japan
Yishui Zhu, Chang'an University, China

Publicity Chairs

Kei Ohnishi, Kyushu Institute of Technology, Japan
Makoto Fukumoto, Fukuoka Institute of Technology, Japan
Ryohei Funaki, Kyushu University, Japan
Soumya Banerjee, Birla Institute of Technology, India
Jindrich Kodl, Authorised expert in security of information systems, Czech Republic
Min-Feng Lee, National Museum of Natural Science, Taiwan
Poonphon Suesaowaluk, Assumption University of Thailand, Thailand
Shing-Chern You, National Taipei University of Technology, Taiwan
Linjing Wei, Gansu Agricultural University, China
Jun-Hong Shen, Asia University, Taiwan
Ching-Ta Lu, Asia University, Taiwan
Goldina Ghosh, Indian Institute of Information Technology, India

Competition Chairs

Chuan-Feng Chiu, Minghsin University of Science and Technology, Taiwan
Yi-Chun Liao, China University of Technology, Taiwan

Website Chairs

Kuan-Hua Lai, University of Aizu, Japan

Local Arrangement Chair

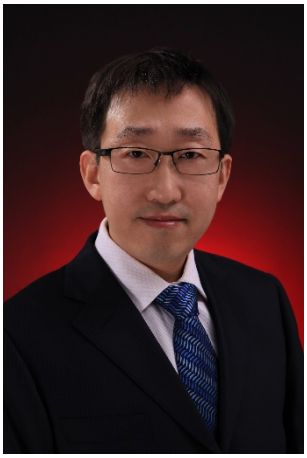
Yung-Hui Chen, LungHwa University of Science and Technology, Taiwan
Yen-Jou Wang, University of Aizu, Japan

International Advisory Board

Jinannong Cao, Hong Kong Polytechnic University, Hong Kong
Su-Ching Chen, University of Florida, USA
Mu-Yen Chen, National Taichung University of Science and Technology, Taiwan
Hamid Fujita, Iwate Prefectural University, Japan

Keynote Speaker
Prof. Chong Fu, Northeastern University, China

Title: Image Encryption Technique Based on Chaos



Abstract

Recent years have seen a fantastic amount of images transmitted over the Internet, raising lots of privacy concerns. Block ciphers (e.g., DES, AES and IDEA), the most commonly used symmetric ciphers, provide a high level of security but can hardly satisfy the growing demand for real-time communications when dealing with image data characterized by large volume. To meet this challenge, a variety of image encryption techniques have been suggested. Among them, the chaos-based technique has been proven to be the most successful. Since 1990s, many researchers have noticed that there exists a close relationship between chaos and cryptography. The intrinsic properties of chaotic dynamical systems such as extreme sensitivity to initial conditions and system parameters, ergodicity and mixing property naturally satisfy the essential design principles of a cryptosystem such as avalanche, confusion and diffusion. The iterative permutation-substitution operations, suggested by Shannon for secure ciphers constructions, are widely adopted in chaos-based image ciphers. In the permutation stage, the pixel positions are scrambled in a secret way, which leads to a great reduction in the correlation among neighboring pixels. In the substitution stage, the pixel values are altered sequentially, and the influence of each pixel is diffused to all its succeeding ones during the modification process. With such a structure, a minor change in one pixel of the plain-image may result in a totally different cipher-image with several overall rounds of encryption. Area-preserving chaotic maps, including the cat map, the baker map, and the standard map, have been widely used in image scrambling. The substitution algorithm consists of three major procedures: 1) generation of a pseudorandom sequence by iterating a chaotic system; 2) extraction of a keystream from the chaotic sequence; 3) mixing plain pixels with the keystream and diffusing the influence of a pixel to its succeeding ones. Theoretical analysis and experimental results have indicated that well-designed chaos-based image encryption algorithms have a high security level, which can effectively resist all common attacks, such as brute force attack, statistical attack and differential attack. In conclusion, the chaos-based image encryption technique has shown to be a promising way for online secure image communication applications.

Biography

Chong Fu received the M.S. degree in telecommunication and information systems and the Ph.D. degree in computer software and theory from Northeastern University, Shenyang, China, in 2001 and 2006, respectively. He joined Northeastern University in 2001, where he is currently a Professor and serves as associate dean at the School of Computer Science and Engineering. In 2010, he spent three months as a Visiting Researcher with the Department of Electronics Information Engineering, The Hong Kong Polytechnic University. His research interests include multimedia security and computer vision.

Table of Content

Page

- **Recognition and Diagnosis of Computed Tomography Images Using Reconstructive Techniques** ····· 1
  Pengzhi Li, Jianqiang Li, Haihua Xie, Yan Pei and Hui Feng
- **Comprehensive Economic Evaluation of Strategic Emerging Industries in China Based on the model of entropy TOPSIS** ····· 12
  Wei Cai
- **Boundary U-Net: A Segmentation Method to Improve Salt Bodies Identification Accuracy** ···· 25
  Yelong Zhao, Bo Liu, Jianqiang Li and Guangzhi Qu
- **Identity Authentication Protocols for Greenhouse Environmental Monitoring System Based on Internet of Things** ····· 37
  Shujie Lu, Xiao Zhang and Ling Li
- **Design of anatomy multimedia E-learning platform** ····· 46
  Mengjiao Liu, Liheng Gong and Xiao Zhang
- **A Study on Learning Effectiness and Satisfaction by Integrating Social Network Analysis into Cooperative Learning -A Case Study of Junior High School Physics and Chemistry in New Taipei City** ····· 54
  Wen-Chih Chang and Jing-Jing Chang
- **Study of the Medical Image Sharing System Based on a RBAC Expansion Model** ····· 63
  Liheng Gong, Mengjiao Liu and Xiao Zhang
- **Quality traceability system for multi-station SMT Manufacturing Process** ····· 72
  Cheng-Hui Chen, De-Wei Hsieh, Ci-Hua Wu, Ci-Yi Lai and Chi-Chin Hsieh
- **An Image Similarity Estimation Approach Based on Weighted Features** ····· 84
  Seok-Woo Jang
- **An Empirical Study on Success Factors of Asian Chinese Companies** ····· 90
  Seung Gyun Yoo
- **Efficient Covering of a Target Object Using a Prediction Technique** ····· 95
  Byeongtae Ahn and Seok-Woo Jang
- **An Empirical Study on Success Factors of Game Industry** ····· 101
  Seung Gyun Yoo
- **Impact of Corruption on Economic Growth: Focusing on Asia of emerging developing countries** ····· 106
  Woo-Sung Cho and Seung Gyun Yoo
- **Keyword Network Analysis on North Korean Children's Rights and Welfare** ····· 111
  Byung-Man Kim
- **Topic Modeling of News articles for Low Birth Patterns in South Korea** ····· 119
  Yoojin Shon and Bosoon Seo
- **Development of Evaluation Scale of Creativity and Personality for College Students Using Delphi Survey Method_Change the title of a paper** ····· 129
  Byung-Man Kim, Jeong-Jin Youn and Jeong Su-Jeong
- **Automatic Pollen Detection Based on Feature Fusion and Self-attention Mechanism** ····· 139
  Quanzeng Wang, Juan Li, Fujiao Ju, Jianqiang Li, Baokai Zu and Caihua Ye
- **A Malicious Web Request Detection Technology Based on Gate Recurrent Unit** ····· 147
  Zhibin Liu, Wenqiang Zhang, Yuanyuan Huang and Qingguo Zhou
- **Using the XGBoost model to predict Santander Customer Trading** ····· 158
  Wen-Chih Chang, Yi-Hong Guo, Ya-Ling Yang, Ming-Chien Hsu, Yi-Hsuan Chu, Ting-Yi Chu and Long-Cheng Meng
- **Fluid Simulation with a Dense Space-time Deformation via L0 Gradient Minimization** ····· 168
  Kun Li, Na Qi and Qing Zhu
- **A Deep Learning Framework for Character Gait Motion Control with Physical Model** ····· 180
  Qing Zhu and Chuanhua Liu

- **A Visualization Analysis of the Paper of Chinese Computational Thinking of China in the Recent Decade** ····· 192
  Haomin Song, Yu Liu, Lihua Ding and Xinghua Sun
- **Painter eMarketplace Platform** ····· 201
  Chatchai Suthapakti, Taminee Shinasharkey and Santithorn Bunchua
- **Acoustic Emission Signal Analysis for Stamping Machine Condition Monitoring and Fault Diagnosis** ····· 213
  Hsiao-Yu Wang, Yu-Hung Chiang, Chih-Yuan Chen and Yu-Shiang Hon
- **Depth Image Based Rendering System by Kinect Sensor** ····· 222
  Beom-Seok Oh, Eungyeol Song and Sunjin Yu
- **LAD Analysis of Policy Reports for Low Birth Patterns** ····· 226
  Mi-Jin Kim and Byung-Man Kim
- **A Semantic Network Analysis of Research Trends on Media Education of Elementary School Students** ····· 236
  Su Jeong Jeong
- **Implementaion for Used Trading Management System Based BlockChain(Case: Used Car)** ··· 246
  Byeongtae Ahn
- **Design of Real Estate Contract Management System based Blockchain** ····· 253
  Byeongtae Ahn
- **A Study for Analysis of Stock Price Information through Extraction of News Articles** ····· 259
  Seok-Woo Jang and Byeongtae Ahn
- **Solid particle dynamics by using Moving Particle semi-implicit Method** ····· 264
  Kyung Sung Kim
- **Research on Remote Control of Self-propelled Lawn Mower** ····· 272
  Wan Rong Chen, Tung-Shou Chen, Jeanne Chen, Fang Rong Hsu and Yu-Hsun Kuo
- **Application model research of visual focus on exhibition item in museum** ····· 280
  Min-Feng, Lee
- **Determinants of Competitiveness in Immersive Technologies with Sensor Networks** ····· 284
  Byun Jeongeun , Hong Jae-pyo, Byun Jeoung-woo
- **Using Computer-based Feedback Help Student Nurses Improving Accuracy of Performing Clinical Operations** ····· 286
  Ying Geng and Po-Sen Huang
- **A Sudy of 2D multi-person Pose Estimation Distance Scaling on Images** ····· 291
  Boney Labinghisa and Dong Myung Lee
- **A Study of Zero-Knowledge Circuit-based for Reduce Transaction Storage Cost** ····· 297
  Soonhyeong Jeong and Byeongtae Ahnbt
- **A Study on the Factors of Performance Improvement in Manufacturing SMEs on Overseas Market** ····· 306
  Seung Gyun Yoo
- **The Design and Implementation of Blockchain-based Supply Chain System with Traceability·** 313
  Chien-Ying Chen, Yu-Wei Chan, Chih-Hung Chang, Tsan-Ching Kang, Chun-Hong Huang and Yin-Te Tsai
- **WiFi Location-based 3D Map for Device Connections** ····· 321
  Chao-Tung Yang, Chen-Kun Tsung, Wei-Chen Chen, Jia-Hao Zhang, Shih-Kuang Chang and Ming-Shang Hsu
- **Using Spark Distributed Deep Learning to Analyze NetFlow in Data Lake System** ····· 327
  Cheng-Tian Jiang, Chao-Tung Yang, Yu-Wei Chan, Endah Kristiani and Jung-Chun Liu
- **Using Long Short-Term Memory Deep Learning for Short-Term PM2.5 Prediction in Taiwan** ····· 333
  Hao Lin, Wen-Yen Lin, Chao-Tung Yang and Jwu-Rong Lin
- **A Container-Based of Edge Device Monitoring on Kubernetes** ····· 340
  Halim Fathoni, Hao-Yi Yen, Chao-Tung Yang, Chin-Yin Huang and Endah Kristiani
- **Flame Recognition System Using YoLo** ····· 346
  Chao-Tung Yang, Wen-Yen Lin, Yi-Chun Chen and Zheng-Yao Wang
- **Comparison of Influenza Disease Prediction Using ARIMA and LSTM models for Central Taiwan** ····· 352
  Kieu Lan Phuong Nguyen, Ho-Wen Chen, Chao-Tung Yang and Endah Kristiani

- **The Design and Implementation of Dynamic Costume Projection System** ····· **361**
  Chuan-Feng Chiu, Han-Yun Hsieh, Wei-Chuan Chung and Shwu-Huey Yen
- **Unidentified PII detection with a k-Nearest Neighbors Approach** ····· **367**
  Tzu-Yin Liao, Yu-Chih Wei, Wei-Chen Wu
- **Artificial Intelligence Identification Model for Chronic Kidney Disease** ····· **376**
  Ya-Fang Cheng, Hsiu-An Lee, Chien-Yeh Hsu
- **Using AI algorithm to Establish the CVD Risk Assessment Model** ····· **387**
  Yin-Chen Chen, Hsiu-An Lee, Chien-Yeh Hsu
- **An Architecture of Real-World Data Database for Real-World Evidence Research**····· **399**
  Hsiu-An Lee, Chien-Yeh Hsu
- **An Innovation Study on Applying Deep Learning to Recognize Gesture in Sign Language** ····· **410**
  Chen-Hu Chou, Yu-Yu Yen, Yu-Chen Su, Horng-Twu Liaw, Wei-Chen Wu
- **Research on Medical Information Exchange Mechanism of Blockchain Combined with Health Passbook** ····· **430**
  Ying-Che Huang, Yu-Yu Yen, Jui-Hung Kao, Horng-Twu Liaw, Wei-Chen Wu
- **Using RoBERTa and Linguistic Features to Detect Fake News**····· **444**
  Jhu-Jyun Huang,Yen-Heng Tsao, Zi-Ying Chen, You-Chuan Yang
- **Developing an Accessibility Assessment System for Mobile Devices** ····· **451**
  Ting-Fang Wu, Chi-Nung Chu, Hui-Shan Lo
- **A Mobile Module Design of Rhythm in Music Based on Practicing Strategy** ····· **457**
  Yu Ting Huang, Chi Nung Chu

Recognition and Diagnosis of Computed Tomography Images Using Reconstructive Techniques

Pengzhi Li[1], Jianqiang Li[1], Haihua Xie[1], Yan Pei[2], and Hui Feng[1]

[1] School of Software Engineering, Beijing University of Technology, Beijing, 100124 China
lipz@emails.bjut.edu.cn, lijianqiang@bjut.edu.cn, heyxhh@emails.bjut.edu.cn, HuiFeng@emails.bjut.edu.cn
[2] Computer Science Division, University of Aizu, Aizuwakamatsu, 965-8580 Japan
peiyan@u-aizu.ac.jp ✉
https://www.u-aizu.ac.jp/~peiyan/

Abstract. In this paper, we propose a diagnosis and classification method of colonography CT Images based on deep learning and image reconstruction, which is processed to increase the difference between pathological features and other healthy tissues and improve the accuracy of pathological images classification. It is difficult to accurately distinguish polyps features owing to the variability of appearances, fuzzy boundaries, heterogeneous densities, shapes and sizes of lesions. Some study results show that the ResNet network has a better performance than other methods and has broad application prospects in the identification of CT images. We used the improved ResNet network as the classification model. We completed the classification experiment on the colonography datasets. Through comparative experiments, we found gradient features play an important role in the classification of colonography CT images. More unexpectedly, for CT images of colonography, the classification effect of CT images with small information entropy is excellent. A new "RGB" image containing two channels of gradient features and one channel of LBP features is very effective for classification, which may be a new attempt.

Keywords: Colonography classification, Deep learning, Image reconstruction, CT image

1 Introduction

Computed tomography (CT) images classification is important in various biomedical applications [2], such as colonography polyps classification. Recently, several computer-aided diagnosis models based on deep learning [12] strategies have been developed for pathological image processing [4], which achieve the automatic extraction of features and improve the accuracy of classification. Many experimental results show that the method of deep neural network has better performance than other methods.

However, it is difficult to accurately distinguish colonography polyp features owing to the variability of appearances, fuzzy boundaries, heterogeneous densities, shapes and sizes of lesions [18]. Therefore, it is necessary to improve the sensitivity and accuracy of automatic classification. And we need to explore new features that can improve the effect of classification.

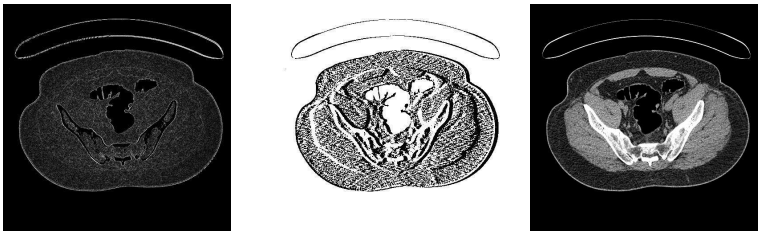
In this paper, a classification method of colonography CT Images based on deep learning and image reconstruction processing is proposed to increase the difference between pathological features and other healthy tissues and improve the accuracy of pathological images identification. And through several groups of comparative experiments, the degree of influence of the characteristics used in the experiment was obtained. First, We extracted the gradient features [1] and local binary pattern (LBP) features [9] of CT images through experiments. We use the method of image reconstruction to generate new RGB images. In other words, the reconstructed image is a newly generated image containing three channels composed of gradient features, grey features and local binary pattern (LBP) spaces. To form a comparative experiment, we performed various combinations of the three features and obtained six sets of the image datasets to do the comparative experiments. Secondly, We used the six sets of datasets to train improved ResNet-34 [6] network to obtain results of classification, as discussed later in section 2.

The remainder of this paper is organized as follows. In Section 2, we introduced a public TCIA colonography datasets and the image processing work in the experiment. We performed a classification experiment on the colonography polyp images. The results of the experiments are in Section 3. Finally, the result analysis and conclusions are summarized in sections 4 and 5, respectively.

2 Materials and Methods

2.1 Dataset Processing

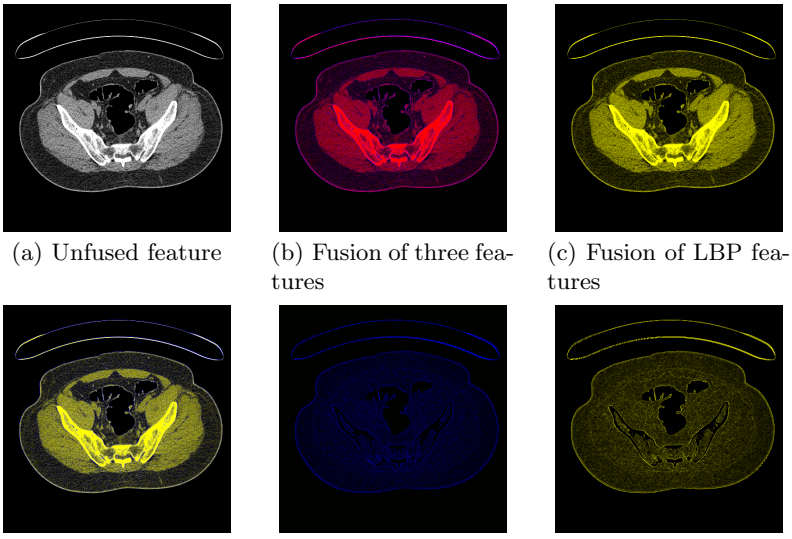
The colonography image datasets used in our experimental experiments are T-CIA. The datasets contain 104 cases where at least one 6mm or larger size polyp was found and 243 cases that were recorded as free of polyps by both CTC and optical techniques. Each 3D CT image consisted of 600 2D slices of 512x512 pixels in size. Three XLS sheets provide polyp descriptions and their location within the colon segments. We linked the XLS polyp tables with the DICOM image studies in TCIA to select image data including polyps and no polyps. Unfortunately, the table may not have the complete number of slice information, so we only selected 170 images that were marked with polyps. We randomly selected 100 images from 170 sheets to use as training datasets. Because the number of images is relatively small, we have performed the methods of image enhancement to increase the number of images, such as random flip [17], random rotation [14], whitening [3]. Finally, the training datasets have 400 positive samples (with Polyp) and 200 negative samples (without polyp). The verification datasets consist of 70 positive samples and 140 negative samples.



(a) Gradient image (b) LBP image (c) Original image

Fig. 1. Extract gradient features and LBP features.

As shown in Fig.1, we extracted the gradient features and local binary pattern (LBP) features from colonography polyps images. For the extracted features, it seems that the images of colonography polyps are clearer than that of colonography.Obviously, gradient images are sensitive to boundaries, and grayscale images contain more information about the original image. The LBP image mainly highlights the texture features of the image. Next, we performed the method of image recombination. In this way, we can get six groups of different datasets of fusion. Fig. 2 shows the six sets of reconstructed images of colonography polyps. We used six groups of datasets to train the improved ResNet-34 model and got the classification results respectively. The details are as follows:



(a) Unfused feature (b) Fusion of three features (c) Fusion of LBP features

(d) Fusion of gradient features (e) Fusion of two LBP features (f) Fusion of two gradient features

Fig. 2. Six groups of features recombined images.

- (a)Unfused feature: Three channels are grayscale images.
- (b)Fusion of three features: The three channels are a grayscale image, gradient image and LBP image.
- (c)Fusion of LBP features: Two grayscale images and one LBP image in three channels
- (d)Fusion of gradient features: Two grayscale images and one gradient image in three channels
- (e)Fusion of two LBP features: Two LBP images and one gradient image in three channels
- (f)Fusion of two gradient features: Two gradient images and one LBP image in three channels

By comparing the experimental results of the group (a) and the group (b), it can be known whether the feature recombination image has a great effect of classification. Groups (c) and (d) can indicate which feature is more important between the gradient features and the LBP features. The purpose of setting the group (f) and the group (e) is to verify whether the original features information is the best features of classification.

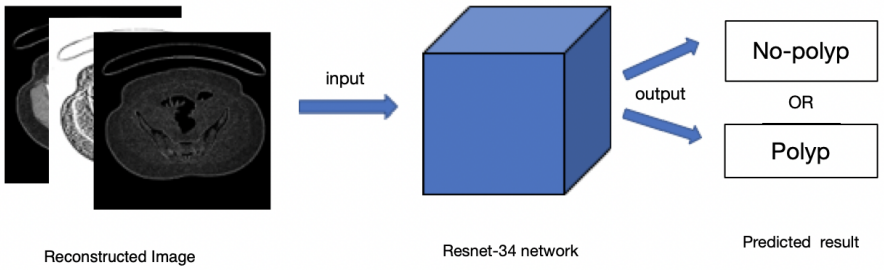


Fig. 3. The reconstructed image is used to train the improved ResNet network model to obtain classification results.

2.2 Improved Training Model

Recently, deep learning strategies are widely applied to the feature extraction of different image classification tasks. Particularly, convolutional neural network (CNN) [5] consists of alternating convolution and subsampling operations. A series of operations were regarded as a process of feature extraction. In the following years, a large number of deep neural network models were proposed, such as AlexNet [11], VGG Net [15], Google net [16], etc, but deeper neural networks are more difficult to train. Some scholars have proposed a residual learning framework to ease the training of networks that are substantially deeper than those used previously. They explicitly reformulate the layers as learning residual functions concerning the layer inputs, instead of learning unreferenced

functions. They provide comprehensive empirical evidence showing that these residual networks are easier to optimize, and can gain accuracy from considerably increased depth.

He et al. proposed some models based on different depths of ResNet network, such as ResNet-18 and ResNet-34 [7]. We conducted some experiments to select the model before the formal experiment. We performed experiments on the original colonography image datasets using VGG-19, Resnet-18 and Resnet-34, respectively. The experimental results show that the Resnet-34 model is the best, so we chose Resnet-34 as the target of an improved model. ResNet-34 is a simple and common deep convolutional neural network [10]. We used the improved ResNet-34 network to complete a classification experiment on colonography CT Images. As is exhibited in the Fig.3, we used six groups of experimental datasets to train and verify the classification results of improved ResNet-34 net respectively. To accurately calculate the image information entropy, we need to retain more image information. So we used the images of the original size to train and verify the model. In general, the size of images inputted into the ResNet model is 224x224, so we made some improvements on ResNet-34 network, to avoid zooming out image [19]. We removed the average pooling layer and full connection layer of the original model, and then we added Relu1 layer, convolution1 layer, Relu2 layer, convolution2 layer, average pooling layer and full connection layer in turn. Six groups of comparative datasets were classified on the improved model respectively.

3 Results and Comparisons

To verify the validity of different feature maps for classification results, four commonly used evaluative criteria are considered as follows in (1)-(4) where ACC means the overall grading accuracy. SEN and SPE are sensitive and specificity for each class. Three different measurements were obtained, all of them in terms of the number of true positives TP, true negatives TN, false positives FP and false negatives FN, and F1 is the weighted harmonic average of precision and recall. As shown in Table 2, we have calculated the results of the four indicators for the six sets of data.

$$ACC = \frac{TP + TN}{TP + TN + FP + FN} \tag{1}$$

$$SEN = \frac{TP}{TP + FN} \tag{2}$$

$$SPE = \frac{TN}{TN + FP} \tag{3}$$

$$F1 = \frac{2 * PR * RE}{PR + RE} \tag{4}$$

From the Table 3, we want to understand the effect of the size of the information entropy [8] on the classification results, so we calculate the information

Table 1. Classification results of six groups of features recombined images. The indicator results show the mean and confidence interval.

Image type	F1	ACC	SEN	SPE
Group(a)	0.88499 (0.86700,0.90299)	0.89600 (0.87311,0.91888)	0.84433 (0.82463,0.86403)	0.95033 (0.92430,0.97636)
Group(b)	0.90400 (90.89587,0.91212)	0.914 (0.90085,0.92714)	0.86366 (0.84978,0.87755)	0.96633 (0.95640,0.97626)
Group(c)	0.89233 (0.87753,0.90712)	0.90766 (0.89319,0.92213)	0.85099 (0.83238,0.86961)	0.958 (0.94314,0.97285)
Group(d)	0.89866 (0.88171,0.91561)	0.91166 (0.89637,0.92695)	0.854 (0.83543,0.87256)	0.96333 (0.95493,0.97173)
Group(e)	0.88733 (0.87862,0.89604)	0.89466 (0.87993,0.90939)	0.846 (0.82414,0.86785)	0.95200 (0.94686,0.95713)
Group(f)	0.90600 (0.89937,0.91262)	0.91300 (0.90270,0.92329)	0.86766 (0.84894,0.88639)	0.96566 (0.95889,0.97243)

entropy of the corresponding feature map. The entropy of an image is a statistical form of a feature that reflects the amount of average information in an image and represents the aggregated features of the grayscale distribution of the image. The image entropy is expressed as the bit average of the set of grey levels of the image, unit bits/pixel, which also describes the average amount of information of the image source. For a two-dimensional image of the discrete form, the calculation method of one-dimensional information entropy is a formula (5) [13], and Pi represents the proportion of pixels with a grey value i in the image.

$$H = \sum_{i=0}^{255} pi \log pi \tag{5}$$

Wilcoxon signed-rank test is a nonparametric test. Under the condition that the two sample spaces do not need to assume a normal distribution, test whether their distribution is the same. If p-value <0.05, there is a significant gap between the two groups of data. To verify that the two sets of data are significantly different, we have introduced this test. We divided the six sets of data into three groups and then performed the Wilcoxon signed-rank test. We used (a) and (b) as contrast group 1, (c) and (d) as contrast group 2, (e) and (f) as contrast group 3. As shown in Table 3, we calculated the P-values of the indicators for the three groups of contrast experiments.

Table 2. Information entropy calculation results of different recombination graphs.

Image type	Image information entropy
(a)Unfused feature	4.0516
(b)Fusion of three features	3.5966
(c)Fusion of LBP features	3.9827
(d)Fusion of gradient features	4.1069
(e)Fusion of two LBP features	3.4973
(f)Fusion of two gradient features	3.6727

Table 3. P-value of three contrast groups

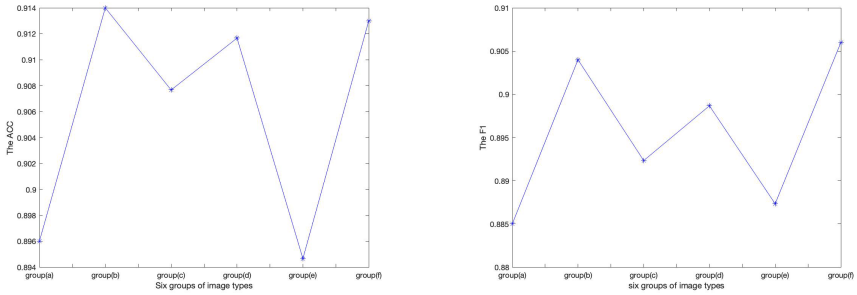
contrast group	F1	ACC	SEN	SPE
P of group 1(a,b)	0.00258	0.00633	0.01058	0.12943
P of group 2(c,d)	0.08222	0.37551	0.24086	0.59750
P of group 3(e,f)	1.931e-05	1.811e-05	0.002607	0.00437

4 Result Analysis

Based on the classification results in Section 3 (Table 1 and Table 3), we have obtained several important feature details as follows:

- Comparison group 1: By observing the statistics of the group (a) and group (b), it can be known that the indicators of the group (b) are higher than group (a). And the p-value of the indicator is less than 0.05, which indicates that there is a significant difference between the two groups of data under the large sample. Therefore, the results show that the fusion of multiple feature methods has a good effect and improves the classification effect.
- Comparison group 2: By observing the indicators of the group (c) and group (d), although the P-value is not greater than 0.05, it can be seen that the indicators of group d are higher than group (c). Both groups are higher than the group (a). The experimental results show that the addition of gradient features is better than the addition of LBP features.
- Comparison group 3: The numerical information in the table clearly shows a phenomenon that the indicators of the group (f) are higher than that of the group (e). The indicators of the group (e) and group (b) are almost the same, and some indicators are even higher than the group (b). And the P-value is less than 0.05. The results show that the data of the two groups are significantly different under the large sample, and the gradient characteristics are better than the LBP characteristics.

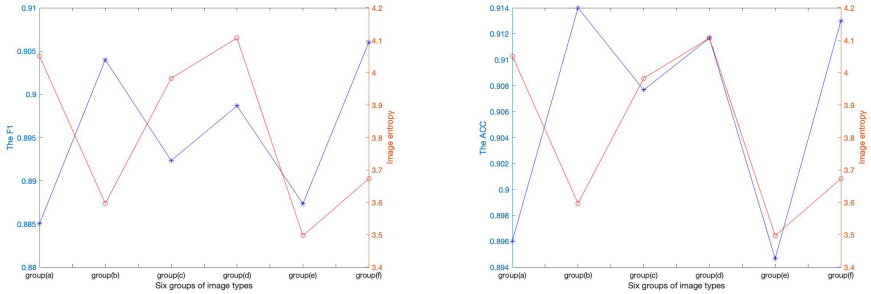
Next, to more intuitively observe the classification effect of the six groups, we selected F1 and Acc from the six indicators to draw a line chart, as shown in



(a) ACC line chart (b) F1 line chart

Fig. 4. (a) is a straight line drawn by ACC of six groups of data. (b) is a straight line drawn by F1 of six groups of data.

Fig.4. From Fig. 4. (a), we can see that the accuracy of group (b)(group (a)) is 91.4% (89.6%), and that of group (f) (group (e)) is 91.3% (89.46%). Group (f)'s accuracy rate is 0.1% higher than Group(b). The accuracy index of the group (b) and group (f) is much higher than that of the group (a), and this trend is obvious. From the Fig. 4. (b), we can know that the changing trend of F1 is similar to that of ACC. Therefore, when we add gradient features, the accuracy of colorectal polyps can improve significantly. We can also see that the F1 with the gradient feature image is high. This also indicates that the gradient feature is an important feature of CT images of colorectal polyps. Gradient features play an important role in improving the classification.



(a) F1 and Information entropy (b) ACC and Information entropy

Fig. 5. (a) is a straight-line drawing of F1 and image entropy of six groups of datasets. (b) is a straight-line drawing of ACC and image entropy of six groups of data.

Next, we analyze the information entropy of the image. From Table 1 and Table 2, we can know the effect of the image information entropy size on the

classification result of the image. As shown in Fig.5, to better see the change of information entropy and classification effect, we selected representative F1 and ACC from the four indicators and then drew a curve graph.

The Fig.5 provides some interesting data regarding the trend of information entropy. From these pictures, we can see that the tendency of information entropy and F1 is inversely proportional. For datasets with large information entropy, F1 obtained by classification experiment is small, such as group(a) and group(b). In other words, images with smaller information entropy will get better classification results. Besides, we can also see that the F1 with the gradient feature image is very high. This also shows that gradient feature is an important feature for CT images of colorectal polyps. The gradient feature plays an important role in the improvement of classification effect.

5 Conclusion

In this paper, we used multi-space image reconstruction method and improved ResNet network model to classify colonography CT images. The image reconstruction mapping method can convert the original CT image into gradient, grayscale and local binary mode (LBP) space. Then, we use three kinds of feature images to perform multiple combinations to generate six sets of data sets. We use the reconstructed images to train and verify Improved ResNet model respectively to obtain six groups of classified data. We used this method to perform classification experiments on the colonography datasets. By comparing and analyzing the results of the experiment, we found that the method of multi-feature fusion is effective. The gradient features play an important role. Besides, A new "RGB" image containing two channel gradients and one channel LBP is very effective for classification. By calculating the information entropy of images, we found that the information entropy and the classification effect show opposite trends of change. The original feature of CT image is not ideal classification feature, the feature of feature recombination is better. In future work, we can try this new "RGB" image for classification experiments on various CT images.

Acknowledgment

This work was supported in part by the Beijing Municipal Science and Technology Project (No.KM201910005028).

References

1. Bouadjenek, N., Nemmour, H., Chibani, Y.: Age, gender and handedness prediction from handwriting using gradient features. 13th International Conference on Document Analysis and Recognition, ICDAR 2015, Nancy, France, August 23-26, 2015 pp. 1116–1120 (2015)

2. Chen, S., Yang, H., Fu, J., Mei, W., Ren, S., Liu, Y., Zhu, Z., Liu, L., Li, H., Chen, H.: U-net plus: Deep semantic segmentation for esophagus and esophageal cancer in computed tomography images. IEEE Access **7**, 82,867–82,877 (2019)
3. Deng, S., Zhang, J., Li, P., Huang, G.: Edge detection from polarimetric SAR images using polarimetric whitening filter. 2011 IEEE International Geoscience and Remote Sensing Symposium, IGARSS 2011, Vancouver, BC, Canada, July 24-29, 2011 pp. 448–451 (2011)
4. Gadermayr, M., Wimmer, G., Uhl, A., Kogler, H., Vécsei, A., Merhof, D.: Fully-automated cnn-based computer aided celiac disease diagnosis. Image Analysis and Processing - ICIAP 2017 - 19th International Conference, Catania, Italy, September 11-15, 2017, Proceedings, Part II pp. 467–478 (2017)
5. Gao, X.W., Hui, R., Tian, Z.: Classification of CT brain images based on deep learning networks. Computer Methods and Programs in Biomedicine **138**, 49–56 (2017)
6. He, K., Zhang, X., Ren, S., Sun, J.: Deep residual learning for image recognition. 2016 IEEE Conference on Computer Vision and Pattern Recognition, CVPR 2016, Las Vegas, NV, USA, June 27-30, 2016 pp. 770–778 (2016)
7. He, K., Zhang, X., Ren, S., Sun, J.: Deep residual learning for image recognition. 2016 IEEE Conference on Computer Vision and Pattern Recognition, CVPR 2016, Las Vegas, NV, USA, June 27-30, 2016 pp. 770–778 (2016)
8. Huo, W., Huang, Y., Pei, J., Zhang, Q., Gu, Q., Yang, J.: Ship detection from ocean SAR image based on local contrast variance weighted information entropy. Sensors **18**(4), 1196 (2018)
9. Jabri, S., Saidallah, M., el Belrhiti el Alaoui, A., Fergougui, A.E.: Moving vehicle detection using haar-like, LBP and a machine learning adaboost algorithm. IEEE International Conference on Image Processing, Applications and Systems, IPAS 2018, Sophia Antipolis, France, December 12-14, 2018 pp. 121–124 (2018)
10. Kemaev, I., Polykovskiy, D., Vetrov, D.P.: Reset: Learning recurrent dynamic routing in resnet-like neural networks. Proceedings of The 10th Asian Conference on Machine Learning, ACML 2018, Beijing, China, November 14-16, 2018. pp. 422–437 (2018)
11. Krizhevsky, A., Sutskever, I., Hinton, G.E.: Imagenet classification with deep convolutional neural networks. Advances in Neural Information Processing Systems 25: 26th Annual Conference on Neural Information Processing Systems 2012. Proceedings of a meeting held December 3-6, 2012, Lake Tahoe, Nevada, United States. pp. 1106–1114 (2012)
12. Liu, X., Hou, F., Qin, H., Hao, A.: Multi-view multi-scale cnns for lung nodule type classification from CT images. Pattern Recognition **77**, 262–275 (2018)
13. Luo, X., Wu, X., Zhang, Z.: Regional and entropy component analysis based remote sensing images fusion. Journal of Intelligent and Fuzzy Systems **26**(3), 1279–1287 (2014)
14. Ma, B., Liu, Z., Jiang, F., Yan, Y., Yuan, J., Bu, S.: Vehicle detection in aerial images using rotation-invariant cascaded forest. IEEE Access **7**, 59,613–59,623 (2019)
15. Simonyan, K., Zisserman, A.: Very deep convolutional networks for large-scale image recognition. 3rd International Conference on Learning Representations, ICLR 2015, San Diego, CA, USA, May 7-9, 2015, Conference Track Proceedings (2015)
16. Szegedy, C., Liu, W., Jia, Y., Sermanet, P., Reed, S.E., Anguelov, D., Erhan, D., Vanhoucke, V., Rabinovich, A.: Going deeper with convolutions. IEEE Conference on Computer Vision and Pattern Recognition, CVPR 2015, Boston, MA, USA, June 7-12, 2015 pp. 1–9 (2015)

17. Varma, P., Iter, D., Sa, C.D., Ré, C.: Flipper: A systematic approach to debugging training sets. Proceedings of the 2nd Workshop on Human-In-the-Loop Data Analytics, HILDA@SIGMOD 2017, Chicago, IL, USA, May 14, 2017 pp. 5:1–5:5 (2017)
18. Yang, X., Ye, X., Slabaugh, G.G.: Multilabel region classification and semantic linking for colon segmentation in CT colonography. IEEE Trans. Biomed. Engineering **62**(3), 948–959 (2015)
19. Zhong, X., Feng, G., Wang, J., Wang, W., Si, W.: A novel adaptive image zooming scheme via weighted least-squares estimation. Frontiers Comput. Sci. **9**(5), 703–712 (2015)

Comprehensive Economic Evaluation of Strategic Emerging Industries in China Based on the model of entropy TOPSIS

Cai Wei

(School of Economics and Management ,ECUT, NanChang, 330013)

Abstract: This paper, by using the entropy TOPSIS model, establishes a comprehensive evaluation system of Chinese Strategic Emerging Industry. Taking 49 industries in 2016 as samples, this paper analies the problem as follows. Competitiveness of these provinces should be divided into four levels by advance level, above-average level, average level and below-average level. As the proportion of capital to regional finance is similar, each industry draws a different result. But obvious differences of output and the affect to the industry economic evaluation. At last, relevant policy recommendations are presented to improve the strategic emerging industries and promote balanced and efficient development.

Key words: Strategic Emerging Industry; New information technology industry; Model of TOPSIS; Comprehensive Economic Evaluation

1. Introduction

Since the strategic emerging industries were first put forward at the conference of science and technology in 2009, the State Council has issued a series of policies and plans to promote development. In *the 12th Five Year Plan* and *the 13th five year plan*, the Party Central Committee has established the development of strategic emerging industries as a strategy to invigorate and strengthen the country, and has successively issued the policy on accelerating the cultivation and development of strategic emerging industries The decision of the industry, the implementation opinions on encouraging and guiding private enterprises to develop strategic emerging industries, the Guidance Opinions on promoting the international development of strategic emerging industries, the national strategic emerging industry development plan of the *12th Five Year Plan* and other macro level strategic emerging industry policy documents. And special industrial plans, such as the opinions of the State Council on accelerating the development of energy conservation and environmental protection industry, the development plan of energy conservation and new energy automobile industry (2012-2020), the *12th Five Year Plan* of new material industry, and the development plan of biological industry.

The development of strategic emerging industries is a major strategic deployment made by

the central government of China based on the development environment at home and abroad, in line with the economic development situation, and seeking to seize the commanding heights of international economic and technological competition. Therefore, the healthy development and effectiveness of strategic emerging industries are directly related to whether China can smoothly change the mode of economic growth, realize the optimization and upgrading of industrial structure, and whether China's economy can take the initiative in the process of globalization. However, with China's large investment in strategic emerging industries, what is the economic efficiency of production of strategic emerging industries in China? What is the effect of relevant factors on economic efficiency? What are the ways and Countermeasures to improve the production and economic efficiency of China's strategic emerging industries? Wait. The correct answer to these questions is a theoretical problem that needs to be solved urgently for the successful implementation of this strategic deployment.

In the development process of strategic emerging industries, how to get rid of homogeneous competition, bad competition, strong dependence on subsidies and other issues have attracted the attention of academia and relevant government departments. What is the economic efficiency of China's strategic emerging industries? Does policy investment bring about efficiency improvement of corresponding industries? In order to answer these questions, we need a mechanism to monitor the economic efficiency of strategic emerging industries. The support from the government plays a huge role in promoting the development of strategic emerging industries, but the test from the market is also indispensable. How to experience the market selection, have industrial competitiveness, survive in the market and further develop? The answers to the above questions are of great practical significance for improving the economic efficiency and economic efficiency of strategic emerging industries and promoting the high-quality development of strategic emerging industries.

As for the research on the economic efficiency of strategic emerging industries, the research at home and abroad mainly focuses on the influencing factors of economic efficiency, efficiency evaluation methods and dynamic monitoring. Foreign scholars' researches on the economic efficiency of strategic industries are mainly focused on the industry standards and characteristics. For example, Hirschman (1966) put forward the concept of strategic industries first. He regarded the strategic industries as the leading industries and called the economic sectors with the largest weight in the input-output relationship as the strategic sectors, i.e. the leading industrial sectors. Krugman (1986) studied the strategic industries This paper introduces the field of international trade protection policy, and puts forward two standards to identify strategic sectors. In recent years, foreign scholars have discussed the formation and characteristics of emerging industries. For example, Potter (1997) believes that the main reasons for the formation of emerging industries are

technological innovation, changes in the relationship between relative costs, and the emergence of new consumer demand. Erickcek and watts (2007) believe that emerging industries are equivalent to structural upgrading. Some scholars have also conducted empirical tests on the influencing factors of the development of emerging industries. Farrell (1957) first conducted a pioneering study on the theory of economic efficiency. Aigner (1977) proposed two measurement methods: non parametric statistical method data envelopment analysis (DEA) and parametric statistical method random forward analysis (SFA). Since then, economic efficiency monitoring has become a research hotspot of foreign scholars. These two kinds of monitoring methods are popular and have a great view. There are many literatures on economic efficiency of related industries. Representative researches such as Elena B (2006) used SFA method to analyze data using financial database and monitored economic efficiency of European banking industry with different financial structures. Eric Wang (2007) used stochastic frontier analysis method to research and develop in economic growth Sharma, et al. (2008) used DEA method to measure the international R&D economic efficiency, and verified the effect of patent number, R&D personnel full-time equivalent and other indicators on R&D use efficiency. However, it is a pity that foreign scholars have not introduced the study of industrial economic efficiency into the field of strategic emerging industries, and also lack the empirical test of the factors affecting the economic efficiency of strategic emerging industries.

In recent years, relevant research in China has been gradually active. Li Qiang (2010) used DEA method to monitor the economic efficiency of small and medium-sized listed companies in Shenzhen Stock Exchange. Yao Yang and Zhang Qi (2001) monitored the technical efficiency of Chinese industrial enterprises. Qifeng and Cheng Guisun used DEA (2015) to respectively monitor the innovation efficiency of China's strategic emerging industries and China's private strategic new industries Li Bozhou and Wang Xue (2019) used DEA method to monitor the economic efficiency of the R&D stage and transformation stage of China's strategic emerging industries.

It can be seen that in terms of influencing factors, scholars' research shows that industrial scale, technological innovation, R&D investment, finance, ecology and corporate culture have an important impact on economic efficiency; in terms of efficiency evaluation methods, they are generally divided into two categories: non parametric evaluation and parametric evaluation; non parametric evaluation (DEA) is represented by data envelopment method, and parametric evaluation is based on Stochastic Frontier Analysis (SFA) ） For and on behalf of. In the aspect of dynamic monitoring, establish a dynamic monitoring application platform to follow up the financial indicators of nine industries under the strategic emerging industry in real time, mainly for the electronic information equipment and other industries closely related to the strategic

emerging industry, and the research samples are mostly concentrated before 2012. There is no doubt that the existing research literature provides valuable enlightenment and help for further research on the economic efficiency of strategic emerging industries in China. However, the existing research literature uses high-tech industry data or randomly selected listed or Shenzhen listed companies to replace strategic emerging industries as research objects, which can not fully reflect the situation of strategic emerging industries, and may lead to the economic efficiency and The analysis of influencing factors is not accurate; using data envelopment analysis (DEA) to measure economic efficiency sacrifices the estimation of production function and is difficult to detect the significance of results; using SFA production function to measure economic efficiency also cannot overcome the assumption of technology neutrality and fixed output elasticity, which is easy to bring estimation bias.

Therefore, this paper intends to improve the monitoring of the economic efficiency of strategic emerging industries from three aspects: first, to build a monitoring index system of the economic efficiency of strategic emerging industries, to conduct in-depth discussion on the influencing factors of strategic emerging industries; second, to optimize the input-output analysis of the economic efficiency of strategic emerging industries, and to use TOPSIS method to improve the economic efficiency of China's strategic emerging industries The weight of economic efficiency monitoring indicators is measured, based on which the industrial statistical monitoring is conducted, which is a useful exploration for the existing research; thirdly, in the selection of research samples, 49 industries of China's strategic emerging industries in 2016 are selected as samples, which makes up for the lack of the existing research samples that mostly replace strategic emerging industries with high-tech enterprises and listed companies, and the research is more specific and real-time.

2. Experiment Design

2.1 Entropy Method

In physics, entropy is a statistical measure of information loss; its entropy function form can completely express the content of information system: in information system, information entropy is a measure of the degree of information disorder, and the higher the degree of information entropy disorder, the smaller the information entropy, the lower the degree of information disorder, and the greater the information utility. Therefore, entropy method can deeply reflect the utility value of index information entropy.

Entropy method is based on the principle of difference driven, focusing on the local differences between the indicator values: for the original data$\{x_{ij}\}$, each indicator x_j has different values about the n evaluated objects, that is, there are local differences between the evaluated

objects.

If the data difference of the evaluated object in an index is very small, the weight coefficient of the index will be very small by entropy method, so we have reason to think that the index has little effect on the whole evaluation index system. On the contrary, if there is a large degree of difference among the evaluation objects, the function of the index is relatively large.

(1) To calculate the i evaluated object on the j index,

$$p_{ij} = x_{ij} \Big/ \sum_{i=1}^{n} x_{ij} \qquad (1)$$

(2) To calculate the j index entropy value,

$$e_j = -\frac{1}{\ln n} \sum_{i=1}^{n} p_{ij} \ln(p_{ij}) \qquad (2)$$

$0 \le e_j \le 1$。

(3) To calculate driven difference x_j,

$$g_j = 1 - e_j \qquad (3)$$

(4) To calculate the weight,

$$w_j = g_j \Big/ \sum_{j=1}^{m} g_j \ , \quad j = 1, 2, \cdots, m \qquad (4)$$

Sum of w_j equals 1。

2.2 Using TOPSIS to evaluate

The central idea of TOPSIS is to select a positive ideal point and a negative ideal point, and select the better evaluated object by comparing the distance between the evaluated object and the positive ideal point and the negative ideal point. The Euclidean distance method is usually used to calculate the distance.

Steps as follows:

(1) Dimensionless evaluation index

Positive index $x_{ij}^{*} = \frac{x_{ij} - m_j}{M_j - m_j} (i = 1, 2, \cdots, n;\ j = 1, 2, \cdots, m)$ (5)

Negative index $x_{ij}^{*} = \frac{M_j - x_{ij}}{M_j - m_j} (i = 1, 2, \cdots, n;\ j = 1, 2, \cdots, m)$ (6)

(2) Weight of dimensionless data

Yij is the evaluated object.

$$y_{ij} = x_{ij}^{*} * w_j \qquad (7)$$

(3)To calculate positive ideal point and negative ideal point

$$y_j^{+} = \max_{1 \le i \le n} (y_{ij}) \qquad (8)$$

$$y_j^- = \min_{1 \le i \le n}(y_{ij}) \quad (9)$$

So,we calculate the matrix of positive ideal point and negative ideal point as $y^+ = (y_1^+, y_2^+, \cdots, y_m^+)$, $y^- = (y_1^-, y_2^-, \cdots, y_m^-)$.

(4) To calculate the Euclidean distance

$$d_i^+ = \sqrt{\sum_{j=1}^{m}(y_j^+ - y_{ij})^2} \quad (10)$$

$$d_i^- = \sqrt{\sum_{j=1}^{m}(y_j^- - y_{ij})^2} \quad (11)$$

(5)Order by weight

Calculating c_i ;

$$c_i = d_i^- / (d_i^- + d_i^+), \quad (12)$$

As c_i gets larger, the evaluated object is better, as c_i gets smaller, the evaluated object is worse.

3.Building Index System

3.1 Principles

(1) Principle of completeness

The economic efficiency evaluation of strategic emerging industries should not only adopt one or several indicators, which is not enough to analyze and monitor economic efficiency, but also establish a comprehensive monitoring index system under the condition of ensuring the rationality and completeness of the selection of indicators.

(2) Principle of index availability

The economic efficiency monitoring system should be accurate. Among many indicators, it should be highly representative, with the same statistical caliber and clear research scope. The indicators must be available. The data sources must be followed to facilitate access and analysis. The data mainly comes from the survey of strategic emerging industries by the National Bureau of Statistics, which can be inquired in the Yearbook.

(3) Principle of comparability

In the design of economic efficiency monitoring system, it is composed of absolute index, relative index and average index, which has strong comparability. Comparability is reflected in the comparability of space (horizontal) and time (vertical). Spatial comparability refers to the standardization of index contents and methods adopted when different regions at the same time are taken as research objects, which can comprehensively compare the economic efficiency of each region. Time comparability refers to the comparison of economic efficiency of different time or

period in the same region as the research base period.

(4) Dynamic principle

As the production and sales process of strategic emerging industries has the characteristics of a certain period objectively, taking the investment in fixed assets as an example, the purchase, use and maintenance of fixed assets are completed in a certain period, and the impact on economic efficiency has a certain lag, which determines that the indicators of comprehensive evaluation of economic efficiency should be dynamic.

(5) The principle of combining quantitative analysis with qualitative analysis

Because there are many indicators to measure economic efficiency, in order to make the indicator system comprehensive, the indicator system needs to include quantitative and qualitative contents.

3.2 Index system

This index system is used to measure the level of economic efficiency in each strategic emerging industry, that is, the economic efficiency monitoring index system. Its construction is based on the "input-output" principle: the input of the country or industry to the elements of the economic production process in China, including human, material and financial resources. At the same time, combined with the output, circulation and distribution in the production process, we can explore from it The internal reason for the low economic efficiency of China's strategic emerging industries is that scientific research achievements have not been converted into productivity.

Therefore, there are two secondary indicators in this paper, namely, the input of economic production and the output of economic production.

Table1 Index System

attribute	name	lable	illustration
input	Average number of employees	X_1	Average number of employees of this industry
	Number of Enterprises	X_2	Number of Enterprises in the industry
	Main business cost	X_3	Main business cost in the industry
	Total debt	X_4	Total debt in the industry
	Manage cost	X_5	Manage cost of the industry
	Financial cost	X_6	Financial cost of the industry
	Marketing cost	X_7	Marketing cost of the industry
	Fixed assets	X_8	Fixed assets of the industry
	Current assets	X_9	Current assets of the industry
output	Main business revenue	X_{10}	Main business revenue of the industry

	Operating profit	X_{11}	Operating profit of the industry
	Industrial sales value	X_{12}	Industrial sales value of the industry
	Owners' equity	X_{13}	Owners' equity in the industry

4. Results

Table 2 weight of the 13 indicators

	X_1	X_2	X_3	X_4	X_5	X_6	X_7
e_j	0.9899	0.9894	0.9869	0.9895	0.9889	0.9882	0.9919
g_j	0.0101	0.0106	0.0131	0.0105	0.0111	0.0118	0.0081
w_j	0.1157	0.0512	0.0704	0.0638	0.0752	0.0828	0.0877
	X_8	X_9	X_{10}	X_{11}	X_{12}	X_{13}	
e_j	0.9794	0.9817	0.9892	0.9925	0.9861	0.9881	
g_j	0.0206	0.0183	0.0108	0.0075	0.0139	0.0119	
w_j	0.1302	0.0745	0.0661	0.0477	0.0669	0.0679	

Among the 13 indicators in the economic efficiency monitoring indicator system, X_8 is the fixed assets of strategic emerging industries with the largest weight of 0.1302. Fixed assets are the most important indicator of the economic efficiency monitoring indicator system and the indicator of investment part. It shows that in the period of production investment in strategic emerging industries, the investment of assets is crucial and the industries will have higher economic efficiency and vice versa.

X_{11} refers to the operating profit of a strategic emerging industry, with the minimum weight of 0.0477, which measures the profit ability generated by the operation of a strategic emerging industry. Operating profit is the output part of a strategic emerging industry. Theoretically, if the operating profit value of a strategic emerging industry is higher, the economic efficiency of the industry will be higher; otherwise, the economic efficiency will be lower. On the whole, the weight of the index system is between 0.0477 and 0.1302. The first three indexes are: X8 fixed assets, X_1 average number of employees, X_7 sales expenses, two input indexes and one output index. The last three indexes are: X_4 total debt, X_{12} industrial sales output value, X_{11} operating profit, one input index and two output indexes Therefore, the weight measured by entropy method is scientific and reasonable.

So, we can get the results followed;

Table3 Comprehensive Economic Evaluation of Chinese

Strategic Emerging Industries

Industry	Evaluation value	Order	Industry	Evaluation value	Order
Tram manufacturing	0.9962	1	Paint, ink and pigment manufacturing	0.9149	26
Low speed truck manufacturing	0.9952	2	Special equipment for prevention and control of environmental pollution manufacturing	0.9143	27
Ferrous metal casting	0.9950	3	Glass products manufacturing	0.9067	28
Radar and supporting equipment manufacturing	0.9860	4	General instrument manufacturing	0.8982	29
Trailer manufacturing	0.9857	5	special ceramic products manufacturing	0.8645	30
Clocks and timekeeping instruments manufacturing	0.9816	6	Chemical API manufacturing	0.8623	31
traditional Chinese medicine drinks manufacturing	0.9797	7	Spacecraft equipment manufacturing	0.8621	32
Veterinary medicine manufacturing	0.9795	8	Audio visual equipment manufacturing	0.8350	33
Printing equipment manufacturing	0.9784	9	Wind power generator equipment manufacturing	0.8305	34
Recycling of waste resources and materials	0.9695	10	Optical fiber and optical cable manufacturing	0.7885	35
special instruments manufacturing	0.9653	11	Chinese patent medicine manufacturing	0.7705	36
Smelting of rare earth metals	0.9650	12	Metallurgical special equipment manufacturing	0.6861	37
Optical instrument manufacturing	0.9645	13	Chemicals manufacturing	0.6823	38
Special equipment for electronic and electrical machinery manufacturing	0.9642	14	Special chemicals manufacturing	0.6773	39
Refractory products manufacturing	0.9569	15	Computer manufacturing	0.6748	40
Radio and television equipment manufacturing	0.9546	16	Organic fertilizer and microbial fertilizer manufacturing	0.6656	41
Refitting car manufacturing	0.9526	17	Non-ferrous metal rolling	0.6634	42
Glass fiber and reinforced	0.9517	18	Synthetic materials	0.6442	43

plastic products manufacturing			manufacturing		
Biochemical pesticide and microbial pesticide manufacturing	0.9466	19	Electronic component manufacturing	0.5713	44
Metal processing machinery manufacturing	0.9465	20	Basic chemical raw material manufacturing	0.4947	45
Optical glass manufacturing	0.9455	21	Electronic device manufacturing	0.4742	46
Biomedical equipment and apparatus manufacturing	0.9361	22	Communication equipment manufacturing	0.4125	47
Nonferrous Metals manufacturing	0.9360	23	Auto parts manufacturing	0.3214	48
Graphite and non-metallic products manufacturing	0.9212	24	Automobile manufacturing	0.2651	49
Biopharmaceutical manufacturing	0.9185	25			

From the table 3, it can be concluded that during the sample period, the average level of strategic emerging industries is 0.83, 34 industries are higher than the average level, and another 15 industries are lower than the average level. There are differences and imbalances between the economic efficiency of various industries in China's strategic industries.

On the whole, tram manufacturing industry ranks first, followed by low-speed truck manufacturing, non-ferrous metal casting, radar and supporting equipment manufacturing, trailer manufacturing, clock and timing instrument manufacturing. The last six provinces in the comprehensive evaluation of economic efficiency are electronic component manufacturing, basic chemical raw material manufacturing, electronic component manufacturing, communication equipment manufacturing, auto parts manufacturing and auto vehicle manufacturing. The comprehensive evaluation is in line with the objective reality as a whole.

This paper divides 49 strategic emerging industries into four categories:

The first category: $C_i \geqq 0.98$; It maintains 6 industries, including tram manufacturing, electronic component manufacturing, basic chemical raw material manufacturing, electronic component manufacturing, communication equipment manufacturing, auto parts manufacturing and auto vehicle manufacturing.

The second category: $0.98 > C_i \geqq 0.90$;It maintains 22 industries.

The third category: $0.90 > C_i \geqq 0.6$. It maintains 15 industries.

The fourth category: $C_i < 0.6$. It maintains 6 industries, Including electronic component

manufacturing, basic chemical raw material manufacturing, electronic component manufacturing, communication equipment manufacturing, auto parts manufacturing, auto vehicle manufacturing.

Table4 Distribution of Chinese Strategic Emerging Industries

Strategic Emerging Industries	Total number	types	Industries number
New information technology industry	1	Average level	1
High-end equipment manufacturing industry	15	Leading level	3
		High level	5
		Average level	4
		Low level	3
New material industry	13	Leading level	1
		High level	8
		Average level	4
Biological industry	11	High level	5
		Average level	5
		Low level	1
New energy automobile industry	5	Leading level	2
		High level	1
		Low level	2
New energy industry	1	Average level	1
Energy saving and environmental protection industry	2	High level	2
Digital creative industry	1	High level	1

From the above table, it can be seen that the overall economic efficiency of high-end equipment manufacturing is 0.84, and more than 50% of strategic emerging industries have higher economic efficiency than the average. Manufacturing industry, as the foundation of building a country, the tool of rejuvenating a country and the foundation of a powerful country, is the main body of the national economy. Its economic efficiency represents the development quality of the whole manufacturing industry. From the empirical results, the high-end equipment manufacturing industry is gradually achieving the goal of high-quality development, which is one of the important objectives of the development of strategic emerging industries.

The overall economic efficiency of the new generation information technology industry and new energy industry is at an average level, and that of the energy conservation and environmental protection industry and digital creative industry is 0.94 and 0.91 respectively, which are at a high level; the overall average economic efficiency of the new material industry and the biological industry is 0.89 and 0.79, which are at an average level; the overall average economic efficiency of the new energy automobile industry is the lowest, It is the industry with the biggest difference among the strategic emerging industries.

5. Advices

In this paper, TOPSIS entropy weight method is used to measure the economic efficiency index system of strategic emerging industries, and then the ideal value method is used to rank 49 industries in the strategic emerging industries catalog. Therefore, improving the economic efficiency of strategic emerging industries needs to be carried out from the following six aspects:

Firstly, consolidate industries with high economic efficiency, increase investment in fixed assets and scientific and technological personnel for industries with weak foundation and low economic efficiency, and form a three-dimensional development pattern with close links between industrial chains, mutual technology promotion.

Secondly, the high-risk industry needs to carry out risk publicity, remind the investors who are about to enter the industry, and improve the internal control and risk evaluation system of the industry. Establish the risk management committee of the industry to control the international risk, financial risk and operational risk. We will improve the ability of strategic emerging industries to prevent and deal with risks, resolutely win three major battles, guard against major risks, and do a good job in "six stability" work, such as risk prevention and stability assurance.

Thirdly, promote the ability of high-end equipment manufacturing industry to absorb and transform advanced technology. In order to consolidate the high economic efficiency of the high-end equipment manufacturing industry, and to follow the important instructions of general secretary Xi Jinping in Henan's research, "to promote China's manufacturing and real economy, and to push China's economy from quantity to quality", we must enhance the manufacturing industry's ability to absorb advanced productive forces and speed up technological upgrading, and achieve the innovation drive of high-end equipment manufacturing industry. High quality development path.

Fourth, we should rationally allocate the input of strategic emerging industry elements, improve the production technology efficiency of the industry, and optimize the management cost, sales cost and financial cost of the industry, especially in the process of selling goods and

materials and providing labor services.

Fifth, improve the policies and regulations of strategic emerging industries. Strategic emerging industry is an important strategy for future development. National ministries and commissions actively promote the development of strategic emerging industry, and plan strategic emerging industry from the macro level of industrial production, deep integration, non-public economy and so on. In order to achieve high-quality development of strategic emerging industries, provinces and municipalities need to further form supporting laws, regulations and relevant policies.

Sixth, create a good financing environment and accelerate the marketization of industry. As a whole, the strategic emerging industry has entered the development period. It needs financial and credit funds to expand the emerging business, and it needs to improve the main business income of the strategic emerging industry. This can not be separated from the local government's market allocation, financing environment, capitalization operation and other aspects to create a vigorous development atmosphere.

Reference

[1]Hirschman, A. O .The strategy of economic development [M]. New Haven：Yale University Press，1966.2-49.

[2]Krugman, P. R. Strategic trade policy and the new international economics [M]. Cambridge Mass：MIT Press，1986.14-69.

[3]Potter.Competitive strategy-techniques for analyzing industries and competitors [M].Beijing：Huaxia Press，1997.43-108.

[4]Erickcek & Watts .Emerging Industries: Looking Beyond the Usual Suspects: A Report to WIRED[R].Kalamazoo MI：W.E.，2007.4-12.

[5]Farell, M. J. The Measurement of Production Efficiency[J].Journal of the Royal Statistical Society，1957（3）：253-290.

[6]Aigner D.，Lovell K. Formulation and Estimation of Stochastic Frontier Production Function Models, Journal of Econometrics，1977（6）：21-37.

[7]Elena B., Barbara C. & Girardone C. (2006)，“Efficiency and Stock Performance in European Banking”, Journal of Business Finance & Accounting, 33(1-2): 245-262.

[8]Eric Wang C. (2007), “R&D Efficiency and Economic Performance: A Cross-country Analysis Using the Stochastic Frontier Approach”, Journal of Policy Modeling, 29(2): 345-360.

[9]Sharma, S. & Thomas, V. J. (2008), “Inter-country R&D Efficiency Analysis: An Application of Data Envelopment Analysis”, Scientometrics, 76:483~501.

[10]Li Qiang. Research on the technological innovation efficiency of small and medium-sized enterprises in China,Science and Technology Management, 2010（10）:43-45.

Boundary U-Net: A Segmentation Method to Improve Salt Bodies Identification Accuracy

Yelong Zhao[1], Bo Liu[1*] , Jianqiang Li[1], Yong Li[1], HL Chen[1], Guangzhi Qu[2]

[1] Faculty of Information Technology, Beijing University of Technology, Beijing 100124, China
[2] Computer Science and Engineering Department, Oakland University, Rochester, MI 48309, USA
zhaoyelong@126.com, boliu@bjut.edu.cn, lijianqiang@bjut.edu.cn, li.yong@bjut.edu.cn, chenhongli666@126.com, gqu@oakland.edu

Abstract. Seismic imaging is a geophysical technique to investigate the subsurface based on seismic wave. Seismic image is a crucial channel to acquire salt deposits positions which are important evidences to locate oil and gas resources area underground. Although the development of geophysical technology makes it easier to collect more seismic images, difficult wave shapes and structures in images still hinder the use and analysis of these valuable data. In this paper, we adopt a novel deep learning method to solve this issue. An individual decoder, named Boundary Decoder, is introduced in U-Net based framework to improve segmentation performance of seismic images. Boundary decoder is supervised by labeled salt body boundaries and further assists original segmentation decoder to generate more precise results with its boundary priori knowledge guidance. As a result, we called this novel network Boundary U-Net (BU-Net). In experimental evaluation, the proposed approach can increase IoU by 2.6% and optimize the recognition of semantic boundary and details of salt bodies greatly.

Keywords: Deep Learning, Semantic Segmentation, Seismic Image.

1 Introduction

As crucial chemical raw materials and energy supplies, demand for oil and natural gas has been increasingly and accelerates the exploitation of natural gas and oil. Before exploitation, professional exploration group will carry out elaborate efforts geological investigation to determine their location. Generally, oil and gas accumulation areas tend to form saline salt deposits in subsurface because salt deposits are impermeable structures and provide entrapment for oil and gas reservoirs [1]. Seismic imaging technology can reflect geological information of deep underground like an ultra-sound and is the most common method to acquire salt bodies information in the these groups [2]. However, due to the abstract structures and complex geological information in seismic data, seismic image interpretation is extremely expensive and highly depends on expertise.

Corresponding Author: Bo Liu (boliu@bjut.edu.cn)

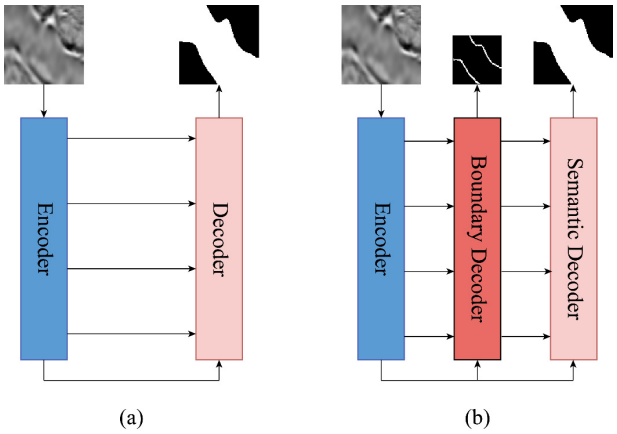


Fig. 1. Main architecture. (a) The original U-Net structure constructed with an encoder and a decoder (b) Our proposed BU-Net, an approach to learn more details. A boundary decoder is employed between the original encoder and decoder. Boundary decoder is aimed to extract boundary information in seismic images and assist semantic segmentation, which is beneficial to locate salt body in complex and blur scenes.

Image semantic segmentation, i.e., assign a semantic label to every pixel in an image, is a quite fundamental task to understand the image scene in computer vision. Over the past decade, a few research groups have already proposed various automated methods. Early work [3-8] primarily based on images properties of seismic image, e.g., gradient, structure and contrast. And many of these methods [3-6] exert boundary detection algorithm and transform segmentation task into boundary detection. Though these methods are lack of robustness and require researchers to grasp professional knowledges, it is worth noting that boundary is effective for salt body segmentation.

Recently, with the improvement of Convolutional Neural Network (CNN), numerous segmentation architecture based on fully convolutional network (FCN) [9] have been proposed. These researches have broken the limits traditional handcrafted features in many fields. More recent works employed CNN in salt deposits segmentation [10, 11]. Although technology advances greatly, these methods ignore the information loss of boundary in networks. As shown in Fig. 2, in seismic image, there is lake of clear boundary between salt body and other geological structure and the texture and structure are quite complex. Hence sometimes models cannot contain enough high-resolution information after down sampling.

As a solution to this challenge, in this paper, a boundary decoder is designed and incorporated in a U-Net [12] architecture which is called Boundary U-Net (BU-Net). The main architecture is shown in Fig. 1. This decoder is parallel to the segmentation decoder and outputs a salt body boundary mask of seismic image. Different from common U-Net architecture, segmentation decoder not only receives the intermediate features in different scale to utilize shallow features but also refine the semantic boundary with the guidance of boundary information. We also employ conditional batch normalization layers [13] in segmentation decoder to merge boundary feature and save more boundary information.

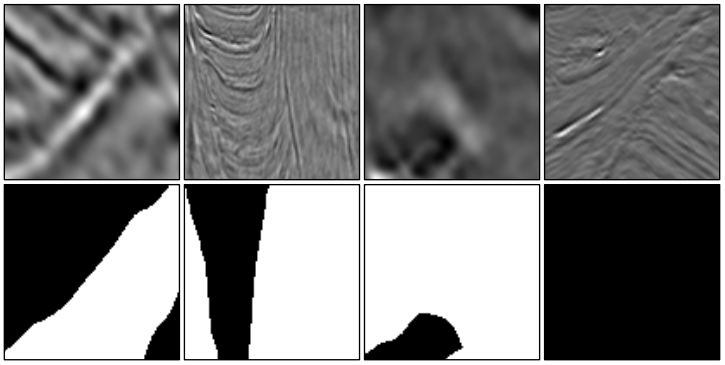


Fig. 2. Several Examples of seismic images and their corresponding segmentation and boundary Ground-Truth.

2 Related Works

2.1 Seismic Image Segmentation

Seismic image segmentation is an important issue and plenty of researches have been devoted in this area. Most of the earlier approaches mainly rely on handcrafted image features. Halpert and Clapp [8] utilize dip variability and different frequency attributes of seismic images to detect salt bodies. Shafiq et al. [7] propose an attribute in seismic images to detect target area with the help of fast Fourier transform. Moreover, plenty of works tend to identify the salt boundaries to locate salt bodies. Lomask et al. [4] transform seismic image to a weighted undirected graph and normalized cut image segmentation [14] is applied to track boundary. Aqrawi et al. [3] apply a Sobel detector to design a salt body detection algorithm involving both amplitude normalization and dimension weighting. Amin and Deriche [6] propose an approach that is more sensitive to variations in seismic data by detecting more directions compared with original Sobel filter. Wu [5] highlights the salt boundaries by computing salt likelihoods and ridges of likelihood values. These achievements indicate that boundary is a crucial information for salt bodies segmentation. However, these methods require researchers understand related geological knowledge and limit the performance.

Nowadays, CNNs achieved extraordinary effect in semantic segmentation. In 2018, TGS, a world leading geoscience data company, hosted a competition on Kaggle [15]. With the help of the large dataset in this competition, lots of approaches were proposed and achieved state-of-the-art. These methods are mainly based on U-Net [12] architecture. To improve the segmentation score, data distillation, semi-supervising and multi-model voting are used in leading approaches [10, 11].

In this paper, inspired by prior methods based on boundary knowledge, we attempt to improve the segmentation accuracy in a novel perspective and introduce boundary detection in this work. With the supervised of boundary labels, the network can amplify the texture dissimilarity between the salt bodies and background.

2.2 Semantic Segmentation based on Boundary

Recently, many semantic segmentation architectures begin to take advantage of boundary information, especially in salient object detection. These works have proven that boundary information of the image is crucial to predicting the detailed output.

Generally, there are mainly two approaches to utilize boundary feature. The first approach uses the losses on edges to preserve the boundaries of object and refine the segmentation result. Zhu et al. [16] proposed a prostate segmentation network for MR Image which designed a boundary-weighted segmentation loss to make the network more sensitive to the boundary and enforce network learning more knowledge near boundary during transfer learning. Chen et al. [17] presented Coutour Loss to leverages object contours for salient object detection which narrows the boundaries gap of ground-true mask and predicted map. Besides, the second approach incorporated boundary information into models directly. Yu et al. [18] proposed Discriminative Feature Network (DFN) containing a sub-network named Border Network. Border Network provides semantic boundary supervision and tackle inter-class indistinction problem. For thermal image semantic segmentation, in [19], a Holisticallynested Edge Detection (HED) [20] network is trained firstly to generate semantic edge features and these features are embedded into DeepLab v3 [21]. Certainly, both approaches can be applied in same model simultaneously. Li et al. [22] wired shape information as a separate processing branch and designed a dualtask loss function to ensure consistency in two tasks, predicted boundary and predicted semantics boundary, ground truth boundary and predicted semantics boundary, respectively.

Following previous works, we proposed a boundary decoder which can be easily incorporated into existing networks to relieve the details loss. Under this architecture, boundary decoder and segmentation decoder share the same encoder and can be trained end-to-end which ensures efficiency of the semantic segmentation network.

3 Methods

The proposed BU-Net mainly consists of three major parts: encoder, boundary decoder, and segmentation decoder. In this section, we elaborate the architecture structure and the effectiveness of these parts. All convolution layers in our architecture are followed by batch normalization [23] and ReLU.

3.1 Main Network Architecture

As discussed above, seismic image is an ultra-sound scan of the subsurface and hence is quite similar to biomedical images. Hence, as shown in Figure 3, our network follows U-Net, a seminal fully convolutional encoder-decoder semantic segmentation architecture for many biomedical images. U-Net is a typical U-shape structure, where high-level semantic information is captured in encoder while the detailed low-level information is recovered in the decoder path. And there has a skip connection between

encoder and decoder block in same resolution to recover more spatial information. We use a universal backbone ResNet [24] pre-trained on ImageNet [25] as base encoder.

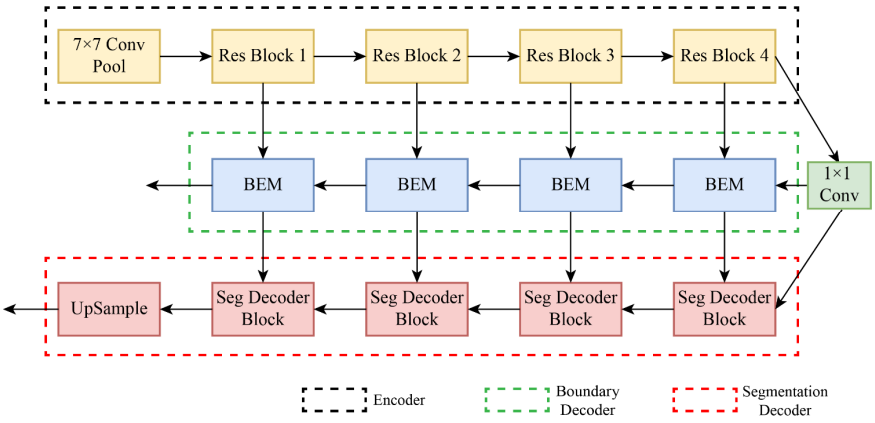


Fig. 3. The main architecture of BU-Net.

In seismic image, the boundary between salt body and background is blurred and hard to recognize. Therefore, we need to amplify the gap of their features and pay more attention on boundary. The tasks of semantic segmentation and boundary detection are closely relevant. We integrate a separate boundary decoder to combine multi-scale boundary priori knowledge into original decoder. For ease of distinction, the original decoder for semantic segmentation is called segmentation decoder.

3.2 Boundary Decoder and Boundary Enhancement Module

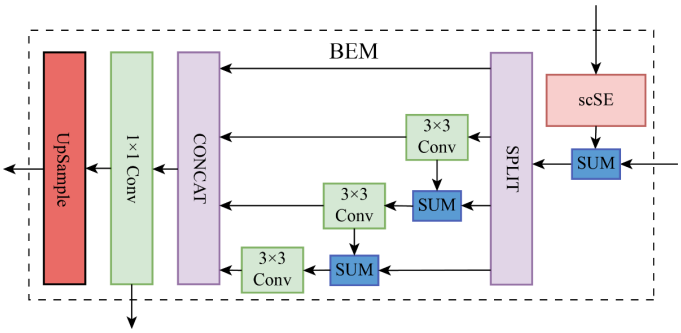


Fig. 4. Structure of Boundary Enhancement Decoder

As shown in Fig. 4, inspired by previous works about boundary detection [26, 27], we stack multiple convolutional modules named Boundary Enhancement Module (BEM) as boundary decoder, which forms a top-down architecture to achieve semantic boundary detection. This architecture can be employed in any U-shape network easily because boundary decoder can directly receive the output of encoder and share the

same encoder with segmentation decoder. Besides, boundary decoder is supervised by boundary ground-truth segmentation map and the supervisory signal can be propagated back to encoder blocks different stage. The distinction of foreground features and background features in encoder can be enlarged greatly. In our work, we acquire these semantic boundary label by Canny [28].

The stacked BEMs are aimed to enhance the boundary feature. The main idea of BEM is stated in Fig. 4. Firstly, the feature maps from encoder is weighted by a scSE [29] attention block to aggregate more related information. The processed encoder feature map is summed with high-level boundary feature, i.e. the output of last BEM. And we append a Res2Net module [30] as the main convolutional operator in BEM to achieve the right trade-off between performance and computational complexity. As shown in Fig. 4, the feature map is divided into four parallel branches. Output features of the previous branches are then sent to next branches. At last, all feature maps are concatenated and normalized with 1×1 convolutional kernel. Res2Net introduces multi-scale features in one layer and increases the range of receptive fields. And this architecture, i.e., group convolution with different receptive field, not only reduces computation dramatically, but also is applied in other works [31, 32] to enhance relative features. Finally, the feature maps are upsampled by bilinear interpolation.

3.3 Segmentation Decoder Block

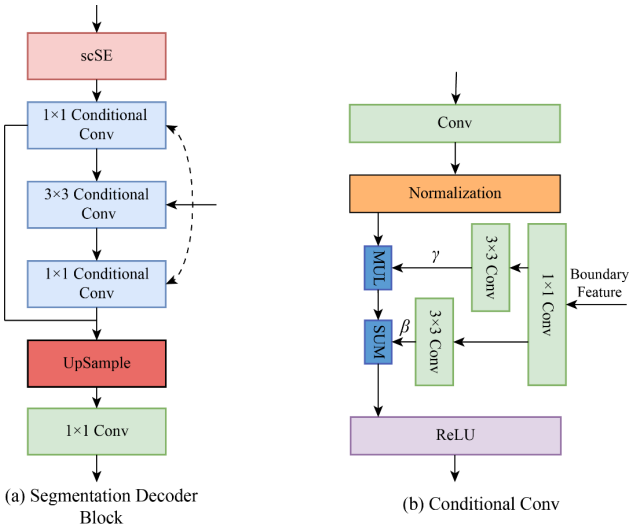


Fig. 5. Illustration of segmentation decoder block and its conditional batch normalization to fuse boundary feature.

Segmentation decoder corresponds to the decoder in U-Net architecture, which recovers the compressed semantic details to segmentation map. Like most U-Shape network, segmentation decoder consists of successive upsample blocks to refine the feature maps in low resolution progressively. As shown in Fig. 5, our segmentation decoder block is based on ResNet blocks where the first 1×1 convolution has n / 4 filters as a

bottleneck layer before the 3×3 convolution and the last 1×1 convolution to restore the number of filters to c / 2, corresponding to the channel in next stages.

Inspired by previous works about Generative Adversarial Network (GAN), the batch normalization in 3 × 3 convolution is replaced with conditional normalization [13] to fuse and utilize the boundary feature maps better in boundary decoder. Firstly, similar to batch normalization [23], the feature map x is normalized by its mean μ and standard deviation σ and then modulated with scale and bias learned from the feature map b of boundary decoder in each stage. It can be given by:

$$\hat{x} = \gamma(b) \cdot \frac{x-\mu}{\sigma} + \beta(b) \tag{1}$$

where γ (b) and β (b) denote the functions that calculates scaling and bias values from boundary feature b. In our network, the two functions are implemented with two-layer convolutional layers stated in Fig. 5. To stabilize training process, we detach backward gradient to boundary decoder in 1×1 convolutional layer illustrated in dotted flows of Fig. 5(a). Compared with other ways in feature fuse, conditional normalization can be integrated in any network and prevent information loss of boundary.

3.4 Loss Function

Our framework contains two end-to-end decoders. We jointly supervise segmentation and boundary map prediction during training. In segmentation decoder, Lovasz loss function [34] is used and recorded as $L_S = Lovasz\ (P^S, G^S)$, where P^S and G^Srepresent prediction segmentation map and ground-truth, respectively. While, to mitigate the imbalance of the positive and negative samples, we use focal loss assigning higher weights to focus on these hard examples to optimize boundary map with low computation payload. The boundary loss function L_B is shown in (2):

$$L_B = FocalLoss(P^B, G^B) = \sum_i \begin{cases} -\alpha(1-P_i^B)^\lambda \log P_i^B & G_i^B = 1 \\ (1-\alpha){P_i^B}^\lambda \log(1-P_i^B) & G_i^B = 0 \end{cases} \tag{2}$$

where $P_i^B \in [0,1]$ and $G_i^B \in \{0,1\}$ represent the label of i-th pixel in the prediction the ground-truth respectively. And we set $\gamma = 2$ and $\alpha = 0.25$ as the parameters of focal loss. Finally, as Equation 3 shows, a parameter λ is employed to balance L_S and L_B:

$$L = L_s + \lambda L_B \tag{3}$$

In next section, we carry out a series of experiment to determine the best value of λ.

4 Experiments

4.1 Dataset

The seismic image dataset is proposed in TGS Salt Identification Challenge, a competition organized on Kaggle [15]. The data is sliced from 3D seismic view of earth to 2D image and then cropped into small patches of 101 × 101 pixels. The competition

provides 4000 labeled images and 18000 unlabeled images, which are used for training and testing respectively.

4.2 Implementation Details

Data augmentation. In our experiment, we split whole labeled data into training set and validation set containing 3200 and 800 examples. All models are trained on training set and the performance of models is evaluated on validation set in each epoch. During training, data augmentation is utilized, which scales each image to 256×256 before random flipping, random changing brightness and contrast slightly, and randomly cropped between 0.8 to 1.2. And corresponding boundary label is acquired by Canny kernel after augmentation operation.

Training. We use Adam [33] to optimize our network for 100 epochs, where the initial learning rate is set to 0.0001 and the batch size is 10. And a poly learning rate policy was employed where the initial learning rate was multiplied by $(1-\frac{iter}{\max_iter})^{0.9}$after each iteration. The best results on validation set are reported during the training process. All the experiments are implemented with PyTorch and performed with one GTX 1080 GPU, CUDA 9.0 and cuDNN 7.1.

Evaluation metric. Following the metric in Kaggle competition, the proposed architecture is evaluated on the mean average precision (mAP) at different intersection over union (IoU) thresholds. The threshold t values range from 0.5 to 0.95 with a step of 0.05:

$$t \in T = \{0.5, 0.55, 0.6, \dots, 0.95\} \tag{4}$$

At each threshold value t, a precision value is calculated based on the following rules:

$$P(t) = \frac{TP(t)}{TP(t)+FP(t)+FN(t)} \tag{5}$$

where $TP(t)$, $FN(t)$ and $FP(t)$ represented number of true positives, false negatives, and false positives resulting from comparing the predicted mask in different threshold t to all ground-truth. The average precision of a single image is then calculated as the mean of the above precision values at each IoU threshold:

$$IoU = \frac{1}{|T|}\sum_{t\in T} P(t) \tag{6}$$

4.3 Ablation Study

In this subsection, we carry out a series of experiments to demonstrate effectiveness of our design. The experiment results are presented in Table I. We also present some sample results in Fig. 6 to visually compare our method with backbone approach.

Table 1. Ablation study for each component on BU-Net. (SD=Segmentation Decoder, R2N=Res2Net, BD=Boundary Decoder, CBN=Conditional Batch Normalization)

Method	IoU
U-Net	0.821
ResNet + SD (ResUNet)	0.837
ResNet + Conv-BD + SD	0.839
ResNet + R2N-BD +SD	0.842
ResNet + R2N-BD + Concat + SD	0.848
ResNet + R2N-BD + Sum+ SD	0.858
ResNet + R2N-BD + CBN + SD	**0.863**

Baseline. We use a U-Net as the most fundamental baseline model. And we construct a Res-U-Net whose encoder is replaced with a ResNet50 pre-trained on ImageNet dataset. Obviously, like many previous works, residual block can enhance model's feature extraction capability and improve convergence rate greatly. Table I shows the experiment results. By employing a pre-trained ResNet backbone, the IoU is increased by 1.6% from 0.821 to 0.837.

Ablation for Segmentation Decoder. To emphasize boundary feature in seismic images, the boundary decoder is employed in network and supervised by boundary labels. A boundary decoder is directly appended in original network and segmentation decoder do not receive its output. This structure is similar to DFN. First, we applied a simple boundary model constructed with a 3×3 convolutional layer and a upsample module which improves the performance by 0.2% slightly. Following previous works [31, 32], we replace convolutional layer with a group convolutional structure, Res2Net block, to detect boundary map. This block further increases IoU by 0.3%.

Ablation for Segmentation Decoder. The proposed boundary decoder extracts the boundary information. We believe that this information is profitable for semantic segmentation. Therefore, we compare several methods to fuse boundary feature. Firstly, same as skip connection, the segmentation features are concatenated with boundary features. The concatenating increases the IoU by 0.6%. Besides, the sum operation improves performance 1.6%. We suppose the reason for this improvement is that sum operation reduces the number of parameters and alleviates over-fitting. Besides, inspired by works of GAN, we employ conditional batch normalization which facilitates the boundary information to guide segmentation. Finally, as shown in Table I, conditional batch normalization achieves highest IOU (0.863). These results indicate that boundary information is beneficial to extract semantic features.

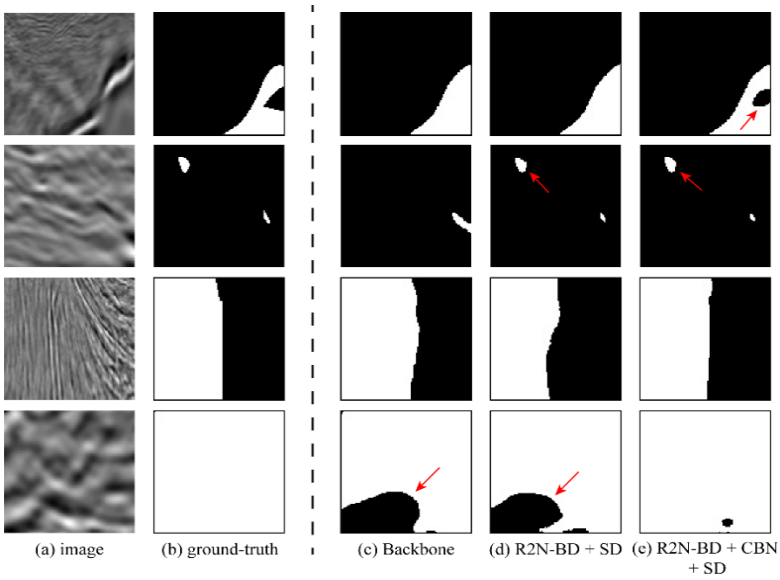


Fig. 6. Example results of salt segmentation. The areas annotated by red arrows show that our methods increase model's sensitivity to detailed information and alleviate segmentation error. The semantic boundaries are also optimized.

4.4 Comparison with Different λ Value

To achieve better performance, we conduct experiments to adjust λ value in loss function. Six different values of {0, 0.1, 0.2, 0.3, 0.4, 0.5} are tested in our experiments. As shown in Table II, the value of 0.1 achieves highest performance. It is worth noting that all results are higher than baseline with one decoder, which indicates that two-stage decoder is a crucial method to improve performance.

Table 2. Result of BU-Net in different λ value.

λ	0	0.1	0.2	0.3	0.4	0.5
IoU	0.861	**0.863**	0.860	0.859	0.856	0.855

5 Conclusions

Seismic imaging is important in oil and gas resources location. In this paper, we present a modified U-Net framework named Boundary U-Net to improve seismic images segmentation. Due to the blur and complicated texture in seismic images, an extra boundary decoder capturing boundary information is integrated in original U-Shape structure to enlarge feature gap between salt bodies and background in a supervised approach. And conditional batch normalization is also employed in segmentation

decoder to receive output of boundary decoder. Our experimental results show that boundary information and two-stage decoder are able to generate more refine segmentation results for seismic images. It is believed that this method can be applied in other segmentation tasks which would be our possible future work.

Acknowledgment

This work is supported by National Natural Science Foundation of China (61702021) and the grant of China Scholarship Council.

References

1. Link, W.K.: Significance of oil and gas seeps in world oil exploration. AAPG bulletin 36, 1505-1540 (1952)
2. Ratcliff, D.W., Gray, S.H., Whitmore Jr, N.: Seismic imaging of salt structures in the Gulf of Mexico. The Leading Edge 11, 15-31 (1992)
3. Aqrawi, A.A., Boe, T.H., Barros, S.: Detecting salt domes using a dip guided 3D Sobel seismic attribute. SEG Technical Program Expanded Abstracts 2011, pp. 1014-1018. Society of Exploration Geophysicists (2011)
4. Lomask, J., Biondi, B., Shragge, J.: Image segmentation for tracking salt boundaries. SEG Technical Program Expanded Abstracts 2004, pp. 2443-2446. Society of Exploration Geophysicists (2004)
5. Wu, X.: Methods to compute salt likelihoods and extract salt boundaries from 3D seismic images. Geophysics 81, IM119-IM126 (2016)
6. Asjad, A., Mohamed, D.: A new approach for salt dome detection using a 3D multidirectional edge detector. Applied Geophysics 12, 334-342 (2015)
7. Shafiq, M.A., Alshawi, T., Long, Z., AlRegib, G.: The role of visual saliency in the automation of seismic interpretation. Geophysical Prospecting 66, 132-143 (2018)
8. Halpert, A., Clapp, R.G.: Salt body segmentation with dip and frequency attributes. SEP-Report 136, 113-124 (2008)
9. Shelhamer, E., Long, J., Darrell, T.: Fully Convolutional Networks for Semantic Segmentation. IEEE Transactions on Pattern Analysis and Machine Intelligence 39, 640-651 (2016)
10. Karchevskiy, M., Ashrapov, I., Kozinkin, L.: Automatic salt deposits segmentation: A deep learning approach. arXiv preprint arXiv:1812.01429 (2018)
11. Babakhin, Y., Sanakoyeu, A., Kitamura, H.: Semi-Supervised Segmentation of Salt Bodies in Seismic Images using an Ensemble of Convolutional Neural Networks. In: German Conference on Pattern Recognition, pp. 218-231. Springer, (2019)
12. Ronneberger, O., Fischer, P., Brox, T.: U-net: Convolutional networks for biomedical image segmentation. In: International Conference on Medical image computing and computer-assisted intervention, pp. 234-241. Springer, (2015)
13. De Vries, H., Strub, F., Mary, J., Larochelle, H., Pietquin, O., Courville, A.C.: Modulating early visual processing by language. In: Advances in Neural Information Processing Systems, pp. 6594-6604. (2017)
14. Shi, J., Malik, J.: Normalized cuts and image segmentation. IEEE Transactions on pattern analysis and machine intelligence 22, 888-905 (2000)

15. TGS Salt Identification Challenge, https://www.kaggle.com/c/tgs-salt-identification-challenge, last accessed 2020/5/5
16. Zhu, Q., Du, B., Yan, P.: Boundary-weighted domain adaptive neural network for prostate MR image segmentation. IEEE Transactions on Medical Imaging (2019)
17. Chen, Z., Zhou, H., Xie, X., Lai, J.: Contour Loss: Boundary-Aware Learning for Salient Object Segmentation. arXiv preprint arXiv:1908.01975 (2019)
18. Yu, C., Wang, J., Peng, C., Gao, C., Yu, G., Sang, N.: Learning a discriminative feature network for semantic segmentation. In: Proceedings of the IEEE conference on computer vision and pattern recognition, pp. 1857-1866. (2018)
19. Li, C., Xia, W., Yan, Y., Luo, B., Tang, J.: Segmenting objects in day and night: Edge-conditioned cnn for thermal image semantic segmentation. arXiv preprint arXiv:1907.10303 (2019)
20. Xie, S., Tu, Z.: Holistically-nested edge detection. In: Proceedings of the IEEE international conference on computer vision, pp. 1395-1403. (2015)
21. Chen, L.-C., Papandreou, G., Schroff, F., Adam, H.: Rethinking atrous convolution for semantic image segmentation. arXiv preprint arXiv:1706.05587 (2017)
22. Takikawa, T., Acuna, D., Jampani, V., Fidler, S.: Gated-scnn: Gated shape cnns for semantic segmentation. In: Proceedings of the IEEE International Conference on Computer Vision, pp. 5229-5238. (2018)
23. Ioffe, S., Szegedy, C.: Batch Normalization: Accelerating Deep Network Training by Reducing Internal Covariate Shift. In: International Conference on Machine Learning, pp. 448-456. (2015)
24. He, K., Zhang, X., Ren, S., Sun, J.: Deep residual learning for image recognition. In: Proceedings of the IEEE conference on computer vision and pattern recognition, pp. 770-778. (2016)
25. Russakovsky, O., Deng, J., Su, H., Krause, J., Satheesh, S., Ma, S., Huang, Z., Karpathy, A., Khosla, A., Bernstein, M.: Imagenet large scale visual recognition challenge. International journal of computer vision 115, 211-252 (2015)
26. Acuna, D., Kar, A., Fidler, S.: Devil is in the edges: Learning semantic boundaries from noisy annotations. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 11075-11083. (2019)
27. Deng, R., Shen, C., Liu, S., Wang, H., Liu, X.: Learning to predict crisp boundaries. In: Proceedings of the European Conference on Computer Vision (ECCV), pp. 562-578. (2018)
28. Canny, J.: A computational approach to edge detection. IEEE Transactions on pattern analysis and machine intelligence 679-698 (1986)
29. Roy, A.G., Navab, N., Wachinger, C.: Concurrent spatial and channel 'squeeze & excitation'in fully convolutional networks. In: International Conference on Medical Image Computing and Computer-Assisted Intervention, pp. 421-429. Springer, (2018)
30. Gao, S., Cheng, M.-M., Zhao, K., Zhang, X.-Y., Yang, M.-H., Torr, P.H.: Res2net: A new multi-scale backbone architecture. IEEE transactions on pattern analysis and machine intelligence (2019)
31. Li, J., Wang, Y., Wang, C., Tai, Y., Qian, J., Yang, J., Wang, C., Li, J., Huang, F.: DSFD: dual shot face detector. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 5060-5069. (2019)
32. Liu, S., Huang, D.: Receptive field block net for accurate and fast object detection. In: Proceedings of the European Conference on Computer Vision (ECCV), pp. 385-400. (2018)

Identity Authentication Protocols for Greenhouse Environmental Monitoring System Based on Internet of Things

Shujie Lu[1] , Xiao Zhang[1]* and Ling Li[1]

[1]School of Information Science and Engineering,Hebei North University,07500 Zhangjiakou, Hebei, China
366588863@qq.com

Abstract. In a greenhouse environment monitoring system designed based on the Internet of Things technology, it is necessary to implement a security policy that implements two-way identity authentication between a user's smart APP and IoT node devices (such as sensors, controllers, etc.) to ensure safety. Therefore, the article designs a two-way authentication strategy using EC encryption n algorithm and cloud server assistance. Experiments prove that the strategy can meet the requirements of two-way identity authentication between the user's smart phone APP and the IoT node device with lower resource requirements.

Keywords: Internet of Things security, mutual authentication, key agreement, elliptic curve, security policy.

1 Introduction

Environmental factors have a greater impact on the generation of greenhouse agriculture [1]. For example, the growth temperature of tea seedlings is 22 ℃ -28 ℃, the humidity is 50% -70%, the optimal values are 25 ℃ and 60% [2]; the humidity of cucumber planting is generally maintained at 60% ~ 80%, the temperature is 18 ~ 30 ℃ is the best. The temperature needs to be increased by 3 ~ 5 ℃ during photosynthesis. Long-term temperature higher than 38 ℃ or lower than 15 ℃ will affect cucumber development [3]. Excessive humidity will cause fungal reproduction, affect cucumber growth, and then affect Cucumber quality and yield [4].

With the rapid development of Internet technology, the Internet of Things technology has emerged as the times require [5]. The greenhouse environment monitoring system designed using the Internet of Things technology is shown in Fig.1. It can realize functions such as automatic collection, storage, analysis, early warning and remote intervention of greenhouse internal environmental parameters such as light, temperature, humidity, and carbon dioxide.

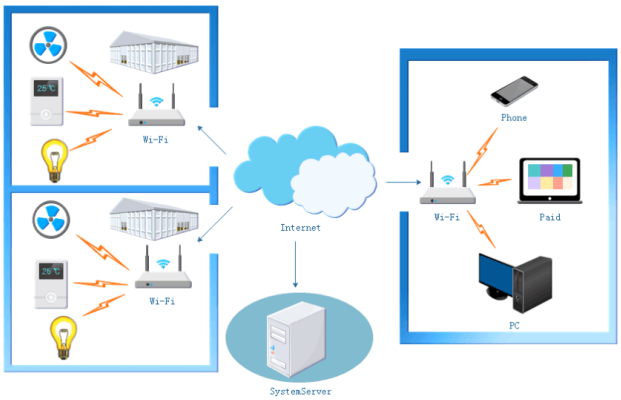


Fig. 1. The Internet greenhouse environment monitoring system

Node devices such as various environmental parameter sensors and environmental intervention controllers in Figure 1 form the sensing layer of the system; the system parameter early warning server and various user terminal devices form the application layer of the system; IoT gateway, cloud authentication The server (Cloud Authentication Server CAS) and data storage server (Data Server DS) constitute the network layer of the system.

The IoT gateway is middleware that converts network data and Internet data within the greenhouse. The device is highly automated and transparent to users. Its core business is to provide real-time status data reporting and intervention instruction distribution services to user apps through the Internet. The core business of the gateway has a low demand for hardware computing capabilities. Using the STM32F412 low-power CPU with a working frequency of 100 MHz can fully meet the demand.

However, the emergence of endless attacks on the Internet of Things has made the security issues of the Internet of Things arouse people's attention. In 2015, hackers launched a malicious attack on the Ukrainian power system, resulting in power outages for more than 700,000 residents. In the Mirai incident in 2016, the attackers used a large number of IoT devices such as web cameras to launch DDoS attacks on domain name servers, resulting in a large number of users Unable to use the network. According to a Gartner survey, nearly 20% of units and departments worldwide have suffered IoT attacks in recent years.

Considering that the IoT gateway enters the normal working state, the smart phone terminal can directly establish a communication link with the gateway, and send data query commands and environmental intervention commands (such as remote control of fans, wet curtain cooling control, etc.). Before establishing a link between the IoT gateway and the user's smart terminal, it is necessary to verify the identity and negotiate the session key to meet the requirements of user control command information not being impersonated, tampered and replayed.

2 Relevant work

Nicanfar et al. [6] Proposed an authentication protocol based on elliptic curve encryption in 2015. Although their solution greatly reduces the computational complexity, because trusted third parties need password tables to hold user information, they are vulnerable to attacks such as lost or stolen tables. Later, Li et al. [7] Proposed a new cryptographic protocol but it also proved to be vulnerable to eavesdropping and vulnerable to imitation attacks due to lack of key agreement. In 2019, Q. Jiang et al. [8] Analyzed a Das protocol [1] and proposed a new security protocol that implements wearable device authentication and key negotiation with the assistance of a cloud server. The protocol uses the ECC algorithm to further enhance the security of the data and reduce the protocol's consumption of computing resources by the device. In 2019, Wang et al. [9] Proposed a protocol that uses ECC algorithm and cloud server-assisted IoT gateway to perform two-way identity authentication and key negotiation with user smartphones.

Therefore, in this paper, most of the current IoT system perception layer devices use symmetric key protocols to build network security infrastructure. There are many key maintenance tasks. Operation and maintenance personnel often come into contact with communication keys, which can easily lead to key leakage and attacks by internal privileged personnel. And the limited computing, storage, and communication resources of the perceptual device, and the attacker is prone to launch a resource exhausted DDos attack. An EC encryption algorithm is used to design a solution that implements the two-way authentication and key negotiation between the mobile APP and the system IoT gateway device with the assistance of the system Cloud Authentication Server CAS.

3 Basic Theory

3.1 Elliptic Curve Encryption Algorithm

Due to the many benefits of ECC [10], it has been used in various environments. One of the most important benefits of ECC is providing the same level of security with a smaller key size compared to other cryptography techniques like RSA. For instance, ECC with 160 and 512 bit keys provide the same level of security as RSA, D-H or ElGamal cryptography with 1024 and 15360 bit keys, respectively. In addition to addressing the resource constraint issue, ECC is also beneficial in enabling an efficient protocol that supports current and future devices with various levels of technology, which is important in emerging SG systems.

Generally, ECC is presented as an Elliptic Curve (EC) nodes/points(x,y) over Zp, and satisfies the formula (1).

$$y^2 \equiv x^3 + ax + b \; mod \; p \; where (x, y) \in Z_p \tag{1}$$

In formula (1), p is a large prime number, and the larger the value of p, the higher the security and the greater the amount of calculation. a, b ∈ Zp are used to determine the specific elliptic curve, and need to satisfy formula (2):

$$4a^3 + 27b^2 \neq 0 \tag{2}$$

In the elliptic curve Zp (a, b) that meets the above requirements, take the point G along with integer t, then you will get K = tG, in which K is also the point of the elliptic curve Ep (a, b). According to the rule of addition: with tQ Q Q Q ... Q and given G, t, it is easy to find K but very difficult to find t if if K and G are known. In the elliptic curve encryption algorithm, G is called the base point, t is the private key, and K is the public key.

3.2 Attacker Model

In this article we assume that the attacker has the following capabilities.

1. The attacker can obtain all the information in the public transmission channel, and can modify, replay, and forge any new information to transmit to the receiver [11].
2. The attacker can know the user names of all users in this communication system and the device ID of the IoT gateway.
3. The attacker can act as a legitimate user in the system.
4. The attacker can establish a physical link with any gateway in the system through wi-fi.

4 Two-way authentication between user's smartphone terminal and IoT gateway

To implement the two-way authentication and key negotiation between the IoT gateway and the user's intelligent terminal device, we need to make the following agreement. (1) A secure communication channel has been established between the IoT gateway and the cloud authentication server CAS. (2) All devices including cloud authentication server CAS, IoT gateway, and user terminals share a common set of parameters. It includes a set of EC parameters E {a, b, p, G, n, h}; a single-term hash function h (); independent identity ID and clock. h () is used to map character information to an elliptic curve. The identity IDi is used to uniquely identify the specified device in the system. Other devices need to access the specified device through the ID. The device clock is used to keep messages fresh.

Considering that operations such as key generation, key negotiation, user authentication, and authentication require more computing and storage resources, we propose that the above operations required by the gateway be completed by the system cloud authentication server to reduce network Resource consumption demand scheme. Figure 2 depicts the strategy of two-way identity authentication and key negotiation between the user's smart terminal and the designated IoT gateway with the assistance of the system cloud authentication server. The basic process is as follows.

Step 1. The user logs in to the system cloud server through the smartphone terminal app, and the server verifies the identity of the user and the legitimacy of the app. Success is to establish a secure communication channel between the user app and the system cloud authentication server using traditional DTLS.

Step 2. The smartphone terminal App randomly generates a number p, which is calculated by formula (3), and the point P (x, y) on the EC. Then, generate request information Requestu-g = {Appid, GWID, P} requesting access to the GWID-designated gateway, and send Requestu-g to the system authentication server through the previously established secure channel.

$$P = p \cdot G \tag{3}$$

Step 3. After receiving the Requestu-g, the cloud server verifies the user's authority to access the gateway specified by GWID. If so, choose a number r, and calculate the EC point R by formula (4). Use formula (5) to calculate the session key K in the form of an EC node.

$$R = r \cdot G \tag{4}$$

$$K = r \cdot P = r \cdot p \cdot G \tag{5}$$

Step 4. The cloud server uses the server-gateway session key to encrypt {Appid, GWID, TimeStamp, K}, and serves as the user ’s smart terminal to specify the temporary identity authentication certificate Au of the GWID. The point R and the certificate are returned to the user's smart terminal App; the {Appid, GWID, TimeStamp, K, random number} is encrypted as the certificate A ′ and sent to the GWID designated gateway.

Step 5. After receiving the certificate A′ , the gateway decodes the certificate, presents the information, and uses formula (6) to calculate the symmetric session key k for communication with the user terminal APP and waits for the user's smart terminal App access request.

$$k = h(K) \tag{6}$$

Step 6. After receiving the point R and certificate A returned by the cloud server, the user terminal APP uses formula (7) to calculate the session key K in the form of an EC node. After formula (6), the session key k is calculated.

$$K = p \cdot R = p \cdot r \cdot G = r \cdot p \cdot G \tag{7}$$

Step 7. The user terminal APP uses the calculated key k and AES128 algorithm to encrypt the certificate A and sends it to the GWID designated gateway.

Step 8. After receiving the access request from the user terminal APP, the gateway uses the key k and the AES128 algorithm to decrypt the operation to obtain certificate A, and compares it with the information in certificate A 'received by the server. If they match, authentication succeeds. Both parties use the key k for subsequent communication.

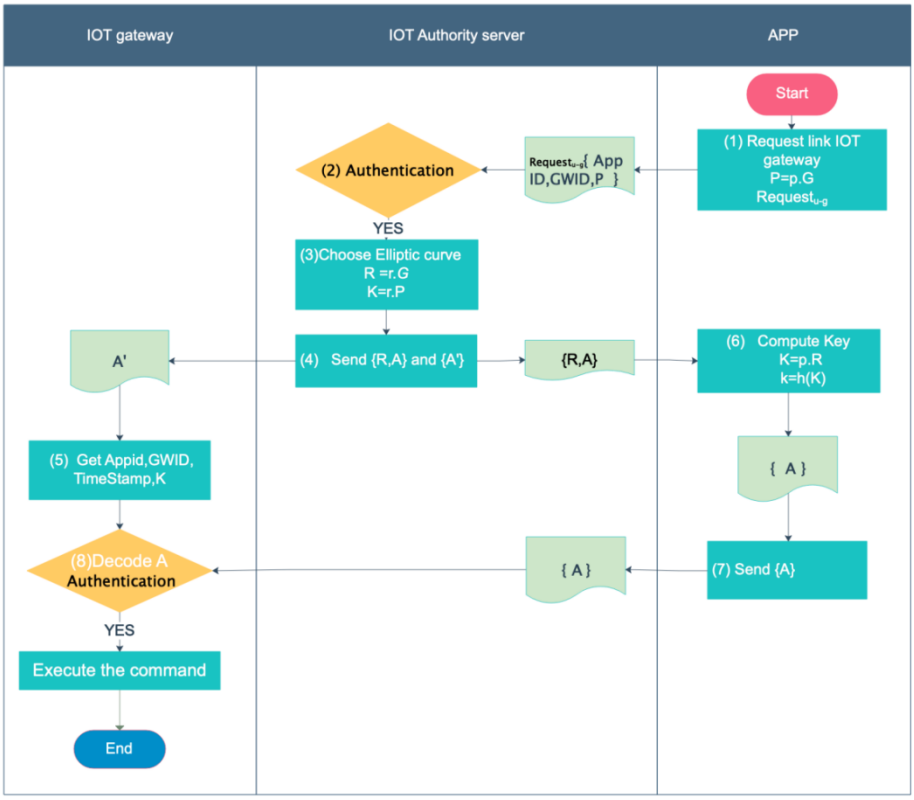


Fig. 2. the strategy of two-way identity authentication and key negotiation between the users' smart terminal and the designated IoT gateway with the assistance of the system cloud authentication server.

5 System Security Analysis

This section will analyze the security of the improved gateway and user two-way authentication protocol under the assistance of the server under the attacker model assumed in this article in section 2.3.

5.1 Anti-eavesdropping analysis.

During communication, the communication channel between the server and the gateway and between the server and the client is a secure encrypted channel. Therefore, eavesdroppers cannot eavesdrop on the true content of the information obtained. When a communication channel is established between the client and the gateway, the listener can obtain the public key Kc-g, but cannot obtain the private key k of both parties. Even if the secret C is obtained, the real information cannot be obtained, so the protocol has anti-eavesdropping capabilities.

5.2 Anti-counterfeiting analysis.

The communication information in the system uses the elliptic curve Ep (a, b), the selected base point G, and the private key t, and the identity IDs of both parties in M cannot eavesdrop on the encrypted information. Therefore, in this system, it has Anti-counterfeiting ability.

5.3 Anti-retransmission analysis.

Retransmission attack is an attack method in which eavesdroppers illegally obtain trust information by resending previous information after illegally obtaining communication information in the network. Information such as timestamp, validity period, and sequence number are added to this system message, so that the protocol has the ability to prevent retransmissions.

5.4 Camouflage attack

If an attacker wants to pretend to be a legitimate user to communicate with the gateway, he must forge legitimate authentication information. From the above analysis, it is known that only the server designated user has a server authentication certificate, so the protocol can prevent spoofing attacks.

5.5 Man in the middle attack

According to the attacker, there is no way to pretend to communicate with other users, and the attacker cannot forge legitimate authentication information, so this protocol can resist man-in-the-middle attacks.

6 Experiments and conclusions

Based on the system architecture diagram of the greenhouse environmental monitoring system shown in Fig． 1, an experimental test platform was constructed. The system server group uses a fixed IP 10MP optical fiber dedicated line to access the Internet. The mobile phone with Android system is used as the client platform. The STM32F412 low-power CPU with a Cortex®-M4 core operating at 100 MHz is used to build an IoT gateway to implement the user authentication protocol designed in this paper. In the experiment, based on the attack model, a computing resource exhausted DDos attack was launched on the IoT gateway. The results show that the solution has higher security, consumes less resources, has less impact on the gateway's core business communications, and runs the gateway more stably.

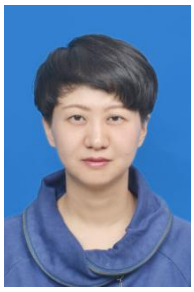
ACKNOWLEDGMENTS

The work was supported by the Population Health Information in Hebei Province Engineering Research Center.

References

1. McSorley R., M. Ozores-Hampton, P.A. Stansly, & J.M. Conner. (1999). Nematode management, soil fertility, and yield in organic vegetable production.Nematropica,29(2), 205.
2. Nicole Kfoury, Joshua Morimoto, Amanda Kern, Eric R. Scott, & Albert Robbat. (2018). Striking changes in tea metabolites due to elevational effects. Food Chemistry, 264, 334.
3. C.-F. Zhou, C.-T. Wu, D.-D. Li, X.-W. Zhang, & X.-Z. Ai. (2018). Hydrogen sulfide promotes chilling tolerance of cucumber seedlings by alleviating low-temperature photoinhibition.Plant Physiology Journal, 54(3), 411-420.
4. JianHua Xu, Sheng Ye, & JiaJun Fu. (2018). Novel sea cucumber-inspired material based on stiff, strong yet tough elastomer with unique self-healing and recyclable functionalities. Journal of Materials Chemistry A, 6.
5. Kopetz, Hermann. "Internet of things." Real-time systems. Springer US, 2011. 307-323.
6. Nicanfar H, Leung V C M. Multilayer Consensus ECC-Based Password Authenticated Key-Exchange (MCEPAK) Protocol for Smart Grid System[J]. IEEE Transactions on Smart Grid, 2013, 4(1):253-264
7. Li D, Aung Z, Williams J R, et al. Efficient and fault-diagnosable authentication architecture for AMI in smart grid[J]. Security & Communication Networks, 2015, 8(4):598-616.
8. Q. Jiang, Y. Qian, J. Ma, X. Ma, Q. Cheng, and F. Wei, “User centric three-factor authentication protocol for cloud-assisted wearable devices,” Int J Commun Syst, vol. 32, no. 6, p. e3900, Apr. 2019.
9. ZHIHUI WANG, JIANLI ZHAO, BENZHEN GUO, JINGJINGYANG, XIAO ZHANG. Mutual Authentication Protocol for IoT-based Environment Monitoring System[j]Journal of Environmental Protection and Ecology 2019,6(2)
10. A. H. Koblitz, N. Koblitzb, and A. Menezes, “Elliptic curve cryptog- raphy: The serpentine course of a paradigm shift,” Elsevier J. Number Theory, vol. 131, no. 5, pp. 781–814, May 2011.
11. K. Mahmood, S. A. Chaudhry, H. Naqvi, S. Kumari, X. Li, and A. K. Sangaiah, “[4]An elliptic curve cryptography based lightweight authentication scheme for smart grid communication,” Future Generation Computer Systems, vol. 81, pp. 557–565, Apr. 2018.

Author



Shujie Lu,graduate student.Hebei North University,Zhangjiakou, China, the main research direction： agricultural informatization.

Xiao Zhang*,Corresponding author.Professor, master's tutor. Hebei North University,Zhangjiakou, China, the main research direction： medical informatics.

Ling Li,master. Hebei North University,Zhangjiakou, China, the main research direction： Internet of Things security.

Design of anatomy multimedia E-learning platform based on Toolbook

Mengjiao Liu[1], Liheng Gong[1] ,and Xiao Zhang[1*]

[1]School of Information Science and Engineering, Hebei North University,07500 Zhangjiakou, Hebei, China
821273966@qq.com

Abstract. The Anatomy Multimedia E-learning (Electronic Learning) platform is a platform for preparing and displaying courseware for lectures, lectures and self-study by anatomy teachers. This platform not only solves the problem of self-adaptability and self-upgradability of courseware, but also enables front-line teachers to realize the free organization of teaching content and multimedia materials without writing a single line of program, and get rid of development caused by non-programming difficult. The Anatomy Multimedia E-learning platform uses the ToolBook multimedia development tool as the development environment and uses the object-oriented Openscript programming language to write programs. It is designed to open two independent "books" on the desktop at the same time based on the common characteristics of morphology courses. One "book" The text script used to input and display the courseware is a "teaching plan" and the other is a morphology "atlas". The two books can be related to each other through hot words or can be viewed independently. Anatomy multimedia courseware helps teachers to actively control and use the E-learning teaching platform, helps to fully realize E-learning's teaching auxiliary functions and mobilize teachers' personalized teaching.

Keywords: E-learning, Multimedia lesson preparation platform, Anatomy, Toolbook.

1 Purpose And Significance Of Development

Anatomy belongs to the curriculum of medical morphology. Medical morphology teaching practice and activities will involve a lot of knowledge. Teachers 'teaching and students' autonomous learning may need to obtain a lot of network information resources [1].This E-learning teaching platform is very suitable for the teaching of medical morphology courses, which makes up for the tediousness of traditional

teaching, not only saves time for preparing lessons, but also improves students' learning interest. To some extent, it relieves the pressure of insufficient teachers, shortage of teaching hours, and various contents.

At present, most of the domestic and foreign curriculum development is based on web-based learning platforms, interactive platforms or multimedia courseware developed by PowerPoint. Due to the complex and scattered mess of course resources, teachers' inefficient preparation of courses and uneven levels of courseware production have affected students. Learning interest and learning effect. The anatomy multimedia E-learning platform can better solve the programming problems in the development of E-learning. Its advantages are not only to reduce the difficulty for teachers to prepare lessons and achieve high-quality lessons, but also to make the teaching content image, intuitive and vivid, improve classroom efficiency, and enhance teaching effects. In addition, the developed courseware can also be circulated or burned to classmates as internal materials, so that students can freely choose the time to study, or save as materials to better play the function of courseware [2].

The anatomy multimedia E-learning platform can better solve the problem of adaptability and self-upgradability of courseware. Any excellent courseware cannot meet the needs of teaching once and for all. As a lesson preparation system, this software completely gives the teachers the right to determine and modify the teaching content, which solves the problems of self-adaptability and self-upgradability of courseware.

In order to further promote teaching reform, improve the corresponding teaching resource library, provide rich graphics, text and audio-visual resources for the teaching of anatomy, reduce the burden on teachers to produce courseware, and build a multimedia lesson preparation platform that meets the needs of practice is very necessary [3]。

2 Main Technologies And Development Tools

TooLBook is an object-oriented multimedia production tool. The biggest feature is the development work under the integrated environment of Windows. Anyone can use Toolbook to develop interactive courseware, testing, evaluation and simulation training in accordance with international standards. ToolBook has powerful menu design functions. Creators can add, delete, and modify menu items in ToolBook's system menu according to the overall concept, requirements, and style of the courseware, or design menus with different styles and multiple functions. At the same time, developers can directly cut into the user layer to watch the production effect, and support dynamic data linking, object embedding and linking.

ToolBook organizes information in the form of books and pages. The page is the main work screen provided by the system to the creator. ToolBook allows creators to open multiple windows (Viewer) for a book, different windows can use different menus [4]. The "page" is composed of text fields, buttons, hot words, graphics and

other objects (see **Fig. 1.**). The "pages" are organically combined through navigational clues.

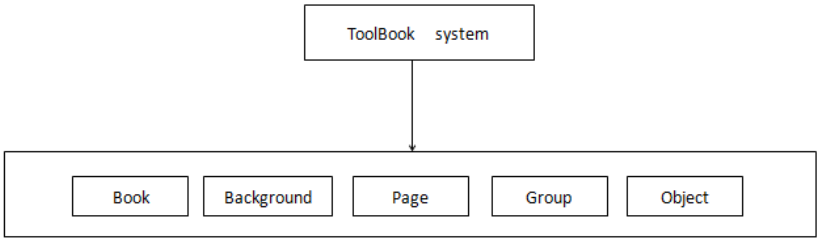


Fig.1. System composition of ToolBook

3 System Design And Function Implementation

The overall design is based on the comprehensive consideration of the teaching purpose of E-learning, so as to complete the overall structure of E-learning software. According to the common characteristics of morphology courses, this platform is designed to open two independent "books" on the desktop at the same time, one is a text script for inputting and displaying courseware, is a "teaching plan", and the other is anatomy "atlas ", Two books can be related to each other through hot words, or can be viewed independently. The content of the picture should be reasonably arranged, and the correct layout of the image and text should be combined to make a unified template. It is recommended to attract students' attention. Through color matching to enhance and enrich the visual experience, a correct and scientific color scheme can reduce students to watch a large number of images Visual and mental fatigue [4].

3.1 System Function Design

This E-learning teaching platform has two working levels, reader level (use) and author level (editor). At the reader level, users can run the developed program, at the author level, they can create new books, and create and modify "pages" Objects (for example: text, images, videos). The main functions of the system are shown in **Fig. 2.**:

4

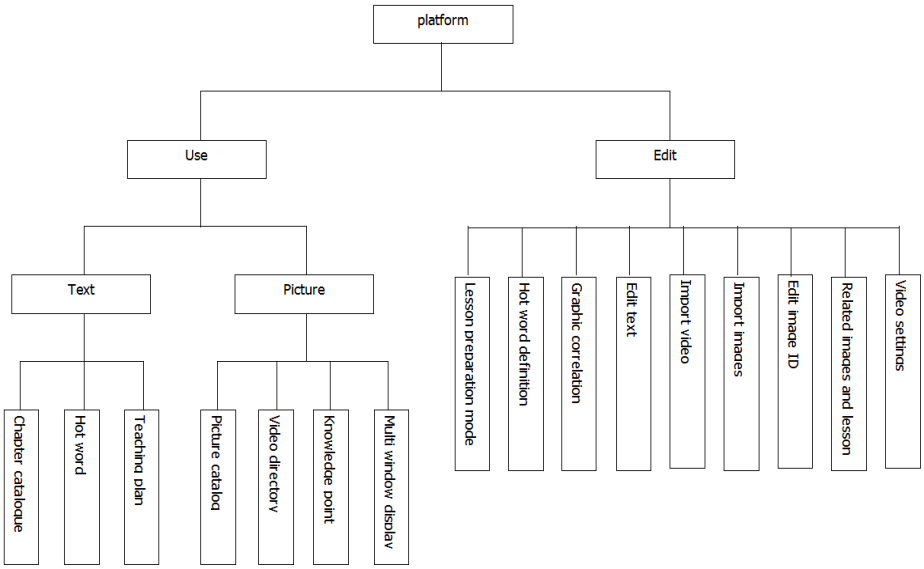


Fig. 2. System function block diagram

1. The "teaching plan" appears in the form of pages. On each page, there are chapters, sections and page numbers. In order to meet the specific needs of teaching, the font, size and color of the text can be adjusted freely through the menu during use. The "teaching plan" can be viewed and searched through three navigation systems: click the "forward" and "backward" buttons to page through the pages; use the pop-up menu to index the chapters and jump to any text you need to find page.
2. Atlas is also displayed in pages. Click the "Forward" and "Back" buttons to turn the screen one page at a time; use the pop-up menu to index the chapters and jump to any picture you need to find; click the hot word in the "teaching plan" to immediately call the picture.
3. In view of the special importance of graphics in the anatomy course, the observation of images is designed in 7 different ways to meet the needs of teaching:
   （1） Preview (small window) mode;
   （2） Graphic (middle window with annotation) mode;
   （3） Full screen single image display (large window) mode;
   （4）Note mode of placement / cancellation;
   （5）Interactive dynamic annotation method (moving the mouse in the image, dynamically displaying the names of various parts in the text box, moving the mouse in the text note, the corresponding image parts will flash);
   （6）Local magnifier mode (local magnifier with mouse movement);
   （7）Drag the image to select the view mode.
4. Software setting lesson preparation / use the switch button, press the button to enter the lesson preparation state. At this time, you can edit all the text in the

courseware, add new annotations and related settings to the image, add, delete, modify, and update the settings Make hyperlinks and so on without having to program a line. Raise the button to enter the use state, at this time all lesson preparation functions are blocked. Different lesson preparation contents can be stored and recalled separately.

3.2 Overall System Design

This lesson preparation system is mainly composed of text and screen. The screen explains the specific parts of the image by sound, text, etc. By running the software to enter the main menu, the user can choose different modes, teaching mode or lesson preparation mode. Application, text, image, video interactive functions have been further enhanced. The main use process of the system is shown in the Fig. 3.:

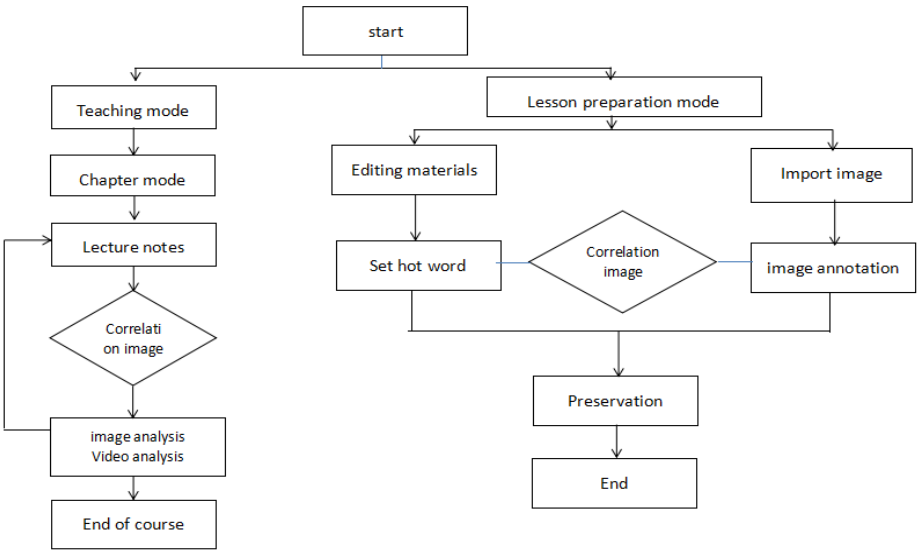


Fig. 3. System flow chart

3.3 Implementation Method

This system uses ToolBook multimedia development tools as the development environment, and adopts the object-oriented Openscript programming language. This language is used to write object scripts and is a shortcut for developing multimedia applications. ToolBook is equal to Windows documentation + SDK + C language programming + object-oriented programming .

This system uses the menued30.exe utility provided by ToolBook to redefine and organize menus according to the needs of anatomy teaching, and edit each menu item and corresponding drop-down menu. This utility program provides specific menu editing methods and help. After the menu editing is saved, you can select the menu through the "Choose Menu Bar" or use the OpenScript command to select the

required menu for each viewer. For the definition of the function of each menu sub-item, it can be given in Book Properties .

System Interface Implementation Method

1. The system uses "teaching plan" as the main window, and uses the record field for text entry. The "page" that can be expanded infinitely meets the needs of entering courseware content.
2. The system opens a " Atlas " subwindow in the main window to achieve the purpose of coexistence and separate management of the two books. The size and position of the sub-window can be adjusted as needed. Multimedia materials such as animations and videos are arranged in the "Atlas" in a "theatre" manner. Open a second subwindow in the main window, as a magnifying glass window.
3. The main interface of the entire program is based on the opening, hiding, and transformation of the above three windows, and its various related and derived functions are implemented by clever programming through object-oriented languages. As a lesson preparation system, since it is used directly for teachers and students, the program design is based on conciseness, easy learning and powerful functions. The "learning / use" state is switched, and the "author layer" provided by the programming environment is no longer used There is also no need for the user to write another line of program, and all are automatically completed by the designed program. The user only needs to press a button or two as needed to achieve the desired function.

Entry Methods For Scripts And Multimedia Materials:

1. Text entry: enter text directly in the text field;
2. Image input: paste image from windows clipboard, supporting multiple graphic modes;
3. Animation and video input: by linking video files;
4. Voice input: realized by linking voice files;
5. Hyperlink: open two "books" at the same time, create or find hot words, turn over the image or multimedia "theater", right-click the hot words, and the association will take effect automatically;
6. The chapter index and navigation system of this system can automatically retrieve the entered content and update the index without user intervention;The operation of the program is basically completed by mouse clicking. The program has designed a simple and clear graphic speedboat button, which makes the operation and use very simple.

4 Concluding remarks

The development of this platform helps teachers to actively control and use E-learning, fully realize the teaching auxiliary functions of E-learning, and mobilize teachers' personalized teaching. Although E-learning is an advanced modern education technology, it must not limit the initiative of teachers in teaching according to their aptitude, in person, in place, and in time. To the direction of intelligence, multi-function and platform, excellent courseware should contain a fairly rich multimedia material library, and the organization and management of these materials should be left to teachers as much choice as possible, E-learning The essence is computer-assisted rather than active teaching, which needs special consideration during the development of teaching assistant courseware.

References

1. Xiao Desheng, Wen Jiyu, Zhou Jianhua, et al. Reform and practice of open teaching model of medical morphology ［J］. China Journal of Modern Medicine, 2014, 24 (21): 110-112.
2. Cao Wenjun. Development and Application of Pathology Multimedia Teaching System. Department of Pathology, Capital Medical University, 2000, 61
3. Chen Bing. Construction of multimedia teaching platform for medical morphology based on MOOC. Basic Medical College of Wannan Medical College, 2017, 151
4. Cheng Kailiang, Zhang Mengchao, Ding Lei, et al. Standardizing teaching courseware and improving the quality of medical imaging teaching. China-Japan Friendship Hospital of Jilin University, 2018, 758

Author

Mengjiao Liu Master graduate student, Hebei North University ,Zhangjiakou,China. Main research direction : Modern educational technology.

Liheng Gong Lecturer, Hebei North University , Zhangjiakou, China, the main research direction z:medical informatics.

Xiao Zhang* Professor, master's tutor. Hebei North University , Zhangjiakou, China, the main research direction: medical informatics.

A Study on Learning Effectiveness and Satisfaction by Integrating Social Network Analysis into Cooperative Learning -A Case Study of Junior High School Physics and Chemistry in New Taipei City

[1]Wen-Chih Chang and [1]Jing-JingChang

Chung Hua University, Hsinchu, Taiwan
[1] Chung Hua University, Hsinchu, Taiwan earnest@g.chu.edu.tw

Abstract. Cooperative learning is an educational approach that makes students cooperate with each other to determine personal ability and fulfill team achievement. One of the most important thing of cooperative learning is how to divide student into proper teams. So that each team member can make contribution to team to improve the effect of cooperative learning. Therefore, good grouping is an important thing that must be considered before co-operative learning. Interpersonal relationships and academic performance have always been the most concerned things for students. Based on interpersonal relationships to apply heterogeneous grouping may improve team's performance and reduce the pressure caused by grouping. Consequently, the purpose of this study is to consider the differences in student learning result and satisfaction with interpersonal relationships when grouping. Taking the Science and Technology of junior high school as an example. Through the implementation of teaching activities, the test of learning results and the survey of satisfaction to understand the results and satisfaction of different grouping methods. And study the factors that affect learning performance of students as a reference for grouping in Science and Technology field. This study contains three classes, total 78 junior high school second grade students from Ming-de high school (complete school). The three classes adopt the S-heterogeneity grouping, the free heterogeneity grouping, the heterogeneity grouping of interpersonal relationships, and three different groupings to conduct cooperative learning teaching activities. The purpose of this study is to investigate whether the heterogeneous groupings combined with social network analysis and interpersonal relationships evaluation could improve students' learning effectiveness and student satisfaction with grouping.

Keywords: Cooperative Learning, Social Network Analysis, Grouping Methods.

1 Introduction

With the advent of the 12th year of the state religion. The way of education is new. The traditional way of telling students is not willing to learn. Insufficient interaction

between teachers and students. And students only know how to recite textbooks, practice test questions, and deal with exams. Lack of ability to think and communicate the ability to solve problems. In order to change this situation. Teachers need to take the student's life as the starting point in teaching. Emphasizing hands-on implementation. Let students pass grouping. In the process of group interaction. Learning communication and cooperation. The way to achieve the goal. Let students learn the ability to take away. Different from the past, the focus of learning has shifted to students as the main body. Causing students to learn motivation and hobby. Therefore, cooperative learning has been widely used in the teaching of various subjects. In order to facilitate teachers to effectively do a group work of cooperative learning. This shows the importance of grouping in cooperative learning. Cooperative learning has always emphasized group cooperation and heterogeneous grouping. Cooperative learning almost always uses heterogeneous groupings. And most studies have pointed out that heterogeneous grouping is more beneficial to learners. Can let high-capacity people to assist low-capacity groups. The low-capacity group can also contribute to the group. Get a sense of accomplishment when learning. But the interaction between people is complex. And interpersonal relationship is one of the most problematic problems for students. Incorporating social relationships into heterogeneous groupings may help improve interpersonal relationships. At the same time, it is also possible to make the team run more smoothly. Reduce the chance of conflict. In contrast, learning outcomes may also increase.

2 Related Works

The social network is connected by many nodes and nodes, Pattison [7] pointed out the type of social network is a collection of interconnected relationships between organizations or individuals throughout the social network. Social network is a very important point for Western social psychologists to study interpersonal interactions.

However, when people are in trouble, If you need help, you can also find people who can help according to the social network [1,2,3], find out key people and loner to understand the important factors that students have when choosing a partner, and for the students who need care, ask the instructors and coaches to pay more attention and care.

2.1 Social Network Analysis

Garton, Haythornthwaite and Wellman [4] believes that the units of social network analysis can be divided into three categories: Node, Relation, and Tie.

Node:. Points and lines constitute the basic elements of the basic elements of social network analysis diagrams. Points are actors. The actors here not only refer to specific individuals, but also to a group, company school, city, country, etc. or other collective society.

Relation: From the feature square, it can be divided into four parts: content, direction, strength, active or passive relationship.

Content: Refers to the reason why two points are linked. For example, two actors have relationships because of friends, couples, and travel together.

Direction: Direction can be divided into Undirected and directional. The relationship between two points is Undirected, expressed in a straight line. The student wants to find other students to become members of the group, or to seek help from others when there is a problem. When choosing another party for one party, it is directional, pointing to the student who wants to find an arrow, and when choosing each other, Expressed by double arrows.

Strength: It is the way the relationship between actors is, the strength criteria between the two are assessed. For example, in the relationship between classmates, the relationship and number of times of co-hosting activities, joint travel, and joint action can be used to measure.

Active and passive relationship: The relationship between the two actors is generated, and there will be an active and passive relationship. In the social network diagram, an actor is pointed by many arrows to indicate a passive relationship. If the actor sends out many arrows, it is a more active relationship.

Tie: . Can be used to explain the combination of multiple relationships between actors and actors. Granovetter (1973) argues that the strength of the contact is the frequency of contact as an indicator.

2.2 Cooperative Learning

Collaboration refers to a work in which two or more individuals or groups work together for a common goal. Hilke [5] also believes that cooperative learning is an organized structure, the process of students working through groups. Learning together, in addition to contributing to their own abilities, also assist others to complete their work. Cooperation can also promote students' relationships, communication skills and high-level thinking skills. It can also foster students' ability to cooperate, enhance inter- group relationships, expand students' self-esteem and promote academic achievement. Parker [6] Cooperative Learning provides a collaborative learning environment in a classroom learning environment where students can learn from peers in heterogeneous groups, support each other, criticize or share each other's perspectives, and finally share the results. In a cooperative learning environment, it is subtle to foster more cooperative behavior.

3 Research Method

This paper explores the impact of different grouping methods in cooperative learning on the effectiveness of natural and life science and technology learning in the eighth grade of the country. It adopts three grouping methods, which are grouped by S-type heterogeneity grouping, free heterogeneity grouping, and interpersonal relationship heterogeneity grouping. Pre-testing and post-testing data analysis to understand the learning outcomes of the three different grouping methods and to explore

the reasons that affect their learning effectiveness. This study controls the variables as teaching methods, teaching time, teaching content, teaching environment, teachers, and former Measurement, post-test, self-variation is grouping, one group is S-type heterogeneous grouping, one group is free heterogeneity grouping, one group is considering heterogeneity grouping of interpersonal relationship, and the variable item is learning effect and group satisfaction. .

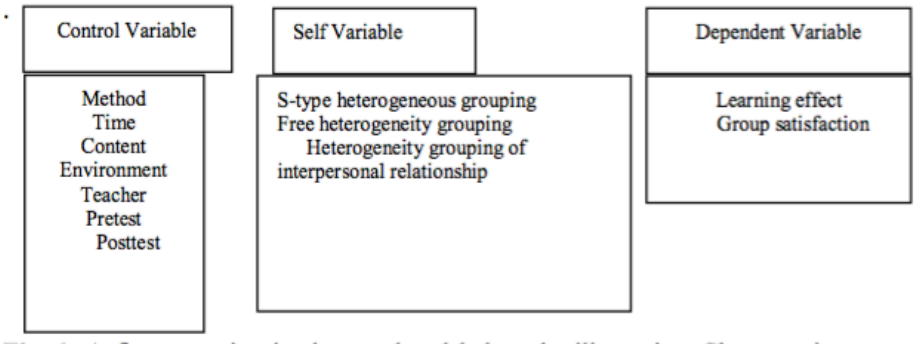


Fig. 1. A figure caption is always placed below the illustration. Short captions are centered, while long **ones** are justified. The macro button chooses the correct format automatically.

First, conduct relevant literature collection and analysis, search for past research, prepare research plans and revisions after determining research topics, and prepare for experiments (including preparation of teaching plans, group questionnaires, pre-tests, post-tests, expert assessments, venue time). Arrangement, study consent, group satisfaction questionnaire, etc., and then pre-test, group questionnaire, teaching experiment operation, post-test, group satisfaction questionnaire, data analysis and writing.

In this study, the second-level students who are taught by the researchers are divided into three groups according to the class. They are divided into different heterogeneous groupings and considering interpersonal heterogeneity grouping for cooperative teaching. The main purpose is to explore social networks. Analyze whether it can improve the effectiveness of grouping and the learning outcomes of students and the satisfaction and feelings of students when grouping into groups.

In the eight-year class A, the S-type heterogeneity grouping is divided into 25 people, 16 boys and 9 girls. The S-type heterogeneity group (Table 1) is arranged by the teacher. There is no free will. A group of 5 people will have 2 The heterogeneity group of the middle degree group and the 1 low degree group in the high degree group should be the same as the male and female ratios, because the girls are 9 people, there are 4 groups of boys, 3 girls and 2 girls, a group of boys, 4 girls and 1 person.

Table 1. Heterogeneity Grouping

Group 1	Group 2	Group 3	Group 4	Group 5
S1	S2	S3	S4	S5
S10	S9	S8	S7	S6
S11	S12	S13	S14	S15
S20	S19	S18	S17	S16
S21	S22	S23	S24	S25

As a result of the eight-year A class grouping according to the S-type heterogeneity, the average strength of each group is equivalent, and the male-female ratio is equivalent. Only the fifth group of one girl, the rest are two girls.

The eight-year B class adopts a free heterogeneity grouping. This group allows students to freely group, but must meet two groups of high-level groups and one low-level group. As shown in the figure, there are two groups of boys and 4 girls. People, two groups of boys, 3 girls, 2 people, a group of 4 boys and 1 girl, can make the number of girls equal, the researchers will give a high degree group green color card before the grouping, medium degree group blue color card, low degree group Red color card, which requires a group of two green cards, two blue cards and one red card, and is equal to the number of men and women, allowing students to freely group according to the rules.

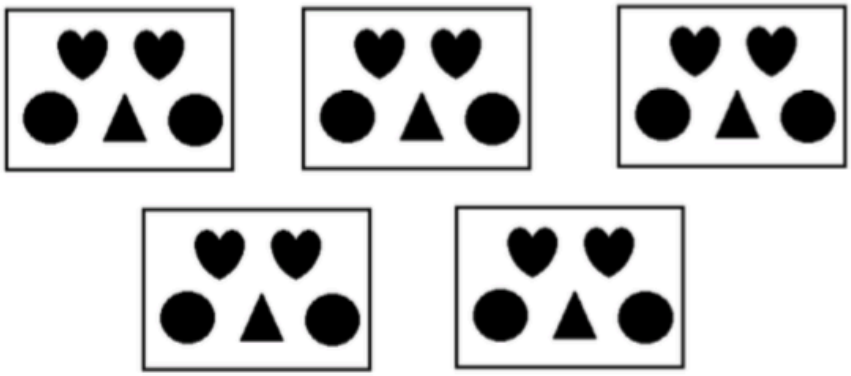


Fig. 2. B class applied freewill for grouping, and high score group ♥, middle score group● and low score group▲.

In the eight-year B class, according to the results of the grouping of free heterogeneity, the average strength of each group is equivalent, and the ratio of male to female is equal. Only the fourth group is a girl, and the other group has two girls.

We applied betweenness centrality to find out the key person. The betweenness centrality of a node v is given by the expression:

$$g(v) = \sum_{s \neq v \neq t} {}^{\sigma_{st}(v)} (1)\, \sigma_{st}$$

Where σ_{st} is the total number of shortest paths from node s to node t and $\sigma_{st}$is the number of those paths that pass through v.

Considering the heterogeneous grouping of interpersonal relationships, it is necessary to use the grouping willingness questionnaire to find the combination of the high degree group and the low level group, and then the middle degree group joins the group according to the free will. The order is as follows: 1. Confirm the length of each group 2. Match the appropriate team leader. The combination with the low degree group 3. Determine the high degree group two and the low degree group 4. The middle degree group is added according to the free will, as described in detail below.

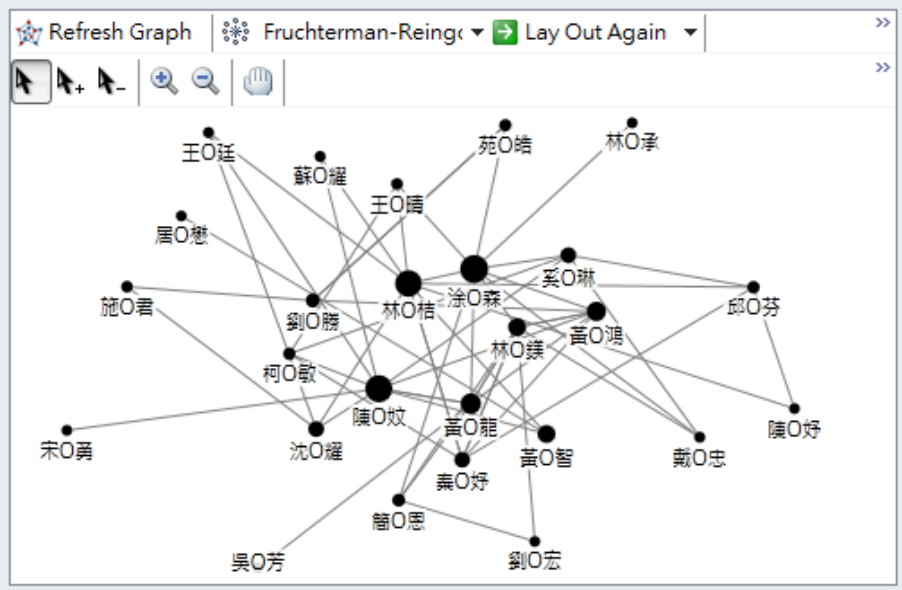


Fig. 3. A figure caption is always placed below the illustration. Short captions are **centered**, while long ones are justified. The macro button chooses the correct format automatically.

Topic 1 (confirmation of the team leader) is the natural class, which group do you most want to write, and why? Then use social network analysis to find the most popular people in the class, find the class with the largest origin in the picture. The larger the origin, the greater the relevance.

Topic 2 (with the appropriate team leader and low ability group) requires a high degree of assistance in the classroom and one-to-one teaching to follow the progress of the teaching, so how to find a high degree group that is willing to teach them and have patience for the low degree group It seems important, so the next step is to find the right team leader and low degree group to make each other work smoothly.

Topic 3 (Determining the combination of two high-level groups and low-level groups) After identifying the team leader and the low-level group, you need to match a high- level group assistant leader. When the leader has problems, you can work with

another high-level group. Discussing the problem together, another high-level group can also teach the low-level group together.

Topic 4 (Chinese ability group joins according to free will) In order to make the grouping smoother, this questionnaire is not open and the teacher assigns a low degree group to the group. Therefore, there are fewer group disputes between students and human pressure, which can also make Low-level groups don't have to feel unwelcome when they are grouped, but they can also match their team leader to alleviate the pressure of low-level groups in heterogeneous grouping.

In the eight-year class C, the average strength of each group is equal, and the ratio of male to female is equal. Only the fifth group is a girl, and the other group has two girls. Only the fourth group has six members, and the rest are five-person groups.

4 Results and Discussion

In this section, present mainly results of study and discussion about the differences in student learning effectiveness and satisfaction with interpersonal relationships when grouping. Taking the Science and Technology course of junior high school as an example.

Table 2. Learning effectiveness analysis

	Test	N	Mean	t test	test interpretation
A Class **the S-heterogeneity grouping**	Pre-test	25	36	0.00*	significant
	Post-test		68.6		
B Class **the free heterogeneity grouping**	Pre-test	27	35.4	0.00*	significant
	Post-test		74.5		
C Class **the heterogeneity grouping of**	Pre-test	25	35.2	0.00*	significant
	Post-test		75.8		
interpersonal relationships					

According to Table 2, we can find that both of three class square scored an average of 35 points higher after cooperative learning. In pair sample t test, all the p-value are less than 0.0001, pointing out that the values are strongly significant.

The studies indicated that cooperative learning and the heterogeneity grouping for students' domain knowledge is useful also the heterogeneity grouping is suitable for cooperative learning.

Additionally, we want to know students' satisfaction with grouping. we design questionnaire to deal with it, after that use the radar chart to show the differences

when grouping (see fig.5.). Adopting the heterogeneity grouping of interpersonal relationships ones' average score is 4.3 point higher than others.

C class has well interpersonal interaction with each other and have sense of teamwork spirit, especially at item of group member will help me. In this experiment, further shows that the heterogeneity grouping interpersonal relationships works well.

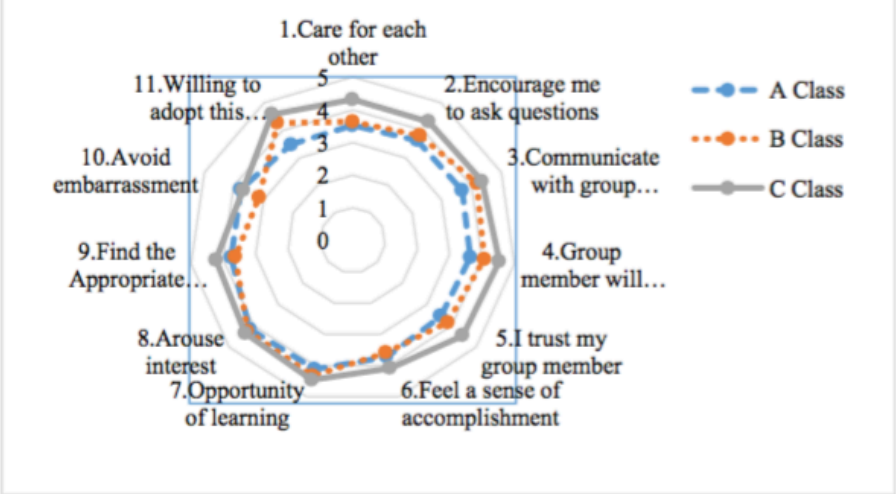


Fig. 4. The satisfaction with A Class, B Class and C Class 5 Conclusions and future recommendations.

The purpose of this study is to explore the impact of collaborative learning on learning outcomes and group satisfaction in different groupings, and to understand the learning situation of students with different abilities in different groups. It is hoped that this will be used for teachers' teaching and future researchers' research. This chapter is divided into two sections, the first section is the conclusion, the second section is the recommendation.

5.1 Conclusion

In order to understand the interpersonal relationship when considering the cooperative learning group, whether it can improve the learning effect, after the analysis and discussion of the research results based on the research results, the following conclusions are obtained.

Cooperative learning uses S-type heterogeneity grouping, Free heterogeneity grouping and Consider interpersonal heterogeneity grouping, all of which have significant learning outcomes.

Cooperative learning uses Consider interpersonal heterogeneity grouping, group satisfaction is higher than S-type heterogeneity grouping and free heterogeneity grouping.

5.2 Recommendations

This study aims to cooperatively learn grouping strategies, add interpersonal factors to grouping strategies, try to find a better heterogeneity grouping, improve learning outcomes and group satisfaction. This study only uses the three questions in the questionnaire as the basis for grouping. Therefore, the topic of the future questionnaire can be designed more to better understand the class social network, and it may also make the grouping effect better.

In order to better understand the situation of group satisfaction, we can increase the interview in the future to understand which part of the grouping method is more satisfactory for students. Which part is the trouble that students use after using this grouping method, and also correct the grouping method to find better heterogeneity.

References

1. Chang Wen-Chih, Wang Te-Hua and Huang Yen-Ching ,Exploring the distribution and relationship between MSN and the Physical World , Networked Computing and Advanced Information Management (NCM), 2010 Sixth International Conference, 716-721
2. Chang Wen-Chih, Chien Yi-Yang, "Learning achievements and after-school friendship interpersonal of social network analysis", Information Science and Digital Content Technology (ICIDT) 2012 8th International Conference on, vol. 3, pp. 604-609, 2012.
3. Chang Wen-Chih, Wu Lai-Chi and Lin Hsiao-Wen, Applied Social Network Analysis to Projext Curriculum, Networked Computing and Advanced Information Management (NCM), 2010 Sixth International Conference on 16-18 Aug. 2010, 710-715
4. Garton, L, Haythornthwaite C, and Wellman, B.(1997). “Studying Online Social Networks”, Journal of Computer-Mediated Communication, 3(1), June 1997.
5. Hilke,E.V.(1990).Cooperativelearning.Bloomington,IND:PhiDeltaKappa educational
6. Parker , R.E. (1985) .Small-group cooperative learning-- improving academic, social gains in the classroom.NassBulletin,69(479),PP.48-57.
7. Pattison, P.(1993).Algebraic Models for Social Networks, Cambridge University Press.

Study of the Medical Image Sharing System Based on a RBAC Expansion Model

Liheng Gong[1] , Mengjiao Liu[1]and Xiao Zhang[1*]

[1]School of Information Science and Engineering, Hebei North University,07500 Zhangjiakou, Hebei, China
1436798605@qq.com

Abstract: Currently the PACS systems are distributed in various hospitals,the interconnection of data cannot be achieved.This paper proposes a RBAC extension model which supports the request-authorization.In this extended model, the concept of authorization role and other related concepts are creatively presented, a new access control system scheme is proposed and was initially applied to the image sharing system. Moreover, according to the demand of data sharing system, this paper designs the feature extraction module of image data.

Keywords: medical, RBAC, image sharing, feature extraction

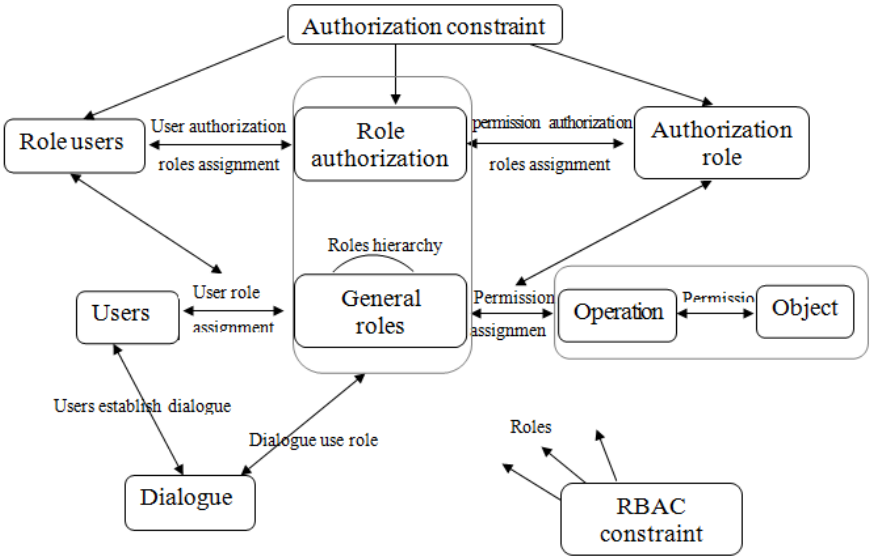
Introduction

With a vast territory and large population, China has great demands for medical resources. In recent years, the medical technology has made astounding advances and medical resources are enriched to a certain extent, but there are still many problems: the accurate clinical diagnosis is benefited from the auxiliary diagnosis of medical equipment, but the high-end image equipment is expensive, costing several millions or even up to tens of millions yuan, so that not all medical institutions are able to buy the needed equipment; most of the medical image equipment can produce hundreds or even tens of hundreds of images by one scan, while the films took by patients are only a few of them, which are unavailable for parameter adjustment and three-dimensional dynamic display, so the value of diagnosis is given a discount, and patients need to accept repeated examinations after they are transferred to a different hospital as the

original image data cannot be acquired, thus increasing the medical burden of patients. In view of the above problems, we plan to propose an image information sharing method among different institutions, so as to integrate the access control authorizations to image information system among different medical institutions.

1 RBAC model supporting user request and authorization used in the method

Firstly, the method makes improvement based on the existing RBAC [1]model and increases the request and authorization ways. The RBAC model supporting user request and authorization is established as shown in the above Figure. Different from the general RBAC model, the General Roles in the RBAC model supporting request and authorization have role hierarchy[2], while the Authorization Roles do not have roles hierarchy. Through the roles assignment, every user in the medical institute has role, and the users with roles form the Role Users; then through permission-role assignment, the role users acquire the permission possessed by the roles and thus the Role Permission is formed. For the role users of some hierarchies[3], they can be given power through assignment to create the authorization roles, so that the role permission of role users can be assigned to the authorization roles through permission authorization roles and then assigned to the requesters through user-authorization roles, and all RBAC constraints defined by the system managers shall be followed in the assignment of permission authorization roles and the assignment of user authorization roles[4]



.

The RBAC model supporting user request and authorization is defined:

The user, permission, general role, conversation and RBAC constraint are in total accordance with RBAC definition.

Role user: the users with roles. Only role users can become requesters and only role users are qualified to become trustees[5].

Role permission: the permission assigned to roles. The requesters can only request the role permission possessed by the requesters, and all permission of the role can be represented by the role name, called as Role Permission. No real permission is stored in the role permission, but Role Permission will be automatically formed in operation according to the actual permission of the role[6]. The Role Permission can take part in the request and authorization as a whole.

Authorization role: created under the request of requesters and by the specific users with creation permission. It has the starting time attribute, end time attribute and activation attribute of the life cycle. Only when the authorization role stays in the life cycle and is activated, can it use its permission. Only the specific users with the permission of creating authorization roles can manage the authorization roles; of course, the system manager can also be set to manage and change the activation attribute of authorization roles in unification.

Permission-role-assignment (PA): a many-to-many relationship of permission assignment and conventional roles, PA belongs to P×GP.

Users-role-assignment (UA): a many-to-many relationship to assign users to general roles, UA belongs to U×GP.

Roles-hierarchy (RH): the division of different hierarchies in a role, RH belongs to GR×GR and is a partial ordering relation of GR.

Permission-authorization-roles-assignment (PAA): a many-to-many relationship to assign the permission of role users to the authorization roles, and the permission assigned to authorization roles is different from the general roles in use, and they can only be used after it's examined that requesters have such permission, PAA belongs to RP×AR.

User-authorization-roles-assignment (UAA): a many-to-many relationship to assign the role users to the authorization roles, UAA belongs to RU×AR.

Roles (R): R=GR×AR.

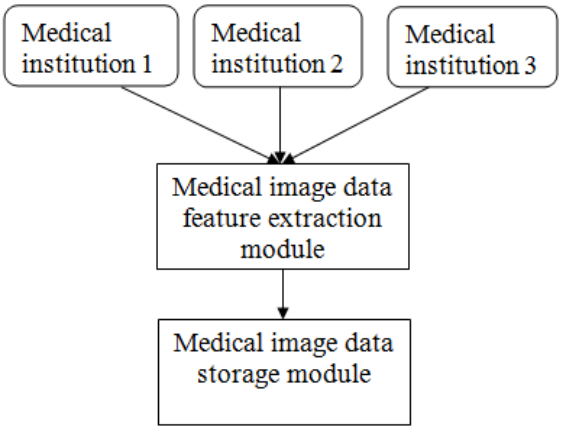
Authorization Constraints: the set of authorization constraints. The authorization constraints are the constraints formulated by medical management organizations aiming at the authorization operations to various requests, used to regulate the authorization behaviors of those requesting authorization.

Authorization Role Constraints: the set of authorization role constraints. The authorization role constraints are the constraints especially created for authorization roles, used to regulate the conditions for authorizers to use permission.

The improved RBAC model adds a request-authorization module based on the traditional RBAC model. The module is operated and managed together by the requesters, authorizers and system managers. The request users can request to create an authorization user and assign certain permission to the authorization user, and then the users with authorization permission decide whether to assign the role to the requester; if it is agreed, then the authorization user is created and the requested permission is

assigned, to complete the request and authorization operations. The authorization user and its permission are managed by the users with authorization permission, and the safety manager can regulate the behavior of requesting authorization through creating overall authorization constraints.

2 The chart of this medical sharing system



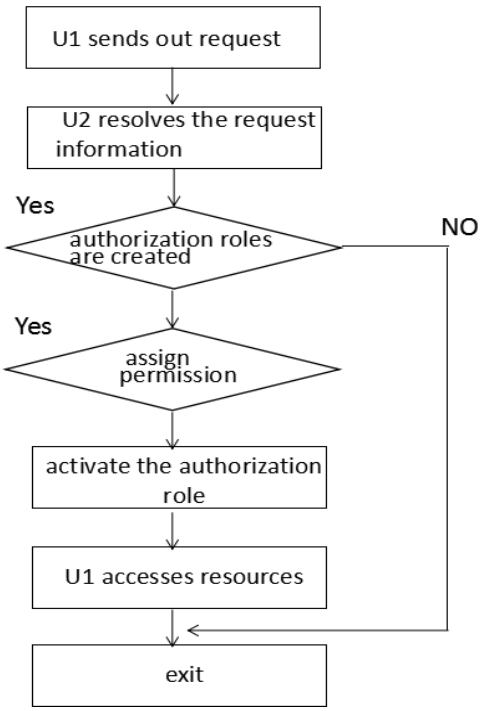
As shown in the above Figure, the system includes medical institution module and medical image data sharing module. The medical institution module is modeled based on the Role-Based Access Control (RBAC) model supporting authorization request and it can include several hospital sub-modules, each of which corresponds to a hospital in reality; the medical image data sharing module includes the feature extraction sub-module and storage sub-module. The module receives the medial image data sent by the medical institution module; also for the subsequent inquiry of data, feature extraction and feature marking are conducted to each medical image data record, and then the data records can be stored in the medical image data storage sub-module after marking, for the inquiry of various hospital sub-modules in the medical institution module.

3 Process of access control method based on the improved RBAC model

Generally, the basic idea of RBAC model is to divide roles according to the different tasks to be completed, the access permission of the data is packaged into different roles, and users can indirectly realize the access to data through role assignment[7].

When resource sharing needs to be realized among different medical institutions, it shall be based on the setting of RBAC model supporting authorization request[8], and also different authorization permission needs to be assigned to different users according

to needs. When the RBAC model supporting authorization request is applied for operation, the specific authorization process can be:



(1) The user U1 of the first institution sends out request for the assignment of a role in the second institution and some permission, and the life cycle information of the authorization role can also be included.

(2) After the second institution receives the request and resolves the request information, it will analyze the users with authorization permission corresponding to the request information, and then transfer the request to the users with such authorization permission[9].

(3) The users with authorization permission review the request and decide whether to give such authorization. If the authorization is agreed, the authorization roles are created, and their permission and life cycle are assigned.

(4) After the system of the second institution makes feedback of the authorization information to the user U1 of the first institution, the authorization role will be activated at the starting time of its life cycle.

(5) After the user U1 of the first institution receives the authorization information, it can make relevant operations through the assigned permission after the starting time of its life cycle.

(6) After the end time of the life cycle arrives, the system manager of the second institution cancels the authorization user and notifies the first institution[10].

Of course, in such process, all the constraint conditions have to be met, including authorization role constraint, authorization constraint and other relevant RBAC constraints[3], etc.

The permission requested by the first institution can be all permission of a role in the second institution, such as that of the attending doctor; or it can be part permission of a role in the second institution, such as the issuance and reservation rights of examination list[11]. Similarly, the second institution can agree the whole request or support the fine grit authorization of part permission, i.e. the authorizer can only assign the necessary permission to complete specific functions to the requester, to conform to the principle of least privilege[12].

4 Medical image data feature extraction sub-module

The medical image data feature extraction sub-module can include: character feature extraction sub-module and image feature extraction sub-module.

4.1 The character feature extraction sub-module

The module can extract character features from the character information in the medical image data. For example, the medical image data of an X-ray image can include the following character description, for example, organ name – lung, image description information – marking increase, doctor diagnosis information – pneumonia. From the above information, the character features that can be extracted from the data record are: lung, marking, pneumonia[13].

4.2 The image feature extraction sub-module

The module can extract image features from the image information in the medical image data. The image feature extraction sub-module includes: segmentation sub-module, to segment the image into several sub-blocks; feature extraction and vector generation sub-module, to extract the gray level and texture features of various sub-blocks and form into vector; cluster and shape feature extraction sub-module[14], to make statistical clustering with K mean value method so as to divide the image into several areas, and extract the shape feature of the areas; feature vector generation sub-module, to combine the cluster center vector and shape feature vector into the feature vector expressing various image areas.

Firstly, the received image data is segmented, the image is divided into several sub-blocks, and the features of the sub-blocks are obtained. The selection of sub-block size shall not only maintain the efficiency of textures and guarantee the acceptable calculation time. If the sub-block size is small, the maintained texture information is fine, but the calculation time is correspondingly added; otherwise, if the sub-block is large, the calculation time is reduced, but certain texture information is lost and the image segmentation effect is poor. For example, the selected sub-block includes 4×4 pixels; if the image size is 256×256, then each image will have 4096 feature vectors; each feature vector fi is composed of 4 features, i.e. $fi \in R4, 1 \leqq i \leqq 4096$, R denotes the segmented images. The first feature is the average gray of the sub-block, while the other three features are the WT HF sub-band energy[15].

To obtain the three wavelet moments, Daubechies-4 WT should be firstly conducted to the image sub-block. After each level of WT, a 4×4 sub-block is decomposed into four frequency bands: LL, LH, HL and HH. 2×2 coefficients are included in each

frequency band, and the features of four frequency bands can be obtained with the current band feature calculation method. The objective of using wavelet HF sub-bands lies in that, they can reflect the texture feature of images, the coefficient matrix in various frequency bands of wavelet can express the image textures efficiently, and the different frequency band coefficients reflect the change of textures in different directions. For example, HL frequency band reflects the activities in horizontal direction. If there are vertical stripes in the image, HL frequency band reflects the high energy, while LH frequency band reflects low energy.

The feature vectors are taken as the training set, and K mean value is applied to make clustering, with each cluster corresponding to one segmented area of the image. That is to say, there is a feature vector set F={fi∈R4, 1≦i≦4096} in each image, and F is divided into C groups {F1, ……, FC}, respectively corresponding to the C areas of the segmented image. As clustering is conducted in the feature space, it is not necessary for the sub-blocks in each cluster to form connected areas of the image[16]. K mean value method adjusts the quantity of clusters C self-adaptively according to whether the loop termination conditions are satisfied.

After the image segmentation, three extra features are extracted to the segmented areas to describe the shape feature of each area, and they can be 1-3 order normalized inertia. For the area Ri in an image plane, r order normalized inertia is a limited set and its calculation method can be:

$$I_{(R,r)} = \frac{\sum[(X - \bar{X})^2 + (Y - \bar{Y})^2]^{r/2}}{V(R_i)^{1+r/2}} \qquad \text{Formula (1)}$$

Wherein, (X, Y) is a coordinate in the image area Ri, $(\bar{X}, \bar{Y})$ is the centroid of Ri, and V(Ri) is the volume of Ri, but for two-dimensional image, V(Ri) is the area. The relevant size zooming and position rotation of normalized inertia are invariants. The minimum normalized centroid is obtained through the calculation to unit ball. The r-order normalized inertia of unit ball is denoted as Ir, i.e.:

$$\bar{h}_i = \left[\frac{I_{(R_i,1)}}{I_1}, \frac{I_{(R_i,2)}}{I_2}, \frac{I_{(R_i,3)}}{I3} \right]^T \qquad \text{Formula (2)}$$

Thus, the area feature Fj (1≦j≦C) of the image is respectively composed of expression area gray, central clustering vector of texture and the inertia vector of expression area shape.

5 Medical image data storage sub-module

The module can be separate data storage unit or the distributed data storage system. We can set the storage strategies of medical image data according to the actual demands.
The various hospital sub-modules in the medical institution module can set no local medical image storage, but store all the medical image data in the medical image data

sharing module; also the medical image database of small capacity can be set; but for the safety and convenient sharing of the data, even the hospital sub-module setting local image database needs to send all data to the medical image data sharing module for storage, for example: in standby unattended time, all data will be sent to the sharing module with day as the cycle. In this case, when the other hospital sub-modules need to retrieve the medical image data, they can visit the sharing module directly but do not need to traverse all hospital sub-modules for search.

6 Conclusion

The system stated in the paper can realize the objectives of sharing advanced diagnosis machines and excellent doctors；, storing the medical image data safely and completely, and using the stored medical image data efficiently, so as to implement the hospital functions effectively, improve the patient satisfaction and strengthen the cooperation among medical institutions.

References

1. Moon Sun Shin, Heung Seok Jeon, Yong Wan Ju, Bum Ju Lee, and Seon-Phil Jeong, "Constructing RBAC Based Security Model in u-Healthcare Service Platform," The Scientific World Journal, vol. 2015, Article ID 937914, 13 pages, 2015. doi:10.1155/2015/937914
2. Chang-Joo Moon,Dae-Ha Park,Soung-Jin Park,Doo-Kwon Baik. Symmetric RBAC model that takes the separation of duty and role hierarchies into consideration.[J]. Computers & Security,2004,23:.
3. Sun, Yuqing,Wang, Qihua,Li, Ninghui,Bertino, Elisa,Atallah, Mikhail. On the Complexity of Authorization in RBAC under Qualification and Security Constraints[J]. IEEE Transactions on Dependable and Secure Computing,2011,86.
4. Information Technology - Information and Data Management; Research Conducted at China Medical University and Hospital Has Provided New Information about Information and Data Management (CBR-based negotiation RBAC model for enhancing ubiquitous resources management)[J]. Computers, Networks & Communications,2017.
5. Jeffrey Fischer,Daniel Marino,Rupak Majumdar,Todd Millstein. Fine-Grained Access Control with Object-Sensitive Roles[M].Springer Berlin Heidelberg:2009-06-15.
6. Subhendu Aich,Samrat Mondal,Shamik Sural,Arun Kumar Majumdar. Role Based Access Control with Spatiotemporal Context for Mobile Applications[M].Springer Berlin Heidelberg:2009-06-15.
7. Lee Hung-Chang,Chang Shih-Hsin. RBAC-Matrix-Based EMR Right Management System to Improve HIPAA Compliance.[J]. Journal of Medical Systems,2012.
8. Peleg Mor,Beimel Dizza,Dori Dov,Denekamp Yaron. Situation-Based Access Control: privacy management via modeling of patient data access scenarios.[J]. Journal of Biomedical Informatics,2008,41(6).

9. Moon Sun Shin,Heung Seok Jeon,Yong Wan Ju,Bum Ju Lee,Seon-Phil Jeong,Jong-Hyuk Park. Constructing RBAC Based Security Model in u-Healthcare Service Platform[J]. The Scientific World Journal,2015.
10. Solomon Berhe,Steven Demurjian,Thomas Agresta. Emerging Trends in Health Care Delivery: Towards Collaborative Security for NIST RBAC[M].Springer Berlin Heidelberg:2009-06-15.
11. Cátia Santos-Pereira,Alexandre B. Augusto,Ricardo Cruz-Correia,Manuel E. Correia. A Secure RBAC Mobile Agent Model for Healthcare Institutions - Preliminary Study[M].Springer Berlin Heidelberg:2013-06-15.
12. Jim Longstaff. Extending Attribute Based Access Control to Facilitate Trust in eHealth and Other Applications[M].Springer Berlin Heidelberg:2013-06-15.
13. Hong Wu,Lele He. Combining visual and textual features for medical image modality classification with ℓ p −norm multiple kernel learning[J]. Neurocomputing,2015,147.
14. Kai Xiao,A. Lei Liang,Hai Bing Guan,Aboul Ella Hassanien. Extraction and application of deformation-based feature in medical images[J]. Neurocomputing,2013,120.
15. Pedro P. Rebouças Filho,Elizângela de S. Rebouças,Leandro B. Marinho,Róger M. Sarmento,João Manuel R.S. Tavares,Victor Hugo C. de Albuquerque. Analysis of human tissue densities: A new approach to extract features from medical images[J]. Pattern Recognition Letters,2017,94.
16. P. Jaganathan,R. Kuppuchamy. A threshold fuzzy entropy based feature selection for medical database classification[J]. Computers in Biology and Medicine,2013,43(12).

Quality traceability system for multi-station SMT Manufacturing Process

Cheng-Hui Chen [1], De-Wei Hsieh [1], Ci-Hua Wu [1], Ci-Yi Lai [1], Chi-Chin Hsieh [1]

[1] Central Industry Research & Service Division (CID), Institute for Information Industry, Nantou, 540, Taiwan

Abstract. Yield is a key indicator in the SMT manufacturing process. To solve the traditional experience-oriented quality control method, we put forward a quality traceability system for SMT. The system includes two parts: firstly, the carrier identification and estimation module, which is used to solve the problem that the production resume cannot be established because of costs and the bar code equipment cannot be set up in the continuous production environment; secondly the quality traceability system, which is used to track the quality under the interaction of more than 300 factors, such as steel plate, solder paste quality, scraper speed, suction nozzle pressure and temperature condition of each area of the oven. In fact, the system has been operating in the actual production line of SMT, and assisting in the tracking of defective products in the manufacturing process.

Keywords: Defect-causing trace in SMT manufacturing, SMT, Quality tracking, Production resume, Imbalanced data

1 Introduction

SMT (Surface Mounter Technology) has always been an important technology in the manufacturing of electronic components. Mostly fully automated production with large output and low gross margin is adopted. Therefore, yield has always been a key indicator in the SMT manufacturing process. In the past, when defective products occurred, in most cases it took 1 or 2 days to perform time & labor consuming manual tracking analysis by relying on the experience of old masters and by judgment on materials, process formulation, equipment status and product types. In addition, this needs a process of long time and knowledge training, and experience inheritance is also a major concern of the SMT production line at present.

The production equipment order in SMT production line is printing, mounter, reflow and AOI, in which AOI is the final quality inspection results, and mounter and reflow vary according to product requirements. There will be 1-3 units in a production line, each equipment will be set according to the product formula. To solve the traditional experience-oriented quality control method, we put forward a quality traceability system for SMT. The system includes two parts: firstly, the carrier identification and estimation module, which is used to solve the problem that the production resume cannot be established because of costs and the bar code equipment cannot be set up in the

continuous production environment; secondly the quality traceability system, which is used to track the quality under the interaction of more than 300 factors, such as steel plate, solder paste quality, scraper speed, suction nozzle pressure and temperature condition of each area of the oven. Moreover, because the yield is usually as high as 95% or more, the data imbalance degree is high, and the algorithm selection also needs to be considered. There are a variety of methods for experimental evaluation, and best results are found out from them. In this paper, Random Forest (RF)-feature importance is used as the feature screening method, RF with traceable algorithmic process is still selected for modeling establishment, and the forest is disassembled to calculate the rules and weights.

In this paper, the quality traceability system will be established, and the currently widely used SMT continuous process can be applied. This paper will describe relevant literature for reference, each module method, discussion of the application in the actual production line and conclusion. In addition, this method has also been applied to the actual SMT production line, and is used to assist in the tracking of defective products in the manufacturing process.

2 Related Works

SMT is the technology that circuit boards of many electronic products need to use, which is mainly divided into the three-station process, i.e. solder printing, mounter and reflow, the pressure and thickness of the solder paste scraper in the first-station process cause defects of PCB board, the best solder paste thickness tolerance is 5.5-8.5 mil [12], so it is necessary to optimize the height of solder paste and to keep solder paste in good condition [4]. The second station is mounting. The use of mounting is to place components on the PCB board. Common problems are component missing, offset, component damage, etc. The occurrence of component missing is due to that insufficient suction nozzle pressure of the mounter and unfulfillment of material feeding by the component feeding frame cause component missing. The occurrence of offset is due to that when the mounter parameters are set, incorrectness of X-axis and Y-axis coordinates of the mounter causes the offset of component placement [5], while the occurrence of component damage is due to that incorrectness of Z-axis coordinate causes the component deformation from component extrusion at the time of component placement, therefore, the amount of solder paste on the template is the key in determining whether components can be placed on the PCB board smoothly. The third station is reflow soldering. Reflow soldering is to heat the solder paste to the melting point and to be joined with components. Therefore, solder paste factors, such as unsealing time, viscosity, oxidation, printing temperature, melting point temperature and so on, are all key factors [1][7]. In addition, in terms of reflow oven equipment, temperature rise curve, solder-dipping temperature, wetting time and solder melting time [3][9] are also one of the factors leading to defective products.

3 Methodology

SMT belongs to fully automatic production equipment with high-speed production efficiency, but there are two major problems for on-site quality control, first of all, there is indeed a great fault in the correspondence between equipment process data collection and products, the production data and products cannot effectively correspond to each other, the second is that the judgment on quality problem needs to be made on the basis of the experience of senior personnel. Therefore, to cope with the above-mentioned problems, the production resume module and the quality tracking analysis module are constructed to solve the current problems. The framework adopted in this system is as shown in the figure, and various elements are described as follows:

- Barcode Reader: In order to increase the accuracy of data correspond, the system is installed two sets of barcode reader to solve the incomplete data collect and overlapping between stations in the existing data corresponding way.
- Edge Server: To reduce the server load and to cope with the large field of working area, the edge server is used to collect data from the barcode readers and transmit the data to the data center.
- MES: The Manufacturing Execution System(MES) which is existing in the field can provide the parameters in each workstation, AOI results and etc. as analysis content.
- Data Center: As the data collection and transfer of the elements, with working hour pushback module, production history corresponding module, and connection in series of formulas, production data and quality inspection results.
- Database: Storage of integrated production resume data and algorithm analysis results.
- Web Server: The interface from which users view production resumes and analysis results, adopting by RWD and with the function of sending periodic reports.
- Algorithm Operator: Five analysis modules are set up, namely Defective Product Fuzzy Trace-back Module, Quality Relationship Mining, Single Factor Analysis, NG Relevance Ranking and Commonality Analysis, to help users to find out the possible parameter correlation, causes and commonality of product defects by means of data.

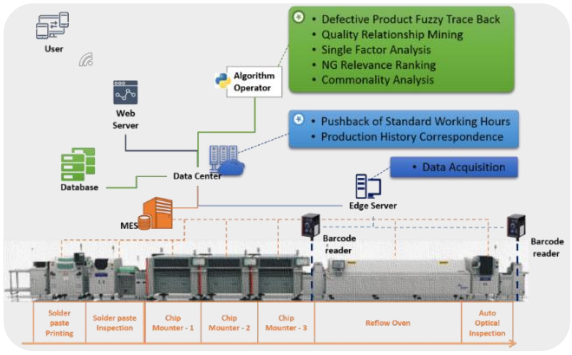


Fig. 1. Quality traceability system architecture

3.1 Product process module

In the production process of the SMT production line, each machine has an Edge Computer to record and to store the production process data. However, some key data, such as work orders, materials, personnel working hours, machines and finished product serial numbers, really cannot be inquired in real time. Linking these data into production history records lacks the establishment of the coding for unique serial numbers of produced objects, in addition, continuous equipment being used for continuous production leads to that barcode readers cannot be installed at inlets and outlets of each station, which are the difficulty points of this module.

Through the import of sensing element barcode, the manufacturing process information including key personnel, machines, materials and methods, etc. is recorded, such information is integrated with the MES system information, so that complete data about the production process can be established in real time to achieve the purpose of accurate information query. In this method, two barcode readers are installed after the mounter and AOI equipment in the SMT production line. When the carrier leaves the equipment from the conveyor belt after the completion of production, the barcode reader scans the barcode. Through the label and time record information, the mounter and AOI information can be traced back, and then through the product process module, the production data of each station can be inquired from each carrier.

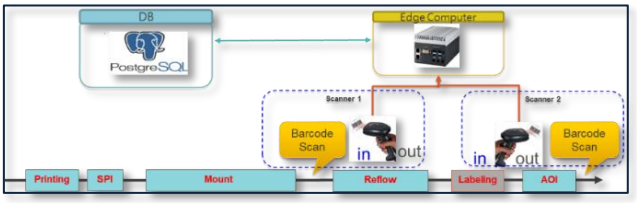


Fig. 2. Barcode scanner system architecture

The product process module makes the correspondence relationship with production data mainly through the barcode ID and time scanned by two barcode readers, its process steps are as follows:

A. Make consolidated treatment of the barcode and time data scanned by two barcode readers, confirm that one carrier has barcode data canned by two barcode readers, and make the correspondence relationship with production data.

B. Production data of Printing, SPI, Mounter (1-3), Reflow, Event and AOI are extracted from the database, after the correspondence relationship between the time canned by Barcode Scanner 1 and the completion time of Mounter 3 is made, the count index of the same carrier is consistent. Therefore, the “count index” columns is used for the primary key correspondence of the carrier's mounter 1-3 data, and the corresponding mounter event data are searched.

C. Make the correspondence relationship between the barcode data scanned by Barcode Scanner 2 and the AOI data table barcode field, and confirm that the two time complies with the working hours range.
D. By taking the time of Barcode Scanner 1 as the starting point, and the completion time of AOI test minus the carrier working time in AOI as the ending time, calculate this period of time from the starting point to the ending time and make the correspondence relationship with the reflow data corresponding to this period of time, as the production data corresponding to the Carrier.
E. Search for corresponding SPI and Print data within the standard working time with the bar code of the Carrier.

Write key data of Print, SPI, Mounter 1-3, Reflow and AOI corresponding to the Carrier into the data table.

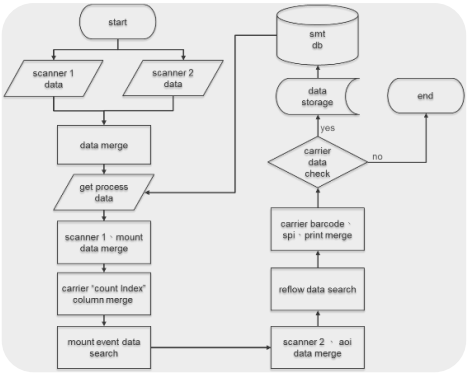


Fig. 3. Production data merge system flowchart

The product process module is executed in a timed and scheduled tasks, and the production data is connected to the data table once a minute, so that the production record can be queried and used in the algorithm.

3.2 Improving the efficiency of data access

In the production process of the SMT production line, huge amount of data generated by different processes and inspection equipment causes some difficulty in the connection of the data table records of the process database. There are OUTER JOIN or INNER JOIN in many tables, usually the amount of such data can reach more than one million, resulting in the slower query speed. Creating an index relational view in the database to link the correlation of various data can save the time and cost of data query without data changes.

With the index key of various equipment production data table that Carrier corresponds to, the data of one Carrier can correspond to one set of production data, several sets of Event data and pickup feeder/pickup nozzle data. Therefore, several sets of data

in the Database View will be included in the View Column through the way of Json Array, so the number of many pairs of database transactions can be reduced.

3.3 Production quality tracking module

In this section, ensemble learning methods based on decision trees will be applied to identify abnormal manufacturing process data, and to derive potential defect source with weighted rules extracted from model training with historical manufacturing data, so as to assist operators in SMT production line to improve the efficiency in analyzing defect-causing factors from different machines.

Model Interpretability and Ensemble learning based on Decision Tree. No matter using machine learning or deep learning methods, it need to be considered the trade-off between accuracy and interpretability carefully. For the purpose of analyzing abnormal parameters, the algorithm with high interpretability is the primary choice. However, [2] shows that the algorithm with high interpretability, such as decision tree and logistic regression, is commonly used, but the model performance usually below user's expectation, refer to the following figure. To solve this problem, ensemble learning is a good choice.

The core concept of ensemble learning is to improve the overall model performance by making decision with several diverse simple models, which usually retains interpretability. Therefore, it's an important step that how to generate multiple diverse simple classifiers from a single data set in ensemble learning.

With decision tree as the basic classifier, bagging and boosting are the main methods to generate simple classifies in ensemble learning. Bagging method is to extract sub-data sets from the original data set in parallel, and construct a new decision tree for each sub-data set. Random forest is one of the popular bagging methods. Boosting method is to set up the first decision tree from the original data set, and to increase the weight of the samples with misclassification. Thus, the distribution of the data set is different from the original one, and it would be applied to build the second tree. This procedure would be repeated until the total number of trees is satisfied. Ada-Boosting is one of classical boosting methods. Independency is the main difference between bagging and boosting methods, where trees in boosting methods are dependent owing that trees are generated in orders, but not for trees in bagging methods.

In this study, ensemble learning with multiple decision tree based classifiers is applied to trade-off a model performance and interpretability.

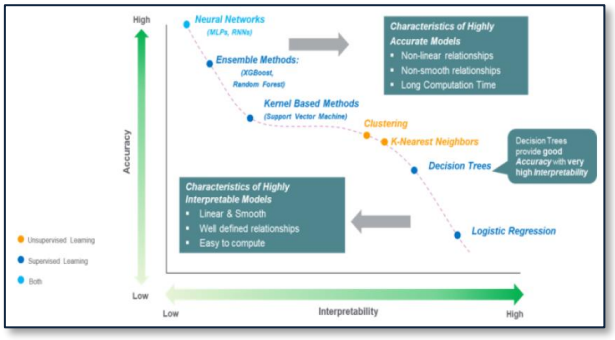


Fig. 4. Trade-off between performance and interpretability while adopting different models

Algorithm 1

Input: train dataset and validation dataset

Output: a random forest model

1. Do pre-process and feature engineer on train and validation dataset.
2. Deal imbalanced data issue with SMOTE and Under Sampling method.
3. Use train and validation dataset to build a random forest model within K-Fold validation.

Rule extraction, refined and weighted. A decision tree starts growing process from a single root node. For each split, the most appropriate characteristics and range is defined to classify data samples into two child nodes. The split procedure will end in leaf nodes, which cannot split, and it must stop under beforehand setting conditions, such as, data samples in the same node all belong to a category, or a setting maximum tree depth is achieved. Thus, each leaf node can trace back to the root node, and each tracing path can be described as an if-then rule, which composes of conditions. For instance, if a leaf node, which only contains class A, split by condition 1 and condition 2 from the root node, then this path can convert into "if condition 1 and condition 2, then class A". Furthermore, if a random forest is composed of M tress, and, for the ith tree in forest, it contains n_i leaf nodes, then $\sum_{i=1}^{M} n_i$rules can be obtained from random forest. Besides, each rule also represents a unique category.

Remove redundant conditions and rules. Since each decision tree is built based on different subset or distribution from the original training dataset, combination of conditions in rules are different. However, it is inevitable that similar rules may be extracted from different trees. [11] provide several view points and methods to get a more compact rule set, include removing redundant condition, the same rules but categories contradiction, overlap on rule dimension, and so on.

Computation on rule weight. After the preliminary screening, various rules represent different characteristic combinations and characteristic ranges to describe their corresponding categories. However, there is rarely a set of rules that can distinguish categories accurately, so the interpretability of rules to training data sets should be considered

before assigning weights. [6] provides a method to rank rules by the importance level of conditions in the decision tree and in the rule, and the proportion of the corresponding category that the rule can interpret. When calculating a weight of a rule, [8] consider the number of conditions, represented the complexity of the rule, and give a higher weight to the rule with fewer conditions (low rule complexity). The weight formula used in this study is provided as follows.

$$Weight_r = F1 - Score = \frac{2}{\frac{1}{Recall} + \frac{1}{Precision}} \quad (1)$$

After classifying the training data set by rule r, the confusion matrix and F1-score can be computed from the prediction and the group truth, and F1-score is also assigned as a weight of a rule.

Algorithm 2

Input: a random forest model
Output: rules with weight
For tree in Random Forest
  For leaf node in tree
    1. Trace and record information in parent node recursively, and end at root node.
    2. Format conditions from parent node into an if-then rule with corresponding class type.
Remove redundant rules from rule set.
Compute weight of each rule, and assign weight to conditions in each rules averagely.

Application of weighted rule. A rule usually consists of multiple conditions. For instance, if a rule is "if Condition 1 and Condition 2 and Condition 3, then Class A", then there are three conditions in this rule, which means that Class A can be described as a union of these conditions. Thus, all conditions must be fulfilled by manufacturing process data of a new product, then the product can be predicted as a particular category. Due to that parameters in conditions come from different machines, weight of a rule can be reassigned to conditions averagely. While multiple rules are fulfilled by manufacturing process data of a product, conditions and weights can be regrouped by machines. Finally, defect causing probability of each machine can be obtained by scaling weights after reassigning to machines, and an intuitive way is presented to help users tracing defect-causing factors.

Algorithm 3

Input1: new data
Input2: rules with weight
Output: Abnormal probability between machines
For rule in Inpute2
  1. If all conditions are fulfilled, then separate the rule to conditions
  2. Append conditions to matched-conditions-set

Classify conditions in matched-conditions-set to corresponding machines
Sum up weights at each machine

Nozzle analysis. In addition to a comprehensive analysis of machines in the SMT production line provided in the previous section. Abnormal nozzle function in mounter also leads to product defects frequently, so defect ratio analysis of nozzles in mounters is provided as following.

$$\text{abnormal nozzle carused defect ratio} = \frac{count\left(\begin{matrix}carrier\ with\ defct\ and \\ \text{with}\ nozzle\ error\end{matrix}\right)}{count(carrier\ with\ defect)} \quad (2)$$

4 Discussion

4.1 Product process module data access efficiency validation

The original SMT production data are stored in multiple tables. If SQL Joins and filtering are screened only when database query is made, the runtime of a single query is about 20.155 seconds, the reaction time is too long, which slows down the reaction speed of the analysis module and the database query module. After the system changes to create the data view in the way of relational database view, the overall query response time is shortened to 3.495 seconds, which accelerates the database query speed of about 80% compared with the original situation, and accelerates the overall efficiency of the system.

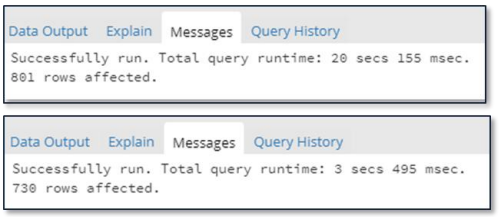


Fig. 5. SQL transaction performance comparison

To verify the correctness and completeness of the production history corresponding, we adopt production data of 14 consecutive days as the verification subject, including 14 products, a total of about 10,000 sets of carriers, with about 1.3 GB data volume. Compared with the situation where the original manual tracking method could not correspond to the product process, the standard working hours pushback module can effectively improve the completeness of production resume data to about 80%, which is more helpful to the subsequent production resume and analysis accuracy.

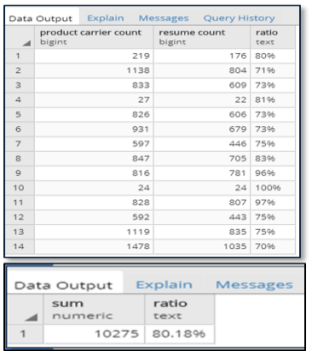


Fig. 6. Production data tracing rate

4.2 Defect-causing trace evaluation

At current stage, operators in production line still improve production conditions through experience, so it is difficult to measure the correction of suggestions on defect-causing trace. To keep the reliability of rules extracted from ensemble learning models, it's necessary to define standard criteria of model performance. Model performance criteria, such as Accuracy, Precision, Recall, F1-Score, etc. are frequently applied to help data analysts evaluating efficiently the applicability of different models. In this study, recall rate and accuracy is adopted while evaluating models. Different ensemble learning methods are evaluated in the following table.

Table 1. Available data for training model

Number of data samples	Number of failure carrier confirm manually	Number of qualified carrier confirm manually
10390	186	10204

Table 2. Model evaluation

Algorithm	Method for Data Imbalance issue	5 Fold Recall	5 Fold Accuracy
Decision tree	Under-Sampling	AVG：59.3% STD： 7.1%	AVG：49.5% STD：3%
	SMOTE	AVG：1.5% STD：1.9%	AVG：49.9 STD：1%
Random Forest	Under-Sampling	AVG：72.3% STD：6.5%	AVG：53.9% STD：4.3%
	SMOTE	AVG：21.5% STD：17.2%	AVG：48.5 STD：4%
Ada-Boosting	Under-Sampling	AVG：58.7% STD：10%	AVG：52.5% STD：2.5%
	SMOTE	AVG：67.5% STD：26.6%	AVG：48% STD：1.9%
Extra - Tress	Under-Sampling	AVG：68% STD：16%	AVG：50% STD：2%

	SMOTE	AVG：61% STD：9%	AVG：50.5% STD：2.5%
Skope Rules [10]	Under-Sampling	AVG：63.9% STD：42.2%	AVG：48% STD：2%
	SMOTE	AVG：33.43% STD：40.9%	AVG：47.9% STD：2.6%

It can be observed that combinations of different ensemble learning methods and resampling methods are evaluated. From the above table, it can be concluded that under-sampling method is more effective than SMOTE in this problem. Besides, Skope Rules, a rule extraction library, in [10] is also included in the experiment. However, compared with other models that have been extensively used in real manufacturing problem, the current model is still not precise enough to assist operators without doubt. In the future, this system will be adopted in other SMT production lines, which may increase the volume of input data, and provide a chance to dig out more useful information to enhance the current model.

5 Conclusion

In this paper, a quality tracking system is applied to a SMT production lines. When defective products are detected by AOI, defect-causing factor of a defect can be point out efficiently, which makes defect tracking issue more efficient. In addition, this method has been applied in indicator SMT vendors.

Moreover, according to the professional knowledge from production lines, final weights can be given based on the known defect categories, named as domain-knowledge weight. For example, among machine A, B and C, it is confirmed by the production line that type 1 defect only caused by machine A and C, then the domain-knowledge weight of machine B can be set as 0. Next time, when a new product is predicted as type 1 defect with high probability, probability derived from weighted rules can multiply domain-knowledge weight weights. Thus, trace suggestion would never direct to machine B, which make trace suggestions more concrete.

Furthermore, when sufficient amount of data and corresponding trace suggestion is verified by production lines, including process data, product quality data and abnormal probability between machines, these data can be used as data input to train a multi-task neural network model to enhance the quality control.

To improve production quality, it’s necessary to rely on information feedback from models, which is highly related to interpretability of models. In near feature, it will become a tendency that applying machine learning and deep learning methods to improve quality control in manufacturing industry.

6 Acknowledgement

This study is conducted under the “Technology of Quality Decision-making support in Intelligent Manufacturing (3/4)” of the Institute for Information Industry which is subsidized by the Ministry of Economic Affairs of the Republic of China.

References

1. Chun-quan Li ,Zhao-hua Wu: SPC System Analysis and Design of Reflow Soldering Process, 2006 7th International Conference on Electronic Packaging Technology
2. Data Science Ninja, The balance: Accuracy vs. Interpretability, http://www.datascienceninja.com/2019/07/01/the-balance-accuracy-vs-interpretability/
3. Heinz Wohlrabe , Thomas Herzog ,Klaus-Jurgen Wolter: Optimization of SMT Solder Joint Quality by Variation of Material and Reflow Parameters, 2008 2nd Electronics System-Integration Technology Conference
4. Jeffrey Len Yung Kwuan , Leon Rao , Evan Yip , Wisdom Qu , Jonas Sjoberg: IMPACT OF STENCIL QUALITY & TECHNOLGY ON SOLDER PASTE PRINTING PERFORMANCE, 2019 Pan Pacific Microelectronics Symposium (Pan Pacific).
5. Kamen E.w., Goldstein A., Creveling D.R., Sahinci E. and Xiong Z.: Analysis of factors affecting component placement accuracy in SMT electronics assembly, In Twenty Third IEEE/CPMT International Electronics Manufacturing Technology Symposium (Cat. No.98CH36205), pp. 50-57, Austin, TX, USA (1998).
6. L. T. K. Phung, V. T. N. Chau and N. H. Phung: Extracting Rule RF in Educational Data Classification: From a Random Forest to Interpretable Refined Rules, 2015 International Conference on Advanced Computing and Applications (ACOMP), pp. 20-27. Ho Chi Minh City, 2015,
7. Li Chun-quan ; Wu Zhao-hua: Process Management System Analysis and Design of SMT Reflow Soldering process, 2005 6th International Conference on Electronic Packaging Technology.
8. Mashayekhi M., Gras R. (2015) Rule Extraction from Random Forest: the RF+HC Methods. In: Barbosa D., Milios E. (eds) Advances in Artificial Intelligence. Canadian AI 2015. Lecture Notes in Computer Science, vol 9091, pp223-237. Springer, Cham
9. P. Svasta ; D. Simion-Zanescu ; C. Willi: Thermal Conductivity Influence in SMT Reflow Soldering Process, 52nd Electronic Components and Technology Conference 2002. (Cat. No.02CH37345).
10. Ronan.Gautier, Gregoire.Jafre, Bibi. Ndiaye: Interpretabbility With Diversified-By-Design Rules; SKOPE-RULEs, A Python Package. http://2018.ds3-datascience-polytechnique.fr/wp-content/uploads/2018/06/DS3-309.pdf
11. Sirikulviriya, N., Sinthupinyo, S: Integration of Rules from a Random Forest, In Proc. of the 2011 Int. Conf. on Information and Electronics Engineering, 2011, pp.194-198
12. Zhen H, Ershi Q, Zixian L.: QUALITY IMPROVEMENT THROUGH SPC/DOE IN SMT MANUFACTURING, Proceedings of the 2000 IEEE International Conference on Management of Innovation and Technology. ICMIT 2000. 'Management in the 21st Century' (Cat. No.00EX457).

An Image Similarity Estimation Approach Based on Weighted Features

Seok-Woo Jang[1][0000-0001-5580-4098]

[1] Anyang University, Anyang 14028, Republic of Korea
swjang7285@gmail.com

Abstract. Recently, many attempts have been made to automatically determine error pixels from various types of input color images through image processing. In this paper, we propose a new method that effectively detects regions different from the reference image from the input image by utilizing the difference calculation and a Gaussian mixture model. The proposed method first detects the color error between the reference image and the input image using the difference operation. Then, the object representation error between the reference image and the input image is detected using a Gaussian mixture model. Finally, in this paper, the final error of the input image is determined by combining the detected color error and the object representation error effectively. The experimental results show that the proposed approach robustly detects error pixels that are not similar to the reference image from the input color image by using the difference calculation and a Gaussian mixture model.

Keywords: Weighting Factor, Color Feature, Difference Calculation.

1 Introduction

With the rapid development of high-speed wireless network functions, high-speed central processing units, low-cost storage space, and smart devices that provide excellent image quality, many types of applications that can be executed on these devices are spreading [1-3]. In particular, applications that can be used in various platforms including Android and iOS and can be used in various industries including artificial intelligence and big data, are being developed [4-6]. As a result, users tend to use these various applications in a mobile environment instead of a personal computer environment.

As the number of applications developed for mobile devices such as smart phones and tablet PCs explodes, the demand to inspect screens executed in the developed programs is also increasing exponentially. In other words, the tester checks whether each screen of the developed program is displayed correctly. Then, if screens with errors appear during the testing process, there is an excessive amount of work for the tester to request for correction.

Therefore, research is needed to intelligently determine whether an input image is a normal or error image through content analysis from various input images [7]. In

other words, it is necessary to extract and integrate main features from the input image, and then determine whether the integrated features are similar to those of the reference image stored in the system.

Existing studies for detecting error pixels different from reference pixels from an input test image can be referred to in related literature. However, the existing error pixel detection methods are not yet relatively high in completeness, and there are limitations due to surroundings such as lighting and noise. In addition, the existing researches currently being conducted in relation to the detection of error regions in an image are relatively few compared to other studies.

Therefore, this study proposes a new method to effectively detect regions different from the reference image by combining and using the difference calculation and the Gaussian mixture model from the input color image. The following Fig. 1 represents the overall flowchart of the weighted feature-based error pixel extraction approach proposed in this paper.

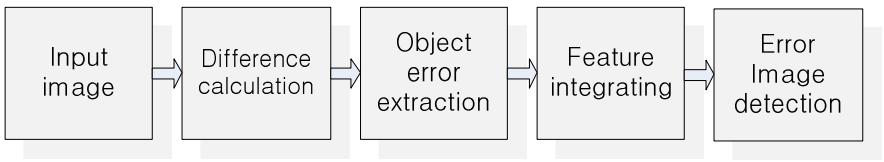


Fig. 1. Flowchart of the suggested algorithm.

As shown in Fig. 1, the algorithm proposed in this paper first detects color errors between the reference image and the input image using a difference operation. A Gaussian mixture model is then used to detect an object expression error between the reference image and the input image. Finally, it is possible to effectively determine the final error of the input image by effectively combining the detected color error and the object expression error.

In Chapter 1, the overall background and motivation for conducting this study were explained. Chapter 2 describes the technique of obtaining a difference image from two images. Chapter 3 explains how to detect an object expression error for an image using a Gaussian mixture model and then determine whether the image contains an error. Chapter 4 describes the experimental results performed to quantitatively evaluate the performance of the proposed image error extraction algorithm. Finally, Chapter 5 describes the conclusions of this study and future research plans.

2 Extraction of Color Difference

In the proposed method, the color difference feature between the reference image and the input image is first extracted [8-9]. To this end, in this paper, the *RGB* color space of the image is separated for each of the *R*, *G*, and *B* channels, and then the difference feature for each color is extracted as shown in Equation (1). In Equation (1), $R_{ref}(x, y)$, $G_{ref}(x, y)$, and $B_{ref}(x, y)$ represent the red, green, and blue color values at the (x, y) position of the reference image, and $R_{inp}(x, y)$, $G_{inp}(x, y)$, and $B_{inp}(x, y)$ represent the

red, green, and blue color values at the (x, y) position of the input image. In this paper, the color difference feature has a value between 0 and 255.

$$\begin{aligned} D_r(x,y) &= \left| R_{ref}(x,y) - R_{inp}(x,y) \right| \\ D_g(x,y) &= \left| G_{ref}(x,y) - G_{inp}(x,y) \right| \\ D_b(x,y) &= \left| B_{ref}(x,y) - B_{inp}(x,y) \right| \end{aligned} \tag{1}$$

Subsequently, when the color difference for each channel calculated in Equation (1) is greater than a predetermined threshold, the proposed method determines that the corresponding pixels of the reference image and the input image are not similar as in Equation (2). In Equation (2), Th_r, Th_g, and Th_b represent threshold values for red, green, and blue colors, respectively. In general, thresholds are determined through repeated experiments in the field of computer vision. Also, e_r, e_g, and e_b mean the total number of pixels that are not similar for each red, green, and blue color.

$$\begin{aligned} &IF\ (D_r > Th_r)\ \ THEN\ \ e_r = e_r + 1 \\ &IF\ (D_g > Th_g)\ \ THEN\ \ e_g = e_g + 1 \\ &IF\ (D_b > Th_b)\ \ THEN\ \ e_b = e_b + 1 \end{aligned} \tag{2}$$

Equation (3) represents the similarity measure for the color difference feature between two images.

$$S_{diff}(I_{ref}, I_{inp}) = \left(1 - \frac{e_r + e_g + e_b}{3 \times W \times H} \right) \times 100 \tag{3}$$

3 Gaussian Mixture Model

The Gaussian mixture model-based method for detecting an error region from an input image is composed of three steps. In the first step, the initial parameters are set and then a Gaussian mixture model for the reference image is generated [10-11]. In general, the Gaussian mixture model can be expressed as in Equation (4). In Equation (4), X_t is the value of the input pixel, and $w_{i,t}$ is the weight of the i-th Gaussian distribution. $\mu_{i,t}$ and $\Sigma_{i,t}$ represent the mean and covariance matrix of the distribution.

$$P(X_t) = \sum_{i=1}^{K} \omega_{i,t} \eta(X_t, \mu_{i,t}, \Sigma_{i,t}) \tag{4}$$

In the second step, the Gaussian mixture model of the input image is generated, and it is matched with the Gaussian mixture model of the reference image to detect an error region. In the third step, the parameters of the Gaussian mixture model are up-

dated according to whether the input images match. In other words, if a match is made, the weight of the corresponding distribution is increased. Conversely, if no match is made, the corresponding weight is reduced. Through this process, the similarity between the reference image and the input image can be measured as shown in Equation (5). In Equation (5), P_{error} is the number of pixels where an error occurred, and W and H are the width and height of the image.

$$S_{GMM}(I_{ref}, I_{inp}) = \left(1 - \frac{P_{error}}{WH}\right) \times 100 \tag{5}$$

In this paper, it is judged that the two images are similar if the weighted similarity between the difference operation and the Gaussian mixture model, as shown in Equation (6), is greater than or equal to the threshold, otherwise they are not. In Equation (6), α and β are the weighting factors that control the importance of Gaussian mixture model feature values and difference feature values.

$$S(I_{ref}, I_{inp}) = \alpha \times S_{GMM}(I_{ref}, I_{inp}) + \beta \times S_{diff}(I_{ref}, I_{inp}) \tag{6}$$

4 Experimental Results

The computer used to perform the experiment in this paper consists of an Intel Core (TM) i7-6700 3.4 GHz CPU, 16 GB main memory, and 256 GB solid state drive (SSD). The Windows 10 operating system (OS) was installed on the personal computer used. In addition, Microsoft's Visual Studio version 2015 was used as the integrated development environment (IDE) to implement the proposed algorithm. In the experiment, the image data set generated by capturing screens of various application programs developed in various platform-based mobile environments was used as our test data set.

$$\Phi_{accuracy} = \frac{IMG_{err}}{IMG_{all}} \times 100(\%) \tag{7}$$

In this paper, the performance of the proposed error image detection algorithm is quantitatively evaluated. In this paper, a measure such as Equation (7) is used, which expresses the ratio of the error images of the program extracted correctly from the input test image to the total number of error images in the test image data. In Equation (7), IMG_{error} represents the number of error images extracted accurately using the proposed algorithm. In addition, IMG_{all} indicates the total number of error images present in the input test image.

Fig. 2 is a graph showing the results of measuring the accuracy performance of the proposed error image extraction algorithm. As can be seen from Fig. 2, the algorithm based on the difference operation and the Gaussian mixture model proposed in this paper accurately detects an error-prone screen in an application program.

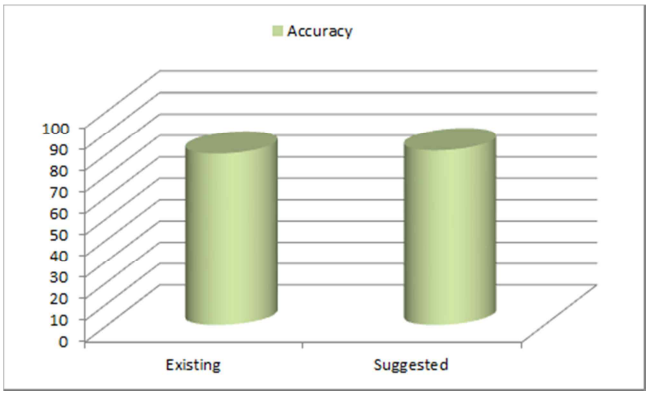


Fig. 2. Performance evaluation

5 Conclusion

In this study, we proposed a weighted feature-based content analysis method to robustly detect an error image that is not similar to a reference image from an input test image. In the proposed method, the color error value and the object expression error value are respectively extracted from the input color image using a difference operation and a Gaussian mixture model. Subsequently, the proposed algorithm integrated the extracted feature values using weighting factors, and then calculated a similarity measure for error pixel detection. Finally, it was determined whether the input image is an error image or a normal image using the calculated measure. The experimental results showed that the proposed method robustly extracts an error image that is not similar to the reference image by using the weighted feature based on the difference image and the Gaussian mixture model from the input image.

In the future, we plan to further improve the performance of the proposed system in terms of accuracy by applying the error image detection algorithm proposed in this study to more various types of color test images. In addition, it is planned to adaptively adjust various internal parameters used in the proposed algorithm to effectively minimize the output error.

Acknowledgement

This work was supported by the National Research Foundation of Korea (NRF) grant funded by the Korea government (MSIT) (2019R1F1A1056475).

References

1. Soldani, J., Binz, T., Breitenbucher, U., Leymann, F., Brogi, A.: ToscaMart: a Method for Adapting and Reusing Cloud Applications. Journal of Systems and Software 113, 395-406 (2019)

2. Garcia, C.S., Meincheim, A., Junior, E.R.F., Dallagassa, M.R., Scalabrin, E.E.: Process Mining Techniques and Applications - A Systematic Mapping Study. Expert Systems with Applications 133, 260-295 (2019)
3. Han, S., Yun, Y., Hwang, E., Kim, W., Kim, Y.H.: Event-based Application Modeling for Analysis of Asymmetric Multicore-based Mobile Systems. Journal of Systems Architecture 97, 477-490 (2019)
4. Alazab, M., Alazab, M., Shalaginov, A., Mesleh, A., Awajan, A.: Intelligent Mobile Malware Detection Using Permission Requests and API Calls. Future Generation Computer Systems 107, 509-521 (2020)
5. Chen, Y., Li, Z., Yang, B., Nai, K., Li, K.: A Stackelberg Game Approach to Multiple Resources Allocation and Pricing in Mobile Edge Computing. Future Generation Computer Systems 108, 273-287 (2020)
6. Kiconco, R.I., Rooks, G., Snijders C.: Learning Mobile Money in Social Networks: Comparing a Rural and Urban Region in Uganda. Computers in Human Behavior 103, 214-225 (2020)
7. Shi, J., Wang, X.: A Local Feature with Multiple Line Descriptors and Its Speeded-up Matching Algorithm. Computer Vision and Image Understanding 162, 57-70 (2017)
8. Gao, H., Hu, M., Gao, T., Cheng, R.: Robust Detection of Median Filtering Based on Combined Features of Difference Image. Signal Processing: Image Communication 72, 126-133 (2019)
9. Fan, C., Zhang, Y., Zhang, H., Hamzaoui, R., Jiang, Q.: Picture-Level Just Noticeable Difference for Symmetrically and Asymmetrically Compressed Stereoscopic Images: Subjective Quality Assessment Study and Datasets. Journal of Visual Communication and Image Representation 62, 140-151 (2019)
10. Ri, H.S., Ri, J.C.: Subspace Gaussian Mixture Based Language Modeling for Large Vocabulary Continuous Speech Recognition. Speech Communication 117, 21-27 (2020)
11. Chen, Y., Wu, W.: Separation of Geochemical Anomalies from the Sample Data of Unknown Distribution Population Using Gaussian Mixture Model. Computers and Geosciences 125, 9-18 (2019)

An Empirical Study on Success Factors of Asian Chinese Companies

Seung Gyun Yoo

Department of Global Economics and Commerce, Dongguk University-Gyeongju, 123, Dongdae-ro, Gyeongju-si, Gyeongsangbuk-do, 38066, Republic of Korea

bluetrade@dongguk.ac.kr

Abstract. The 21st century is called the era of Asian economy. The success factors of Chinese enterprises that emerged as the new economic power are corporate culture, entrepreneurship, tradition and innovation, and response to rapid environmental changes. Through this study, entrepreneurship, corporate culture, and response to environmental changes were identified as major factors affecting corpo60rate performance. In a market environment where global competition is intensifying, entrepreneurship and response to change are very important. In addition, innovation is required along with tradition. In future studies, we will study the difference between success factors between Chinese and Korean companies.

keyword : Chinese, Company Culture, Performance, Adaptation, Management Strategy

1. Introduction

The 21st century is called the era of Asian economy. In the new era of Asian economy, Chinese and Chinese are the main players, and the development of China's economic zone connects regions of Southeast Asia and Northeast Asia, contributing greatly to economic growth and trade expansion in East Asia. In Southeast Asia, mutual investment using the Chinese network is actively developing, and in Northeast Asia, the Chinese economy plays a central role in economic integration.

To grow into a global company, an executive with excellent corporate culture and entrepreneurship is needed. Otherwise, it is difficult to occupy an advantage in the world's large markets. Chinese companies that are successful abroad have this special success factor. A close analysis of these success factors and empirical analysis can provide implications for Korean companies. It can also help to find a way to secure a more favorable position in competition with Chinese companies in the global market.

2. Characteristics and Success Factors of Chinese Enterprises

2.1 Characteristics of Chinese Enterprises

Those who affirm Confucian capitalism insist on the key to Chinese companies' management and operational success. First, family management. The majority of Chinese companies were founded on the basis of family funds, family members' workforce and relationships. Second, it is high power. Companies adopt a centralized approach to policy making, where one or a few leaders make all the important decisions here as the heart of the company. Staff will provide feedback on policy decisions. Third, it is the succession of the rich. The management of the company is based on the family. Fourth, regionalism. Chinese enterprises are thoroughly operated in a localized form, and refrain from large-scale management methods beyond the region. Fifth, it is term response. Businesses are very sensitive to environmental changes and have flexible responses. Sixth is single management. Chinese enterprises are operated as small and medium-sized business models, and there is little demand for funds. Seventh, confidentiality. Family-centered management is becoming a secret to others. In addition, information is acquired

through personal relationships where financial information is restricted.

2.2 Success Factors of Chinese Enterprises

2.2.1 Cultural Aspects

The biggest feature of the globalization of Chinese companies is that they have the same cultural background. That is, they use the same language and share the same cultural foundation, and help distribute their goods across borders or raise the necessary funds to other regions, creating a favorable environment for the internationalization of companies. Cultural factors affecting the internationalization of Chinese companies include ethnic ties, homogenous cultural bases, guanxi, and the tendency to hedge.

First, ethnic ties are special. Investing near the home of a manager or dispatched worker has an economic advantage that can significantly reduce the cost of obtaining market information. Second, it has a homogenous cultural base. Because they share the same heritage of Chinese civilization as other regions and have similar values, they feel familiar rather than confrontation and form a cooperative system. Third, the relationship is formed through the poem. One of the main factors that allows Chinese companies to easily expand their markets across the border is guanxi. This network features individual companies that are not limited to a particular country or continent but spread worldwide. Fourth, the ability to hedge. The Chinese entrepreneur lives in full readiness to prepare for the dangers of his or her own personal or property. This situation is also well illustrated by the epigram that you should always pack your bags and be ready to go bankrupt at any time.

2.2.2 Managerial Aspects

Analysis of the advances made by most Chinese companies from traditional fields such as distribution and finance to new market areas shows that profit preference for short-term investments takes precedence over long-term investments. In addition, if you look at Chinese companies' overseas expansion, it can be summed up in four major stages: small investment, advanced countries, Southeast Asian director markets and global market expansion.

In addition, the general characteristics of Chinese companies' overseas investments are as follows. First, it will proceed with overseas investment after financing, focusing on financial hubs such as Hong Kong and Singapore. Second, they prefer joint venture with Chinese. Third, we prefer the service industry to manufacturing. Fourth, there are many cases of joint ventures and M&As rather than greenfield investments.

Many of the companies competing in the global market are creative. But such creative reasons and adventurous spirit are lacking in Confucian culture. These creative knowledge and adventurous market competition make Chinese companies successful overseas constantly trying to innovate. The manager of a Chinese firm also emphasizes that the personal development and benefits of its employees should be subordinate to their organizational needs. In other words, employees tend to think of them as organized human beings.

Promotion within Chinese companies is more important than employee performance reviews and competencies. However, Chinese companies that succeed in the global market reduce middle managers, emphasize efficiency, and emphasize the flexibility and cooperation of tradition and the future.

3. Model and Results

3.1 Model.

The model of this study is as follows.

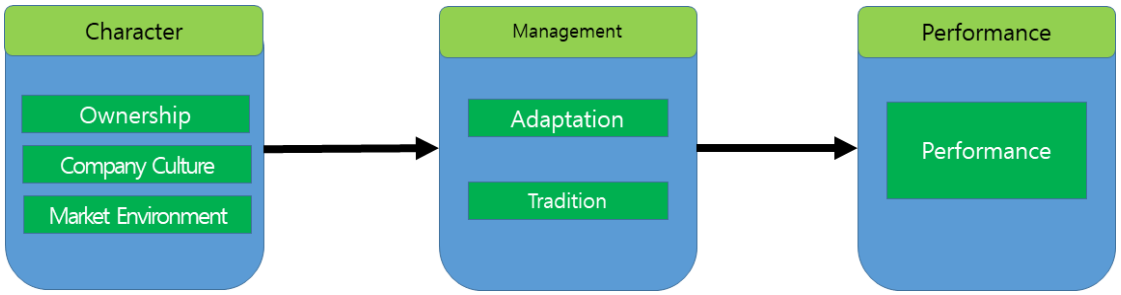


Fig.1. Reserch model

3.2. Results.

The correlation by factor is as follows.

Table 1. Correlation by factors

Classify	A	B	C	D	E	F	G
Ownership(A)	1						
Company Culture(B)	.305(**)	1					
Market Environment(C)	.099	.206(*)	1				
Industry Environment(D)	.268(*)	.041	.112	1			
Adaptation(E)	.308(**)	.306(**)	.336(**)	.141	1		
Tradition(F)	-0.32	.255(*)	.198	.302(**)	.138	1	
Performance(G)	.146	.250(*)	.304(**)	.194	.313(**)	.205(*)	1

**(Significance at the 0.01 level), *(Significance at the 0.05 level)

The conformity assessment index of the research model is as follows.

Table 2. Conformity Assessment Index

Classify	ModelConformityAssessmentIndex		Result
Absolute ConformityIndex	χ^2	Chi-square(degreeoffreedom)	48.634(39df)
	p	Significantprobability	*0.152≥.05*
	Q	Chi-squared/degree-of-freedomratio≤3	1.291≤3
	GFI	GoodnessofFitIndex≥0.9	0.907≥0.9
	AGFI	AdjustedGFI≥0.9	0.846≤0.9
	RMR	RootMean-SquareResidual	0.041≤0.05

	RMSEA	RootMeanSquareerrorofapproximation	0.05≤0.05
Incremental ConformityIndex	NFI	NormedFitIndex≥0.9	0.902≥0.9
	RFI	RelativeFitIndex≥0.9	0.8465≤0.9
	CFI	ComparativeFitIndex≥0.9	0.930≥0.9
Simplicity ConformityIndex	PNFI	ParsimoniousNormed-of-FitIndex,	0.553

The route analysis results are as follows.

Table 3. Route Analysis Results

Theory	Path	Path coefficient	Standard error	t	p	Result
1	Ownership→Adaptation	.242	.139	1.908	.046	Accept
2	Ownership→Tradition	-.357	.143	-2.308	.118	Reject
3	CompanyCulture→Adaptation	.149	.248	.967	.319	Reject
4	CompanyCulture→Tradition	.753	.330	2.405	.013	Accept
5	M.E.→Adaptation	.345	.342	2.069	.042	Accept
6	M.E.→Tradition	-.141	.334	-.401	.645	Reject
7	I.E.→Adaptation	.051	.138	.204	.784	Reject
8	I.E.→Tradition	.556	.236	3.705	.130	Reject
9	Adaptation→Performance	.258	.144	2.806	.004	Accept
10	Tradition→Performance	.259	.139	2.051	.007	Accept

4. Conclusion

In today's rapidly changing world of internal and external business environments surrounding companies, all global companies have a strong commitment to management innovation. In particular, after the 4th Industrial Revolution, the intensification of competition among companies has changed rapidly.

Asian video companies maintain the identity of video companies by adhering to the new environment of management and adhering to the traditional way of corporate culture. In addition, the company is seeking to improve the performance of companies by reacting sensitively to changes in the market environment. Through this, it can be seen that successful video companies are pursuing constant change and flexibly responding to market changes. In addition, it is required to share culture between managers and members and share strategies to improve corporate performance.

This paper empirically analyzed the success factors for Asian video companies. However, it is a limitation of the study that only the accessible parts of the success factor analysis are examined. Future research will focus on the differences between Chinese-based companies and overseas video companies.

References

1. Amit Dhir, Wireless Homenetworks-DECT, Bluetooth, HomeRF, and Wireless LANs,XILINX(2001).
2. "Bluetooth HCI", ATMEL(2000).
3. C. Bisdikian, "An Overview of the Bluetooth wireless Technology", IEEE Comm' Mag, pp. 86~93(2001).
4. Dietmar Loy, Atsushi Murase and Andreas doederlin, "System Level Prototyping for Embedded Networking Applications", IEEE 12th International Workshop on Rapid System Prototyping, pp. 12~16(2001).
5. Gcommkorea, "Bluetooth", pp. 431~433(1999).
6. Hesse, R. B. Two-Sided Platform Markets and the Application of the Traditional Antitrust Analytical Framework. Competition Policy International(2007).
7. IEEE802.11 Standard, "Wireless LAN Medium Access Control(MAC) and Physical Layer(PHY) Specifications"(1997).
8. Jennifer Bray, Charles F Sturman, Bluetooth:Connect Without Cables, Prentice Hall(2001).
9. Joo, Hyun Woo. A Study on the Strategic Development Model of the Logistics Platform. Chung-Ang University (Master's Thesis)(2018).
10. Jean J. Labrosse, Micro C/OS-II The Real-Time Kernel, CMP Books, pp. 9~464(2002).
11. Kim, Sun-Nam. An Empirical Study on the Important Factors Increasing the Acceptance Intention of Online International Logistics Platform; (Doctoral Dissertation) Myongji University(2016)
12. LXE Inc., Bluetooth Basics(2002).
13. Moberg, C. R.; Speh, T. W. Evaluating the relationship between questionable business practices and the strength of supply chain relationships. Journal of Business Logistics, 24, pp. 1-19(2003).
14. Parvin Bhagwat, Adrian Segall, "A Routing Vector Method(RVM) for Routing in Bluetooth Scatternets", MoMuC' 99, pp. 375~379(1999).
15. R. Berezdivin, R. Breinig and R. Topp, "Next-generation Wireless Communication concepts and Technologies", IEEE Comm' Mag, pp. 188~116(2002).
16. Ryu, Yejun and Yeo, Gi-Tae. A Study on the Evaluation of the Functionality of Shipping Logistics Platform Using IPA Analysis. Journal of Korean Navigation and Port Research, 44, pp. 32-43(2020)
17. Takeshi Saito, Ichiro Tomoda and Yoshiaki Takabatake, "Gateway Technologies for Home Network and Their Implementations", IEEE 2001 International Conference on Distributed Computing Systems Workshop, pp. 175~180 (2001).

Efficient Covering of a Target Object Using a Prediction Technique

Byeongtae Ahn[1[0000-0003-3431-9493]] and Seok-Woo Jang[2[0000-0001-5580-4098]]

[1,2] Anyang University, Anyang 14028, Republic of Korea
ahnbt@anyang.ac.kr, swjang7285@gmail.com

Abstract. In recent years, as the shooting of video and the spread and viewing of video over the Internet have increased exponentially, the damage caused by the exposure of personal information has occurred. In this paper, we propose a method that robustly extracts the target object exposed personal information from continuous input video and then blocks the image using the image blurring technique while quickly tracking the extracted object using a prediction algorithm. The proposed method first extracts the target object exposed the personal information region from the input video using a neural network-based learning algorithm. The detected object is then blocked by applying image blurring while being quickly tracked using a position prediction algorithm. Experimental results show that the proposed method effectively protects personal information by quickly tracking and blurring the target object exposed personal information from the received video. The proposed method is expected to be useful for many practical applications such as privacy protection, video surveillance, object detection and tracking.

Keywords: Prediction Technique, Object Covering, Learning algorithm.

1 Introduction

The Internet, which is capable of high-speed wired and wireless network functions, serves as an important data repository that provides various types of information for many users [1]. Therefore, users can easily obtain their desired multimedia materials through the Internet. As such, since the Internet provides useful functions for people, it is one of the indispensable technologies in related fields of application [2]. Moreover, the Internet has become an indispensable and important factor in our recent lives where information and communication technology is rapidly developing.

However, multimedia content containing personal information such as a person's face, body part, social security number, mobile phone number, etc. can be easily obtained through the Internet, causing serious social problems. For example, a video of one's exposed body region may be uploaded and distributed on social network services (SNS) regardless of his or her intention. In addition, images of cell phone contacts recorded on the car window may be automatically collected and used to send spam texts.

Therefore, there is a need for a study that automatically detects areas where personal information is exposed from various kinds of video contents and then effectively blocks the detected areas using image processing techniques such as blurring or mosaic processing [3].

Existing studies for detecting or blocking a target region of interest from an input color image can be found in related literature. However, the existing blocking methods are not so complete yet, and there are various limitations. Furthermore, existing studies related to the detection and blocking of target objects containing personal information are relatively few compared to other research methods.

Therefore, in this paper, we propose a method of robustly detecting a target area that contains personal information from various color images and then effectively blocking the target area using image blurring technique. Fig. 1 shows the overall flow chart of the blurring-based algorithm for the target area including personal information introduced in this paper.

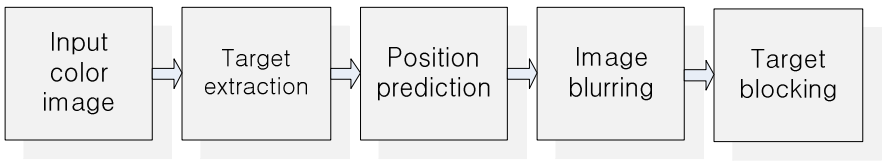


Fig. 1. Overall flow of the proposed method.

As shown in Fig. 1, the novel method proposed in this paper first removes the background regions from the received input color image data and then robustly segments only the target area containing the personal information using the learning algorithm. Then, the proposed method can protect the personal information from being exposed to the outside by effectively blocking the object using image blurring technique while quickly tracking the segmented target object by applying the position prediction algorithm.

2 Learning-based Target Detection

In this paper, only human skin color regions [4] are detected from an input color image except background regions. Then, the face region of the person representing the personal information is extracted from the detected skin color region.

In this paper, we first change the *RGB* color space of the input image to the YC_bC_r color space, which is more suitable for skin color extraction [5]. In general, the YC_bC_r color model is a color space used in an image system of a modern display device and is one of methods of encoding *RGB* information. In order to convert the *RGB* color space to the YC_bC_r color space, scaling and offset adjustment are performed, and can be converted using Equation (1).

$$
\begin{aligned}
Y^{'} &= K_R \times R^{'} + K_G \times G^{'} + K_B \times B^{'} \\
C_B &= \frac{1}{2} \times \frac{B^{'} - Y^{'}}{1 - K_B} \\
C_R &= \frac{1}{2} \times \frac{R^{'} - Y^{'}}{1 - K_R}
\end{aligned} \tag{1}
$$

After transforming the color space of the input image to the YC_bC_r space, the human skin color distribution model, which is created using an elliptical model defined by learning, is used to robustly extract only skin color regions from the input image. An artificial deep neural network is then applied to the detected skin region to robustly extract only the human face region. In general, deep learning is one of the machine learning algorithms in the field of artificial intelligence, and has been actively used in the field of image recognition and understanding in recent years because of its excellent performance. In this paper, we use a deep learning algorithm based on the convolutional neural network (CNN) [6]. Since the CNN model can use two-dimensional data of an image as an input, features related to position in the image are not lost. In this paper, we learn by inputting the color image into a pyramid structure in the learning phase so that the target object can be detected at high speed from the image input in real time.

The rectified linear unit (ReLU) function is used for the activation function [7] instead of a sigmoid function. This is because using the sigmoid function produces a vanishing gradient where the gradient converges to zero [8]. In other words, the sigmoid function has a value between 0 and 1, as shown in Equation (2). Therefore, the deeper the layer of the artificial neural network, the smaller the weight becomes, which causes a problem that the transmitted values disappear. On the other hand, the ReLU activation function has the advantage that the calculation is relatively efficient and the convergence speed is relatively faster than other activation functions.

$$
ReLU(x) = \max(0, x) \tag{2}
$$

In the deep learning model used in this paper, the affine layer means that each neuron of the previous layer is connected to each neuron of the current layer. The affine layer is added before performing the final prediction at the highest output of the CNN. Usually, the affine layer is expressed in the form of $y = f(Wx + b)$, where f is a nonlinear activation function, W is a weight, x is an input layer, and b is a bias. In the proposed CNN model, max-margin object detection (MMOD) does not perform subsampling and optimizes all sub-windows. That is, the object is detected by applying the window scoring function to all sub-windows. Finally, the location of the final target object detection is obtained by summing the areas having the highest score in the six-level pyramid image.

3 Fast Target Blocking

In this paper, we first apply an image blurring technique to block the extracted target area. The proposed method uses a Gaussian function to blur the corresponding area. In general, the Gaussian function replaces the value of the current pixel by using a weighted average of the current pixel value and neighboring pixel values. Therefore, if the blurring using the Gaussian function is applied to the personal information area, the blocking result with less heterogeneity can be obtained. A Gaussian function with an average of (0, 0) in a two-dimensional coordinate space to be applied to an image is defined as Equation (3).

$$G_\sigma(x,y) = \frac{1}{2\pi\sigma^2} e^{-\left(\frac{x^2+y^2}{2\sigma^2}\right)} \tag{3}$$

In this paper, we develop a tracking model of an object to continuously track the target area from the video input in real time. In other words, the proposed method first generates a model for tracking from the personal information area detected from the input image. Then, we use the kernelized correlation filters (KCF) algorithm [9], which is robust to the shape change of objects for continuous tracking. The algorithm utilizes the properties of the cyclic matrix and fast Fourier changes to quickly and accurately track the target area.

Most of the existing studies generate random sub-windows in a certain region from the object position of the previous frame for tracking, extracts features, and apply an algorithm for tracking. In this paper, we extract the feature of the input image and perform tracking using all sub-windows through dense sampling. In dense sampling, a moved sample of an object can be obtained using the Fast Fourier Transform (FFT) without the complicated process.

$$\hat{y} = F^{-1}(F(\bar{k}) \oplus F(\alpha)) \tag{4}$$

The target object model α obtained by training is used to track the target object in the next input image. In Equation (4), $\bar{k}$ is the vector of elements and the kernel correlation between the training image x and the input image z. $\hat{y}$ is the response vector at all positions. The position with the maximum value in the final response map is the tracking coordinate where the next object is located.

4 Experimental Results

The personal computer used for the experiment in this paper consists of an Intel Core (TM) i7-6700 3.4 GHz CPU, 16GB of memory, 256GB Solid State Drive (SSD), and a Galaxy Geforce GTX 1080 Ti graphics card with NVIDIA GPU GP104. The personal computer comes with Microsoft's Windows 10 operating system. In addition, Microsoft's Visual C ++ version 2015 was used as the development tool for the application. In this paper, we collected and utilized various types of images with personal

information to evaluate the performance of the proposed algorithm. These images were taken in various indoor and outdoor environments with no specific constraints.

$$M_{correctness} = \frac{No_{correct}}{No_{total}} \times 100(\%) \tag{5}$$

In this paper, we compared and analyzed the performance of the proposed target area blocking method in terms of accuracy. In this paper, we used the same measure as Equation (5) defined as the ratio between the number of target areas that are robustly extracted and blurred and the number of target areas included in the entire image content. In Equation (5), $No_{correct}$ means the number of target areas correctly blocked using the proposed method. No_{total} means the total number of target areas including personal information in the video content under test. The quantitative measure defined is expressed as a percentage.

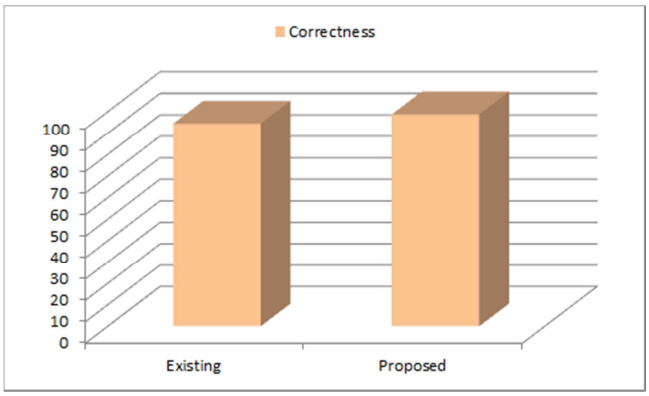


Fig. 2. Performance comparison.

Fig. 2 graphically shows the performance measurement results of the blocking method for the target area in terms of accuracy. As can be seen in Fig. 2, the prediction-based algorithm proposed in this paper blocks the target area containing personal information more effectively than the conventional method.

5 Conclusion

In this paper, we propose a robust method of extracting a target area containing personal information from a variety of video contents input into the system and rapidly protecting the extracted target area by applying a prediction algorithm. In the proposed algorithm, the background area is removed from the received image data and only the target area containing personal information is robustly segmented based on a learning algorithm. Then, the proposed method quickly tracks the segmented target area using the position prediction algorithm and simultaneously blocks it with an

image blurring technique. Therefore, the proposed method could effectively protect personal information from being exposed to the outside. Experimental results show that the algorithm presented in this study robustly extracts the target area including personal information from the input color image content and blocks the extracted area quickly and accurately.

In the future, we plan to stabilize the performance of the proposed algorithm by applying the blocking approach of the target area presented in this paper to CCTV images or adult images. In addition, we plan to protect the content more stably and reliably by varying the area to be blocked in the input image content according to the importance of exposed personal information.

Acknowledgement

This work was supported by the National Research Foundation of Korea(NRF) grant funded by the Korea government (MSIT) (2019R1F1A1056475).

References

1. Yang, X., Wen, Y., Yuan, D., Zhang, M., Zhao, H., Meng, Y.: 3D Compression-Oriented Image Content Correlation Model for Wireless Visual Sensor Networks. IEEE Sensors Journal 18(152018), 6461-6471 (2018)
2. Zhang, S., McCullagh, P., Zheng, H., Nugent, C.: Situation Awareness Inferred from Posture Transition and Location: Derived from Smartphone and Smart home Sensors. IEEE Transactions on Human-Machine Systems 47(6), 814-821 (2017)
3. Dahl, R., Norouzi, M., Shlens, J.: Pixel Recursive Super Resolution. In: International Conference on Computer Vision (ICCV), vol. 1, pp. 5449-5458. IEEE, Venice, Italy (2017).
4. Chakraborty, B.K., Bhuyan, M.K., Kumar, S.: Combining Image and Global Pixel Distribution Model for Skin Colour Segmentation. Pattern Recognition Letters 88, 33-40 (2017)
5. Zhu, S.Y., He, Z.Y., Chen, C., Liu, S.C., Zhou, J., Guo, Y., Zeng, B.: High-Quality Color Image Compression by Quantization Crossing Color Spaces. IEEE Transactions on Circuits and Systems for Video Technology 29(5), 1474-1487 (2019)
6. Amin, S.U., Alsulaiman, M., Muhammad, G., Mekhtiche, M.A., Hossain, M.S.: Deep Learning for EEG Motor Imagery Classification Based on Multi-Layer CNNs Feature Fusion. Future Generation Computer Systems 101, 542-554 (2019)
7. Li, J., Zhang, T., Luo, W., Yang, J., Yuan, X.T., Zhang, J.: Sparseness Analysis in the Pretraining of Deep Neural Networks. IEEE Transactions on Neural Networks and Learning Systems 28(6), 2017.
8. Wang, X., Qin, Y., Wang, Y., Xiang, S., Chen, H.: ReLTanh: An Activation Function with Vanishing Gradient Resistance for SAE-based DNNs and Its Application to Rotating Machinery Fault Diagnosis. Neurocomputing 363, 88-98 (2019)
9. Gong, L., Mo, Z., Zhao, S., Song, Y.: An Improved Kernelized Correlation Filter Tracking Algorithm Based on Multi-Channel Memory Model. Signal Processing: Image Communication 78, 200-205 (2019)

An Empirical Study on Success Factors of Game Industry

Seung Gyun Yoo

Department of Global Economics and Commerce, Dongguk University-Gyeongju, 123, Dongdae-ro, Gyeongju-si, Gyeongsangbuk-do, 38066, Republic of Korea

bluetrade@dongguk.ac.kr

Abstract. With China and Southeast Asia at the center, Korea's game industry is showing meaningful results. This study demonstrated and analyzed the relationship between the characteristics of the basic environment, such as management, technology, market and industry, among Korean game companies. Through this, the company attempted to propose the growth of game companies and the advancement of global markets. As a result of this study, these achievements are being made through the development of leading and technological development in the game industry by identifying the competence of managers and the flow of markets and industries. In addition, securing intellectual property rights to sustain performance and expand the market was one of the important strategies. In other words, the performance of a game company depends on the ability of the manager to provide new story and user services and its success or failure to apply and research and development in technology, market and related industries. If the research so far has been on the external part of the game, i.e. the effectiveness and impact of the game, this study is different in that it comprehensively considers the internal part of the game company and the market and industry.

keyword : Game Industry, R&D, Intellectual Property, Performance, Adaptation

1. Introduction

Like other industries, it is important for the game industry to secure leadership in the market to improve performance. These initiatives arise from a variety of strategies that apply and extend the enterprise's internal and external capabilities. A company's discriminatory strategy lures performance and becomes a driving force to expand the market. Furthermore, achieving standardization through technology development can reduce costs and expand user services.

The purpose of this study is to find success factors to improve management performance, mainly Korean game companies. To this end, related factors were explored by setting up environmental, strategic and performance models. Management, technology, market and industry were established as environmental characteristics, and market application and technology development were established as strategic characteristics. This demonstrated and analyzed the impact of corporate performance using the SPS statistical program.

It is meaningful in that existing studies took into account management and strategic aspects as opposed to exploring the environment and performance, and not just simple analysis of the status quo, but also actual analysis.

2. Literature Reviews

Although we explored the environmental factors of the game industry and the relationship between management strategy and management performance of companies through consideration of prior research, we showed various results rather than general results. This means that a variety of factors, including management capabilities, technical characteristics and environmental characteristics, should be carefully considered in the relationship

between variables that affect corporate performance. Government policy aspects also need to be considered, although they were not included in this study.

The study by JOO (2018) illustrates the differences between content and entertainment businesses and identifies differences in performance by individual businesses. Because content business has a strong nature of knowledge, it suggested that management's will and management strategies are important to improve performance. It also suggested that the issues of technology development and market adaptation are important in the entertainment business. By doing so, relatively small companies are encouraged to expect the effects of their concentration through collaboration and mergers, as qualitative development is difficult. In addition, it proposed multi-use management through the establishment of a consumer-oriented gaming network and distribution platform, rather than a supplier-oriented management method. Kim (2016)'s study argued that building an online platform is necessary for the gaming industry to become globally competitive. He confirmed that the world's top companies have read the future of the game content industry and invested in technology development, and are leading the improvement of performance through the establishment of online platforms. Choi (2016)'s research suggested that the online payment system needs to be overhauled to improve the performance of the gaming industry and technology development that enables mobile-based payment. Choi (2014) argued that the establishment of a transaction-based platform for trading game items is affecting performance improvements. He said that based on analysis of the Chinese market, creative platforms based on transaction base should be developed, and factors of consumer-oriented strategy, market distribution and technology development should be considered.

Lee (2016)'s study argued the importance of technology development for performance improvement, referring to the ability of managers to apply new technologies, such as mobile payment platforms, and adaptations to market flows, for the sake of corporate competitiveness.

Zhang (2011) stressed the importance of forming a relationship between game companies and consumers and argued that building a consumer-oriented management strategy is important.

3. Model and Result

3.1 Model.

The model of this study is as follows.

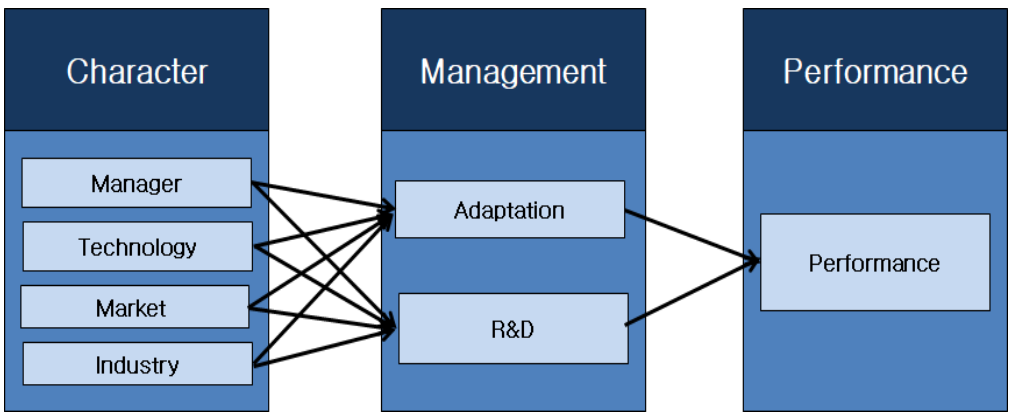


Fig.1. Reserch model

3.2. Results.

The correlation by factors is as follows.

Table 1. Correlation by Factors

Classify	A	B	C	D	E	F	G
Mnanger(A)	1						
Technology(B)	.105(**)	1					
Market(C)	.392	.006(*)	1				
Industry(D)	.068(*)	.541	.462	1			
Adaptation(E)	.008(**)	.006(**)	.036(**)	.545	1		
R&D(F)	-0.732	.054(*)	.598	.002(**)	.635	1	
Performance(G)	.196	.050(*)	.004(**)	.694	.013(**)	.005(*)	1
**(Significance at the 0.01 level), *(Significance at the 0.05 level)							

The conformity assessment index of the research model is as follows.

Table 2. Conformity Assessment Index

Classify	ModelConformityAssessmentIndex		Result
Absolute ConformityIndex	χ^2	Chi-square(degreeoffreedom)	44.462(39df)
	p	Significantprobability	*0.101≥.05*
	Q	Chi-squared/degree-of-freedomratio≤3	1.001≤3
	GFI	GoodnessofFitIndex≥0.9	0.942≥0.9
	AGFI	AdjustedGFI≥0.9	0.793≤0.9
	RMR	RootMean-SquareResidual	0.029≤0.05
	RMSEA	RootMeanSquareerrorofapproximation	0.04≤0.05
Incremental ConformityIndex	NFI	NormedFitIndex≥0.9	0.992≥0.9
	RFI	RelativeFitIndex≥0.9	0.805≤0.9
	CFI	ComparativeFitIndex≥0.9	0.952≥0.9
Simplicity ConformityIndex	PNFI	ParsimoniousNormed-of-FitIndex,	0.593

The route analysis results are as follows.

Table 3. Route Analysis Results

Theory	Path	Path coefficient	Standard error	t	p	Result
1	Manager→Adaptation	.231	.238	1.942	.036	Accept

2	Manager→R&D	-.257	.241	-2.008	.118	Reject
3	Technology→Adaptation	.289	.148	.902	.019	Accept
4	Technology→R&D	.651	.232	2.721	.011	Accept
5	Market→Adaptation	.245	.431	2.541	.032	Accept
6	Market→R&D	-.191	.435	-.491	.515	Reject
7	Industry→Adaptation	.101	.145	.254	.581	Reject
8	Industry→R&D	.513	.269	3.375	.030	Accept
9	Adaptation→Performance	.393	.171	2.571	.004	Accept
10	R&D→Performance	.338	.145	2.581	.007	Accept

4. Conclusion

This study identified the success factors affecting the direction of the market status and future strategies of Korean game companies by exploring the performance based on internal and external environmental factors of the companies. This could be a guide in determining the application of success factors to improve performance and their activities to protect related factors: how to protect intellectual property rights and continued investment in technology development.

Implications through empirical analysis are as follows.

First, managers are required to have state-of-the-art technology, management know-how and competence for cooperation with foreign companies. It should also have the will to develop innovative technologies that can drive the emergence of new consumers. It should support developers' creative activities and invest more in the protection of intangible assets such as intellectual property rights.

Second, differentiated technology capabilities can expand R&D and enhance adaptability in the market. The company's technology concentration has a significant impact on performance improvement and is useful for gaining an edge in collaboration with other companies. However, it is urgent to improve the size of game companies as the vulnerability of game companies is seen as a lack ofprofessional technical manpower.

Third, in order to improve the performance of game companies, market adaptation should be strengthened. This leads to the quick release of new items and the timely release, allowing the removal of psychological distance from consumers. In addition, the game industry must respond quickly to changes in the market environment because it is sensitive to customer differentiation and product awareness.

Fourth, there is a need to demand the government's policy on technology development. Since most of the game industry consists of small and medium-sized enterprises, long-term investments and highly technical applications are difficult. To address this, the integration and sharing of research and development through the government's policy support is required. In particular, R&D on games has a very high degree of development concentration, and the characteristics of production and design methods vary depending on ideas and materials, so we should be able to respond to them through sharing.

Future research seeks to explore the factors that improve a company's performance with a focus on protecting intellectual property rights, which are intangible assets.

References

1. Choi Chang Youl·Hyung Bum Ham, "A study on risk management method for mobile payment", e-Business Review, Vol.17, No.1, 201-221(2016).
2. Choi, Byung-hun, "Taobao Transaction Plateform Analysis: Comparison with eBay Each in China", The Korean

Society of Management Consulting Review, Vol. 5, No.1, 35-54(2014).

3. Creyer, E. H. & W. T. Ross., The influence of firm behavior on purchase intention: do consumers really care about business ethics. Journal of Consumer Marketing, 14(6), 421-32(1997).
4. Fei, Li·Jongho Lee, "A Study on the Customer's Using Intentions of Mobile Payment Services Using NFC Function-Focused on China Market-", e-Business Review, Vol.17, No.1, 51-70(2016).
5. Harish Patil, Brig. Rajiv Divekar, "Inventory Management Challenges for B2C E-Commerce Retailers", Procedia Economics and Finance, Vol. 11, 561-571(2014).
6. Hyukjun Choi, "An Entry strategies on Chinese Online Market of Korean SME's", e-Business Review, 16(3), 221-242(2015).
7. Joo, Hyun Woo. A Study on the Strategic Development Model of the Logistics Platform. Chung-Ang University (Master's Thesis)(2018).
8. Kim, Sun-Nam. An Empirical Study on the Important Factors Increasing the Acceptance Intention of Online International Logistics Platform; (Doctoral Dissertation) Myongji University(2016)
9. Kang, Byung-Mo, "Focus Plan: Using Status of Internet Shopping Mall", The Consumer Age, Vol. 10, 6-7(2000).
10. KeyongSeog Song, "Investigation of Business Model on Fintech Payment System", e-Business Review, Vol.16, No.6, 65- 94(2015).
11. Korea Consumer Agency, Overseas Direct Purchase is more cheaper around 30% but considerable members have dissatisfaction, experienced damage- Preparation of 'User Guideline' for safe countermeasure of Overseas Direct Purchase, Press Release(2014).
12. Maeil Business Newspaper, Cover Story, Attack a Differentiation of Overseas Consumer(2015).
13. Park, Phil Jae, Kim, Jung Duk,「Cross-Border e-Commerce, Is it Silk Road or Mirage?」, Trade Focus Vol.12 No. 40(2013).
14. Yang, Yoon and Cho, Moon-ju, "Effect of word-of mouth communication on Consumer's attitude change", The Korean Journal of Advertising, Vol. 11, No.3, Autumn(2000).
15. Yixiang Zhang, Yulin Fang, Kwok-Kii Wei, Elaine Ramsey, Patrick McCole, Huaping Chen, "Repurchase intention in B2C e-commerce – A relationship quality perspective", Information & Management Vol. 48, 192-200(2011).
16. Young-Shin Sung, Soo-Jung Jung, Sei-Wook Jang, "What do Netizens regard as important when they evaluate the website?", Korean Journal of Consumer and Advertising Psychology, Vol. 2, No. 2, 101-123(2001).

Impact of Corruption on Economic Growth: Focusing on Asia of emerging developing countries

Woo-Sung Cho*, Seung Gyun Yoo**

*Department of International Trade, Dongguk University, Pildong-ro 1gil 30 Jung-gu Seoul, 04620, Republic of Korea
threeonthree@dongguk.edu
** Department of Global Economics and Commerce, Dongguk University-Gyeongju, 123, Dongdae-ro, Gyeongju-si, Gyeongsangbuk-do, 38066, Republic of Korea
bluetrade@dongguk.ac.kr

Abstract. Many countries are making efforts for economic growth. Thus, analyzing the factors of economic growth has become the nation's most important task, regardless of advanced, developing and developing countries. GDP, ex- ports, public aid (ODA), remittances from abroad, educational investment, in- vestment in human resources and direct overseas investment are important economic indicators to be known. Religion, political risk, and many other areas al- so serve as important variables for economic growth. Corruption is also another important area to know for economic growth. Emerging markets in Asia are developing rapidly, modeled after Korea. And there is a situation in which a lot of capital is flowing from foreign countries. Many researchers are criticizing the lack of efficient use of such capital due to corruption in Asian markets. Thus, this paper was launched to find out the relationship between corruption and economic growth in Asia. The paper shows. The decline in corruption has a positive effect on economic growth. This can be said to be a result of the lack of lubrication effect of corruption shown in under- developed countries. Also FDI. ODA, money transfers also have a positive effect on economic growth. This could be seen as a result of an inflow of funds from abroad being used for constructive purposes. In light of these results, emerging economies in Asia will continue to be highly viable markets.

Keywords: Economic Growth, ODA, Corruption, Asia Emerging Market, Developing Country

1. Introduction

Many countries are making efforts for economic growth. Thus, analyzing the factors of economic growth has become the nation's most important task, regardless of advanced, developing and developing countries. GDP, exports, public aid (ODA), remittances from abroad, educational investment, investment in human resources and direct overseas investment are important economic indicators to be known. Religion, political risk, and many other areas also serve as important variables for economic growth. Corruption is also another important area to know for economic growth, many examples show. In most cases, corruption is said to have a negative impact on economic growth. Especially in the case of advanced countries, corruption has a negative impact on economic growth as it undermines the rise of social costs and fair competition. However, in some cases, corruption is not only negative for economic growth. If you do things over your competitors at a certain price, like a ride in an amusement park, these costs could actually help your business. At times like this, corruption acts like a lubricant, which takes a lot of time. Or, it will be a means of smoothly solving areas that are not resolved in principle due to bureaucracy, which will ultimately help economic growth. Thus, corruption can have a negative

or positive impact on economic growth, depending on the perspective of the view, or on the economic conditions of the developing country.

Basically, corruption can have a wealth effect on economic growth, which is that corruption plays a role in one extra cost. According to Mauro (1995), corruption causes low investment and is therefore negative for economic growth. On the contrary, however, corruption sometimes has a positive effect on economic growth. Rather, according to Leff (1964), government bureaucracy has a negative impact on economic growth. In such cases, corruption rather serves as a lubricant to the slow progress made by bureaucracy and various regulations, thus facilitating economic growth.

First of all, we would like to use the Corrupt Perception Index (CPI) released by Transparency International as a variable that affects economic growth. Another variable is FDI, ODA, and remittance.

2. Literature Review

2.1 Corruption

Since 1995, Transparency International (TI), based in Germany, has annually pub- lished figures and rankings for the Corrupt Perception Index (CPI). The CPI is con- ducted as a survey of corruption levels, divided from zero to 100, with the number 100 indicating less corruption, and the number 0 indicates the most corrupt. According to Oh Pil-hwan (2013), the top 10 countries in the CPI index as of 2012 show that, with the exception of Chile and Barbados, per capita income is over $30,000. Previously, data from businesses in charge of business International (BI) and credit evaluations have been mainly used, but in recent studies the CPI, which is now being published by Transparency International, is being used as the most common measure of corruption.

2.2 Literature Review

Mauro (1995) conducted a regression analysis on the impact of investment rates, per capita GDP and corruption on economic growth by dividing bureaucracy into countries with complex forms and not. Both, which were divided into bureaucracy, showed that corruption has a wealth effect on economic growth.

In early research, many studies showed that corruption played a role as a lubricant in implementing government policies, which had a positive effect on economic growth. This was especially true for developing and emerging economies in Asia.

According to Leff (1964), government bureaucracy has a negative impact on economic growth. In such cases, corruption rather serves as a lubricant to the slow progress made by bureaucracy and various regulations, thus facilitating economic growth.

The study by Seo Woo-taek (2012) conducted a study on the impact of ODA on economic growth based on panel data from the Philippines, Indonesia and Vietnam. Vietnam has had a positive impact on economic growth, Indonesia has had no major or statistically significant impact, and in the Philippines, ODA has had a negative impact on economic growth.

In the paper of Cho(2018), the foreign remittance of Mexico and economic growth were analyzed through the Granger causal relationship. The results showed that remit- tances from abroad were significant to economic growth, and that remittances also affected education. This has a positive impact on economic growth and education, as the amount of money that is being transferred from

Mexico is so large that it is similar to FDI.

3. Methodology and Data

3.1 Methodology

$$logGDP_{it} = \alpha + \beta_1 logCPI_{it} + \beta_2 logFDI_{it} + \beta_3 logODA_{it} + \beta_4 logRE_{it} + \varepsilon_{it} \quad (1)$$

The research model used four variables: CPI, GDP, FDI, ODA, and remittance. CPI used CPI index by Transparency International. GDP and ODA used per capita. And FDI and remittance used the entire national amount. Of these, GDP, FDI, and ODA were converted into logs. And i means country, t means time.

3.2 DATA

Seven Asian countries, including the Philippines, Vietnam, Malaysia, Thailand, My- anmar, Indonesia and Cambodia, targeted the Emerging Countries and used 14 years of panel data between 2005 and 2017. Empirical Results

Table 1. Definition of variables

variable	Abbreviations	Unit	Source
logGDP	GDP	per Capita	World Bank Data
CPI	CPI	Index of CPI	Transparency International
logFDI	FDI	total amount	World Bank Data
logRemittance	RE	total amount	World Bank Data
ODA	ODA	per Capita	World Bank Data

A total of five variables were classified as being used in the study. The use of direct overseas investment (FDI), public assistance (ODA) and remit from abroad as varia- bles related to GDP capital inflows indicating economic growth, and finally the Cor- ruption Perceptions Index (CPI) index as indices related to corruption and integrity. The CPI index used data from 1995 by Transparency International and the other four variables used data from the World Bank.

4. Empirical Results

The test results showed that the CPI had a positive effect on economic growth. This is a more transparent society that helps economic growth. Due to the above results, de- veloping countries in Asia have shown no lubrication effect. FDI has been shown to help economic growth. This means that investments from foreign countries are in- vested in industries, which help economic growth. In the case of ODA, the Random effect effect was shown to be conducive to economic growth, but was not statistical- ly significant. This means that ODA from abroad is used for productive parts. In the case of remittances, it was also found to have a positive effect on economic growth

Table 2. Results

Hausman Test	Model	C	CPI	FDI	ODA	RE
	Pooled	0.8188*	0.0263**	0.1841**	-0.0007	-0.0032

43.9975**	Random Effect	-0.3383	0.0109**	0.2448**	0.0006	0.1013**
	Fiexd Effect	-0.8463*	0.006162*	0.24879**	-0.00072	0.1700**

Notes: The p-values are in parentheses.
* indicates 5% level of the statistical significance. ** indicates the 1% level.

5

It can be seen that remittances are used in constructive areas, such as education and family business, rather than simply spending on consumption.

5. Conclusions

Emerging markets in Asia are developing rapidly, modeled after Korea. And there is a situation in which a lot of capital is flowing from foreign countries. Many researchers are criticizing the lack of efficient use of such capital due to corruption in Asian markets. Thus, this paper was launched to find out the relationship between corruption and economic growth in Asia. The paper shows. The decline in corruption has a positive effect on economic growth. This can be said to be a result of the lack of lubrication effect of corruption shown in underdeveloped countries. Also FDI. ODA, money transfers also have a positive effect on economic growth. This could be seen as a result of an inflow of funds from abroad being used for constructive purposes. In light of these results, emerging economies in Asia will continue to be highly viable markets.

References

1. Amit K. Khandelwal and Matthieu Teachout (2016). IGC Policy Note: Special Economic Zones for Myanmar, *International Growth Centre*, pp 1-39(2016).
2. Balassa, B.: Exports and economic growth: Further evidence. Journal of Development Economics 5(2), 181~189 (1978).
3. Barro R. Education as a Determinant of Economic Growth. In: Lazear EP Education in the Twenty-First Century. Hoover Institution Press (2002).
4. Cho, W. S.,: A Study on the Effect of the open Macroeconomics Variables on the Economic Growth in Mexico, International Commerce and Information Review, 20(3), pp 149- 164(2018).
5. Christian Ksoll, Aung Khin Myint and Thein Gi Lwin., Business Process Analysis: Export of Rice and Mango, and Import of Palm Oil in Myanmar, Economic and Social Commission for Asia and the Pacific, 1-184(2013).
6. Dosi, G., Pavitt, K. L. R. and Soete, L. L. G., "The economics of technical change and inter- national trade". Harvester Wheatsheaf, Hemel Hempstead(1990).
7. French, S., "Revealed comparative advantage: What is it good for?", Journal of International Economics, Vol. 106, pp. 83-103(2017).
8. Gujarati, D. N.: Basic Econometrics,New York: McGraw Hill Book Co, (2003).
9. Heckscher, E. F. and Ohlin B., "Heckscher- Ohlin Trade Theory, translated, edited, and introduced by Harry Flam and M. June Flanders Cambridge, Mass. MIT Press(1991).
10. Hla Theingi, Logistic Competitiveness and Export Performance, The Journal of American Academy of Business, 20(2), 101-108(2015).
11. Kim Young-Chun and Ha Jeong-Ju., Business Process Analysis (BPA) of Trade Procedures in the Republic of Korea, 1-170(2015).
12. Leromain, E., and Orefice, G., New revealed comparative advantage: Dataset and empirical

distribution″, CEPII Working Paper(2013).

13. Leromain, E., and Orefice, G., New revealed comparative advantage: Dataset and empirical distribution″, CEPII Working Paper(2013).
14. OECD., ″Globalization, comparative advantage and the changing dynamics of trade. Paris: OECD(2011).
15. Pike, J., Bogich, T., Elwood, S., Finnoff, D., and Daszak, P., ″Economic Optimization of a Global Strategy to Address the Pandemic Threat.″ Proceedings of the National Academy of Sciences of the United States of America Vol. 111 No. 52: pp. 18519-18523(2014).
16. Mauro, P.: Corruption and Growth", Quarterly Jorunal of Economics, 110(3),(1995).
17. Leff, N., "Economic Development through Bureaucratic Corruption", American Behavior- al Scientist. 8, (1964).
18. Suh, W. T.: An Empirical Analysis of the Effects of ODA on the Economic Growth of Developing Countries –with a Focus on Indonesia, the Philippines and Vietnam-, International Developing and Cooperation review4(1), pp57-87(2012).
19. van Hulst, N., Mulder, R., and Soete, L. L. G., (1991), ″Exports and technology in manufacturing industry. Weltwirtschaftliches Archiv, Vol. 127, pp. 246-264(1991).
20. World Bank Group., Economic Rankings, Available from
21. http://www.doingbusiness.org/rankings(2017).
22. World Bank., The economic impact of 2014 Ebola epidemic: Short and medium term esti- mates for Guinea, Liberia, and Sierra Leone. Accessed: http://documents.worldbank.org/curated/en/627851468102871113/pdf/907480REVISED.pdf (2014)

Keyword Network Analysis on North Korean Children's Rights and Welfare

Byung-Man Kim[1]

[1] Kyungnam University, Kyungnamdaehakro 7, Changwon 51767, South Korea
bmkim@kyungnam.ac.kr

Abstract. The purpose of this study is to investigate the rights and welfare of North Korean children through keyword network analysis For this purpose, big data analysis and keyword network analysis were utilized. The results of the study are as follows: it was found that 'North Korea', 'child', 'protect', 'support', 'government', 'children', 'youth', 'relation', 'nation', and 'society' were common keywords in the social awareness of North Korean children's rights and welfare. On the other hand, 'human rights', 'South Korea', 'United Nations', 'UN Convention on the Rights of the Child', 'female', 'world', 'convention', 'Japan', and 'discrimination' were important keywords in North Korean children's rights, while 'North Korean defectors', 'symbol', 'disabled', 'song', 'welfare', 'business', 'history', 'mother-child', 'user', and 'object' were important keywords in North Korean children's welfare.

Keywords: Keyword Network Analysis, North Korea, Children, Rights, Welfare

1 Introduction

South Korea achieved both democratization and industrialization in a short period of time despite the hardships of colonial rule and national division caused by war. However, instability under the continued decentralization system is causing delays in the South's sustainable development, making unification a task that can no longer be delayed. The continuation of the divided situation is damaging the identity of the people by unnecessary waste of national power, the suffering of separated families, the heterogeneity of the two Koreas, and the economic gap between the two Koreas [1]. Furthermore, the issue of social integration caused by the movement of North Korean refugees is causing a sense of alienation and economic burden due to social and cultural differences, thus weakening the perception of unification [2]. This reality will make social integration between the two Koreas difficult after reunification [3]. Therefore, for a smooth reunification, an understanding of the rights and welfare affecting the lives of the people is called upon, along with political, economic and educational preparations [4].

Children are typical social underdogs, and their rights to them are generally thicker than any other object in society. The United Nations signed the International Convention on the Rights of Children in 1989 to specify that the Non-Compliance should ensure active rights for children [5]. The realization of children's rights, especially in the welfare sector, needs to be guaranteed first in terms of the viability level, which is the basic right as a precondition for the realistic functioning of the welfare system.

The reality facing children in different countries is quite different. Dividing the rich children of advanced countries into dichotomy of children and starving North Korean children is simply the result of paying attention to the economic gap. There are also various aspects of children's lives depending on the ideology of running the country and the view of welfare. Even if the system is established based on the same principle for children under the United Nations Convention on the Rights of the Child, the actual life of a child varies depending on the ideology and perspective pursued by a country. The reality of children's lives is largely attributable to the ideology and perspectives that society pursues. This does not mean that children's rights do not have to be protected according to the ideology and perspective of each society. However, social awareness of welfare and systems that guarantee children's rights in each society can be measured in common.

The study aims to look at the social perception of children's rights and welfare in North Korea based on big data analysis. Through this research, we will understand and share the social perception of children who are key members of the peaceful reunification of the two Koreas and provide basic data to related research fields.

2 Research Method

2.1 Research Data

In this study, raw DATA was collected around web DATA, which is provided in three domestic and foreign portal sites and two SNS sites using Textom, a big data analysis solution of The IMC. Since the Moon Jae-in government with high interest in North Korea policy was established in 2017, the collection period was 3 years from 2017 to 2022. Raw DATA was collected as a core keyword of 'North Korea, children, rights' and 'North Korea, children, welfare'. The data collected by 'North Korea, children, rights' is large-scale node (Node = 7,478) data with 2,341 web pages of portal sites, 1,577 blogs, 1,713 news, 1,523 cafes, 94 jisig-IN, and 230 knowledge books. The data collected by 'North Korea, children, welfare' is large-scale node (Node = 8,236) data with 2,094 web pages of portal sites, 1,622 blogs, 1,704 news, 1,303 cafes, 93 jisig-IN, and 1,420 knowledge books [6].

2.2 Analysis Tools

This study used Textom, a big data analysis solution developed by 'The IMC', to collect and refine data on North Korean children's rights and welfare [7]. In addition,

UCINET and NetDraw were used to analyze the network structure between North Korean children's rights and welfare related keywords [8]

2.3 Data Refining

For the purposes of this study, data cleaning was performed to derive meaning from raw DATA, and text mining was performed on the first refinement. In addition, Excel 2016 is used to remove keywords that are not relevant to core keywords, or that are inappropriate for analysis due to spacing. In the second refinement, synonyms or similar words were nominalized and used for analysis based on representative keywords [9, 10].

2.4 Data Analysis

In this study, data on 'North Korean children's rights' and 'North Korean children's welfare' were collected using Textom and the 1st and 2nd refinements were conducted through text mining. The frequency analysis was performed based on the refined data through data cleaning, and the top 20 nodes were selected and a 20 × 20 1-mode matrix data set was created [8]. The network analysis was divided into a micro level and a macro level [9], and the procedure is as follows. First, node, density, average connection distance, number of components, diameter, and network centralization were analyzed to identify network attributes. Second, a single sample mean test using bootstrapping was conducted to test the statistical significance of 'North Korean children's rights' and 'North Korean children's welfare'. Third, in order to understand the structural characteristics among the nodes in the network, we analyzed degree centrality, concor [7, 8]. Fourth, NetDraw was used to visualize each network [9, 10].

3 Results

3.1 Frequency of keywords related to North Korean children's rights and welfare

The frequency of keywords related to North Korean children's rights and welfare is shown in Table 1.

Table 1. Frequency analysis result (top 20 node)

R	**North Korean children's rights**	***N***	***R***	**North Korean children's welfare**	***N***
1	North Korea	1,652	1	support	1,520
2	human rights	823	2	North Korean defectors	1,397
3	South Korea	739	3	North Korea	972
4	child	669	4	youth	819
5	protect	619	5	child	786
6	support	599	6	symbol	660

7	United Nations	577	7	disabled	660
8	UN Convention on the Rights of the Child	565	8	song	651
9	government	554	9	welfare	498
10	children	520	10	children	486
11	China	452	11	protect	478
12	youth	449	12	business	466
13	female	426	13	history	438
14	relation	425	14	government	426
15	world	419	15	mother-child	406
16	convention	405	16	user	406
17	Japan	403	17	relation	405
18	nation	393	18	object	404
19	discrimination	389	19	society	394
20	society	377	20	nation	381

As shown in Table 1, the result of frequency analysis on keywords related to North Korean children's rights, a total of 25,852 keywords were extracted, and among them, target nodes were selected based on 20 keywords. In addition, as a result of frequency analysis on the keywords related to North Korean children's welfare, a total of 29,255 keywords were extracted, and among them, target nodes were selected based on 20 keywords. Table. 1 shows the top 20 keywords in the frequency analysis results for keywords excluding search terms. As a result of frequency analysis on North Korean children's rights, North Korea (1,652) was the most frequent, followed by human rights (823, South Korea (739), and child (669). On the other hand, in North Korean child welfare, support (1,520) was the most frequent, followed by North Korean defectors (1,397), North Korean (972), and youth (819).

3.2 A keyword network analysis of North Korean children’s rights and welfare

The macro-level network structure for North Korean children’s rights is shown in Fig. 1.

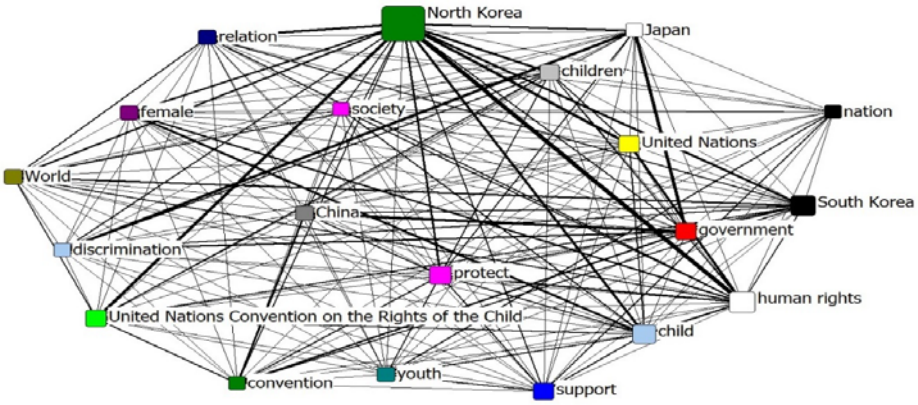


Fig. 1. North Korean children's rights

As a result of analyzing the structural characteristics of network analysis in North Korean children's rights, the nodes were 20, the density was .979, the average degree was 18.600, the average connection distance was 1.021, the number of components was 1, the diameter was 2, and the network centralization was 26.597%.

As a result of testing the statistical significance of the network of North Korean children's rights, the average sampling distribution of the network data was 56.6579 and the standard error was 3.2331. As a result of calculating the Z-score, the probability that the network data for integration early childhood education and childcare are greater than the Z-score is 0.0002 at Z = 6.9462, and the relationship between the network data at the significance level of 5%.

The macro-level network structure for North Korean children's welfare is shown in Fig. 2.

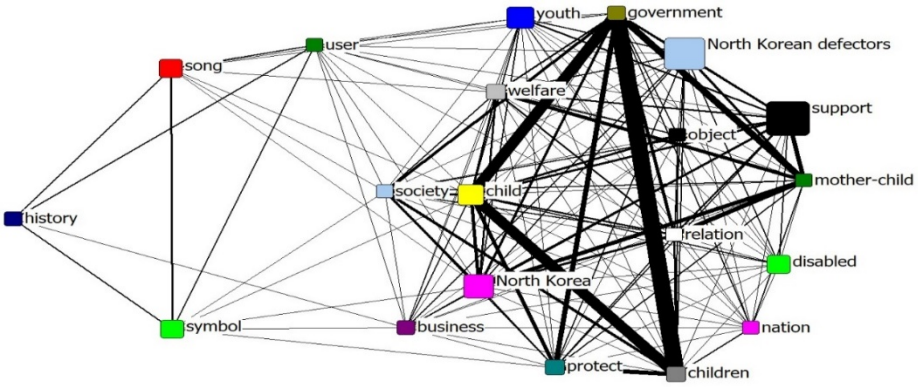


Fig. 2. North Korean children's welfare

As a result of analyzing the structural characteristics of network analysis in North Korean children's welfare, the nodes were 20, the density was .779, the average degree

was 14.800, the average connection distance was 1.221, the number of components was 1, the diameter was 2, and the network centralization was 12.169%.

As a result of testing the statistical significance of the network of North Korean children's welfare, the average sampling distribution of the network data was 295.6526 and the standard error was 37.2516. As a result of calculating the Z-score, the probability that the network data for integration early childhood education and childcare are greater than the Z-score is 0.0044 at Z = 3.3800, and the relationship between the network data at the significance level of 5%.

In order to examine the micro-level network characteristics of North Korean children's rights and welfare, we conducted a centrality analysis of the top 20 keywords, and the results are shown in Table 2.

Table 2. Frequency analysis result (top 20 node)

N	**North Korean children's rights**	***Degree***	***N***	**North Korean children's welfare**	***Degree***
1	North Korea	40.022	14	government	16.149
2	human rights	23.506	10	children	12.283
4	child	19.627	5	child	12.118
9	government	18.325	15	mother-child	8.850
3	South Korea	17.119	3	North Korea	6.041
7	United Nations	16.379	1	support	5.553
5	protect	15.529	11	protect	5.521
10	children	14.021	17	relation	5.221
17	Japan	13.788	19	society	4.151
19	discrimination	13.309	9	welfare	4.120
8	UN Convention on the Rights of the Child	13.213	2	North Korean defectors	2.718
13	female	12.843	18	object	1.756
11	China	12.719	4	youth	1.415
15	World	12.034	20	nation	1.261
16	convention	11.225	7	disabled	1.047
14	relation	10.170	12	business	0.934
6	support	9.471	6	symbol	0.774
12	youth	8.374	8	song	0.729
20	society	8.292	16	user	0.609
18	nation	5.126	13	history	0.524

As shown in Table 2, the results of the centrality analysis of keyword network on for North Korean children's rights and welfare were based on the top 20 keywords and the standardized values were used to take into account the influence of the network scale.

The centrality analysis excluding the search terms showed that the highest degree centrality was North Korea (40.022), followed by human rights (23.506), child (19.627), government (18.325), South Korea (17.119), United Nations (16.379), protect (15.529), children (14.021), Japan (13.788), and discrimination (13.309) in North Korean children's rights, while the highest degree centrality was government (16.149), followed by children (12.283), child (12.118), mother-child (8.850), North Korea (6.041), support (5.553), protect (5.521), relation (5.521), society (4.151), and welfare (4.120) in North Korean children's welfare.

4 Conclusions

The purpose of this study is to provide basic data for policies related to North Korean children's rights and welfare by looking at the social perception of North Korean children's rights and welfare through keyword network analysis. The results of this study are as follows.

First, the result of frequency analysis North Korea was the most frequent, followed by human rights, South Korea, and child in North Korean children's rights, while support was the most frequent, followed by North Korean defectors, North Korean, and youth in North Korean child welfare.

Second, the results of network analysis on North Korean children’s rights and welfare are summarized in two. First of all, analysis of structural attributes for network analysis in North Korean children's rights showed 20 nodes, density of .979, average degree of 18.600, average connection distance of 1.021, component number of 1, diameter of 2, networks centralization of 26.597% respectively. ‘North Korea’, ‘human rights’, ‘child’, ‘government’, ‘South Korea’, ‘United Nations’, ‘protect’, ‘children’, ‘Japan’, and ‘discrimination’ among the keywords that appeared in North Korean children's rights are high in the centrality indicators such as connection centrality. In other words, the main keywords of North Korean children's rights can be seen as ‘human rights’, ‘government’, ‘United Nations’, ‘protect’, and ‘discrimination’. Based on these keywords, we can identify keywords with high connection strength to North Korean children's rights. Therefore, it was found that the theoretical discussion of North Korean children's rights was centered on ‘human rights protect’, ‘Anti-discrimination policy’.

Next, analysis of structural attributes for network analysis in North Korean children's welfare showed 20 nodes, density of .779, average degree of 14.800, average connection distance of 1.221, component number of 1, diameter of 2, networks centralization of 12.169% respectively. In addition, ‘government’, ‘children’, ‘child’, ‘mother-child’, ‘North Korea’, ‘support’, ‘protect’, ‘relation’, ‘society’, and ‘welfare’ among the keywords that appeared in North Korean children's welfare are high in the centrality indicators such as connection centrality. In other words, the main keywords of North Korean children's welfare can be seen as ‘mother-child’, ‘support’, ‘protect’, ‘relation’, ‘society’, and ‘welfare’. Based on these keywords, we can identify keywords with high connection strength to North Korean children's welfare. Therefore, it was found that the theoretical discussion of North Korean children's welfare was centered on ‘Promoting mother-child relationships’, ‘welfare support’ , and ‘social protection’.

These results suggest that ‘North Korea’, ‘child’, ‘protect’, ‘support’, ‘government’, ‘children’, ‘youth’, ‘relation’, ‘nation’, and ‘society’ are key keywords in both North Korean children's rights and welfare. Therefore, North Korean children and government protection and support should be considered important in making North Korean children's rights and welfare policies. In this study, it is meaningful to grasp the key keywords related to North Korean children's rights and welfare by using keyword network analysis and to grasp specific tendency of North Korean children's rights and welfare.

Acknowledgments

This work was supported by Kyungnam University Foundation Grant, 2019

References

1. Institute for Unification Education. I understand the unification issue. Seoul: Department of Education Development at the Unification Education Center (2017).
2. Lim, S. B.: Unified Unification and the Challenges of Christians 2. Yeyeong Communication, Seoul (2003).
3. Heo, E. Y.: Development of a career guidance program for the newly settled youths. Doctoral dissertation. Korea University of Technology and Education, Chungcheongnam (2009).
4. Lee, J. R.: Education vision of a unified Korea, an education welfare state where the next generation is happy. Hanul Academy, Gyeonggi (2016).
5. Hwang, H. I., Nam, M. K., Seo, B. S., Kim, B. M.: Children’s Rights and Welfare. Jungminsa, Gyeonggi (2016).
6. Yousefi, N. R., Sale, J. E. M., Marin, A., Ross, L. E.: Social network analysis: An example of fusion between quantitative and qualitative methods. Journal of Mixed Methods Research 14(1), 110-124 (2020).
7. Kang, S. J., Lee, J. E., Jung, H. Y., Lee, Y. S.: A study on exploring educational implications through post-human trend analysis: Focusing on keyword network analysis. Korean Society of Children’s Literature & Education 19(3), 399-421 (2018).
8. Borgatti, S. P., Everett, M. G., Johnson, J. C.: Analyzing social networks. Thousand Oaks: Sage Publications (2013).
9. Kim, B. M.: Network Analysis of the Social Discourse of Integration of ECEC, Based on Big Data. Journal of Educational Innovation Research 29(2), 17-39 (2019).
10. Kim, B. M., Youn, J. J., Jeong, S. J., Lee, K. H.: Keyword network analysis for creativity and personality of early childhood teachers: Focus on big data. Indian Journal of Public Health Research & Development 9(11), 850-856 (2018).

Topic Modeling of News articles for Low Birth Patterns in South Korea

Yoojin Shon[1] and Bosoon Seo[2]

[1] Dong-eui University, Umgwangro 176, Pusan 47340, South Korea
[2] Dong-eui University, Umgwangro 176, Pusan 47340, South Korea
sbsboss@hanmail.net

Abstract. The purpose of this study was to analyze the main topics of media reports reported in the last 15 years (2005-2019) regarding low birth rates. To this end, 70,488 related news reports reported in 15 years were used. By applying the LAD technique, we extracted the topic of low fertility phenomenon and analyzed the major issues every 5 years. As a result of the study, the main words cited in the low birth rate media articles were fertility rates, children, housing, jobs, services, etc., and divided into 3 cycles divided into 5 year cycles from 2005. The second cycle passes women, children, gender equality, and future society, but the third cycle expands into an aging society, life cycle, and jobs, and it can be seen that the contents related to low birth rate gradually expand to problems related to life in general. It was found that women-related problems were prominent, and that this suggests that the low birth rate policy is a national system that improves human happiness and the quality of life throughout life, beyond simply supporting childcare.

Keywords: Low Birth Rate, Press, Big Data, Topic Modeling Analysis

1 Introduction

Koreans have been experiencing low fertility rates, where the fertility rate is below the population replacement level for more than 30 years, and entered ultra-low birth countries in 2001 [1]. In the future, the population wall will become a reality, and the total population will decrease within a few years, and the decline will accelerate, and in 2065, the elderly population is expected to reverse the production-age population [2]. The progress of low fertility leads to a decrease in the size of the population and a weakening of the national growth engine and competitiveness due to the imbalance of the demographic structure, a slowdown in the national economic growth due to a decrease in the productive population, an increase in the burden of social security equipment for the production age group, and social problems due to imbalance between generations. It sparked concerns over a wide range of crises [3-5].

As the national crisis theory of low birth rate was raised, the movement to find a cause for not giving birth and a solution to give birth was heated. Although approaches to countering low birth rates have been around for a long time, the clues to solving them

are becoming more unclear. In this regard, low fertility is mentioned as one of the social phenomena [6, 7] that reveals complexity, uncertainty and instability. Social phenomena and problems have many components constituting the system create new collective traits through self-organization, and constantly fluctuate and dynamic evolution occurs in the network [8]. The low birth rate is moving through our environment, institutions, organizations, the media, policies, institutions, research, and networks and communication among members of society, and our society is constantly generating abundant discussions on this phenomenon.

Discussion of the low birth rate should be conducted in terms of structure and function of the interpretation and opinion of the members. In our society, it is necessary to try to improve the understanding of phenomena by comprehensively reviewing the numerous issues that are created in relation to the low birth rate and the objects, concepts, and languages discussed. It is necessary to pursue an objective basis such as technical statistics, a subjective basis starting from personal experience, and a mutually subjective basis based on social consensus, to pursue mutual objectivity that is dynamic, communicative, and mutually defined in a huge context [6].

To understand the low birth rate, a dynamic and complex situation, the study seeks to analyze news articles [9] that show how these issues develop, with a focus on various problems and conflicts in society. Topic modeling technique is an analysis method for identifying topics presented in a document, assuming that one document is a mixture of topics [10]. It divides a lot of data into minimal meaningful units to identify the central theme [11-13], and can help you understand the phenomena through discovery of hidden structures in documents, extracting related words, identifying similarities and differences in topics, and comparing them by timing.

Thus, this study will draw the topical features of the press reports related to the low birthrate, examine the overall trends of issues, policies and discourse in our society, analyze the characteristics and structure of objects or concepts that are focused, and compare them over time. By analyzing the trend of topics related to low birth rates, we want to expand our understanding of the status quo and contribute to the creation of multifaceted policies and follow-up research.

The research questions for the purpose of this study are as follows.

1. What is the main topic of media coverage regarding the low birthrate?
2. What changes do media reports indicate depending on the timing of the low birthrate?

2 Research Method

2.1 Data Collection

The study collected news articles related to the low birthrate reported for 15 years (January 2005 to March 2019) through the Korea Press Promotion Foundation's Big Kinds System (https://www.kinds.or.kr/), a media information database. The database is building a news gathering system that integrates articles from 53 media outlets, including the comprehensive daily, economic, regional daily and broadcasting companies. After searching for articles including the words "low birth" in the Big system, they

were saved in the R statistical package, and were selected for analysis except for articles with overlapping titles. In addition, for the purpose of comparing by timing, the five-year cycle of the "Basic Law on Low Birth and Aging Society" was set up in 2005 and the Presidential Committee on Low Birth and Aging Society was launched, the first "Basic Plan for Low Birth and Aging Society" (2006-2010), the second Plan for Low Birth and Aging Society (2016-2020), and the announcement of a roadmap for low birth rate [14]. The number of articles to be analyzed by timing totaled 70,488 with 13,550 in 2005-2009, 22,492 in 2010-2014, and 34,446 in 2015-2019.

2.2 Data processing

The collected article was entered after word refining. First, words were refined in the calibration work, and the second control work had the same words in one word (e.g., daycare, childcare facilities/education costs, tuition/welfare, health and welfare, health and welfare ministries, etc.) that were different in detail. The 362 words list, corrected and controlled through the first refining process, was reviewed for appropriateness by three professors. As a result of the review of appropriateness, experts presented unified opinions on 60 types of words, reflecting them, and incorporated them into 28 types and then modified them to the final 330 words.2.

2.3 Data analysis

Data analysis was performed using the R program for Topic Model Topic modeling is a non-supervised method that minimizes researchers' perception of what topics the article contains without prior knowledge of the content, making it useful to organize long-term, vast amounts of data around the subject [12]. To implement the topical modeling of big data, the study leveraged the Latent Dirichlet Allocation (LDA) to deduce the parameters of the probability distribution from which the raw data was generated from the observed data to discover the key topics of vast text [13] and to identify the interrelationships between topics. The number of topics was chosen by the researchers to specify random numbers, taking into account the interpretability and suitability of the field. After reviewing the correlation between the number of topics and the level of confusion, the number of topics was set to 17 for which the level of confusion decreased continuously and analyzed. Of the 17 topics, the top 7 topics of the Topic Probability Distributions ($p > 0.45$) were selected and the topic names were given by reviewing the configuration words for each topic.

3 Results

We looked at the top 20 words detected throughout 15 years and by timing to look at the words that often appear in the low birthrate articles. The analysis of the words "Korea" (total 18,062 times), "birth rate" (total 14,233 times), "affiliate" (total 11,342 times), "child" (total 10,767 times) and "president 10,609 times) showed high frequency over the entire period. The following are the topics that analyze the words that appear in the article on the low birthrate by cycle.

3.1 Analysis of the Topics of the 1st, 2005-2009 Articles.

The main words for each topic extracted from an article reported from 2005 to 2009 are shown in Table 1.

Table 1. Key concepts and topics in 2005-2009

	Concept	Words
Topic1	National challenges	President, government, nation, candidate, Korea, society, economy, Prime Minister, politics, Grand National Party, Commitment, Person, country, election, lawmaker, solution, future, World, representation
Topic2	World trends	Korea, economy, Japan, Enterprise, government, invest, America, society, growth, World, market, Aging society, labor, country, dollar, future, China, industry, View
Topic3	Birth-rate and Population Outlook	population, increase, decrease, Fertility rate, Korea, female, furniture, Aging society, World ration, View Elderly, newborn baby, Research, prediction, Japan Seoul, birth
Topic4	birth support system	Support, Child, provision, infertility, Insurance, descendant, Expansion, Pregnancy, hospital, a newborn baby, family, Business, Work, Government, health, mothers, Childbirth incentive, Day care Center, Burden
Topic5	the expansion of government-led policies	Government, counterplan, Budget, Support, Propulsion, Expansion, minister, plan, the public, resources of money, Ministry of Health and Welfare, Society, President, Burden, National Assembly, system, Work
Topic6	a change in the educational population	school, student, Education, Reduction, university, Class, kindergarten, teacher, plan, Area, High school, the Department of Education, garden, Education Office, Work, school year, elementary school student, operation

		Child, family, Society, the female sex, descendant, overcome
Topic7	Family and Women	upbringing, Korea, marriage, man and wife, Government, parents, the birth rate, Pregnancy, Support, Day care Center,. school, a person, a newborn baby

The main topics from 2005 to 2009 were 'National Solving Challenges', 'Preparation for the World's Future Society', 'Analysis of Birth Rate and Population Forecast', 'Birth Assistance System', 'Extension of Government-led Measures', 'Changing the Population' and 'Family and Women'. Media reports from 2005 to 2009 highlighted the advent of the low birthrate era due to statistical surveys on the birth rate and a decrease in the number of people in education. Seeing this phenomenon as a crisis facing the nation, it represented the trend of exploring other countries' social movements and response strategies amid the global trend toward an aged society. At the same time, it was discussed to focus on families and women, and to establish a system to support childbirth itself, which could confirm that there were many voices calling for the government to take the lead.

3.2 Analysis of the Topics of the 2nd, 2010-2014 Articles.

The main words for each topic extracted from an article reported from 2010 to 2014 are shown in Table 2.

Table 2. Key concepts and topics in 2010-2014

	Concept	Words
Topic1	politics issue	Government, President, Welfare, Candidate, Budget, Korea, Support, Society, National Assembly, Democratic Party, child care expenses, pledge, Country, Grand National Party Representative, Expansion, Child care
Topic2	Support System for Birth and Child Care	Support, Child, Expansion, family, descendant, provision, child care expenses, Business, Budget, Target, Work, Income, upbringing, Burden, a newborn baby, Day care Center, Government, furniture, man and wife, Childbirth incentive

Topic3	consideration of quality of life	Child, family, marriage, Korea, Society, children, the female sex, descendant, a person, man and wife, upbringing, parents Happiness, Pregnancy, Work, love, health, Mom, Education the masculine gender
Topic4	the pursuit of housing stability	Housing, real property, furniture, Population, Government, Society, Korea, Supply, Area, an aged society, City, Increase market, Business, Seoul Metropolitan Area, Seoul, Work, Propulsion, Welfare, Change
Topic5	Birth Rate Analysis and Population Forecast	Population, Reduction, Increase, Korea, an aged society an old man, the birth rate, the female sex, furniture, Outlook, Ration, Society, marriage, single tax, employed person, expectation, a numerical statement, Record, a newborn baby, productive population
Topic6	Preparing for the Future of the World	Korea, Government, Economy, Society, Japan, the public, Welfare, Country, an aged society, job, President, world, Growth, Corporation, Population, Future, the United States, Increase, Labor, Expansion
Topic7	the foundation of an aged society	Society, Government, Korea, Insurance, an aged society, an annual pension, Corporation, the public, National pension, an, old man, Service, Economy, Welfare, Country, Investment, Support, President, the female sex, job, Expansion

The topics derived from 2010 to 2014 are "key political issues," "birth and child support systems," "consideration of quality of life," "seeking housing stability," "an analysis of fertility rates and population prospects," "preparing for the future of the world" and "the foundation of an aged society." Similar to 2005-2009, the analysis of the birth rate and future population trends was frequently mentioned in articles from 2010 to 2014, and the current situations and countermeasures of each country were introduced in conjunction with the areas of politics, economy, society and labor. The low birthrate has also been found to be a key topic in the government's welfare policy and party pledges. Meanwhile, reports on the low birth rate in this period expanded

from support for women, pregnancy and childbirth, and the whole life was discussed, including upbringing, housing stability, health, work and happiness. In addition, as the aging phenomenon was mentioned in line with the low birthrate, the topic was composed of keywords on the basis of an aged society, such as welfare, services and jobs.

3.3 Analysis of the Topics of the 3rd, 2015-2019 Articles.

The main words for each topic extracted from an article reported from 2015 to 2019 are shown in Table 3.

Table 3. Key concepts and topics in 2015-2019

Concept		Words
Topic1	the foundation of an aged society	National pension, Insurance, an annual pension, Government, an old man, the public, Income, Burden, an aged society, Society, Korea, Assurance, System, descendant, Welfare, inside the room, Impression, Criteria, Fund, provision
Topic2	a change in the educational population	school, Student, Education, Day care Center, university, kindergarten, Child, Reduction, Support, the Department of, Education, Class, Area, teacher, Government, Place, children, High school, operation, classroom, educational expenses
Topic3	a government-led support system	Budget, Government, job, Support, President, Expansion, National Assembly, counterplan, the public, a budget bill, Welfare, Compilation, Business, provision, Child, Income, Youth, Work, child allowance
Topic4	efforts for gender equality	the female sex, upbringing, Government, Child, Parental, leave, family, Support, Society, marriage, descendant, time, the birth rate, Korean, the masculine gender, System, counterplan, Work, Expansion, Corporation, work
Topic5	the pursuit of housing stability	a newly married couple, Housing, Support, Supply, Residence, descendant, Seoul Metropolitan Area,, real prop-

		erty, Youth, Government,, Expansion, Business, Happiness House, marriage, plan, Propulsion
Topic6	Analysis of Birth Rate and Population Forecast	Population, Reduction, the birth rate, a newborn baby, Korea, Increase, an aged society, the female sex, marriage, Record, Society, Outlook, Criteria, an old man, Area, a, numerical statement, counterplan Marriage, Nationwide
Topic7	life-cycle support	job, Business, Support, Area, plan, Society, Government, Work, Youth, Expansion, Creation, Enhance, City,, Population, operation, Welfare, an aged society, Corporation, Economy

Topics derived from media coverage during this period were "the basis of an aged society," "the change in the number of people in education," "the government-led support system," "the effort for gender equality," "the pursuit of housing stability," "the analysis of fertility rates and the prospect of population" and "the support based on the life cycle." In 2015-2019 articles, findings on birth and population trends were also presented, and basic data explaining population reduction and aging were presented. At the same time, a series of articles on the subject of changes in educational institutions and future prospects in response to the declining number of education people formed a pulse. Despite the continuous coverage of articles on housing stability and the laying of the foundation for an aged society, the view was expanded to support young people, jobs, welfare, the region, and the economy. And while "female" is still a major keyword, the trend to deal with the movement toward establishing gender equality culture, such as "male," "child leave," "institution," "system," "work" and "work," is also found, indicating that the view of low birth rates is widening. The topics on government-level support systems also confirmed that access to the macro system is continuously pursued.

4 Conclusions

The South Korean government defines the low birth rate as a national crisis to overcome the low birth rate, and despite a flurry of policies and support, Korea is still a very low birth rate country. In this study, we hoped to study the low birthrate from a macro perspective and explore the relevance between various factors. To this end, the government aims to analyze the major topics of media reports on the low birth rate for the past 15 years (from 2005 to 2019). To this end, the Latent Dirichlet Allocation (LAD) technique was applied to 70,488 related news reports for 15 years to extract topics of low birth rate phenomenon and to analyze major issues every five years. Based on the research results, the following are discussed.

First, the most frequently mentioned words in 70,488 articles were "Korea," "job," "birthrate," "president," "relative," "child," "service," "aged society," "newborn" and "story" words. And as 2019 progressed, the word 'difficulty' was mentioned. Second, the topics reported in the press from 2005 to 2009 were named 'National Solving Challenges', 'Preparation for the World's Future Society', 'Birth Rate Analysis and Population Forecast', 'Birth Support System', 'Extension of the Government-led Measures', 'Changing the Population' and 'Family and Women's Ideology'. Taking this phenomenon as a crisis facing the country, the establishment of a system that focuses on family and women and supports childbirth itself was discussed. Third, the topics on low birth rates reported by the press from 2010 to 2014 were named 'key political issues', 'birth and child support system', 'consideration of quality of life', 'seeking housing stability', 'an analysis of fertility rate and population outlook', 'preparing for the future of the world's society' and 'the foundation of an aged society'. Analysis of the birth rate and future population trends were frequently mentioned, and the current situation and countermeasures of each country were introduced in addition to the areas of politics, economy, society and labor. In addition, as the aging phenomenon was mentioned in conjunction with the low birthrate, the topic was composed of keywords on the basis of an aged society, such as welfare, services and jobs. Fourth, the topics related to the low birthrate, reported in the media from 2015 to 2019, were "the basis of an aged society," "the change in the education population," "the government-led support system," "the pursuit of gender equality," "the prospect of housing stability," "the analysis of fertility rates and the prospect of population" and "support according to the life cycle."

If there were many topics related to changes in educational institutions, such as the birth rate, children, parents, support, and policies, such as care, childcare, and family support through the prospects of education institutions, the views are gradually expanded to support the entire life of young people, jobs, welfare, regions, and the economy, even as articles on housing stability are continuously reported. And many keywords related to "women" have been found, so have the flow to deal with the movement to establish a gender equality culture. This may suggest that the research on low birth rates should not simply support child care and care, or raise the birth rate, but should come up with policies from a life-long macro-level perspective that will improve the quality of life and ensure happiness throughout life.

Acknowledgments

This work was supported by the Ministry of Education of the Republic of Korea and National Research Foundation of Korea (NRF-2018S1A5A2A03033511)

References

1. Statistics Korea.: Census. 15 January 2018 withdrawal from http://kostat.go.kr (2018).
2. Statistics Korea.: Future Population Special Estimate 2017-2067. Daejeon: Statistics Korea (2019).

3. Spodek, B., O. N. Saracho.: International Perspectives on Research in Early Childhood Education (Vol, VII). Greenwich, Connecticut: Information Age Publishing (2005).
4. Park, H. Y., Yeo, Y. K.: A Study on the Causes of Low Birth Rate in Korea and the Direction of Child Care and Child Care Policy. Korea infant and child care, 85, 69-95 (2014).
5. Uchida, H: Age, period and birth-cohort effects on marriage rates in Japanese women between 1985 and 2005, and comparison of trends of effects between marriage and birth rates. Japanese Journal of Public Health 55(7), 440-448 (2008).
6. Davis, D., Baddock, S., Pairman, S., Hunter, M., Benn, C., Wilson, D., Dixon, L., Herbison, P.: Planned place of birth in New Zealand: Does it affect mode of birth and intervention rates among low-risk Women?. Birth 38(2), 111-119 (2011).
7. Kneer, G., Nassehi, A.: Invitation to Niclas Roman: Nicklas Luman's Social System Theory. Galmuri, Seoum (2008).
8. Niklas Luhmann. Law, justice, society. Routlege, London (2009).
9. Shoemaker, P. J., Reese, S. D.: Mediating the message in the 21st century: A media sociology perspective. Routledge (2013).
10. Hwang, H. J.: Exploratory User Experience Analysis Using Topic Modeling: Focused on Amazon Echo Online Review Analysis Case. Master's thesis. Yonsei University, Seoul (2016).
11. Kim, H. J, Kim, S. W., Kim, S. T.: Media coverage topical modeling analysis on smartphone addiction: centered on government policy comparison from 2010 to 2018. Korea Communications Research 104, 38-69 (2018).
12. Kim, Y. K. A newspaper discourse on low birth rates Civil Humanities Literature, 36, 43-100 (2019).
13. Blei, D. M.: Probabilistic topic models. Communications of the ACM 55(4), 77-84 (2012).
14. The Commission on Low Birthrate and Aging Society.: A policy road map for low birth rate and aging society Seoul: Low Birthrate and Aging Society Committee (2018).

Development of Evaluation Scale of Creativity and Personality for College Students Using Delphi Survey Method

Byung-Man Kim[1], Jeong-Jin Youn[2] and Su-Jeong Jeong [3]

[1] Kyungnam University, Changwon 51767, Republic of Korea
[2] Tongmyong University, Pusan 48570, Republic of Korea
[3] Tongmyong University, Pusan 48570, Republic of Korea
jsjs@tu.ac.kr

Abstract. The purpose of this study is to develop a college student's creativity and personality evaluation scale using Delphi survey. For this study, 20 experts in the field of creativity and personality were selected as panels, and the 4th Delphi survey was used to develop the scale for the evaluation of creativity and character for college students. The Delphi survey was conducted three times to calculate the descriptive statistics of the mean, standard deviation, skewness, and quartiles. In addition, content validity, coefficient of variation, degree of consensus, and degree of convergence were derived to correct or delete the content of the evaluation scale based on the opinions of the expert panel. This study analyzed the opinions of the expert panel to derive two evaluation areas and five evaluation criteria consisting of creativity (cognitive factor, disposition factor, motivation factor) and personality (virtuality, core competencies). In addition, through expansion of thinking, convergence of thinking, problem solving ability, openness, independence, curiosity and interest, immersion, eg, filial piety, honesty, responsibility, respect, consideration, communication, cooperation, knowledge of personality practice, communication skills, conflict resolution. A final draft of the creativity and personality scale for college students, consisting of 18 evaluation indicators and 90 evaluation questions, was developed. The creative and personality evaluation scale for college students developed through this study is expected to provide important implications for the academic development and qualitative improvement of the creative and personality areas.

Keywords: creativity and personality, evaluation scale, college students, creativity and personality experts, Delphi survey

1 Introduction

We have to live in the reality of the first industrial revolution that began with mechanization, the second industrial revolution in which mass production using electricity began in earnest, and the third industrial revolution led by the computer information and automation production system, through robots or artificial intelligence. As such, we are at the center of the fourth industry and it is time for us to make efforts to improve our capabilities appropriate to this era. With the expansion of new technologies such as AI and big data in the era of the fourth industrial revolution, AI cannot be acquired, but humans must develop their warm personality and creativity. In other words, the ability to communicate and empathize with each other based on respect and consideration has become a key age.

Future scientists, in particular, cite the "creativity" of the human resources of the Fourth Industrial Revolution, such as the sensitivity of discovering and detecting widespread thinking and problems such as fluency, flexibility, originality, and sophistication, the "creative ability" to transform and reinterpret existing concepts, and the "creative nature and self-regulation" of task-obsession, consideration, challenge, leadership, etc. [1, 2].

These capabilities were generally used to mean the ability to perform a specific job role, but the meaning was extended to a wide range of capacity to cope with the complex problems that arise in modern society while expanding the scope of the OECD DeSeCo Project to lead a successful life in modern society [3]. Thus, the core competency of future social talent was changed to fostering talent with all the capabilities required by modern society.

In response, college education is focusing more on fostering talent with creativity and personality, which is the core competency of future talent, in traditional education centers that focus on knowledge and skill acquisition. Creativity and personality education has become a key part of education in Korea since 2009 when the Ministry of Education, Science and Technology presented creative and personality education as the main policy of national development strategy [4, 5]. Creativity and personality education are fostering both creativity and character so that creativity can be expressed in the right persona, stressing that the direction of education policy should be transformed into education that fosters future talent living with the creation of new values [6]. As the pace of change in society changes rapidly, creativity and personality skills are becoming more prominent in the field of college education.

Just as creativity and character were already having an important influence on college education, it is also necessary to consider whether or not creativity and personality areas should be evaluated and desirable for future direction of creative personality education. In other words, in order for creative and personality nature education to be reborn as a quality education in college education, it is necessary that a reasonable and reliable evaluation should be made based on a correct understanding of the evaluation and diagnosis of creativity and character.

In the preceding study, the study on creativity of college students is mainly about developing and applying programs related to creativity and character [7], and when looking at evaluation scale for creativity and character measurement,

it is mainly about evaluating the general areas of creativity and personality based on a general perspective [8-10]. In other words, few studies have been conducted on the development of evaluation scale that can directly measure and evaluate the concept of 'creative' character' for college students [11].

The purpose of this study is to develop a measure of creativity and character evaluation for college students using Delphi survey. The measure of creativity and personality for college students developed through this study is judged to be meaningful in finding the direction and value that creative and personalityistic academic fields should pursue in the future. It is also expected that this study will provide implications for the development of the creative and personality nature areas and for the enhancement of quality levels, and provide significant data on the academic development of the creativity and personality evaluation fields. Specific research details to achieve the objectives of this study are as follows.

First, check the development direction of creative and personality evaluation measures for college students.
Second, the final measure of creativity and personality evaluation for college students will be developed.

2 Research Method

2.1 Research Subject

Based on the Delphi survey, this study selected experts in creativity, character and evaluation as the panel to be studied. The criteria for organizing the panels that are studied in this study are as follows. First, the panel of experts, the subjects of the study, based on their creativity, personality, and academic nature of the evaluation, involved authoritative experts from various fields of academia. Second, the panel included professors majoring in the above field and field experts to develop evaluation measures that reflect theory and practice on creativity and character. Third, experts' panel was selected through an in-depth review based on their expertise, performance and experience. In addition, 20 panel members were selected to reduce the recovery rate and minimize errors in the survey results. A number of evaluation measures developed using Delphi survey [1-3] selected 20 expert panels and selected 20 expert panels in this study.

First, after drawing up a list of experts suitable for the panel selection criteria of this study, five creative and personality experts with PhDs were reviewed for suitability. Subsequent calls and e-mails were used to guide the purpose of this research and the procedure for conducting the research and then asked for participation. From the first Delphi survey, a total of 20 experts participated in this study, but the subsequent two panels dropped out of the study and 18 (90.0%) participated in the first and fourth Delphi surveys. The general background of the Delphi survey study is shown in Table 1.

Table 1. General characteristics of the expert panel.

	Age	Gender	Gob	Major	Final education		Age	Gender	Gob	Major	Final education
1	49	male	researcher	sociology	doctor	10	43	female	Professor	Child Studies	doctor
2	37	female	researcher	Social welfare	doctor	11	51	female	Professor	Social welfare	doctor
3	38	female	Professor	Early Childhood Education	doctor	12	53	female	Professor	Fashion Marketing	doctor
4	40	female	Professor	Early Childhood Education	doctor	13	35	female	researcher	Social Policy	doctor
5	49	female	Professor	Social welfare	doctor	14	54	female	Professor	Child Family Studies	doctor

6	45	female	Professor	Child Studies	doctor	15	44	female	Professor	Early Childhood Education	doctor
7	35	female	Professor	education	doctor	16	49	female	Professor	education	doctor
8	42	female	Administrator	Social welfare	doctor	17	42	female	Professor	education	doctor
9	57	male	Professor	Computer Science	doctor	18	45	male	Professor	Art engineering	doctor

2.2 Research Tools

2.2.1 Primary Delphi Survey Questionnaire

In the first Delphi survey, the questionnaire was organized with open and closed questions to identify the direction of development of creative and personality evaluation measures for college students. The free speech preparation and suggestion of expert panels is essential for developing the evaluation scale, and for an in-depth understanding of expert panels [1-3]. The open question in the Delphi survey provided an opportunity for various and widespread perceptions of the questions on the panel of experts, thus meeting the purpose of the development direction of the primary Delphi survey in this study [4]. The first Delphi survey questionnaire was composed of questions such as evaluation categories and criteria for developing creative and personality evaluation measures for college students, reasons for selecting evaluation elements, and opinions on developing evaluation measures for college students. The first Delphi survey questionnaire consisting of the above contents was commissioned by five creative and personality experts with Ph.D.s, and examined the validity of the content and composition of the first Delphi survey questionnaire, appropriateness of the response method, suitability to the use of terms, and understanding of the questions.

2.2.2 Second Delphi Survey Questionnaire

Based on the response results from the first Delphi survey, the first draft of the evaluation scale was constructed and the second Delphi survey questionnaire was produced to collect opinions from the expert panel. In the second Delphi survey questionnaire, two evaluation categories, five evaluation criteria, and 18 evaluation indicators were constructed based on the opinions of expert panels replying through the first Delphi survey, related literature, and the results of prior research, and 647 evaluation questions were configured to measure them. The second Delphi survey was required to have a panel of experts evaluate the appropriateness of developing creative and personality evaluation measures for college students, which were constructed through the first results. The second Delphi survey questionnaire consisted of flat questions and was designed to be flat on a five-point Likert scale, and included other comments to provide additional expert panel feedback for each flat-set table. In addition, the evaluation scale also helped to improve understanding of expert panels by summarizing the composition and content of the first draft, and terms that could be used to refer to the evaluation. The second Delphi survey questionnaire, consisting of the above contents, was commissioned by five creative and personality experts with Ph.D.s to review their validity.

2.2.3 Third Delphi Survey Questionnaire

The third Delphi survey was intended by a panel of experts to share the results of the second Delphi survey and to reach an agreement on the results of the survey. In other words, a measure of creativity and character evaluation for college students whose appropriateness was verified through an expert panel was also implemented to develop a second draft. The third Delphi survey questionnaire was constructed based on the results of the second Delphi survey responded by a panel of experts. Technical statistics (average, standard deviation, skewness, quartile) for each item were returned to individual expert panels in a questionnaire, and the expert panel corrected, incorporated, or deleted the lower-biased question (skewness > .0) than the average or the distribution of responses [1-3]. The response results of the expert panel and the opinions of the expert panel raised through other comments have been compiled to modify the question on the scale, and overlapping questions have gone through the process of integration and reorganization process. The third Delphi survey questionnaire comprised two evaluation categories, five evaluation criteria and 18 evaluation indicators, and 159 evaluation questions were constructed to measure them. The third Delphi

survey questionnaire consisted of flat questions and was designed to be flat on a five-point Likert scale, and included other comments to provide additional expert panel feedback for each flat-set table. The third Delphi survey questionnaire, consisting of the above contents, was commissioned by five creative and personality experts with PhDs to review their validity.

2.2.4 Fourth Delphi Survey Questionnaire

The fourth Delphi survey was intended by a panel of experts to share the results of the third Delphi survey and to reach an agreement on the results of the survey. Although evaluation tools are usually developed through the third Delphi survey process, this study had difficulties in the process of integrating and reorganizing the questions, as many details of developing questions on expert panels were presented in the first Delphi ancestor. This led to four rounds of Delphi surveys. The 4th Delphi survey questionnaire comprised two evaluation categories, five evaluation criteria, and 18 evaluation indicators, and 90 evaluation questions to measure them. The 4th Delphi survey questionnaire consisted of flat questions and was designed to be flat on a five-point Likert scale. The third Delphi survey questionnaire, consisting of the above contents, was commissioned by five creative and personality experts with PhDs to review their validity.

2.3 Research Procedure

Using prior research related to this study and focus group interviews, preliminary research was conducted to check the validity of the contents of the preliminary 1st Delphi survey questionnaire, which was composed of the preliminary 1st Delphi survey questionnaire [5] based on the expert's opinion on the direction of developing creative and personality evaluation measures for college students. The preliminary survey was conducted in February 2019 on five early childhood education experts with PhDs. After collecting opinions from childhood education experts through preliminary surveys, the first Delphi survey questionnaire was finally formed by modifying and supplementing the questionnaire.

The first Delphi survey produced an open questionnaire related to the development of a scale of creativity and personality evaluation for college students and distributed it to the subjects by e-mail. The first Delphi survey was conducted for about two weeks from April 1, 2018 to April 14, 2018.

The response results of expert panels collected from the first Delphi survey were analyzed to nominate the evaluation scale, and the secondary questionnaire was constructed by modifying and supplementing the indicator content. The second questionnaire was designed to be flat on a five-point Likert scale, and the conformity and appropriateness of the evaluation area, evaluation criteria, evaluation indicators, and evaluation questions for college students were verified. The second Delphi survey was conducted for about two weeks from May 10 to May 23, 2018 and was distributed in E-mail the same way as the first Delphi survey before receiving the experts' responses. During the investigation, experts' panel participation was frequently encouraged by using methods such as telephone, text and e-mail to prevent omission of expert panels.

The results of the second Delphi survey were analyzed to modify and supplement the evaluation area, evaluation criteria, evaluation indicators, and evaluation questions, and the response results for each item were statistically processed and the third Delphi survey questionnaire was organized. The third questionnaire was also designed to be flat on a five-point Likert scale, and the evaluation area, category, evaluation criteria, and evaluation index representation, clarity, and integration of the measures of creativity and personality for college students were to be responded to. The third Delphi survey was conducted for about two weeks from June 12, 2018 to June 25, 2018, and the responses of experts were retrieved after distributing them in E-mail the same way as the first and second Delphi surveys. Similar to the second Delphi survey, experts' panel participation was encouraged from time to time by using methods such as telephone, text and e-mail to prevent missing expert panels during the investigation.

The fourth Delphi survey was conducted for about two weeks from August 5 to June 18, 2018, and was then distributed in E-mail the same way as the first and third Delphi surveys, and the responses of experts were retrieved. The results of the fourth Delphi survey were analyzed to modify and supplement the evaluation area, evaluation criteria, evaluation indicators, and evaluation questions, and the final plan of the creative and personality nature evaluation measures developed through the above process was finalized.

2.4 Data Analysis

In this study, the contents analysis and the SPS 23.0 Program for Window were used for analysis of data according to the development process of creative and personality evaluation measures for college students. First, the qualitative data collected through the first Delphi survey was compiled by the contents analysis

through SWOT analysis. Quantitative data collected from the 2nd and 4th Delphi surveys were calculated by scoring and statistically processing the equilibrium values of each question. In addition, as proposed by [1], the dodo was calculated to check the tendency of the distribution to the response, and the degree of agreement of the response was confirmed by checking the quartile to determine the density of the response distribution. In addition, the results of the 2nd or 3rd Delphi survey were verified and the 5th Delphi survey results were deleted, corrected, and incorporated questions of less than 4.00 average or greater than or equal to or greater than a five-point scale, questions not compatible with each other as proposed by [2], and other opinions on the panel of experts that were not appropriate, if the CVR was lower than .42 or greater than or greater than or greater than or higher.

3 Results

3.1 Research on the direction of development of creative and personality evaluation measures for college students

Experts' opinions for developing a scale of creativity and personality evaluation for college students presented through the first Delphi survey are summarized as content analysis through the SWOT analysis method. Experts' opinions on the direction of developing creative and personality evaluation scale for college students were analyzed as follows.

First, the strengths of developing creative and personality evaluation scale for college students were 'strength to present inspection norms that fit the steps of college students', 'strength of presenting objective inspection indicators for creative and personality nature education contents', 'strength of directional trend of creative and personal convergence educational contents and strength of result exchange', 'strength of cultural, artistic and creative, and personality factors' factors of multi-faceted characterization, and characterization factors.

Second, weaknesses in the development of creative and personality evaluation scale for college students were 'weakness on the absence of a clear concept of creativity and character integration', 'weakness on the subjectivity of experts' intervention in the evaluation of creative and personality education content', 'weakness on the spot difficulty of inspection and evaluation of specific areas', difficulty of 'weakness of the evaluation scale developed by choice of specific areas' and analysis. Third, the opportunity to develop creative and personality evaluation scale for college students was the opportunity to 'enhance diverse perceptions of culture, art, creativity, and character', 'provide a new paradigm for establishing new surveyors', 'creation, personality evaluation and inspection', 'opportunity to enhance the speciality of creative and personality nature education contents', 'opportunity for the expansion of creative and personality nature education and research' and 'option of social necessity' functions. Fourth, the threat of developing creative and personality evaluation scale for college students was analyzed as 'threats of the possibility of forming stereotypes and misconceptions about creativity and character', 'threats of constant complementation and refluxing of developed creative and personality education evaluation scale', 'efforts to find common denominators of creativity and character' and 'standardized work to increase the credibility of evaluation scale' by college students'.

3.2 Delphi Survey on Developing a Scale of Creativity and Personality for College Students

The first draft of the measures of creativity and character evaluation for college students was derived. The evaluation index adequacy evaluation of the first draft shown in the second Delphi survey comprised two evaluation categories, five evaluation criteria, and 18 evaluation indicators, and 647 evaluation questions were constructed to measure them. The third Delphi survey was conducted after confirming the expert panel's opinions on the evaluation categories, evaluation criteria, evaluation indicators, and evaluation questions, and the expert panel's views on the appropriateness of each question. The appropriateness of the expert panel for the secondary draft was reviewed by the subsequent analysis of the third Delphi survey.

Expert panel, which confirms the results of the 2nd and 3rd Delphi survey analysis, deleted, corrected, and integrated items with an average of less than 4.00 or more than .0 among the five-point scale indicators, which are inconsistent with mutual exclusivity. In case of other opinions, if the CVR is less than .42, consensus is less than .75, and convergence is more than .50, the content has been modified or deleted.

In the fourth Delphi survey, two evaluation categories, five evaluation criteria, 18 evaluation indicators, and 90 evaluation questions were derived from the 3rd Delphi survey, and the suitability of the survey was as follows. First, the fittedness analysis of creativity showed that the mean (M) was 4.17-4.72, and all the wagons were ineligible. The Quartet (Q1-Q3) distribution was 3.00-500 and the CVR was .56-1. The consensus was .75-.94, the Convergence.The 50-63 result showed a high degree of suitability assessed by experts. In addition, the average (M) was high between 4.11 and 4.92, and all the wagons were ineligible. The distribution of quartiles (Q1-Q3) was shown at 3.00-5.00 and the CVR was at .56.00 and the consensus was at .75-.80, and the convergence was .50-.76. The suitability assessed by experts was high.

Through the above four rounds of Delphi survey, the study developed creative and personality evaluation measures for college students, consisting of two evaluation categories, five evaluation criteria, 18 evaluation indicators, and 90 evaluation questions. The final draft of the measure of creativity and character evaluation for college students is shown in Fig. 1, and the details are shown in Table 2.

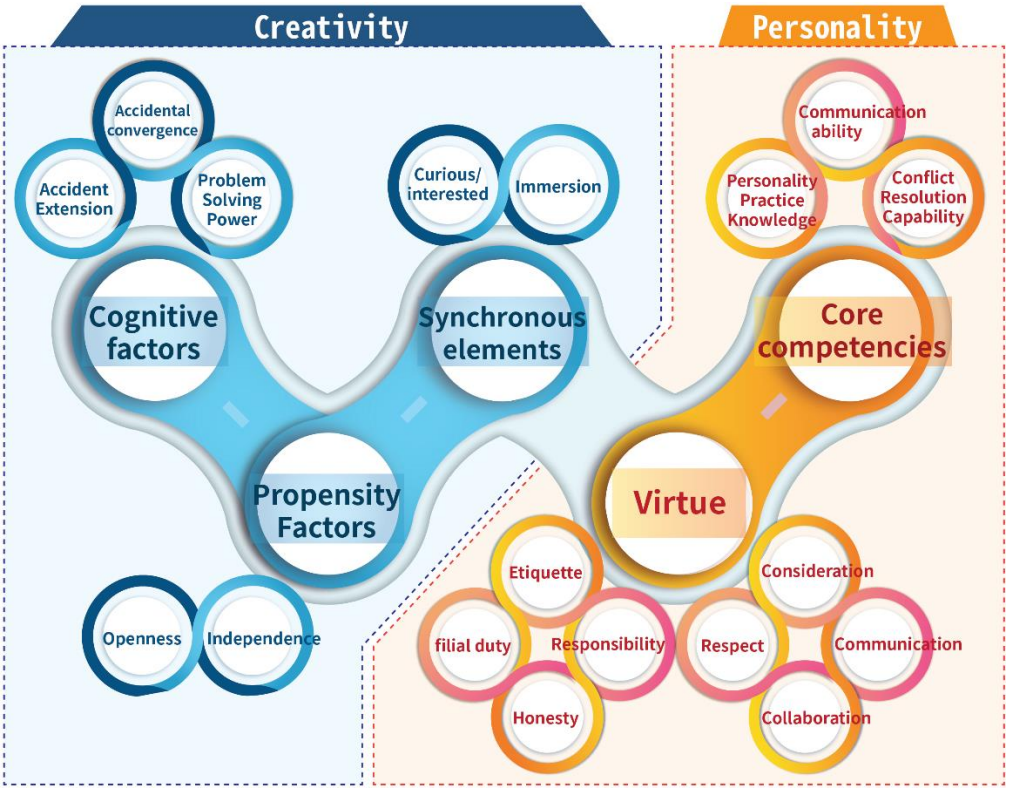


Fig. 1. A Schematic design for the evaluation of creativity and personality for college students.

Table 2. A Final Proposal on the Evaluation Scale of Creativity and Personality for College Students.

Evaluation area	Evaluation standard	Evaluation index	Evaluation question
1. Creativity	1-1. Cognitive factors	1-1-1. Accident Extension	1-1-1-1. When I face a problem situation, I think of several solutions. 1-1-1-2. I think of ways to use things differently when observing things. 1-1-1-3. When I look at an object, I can conjure up various ideas related to it. 1-1-1-4. I often think from a different perspective than the way I think so far. 1-1-1-5. I like to think in various ways.
		1-1-2. Accidental convergence	1-1-2-1. I can effectively classify and organize the various information given. 1-1-2-2. I can sort out the ideas that I need out of the various thoughts. 1-1-2-3. I know how to determine the importance and priority of the information collected. 1-1-2-4. I tend to find the heart of events and problems. 1-1-2-5. I am good at finding the most appropriate of the various information.
		1-1-3. Problem Solving Power	1-1-3-1. I can solve the problem by utilizing the alternatives given when faced with a problem. 1-1-3-2. I actively exchange views with others to solve problems. 1-1-3-3. I integrate the core content or information necessary to solve the problem and obtain specific results. 1-1-3-4. I can solve complex problems or processes through discussion or discussion and easily produce results. 1-1-3-5. I set priorities for solving problems and make plans to complete them one by one.

	1-2. Propensity Factors	1-2-1. Openness	1-2-1-1. Accept others' ideas or opinions without prejudice. 1-2-1-2. Respect someone who has different values and beliefs from me. 1-2-1-3. I'm also interested in the stories of people who think differently. 1-2-1-4. I take other cultures with an open mind. 1-2-1-5. I can fully accept my opinion even if someone else criticizes it.
		1-2-2. Independence	1-2-2-1. I tend to make important plans myself. 1-2-2-2. When I have a choice, I usually decide by myself. 1-2-2-3. I try to carry out my own plans actively. 1-2-2-4. I don't rely on others and I solve it myself. 1-2-2-5. I am not easily swayed by criticism or opposition from others.
	1-3. Synchronous elements	1-3-1. Curious/interested	1-3-1-1. I am interested in new things. 1-3-1-2. I try to find out if I don't know anything. 1-3-1-3. tend to explore the principles or characteristics of things or phenomena. 1-3-1-4. I'm very interested in the various issues around me. 1-3-1-5. I like to have various experiences.
		1-3-2. Immersion	1-3-2-1. I'm lost in time on things I like. 1-3-2-2. When I start one thing, I don't finish it until I finish it. 1-3-2-3. I tend to concentrate well without being disturbed by the surroundings. 1-3-2-4. I don't give up till the end, even if it's a difficult task. 1-3-2-5. I don't have a time limit when I do something interesting.
2. Personality	2-1. Virtue	2-1-1. Etiquette	2-1-1-1. I value the etiquette and norms needed in everyday life. 2-1-1-2. I often hear people around me say I'm polite. 2-1-1-3. I try to abide by public morals and order. 2-1-1-4. When I see an adult on a bus or subway, I give up my seat. 2-1-1-5. I always maintain a polite attitude in relationships with others.
		2-1-2. Filial duty	2-1-2-1. I want to make my parents feel at ease. 2-1-2-1. I am grateful to my parents. 2-1-2-3. I want to understand and practice my parents. 2-1-2-4. I respect my parents' wishes. 2-1-2-5. I support my parents with action to make them feel comfortable.
		2-1-3. Honesty	2-1-3-1. I tend to speak and act truthfully. 2-1-3-2. I think words and actions should match. 2-1-3-3. I would like to express my opinion to others frankly. 2-1-3-4. I'm being honest even if I lose. 2-1-3-5. I'm honest when I make a mistake.
		2-1-4. Responsibility	2-1-4-1. I tend to get my job done. 2-1-4-2. I know what I have to do. 2-1-4-3. I try to keep even the slightest promise. 2-1-4-4. I try to do my job and task responsibly. 2-1-4-5. I try to finish my work to the end, even if the sacrifice comes.
		2-1-5. Respect	2-1-5-1. I think everyone is worth it. 2-1-5-2. Try to understand people who disagree with me rather than criticize them. 2-1-5-3. I care about myself and cherish myself. 2-1-5-4. Regardless of age, gender, nationality, religion, disability, I consider others as precious beings as I am. 2-1-5-5. I treat everyone with an attitude of respect, regardless of rank or position.
		2-1-6. Consideration	2-1-6-1. I always try to help others first. 2-1-6-2. I'm well aware of what others need. 2-1-6-3. When I carry out a common task, I tend to think a lot about other people's 2-1-6-4. I try to help people who are in trouble 2-1-6-5. I take care that my words and actions do not hurt others.

	2-1-7. Communication	2-1-7-1. I usually share a lot of opinions with others on a certain issue. 2-1-7-2. I tend to convey my thoughts or feelings according to the target and situation. 2-1-7-3. I try to have a conversation when I have an issue to solve in my relationship with others. 2-1-7-4. I'm a good listener. 2-1-7-5. I use various channels to exchange views with others.
	2-1-8. Collaboration	2-1-8-1. I try to join forces with members to achieve a common goal. 2-1-8-2. As a member of the community, I tend to carry out my duties well. 2-1-8-3. I enjoy working with other people. 2-1-8-4. When I work with other people, I try to solve them together if I have a problem. 2-1-8-5. I think working together with others will solve difficult problems better.
2-2. Core competencies	2-2-1. Personality Practice Knowledge	2-2-1-1. I tend to adjust and apply my ideas according to the circumstances. 2-2-1-2. I fully exercise my ability when the actual situation occurs. 2-2-1-3. I properly execute my thoughts in real life. 2-2-1-4. I can make a sari judgment on the occasion. 2-2-1-5. I try to do what I think is right.
	2-2-2. Communication ability	2-2-2-1. I have the ability to communicate my intention accurately by using contextual words, expressions, and gestures 2-2-2-2. I'm a good judge of what other people say and do. 2-2-2-3. I think nonverbal attitudes (expression, gesture, etc.) as well as language are important in communicating. 2-2-2-4. I can logically express my opinion to others. 2-2-2-5. I can integrate well into my own group through appropriate language and behavioral expression.
	2-2-3. Conflict Resolution Capability	2-2-3-1. If there is a conflict, I want to resolve the conflict reasonably through dialogue. 2-2-3-2. I try to solve problems with other people without avoiding them. 2-2-3-3. I can coordinate opinions even in a conflict of opinions. 2-2-3-4. I am good at persuading and mediating others in case of trouble. 2-2-3-5. I tend to suggest ways to accommodate the needs of several people in conflict situations.

4 Conclusions

The purpose of this study is to develop creative and personality evaluation scale for college students using Delphi survey. To this end, 18 panels of experts in creative and personality nature were conducted on four occasions. Based on the main results of this study, the following is the conclusion.

This study analyzed the opinions of expert panels to derive two areas of evaluation and five evaluation bases consisting of creativity (intellectual, propensity, synchronous) and personality (goods, core competencies). In addition, the final draft of the creative and personality nature test district for college students consisting of 18 evaluation indicators consisting of the ability to expand thinking, converge thinking, problem solving, openness, independence, curiosity/interest, interest, honesty, responsibility, respect, consideration, cooperation, personality practice knowledge, communication ability, conflict resolution, and 90 evaluation questions was developed.

As experts' opinions on developing a measure of creativity and personality for college students show that the development of creative and personality test tools for college students can be analyzed from the strengths of developing tools for college students to the strengths of the test standards for college students, [16] also showed that efforts to measure creative talent capabilities for college students should be considered essential in designing college education. In other words, it suggests that the concept and measure of "creativity and personality nature" of college students should be presented so that they can be understood simultaneously in social context and social relationships, and established as sustainable indicators that can be accepted in the future.

It is also meaningful to develop creative and personal evaluation scale for college students, as suggested by the preceding research [17] that the tools for measuring creativity and character of college students are necessary and reliable.

Creativity and personality test tools developed in this study do not separate creativity and personality,

but measure 'creativity and personality nature' in one place, so they have special features as tools to measure creativity and character of college students directly. The existing tools, in particular, were difficult to measure creativity and character properly due to the diverse spectrum of concepts for creativity and character by examining creativity and character separately, but as this tool was developed, they were able to solve the difficulties that could be encountered in measuring creativity and character of college students.

Looking at the preceding study [1] compared to the prior study of the composition of creative talent competences and sub-capabilities for college students, we can infer that the sub-indexes are composed of diffuse thinking, problem-solving, cognitive characteristics, curiosity, openness, collaboration and consideration, and that there are good indicators that can evaluate creative aspects compared to the evaluation criteria and evaluation metrics of creativity presented in this study.

In addition, this is the same context as the results of a preceding study [16] that developed a tool for examining creativity for 5-year-old infants whose research was conducted in an integrated way, but not for college students. Creativity factors derived from the study in [16] of the preceding study and the evaluation index results derived from this study and the areas of creativity that are consistent between the two studies can be identified. For example, the scale of 'extension of accidents' derived from this study could be found to be similar to a sub-factor called 'multi-differential thinking'. In addition, the convergence of thinking measures creativity with the ability to accept others' thoughts, the ability to solve problems autonomously, the attitude to focus on immersion, and the ability to communicate with each other in an effort to understand each other, thus confirming that the results of prior research and evaluation indicators are in line with the results of the index, and thus suitable indicators have been derived for measuring creativity and human nature. Based on the above results, creative and personality evaluation scale for college students are available for use as basic data for creative and personal education research for college students by measuring creativity and personality of college students. It is also expected to provide major implications for the academic development and quality of the creativity and personality nature areas in the future as it is a tool that can be useful in planning creative and personality education for each college student.

Based on these conclusions, suggestions for further research are as follows. The indicators developed in this study will further require efforts to develop high-quality creative and personal education programs to foster creativity and character among college students. In other words, active efforts by the university field to provide students with quality education programs that combine creativity and personality organically. In particular, each college will be able to provide more detailed creative and personal programs to students' needs, considering the strengths and weaknesses of each college, such as cognitive factors, propensity elements, motivational elements, and core competencies.

Acknowledgments

This work was supported by the Ministry of Education of the Republic of Korea and the National Research Foundation of Korea (NRF-2019S1A5C2A04082033)

References

1. Kim, E. Y., Han, Y. Y.: Analysis of Student Needs in General Education for Fostering Creative Leader Competency. Korean Journal of General Education 12(3), 115-139 (2018).
2. Youn, J. J., Kim, B. M., Kim, S. W., Kim, J. J., Kim, K. J., Eom, S. J.: Theory and Practice of Personality Education: Personality and Media. Gongdongche, Gyeonggi (2017).
3. OECD Homepage, https://www.oecd.org (The Definition and selection of key competencies: executive summary), last accessed 2020/03/19.
4. Kim, B. M., Kim, J. J., Youn, J. J.: Mediating effects of learning flow on the relationship between academic self-efficacy and creative personality in pre-service early childhood teachers. The Journal of Korea Open Association for Early Childhood Education 21(1), 403-424 (2016).
5. Lee, K. M., Choi, Y, J., Park, S, J.: Development of creativity and personality educational program for pre-service early childhood teachers and its effects. Korean Joural of Children's Media 12(2), 73-96 (2013).
6. Ministry of Education, Science and Technology Homepage, https://www.moe.go.kr/main. do, last accessed 2020/03/19.
7. Kang, S. H., Yune, S. J.: Development and validation of a scale to measure engineering creativity. Journal of Creativity and Problem Solving 11(3), 19-44 (2015).
8. Lee, S. K.: The comparison of standard setting methods : extended-angoff, rasch, and the body of work methods. Doctoral dissertation. Kyunghee University, Seoul (2017).
9. Baek, Y. S., Lee, J. H., Kim, E. T., Oh, K. J., Park, C. S., Chung, J. B.: Achievements in the Creativity Education through Freshmen Engineering Design. Journal of Engineering Education Research 9(2), 1-16 (2006).

10. Lee, G. Y.: Critical considerations on issues and methods of character assessment. Master's thesis. Seoul National University, Seoul (2017).
11. Kim, B. M., Cho, M. G., Youn, M. S.: Exploring the Development Direction of Creativity and Personality Testing Tool for College Students. Asia-pacific Journal of Multimedia Services Convergent with Art, Humanities, and Sociology 8(12), 113-124 (2018).
12. Kim, B. M.: Development of Early Childhood Education Policy Evaluation Criteria. Doctoral dissertation. Pusan National University, Pusan (2014).
13. Kim, B. M.: Development of Evaluation Scale for Early Childhood Education Policy for Parents. Journal of Educational Innovation Research 28(4), 1-26 (2018).
14. Seo, B. S.: Development of Evaluation Scale for the Inclusive Instruction in Preschool. Pusan National University, Pusan (2013).
15. Mead, D., Moseley, L.: The use of the Delphi as a research approach. Nurse Researcher 8(4), 2-23 (2001).
16. Seo, H.J., Choi, Y. J., Kim, K. C., Choi, H. J., Cho, B. K.: The development of creativity-character rating scale for 5-year-old children. Korean Joural of Children's Media 16(4), 133-156 (2017).
17. Kim, K. E., Jeong, S. J., Youn, J. J.: Student about Creative Personality. Journal of Learner-Centered Curriculum and Instruction 19(18), 433-453 (2019).

Automatic Pollen Detection Based on Feature Fusion and Self-attention Mechanism

Quanzeng Wang[1(✉)], Juan Li[1], Fujiao Ju[1], Jianqiang Li[1], Baokai Zu[1], and Caihua Ye[2]

[1] Information Department, Beijing University of Technology, Beijing 100124, China, wangquanzeng2018@gmail.com

[2] Beijing Meteorological Service Center, Beijing 100089, China

Abstract. Accurate and timely pollen prediction has great significance for pollen allergy patients. Pollen detection is one of the basic techniques for pollen prediction. In this work, we propose an automatic pollen detector based on feature fusion and self-attention mechanism, which achieves a properly balance between efficiency and precision. A simple backbone is designed to generate a relatively small model to improve the detection speed. We fuse the low-level and high-level features of smaller pollen grains to improve the detection accuracy. As pollen grains are easily broken and deformed, and usually blend with the background, thus to address these issues, we introduce the self-attention module in the proposed model to improve the detection accuracy. The experiment is operated on the pollen dataset total of 1300 images scanned by electron microscopy, and our method obtained well experimental results.

Keywords: pollen detection · feature fusion · self-attention mechanism

1 Introduction

In recent years, hay fever has become a seasonal disease with high incidence [1]. Accurate and timely pollen prediction can help patients with pollen allergies to minimize the impact of daily life. In the traditional process of pollen detection, pollen samples are processed by microscope and image processing software to obtain pollen images. Then, professionally trained technicians manually labeled the pollen grains in the image [2]. There is no doubt that this time-consuming and laborious operation cannot meet the real-time requirements of pollen prediction. Most of the existing pollen research mainly focuses on the automatic classification of pollen, however, the studies on the automatic pollen detection is still relatively few. The current pollen detection algorithms only use a few common identifying features of pollen, such as shape and texture, which makes the algorithms difficult to process pollen images with complex background information. In the field of object detection, many general object detection algorithms have been proposed and achieved good performance [3]. However, the application of these algorithms has been limited in pollen detection because of the particularities of the pollen images. As shown in Fig. 1, the pollen grains in

the image generally have small volumes and confuses with the background easily. Besides, the pollen grains are easily broken and deformed as shown in Fig. 2. Therefore, the general object detection algorithms are limited to apply on the pollen datasets. Moreover, in order to process a large number of pollen images generated in natural environment, pollen detection algorithm should have a faster processing speed. To improve the detection accuracy, the general object detection algorithm often uses complex network structure, which usually leads to the large network model and low detection speed. In this paper, to balance the accuracy and running speed, we propose an automatic pollen detector based on feature fusion and self-attention mechanism. In the proposed model, a simple backbone is used to ensure the detection efficiency. The fusion of low-level feature and high-level feature is realized by dense connection. Feature fusion can improve the detection accuracy of small pollens. The self-attention module integrates the global context information to enhance the feature expression and improve the detection accuracy.

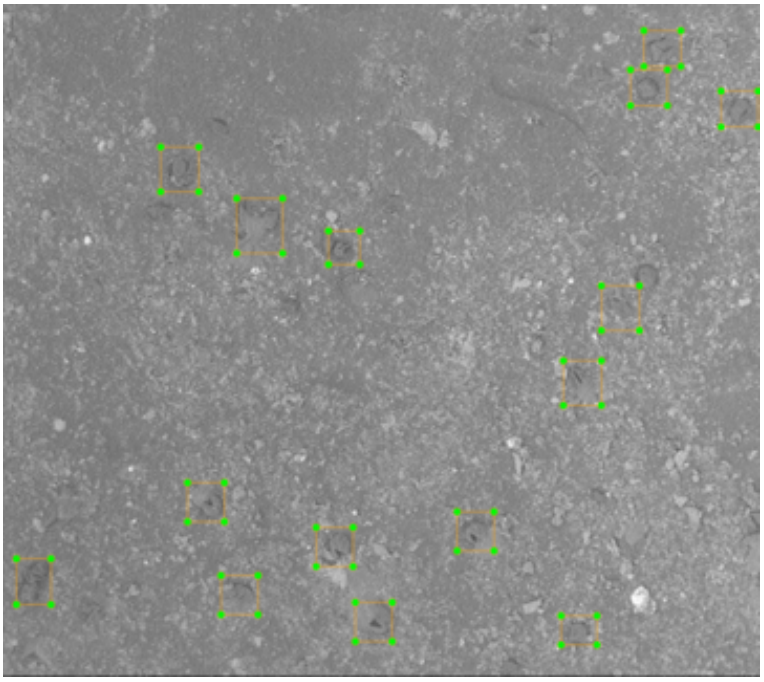


Fig. 1. Pollen image scanned by electron microscopy and the pollen grains are labeled with a green border

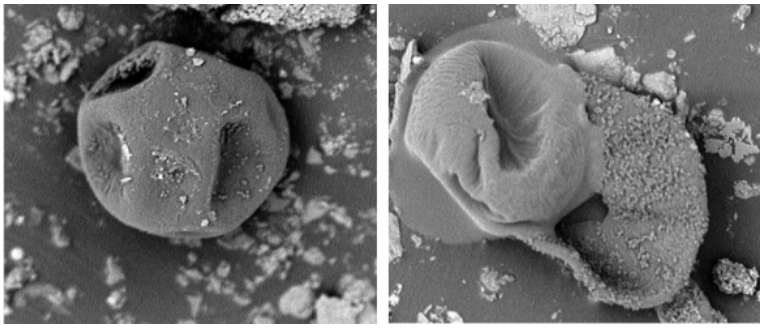


Fig. 2. The image on the left shows an intact pollen grain, and the other one shows a fractured and deformed pollen grain

2 Related work

The pollen detection task is to locate the pollens in an image. Some pollen detection algorithms detect pollen as a circle. As a well-known method to detect circles or circular arcs, Hough transform was used to detect pollen in [4] and [5]. However, the background information of pollen images in real environment is complicated. Many round pebbles and sand grains are mistaken for pollen grains.

In recent years, some seminal works have been proposed in the field of object detection[6], such as RCNN [7] series detectors and SSD [8]. These general-purpose object detectors perform well on many datasets. However, for the small object detection, the performance of these algorithms is unsatisfactory[9]. In addition, SSD and RCNN-based algorithms usually use complex network structures to ensure high-precision detection, which result in large models and low model efficiency. For the sake of running efficiency, the author in [10] proposed a Light and Fast Face Detector (LFFD) inspired by the one-stage and multi-scale object detector SSD. Instead of pre-defined anchor boxes, LFFD is manually designed for each detection branch. In LFFD, receptive fields of neurons in feature maps are regarded as natural anchors. A receptive field is matched to a groundtruth bbox if and only if its center falls in the groundtruth bbox [10]. The backbone of LFFD only consists of common layers (conv3 $\times$ 3, conv1 $\times$ 1, ReLU and residual connection), which is much lighter than VGG16 [11], Resnet50 [12] and Densenet121 [13]. LFFD achieves a considerable balance between efficiency and accuracy. However, the method is limited to using the feature information of the local receptive field to detect the target. Global context information is rarely used during object detection[14], although context information is very important to improve the accuracy of pollen detection.

Considering all the above problems comprehensively, we have added feature fusion and self-attention mechanism in the proposed method while absorbing the advantages of LFFD, to make the pollen detector perform better.

3 The Proposed Method

3.1 Network Architecture

With reference to LFFD, we designed a plain backbone for pollen detection. The overall network architecture is illustrated in Fig. 3. The backbone consists of 15 convolution layers in series and is divided into two parts. The kernel size of each convolution layer of backbone is 3×3. The difference of the two parts is that the number of filters for all convolutional layers are 64 and 128, respectively. As the network deepens, the size of receptive field gradually increases. Pollen grains with different sizes are detected by detection branches connected by the corresponding convolution layers. In the small part, the smaller pollen grains are detected because of the smaller receptive field size. Larger pollen grains are detected in the large part. Similar to DenseNet [13], each part is a dense block and has two detection branches. Since the feature maps generated by the convolution layers of each part have the same size, thus they can be directly concatenated. The first branch of the small part is connected with c8, whose receptive field size is 55 for pollen scale 10-15. The second branch stems from c10 with receptive field size 71 for pollen scale 15-20. In the large part, the first branch come from c13 with receptive field size 111 for pollen scale 20-40, and the second branch is from c15 with receptive field size 143 for pollen scale 40-70. The pollen scale is measured by its longer side. In the experimental datasets, the pollen scale ranges from 10 pixel to 70 pixels. All detection branches have a similar structure which contains an attention module and a detection module. In each detection module, there are two sub-branches, one for pollen classification and the other one for bbox regression. The pollen classification branch contains two 1×1 convolution layers, and the output is the probability that the current receptive field is a pollen grain. The probability exceeding 0.5 indicates that it is a pollen. At this point, the bbox regression branch is activated to output the top-left corner and bottom-right corner coordinates of the pollen prediction bbox.

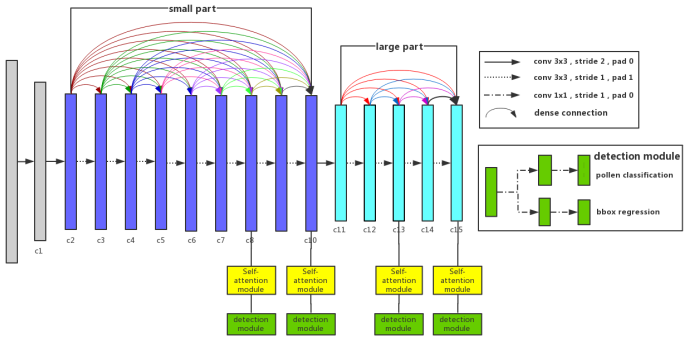


Fig. 3. The overall architecture of the network

3.2 Feature fusion

Pollen images have their own characteristics as contrast to normal life scenes. In pollen images, most pollen grains are small in size. Many general-purpose object detectors are not satisfactory for small objects because of utilizing of pooling layer. The small object loses most of its feature information in the deep layers. It is generally believed that the low-level feature map saves more location information of small object and the high-level feature map contains more semantic information [15]. So, a variety of feature fusion algorithms as effective means are proposed to improve the detection performance for small objects. Considering the efficiency of the model, complex methods are not used in our algorithm. We tend to use dense connections to merge low-level feature map and high-level feature map directly. Finally, the fused features are used to detect pollen grains.

3.3 Self-Attention module

Our algorithm uses receptive fields as natural anchors. The detection accuracy is limited by the information of local receptive fields. Global context information is difficult to capture to improve detection quality. However, global information is distinctly important in pollen detection. During the pollen collection process, all pollen grains in the same image are usually affected by similar environmental factors, such as temperature, wind and sand. Therefore, these pollen grains always show similar features, such as similar degrees of rupture, and the conditions covered by sand and dust. These features may vary greatly between different images. So, the feature information of different parts in the same image has great reference value. Drawing on the ideas of [16], in our algorithm, we added the self-attention module to obtain context information. The goal of self-attention module is to get the correlation between all pixel spaces. As shown in Fig. 4, we use 1×1 convolution to reduce the channels of input feature map. Reducing the number of channels makes the amount of calculation decrease and the efficiency improve. The input feature map is a tensor of size C×H×W, where C is the number of channels, H and W are the height and the width respectively. The number of channels of the processed feature map is changed to C', and C'=C/8.

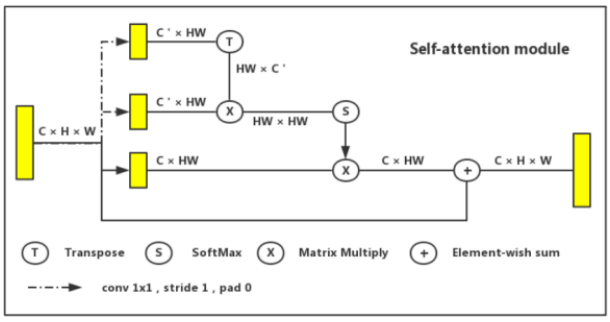


Fig. 4. Self-attention module

The processed feature map is transposed and then multiplied by another feature map, thus a matrix is obtained. The SoftMax operation is performed on each row of the matrix. Furthermore, we obtain the self-attention matrix of size HW×HW, where HW represents the product of H and W. By multiplying the input feature map and the self-attention matrix, we can get the feature map which integrates global context information. The element-wish sum of the obtained new feature map and the input feature map is the final output feature map. This feature map will be used as the input of the detection module to predict the class information and specific location of pollen. The calculation method is shown in formula (1)

$$s = f + \lambda Softmax\left((W_1 f)^T (W_2 f)\right) f \tag{1}$$

In the formula, f represents input feature map. W_1 and W_2 is 1×1 convolution. s is the output feature map. λ represents the proportion of features that incorporate spatial information in the output feature map. Initializing λ to zero, the network pays more attention to relatively simple neighborhood information, and then gradually assigns weights to other long-distance features. Therefore, the output feature map selectively integrates the global context information and improves the network expression ability.

4 Experiments

4.1 Datasets

There are 1300 pollen images in the datasets. All images are taken by a scanning electron microscope. The numbers of pollens in all images are evenly distributed between 0 and 42. The images are uniformly cropped to 512×512 pixels, which can improve the efficiency of model training. The numbers of pollen images are 900, 200, and 200 in the training set, validation set, and testing set, respectively.

4.2 Training Details

Tests are implemented by MXNet [17] on a server with Nvidia GeForce RTX 2070 GPU and 8G memory. We use Xavier method to initialize all the parameters. The optimization method is SGD with 0.9 momentum. The initial learning rate is 0.1. For pollen classification, we use cross-entropy as loss function. For bbox regression, we adopt L2 loss directly. Refer to [10], the regression groundtruth is defined as formula (2):

$$\frac{RF_x - b_x^{tl}}{RF_s/2}, \frac{RF_y - b_y^{tl}}{RF_s/2}, \frac{RF_x - b_x^{br}}{RF_s/2}, \frac{RF_y - b_y^{br}}{RF_s/2} \tag{2}$$

where RF_x and RF_y are center coordinates of the receptive field, b_x^{tl} and b_y^{tl} are coordinates of top-left corner of the bbox, b_x^{br} and b_y^{br} are coordinates of bottom-right corner of the bbox. RF_s is the size of the receptive field, which plays a role of normalization.

4.3 Result

The accuracy of pollen detection in the experiments is calculated as formula (3). In this formula, TP represents true positive example, TN represents true negative example, FP represents false positive example, and FN represents false negative example.

$$Accuracy = \frac{TP + TN}{TP + FP + TN + FN} \tag{3}$$

We used Fast RCNN, SSD and LFFD as the comparison experiment in this paper. Evaluation results on the validation set and test set are shown in Table 1. Our method performs best on both validation and testing set, and consistently outperforms LFFD. Although the gap with SSD is very close, it is worth noting that our model size is 7M and SSD is 285M. Obviously, our algorithm achieves a better balance between efficiency and accuracy.

Table 1. Evaluation results on datasets

Method	validation set	testing set
Faster RCNN	0.655	0.632
SSD	0.705	0.691
LFFD	0.643	0.625
Ours	**0.722**	**0.709**

5 Conclusion

Relying on a simple backbone, we have proposed a fast pollen detector for dealing with mass produced pollen images. We have combined the low-level features and high-level features to improve the detection accuracy of small pollen grains. Considering the characteristics of pollen data, self-attention module was added to improve the detection quality by fusing global context information. Our method achieves a properly balance between efficiency and precision. It lays a good foundation for accurate and timely prediction of pollen concentration. However, in the experiments, we found that the highly deformed pollen is difficult to detect, which will be our next research direction.

Acknowledgements. This study is supported by Beijing Municipal Science and Technology Project with no. Z191100009119013.

References

1. Marcos, J. Víctor, et al. "Automated pollen identification using microscopic imaging and texture analysis." Micron 68 (2015): 36-46.

2. Daood, Amar, Eraldo Ribeiro, and Mark Bush. "Pollen grain recognition using deep learning." International Symposium on Visual Computing. Springer, Cham, 2016.
3. Imran A, Li J, Pei Y, et al. Comparative Analysis of Vessel Segmentation Techniques in Retinal Images. IEEE Access , 2019, 7: 114862-114887
4. Rodriguez-Damian M , Cernadas E , Formella A , et al. Automatic detection and classification of grains of pollen based on shape and texture[J]. IEEE Transactions on Systems Man and Cybernetics Part C (Applications and Reviews), 2006, 36(4)(4):531-542.
5. Chica M . Authentication of bee pollen grains in bright-field microscopy by combining one-class classification techniques and image processing[J]. Microsc Res Tech, 2012, aop(aop).
6. Jianqiang Li, Fei Wang, Semi-Supervised Learning via Mean Field Methods, Neurocomputing , Vol. 177, 385-393, 2016
7. Girshick, Ross, et al. "Rich feature hierarchies for accurate object detection and semantic segmentation." Proceedings of the IEEE conference on computer vision and pattern recognition. 2014.
8. Liu, Wei, et al. "Ssd: Single shot multibox detector." European conference on computer vision. Springer, Cham, 2016.
9. Xu X, Zhang L, Li J, et al. A hybrid global-local representation CNN model for automatic cataract grading. IEEE Journal of Biomedical and Health Informatics, 2019
10. He, Yonghao, et al. "LFFD: A Light and Fast Face Detector for Edge Devices." arXiv preprint arXiv:1904.10633 (2019).
11. Simonyan, Karen, and Andrew Zisserman. "Very deep convolutional networks for large-scale image recognition." arXiv preprint arXiv:1409.1556 (2014).
12. He, Kaiming, et al. "Deep residual learning for image recognition." Proceedings of the IEEE conference on computer vision and pattern recognition. 2016.
13. Huang, Gao, et al. "Densely connected convolutional networks." Proceedings of the IEEE conference on computer vision and pattern recognition. 2017.
14. Li J, Sun J, Liu L, et al. Improved Maximum Margin Clustering via the Bundle Method. IEEE Access 2019, 7: 63709-63721
15. Yang, Xue, et al. "Scrdet: Towards more robust detection for small, cluttered and rotated objects." Proceedings of the IEEE International Conference on Computer Vision. 2019.
16. Zhang, Han, et al. "Self-attention generative adversarial networks." arXiv preprint arXiv:1805.08318 (2018).
17. Chen, Tianqi, et al. "Mxnet: A flexible and efficient machine learning library for heterogeneous distributed systems." arXiv preprint arXiv:1512.01274 (2015).

A Malicious Web Request Detection Technology Based on Gate Recurrent Unit

Zhibin Liu[1] §, Wenqiang Zhang[2] §, Yuanyuan Huang[2], and Qingguo Zhou[2] (✉)

[1] North China Brach of State Grid Corporation of China, Beijing, China
liu.zbin@nc.sgcc.com.cn,
[2] School of Information Science and Engineering, Lanzhou University, Lanzhou, China
{zhangwq18, huangyy19, zhouqg}@lzu.edu.cn

Abstract. With the development of information technology and the rapid growth of Internet of things terminals, the number of web requests is increasing rapidly, and the proportion of malicious requests is also increasing. In order to detect and filter malicious HTTP requests quickly and correctly, this paper proposes a machine learning model to Gate Recurrent Unit (GRU) as the core. The URI extracted from the HTTP requests was encoded into sequence of codes using the indices of the vocabulary list, and then the distributed vector representations are produced by embedding and then input to the GRU and the full-connection neural network. In this paper, the characteristics of datasets are analyzed and the hyper-parameters are determined. Then two different datasets are trained independently. The experiments result shows that the proposed method is effective. This technology can be used in WAF application and malicious web traffic monitoring.

Keywords: Web Attack, Gate Recurrent Unit, Anomaly Detection

1 Introduction

In recent years, a variety of new technologies promote the proliferation of network traffic and applications. People enjoy the convenience form web applications, but also take many risks. With the continuous evolution of network attack technology, attack methods become more various, which not only leads to serious security risks for users and service providers but also makes network security defense more and more complex and difficult. According to the 2019 Internet security thread report [1] released by Symantec in 2019, web attacks on endpoints increased by 56% in 2018. And Symantec was blocking more than 1.3 million web attacks on endpoint machines every day. Besides, 10% of the analyzed URLs were identified as malicious URLs, compared with only one-sixteenth in 2017. The purpose of the network attack is often to collect private or confidential information and obtain control of the target application or server. Many websites

§Z. Liu and W. Zhang contributed equally.

or network applications have security vulnerabilities, so these websites and applications will be vulnerable to attack and intrusion, once penetrated, they will face significant losses.

Due to the complexity of network applications, network engineers often only pay attention to the design and implementation of applications and thus ignore the security issues of web applications. At the same time, attackers always try to attack all kinds of web applications and services all the time, and they constantly improve attack technology and means. Therefore, the network administrator needs to configure a set of intrusions detection systems without affecting the normal user experience. The intrusion detection system can identify the intruder's intrusion and penetration, detect the user's illegal behavior and abnormal operation accurately and quickly, and alert the administrator. The current detection methods are mainly divided into two categories, signature-based detection, and behavior-based detection. Generally, the signature-based detection method is to extract and model the known attack features. In the use of this detection method, it needs to compare the current user behavior with the defined attack mode. Therefore, it can only detect known attacks and is helpless for unknown attacks. Zhou et al. [19] presented that different types of users can be distinguished by user behavior, so malicious requests can be detected by the characteristics of user behavior. This detection method needs to calculate the deviation between the current user or system behavior and the defined normal behavior in use. The larger the deviation means the greater the possibility that the current behavior is abnormal behavior. Behavior-based detection can detect attacks that have never been encountered before, but it also has the disadvantage of a high false alarm rate.

The method proposed in this paper is signature-based detection, trying to improve the accuracy and detection speed as much as possible. It can be used as a sub-module in a system with a behavior-based detection method to work together. In the process of running web applications, HTTP requests will be produced. URI part of web requests is extracted for encoding and then generate distribution vector representation through embedding. The encoded URI sequence will output prediction results through the GRU network and full connection neural network. It can use the data in the HTTP request log of the server to train the model, then detect the malicious requests hidden in the normal network requests in real-time when the server is working.

In fact, most of the words in the URI are the names of variables and variables, which are rarely repeated. The content of Sect. 4.2 of this article can be seen. Therefore, for embedding, the size of the dictionary needs to maintain is actually very small, and the weight of the corresponding project layer will be much less, and the computational cost of using the model as a sub-model will be acceptable. The size of the hidden output of the GRU network is much smaller than its output. There are only two types of classification. In comparison, the information loss caused by the hidden output becomes acceptable. However, using hidden output can reduce the weight size of the prediction layer. The architecture of

the model is introduced in Sect. 3.3 in detail, this paper will not do detailed derivation, and the reference cited has a detailed implementation process.

The technology proposed in this paper has the following characteristics and advantages:

1. It parses URIs and extracts the features of URIs in the time and frequency domain to judge the possibility of malicious web requests.
2. The proposed model is simple in structure and small in scale, so it requires little computational cost.
3. It is based on machine learning, so it can expand the training set to improve accuracy and deal with new types of attacks without manually updating rules.
4. It can be used in malicious traffic real-time warning systems and WAF applications and can work together with a behavior-based model.

2 Related work

For the behavior-based detection method, Zolotukhin's team proposes to analyze the normal requests of HTTP logs by data mining [20], use anomaly detection algorithm and clustering algorithm to model the normal user behavior, and then judge the malicious requests by the deviation value of users' behavior. Tian's team proposes a real-time network attack detection method based on stream data analysis [18]. The work of their paper creates a hash method framework for caching log data and proposes a multi-pattern matching algorithm based on stream data analysis. The system proposed in Csubák's paper detects malicious traffic through the characteristics of big data statistical normal traffic and malicious traffic [5]. For signature-based detection methods, a decision tree method is proposed in García's work [7], but nowadays, the accuracy of this method is not high enough. Generally speaking, due to the more complex request mode of the legal script, the risk of malicious requests can be estimated by Shannon entropy [17]. HMM can obtain the characteristics of the web request URI sequence in the time domain. Three different processes are used to detect malicious requests using HMM model [4, 13, 6]. Goyal and Bansal propose to use a multi-layer neural network to detect malicious URLs [9]. However, their work mainly focuses on phishing URLs.

LSTM (long short term memory) network is an improvement of RNN [10]. LSTM has the ability of long-term memory. GRU is a variant of LSTM, which has a similar effect with LSTM by experiment. However, GRU has a simpler structure and requires less calculation than LSTM. Cho et al. [3] are the first proposed GRU, with the structure and the algorithm of GRU are implemented. The Encoder-Decoder model is used to process the word sequence in this paper. Distributed vector representations are the output of embedding. Mikolov and his partners' research shows that a continuous skip-gram model has a good effect on learning high-quality distributed vector representations(see [12] and [11]), while distributed vector representations can capture a large number of accurate syntactic and semantic relationships.

3 Proposed work

3.1 Process data

The scheme proposed in this paper is used to detect malicious web requests, but before training or prediction, redundant information needs to be removed to reduce interference information and computing resource. The length and number of words contained in the URI in the request will also be very different in different web applications, resulting in a large difference in the length between the URIs. To keep important information and simplify the calculation, we only keep the words with high frequency of occurrence. The URI is then transformed into a sequence of codes by querying the words list. The specific procedure is divided into four steps, which are detailed as below.

Step 1 (preprocess data). For the POST and PUT requests, the URI and form data are merged to make up a new URI. To remove redundant information, the host address or domain name and port of the URL part are removed. Each item in the dataset needs to be marked as malicious requests or regular requests, where malicious requests are positive samples and regular requests are negative samples. We take 70% of positive samples and negative samples as the training set and 30% as the test set.

Step 2 (split item into words). For each item in the dataset, it is split into several words by symbols as a division point. In order to partition correctly, the hexadecimal data starting with the percent sign needs to be separated into rules. In addition, the symbol, number, and hexadecimal data are also calculated as words.

Step 3 (limit words number). In the previous step, theoretically, a tremendous amount of unique words will be generated. Too many words will make vocabulary list querying difficult, and there will be words outside the vocabulary list in the test set. Therefore, it is necessary to limit the number of words but able to contain all split words. The frequency of the words in the training set are counted, and then sort them according to the frequency from large to small. The former part supposed to add it to the vocabulary list directly, and then classify the latter part into several words and add them to the vocabulary list. The vocabulary list is then reordered by frequency.

Step 4 (encode item). According to step 2, the items are split into words sequence, then the corresponding codes are found by querying the vocabulary list, and the sequence of the words is converted into a numerical sequence of codes. The size of the sequence needs to uniform. If the size is insufficient, PAD (padding) is used to fill it. PAD is also used as a word in the vocabulary list. If it exceeds the limited size, the redundant part will be discarded. In theory, the

Table 1. Example of data processing

Step	Content
1	[/index.jsp?id=927&n=1+AND+%22;alert(\Stay!\)]
2	[/ index . jsp ? id = 927 & n = 1 + AND + %22 ; alert (\ Stay ! \ )]
3	[/ index . jsp ? id = -d & n = 1 + AND + -h -s alert -s \ -t ! \ -s]
4	[5 2 7 10 11 4 1 12 2 15 1 16 3 14 3 13 9 8 9 5 6 17 5 9]

longer the sequence length is, the more complete information items are retained in the dataset, and the better the theoretical effect will be. However, if the size of the sequence is too large, the operation time will be prolonged, so a compromise is needed here.

3.2 Example of process data

Here is a simple example is shown in Table 1. The left column is the step number, and the right column is the corresponding result of each step. The words out of the vocabulary list was split into 7 classes: -s (symbol), -d (digit), -u (upper case), -l (lower case), -t (title case), -h (hexadecimal data), -o (other).

3.3 Model architectures

The main body of the model is GRU, which can keep a long memory of the whole sequence and take into account the relevant contents of the whole sequence. The hidden output of GRU input to the fully connected neural network, and then predict the possibility of the sequence as a malicious request and the judgment results, the model architecture is shown in Fig. 1.

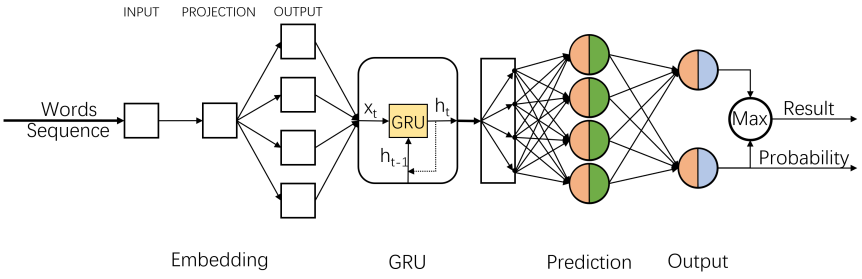


Fig. 1. The model architectures.

Embedding layer. Its function is embedding the words through the encoded-word sequence, and output the distribution vector representations. As this layer of training, distribution vector representations of the model can capture a large number of accurate grammar rules and semantic relations [12], so the distribution vector representation is the features of the words. The embedded internal dimension is a lookup table that stores embeddings of a fixed dictionary and size, and the number of features in each word is determined by the training effect and training speed. The size of the dictionary in the embedding layer can be determined by the frequency distribution of words. Rong [14] explains in detail the process of using the continuous skip-gram model to implement the transformation of the word to vectors (word2vec), including detailed algorithm derivation and principle.

GRU layer. The main structure of GRU is a reset gate and an update gate. Reset gate determines whether to update the memory weight of GRU. Update gate control the output that how many past memories should be included. Cho et al. [3] also completed the GRU implementation and algorithm derivation. GRU is responsible for extracting the time-domain features of entries. The whole sequence vector output from the first layer directly enters the GRU layer. After iterating the sequence length times, all outputs are ignored and all hidden outputs are taken as the input of the next layer. The number of features in the hidden state and the number of features in each word should have the same value.

Prediction layer. This layer is a general-purpose neural network layer, the number of neurons is the same as the size of the dictionary in the nesting. Using the ReLU(Rectified Linear Unit) function as the activation function [8], the calculation of the ReLU function is simple and the output is stable, which can allocate the gradient to the weight well.

Output layer. The previous data stream outputs two data through the processing of two neurons, which correspond to normal requests and malicious request respectively. As shown in the right part of Fig.1, after softmax function, the probability of two request types can be obtained, and the most probability is usually selected as the final result.

Backward. The goal of the model is to solve a classification problem, so the Cross-Entropy loss function is used in this model. The essence of this model is the combination of a series of linear functions, sigmoid functions, tanh functions, softmax functions and so on, we can calculate the gradient that any weight should be updated by deriving the above functions [16]. Reference [15] summarizes many gradient descent methods, including some optimization methods and applications. In this paper, we use the Mini-Batch Gradient Descent (MBGD) method to train the model and update the model weight.

Table 2. Dataset information

Dataset	Normal	Malicious	Train	Test
CSIC2010	18620	15964	24208	10376
XSS_DS	12000	11744	16620	7124

4 Experiments

4.1 Dataset introduction

Two datasets are used in the experiment, they are CSIC2010 and XSS_DS. CSIC2010 was made by the Information Security Research Institute of the Spanish National Research Council (CSIC) [2], which contains 25000 abnormal requests, including multiple attack types such as SQL injection, CRLF injection, XSS, buffer overflow, etc. it is mainly used to test the network attack protection system. In the experiment, the URI is extracted for each item, and then the duplicate items are removed. XSS_DS is a URI dataset that is consists of XSS requests and normal requests from the server collected by the author team. The scale of the processed dataset is shown in Table 2.

4.2 Parameters analysis

Taking CSIC2010 dataset as a test object, here we determine the size of the dictionary in the nesting layer, count the frequency of all words in the whole training set after segmentation, sort the word frequency by frequency, and then make the frequency distribution diagram as shown in Fig. 2. According to the statistics result, we can see that the maximum single word frequency in the training set can reach 176869, and the number of words can reach 37829. After zoom in the main position, shown as Fig. 2, it can be seen that the curve gradient changes the most in the position of the abscissa 250, and the frequency of the first 450 words accounts for 87.03% of the total frequency, which is a high proportion, but a small proportion in the number of words. Therefore, in CSIC2010 training, the dictionary size of the nesting layer is set to about 450.

We count the length and frequency of the encoded sequence in the training set. The frequency distribution is shown in the blue area of the left of 3. when there are enough samples, the distribution should meet the Gaussian distribution. According to the frequency distribution, the parameters of the Gaussian distribution can be calculated, as shown in the red curve in the left of Fig. 3. According to the proportion of red area and blue area in the right of Fig 3, when the maximum length is 130, about 99.5% of items can retain complete information.

However, other super parameters need to be determined through evaluation performance and even experience in the training process. See Table 3 for some super parameters used in the experiment.

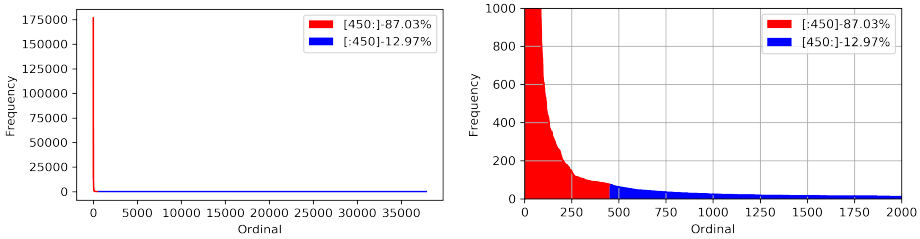


Fig. 2. The frequency distribution of words in the CSIC2010 dataset, the left one shows all words and frequency, while the right one shows a zoomed-in view of the left one.

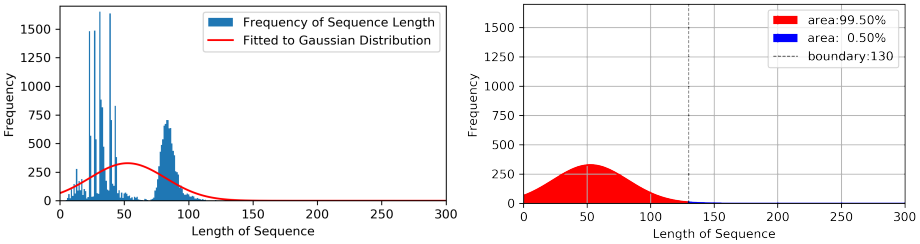


Fig. 3. The frequency distribution of encoded sequence length, the left graph shows real distribution with a red curve fitted by Gaussian distribution, the right shows a proportion between more than 130 words and no more than 130 words of sequences

4.3 Experiment result

To achieve the best effect, the learning rate of training is gradually reduced. The two datasets are trained by 30 epochs respectively, and the resulting loss curves are shown in Fig 4. Both of them basically drop to the lowest point when approaching 30 times of training. Because of the gradual decline of the learning rate, the first few epochs of the curve decrease very fast, and the later ones slow, which is to find the lowest loss to prevent overfitting.

Fawcett explained several evaluation criteria and calculation methods in their work [6]. This paper will use Accuracy, Precision, and Recall to evaluate the model. Accuracy indicates the proportion of correctly predicted samples to the total number of samples. Precision represents the proportion of the correct number of items predicted as positive samples. Recall indicates that the number of positive samples correctly predicted accounts for the number of all positive samples. In this paper, malicious web requests are taken as positive samples. When the loss is reduced to the lowest level, the training can be evaluated with the test set of two datasets, and the results shown in Table 4 can be obtained. It can be seen from the Accuracy that XSS_DS has achieved excellent results. The main reason is that the XSS_DS dataset only contains XSS attacks. There are many types of malicious requests in the CSIC2010 dataset, so the training results are a little poor. It can be seen from Precision and Recall that the detection accuracy of malicious requests we are concerned about is very high. The values of Accu-

Table 3. Hyper-Parameters in the experiment

Hyper-Parameters	Value
Length of Sequence	130
The size of dictionary in embedding	450
The size of the hidden state in GRU	130
Probability of dropout after GRU	0.3
The number of neurons in prediction layer	450
Probability of dropout after prediction layer	0.2
Mini-batch size	10

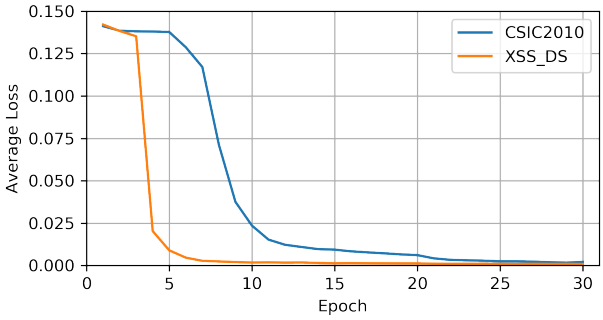


Fig. 4. Loss curves of dataset CSIC2010 and XSS_DS.

racy, Precision, and Recall are very similar, which indicates that the quality of the model is good, and there is no high false alarm rate for some class samples.

5 Conclusion

This paper proposed an innovative technology to detect malicious web requests based on machine learning. In the content of this paper, the process of transforming the items in the dataset into the trainable sequence of codes is described in detail. Then sequences of codes enter a simple machine learning model. This model consists of three parts: embedding, GRU, and fully-connected neural network. Finally, two neurons output the prediction results. Through the analysis of the two datasets, the important super parameters are determined. After using two datasets to train and evaluate the model, the results show that technology has a significant effect on the detection of malicious web requests. The result of the XSS_DS dataset shows that the model is effective for the classification between XSS type attacks and regular requests. Because the detection method is based on machine learning, we can add new training data to form detection rules for new types of attacks and improve the accuracy of detection.

Table 4. Dataset information

Dataset	Accuracy	Precision	Recall
CSIC2010	0.9704	0.9669	0.9691
XSS_DS	0.9992	0.9994	0.9991

However, according to the training results of the CSIC2010 dataset, the accuracy of the classification of malicious request models containing multiple types is not high enough, and there is still the possibility of improving the accuracy. In the future, we can identify the type of malicious request through more accurate type tags, and can also cooperate with other Web attack detection systems to improve the accuracy rate.

Acknowledgment

This work was partially supported by National Natural Science Foundation of China under Grant No. 61402210, The Fundamental Research Funds for the Central Universities under Grant No. lzujbky-2019-kb51 and lzujbky-2018-k12, Ministry of Education - China Mobile Research Foundation under Grant No. MCM20170206, Major National Project of High Resolution Earth Observation System under Grant No. 30-Y20A34-9010-15/17, State Grid Corporation Science and Technology Project under Grant No. SGGSKY00FJJS1800403 and No.522722160071, Program for New Century Excellent Talents in University under Grant No. NCET-12-0250, and Double first class Funding-International Cooperation and Exchange Program under Grant No. 227000-560001, and Strategic Priority Research Program of the Chinese Academy of Sciences with Grant No. XDA03030100. Google Research Awards and Google Faculty Award.

References

1. ISTR Internet Security Threat Report. https://cdn2.hubspot.net/hubfs/5156294/SED/SED%20SYMC/6819_SED_SYMC_ISTR_24_2019_April_en.pdf (2019)
2. Carmen Torrano, G., Alejandro Pérez, V., Gonzalo Álvarez, M.: HTTP DATASET CSIC 2010. https://www.isi.csic.es/dataset/
3. Cho, K., van Merriënboer, B., Gulcehre, C., Bougares, F., Schwenk, H., Bengio, Y.: Learning phrase representations using rnn encoder-decoder for statistical machine translation. In: Conference on Empirical Methods in Natural Language, pp. 1724–1734 (2014). DOI 10.3115/v1/D14-1179
4. Corona, I., Ariu, D., Giacinto, G.: Hmm-web: A framework for the detection of attacks against web applications. In: 2009 IEEE International Conference on Communications, pp. 1–6 (2009). DOI 10.1109/ICC.2009.5199054
5. Csubák, D., Szuecs, K., Vörös, P., Kiss, A.: Big data testbed for network attack detection. Acta Polytechnica Hungarica **13**, 47–57 (2016)

6. Fawcett, T.: Roc graphs: Notes and practical considerations for researchers. Machine Learning **31**, 1–38 (2004)
7. García, V.H., Monroy, R., Quintana, M.: Web attack detection using id3. In: J. Debenham (ed.) Professional Practice in Artificial Intelligence, pp. 323–332. Springer US, Boston, MA (2006)
8. Glorot, X., Bordes, A., Bengio, Y.: Deep sparse rectifier neural networks. Proceedings of the 14th International Conference on Artificial Intelligence and Statisitics (AISTATS) 2011 **15**, 315–323 (2011)
9. Goyal, B., Bansal, M.: A competent approach for type of phishing attack detection using multi-layer neural network. International Journal of Advanced Engineering Research and Science **4**, 210–215 (2017). DOI 10.22161/ijaers.4.1.34
10. Hochreiter, S., Schmidhuber, J.: Long short-term memory. Neural computation **9**, 1735–80 (1997). DOI 10.1162/neco.1997.9.8.1735
11. Mikolov, T., Chen, K., Corrado, G., Dean, J.: Efficient estimation of word representations in vector space. Proceedings of Workshop at ICLR **2013** (2013)
12. Mikolov, T., Sutskever, I., Chen, K., Corrado, G., Dean, J.: Distributed representations of words and phrases and their compositionality. Advances in Neural Information Processing Systems **26** (2013)
13. Ren, X., Hu, Y., Kuang, W., Souleymanou, M.B.: A web attack detection technology based on bag of words and hidden markov model. In: 2018 IEEE 15th International Conference on Mobile Ad Hoc and Sensor Systems (MASS), pp. 526–531 (2018). DOI 10.1109/MASS.2018.00081
14. Rong, X.: word2vec parameter learning explained. Computer Science pp. 1–21 (2014)
15. Ruder, S.: An overview of gradient descent optimization algorithms. https://ruder.io/optimizing-gradient-descent/index.html (2016)
16. Rumelhart, D., Hinton, G., Williams, R.: Learning representations by back propagating errors. Nature **323**, 533–536 (1986). DOI 10.1038/323533a0
17. Threepak, T., Watcharapupong, A.: Web attack detection using entropy-based analysis. In: The International Conference on Information Networking 2014 (ICOIN2014), pp. 244–247 (2014). DOI 10.1109/ICOIN.2014.6799699
18. Tian, J., Zhu, H., Li, X., Tian, Z.: Real-time online detection method for web attack based on flow data analysis. In: 2018 IEEE 9th International Conference on Software Engineering and Service Science (ICSESS), pp. 991–994 (2018). DOI 10.1109/ICSESS.2018.8663848
19. Zhou, X., Wu, B., Jin, Q.: User role identification based on social behavior and networking analysis for information dissemination. Future Generation Computer Systems **96** (2017). DOI 10.1016/j.future.2017.04.043
20. Zolotukhin, M., Hämäläinen, T., Kokkonen, T., Siltanen, J.: Analysis of http requests for anomaly detection of web attacks. In: 2014 IEEE 12th International Conference on Dependable, Autonomic and Secure Computing, pp. 406–411 (2014). DOI 10.1109/DASC.2014.79

Using the XGBoost model to predict Santander Customer Trading

Wen-Chih Chang, Yi-Hong Guo, Yang,Ya-Ling, Hsu, Ming-Chien, Chu, Yi-Hsuan, Chu,Ting-Yi, Meng,Long-Cheng

[1]Department of Information Management, Chung Hua University, Taiwan
Chung Hua University, Hsinchu, Taiwan
earnest@g.chu.edu.tw

Abstract. We participated the customer trade prediction, we could upload three result one day. The competition is 2019/2/14~2019/4/3. The end date is 2019/4/10. Our object is to predict each customer if he/she will trade or not. In this paper, we learn how to process the data, analysis and prediction. We also try different machine learning model to realize the efficiency. Let Santander company understand the customer's trading habits and transaction volume, as well as let the company understand the operating status, and establish the transaction value of potential customers. First of all, we utilized visualization to observer the dataset and to understand the imbalance of the target attributes. We separated the data into 9 parts and training individually. The trained model can be integrated, and voted. However, the voted result is not good enough. We used the XGBoost model. This model has higher prediction and training speed is faster. It decreased the training time.

Keywords: Predict, XGBoost model, machine learning, the customer prediction

1 Introduction

Big data, also known as huge data, refers to a large amount of data sets that traditional data processing application software cannot directly process. With the rapid development of the Internet, the amount of data on the Internet is growing rapidly today, and various types of data are waiting for us. Analytical research, and big data analysis can effectively integrate and use these huge data sets.

With the rapid development of Internet, many transactions in many enterprises can be completed on the Internet, and transactions on the Internet will leave records in the database. And we hope to help companies analyze and predict this huge data set. We used Santander Bank customer prediction competition data set. If the data set can be accurately predicted, the prediction result will become a very valuable data. After effectively using these data, companies can better understand their customers' trading habits and transaction volume, and let companies understand the operating experience.

In order to learn how to analyze and predict data, and to develop a keen sense of data observation, we decided to participate in the Kaggle [3,4] competition. There are

two major goals in Kaggle competition [3,4]: predictive analysis and data classification. Santander Bank assists clients in understanding their financial situation. Are they satisfied? Will customers buy this product? Can customers pay this loan? Regardless of the transaction amount, will the customer conduct the transaction?

The data type of the data set is less complicated, and the variable distribution of the test set and training set is more balanced. The goal of this forecast is to analyze whether the customer will trade, which is suitable for us to learn how to perform data processing and analysis, and use all Learn, compete with participants from all over the world to improve the ranking, and improve our ability to process data.

This topic mainly predicts Santander's customer transactions, helping customers understand financial status, customer satisfaction, customer needs, and whether customers will conduct transactions. The data provided for this competition can be used to solve related problems. The purpose of this research as follows:

1. Learn how to conduct in-depth analysis and forecasting of data.
2. Develop a keen sense of data observation.
3. Explore the analytical power of different machine learning models.
4. Assist companies to establish transaction value for potential customers.

The research steps are mainly divided into four stages. The first stage mainly uses different methods for data pre-processing. The second stage uses exploration and visual design to display and observe important attributes of the data in a visual way such as drawing charts to help. We clearly understand the characteristics and correlation between attributes. In the third stage, we build models and use different models for training. In order to train a better model, the fourth stage evaluates the model. After the model is established, we use charts and chaotic matrices and classification reports. Evaluate the trained model.

Each competition in the Kaggle competition is operated independently, so different competitions will have different qualifications. To participate in the competitions in the Kaggle competition, you must first register a Kaggle account, and each member of the team can use their own Account to participate in the same competition, but usually the competition will stipulate the maximum number of uploads of forecast results per day, and also limit the maximum number of uploads of predicted results for the entire competition. At that time, Kaggle will require the accounts of the same team to be merged. The maximum number of merged members will vary depending on the competition. The number of uploads for each account will also be added after the team merges. Pay special attention to the rules and restrictions of the game.

The results uploaded in the competition must be submitted in the Excel file generated by the forecast. No manual or manual data changes can be made in the Excel file obtained after the forecast. Otherwise, it will be considered cheating and disqualified.

After uploading the prediction results, the system will score in time, and the ranking will be based on the best results in the upload results. When the prediction accuracy is higher, the ranking is more advanced. Kaggle's competitions are continuous, so I want to get good results, Must continue to train and modify the model daily, because the predicted results will be surpassed at any time, and the contestant who can maintain good results until the end is the biggest winner.

Kaggle has data science teams from all over the world that constantly challenge machine learning algorithms and cooperate with the global data science community. We will also participate in it and refer to different analysis methods to ensure that we can more accurately solve the difficulties encountered by enterprises, Binary classification questions, such as: Are customers satisfied? Will customers buy this product? Can customers pay this loan?

In this challenge, we must determine which customers will conduct specific transactions in the future, regardless of the transaction amount. We can use the data we predict to judge and analyze, and reach the expected transaction.

Table 1. Prediction Tool and Model.

Platform	Development Software
Operation System	Microsoft Window 10 Professional Verison
Development Environment	Anaconda 3
Language	Python 3.7 Version
Data Prediction Tool	Jupyter Notebook, Spyder
Applying Model	Random Forest, XGBOOST
Data Analysis Tool	Excel2016

2 Related Works

This chapter is divided into five sections, mainly focusing on decision trees, random forests, support vector machines, neural-like networks, and the XGBoost model.

2.1 Decision Tree

Decision tree is proposed by Hunt and his colleges in 1996 [2], it is a tree branch concept to be a decision tree model. This model almost used in classification prediction, which is divided into Classification tree, regression tree, classification regression tree.

Classification Tree is used to predict the strain number of a category type. Regression Tree predicts values using regression analysis. ， Classification and Regression Tree (CART) is proposed by Leo Breiman in 1984 [5]. It combines the characteristics of classification trees and regression trees. It can predict category types and numerical data. It is characterized in that only two branches are generated for each classification for planning and analyzing data sets, and there is no restriction on the type of variables. There is greater flexibility in analysis.

2.2 Support Vector Machine (SVM)

Support Vector Machine (SVM for short), proposed by Corinna Cortes and Vapnik in 1995[1], mainly classifies binary data, and divides into two categories based on the

principle of minimum classification error rate. SVM uses the current data as training, selects several support vectors from the analyzed training data to represent all the data, removes some extreme values, and packs the selected support vectors into a model.

SVM is a supervised learning classifier. It can be used to analyze classification patterns in data. The main purpose is to perform classification and regression analysis. The following figure is an example. Given a training data set, each of the data (black points, white points) is given a classification answer 0 and 1, each data is a vector, and the length of the vector is 2), The SVM can construct a classification model to classify all the data into 2 categories. Later, if new test data is input into the model it can also be automatically classified into 0 or 1 by the model.

2.3 Artificial Neural Network (ANN)

Artificial Neural Network (ANN) [6] is a mathematical or computational model that simulates the structure and function of the human brain and is inspired by biology. Each neuron is connected to other neurons, and will be impacted by the state of other neurons, thereby determining whether it will fire itself. ANN is a black box model that can find the relationship between input data and output data, and form a pattern to make predictions on different input data.

2.4 eXtreme Gradient Boosting (XGBoost)

Extreme gradient boosting (eXtreme Gradient Boosting, referred to as XGBoost) was proposed by the promoter Chen Tianqi in 2014 [7,8,9,10]. Gradient means that the XGBoost module is developed in C ++ of the Gradient Boosting Machine. Boosting belongs to the overall learning classifier, which classifies many classifications with low accuracy. The model trees are combined to become a model with higher accuracy, which will be continuously revised, and a new tree is generated after each correction. XGBoost [7,8,9,10] proposes the concept of blocks, and sorts the data before training. It is then saved as a block structure, and the data in the structure is reused in subsequent processes to allow the modules to be parallelized and reduce training time.

3 The Analysisi Model

3.1 Data Set

This data comes from the data center in the Kaggle Santander trading prediction competition. The competition data is divided into three parts. The first part is the training set, the second part is the test set, and the last part is the prediction result data set. The training set data can be used for data pre-processing and visual design and modeling. When the model is created, the processed training set is used to predict the data in the test set, and finally the prediction result data set is generated.

Table 2. Training Data Attributes.

Attribute Name	ID_CODE	Target	Var_0 – Var_199
Attribute Type	object	Int64	Float64
Attribute Sum	200000	200000	200000
Is there a null value	None	None	None
Attribute Meaning	Code	target attribute	Attribute Value

Table 3. Test Data Attributes.

Attribute Name	ID_CODE	Var_0 –Var_199
Attribute Type	object	Float64
Attribute Sum	200000	200000
Is there a null value	None	None
Attribute Meaning	Code	Attribute Value

This data analysis has so far used two models, three modules and four kits: first, the model uses the XGBOOST model and the Random Forest model, and the modules are the Csv module, Numpy module, and Sklearn module. Finally, there are Pandas Suite, Matplotlib Suite, Seaborn Suite and Os Suite.

Models

1.XGBoos model: It is used to predict the accuracy rate. The higher the accuracy rate, the faster the training speed, and the XGBClassifier type is selected.

(1)max_depth parameter: is the maximum depth, the greater the depth of the tree, the higher the degree of overfitting to the data, which is used to prevent the fitting problem

(2)n_estimators parameter: the number of trees, which affects the training speed. The larger the module, the longer the training time.

(3)colsample_bytree parameter: The ratio of random sampling of features when the tree is established. If the ratio is larger, each tree is more similar. If the ratio is smaller, the diversity of the tree can be increased.

(4)Learning_rate parameter: For the learning rate, the larger the value, the faster the model learns the training set. If it is too small, the training is too slow. If it is too fast, the advantage is missed.

(5)Objective parameter: The binary: logistic parameter is used for the logistic regression of the two classifications, which returns the predicted probability (not the category).

(6)n_jobs parameter: how many CPUs are used for execution, -1 is to use 100% CPU.

Random ForestModel :

(1)n_estimators parameter: The number of trees, which affects the training speed. The larger the module, the longer the training time.
(2)random_state parameter: random seed. If it is set to 1, the data of the construction model is fixed. If it is set to other numerical cuts, the data of the construction model is different.
(3)max_depth parameter: is the maximum depth, the greater the depth of the tree, the higher the degree of overfitting to the data, which is used to prevent the fitting problem
(4)min_samples_leaf parameter: Limit the minimum number of samples of leaf nodes, suitable for large sample data sets
(5)Criterion parameter: The formula used to decide whether to use entropy or gini.
(6)max_features parameter: The maximum number of features to consider when dividing the data

Models

1.Csv module: It is used to read and write CSV format. CSV file format is a universal spreadsheet and database import and export format.
(1)Csv.reader (fin, delimiter = `,'): used to read the fin variable delimiter as ","
2.Numpy module: It is mainly used for data processing. It can quickly operate huge data of multi-dimensional array.
(1)x_train = np.array (x_train): Introduce numpy module, create an array by passing list to np.array (x_train).
3.Sklearn module: used for machine learning.
(1) train_test_split: used to divide the training set and test set.
(2) classification_report: Make a classification report showing the accuracy, recall rate, and F1 value of each class.
(3) confusion_matrix: Make a confusion matrix, reflecting the classification accuracy of the module from different angles, it can be easier to determine whether the module confuses two different classes.
(4) accuracy_score: Calculate the percentage of classification accuracy score.
1. Pandas package: supports multiple text and data loading, and can quickly understand the data structure
(1) read_csv: used to read csv files.
(2) Pandas.info (): Check out several data and attributes.
(3) Pandas.shape (): View the number of rows and columns in train.
(4) Pandas.head (): View the data in train.
(5)Pandas.isnull (): Check if there are null values in train.
(6) Pandas.sum (): Calculate the total number of null values in train.
(7) Pandas.describe (): View the value of each field in train.
2. Matplotlib suite: used to build charts, can be used with numpy.
(1) Matplotlib.figure (): Set the chart format.
(2) Matplotlib.subplot: print the graphs in order.
(3) Matplotlib.pie: make pie charts and bar charts Figure 4-7 Matplotlib Suite: From this figure, it can be seen that the target attribute ratio is too large.
(4) Matplotlib.scatter: Make a scatter plot.
(5) Matplotlib.xlabel: Adjust the font size.

(6) Matplotlib.tick_params: Adjust the table scale and font size.

```
def plot_feature_scatter(df1, df2, features):
    i = 0
    sns.set_style('whitegrid')
    #引入Seaborn中的set_style()功能，設定表格樣式為白色網格
    plt.figure()
    fig, ax = plt.subplots(12,5,figsize=(14,14))
    #引入matplotlib中的figure()功能，設定圖表格式，做12*5=60筆資料，寬和高設為14*14(英寸)
    for feature in features:
        i += 1
        plt.subplot(4,5,i)
        #引入matplotlib中的subplot()功能，依序印出圖表，並且用5行4列作排列
        plt.scatter(df1[feature], df2[feature], marker='+')
        #引入matplotlib中的scatter()功能，製作散佈圖，圖案形式為'+'號
        plt.xlabel(feature, fontsize=9)
        #引入matplotlib中的xlabel()功能，調整字體大小
    plt.show();
#看密集度
```

```
features = ['var_0', 'var_1','var_2','var_3', 'var_4', 'var_5', 'var_6', 'var_7', 'var_8', 'var_9', 'var_10',
            'var_11','var_12', 'var_13', 'var_14', 'var_15','var_16', 'var_17', 'var_18', 'var_19'
           ]
plot_feature_scatter(train[::],test[::], features)
```

```
<Figure size 432x288 with 0 Axes>
```

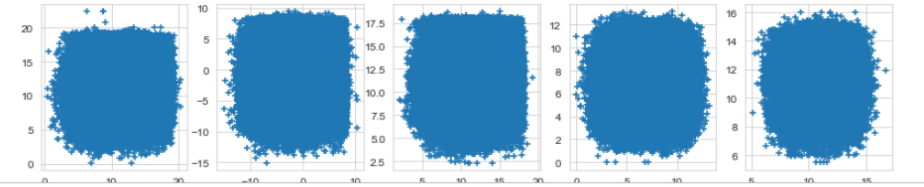


Fig. 1. Matplotlib: the figure shoed the distribution difference is not obvious.

3. Seaborn Suite: used to create charts, more powerful than Matplotlib suite.
(1) Seaborn.countplot: used to draw bar charts.
(2) Seaborn.set_style: Set the table style.
(3) Seaborn.distplot: Only the curve is drawn to draw the histogram.

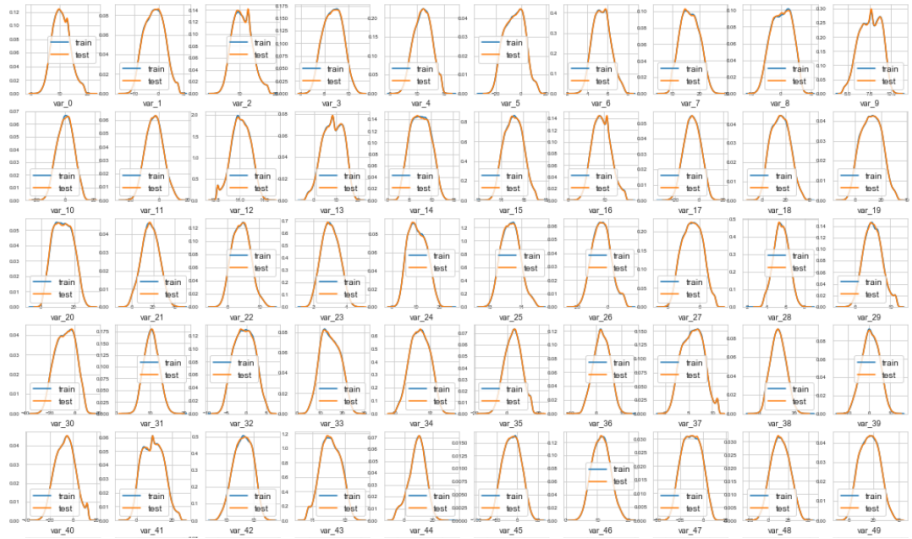


Fig. 2. The training set and test set comparison table. The correlation is high.

4. So Suite: Used to handle external files and directories. Os.listdir: View the list of materials contained in the folder.

Data analysis process this chapter explains the data analysis process of this topic in detail, and explains the use of the module and the parameters of each module. The following 16 steps are used for data analysis.

1.Confirm purpose, 2. Data collection, 3. Data pre-processing, 4.Data Exploration & Visual Design, 5. Model establishment, 6. Model evaluation, 7. Results presented, 8. Data pre-processing, 9. Model establishment, 10. Model evaluation, 11. Data pre-processing, 12. Module creation, 13. Model evaluation, 14. Building the model, 15. Model Evaluation and 16. Results Presentation

Module prediction results After model training, in order to check whether the model has overfitting problems, we use visual analysis to evaluate the model, and continuously modify and adjust parameters. This chapter will introduce the use parameters, chaos matrix, and classification report of each model. Confusion matrix, classification report introduction

Classification report Precision: For prediction accuracy

Forecast accuracy / scoring process This section describes how the prediction accuracy is generated and how to perform module scoring.

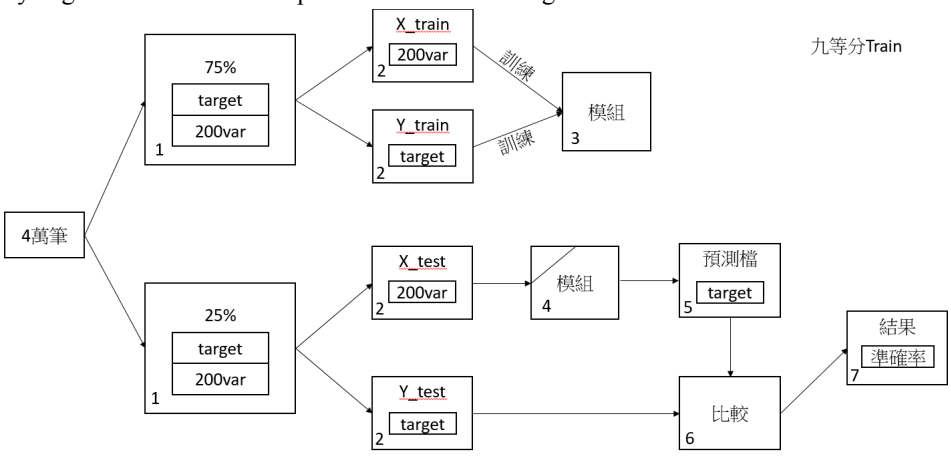


Fig. 3. Prediction accuracy flow chart 1.

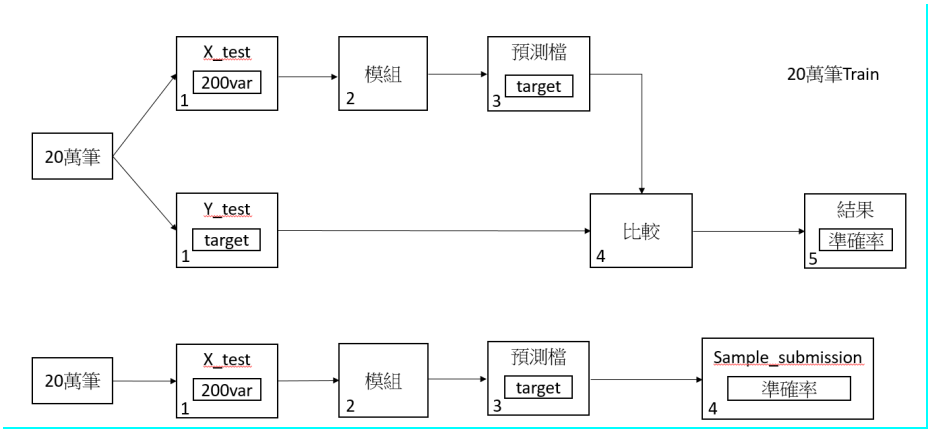


Fig. 4. Prediction Accuracy flow chart 2.

Best prediction results for each model After just processing the data, check the accuracy benchmark value and upload the results. The accuracy rate is only 50%. After 105 uploads, the best prediction result is 81.3%. In this section, the Criterion in each model is divided into entropy and gini for comparison.

Forecast results growth Over time, our prediction results have also broken through. The fourth model line chart can be seen The prediction results from the beginning to the end of the competition have grown significantly, and the reason why model4 was selected as the observation model is that this model's prediction results are better than other models and the growth rate is also superior.

4 Conclusion

In the end, we selected individual training modules and found that the fourth module has the highest accuracy rate, which can reach 83%, but we believe that there is still room for improvement, so we are currently focusing on modifying the fourth module, hoping for the future It can improve the accuracy of nine modules. Use these nine modules to estimate the original 200,000 records, reduce the misjudgment, and improve the performance ranking of each module. After the competition, there are many discussion forums on the platform, including machine learning algorithms used by participants from all over the world. We have learned that the calculus model we have learned is not enough to achieve good results in the competition. Therefore, actively learn more practical machine algorithms from other participants in the discussion forum, and use the newly learned analysis methods to predict the second competition. I hope that in the future, we can create better results for us.

10

References

1. Corinna Cortes and Vladimir Vapnik, (1995). Suppport-Vector Netowrks, Machine Learning , 20, 273-297. https://www.Kaggle.com/c/santander-customer-transaction-prediction/kernels.
2. Hunt, E. B., Marin, J., & Stone, P. J. (1966). Experiments in induction. New York: Academic.
3. Kaggle-Santander Customer Transaction Prediction Discussion. https://www.Kaggle.com/c/santander-customer-transaction-prediction/discussion.
4. Kaggle-Santander Customer Transaction Prediction Kernels.
5. Leo Breiman, Jerome Friedman, Charles J. Stone, R.A. Olshen, ”Classification and Regression Trees”, Taylor & Francis, 1984
6. McCulloch, Warren S.; Pitts, Walter. A logical calculus of the ideas immanent in nervous activity. The bulletin of mathematical biophysics. 1943-12-01, 5 (4): 115–133. ISSN 0007-4985. doi:10.1007/BF02478259
7. Story and Lessons behind the evolution of XGBoost. [2016-08-01]
8. Story and Lessons behind the evolution of XGBoost. [2016-08-01]
9. XGBoost on FPGAs. [2019-08-01].
10. XGBoost4J. [2016-08-01].

Fluid Simulation with a Dense Space-time Deformation via L_0 Gradient Minimization

Kun Li[1,2], Na Qi[1,3], and Qing Zhu[1,4]

[1] Faculty of Information Technology, Beijing University of Technology, Beijing 100124, China
[2] lkun@emails.bjut.edu.cn
[3] qina@bjut.edu.cn
[4] ccgszq@bjut.edu.cn

Abstract. Fluid simulation can be automatically interpolated by using data-driven fluid simulations based on a space-time deformation. In this paper, we propose a novel data-driven fluid simulation scheme with the L_0 based optical flow deformation method by matching two fluid surfaces rather than the L_2 regularization. The L_0 gradient smooth regularization can result in prominent structure of the fluid in a sparsity-control manner, thus the misalignment of the deformation can be suppressed. We adopt the objective function using an alternating minimization with a half-quadratic splitting for solving the L_0 based optical flow deformation model. We evaluate our proposed method on the fluid simulation of both flying droplets and dam break scenarios. Experiment results demonstrate that our proposed method can generate more realistic fluid surface with the optimal space-time deformation under the L_0 gradient smooth constraint than the L_2 one, and outperform the state-of-the-art methods in terms of both objective and subjective quality.

Keywords: Fluid simulation, space-time deformation, L_0 regularization, sparsity, data-driven

1 Introduction

Fluid simulations are increasingly important in computer graphics and have made an impact in real-time games. The simulation algorithms typically include the physics-based methods and the data-driven approaches. The physics-based fluid solvers[1][2][3][4] for working with the large amounts of simulation data are challenging, as the simulation data is typically just passed on to a rendering stage and any required change means re-starting a new simulation from scratch[5]. A categories of data-driven fluid simulations[6] interpolate between two existing fluid animations based on a space-time deformation. Once the matches of two 4D shapes are computed, a user can freely choose any version in between the two extremes efficiently, without starting a new simulation.

Matching two fluid surfaces means finding the corresponding points between the source surface and the target surface, and this procedure needs to deform the

source onto the target. To choose the suitable overall match, the deformation should be regularized with some constraints. A novel optical flow approach[5] is to register deforming space-time source surfaces and target surfaces for the robust registration and blend the complex volumetric phenomena by minimizing energy function with the L_2 norm smoothness regularizer and the Tikhonov regularizer, which is inherently based on closest distances in four-dimensional Euclidean space. The L_2 norm based smoothness regularization in [5] can be enforced on the gradient of the deformations to suppress the differences of the neighbors, however, this deformations can not always require an ideal deformation for the large differences of the source and target surfaces.

We propose an L_0 based optical flow deformation optimization framework to reduce the undesirably artifacts by employing sparse smoothness constraint, and then to match the appropriate points of each shape. Because some large differences between source and target shapes have non-zero deformation gradient values while other regions keep zero, which indicates the sparse property, we present a new measure via L_0 norm as smoothness regularization constraint. The sparse gradient counting measure can confine the number of deformation gradients changes, which allows large gradient magnitude by nature. In addition, to retrieve the motion optimally, this sparse property shows that different degrees of smoothness deformation priors should be introduced to different regions. As a consequence, our proposed L_0 based method can retrieve an optimal deformation to yield better motion than the L_2 based method [5] and avoid suffering from the error accumulation.

We demonstrate that our proposed L_0 based optical flow deformation method can robustly and automatically match the surfaces of the fluid simulation data produced by semi-Lagrangian advection method than the L_2 ones [5, 6]. Our proposed method with the sparsity measure can outperform both the user guidance registration and the third data-point support in term of accuracy with small details and moderate correspondences.

2 Related work

In this section, we briefly introduce the related work about the motion reconstruction method with sparse constraints.

Several works have used guide behaviors to make liquid surfaces follow a desired shape or motion [7, 8]. Unlike this modification approaches on existing fluid simulation, the work-flow that concentrates on align pre-computed space-time surfaces and interpolates between them for intermediate motions is mainly based on non-rigid ICP algorithms [6]. The non-rigid ICP algorithm has been applied to reconstruct the temporally coherent dynamic objects from data captured by the multi-view acquisition system [9] or the monocular depth camera [10]. This registration technique has also been applied to the tracking of changing topology shapes such as liquid surfaces [11].

Regular optical flow aims at recovering the motion of dynamic objects, especially for consecutive frames. Up to now, the Horn-Schunck algorithm [12] is

extended to estimate high quality flow field. The connection between the optical flow and fluid flow are established for typical flow visualizations [13]. The optical flow techniques have also been used to acquire the fluid velocities based on physical priors [14]. Optical flow estimation commonly minimizes an energy function to obtain an optimal result, where the smoothness regularization is incorporated as a convex regularizer. The optical flow deformation approach [5] is essentially similar to the non-rigid iterative closest point algorithm that incorporates user guidance [6] in term of the registration process. Likewise, this optical flow deformation approach robustly recovers a match among the inputs based on local smoothness constraints and obtains a deformation field to align the surfaces. The L_2 based smoothness regularization [5] may cause the error propagation and the accumulated errors. In this paper, we utilize a specific sparse regularization as a constraint to improve the results.

In fact, sparse representation [15] has successfully applied in field of image processing [16], computer vision [17], and graphics [10]. Especially, the sparsity priors have been utilized in the optical flow estimation and motion reconstruction. For example, a new sparsity prior of the flow field in certain domains is introduced to estimate dense flow fields, and a stronger additional sparsity constraint on the flow gradients is incorporated into the model to cope the measurement noises [17]. While [18] decomposes the motion field into sparse and non-sparse components to restrain the over-sparsity. In the motion tracking [10], the L_0 based motion regularizer is incorporated into the non-rigid deformation to propose the L_0-L_2 non-rigid motion tracking method to efficiency stop the error propagation.

3 Method

The 4D registrations and interpolation pipeline is processed automatically as illustrated in figure 1. The optical flow solve computes a dense field of four-component deformation vectors $\mathbf{u}$, which maps points of input A onto input B. The residual optical flow is performed until the refinement does not yield gains in quality. Then the deformation sequences from each iteration are aligned into a single deformation field. It is observed that the source surface should ideally deform to end up on the target surface. Based on this goal, the projection step [11] is adopted to retrieve the fine details at the surface that prevented by regularization. It is notable that the projection only gives good results when a suitable overall match has been obtained by optical flow step. After deformations have been calculated, interpolation step [5] is further utilized to produce intermediate surface animations. Given two inputs sequences ψ_1, ψ_2 with deformations $u_{1\to2}$ and $u_{2\to1}$, the linear interpolation can significantly generate the final quality.

3.1 Optical flow deformation

The work of [5] uses the L_2 norm based optical flow solve to perform a registration step. In this setting, two animations from different simulation setups

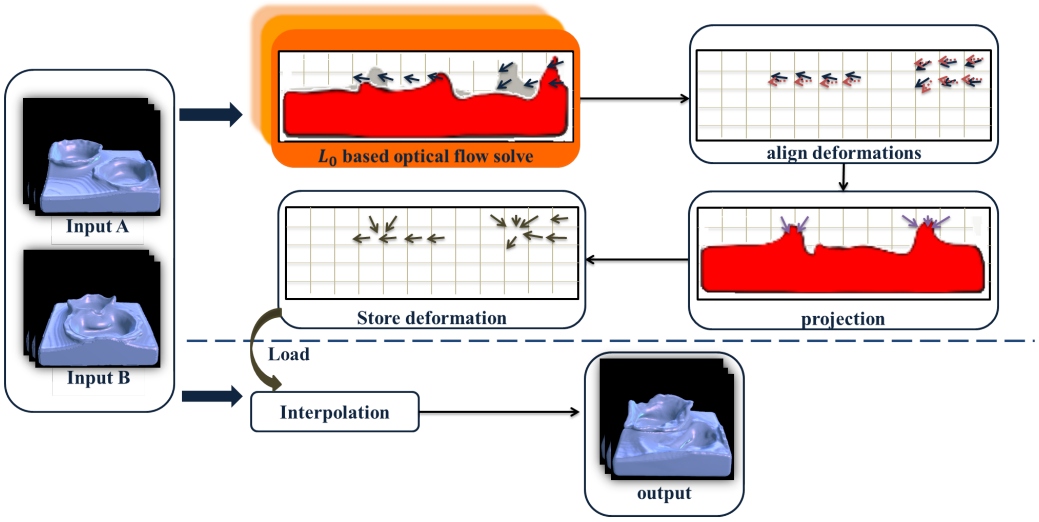


Fig. 1. The pipeline of the registration and interpolation scheme. The inputs animation show two crown splashes. Our proposed L0 based optimization is applied into the optical flow solve.

are taken as the inputs: a source and a target. Each animation is a sequence of surfaces over time, where the surfaces can be concatenated as time slices of a 4D volume. The registration mainly transforms source to further fit it into target through deformations u.

Throughout this paper we will use symbol (e.g.,u) to denote continuous value, and bold symbol (e.g.,$\mathbf{u}$) to denote the discrete counterpart. The symbol with subscript (e.g.,ϕ_r) is usually used to represent the derivative of the corresponding variable. Thus, subscript will not denote the component of a vector except for the special statement.

A flow simulations can be represented by 4D signed-distance values ϕ, and the parameter r is an external control knob to modify the simulation from source to target. The regular optical flow usually recovers the motion of dynamic object from two images taken at different times. [5] instead uses optical flow to recover the motion among surfaces of different simulations. The algorithm employs the brightness-constancy assumption to recover the flow with motion discontinuities.

Applying the optical flow alone is underdetermined because of the non-linearity, so that various additional constraints or priors are imposed to regularize the problem. The Tikhonov regularizer term restricts the deformation to be as small as possible. E_{smooth} defines a L_2 regularizer which smoothes the consistent motions along spatial as well as temporal dimensions. Given the optical flow constraint and some common regularizers, the estimation of deformation field u is achieved by minimizing the following energy

$$\min_u E_{data}(u) + \beta_T E_{Tikhonov}(u) + \beta_S E_{smooth}(u) \tag{1}$$

Where β_S and β_T are the weights directly controlling the importance of each energy term.

3.2 L_0 based optical flow deformation

The goal of optical flow in this work is to compute deformations and to recover the discontinuous fluid motions. In the following, a sparsity constraint on the deformation gradients is incorporated into the optical flow estimation model to reduce the undesirable errors.

The L_2 based optical flow solve [5] for the certain inputs produced by semi-Lagrangian advection algorithm cannot robustly and accurately match surfaces in the nearest-neighbor fashion. Especially in the inputs with little similarities because of the surface flickering, the deformations will try to match pieces in proximity but not necessarily the whole target shape. We propose an L_0 based deformation regularizer over the existing optical flow solve to implicitly define the appropriate correspondences and to prevent motion changes on other regions.

Specifically, the deformation is denoted by $\mathbf{u}$ to transform the source surface closer into the target surface. We formulate the L_0 based deformation by minimizing the following energy function

$$\min_{\mathbf{u}} E_{data}(\mathbf{u}) + \beta_T E_{Tikhonov}(\mathbf{u}) + \beta_S E^{'}_{smooth}(\mathbf{u}) \tag{2}$$

where a new smoothness regularizer $E^{'}_{smooth}$ is introduced to the general optical flow solve to replace the L_2 norm smoothness term. $E^{'}_{smooth}$ measures the deformation differences between all pairs of neighboring points, and is given by

$$E^{'}_{smooth} = S(\mathbf{u}) \tag{3}$$

$$S(\mathbf{u}) = \#\left\{\ p\ \middle|\ |\partial_x \mathbf{u}^p| + |\partial_y \mathbf{u}^p| + |\partial_z \mathbf{u}^p| + |\partial_t \mathbf{u}^p| \neq 0\right\} \tag{4}$$

where $S(\mathbf{u})$ is the L_0 norm for the deformation gradient $\nabla \mathbf{u}$. The gradient $\nabla \mathbf{u}^p = \{\partial_x \mathbf{u}^p, \partial_y \mathbf{u}^p, \partial_z \mathbf{u}^p, \partial_t \mathbf{u}^p\}$ for each point p of the deformation field is calculated based on the two 4D SDFs between neighboring points in space and time. The gradient magnitude $|\partial \mathbf{u}^p|$ is defined as the sum of gradient magnitudes over four dimensions. Our smoothness constraint with L_0 norm regularization is used to control the number of non-zero gradients globally and present the sparsity of deformation gradients.

We estimate the optimized deformation by minimizing the energy function

$$\min_{\mathbf{u}} \left\{ \sum_p |\boldsymbol{\psi}_r^p + \nabla \boldsymbol{\psi}^p \cdot\ \mathbf{u}^p|^2 + \beta_T |\mathbf{u}^p|^2 + \beta_S S(\mathbf{u}) \right\} \tag{5}$$

where the data derivative $\boldsymbol{\psi}_r$ corresponds to the time derivation for optical flow applications. The spatial gradient $\nabla \boldsymbol{\psi}$ is computed with the target surface. The $E^{'}_{smooth}$ term of L_0 norm regularizer in (2) is a non-convex problem. Due to the computationally intractability of the L_0 norm, the half-quadratic splitting strategy utilized in image smoothing achieves the high qualitative effect [16].

Inspired by the strategy described in [16], we split the objective function (5) into two sub-problems by introducing auxiliary variables into the energy function. We argue that this method finds appropriate matches and retrieves the surface motions robustly.

3.3 solver

We introduce an auxiliary variable $\mathbf{v}^p$, corresponding to $\nabla\mathbf{u}^p$, and reformulate the energy function as

$$\min_{\mathbf{u},\mathbf{v}} \left\{ \sum_p |\boldsymbol{\psi}_r^p + \nabla\boldsymbol{\psi}^p \cdot \ \mathbf{u}^p|^2 + \beta_T|\mathbf{u}^p|^2 + \beta_S S(\mathbf{v}) + \beta_A|\nabla\mathbf{u}^p - \mathbf{v}^p|^2 \right\} \tag{6}$$

where $S(\mathbf{v}) = \#\left\{ p \ \middle| |\partial_x\mathbf{v}^p| + |\partial_y\mathbf{v}^p| + |\partial_z\mathbf{v}^p| + |\partial_t\mathbf{v}^p| \neq 0 \right\}$, and β_A is an automatically adapting parameter to control the similarity between the gradients $\nabla\mathbf{u}^p$ and their corresponding variables $\mathbf{v}^p$. To solve this problem, we alternatively fix variables $\mathbf{u}^p$ to solve $\mathbf{v}^p$ and fix variables $\mathbf{v}^p$ to solve $\mathbf{u}^p$ in each iteration.

minimizing v The terms not involving $\mathbf{v}^p$ in (6) are removed to estimate $\mathbf{v}^p$

$$\min_{\mathbf{v}} \left\{ \sum_p |\nabla\mathbf{u}^p - \mathbf{v}^p|^2 + \frac{\beta_S}{\beta_A} S(\mathbf{v}) \right\} \tag{7}$$

As $\nabla\mathbf{u}^p$ is pre-fixed, (7) has a close form solution

$$\mathbf{v}^p = \begin{cases} (0,0,0,0), & (\partial_x\mathbf{u}^p)^2 + (\partial_y\mathbf{u}^p)^2 + (\partial_z\mathbf{u}^p)^2 + (\partial_t\mathbf{u}^p)^2 \leq \frac{\beta_S}{\beta_A} \\ (\partial_x\mathbf{u}^p, \partial_y\mathbf{u}^p, \partial_z\mathbf{u}^p, \partial_t\mathbf{u}^p), & otherwise \end{cases} \tag{8}$$

minimizing u As $\mathbf{v}^p$ is pre-fixed, the $\mathbf{u}^p$ estimation can be represented in a quadratic function

$$\min_{\mathbf{u}} E(\mathbf{u}) = \min_{\mathbf{u}} \left\{ \sum_p |\boldsymbol{\psi}_r^p + \nabla\boldsymbol{\psi}^p \cdot \mathbf{u}^p|^2 + \beta_T|\mathbf{u}^p|^2 + \beta_A|\nabla\mathbf{u}^p - \mathbf{v}^p|^2 \right\} \tag{9}$$

(9) formulates a L_2 based minimization. The solution is similar to the method of L_2 based optical flow [5]. Setting the gradient of $E(u)$ to be zero gives a closed solution for (9), which can be expressed as a linear equation

$$\left([\nabla\boldsymbol{\psi}]^T [\nabla\boldsymbol{\psi}] + \beta_T\mathbf{E} + \beta_A \mathbf{L}\right) \cdot \mathbf{u} = \beta_A\mathbf{v} - [\nabla\boldsymbol{\psi}]^T \boldsymbol{\psi}_r \tag{10}$$

The $\boldsymbol{\psi}_r$ is computed with $\boldsymbol{\psi}_1 - \boldsymbol{\psi}_2$, where the source and target inputs are denoted by $\boldsymbol{\psi}_1$ and $\boldsymbol{\psi}_2$ respectively. $\mathbf{L}$ denotes the discrete Laplacian.

The optical flow based on L_0 regularizer is summarized in Algorithm 1. Initially, the algorithm is performed with two SDF inputs ψ and a zero deformation

Algorithm 1: Optical Flow based on L_0 regularizer

Input: *inputs* $\boldsymbol{\psi}_1$ *and* $\boldsymbol{\psi}_2, \mathbf{u}$, *smoothing weight* β_S, *parameters* $\beta_T, \beta_A, \alpha, \kappa$
Output: *Optimized* $\mathbf{u}$
Initialization : $\beta_A \leftarrow 2\beta_S, \mathbf{u}_0 \leftarrow 0, \mathbf{v}_0 \leftarrow \mathbf{u}_0$
for $i = 1 : max$ **do**
  $\boldsymbol{\psi}_1^{'} = deform(\boldsymbol{\psi}_1, \mathbf{u}^{(i-1)}, \alpha)$;
  $\mathbf{v}^{(i)} \leftarrow \mathbf{min}_{\mathbf{v}}(\beta_S, \beta_A, \mathbf{u}^{(i-1)})$
  $\mathbf{u}^{tmp(i)} \leftarrow \mathbf{min}_{u}(\boldsymbol{\psi}_1^{'}, \boldsymbol{\psi}_2, \mathbf{v}^{(i)}, \beta_T, \beta_A)$
  begin
    $\mathbf{u}^{tmp} = \mathbf{u}^{(i-1)} + alignVelocity(\mathbf{u}^{tmp(i)})$;
    if $error(deform(\boldsymbol{\psi}_1, \mathbf{u}^{(i)}, \alpha), \boldsymbol{\psi}_2) \geq error(deform(\boldsymbol{\psi}_1, \mathbf{u}^{tmp}, \alpha), \boldsymbol{\psi}_2)$ **then**
      $\mathbf{u}^{(i)} = \mathbf{u}^{tmp}$;
      $\beta_A = \kappa \cdot \beta_A$
    else
      break;
    end
  end
end

field $\mathbf{u} = 0$. The function $deform(\boldsymbol{\psi}, \mathbf{u}, \alpha)$ applies the deformation $\mathbf{u}$ to input $\boldsymbol{\psi}$ with weight $\alpha \in [0...1]$. A single, smooth linear deformation from source to target is required rather than deformation sequences produced by the iteration runtime. We follow [5] to align multiple deformations with the $alignVelocity(\mathbf{u})$ function. The reliable error metric [5] is adopted to quantitatively evaluate whether the current optical flow solve leads to an improvement. More emphasis are putted on the surface region rather than the volume itself. Given two inputs $\boldsymbol{\psi}_1$ and $\boldsymbol{\psi}_2$(which we define them in a discrete setting), the error metric becomes

$$error(\boldsymbol{\psi}_1, \boldsymbol{\psi}_2) = \sum_x h(\boldsymbol{\psi}_1(x), \boldsymbol{\psi}_2(x))V \tag{11}$$

where V is the volume of a cell. The h represent an indicator function to detect misalignment surfaces.

4 Experimental Results

We conduct the experiments on the fluid simulation of both the flying droplets and dam break scenarios. The standard fluid simulation, surface matching process and the interpolation step are run on the extensible fluid simulation framework, called mantaflow [19]. When the deformations have been calculated, new fluid animations can be produced in the interpolation weight range from 0 to 1. In order to speed up the convergence, we monotonically increase β_A by multiplying the parameter $\kappa = 2$ in the Algorithm 1. Thus, in our experiments,

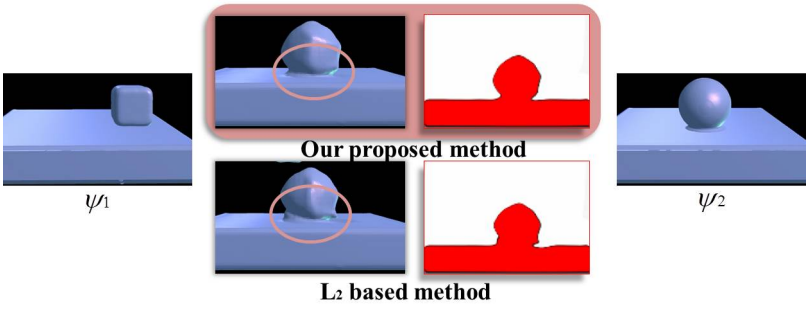


Fig. 2. The visual quality comparison of flying droplets via different deformation methods based on L_2 and L_0 smooth regularization. Left: The source input ψ_1. Middle: The results of L_2 based method [5] and L_0 based method, and the sectional view images correspond to the results. Right: The target input ψ_2.

we set β_A to be $1E-3$ in the first iteration and increase it by multiplying parameter κ after each iteration when error is reduced. In the following, we compare our L_0 based refinement results with those of the state-of-the-art [5] to evaluate the performance of these two different methods.

4.1 Deformation Results

We evaluate the quality of the deformation through the visual quality of the interpolated results, where the uni-directional nearest-neighbor interpolation is used, such as the deformation direction from ψ_1 to ψ_2.

We consider two flying droplets with different initial shapes and different positions. They drop in the pool at different times to create different water ripples. We try to deform the source one to the target one by our proposed fluid simulation method and the state-of-the-art method [5] respectively, with the interpolation weight as 1. Figure 2 shows the interpolated results of the L_2 based and the L_0 based deformation methods respectively. In order to easily observe, the sectional view images at a fixed frame are shown next to the interpolated results. We can observe that our results are closer to the target input ψ_2 and with less artifacts than those of the L_2 based method. In addition, the interpolated results of these two different methods with the deformation from the input ψ_2 to the input ψ_1 are shown in Figure 3. The L_2 based deformation method may create some ambiguous correspondences between the input ψ_1 to the input ψ_2, and the undesired ripples appear at the surface due to the suboptimal matching. Our proposed method can alleviate the artificial motions brought by the L_2 based method effectively. Three examples of interpolated results at different frames for flying droplets are demonstrated in Figure 4, where the droplets falling into the pool and causing the varying splashes. We observe that our interpolated results can preserve the large-scale structures and recover the small-scale details.

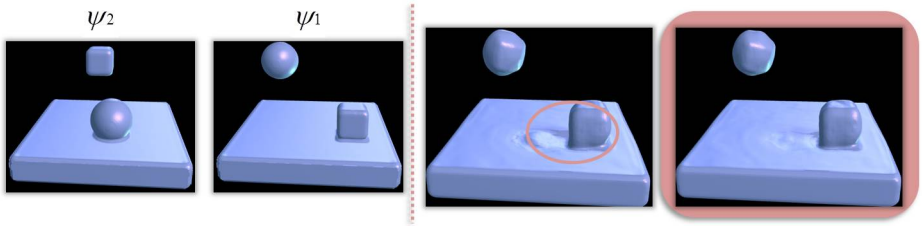


Fig. 3. The results of flying droplets with the deformation from source to the target. Left: The source input ψ_2 and the target inputs ψ_1 are as the input. Right: The interpolated results of [5] and our proposed L0 based deformation method, respectively.

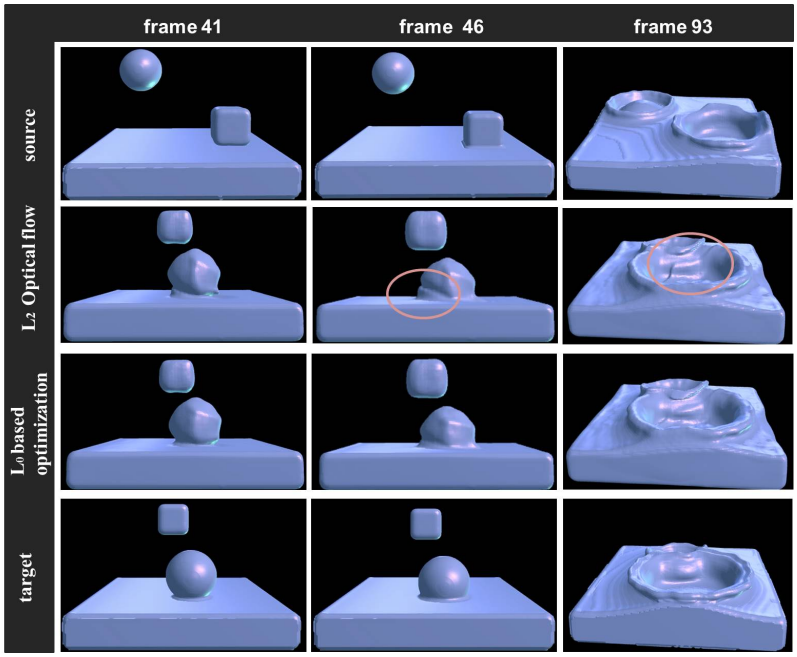


Fig. 4. A visual comparison between the L_2 based optical flow deformation, and our proposed L_0 based optimization method. From top to bottom: the source fluid, the interpolated fluid results with deformation solved by the L_2 based method [5] as well as our L_0 based optimization method, and the target fluid respectively. The figure shows two drops falling into a pool, which causes the crown splash.

We take the dam break as an example to demonstrate the accuracy of our optimization method. In this set of animations, a block of water is released at one side of a pool, and this forms a wave that sweeps the pool. As shown in Figure 5, we produce fluid animations for two different widths for this block of water. The interpolated results are generated with the interpolation weight as 1 to compare the generated surfaces with the target surfaces visually. As shown in

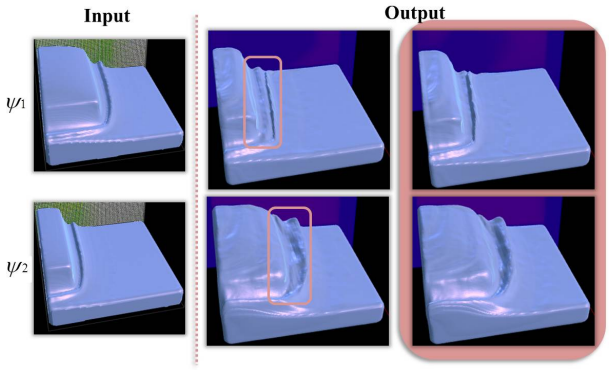


Fig. 5. A comparison of the L_2 based optical flow against our L_0 based optimization output for the dam break scenario. Left: the input ψ_1 and ψ_2. Middle: the interpolated results generated by using the L_2 based method with two different directional deformations including $\psi_1 \rightarrow \psi_2$ and $\psi_2 \rightarrow \psi_1$ with the interpolation weight 1, respectively. Right: the corresponding interpolated results by using the L_0 based method.

the right columns of Figure 5, the L_2 based method causes the undesired ripples at the surface, and the L_0 optimization method captures the accurate motions of the waves. Therefore, the surface of our results based on the L_0 regularization is closer to the target one than that based on the L_2 optimization.

4.2 Evaluation

The error metric mentioned in sec.3.3 can quantitatively and significantly detects the differences between the deformed surface and the target surface. We utilize the error metric (11) to calculate the errors between the target surfaces and the generated surfaces which are deformed with the L_2 based and the L_0 based deformation methods respectively. The error values can be found in Table 1, which includes two aforementioned scenarios (droplet, dam break). We show the results with the deformation direction from ψ_1 to ψ_2 and the inverse one. We can observed that our proposed L_0 based fluid simulation method can obtain less errors than the L_2 based method [5], and generate the desired correspondences.

Table 1. Errors between the deformed surface and the target surface.

Fluid Example	Directions	initial error	the error after deformation	
			L_2	L_0
Flying Droplets	$\psi_1 \rightarrow \psi_2$	4.1207	1.1113	**0.9440**
	$\psi_2 \rightarrow \psi_1$	3.9823	1.9747	**1.8369**
Dam Break	$\psi_1 \rightarrow \psi_2$	4.0991	0.6768	**0.6082**
	$\psi_2 \rightarrow \psi_1$	5.0944	1.5366	**1.3622**

5 Conclusions

In this paper, we proposed a L_0-norm based deformation for fluid simulation to obtain a good alignment of two inputs. The main contribution of our work is a L_0-norm optimization approach to align fluid simulation with changing topologies. The L_0 based optical flow takes advantage of the sparsity property of the fluid motion of neighboring frames to constrain the deformation. Moreover, we propose a novel effective approach to solve the L_0-based optimization problem by making use of the L_0 gradient minimization, which can globally control the non-zero gradients to approximate the deformation in a sparsity-control manner. Experimental results demonstrate that our method outperforms the state-of-the-art L_2 based optical flow and accurately recovers motions between the space-time surfaces of different simulations.

6 Acknowledgement

This work was partially supported by the National Natural Science Foundation of China under Grants (61906009 and 61976011), and the scientific Research Common Program of Beijing Municipal Commission of Education (KM202010005018).

References

1. Foster, N., Metaxas, D.: Realistic animation of liquids. In: Proceedings of the Conference on Graphics Interface '96, GI '96, pp. 204–212. Canadian Information Processing Society, Toronto, Ont., Canada, Canada (1996). URL http://dl.acm.org/citation.cfm?id=241020.241077
2. Müller, M., Charypar, D., Gross, M.: Particle-based fluid simulation for interactive applications. In: Proceedings of the 2003 ACM SIGGRAPH/Eurographics Symposium on Computer Animation, SCA '03, pp. 154–159. Eurographics Association, Aire-la-Ville, Switzerland, Switzerland (2003). URL http://dl.acm.org/citation.cfm?id=846276.846298
3. Solenthaler, B., Pajarola, R.: Predictive-corrective incompressible sph. In: ACM SIGGRAPH 2009 Papers, SIGGRAPH '09, pp. 40:1–40:6. ACM, New York, NY, USA (2009). DOI 10.1145/1576246.1531346. URL http://doi.acm.org/10.1145/1576246.1531346
4. Stam, J.: Stable fluids. In: Proceedings of the 26th Annual Conference on Computer Graphics and Interactive Techniques, SIGGRAPH '99, pp. 121–128. ACM Press/Addison-Wesley Publishing Co., New York, NY, USA (1999). DOI 10.1145/311535.311548. URL http://dx.doi.org/10.1145/311535.311548
5. Thuerey, N.: Interpolations of smoke and liquid simulations. ACM Trans. Graph. **36**(1), 3:1–3:16 (2016). DOI 10.1145/2956233. URL http://doi.acm.org/10.1145/2956233
6. Raveendran, K., Wojtan, C., Thuerey, N., Turk, G.: Blending liquids. ACM Trans. Graph. **33**(4), 137:1–137:10 (2014). DOI 10.1145/2601097.2601126. URL http://doi.acm.org/10.1145/2601097.2601126

7. Raveendran, K., Thuerey, N., Wojtan, C., Turk, G.: Controlling liquids using meshes. In: Proceedings of the 11th ACM SIGGRAPH / Eurographics Conference on Computer Animation, EUROSCA'12, pp. 255–264. Eurographics Association, Aire-la-Ville, Switzerland, Switzerland (2012). DOI 10.2312/SCA/SCA12/255-264. URL http://dx.doi.org/10.2312/SCA/SCA12/255-264
8. Pan, Z., Huang, J., Tong, Y., Zheng, C., Bao, H.: Interactive localized liquid motion editing. ACM Trans. Graph. **32**(6), 184:1–184:10 (2013). DOI 10.1145/2508363.2508429. URL http://doi.acm.org/10.1145/2508363.2508429
9. Li, H., Luo, L., Vlasic, D., Peers, P., Popović, J., Pauly, M., Rusinkiewicz, S.: Temporally coherent completion of dynamic shapes. ACM Trans. Graph. **31**(1), 2:1–2:11 (2012). DOI 10.1145/2077341.2077343. URL http://doi.acm.org/10.1145/2077341.2077343
10. Guo, K., Xu, F., Wang, Y., Liu, Y., Dai, Q.: Robust non-rigid motion tracking and surface reconstruction using l0 regularization. In: Proceedings of the 2015 IEEE International Conference on Computer Vision (ICCV), ICCV '15, pp. 3083–3091. IEEE Computer Society, Washington, DC, USA (2015). DOI 10.1109/ICCV.2015.353. URL http://dx.doi.org/10.1109/ICCV.2015.353
11. Bojsen-Hansen, M., Li, H., Wojtan, C.: Tracking surfaces with evolving topology. ACM Trans. Graph. **31**(4), 53:1–53:10 (2012). DOI 10.1145/2185520.2185549. URL http://doi.acm.org/10.1145/2185520.2185549
12. Horn, B.K.P., Schunck, B.G.: Determining optical flow. Artif. Intell. **17**(1-3), 185–203 (1981). DOI 10.1016/0004-3702(81)90024-2. URL http://dx.doi.org/10.1016/0004-3702(81)90024-2
13. LIU, T., SHEN, L.: Fluid flow and optical flow. Journal of Fluid Mechanics **614**, 253291 (2008). DOI 10.1017/S0022112008003273
14. Gregson, J., Ihrke, I., Thuerey, N., Heidrich, W.: From capture to simulation: Connecting forward and inverse problems in fluids. ACM Trans. Graph. **33**(4), 139:1–139:11 (2014). DOI 10.1145/2601097.2601147. URL http://doi.acm.org/10.1145/2601097.2601147
15. Zhang, Z., Xu, Y., Yang, J., Li, X., Zhang, D.: A survey of sparse representation: algorithms and applications. CoRR **abs/1602.07017** (2016). URL http://arxiv.org/abs/1602.07017
16. Xu, L., Lu, C., Xu, Y., Jia, J.: Image smoothing via l0 gradient minimization. In: Proceedings of the 2011 SIGGRAPH Asia Conference, SA '11, pp. 174:1–174:12. ACM, New York, NY, USA (2011). DOI 10.1145/2024156.2024208. URL http://doi.acm.org/10.1145/2024156.2024208
17. Shen, X., Wu, Y.: Sparsity model for robust optical flow estimation at motion discontinuities. In: The Twenty-Third IEEE Conference on Computer Vision and Pattern Recognition, CVPR 2010, San Francisco, CA, USA, 13-18 June 2010, pp. 2456–2463 (2010). DOI 10.1109/CVPR.2010.5539944. URL https://doi.org/10.1109/CVPR.2010.5539944
18. Chen, Z.: Decomposing and regularizing sparse/non-sparse components for motion field estimation. In: Proceedings of the 2012 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), CVPR '12, pp. 1776–1783. IEEE Computer Society, Washington, DC, USA (2012). URL http://dl.acm.org/citation.cfm?id=2354409.2354857
19. Thuerey, N., Pfaff, T.: MantaFlow (2018). *http://mantaflow.com*

A Deep Learning Framework for Character Gait Motion Control with Physical Model

Qing Zhu[1] and Chuanhua Liu[2]

[1,2] Beijing University of Technology, No.100, Pingleyuan, Chaoyang District, Beijing, China
ccgszq@bjut.edu.cn

Abstract. We present a framework for generating gait actions of a virtual character in real-time. Integrate physical IK into the model make the neural network to learn the movements of the main joints more specifically. Our method takes the action state of the character's previous frame and the terrain of the scene as input user controls, and automatically generates high-quality motions to achieve the required user controls. The entire network is trained end-to-end on a large data set consisting of gait movements. Therefore, our method can automatically generate characters to adapt to the movement of different terrains. Compare with the Gaussian process and EDR model, our network structure can produce higher quality results. Our work is best suited for controlling characters in interactive scenes, such as computer games and virtual reality method.

Keywords: machine learning, character motion, neural network

1 Introduction

Generating real-time data-driven controllers for virtual characters remains a challenging task. Even if there is a large amount of ready-made high-quality motion capture data, it requires a long manual process before it can be used. As animation complexity increases, a very large animation database is needed to ensure that the virtual character can move freely in the scene. This method not only consumes a lot of manpower and material resources, but also requires a lot of storage space during operation.

Therefore, we research and explore the construction method of character gait movement model in virtual scene. A large amount of data is used to pre-train the motion model to solve the deficiencies of the previous method in practical applications, enhance the interaction of the characters in the virtual reality scene, improve the user's immersion in the virtual reality world, and seek better user experience.

We propose a virtual character real-time gait motion control based on deep learning and physical IK. Construct a physics-based motion model for the character, and learn a large amount of human motion data to obtain an efficient real-time motion model.

2 Related Work

There are many existing virtual character motion models, the most basic of which are a data-driven motion synthesis model. The data-driven motion synthesis model requires a large amount of motion data and synthesizes it to ensure that the output motion looks natural and smooth. Commonly used data-driven motion synthesis technologies include linear-based technology and kernel-based technology.

Linear-based Techniques. Howe [1] and Safonova [2] used principal component analysis (PCA), which is widely used to reduce the dimension of motion data, to predict whole body motion from a small number of inputs. However, because the global PCA may have the problem of representing various motions in the low-dimensional latent space, it is improved on the later stage, and the local PCA are used to process any type of motion. Chai [3] applied local PCA to synthesize whole body motion with sparse marker sets. Tautges [4] designed a similar local structure of predicting whole-body motion from sparse inertial sensors. The problem with this structure is that it requires a large amount of data preprocessing and calculation for motion segmentation, classification and alignment during training, and nearest neighbor search at runtime.

Kernel-based Methods for Motion Blending. Radial basis function (RBF) is effective for mixing multiple motions of the same class. Rose [5] designed to classify the same type of actions and use RBF to insert according to the direction of the character's phase movement. Rose [6] proposed a method for classifying motion and aligning it along the timeline to make the mixed motion look more reasonable, and used inverse kinematics to map the joint position to the character's pose. Kovar and Gleicher [7] calculated the similarity of motions and used dynamic time warping to align the motions and let others automatically pair. However, due to the lack of a mechanism to deal with noise and variance, the RBF method can easily overfit the data. Mukai and Kuriyama [8] solved this problem by using a Gaussian process (GP), in which meta-parameters are optimized to fit the model to the data. Therefore, the number of sports that can be mixed is limited. This limitation is difficult to break through.

Physics-based. This is a technique for automatically generating interactive natural motion. It has been a long-term research goal to extend this method to the virtual role of active control, and great progress has been made in recent years. Motion controllers have been developed that can powerfully handle character shapes, external disturbances, and changes in uneven terrain. Geijtenbeek [9] designed a muscle-based control method to simulate bipeds. This method can support various creatures' speed-shifting and turning behaviors, and is robust to external disturbances. However, this method has low fidelity to human gait movements, especially upper body movements, and it is difficult to simulate them with pure biological modules.

Even with a large amount of readily available high-quality motion capture data, making real-time data-driven controllers for virtual characters is still very difficult. It is often difficult to ensure low storage requirements when real-time performance is met, and manual pre-processing in the early stage will also take a lot of manpower and time. Especially when the virtual environment consists of uneven terrain and large obstacles, the problem will be more difficult. In this case, a framework that can be learned from a very large amount of high-dimensional motion data is needed, as there are large

combinations of different motion trajectories and corresponding geometries that can exist. Recent developments in deep learning and neural networks have shown some promise that may solve these problems. Neural networks are able to learn from very large, high-dimensional data sets, and once trained have a low memory footprint and fast execution time. The question remains how neural networks can best are applied to motion data in a way that produces high-quality output in real time with minimal data processing.

Deep Learning. Holden [10] [11] and others used CNN convolution models to implement character movements, perform temporal local transformation on each layer of CNN, and gradually transform input signals until a desired output signal is produced. However, this structure is suitable for offline and cannot meet the real-time nature. The RNN and encoder-recursion-decoder (ERD) networks proposed by Taylor [12] and Fragkiadaki [13] are real-time, and these models only need a single frame of content input to get the next frame Output, but once the time increases, the errors in the prediction will be fed back to the input and accumulated quickly, and high frequency noise will be feedback into the system and the final model can no longer continue to run normally.

3 Data Acquisition

3.1 The Motion Dataset for Deep Learning

In order to train the model, we chose a large-scale motion capture online database CMU [14], which is provided free of charge. The database contains data such as walking and jogging to ensure that characters perform gait movements at different speeds and orientations. At the same time, to increase the richness of the data, we performed motion capture. First of all, the data of the data set needs to be normalized, and the data is relocated to data with a uniform skeleton structure with a single proportion and the same bone length. This requires first copying all the corresponding joint angles in the source skeleton structure to the target skeleton structure, then scaling the source skeleton to the same size as the target skeleton, and finally referring to the method of Yamane and Nakamura [15] to perform inverse dynamics on the target skeleton Bind the joints of the moving target skeleton to match the joint positions in the original skeleton. The final data set contains about 2.5 million frames of high-quality character motion capture data, all using 120 frames per second of sampled data.

3.2 Data Format for Training

After obtaining the available motion capture data, we converted the dataset to a format suitable for training. First, we down sampled the data set to 60 frames per second, because we usually cannot reach a frame rate of 120 frames per second when using virtual characters in games, and 60 frames per second is a relatively reasonable and stable frame rate. Also consider actual use cases. In game engines, objects are usually defined using world coordinates. Convert the position and rotation data of the

character's bones from the relative coordinates in the motion capture file to the world coordinate system for subsequent calculations.

Character walking is a complex action, which contains many details. The generation of character gait actions needs to extract features of the character's movement, thereby enhancing the model training efficiency and model operation efficiency.

Character Orientation. The character's orientation is a very important feature data, it can help us analyze the movement track of the character. The orientation of the character should be a two-dimensional vector in the XZ plane. Generally, the orientation of the character should be determined by the orientation of the trunk. But walking is a whole-body exercise. Each joint is doing a relative circular movement. It is a separate joint. The key joints of the trunk, such as the waist or crotch, are not used to represent the orientation of the character. The calculation method selected in this paper is to select the left and right shoulder and left and right thighs root bone positions, and calculate the character orientation from these four data. Then we perform Gaussian filtering on the calculated data to reduce some noises that occur during motion capture, and finally obtain the required character orientation result.

Footsteps Match the Ground. During the walking cycle of a character, the foot is in contact with the ground for a certain period of time. In this paper, the method of reference [16] is used to determine whether the foot is in contact with the ground by detecting the height and speed of the foot or heel of the character.to define the label. Because this method can sometimes go wrong, we need to manually check the results and correct them.

Definition of Gait Semantics. For the category of exercise, this paper uses binary to define several semantic labels for gait: slow walking (00), walking (01), fast walking (10), and running (11). Defining semantic labels can eliminate motion ambiguity. For example, many joints in brisk walking and running have similarities. It is difficult to determine the correct movement based on the calculation of joint positions alone, so artificial preliminary operations need to be added. Secondly, in the application scenario, the user generally has requirements for the gait of the character. In the process of motion capture, in order to ensure the correctness of the action, the actor is usually required to continue a movement state for a period of time, which is also convenient for us to define the semantic label.

4 Gait motion model

4.1 Physical Model Description

We take game engine as an example to study and analyze the real-time control methods of gait movements that were commonly used before. In the game engine, a large amount of animation data can be connected together through the animation controller to achieve animation state fusion and switching. Among them, some parameters define the state of the animation, such as controlling the speed of running or walking, controlling the rotation angle of the steering animation and straight motion, and controlling whether to jump or not.

In the case of focusing only on gait movement, we found that the user's input data only has the moving speed of the character in the vertical (back-and-forth) direction and the moving speed of the horizontal (left-right) direction. All that controls skeletal motion and elevation are the animation controller and rigid body physics engine. Here we consider letting the neural network replace the animation controller and part of the rigid body physics model. We calculate the motion trajectory in real time through user input, then we can build a real-time gait motion control model.

In the previous character gait motion control, when we wanted to change the defined animation, we usually defined the physical IK of the bone. After defining the IK chain for the character model, we can drive a series of nodes to move to the last node, and can maintain relatively correct dynamic motion.

We consider that by analogizing the method to define physical IK of the model, the remaining related joints (such as knees and elbows) can also move reasonably at the same time when only some nodes are controlled.

4.2 Combining with Neural Networks

Inspired by the physical model, we combined the physical model with the neural network to construct a real-time gait motion control model. Combining the advantages of neural networks and physical models, a more lightweight character gait motion model is trained, as shown in Fig. 1.

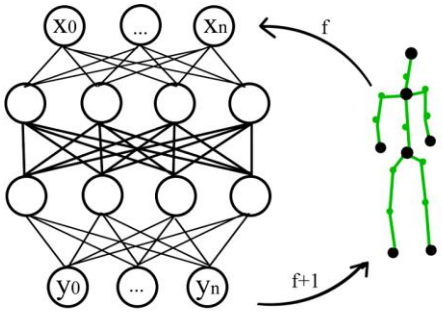


Fig. 1. Structure of this model

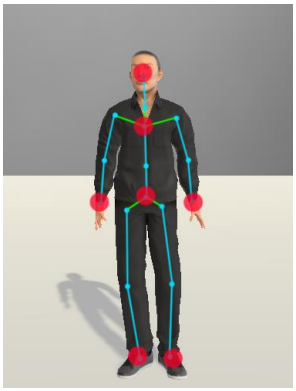


Fig. 2. The definition of character IK

Unlike other methods that use neural networks to generate all joint positions, we reduce the number of trained joints, control the number of joints to the minimum numbers of nodes, and define the remaining joints with IK. This can effectively reduce the size of the neural network and make it more effective to train the joints with high necessity. References [17] [18]Here the following joints are selected and defined in this model: pelvic center (root node), shoulder center, head, left hand, right hand, left foot, right foot. As shown in the Fig. 2, after defining the 7 key joints, IK bind the remaining keys on the physical model to make them driven by IK.

Correspondingly, the definition of neural networks must be changed. First of all, on the input and output data, unnecessary data is first deleted from the data set, leaving only relevant data for 7 joints.
Since the character's use at this moment is related to his movement trajectory, we will further analyze the movement here. The movement of the i frame in a movement is defined by the state of the first n frames and the trend of the next n frames, such as a person from stationary to brisk walking, or a person who is about to turn right, and they are currently The actions of the frame is made up of the previous action and the next action. So, we also use the first n frames to the last n frames of the current frame as input. In actual use, the trajectory of the next n frames can be calculated by the current character's orientation and movement speed and user input. In this model, n is 24, which is the state of one second before and after we will consider this action.

4.3 Neural Network Structure

The input and output of the neural network are defined as $X = \{t^p, t^d, t^h, t^g, j^p, j^v\} \in \mathbb{Q}^n$, $Y_i = \{t^p, t^d, j^p, j^v, j^a, r^x, r^z, r^a, c\} \in \mathbb{Q}^n$.where t^p are the trajectory positions in the 2D horizontal plane, , t^d are the trajectory directions in the 2D horizontal plane, t^h are the trajectory heights points, t^g are the trajectory semantic variables indicating the gait of the character defined in Section 4.2. j^p and j^v are the local joint positions and velocities, j^a are the joint angles local to the character root transform expressed using the exponential map which Grassia proposed[19], r^x, r^z are the root transform translational x velocity and z velocity relative to the forward facing direction, r^a is the root transform angular velocity around the upward direction, c are the foot contact labels defined in Section 4.2

Given input parameters X, output parameters Y, we start by building a simple three-layer neural network Φ as follows:

$$\Psi(\mathrm{X}) = \mathrm{ELU}(\mathrm{ELU}(\mathrm{ELU}(X * \mathrm{W}_1 + \mathrm{b}_1) * \mathrm{W}_2 + \mathrm{b}_2) * \mathrm{W}_3 + \mathrm{b}_3) \quad (1)$$

where, network parameters are defined by $\phi = \{\mathrm{W}_1 \in \mathbb{Q}^{h\times n}, \mathrm{b}_1 \in \mathbb{Q}^{\mathrm{h}}, \mathrm{W}_2 \in \mathbb{Q}^{h\times h}, \mathrm{b}_2 \in \mathbb{Q}^{\mathrm{h}}, \mathrm{W}_3 \in \mathbb{Q}^{\mathrm{m}\times\mathrm{h}}, \mathrm{b}_3 \in \mathbb{Q}^{\mathrm{h}}\}$, h is the number of cells in the hidden layer. which are set to 256.
To train the network we must normalize the input and output sets, ensure that for a given set of control parameters X, we can produce the corresponding output variables Y as a function of the neural network Ψ. The cost function is defined as the following and consists of two terms:

$$Cost(X, \phi) = \|Y - \Psi(\theta)\|_2^2 + \alpha\|\phi\|_1 \quad (2)$$

where, the first term calculates the mean square error of the regression, and the second term is a sparse term to ensure that the minimum number of hidden units is used for regression. Here we set α to 0.05. We train the neural network and use Adam[20] to minimize the Cost function. We set Dropout[21] to 0.5 to reduce the overfitting of the model and enhance the model's resistance to noise. Trained 30epochs and spent about 12 hours

5 Runtime

After training the neural network, we solidify it into a plug-in that can be calculated in real time. After binding with the defined physical model in the game engine, as long as the user inputs the movement direction and speed of the moving character in real time, real-time gait motion control model is ready to run.

Here is a detailed explanation of how each input is calculated during real-time operation. In the game engine, the position coordinates of the character's joints can be read directly. Here we save the joint position vector of the previous frame and subtract it from the current frame to get the speed instantly. Next, we record the coordinates of the root node and the orientation of the person in the first 24 frames, so that the trajectory data before the current frame can be easily obtained. Next, define the trajectory data after the current frame according to the user's input.

We define the user's input as (v, h) representing the forward and backward displacement, respectively, with a value range of [-1,1]. Next, we convert the user's input into three-dimensional coordinating, so that we can get the user's displacement in the world coordinate system.

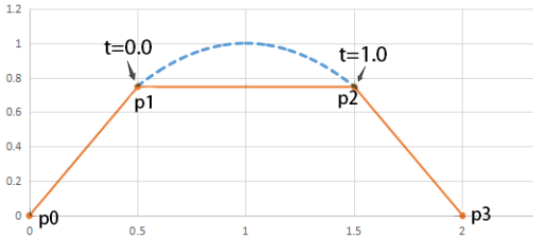


Fig. 3. The illustrations of Catmull-Rom Spline

Because human movement is smooth, here we introduce Catmull-Rom Spline interpolation method to run through the path. Catmull-Rom Spline's interpolation method requires us to find the four points p0, p1, p2, and p3 on the path before generating a smooth point between p1 and p2. as shown in Fig. 3.

$$OutPoint = P_0 * (-0.5t^3 + t^2 - 0.5t) + P_1 * (1.5t^3 - 2.5t^2 + t) + P_2 * (-1.5t^3 + 2t^2 + 0.5t) + P_3 * (-0.5t^3 + t^2) \quad (3)$$

First, we can get the track point (p1) of the current frame and the track point (p0) of the previous frame. Next, we calculate the displacement result of the current user input. For example, if the character makes it turn right as it moves forward, the displacement result should be the front right of the current frame and the character facing right. The simple vector addition can get the displacement result (p2). Next, we calculate the result after 1 second (p3) while keeping the current input. After being incorporated into the formula, the points on the predicted trajectory can be obtained.

6 Result and Discussions

We designed two experimental comparisons, Error1 and Error2. Error1, which calculates the distance difference between each joint of the motion generated by this method and the motion fusion method, and Error2, which is the distance difference between the foot and the ground. Because the animation fusion method and the motion generated by this method cannot be completely consistent. This experiment chooses to be on the same trajectory as the gait motion generated by the animation fusion method, with several cycles of the same step and speed as the method Perform analysis. It does not analyze the movement of the entire trajectory. We define these two parameters for comparison because the motion generated by the animation fusion method can be regarded as a result generally recognized after artificial processing. Compared with it, it can be regarded as a comparison with the subject's actual motion result. The second parameter can be regarded as a criterion for judging the correctness of the action. The gait movement of the character generates effective movement on the one hand, and on the other hand, it is necessary to consider the correctness of the action of the actual landing of the foot.

Table 1. Comparison of four algorithms under different actions

Motion	Method	Error1	Error2	Runtime	Memory
Straight Forward-Start	GP	0.176	0.061	79	2500MB
	ERD	4.193	0.048	324	17.9MB
	NN	0.157	0.034	272	8.4MB
	NN & physical IK	0.163	0.036	96	57.1MB
Straight Forward-Stable	GP	0.053	0.011	161	2500MB
	ERD	0.116	0.037	372	17.9MB
	NN	0.142	0.022	301	8.4MB
	NN & physical IK	0.137	0.019	130	57.1MB
Curve Forward-Slow Turn	GP	0.359	0.064	123	2500MB
	ERD	0.241	0.043	357	17.9MB
	NN	0.265	0.035	288	8.4MB
	NN & physical IK	0.244	0.037	135	57.1MB
Curve Forward-Sharp Turn	GP	0.528	0.072	83	2500MB
	ERD	1.283	0.064	349	17.9MB
	NN	0.262	0.057	265	8.4MB
	NN & physical IK	0.217	0.051	119	57.1MB
Bumpy Ground	GP	-	-	-	-
	ERD	2.872	3.472	261	17.9MB
	NN	0.538	1.756	183	8.4MB
	NN & physical IK	0.706	0.112	76	57.1MB

We chose to compare with three algorithms: autoregressive Gaussian process (GP) [8], ERD (encoder-loop-decoder) network [12] and neural network without adding physical model, the structures are shown in Fig. 4. In order to ensure the fairness of comparison, the sample-based training method uses the same data set training and the same parameter settings as this article. In the experimental scenes, we chose three experimental scenes, which are straight forward, curved forward, and walking on the ground with the height difference.

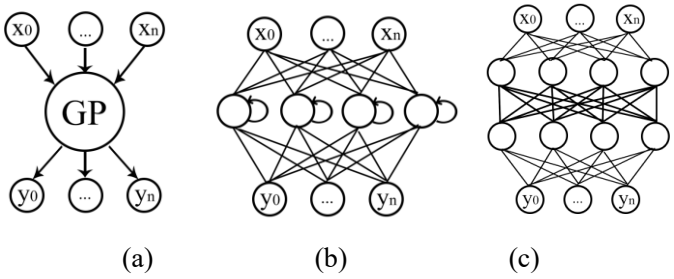


Fig. 4. The configurations of Neural network structures and Gaussian process structures evaluated in our comparison: (a) Gaussian process (b) ERD network (c) Neural network without adding physical model

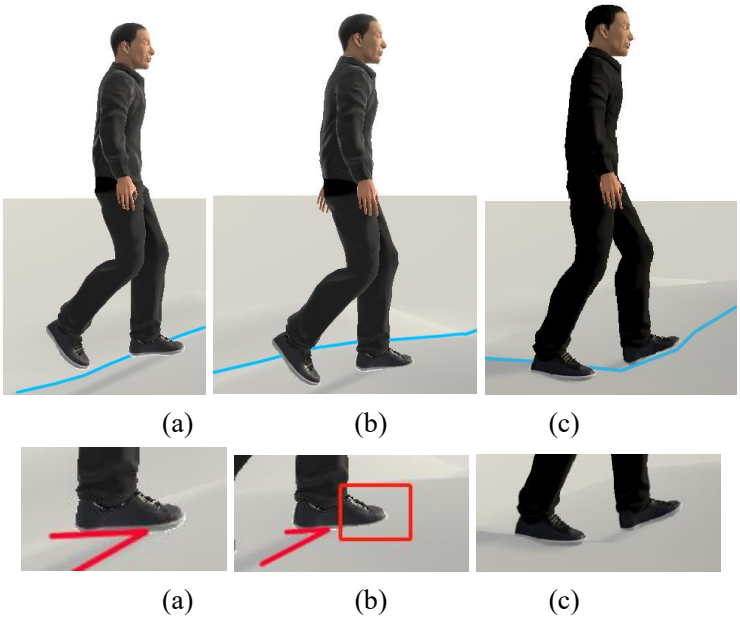


Fig. 5. Results of different methods on the bumpy ground:(a) ERD network (b) Network without physical model (c) This method

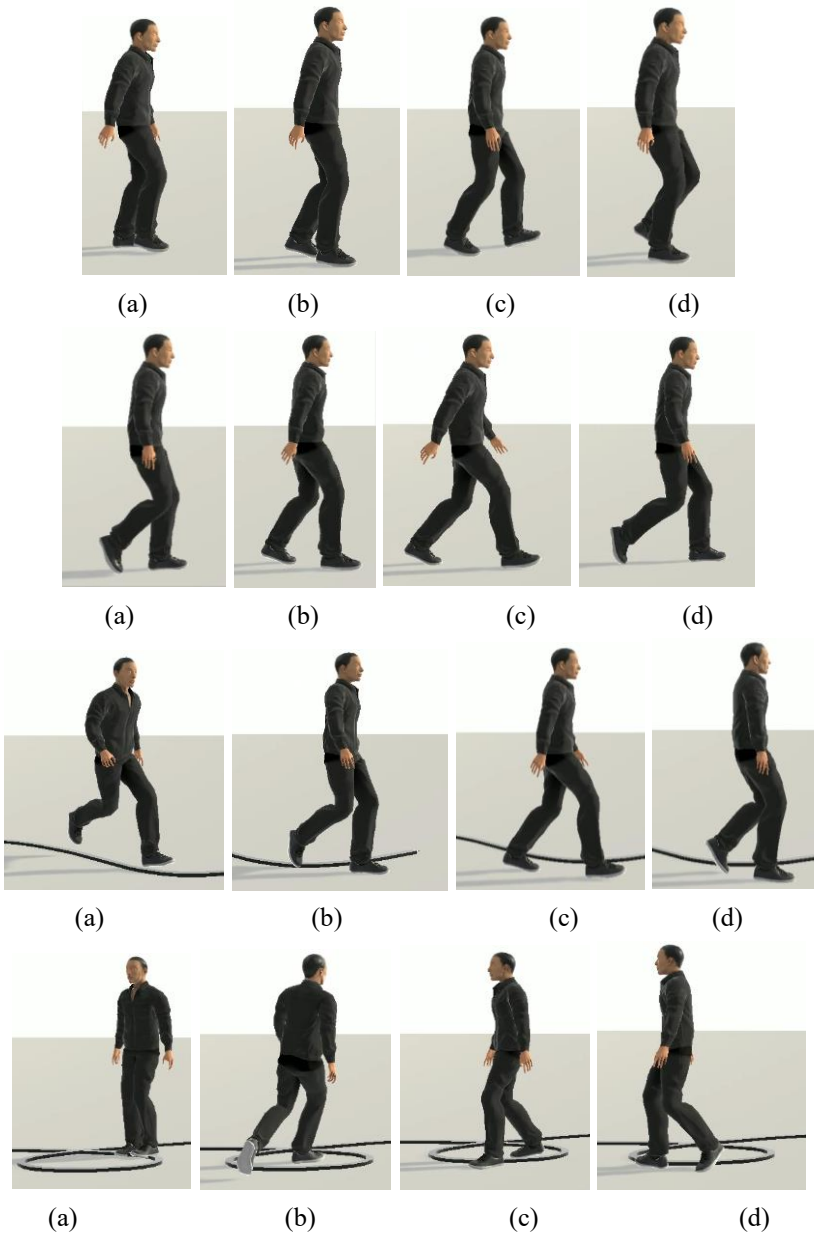


Fig. 6. Results of starting, walking, turning slowly, turning sharply in different methods:(a) Gaussian process (b) ERD network (c) Network without physical model (d) This method

From Table 1, it can be seen that the model error of the autoregressive Gaussian process is the largest among all the models that generated the action. Because the Gaussian process cannot learn data, the results of the Gaussian process are often overfitted to

some data. When the input data is data that is not on the input data set, normal motion cannot be generated, as shown in Figure 6 (a) in the case of sharp turns. And if the model of the autoregressive Gaussian process is to adapt to complex situations, it must increase the storage space. And because the growth of space is too large, it has far exceeded the method of animation fusion, so in the end we only tested the straight and curved gait.

The ERD network can actively learn the motion characteristics that represent the cycle of footsteps according to the data, so when it is in the walking state, the ERD network often performs well. However, as shown in Figure 6 (b), when the character is in the process from stationary to moving, the performance of the ERD network is not good. The ERD network model does not know which leg should be moved at this moment, so it will cause the two legs to float in place. Similarly, in place where the height difference is large, as shown in Figure 5 (a), the joint state of some frames in the middle will be confused with peaceful walking movements, causing the feet to float seriously. Comparing the method in this chapter with the method without a physical model, the difference between the two Error1 is not large. Observing Error2, the method in this chapter is far superior to the method without a physical model. Comparing groups (b) and (c) of Figure 5, it can be found that when there is a height difference, the method in this chapter can basically fit the feet to the ground, and the method without a physical model will still appear. The experimental results are similar to the phenomenon of wearing the mold and floating the feet. This is because the method in this chapter greatly reduces unnecessary joints. Only the joints with obvious features are used as input and output. The network is more targeted and the learning can be more targeted. However, it is difficult to reconstruct the motion of individual foot joints by learning the whole-body joints without a physical model network. The reason for the error1 method in this chapter is slightly higher than the method without a physical model because the physical model is not very stable, especially when the angle is relatively small, it is prone to jitter.

7 Conclusion

We established a real-time gait motion control method by combining neural network and physical IK, and successfully generated gait movements. Compared with other common models, it also proves the advantages of the method in this chapter. As a real-time model, the real-time operation efficiency and model sizes advantages of this chapter are also great. It is a gait motion control model with high stability that can generate effective real-time actions.

However, there are still some problems because the physical model is unstable (this is also related to the physical model of the game engine itself), so joint jitter still occurs sometimes. Due to time issues, we have not solved this problem well, and this is the work that will be carried out in the future. On the other hand, due to the definition of the physical model, many action styles cannot be restored. Whether there are other better physical model definition methods to ensure the stylized retention of character movements is also a problem to be solved in the future.

References

1. Nicholas R. Howe Michael E. Leventon William T. Freeman.: Bayesian Reconstruction of 3D Human Motion from Single-Camera Video. Advances in Neural Information Processing Systems, 820-826(1999).
2. Safonova, A. , Hodgins, J. K. , & Pollard, N. S. :Synthesizing physically realistic human motion in low-dimensional, behavior-specific spaces. ACM Transactions on Graphics, 23(3), 514-521(2004).
3. [Chai J , Hodgins J K : Performance animation from low-dimensional control signals ACM Press ACM SIGGRAPH 2005:686(2005)
4. Tautges, Jochen , et al.:Motion Reconstruction Using Sparse Accelerometer Data, ACM Transactions on Graphics 30.3: p.18.1-18.12. (2011)
5. Rose, C. , M. F. Cohen , and B. Bodenheimer .:Verbs and adverbs: multidimensional motion interpolation. IEEE Computer Graphics & Applications 18.5:32-40. (1998)
6. Iii, Charles F. Rose , P. P. J. Sloan , and M. F. Cohen .:Artist-Directed Inverse-Kinematics Using Radial Basis Function Interpolation. Computer Graphics Forum 20.3:239-250. (2001)
7. Kovar L, Gleicher M.: Automated extraction and parameterization of motions in large data sets, Acm Siggraph. ACM. (2004)
8. Mukai, Tomohiko , and S. Kuriyama .:Geostatistical motion interpolation, Acm Transactions on Graphics 24.3:1062-1070. (2005)
9. Geijtenbeek, Thomas , M. V. D. Panne , and A. F. V. D. Stappen .:Flexible Muscle-Based Locomotion for Bipedal Creatures. ACM Transactions on Graphics 32.6CD:206.1-206.11. (2013)
10. Holden D , Saito J , Komura T .: A Deep Learning Framework for Character Motion Synthesis and Editing. ACM,(2016)
11. Holden D , Saito J , Komura T . Learning Inverse Rig Mappings by Nonlinear Regression[J]. IEEE Transactions on Visualization and Computer Graphics, PP(99):1-1(2016)
12. Graham W Taylor and Geoffrey E Hinton.: Factored conditional restricted Boltzmann machines for modeling motion style. In Proc. ICML. ACM, 1025–1032(2009)
13. Fragkiadaki K , Levine S , Felsen P , et al.: Recurrent Network Models for Human Dynamics. (2015)
14. CMU. Carnegie-Mellon Mocap Database. http://mocap.cs.cmu.edu/.
15. Yamane K , Nakamura Y . :Synergetic CG choreography through constraining and deconstraining at will. IEEE International Conference on Robotics & Automation. (2002)
16. Xie L , Pan W , Tang C , et al. :A pyramidal deep learning architecture for human action recognition[J]. International Journal of Modelling Identification & Control, 21(2):139-146(2014)
17. Tolani, Deepak , A. Goswami , and N. I. Badler .:Real-Time Inverse Kinematics Techniques for Anthropomorphic Limbs. Graphical Models 62.5:353-388 (2000)
18. Tsai, Yao Yang , et al. :Real-Time Physics-Based 3D Biped Character Animation Using an Inverted Pendulum Model. IEEE Trans Vis Comput Graph 16.2:325-337. (2010)
19. Grassia, and F. Sebastian . "Practical Parameterization of Rotations Using the Exponential Map." Journal of Graphics Tools 3.3:29-48 (1998)
20. Kingma D P , Ba J .: Adam: A Method for Stochastic Optimization. Computer ence (2014)
21. Srivastava N , Hinton G , Krizhevsky A , et al.: Dropout: A Simple Way to Prevent Neural Networks from Overfitting[J]. Journal of Machine Learning Research, 15(1):1929-1958(2014)

A Visualization Analysis of the Paper of Chinese Computational Thinking of China in the Recent Decade[1]

Haomin Song, Yu Liu, Lihua Ding, and Xinghua Sun[2]

School of Information Science and Engineering,Hebei North University,07500 Zhangjiakou, Hebei, China

562423588@qq.com,sunxinghua@189.cn

Abstract. In this paper, the Chinese papers of computing thinking published in 2009-2019 are sorted out and analyzed, and the key words co-occurrence map and time zone map are drawn by using the visual map tools VOSviewer, CiteSpace, Gephi and other tools, and the current situation and development trend of computing thinking research in China are analyzed according to the generated charts. The research shows that the research of computational thinking of china are mainly focuses on education reform, program design and information technology courses of college students. There are few researches on the education of Computational Thinking of K12, and many teachers of computer science also lack professional knowledge of computational thinking. The conclusion of this paper provides some reference for the future research of Computational Thinking and the construction of Computational Thinking Curriculum.

Keywords: Computational Thinking, Ability Development, Visual Analysis, Atlas, Educational Reform.

1 Introduction

In 2017, the State Council of China released the "New Generation Artificial Intelligence Development Plan" and artificial intelligence has become a national development strategy. Artificial intelligence will become the biggest opportunity in the future world, and the development of related industries will bring more new jobs and new career opportunities. The plan clearly puts forward: "Set up artificial intelligence-related courses at the elementary and middle school stage, and gradually promote programming education. I think that the development of artificial intelligence in the future will carry out talent upgrade reserves." China is still facing a shortage of talents. Therefore, it is imperative to strengthen the cultivation of computational thinking ability, arouse people's attention to computational thinking, and promote the widespread application of computational thinking, so that computational thinking becomes everyone's basic ability. On January 16, 2018, the Ministry of Education issued the "General High School Information Technology Curriculum Standards" (2017 Edition) (hereinafter referred to as the new curriculum standard). In the new curriculum standard, computational thinking is included in the field of information technology teaching for the first time

[1] This research is supported by grants from Hebei Medical Science Research Project (20200488) and The 13th Five Year Plan of Educational Science Research in Zhangjiakou (193113).

[2] Correspondence author

as one of the four core literacy, which indicates that the cultivation of computational thinking will occupy an important place in the teaching schedule.

Programming education and the cultivation of computational thinking ability have become the core of information technology courses. Although there are extensive interests and concerns at home and abroad, how to integrate computational thinking into education still faces many unresolved problems and challenges.Using the visual software system analysis can study Chinese computational thinking. There are few visual map analyses on the development of computational thinking ability, so this research has certain guiding value for the cultivation of computational thinking ability in the future.

2 Data sources and processing

Firstly, select the journals related to "cultivation of computational thinking ability" related journals published by "China National Knowledge Infrastructure" as the research sample of this topic. The selection method is: use the CNKI advanced search function, the category source is set to "Journal", the content search condition is set to "theme", the keyword "cultivation of computational thinking ability" is entered and the unlocking date is from 2009 to 2019. A total of 2154 articles were retrieved. The search results were transferred in EndNote, Refworks and Excel formats. Then use VOSviewer, CiteSpace and Gephi data to convert to CHKI format and use VOSviewer, CiteSpace and Gephi for analysis.

3 Results Analysis

3.1 Analysis of VOSviewer Visual Atlas: Research on the Cultivation of Computational Thinking Ability

VOSviewer is a visual software that can draw maps of various fields, which has significant advantages in mapping, clustering technology and so on. VOSviewer mainly analyzes co-occurrence maps such as keywords, authors, and journals, and displays them in three graphics: network visualization graph, overlay visualization graph, density view. The clustering relationship between nodes[1] is mainly aggregated by distance and density in the graph.

Frequency of keywords occurrence

After removing some advertisements, meeting reports and other items from the selected literature, import VOSviewer and select co-occurrence keywords, the keywords are set to 5, a total of 193 keywords, the threshold is set to 193, and the first 15 keywords obtained as shown in Fig. 1.. From 2009-2019, the most frequently appeared in the field of computational thinking was "computational thinking" 1870 times, with a total connection strength of 1110; followed by "teaching reform" 269 times, with a total connection strength 554; "program design" 133 times, with a total connection strength 299; "University Computer Basics" 140 times, with a total connection strength 287; "Teaching mode" 130 times, with a total connection strength 281; "Computer Basics" 127 times, with a total connection strength 261; "Information Technology" 89 times, with a total connection strength 193; "Computer Basic Teaching" 98 times, with a total connection strength 193; "Ability training" 77 times, with a total connection strength 174; "Computational thinking ability" 83 times, with a total connection strength 148; "teaching method" 59 times, with a total connection strength 140; "university computer" 67 times, with a total connection strength 138; "teaching" 50 times, with a total connection strength 124; "computer" 55 times, with a total connection strength 117; "cultivation" 48 times, with a total connection strength 115.

The high-frequency word "Computational Thinking". In March 2006, Professor Jeannette M. Wing published the authoritative academic paper Computational Thinking in Communications of the ACM, an American computer authoritative journal. He pointed out that computational thinking is the basic concept of applying computer science. Views on a series of thinking activities covering the breadth of computer science, such as problem solving, system design, and human behavior understanding. These views help people to understand the nature of computing and the core ideas of computer problem solving, so that computational thinking begins to become a research hotspot.

The high-frequency word "Teaching Reform". Cultivating students' computer thinking should begin with education reform. Not only should the traditional computer basic teaching be changed, but also primary and secondary school students should be proficient in computer operation skills and be able to solve practical problems based on computer thinking. The effect of this will lay a solid foundation for the university and the future training of innovative thinking.

The high-frequency word "Programming". To cultivate learners' creative thinking, programming courses are an indispensable part. The focus of teaching is not programming itself, and special attention needs to be given to the method of programming, so that learners can experience the method and thinking of solving computer problems in person， and enhances their learning interest and ability to innovate thinking[2]. It can be seen that the emergence of computational thinking, improving the progress of teaching reform, and promoting the development of education are an essential part of learning.

Verify selected keywords

Selected	Keyword	Occurrences	Total link strength
☑	计算思维	1870	3173
☑	教学改革	269	554
☑	程序设计	133	299
☑	大学计算机基础	140	287
☑	教学模式	130	281
☑	计算机基础	127	261
☑	信息技术	89	193
☑	计算机基础教学	98	187
☑	能力培养	77	174
☑	计算思维能力	83	148
☑	教学方法	59	140
☑	大学计算机	67	138
☑	教学	50	124
☑	计算机	55	117
☑	培养	48	115
☑	课程改革	51	107

Fig. 1. Top 15 keywords in the field of computational thinking ability cultivation

Visualization of keywords

After the selected keywords are extracted and processed in the background, cluster analysis is performed by VOSviewer to generate a visual map as shown in Fig. 2.. In the figure, each circle and the label above

represent a keyword, and the size of the circle and the label represents its appearance. The larger the circle, the more important it is. The size of a node depends entirely on the importance of each keyword. The more important it is, the larger the nodes and fonts. The denser the lines around the nodes, the closer the relationship with other nodes. That is, the greater the degree of relevance. The connection between keywords is also important. The smaller the distance between connected points, the more the number of citations, and the closer the keywords related to closeness are to each other on the graph. Different colors represent different clustering.

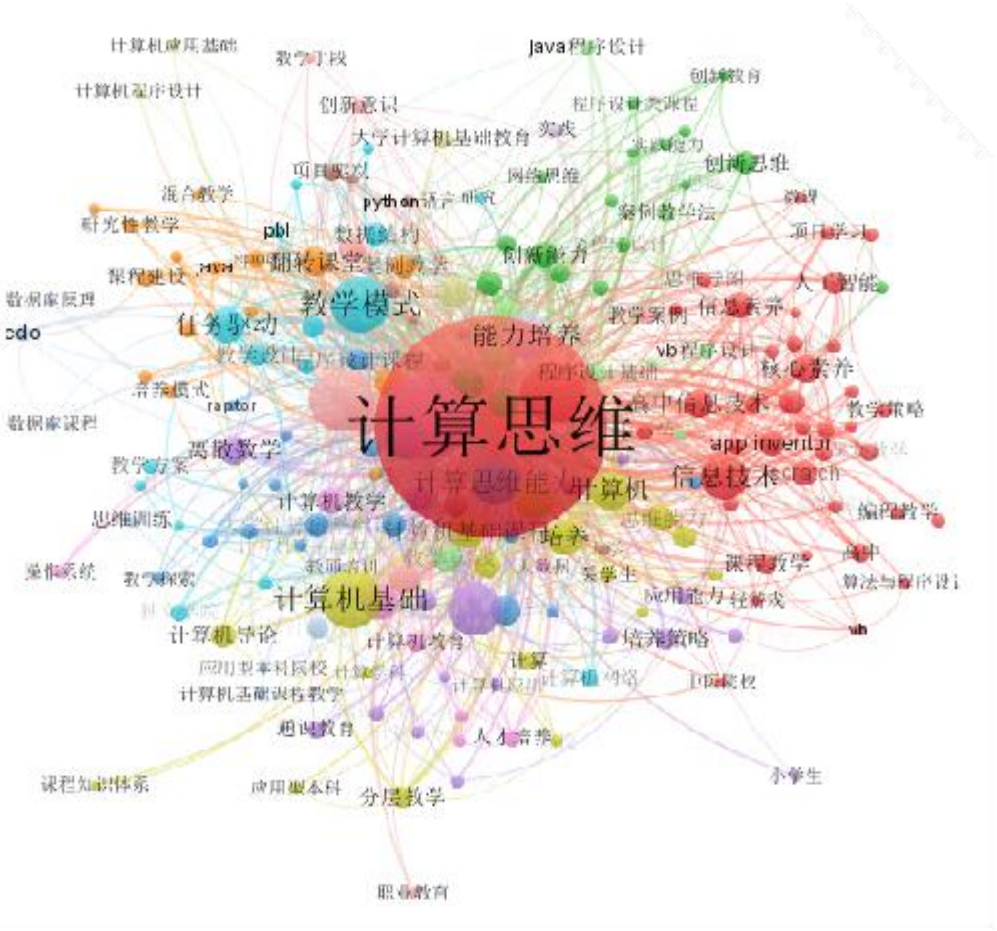


Fig. 2. Visualization of key words for cultivation thinking ability

From the visual map analysis of the key words of computational thinking ability training, we can know that the most frequent keywords are "Computational Thinking", then "Teaching Reform", "Program Design", "University Computer Foundation", "Teaching Mode", "Computer Basics", "Information Technology", "Computer Basic Teaching", "Ability Training", "Computational Thinking Ability", "Teaching Methods", "University Computer", "Teaching", "Computer", "Cultivation". By It can be seen that computational thinking is closely related to education reform, education, and education methods. The field of computational thinking is still a research area that has attracted much attention, and a large number of scholars have contributed to this field. Computational thinking can improve students' ability to thinking and solve problems. The ability to solve problems can also cultivate innovative ability and practical ability. At the same time, with the continuous improvement of education reform hotspots, scholars' research interest in "education reform" has also continued to rise. Computational thinking is also an important part of education reform. Thinking skills have become a trend in education reform.

Incorporating computational thinking into practical teaching models, helping learners to gradually build computational thinking abilities; cultivating their computational thinking abilities; promoting the cultivation of their computational thinking abilities, and improving learners' innovative abilities. Through programming methods and data processing models, let learners produce realistic works such as websites or dynamic videos, from tool users to creators, and truly use computational thinking to solve practical problems. Graphical programming tools are compared to traditional text programming tools that it can improve students' computational thinking ability and provide practical guidance for the training of computational thinking ability[3].

3.2 Analysis of CiteSpace Time Area View: Research on the Development Trend of Computational Thinking

CiteSpace is a citation network visualization software that can draw keyword maps, time zone views, and co-citation maps, paying more attention to tree maps, connection maps, and others to indicate the strength of the relationship between each topic[4]. Using CiteSpace software to load the calculation file that meets the requirements, set the time slice to 1, divide 2009-2019 into 10 time intervals, set the threshold to 50, select keywords to perform, draw a time zone view, and draw a clearer area of computational thinking ability to the continuous evolution of each high-frequency cluster is shown in Fig. 3..

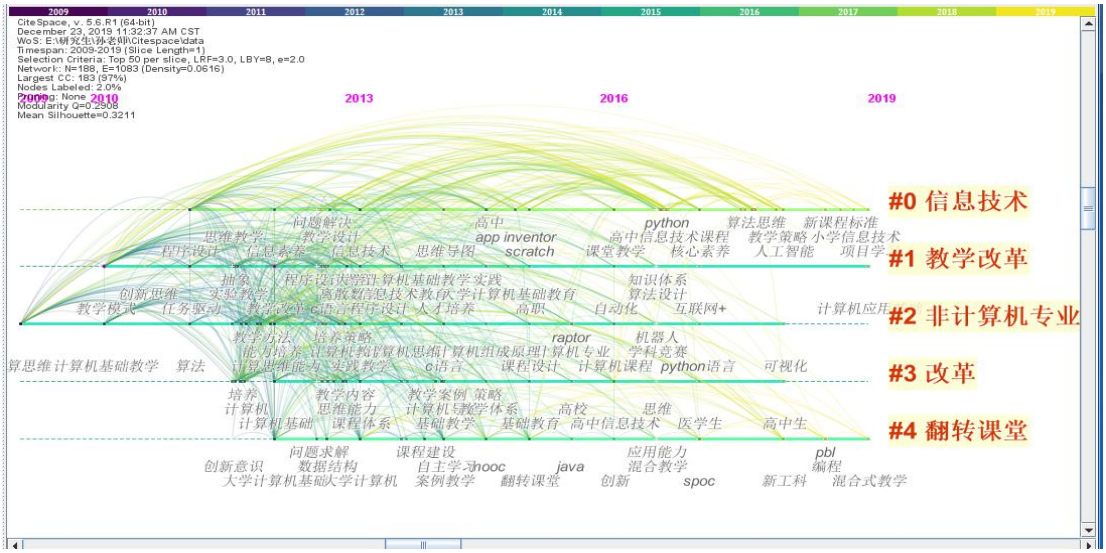


Fig. 3. Map of the time zone in the field of computational thinking capacity analysis

Cluster # 0 Information technology, starting in 2010, it has become increasingly popular in areas such as program design, information literacy and mind maps, and has since appeared in education strategies, instructional design, core literacy, high school, and elementary education. The popularity is continuously increasing. According to the core literacy development learning standards, the effective connection between computer theory and technology in primary and high schools is achieved, and programming tools are used in specific environments or mind maps to provide learners with some learning methods to improve their computational thinking skills.

Cluster # 1 Teaching reform, starting from innovative thinking to educational practice in 2011, and then continuously developing in the direction of Internet + and computer applications, but the development trend continues to decline. It is imperative to change the traditional teaching methods, update teaching modes, and

conduct experimental teaching, and other teaching activities to develop learners' computational thinking skills.

Cluster # 2 Non-computer major. Computational thinking, computer basic teaching, and practical teaching have been widely concerned by scholars since 2009. They have also made relevant research, continued to develop in different directions and other fields. Practical teaching allows students to participate in program design, experience the implementation process and feel computational thinking. In combination with visual programming tools such as Scratch and app inventor in the cluster # 0 information technology. Enhance learners' learning interest and guide learners to apply computational thinking to learning, life, and work.

Cluster # 3 reform, from the International Society for Technology in Education (ISTE) in 2011 proposed the definition of computational thinking operation[5], which caused a lot of research interest, and then introduce to hotspots such as the curriculum system, basic education, medical students, high school students have begun to increase, but the depth of domestic research on computational thinking needs to be improved.

Cluster # 4 Flipping the classroom, in January 2012, Professor Lian Li pointed out that computational thinking is one of the most basic thinking modes of modern science buildings in the article "Computational Thinking: Concepts and Challenges". And that the focus will be on the construction of the teaching system and teaching methods of computational thinking[6]; at the end of 2012, the Ministry of Education took computational thinking as the focus of the "Computer Curriculum Reform Project for College Students" for curriculum reform[7]. It has been developing towards autonomous learning, mooc, flip class, innovation, mixed teaching of applied ability, new engineering, programming, mixed teaching, etc. and has become a new hot topic, and the attention of scholars has been rising. At the same time, in the elementary and secondary education stage, as far as possible, students can experience relevant computer activities to guide and stimulate their interest in learning.

From the time zone map, you can clearly see the correlation between each time periods, sort out the development direction and change trend of the field of computational thinking. Computer thinking, on the one hand, is the reform trend of contemporary education, an important part of education reform. On the other hand, computational thinking can be combined with other fields to solve more problems and open up new perspectives. In teaching practice, attention should be paid to motivating student learners to use computational thinking as a method to solve problems and use it for learning, work, life and other aspects. The education of computational thinking should not just stop at the application of programming tools, but should pay more attention to training learners' programming methods and thinking modes.

3.3 Analysis of Gephi View: Analysis of Publication Year and Journal Distribution of Computational Thinking Literature

Gephi is an open source tool for interactive visualization and detection of various networks, complex systems, dynamic and hierarchical graphs. It is used as exploratory number analysis, link analysis, social network analysis, and biological network analysis. It is a visualization software that depicts large amounts of data with concise points and lines. Its advantages are complex network analysis software based on Java, that is, you can call Gephi's package with Java; you can explore some data relationships, key points, etc; and you can tabulate, draw, and process the data in the diagram and pre-processing of simple data; at the same time it supports DOT language and graphml and GML formats[8].

Create a matrix through Excel and import it into Gephi for analysis. A total of 473 nodes and 396 edges are imported. The average degree of each node is 1.979, the average weighting degree is 4.581, the network diameter is 4, and the average path length is 3.17. As shown in Fig. 4., in the past ten years, the attention of the development of computational thinking ability has shown an increasing trend. The basic education reform that started in 2004 has promoted the development of information technology, but there are still some

deficiencies. Therefore, research on computational thinking cultivation in primary and middle schools has been carried out since 2012, and a small amount of achievements. In 2014, the education sector began to implement the revision of the General Senior High School Curriculum Standards. From 2015 to 2016, computing thinking was identified as the core literacy of information technology, so articles published since 2014 have begun to grow rapidly[9]. New Curriculum Standard in 2017 has taken computational thinking as the core literacy of computer courses[7]. In 2018, the most circulated papers, to a certain extent, showed the research enthusiasm for the cultivation of computational thinking in this field. It reflects the social and academic attention to computational thinking. In the curriculum reform, we should pay attention to the training of teachers' professional literacy, and train teachers in computational thinking to help them master the teaching practice.

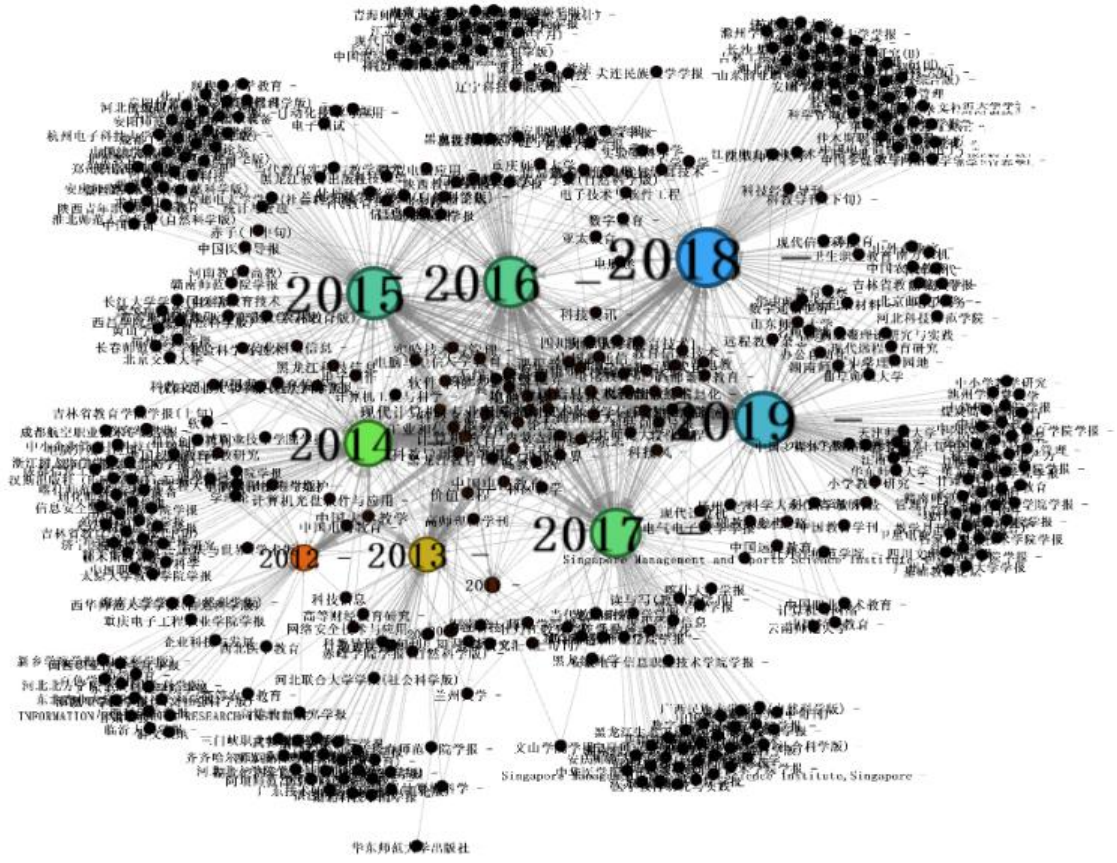


Fig. 4. Analysis on publication year and journal distribution of computational thinking literature

4 Problems and Suggestions on the Cultivation of Computational Thinking Ability in China

4.1 The Problem of the Cultivation of Computational Thinking Ability

At present, the artificial intelligence talent gap in China has reached 5 million, so programming education must be started from childhood. With the rise of artificial intelligence, programming has been pushed down, and the country is promoting the development of the programming education industry. Helping children catch in the future, we must pay attention to programming from the elementary school stage, improve their logical thinking ability, abstract thinking ability, and problem solving ability. Arrange teaching content that is close to life, and let students learn how to organize and algorithmize them. Realize their own creativity, so

that students can really learn something useful for life. In addition, teachers' lack of professional ability, how to effectively develop the ability of computational thinking is also a major problem at present. The most urgent task is to master the relevant professional knowledge and improve the teacher's computational thinking ability. Only if the teacher deeply understands the relevant knowledge, Changing traditional teaching methods, and enables students to participate in teaching experiments, can students understand and master the ability of computational thinking. Moreover, there is a lack of evaluation systems for computational thinking. It is necessary to develop evaluation indicators that are localized and easy to judge.

4.2 Suggestions on the Cultivation of Computational Thinking Ability

Strengthening the focus on the development of computational thinking ability is not limited to traditional information technology teaching, but is based on a mixed teaching model, through reasonable teaching guidance, teaching process design, teaching content, teaching cases, and project-driven[5], mobilize students interest, motivate students to participate, cultivate students' teamwork, and share expression skills. Through interesting content, cheerful forms, effective design, and excellent processes, students can improve their computational thinking ability. At the same time, the urban-rural gap related to information technology is reduced, and the hardware and equipment conditions for computing education are improved[10].

And with the development of information technology, the ability of computational thinking will become our necessary skills and the direction of future development. Continuously summarize teaching practice experience to improve the ability of teachers and students in computational thinking. Of course, the ability of computational thinking is not limited to the computer field. It can realize the cultivation of interdisciplinary, such as agronomy, botany and zoology, geography, bioscience, physics, clinical medicine and other fields.

5 Summary

Computational thinking is a new product in the field of computer science, and it is also a necessary basic skill. Through the methods of "from big to small", "deep into details" and "expanding capabilities", we can decompose, analyze and big problems and difficulties. And solve them, cultivate students' ability to analyze problems, solve problems, teamwork, and innovative thinking from the system, make students experience "technology is fun, technology is useful". Let students have a new, comprehensive and forward-looking understanding of information technology. Enhancing students' interest in computational thinking and improving their learning efficiency. However, China's research on the development of computational thinking ability has not yet matured, and it has little guidance to education. Therefore, scholars in the future should strengthen their research in this area can make up for the shortcomings in this area, and then promote the rapid development of China's computational thinking ability. Of course, the visual analysis software, the data analyzed are mainly published papers, and there are problems such as incomplete analysis. Therefore, it must be combined with the actual situation of teachers and students in China, discuss the training method of computing thinking suitable for China, and promote the rapid development of computing thinking[11].

References

1. Lijun Ji. Analysis of Information Literacy Hotspots at Home and Abroad in 2016-2018 Based on VOSviewer [J] .Contemporary Library, 2019 (03): 23-28.

2. Shuang Ye , Jianhong Ye, Wenxian Jiang. Exploration of Computer General Education Reform Focusing on Cultivating Computational Thinking [j] .Journal of Jiamusi Vocational College, 2018 (07): 414-415.

3. Yi Fu, Bochao Xie, Yafeng Zheng.An Empirical Study of Programming Teaching Based on Graphical Tools to Promote the Development of Computational Thinking in Junior Middle School Students[j] .Electrical Education Research, 2019,40 (04): 122-128.

4. Maohua Gan.Research Hotspots and Trends in the Development of Computational Thinking Ability in China——Visual Analysis Based on CiteSpace [J] .Computer Applications and Software, 2019,36 (06): 1-6 + 42.

5. Peng Chen, Ronghuai Huang, Yue Liang, Jinbao Zhang.How to Cultivate Computational Thinking——Based on the Research Literature of 2006-2016 and the Latest International Conference Papers[j] .Research on Modern Distance Education, 2018 (01): 98-112.

6. Lian Li.Computational Thinking——Concepts and Challenges [j] .China University Teaching, 2012 (01): 7-12.

7. Tao Xue.Analysis of Current Situation of Research on Computational Thinking Education Based on China HowNet Literature [j] .Wireless Internet Technology, 2019,16 (21): 151-153.

8. Yi Jin.Review of the Research on the Relationship between Employment Pressure and Social Adaptation of University Graduates Based on Gephi [J] .Journal of Dali University, 2018,3 (05): 108-114.

9. Jian Yang.Study on the Current Situation of Cultivation of Computational Thinking in Information Technology Teaching in Primary and Middle Schools in China——A Literature Analysis Based on China HowNet [j] .China Education Informationization, 2017 (23): 85-88.

10. Youqun Ren, Fengwei Sui, Feng Li. How is Digital Indigenous Possible? —— Also on the Necessity and Possibility of Computing Thinking Entering Information Technology Education in Primary and Middle Schools [j] .China Educational Technology, 2016 (01): 2-8.

11. Wenxiang Fan, Yichun Zhang, Yi Li. Summarization of Research and Development of Computational Thinking at Home and Abroad [j] .Journal of Distance Education, 2018,36 (02): 3-17.

Painter eMarketplace Platform

Chatchai Suthapakti [1,] Taminee Shinasharkey[2] and Santithorn Boonchua[3]

[1] Assumption University, Samuthaprakarn, Thailand
[2] Assumption University Samuthaprakarn, Thailand
[3] Assumption University, Samuthprakarn, Thailand
c.kiss77@gmail.com
tamineeshn@au.edu
keh@cide.au.edu

Abstract. In Thailand, homeowners have faced the problem of unable to find the skilled painter to paint their home. Unskilled workers could cost some damages with the painting. The e-marketplace system could serve as a hub for customers who needed to find the right qualified painters with reasonable price and painter members would be able to get more jobs and income from this platform.

The platform was similar to e-commerce system, but main products were paint services in various skills e.g. mortar, wood, construction chemical, steel, decorative paint. Core duty of e-marketplace was to search and match demand and supply with matching criteria in service area and painter skills. After painters completed their project, they could add more experience or services in their portfolio and customer could review their quality of work and service.

Keywords: painter platform, ecommerce, eMarketplace

1 Introduction

The Internet based system as known as e-Marketplace has enable automated transactions and collaboration between buyers and sellers. The total volume of e-commerce in the business-to-business or B2B sector was at US$ 2.77 trillion in 2003. The e-commerce market in North America accounted for US$ 1.6 trillion of the US$ 2.77 trillion, follows by Asia-Pacific, which accounted for US$ 0.3 trillion, Europe, US$ 0.8 trillion, Latin America , US$ 0.58 trillion, and Africa and the Middle East together accounting for US$ 0.177 trillion. In Taiwan, the market intelligent center (MIC) reported that the value of e-Marketplace transactions was at US$ 56 million in 2000, US$ 125 million in 2001, US$ 406 million in 2002, and US$ 750 million in 2003. In 2007, 50% of B2B e-commerce in Taiwan had been transacted through e-Marketplaces, compared to an average of 55% of B2B e-commerce worldwide via e-Marketplaces, according to Internet Data Center (IDC) Report [1]. Thailand has also increased in the number of Internet users as well. Figure 1 shows the increasing use of Internet user in Thailand.

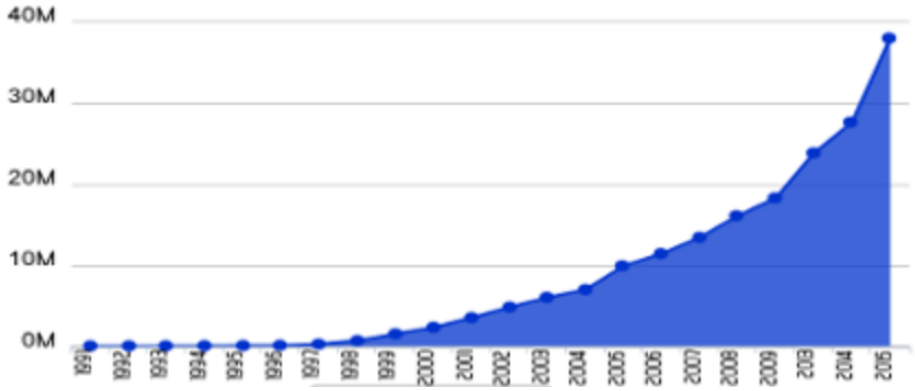


Fig. 1. The trend of internet user in Thailand during 1991 -2016 [2]

The concept of painter e-Marketplace was for painters to update their profile, portfolio, and contact information in order to increase their chance of getting hired. The home or project owner could also find the skilled painters that match their job requirements. This painter e-Marketplace also offered painting DIY forum and community system for sharing knowledge and experience.

This e-Marketplace would help customers to find the qualified painters by searching painters portfolio and reviewing the comments from previous users. Meanwhile, painters could generate more jobs by creating their profile and portfolios via this platform. The significance of the e-Marketplace was traffic and transaction generated by all visitors with great satisfaction in rating and reviewing that would make platform magnetized.

2 Literature Review

The power of Internet of thing (IoT) concept was already deployed in many business and prevalent in day to day activities. It served people to find product and service. Many firms also faced high competition in online area. More people have used Internet to find products and services [3]. Therefore, the painting industry also needed to focus and jump into online competition [4]. The e-Marketplace has two big components, which are platform design and environment support.

For platform design, the responsiveness must be fast. The information must be reliable and accurate. The graphic user interface should be attractive and easy to navigate. The presentation of the page could be listed according to the customer search interest. Therefore, the platform formulation which based on the customer searching record could improve the market share. The eleven sub-components of platform design are Search engine optimization, User friendly interface, Responsive design, User profile and registration process, Search & matching system, Rating & review, Knowledge sharing & Community system, News & event, Content management interface, and Reporting dashboard [5]

For Environment support, Internet was used as a tool to provided e-commerce growth but it could be insecure for customers. Customers were mainly concerned about trust and security. E-vendors have to enhance trusting beliefs and trusting intentions, and strengthen the perception of authentication, nonrepudiation, confidentiality, privacy protection and data integrity to achieve the customer's comfort. E-businesses must provide a trustworthy and secure environment in order to success [6-7].

The platform might not only survive if it did not have any support from backend. E-Marketplace also required the same support. The required components for environmental support include Social network & messenger application channel, Search engine marketing (SEM), Customer support, Content production, and Painter academy.

The main concept of the painter e-commerce was to search and match painters. This function was used by customers to find the providers which met the specification and matched their requirements. Thus, this platform needed to provide several criteria to search base on customer requirements.

For example, in the paint industries, the painter type (mortar, wood, constructed chemical, steel, decorative paint) and service area were the most important keywords to search. The matching system provide sequence by match, potential match and non-match. Customer might have an idea from final complete product that were listed on the site. Customer might get interested and search for the painter who did that job. The platform would provide the tools that could update sample pictures of complete work. The result of search would find from any fields of provider profiles e.g. name-surname, experience, explanation text in portfolio [5,8-10].

Rating & review is also important for users. The feedback and experience from users would be benefit to those interested in using the same service or painters. Otherwise rating and review required to ensure whether that reviews were accurate from customers as shown the sequence in the Figure 2 [11-12].

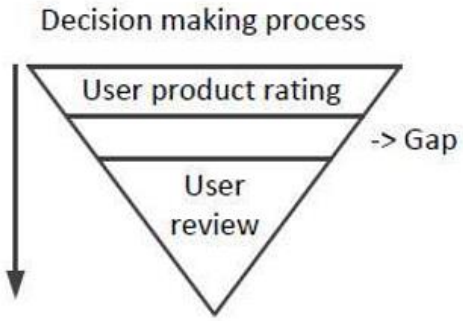


Fig. 2. The decision making process [12]

In Germany, Online shoppers were conducted to examine how consumer reviews and ratings were used to support buying decisions. The survey showed that reviews and ratings were an important source of information for consumers. Whereas the qualitative feedback from the survey indicated that the perceived helpfulness of rating systems varies. Especially the comparison of user reviews was important information to customer due to the unstructured nature of textual user reviews [12].

To prevent rating & review information from malicious people or spam, this function were designed from a painter update their portfolio. After the painter finished updating the portfolio, the painter would require entering the email. The system automatically send rating template to customer via email and response by click URL in email and feedback rating & review with privacy setting to show or hide the reviewer name. Once they submitted the rating & review system will automatically publish that information [13-14].

3 Research Methodology

The framework was based on existing problem for homeowner to find the qualified and right painter. Therefore, the platform was called ecosystem to serve paint service by collect painter profile in platform that provided special promotion that painters may offer. Meanwhile, customers could search and match their requirement with information and knowledge support from community as shown in Fig 2. The template also provided the place for dealer to leave their profile e.g. contact info, location map, available paint product and their offering promotion. Below diagram explains overall of this project.

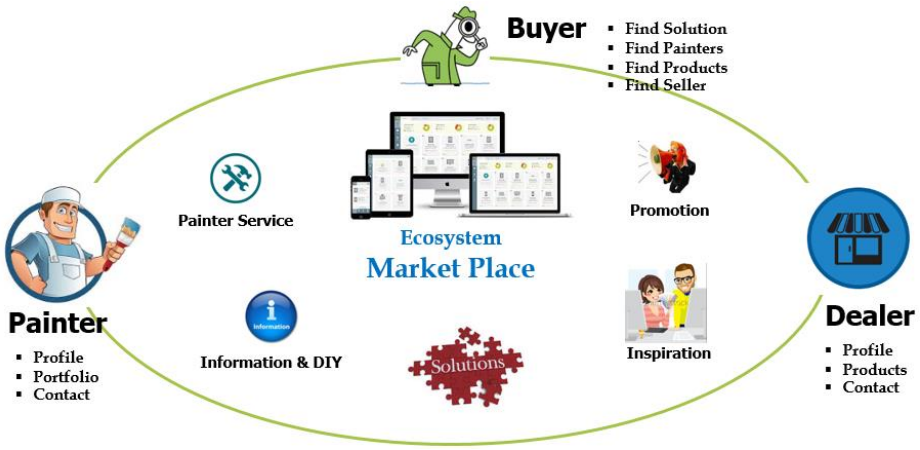


Fig. 3. Conceptual diagram of painter e-marketplace

3.1 Design

The design must be user friendly. Any person without strong computer could be easily used the site. The website needed to have clear appearance such as text size, color and menu position with two languages support in both local language and English.

- **Responsive design:** The website needed to support variety of user's display devices e.g. desktop PC, laptop, smartphone and tablet.
- **Search engine optimization:** The website needed to have option to customize the meta data for search engine especially in google to appear in the best position result of desirable keyword.
- **Social login**: The registration and login process needed to support social login especially face-book.com that's the most user-friendly registration process at this time and easily to store the user profile with reliable data source.
- **Help menu:** This platform needed to provide help menu for user guidance start from registration process and method to use any functions of platform step by step.
- **Frequently ask question (FAQ**) The site also needed to provide help menu that users may have question with. All contents in this help menu had to be available to find by search bar.
- **Information to contact administrator:** Even though platform already provided the good guidance, users may have more enquiries apart from provided information. This function could provide customer support contact information e.g. telephone number, email, social network or chatter inbox.

3.2 Registration and user profile

This platform had three types of users: customer, painter, and dealer. Each of them had difference required information, but each user might possibly be all those three user types. The first access required only basic information and provide Facebook login that might be easy for user not having to fill in the information.

If user signed up without social login, the system needed to re-confirm by sending confirmation link to email address that user was gave to platform and user must click that link to complete it. Registration flowchart can be explained as Figure 4.

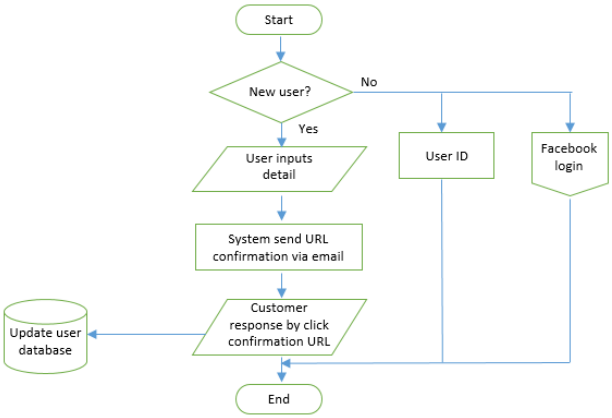


Fig. 4. User registration flow

After that, system would direct user into signing up their account with either customer, painter, or dealer, which would require different information.

3.3 Service Flow

Once a customer needed to find the painter, a customer needed to apply request form. The information would include service address or contact information, type of service, area, scale of area, attach picture, type of building, budget, start date, and other job detail.

Once the requirement of the job was entered, the painter could also view the job detail. Painters would be able to find the matching requirement and skill from customer request. The request would be separated into two types, which customer could request specific painter or request any painter. Jobs display to each painter were difference views by filter from skill. It would start by showing the request nearby area of the painter. Once painter was interested in any job, they could click to accept job and system alert to customer afterward.

To find the accuracy job detail, painter might find from advance search. They selected their criteria which were: area (Province, Amphure, and zip code), job type (new paint, re-paint and fix problem), type of building (home, condo and tower with input range of floor) and budget range in baht.

After painter accepted the requested job, the system would send alert to customer and show the list of painters who accepted their request and customer would check the detail and condition. Customer had to finalize painter selection to match the job. After matching demand and supply, the contact information would appear to each other to contact directly and started work from here. During operation period, if the painter could not complete the job, customer can re-select new painter anytime with provided reason. The comment would appear in unfinished review. Once the job was completed, the painter could update their portfolio, add picture of their work, and adding description. Painter could select to hide or show their profile.

After painter finished updating their portfolio, the detail regarding the job would be sent to customer to review and verify. The system would then lead customer to rating & review page, which customer would be able to rate their experience from this portal with any comment that they wanted to publish and this information would stamp to each painter portfolio. If customer did not submit rating and review, the portfolio information would not be published. In that case, the painter would need to contact customer for submit the rating & review that other customers would be able to see the detail from painter's portfolio.

All process for painter service flowchart can be explained in Figure 5.

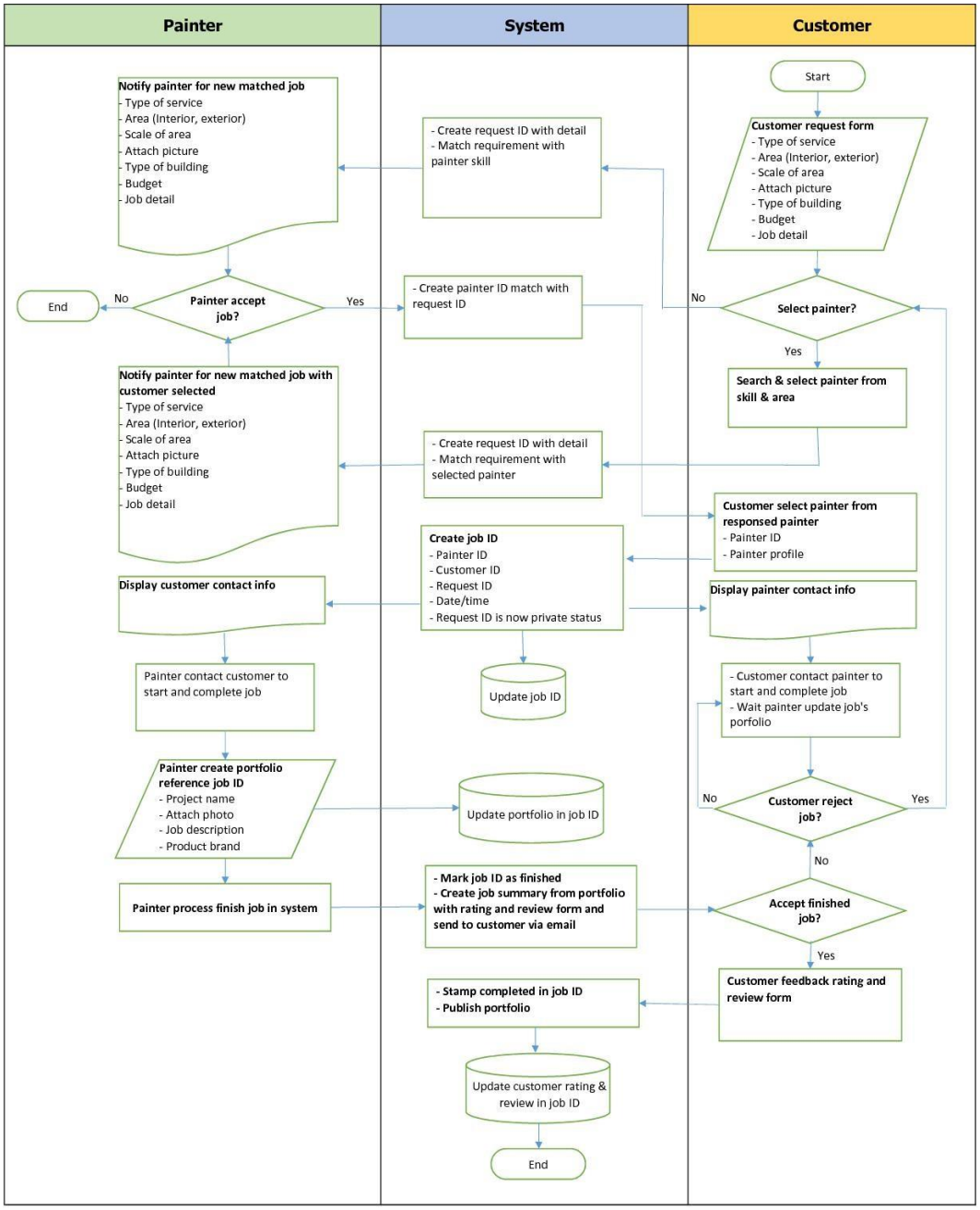


Fig. 5. Request painter service flowchart

4 Result

The implementation of the painter e-commerce started with creating a user profile. The user profile required customer information such as name, last name, email, phone number or they could login via Facebook account to skip filling out required data as shown in Figure 6. After user filled all required detail, the system would send confirmation to email. The customer then needed to response by clicking the confirmation link in the email. After customer clicked confirmation link that means registration process was completed and user needed to key the username and password for the first login as below screen shown in Figure 7.

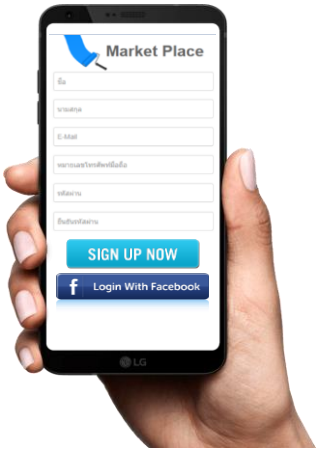


Fig. 6. Require basic user profile to access to system

Fig. 7. User login page

For painter and dealer profile, the layout of painter details was shown in figure 8 and the profile would be published after e-marketplace administrator approved. For the dealer, the layout is shown in figure 9.

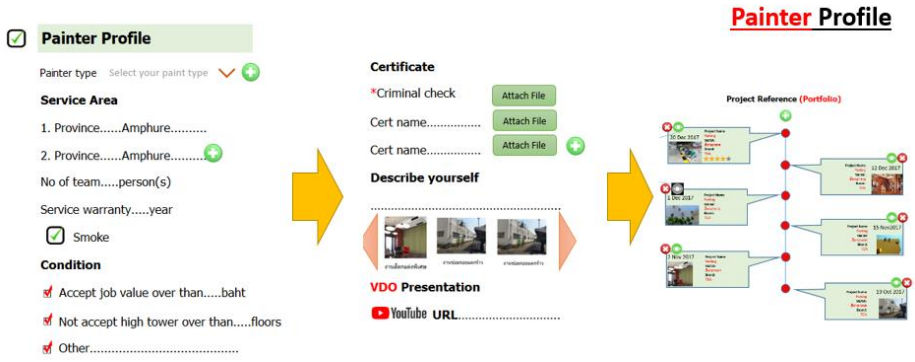


Fig. 8. System required detail for painter

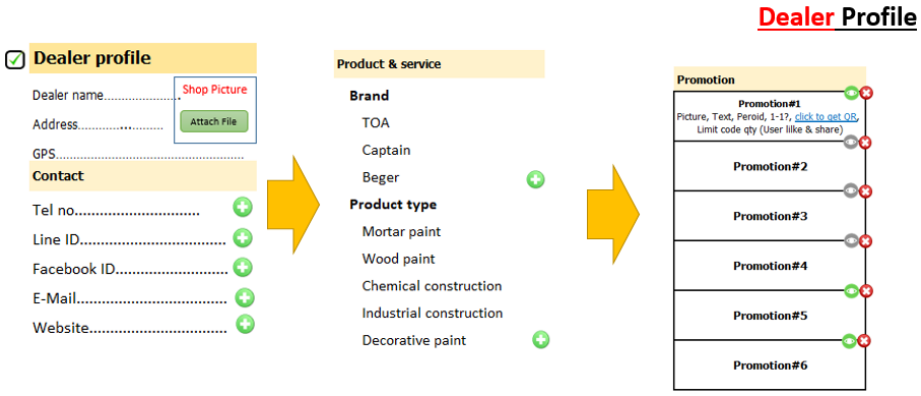


Fig.9. System required detail for dealer

To find a matched painter, customer needed to fill out the requirement for the job, such as painter type (new paint or re-paint), type of area (interior or exterior), how big of the area, job type (wood, mortar, or steel), attach file (for the painter to see the site job), and type of building (home, condo, office) as shown in Figure 10. Once the information was entered, the customer could either chose the painter whose skills has matched the requirement. The customer could review their requirement as shown in Figure 11 as well.

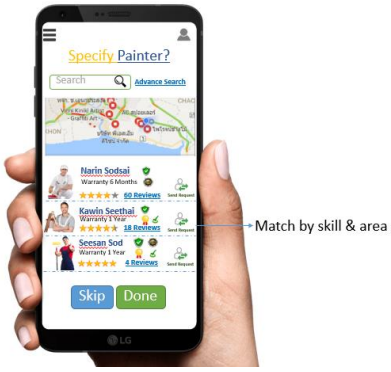


Fig. 60. Customer select desirable painter

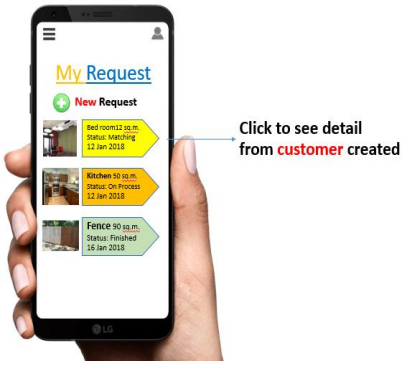


Fig. 11. Customer view their request with status

Painter could response to the request and accept the job as shown in figure 12. Once the painter accepted the request, the customer would be notified and could confirm that the painter has been hired for the job as shown in figure 13.

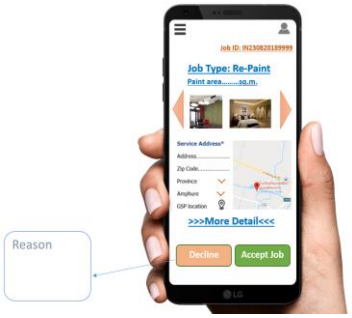


Fig. 12. Display for painter response to customer who specified painter

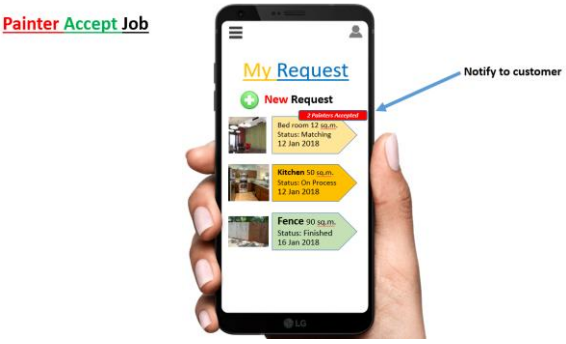


Fig. 13. Notification display from painter accept the customer request

Once their agreed upon the contract, both sides would receive contact information of each other. After completing the job, the painter could add their experience in their portfolio and the customer could review and rate the service of the painter in the system.

5 Conclusion

This chapter concludes the result of painter e-marketplace system to aid the painting industry. It would be useful for user to search and match the right qualified painters. Furthermore, the platform also has a benefit for painters to find jobs and keep their own service experience with rating and review to improve themselves. Moreover, the

platform provided dealer space to promote their shop and generate their own promotions which have benefit for both dealer and customer, The dealer could increase their sales volume from the promotions and customer or painter could get benefit from those promotions.

To improve this platform more effective, the platform managers might need to make sure that the system is reliable. Customer support team should be ready. The site could provide painter academy to improve painter skill and data analysis to improve platform.

In this platform, one of key success was the skill of painters. If the platform running well, management team would need to consider building the painter academy system by gather the knowledge in painting industry and categorize them into digital platform. Thus, the painter development would be easy to search and learn anywhere with comprehensive test to ensure they have well understood the content and that would bring more user to use the painter e-marketplace.

6 References

1. Yu, C.S.: What Drives Enterprises to Trading via B2B E-marketplaces?, Journal of Electronic Commerce Research. Vol.8, No.1, pp. 84-100 (2017).
2. National Electronics and Computer Technology Center (NECTEC) Homepage, "Domain under Thailand", http://internet.nectec.or.th/webstats/home.iir. last accessed 2020/3/5.
3. Internet World Stat Homepage, "Internet growth statistics", http://www.internetworldstats.com/emarketing.html, last accessed 2020/4/13.
4. Shinn, A., Nakatani, K., Rodriguez, W.: Analyzing the role of the internet-of-things in business and technologically-smart cities, International Journal of Internet of Things, Vol. 6, No. 4, pp.149-158 (2017).
5. Dinerstein, M., Einav, L., Levin, J., Sundaresan N.: Consumer price search and platform design in internet commerce, Discussion Papers from Stanford Institute for Economic Policy Research, No. 13-038 (2014).
6. Maniam, B., Naranjo, L., Subramaniam, G.: E-commerce best practices: how to achieve an environment of trust and security, International Journal of Innovation, Management and Technology, Vol.3, No.4, pp.397-401, (2012).
7. Tiwari, S., Singh, P.: E-commerce: prospect or threat for environment, International Journal of Environmental Science and Development, Vol.2, No.3, pp.211-217, (2011).
8. Spink, A., Jansen, B.J., Trends in searching for commerce related information on web search engines, Journal of Electronic Commerce Research, Vol. 9, No. 2, pp. 154-161, (2008).
9. Pu, P.,Chen, L., Kumar P.: Evaluating product search and recommender systems for e-commerce environments, Electronic Commerce Research, Vol.8, pp.1-27, (2008).
10. Barnes, D., Hinton, M.: Searching for e-business performance measurement systems, Information Systems Evaluation, Vol.10, Issue1, pp.1-8, (2007).
11. Hajli, M. N.: A study of the impact of social media on consumers, International Journal of Market Research, Vol.56 Issue3, pp.387-404, (2014).
12. Lackermair1, G., Kailer, D., Kanmaz, K.: Importance of online product reviews from a consumer's perspective, Advances in Economics and Business, Vol.1, pp.1-5, (2013).

13. Ashby, N.J.S., Walasek, L., Glöckner, A.: The effect of consumer ratings and attentional allocation on product valuation, Judgment and Decision Making, Vol.10, No.2, pp. 172-184, (2015).
14. Mo, Z., Li, Y.F., Fan, P.: Effect of online reviews on consumer purchase behavior, Journal of Service Science and Management, Vol.8, pp.419-424, (2015).

Acoustic Emission Signal Analysis for Stamping Machine Condition Monitoring and Fault Diagnosis

Hsiao-Yu Wang[1], Yu-Hung Chiang[2], Chih-Yuan Chen[3] and Yu-Shiang Hon[4]

[1]Central Industry Research & Service Division, Institute for Information Industry, Nantou, 540, Taiwan
shonwang@iii.org.tw

[2]Central Industry Research & Service Division, Institute for Information Industry, Nantou, 540, Taiwan
yuhungchiang@iii.org.tw

[3]Central Industry Research & Service Division, Institute for Information Industry, Nantou, 540, Taiwan
hennessychen@iii.org.tw

[4]Central Industry Research & Service Division, Institute for Information Industry, Nantou, 540, Taiwan
garhung@iii.org.tw

Abstract. A quality inspection is often carried out after the stamping processing for a relay contactor. Once a failure occurs in the process, it may be difficult to detect the defective products due to the fast stamping speed. Previous studies employed the acoustic emission (AE) technique to detect tool wear, to ensure the stability and reliability of the machining process. Moreover, there haven't been many studies on using AE signals to predict the state of the machine and product quality. This study collects AE signals, and manually classifies them into 5 types; abnormal sounds (slug marks), abnormal sounds (raised burr), abnormal sounds (buckling), and environmental sounds. Then, an anomaly sound classifier is trained using Hidden Markov model (HMM) algorithm to predict abnormal behaviors of the stamping machine. The experiment results show that the classification accuracy for normal sounds, abnormal sounds (raised burr), abnormal sounds (buckling) and environmental sounds are all over 91.66% except for abnormal sounds (raised burr). It means that the proposed classifier is able to precisely identify AE signals to help engineers immediately determine whether a stamping machine is in a normal or abnormal state in most cases.

Keywords: Acoustic Emission, Hidden Markov Model, Press Processing, Relay Terminal

1 Introduction

Nowadays, the sensing technology has been emphasized as one of the key technologies for the new trend of manufacturing industries such as IoT (Internet of Things) and Industry 4.0 [11]. How to efficiently monitor the quality and detect the abnormal behavior for are lay contactor during the stamping process has been an important issue for engineers. The abnormal behaviors of stamping machines are closely associated with the quality of the relay contactor [6].In practice, quality inspection is performed in the middle of production lines or at the end of the production stage where engineers visually inspect the quality of the relay terminals. However, due to the large number of relay terminals, it may be difficult for engineers to detect defective products accurately.
To resolve this issue previous studies had applied motion control, blank holding force control, prediction of defect occurrence by simulation, quality evaluation, quality prediction of products, and optimization of forming conditions, to find out the quality of products. However, these methods were post analysis methods. They are not able to report the machine status in real-time [2]. Therefore, the acoustic emission (AE) monitoring methods, including ultrasonic waves and mold deformation/forming load, are proposed to monitor the forming state during in-process [4]. With AE signals, engineers can diagnose whether the machine behavior is abnormal in real-time. AE signals are transient stress waves generated by the rapid release of energy in the material. They are related to abnormal behaviors of press machines, and can be measured using relatively simple instrumentation. Moreover, AE sensors can often be mounted innocuously on machines.

Skåre et al. [7] studied the proportionality between the AE signal during stamping and the tangential force in order to understand galling, cracking and tool wear. Shanbhag et al. [5] used AE sensors to learn the features of the signal that can be used to identify unworn and worn parts and it was shown that AE RMS, peak, kurtosis, and skewness could be used to identify unworn and worn parts and that the “mean frequency” feature may be used to identify wear initiation. Ubhayaratne et al. [9] developed an approach to tool wear monitoring in sheet metal stamping using an audio signal processing technique. A semi-blind signal extraction technique was used to preprocess and denoise the audio signals collected from stamping operations. Seemuang et al. [4] developed a tool condition monitoring system that predicts tool wear by measuring audio signals generated by machine spindles. A low-cost microphone was used to collect the audio signals in drilling. Kothuru et al. [3] developed a tool wear monitoring approach that can classify tool wear conditions using support vector machines. Experimental results have shown that the prediction accuracy is above 90%. Wang et al [10] used a Hidden Markov Model (HMM) approach for detecting tool wear condition in a turning process. Ertunc et al [1], applied an industrial drilling process, and used the cutting force and torque signals; and HMM derived to model different wear conditions, namely “sharp”, “workable”, and “dull”. The results showed HMM with the best probability score was selected as the classifier of tool wear condition.

Most of the existing methods have used with AE signals to monitor galling, cracking, and tool wear for different press processing, yet only few studies directly predicted the state of stamping machine using AE signals. To be more precise, there isn't sufficient relevant research on analyzing different AE signals to predict the state of stamping machine for relay contactors. In addition, according to the literature review, this study suggests that with AE signals, the HMM could have favorable predictability. Therefore, a supervised method is proposed in this study which constructs anomaly sound classifier using HMM for various types of AE signals. There are three main purposes in this paper:

1. AE signals are captured from the relay terminals, and the signals are manually classified into five types (normal, abnormal sound - raised burr, abnormal sound - slug marks, abnormal sound - buckling, and environmental signals).
2. The framing for the AE signals, and then Mel-Frequency Cepstral Coefficients (MFCC) is applied to feature extraction.
3. The HMM anomaly sound classifier performance based on AE signals is examined to predict the state of the stamping machine.

2 The Sensors and the Data Acquisition System

This study aims to predict whether a machine behavior would be abnormal using AE signals from the stamping machine with the relay terminal. First, through the sensor, the real-time AE signals from the stamping machine are collected. And the storage and exchange formats of the collected data must be consistent. Then, all the collected AE signals are manually labeled into different types. Then, feature selection is performed using MFCC, and an anomaly sound classifier is built by the HMM algorithm.

2.1 Collecting and Capturing AE signals

Regarding data collection and extraction, this study first performs sensor optimization and signal extraction calibration development for stamping dies and parts and collects the AE signals from the stamping machine. The thickness of the relay terminal used for this experiment is 0.2mm and 0.5mm. AE signals generated with relay contractors of different thicknesses are collected as well as the cycle time from each stamping machine. Moreover, because microphones used to collect AE signals available in the market cannot collect the AE signals of high and low frequencies as required by this study, we develop two sets of microphones to collect AE signals of low frequencies (20Hz ~ 15kHz) and high ones (10kHz ~ 1000kHz). The hardware sensors for collecting AE signals are shown in Table 1-4.

Table 1. Microphone Sensor (Low Frequency)

Type	Back Electret Type Condenser Microphone
Frequency	20~15000Hz
Open Circuit Sensitivity	-34Db (0dB=1V/Pa,1kHz)
Output impedance	1.4kΩ

Table 2. Ultrasonic Sensor (High Frequency)

Type	Hyper cardioid
Frequency	10k~1000kHz
Open Circuit Sensitivity	>65dB
Operating temperature	-20℃~50℃

Table 3.Data Acquisition(DAQ)

Type	Specification
Channels	3
SignalRanges	±30 V
SampleRate	102.4kS/s/ch
IsolationContinuous	60 VDCCh-Earth
Connectivity	ScrewTerminal ,BNC

Table 4.Industrial Personal Computer (IPC)

Type	Specification
CPU	Intel® Celeron® J1900 Quad Core 2.0GHz
Memory	4GB (1 x 4GB) DDR3
Hard Disk Drive	32G (SSD) +550G
Operating System	Microsoft® Windows® 10

Moreover, microphone sensors for low frequency and ultrasonic sensors for high frequency are installed at the front and the side of the relay terminal separately. Figure 1 shows the installation position of the hardware sensors on the relay terminal.

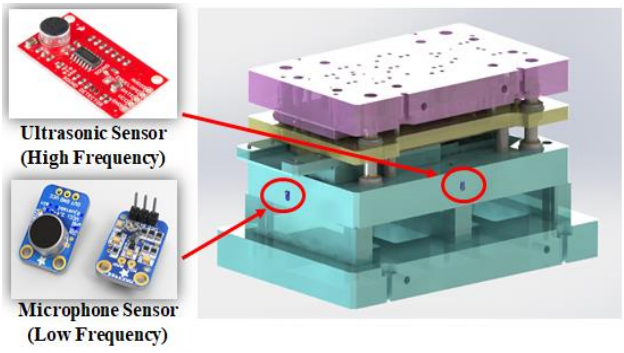


Figure 1.The installation position of the hardware sensors on the relay terminal

2.2 AE Signal Processing and Labeling

The original AE signals collected through the sensor are filtered through Zoom UAC 8, a professional audio signal processing software package, and transformed from analog signals to digital signals (A/D). Cool Edit Pro, a commercial software package, is used to create spectrograms with the original data. In addition, all AE signal are manual labeled into normal sound, abnormal sound (raised burr), abnormal sound (slug marks), abnormal sound (buckling), and environmental sound by Cool Edit Pro. These signals and label information are used in follow-up analyses and predictions. Figure 2 shows the transformation between the AE signal of abnormal sound (buckling) and voiceprints.

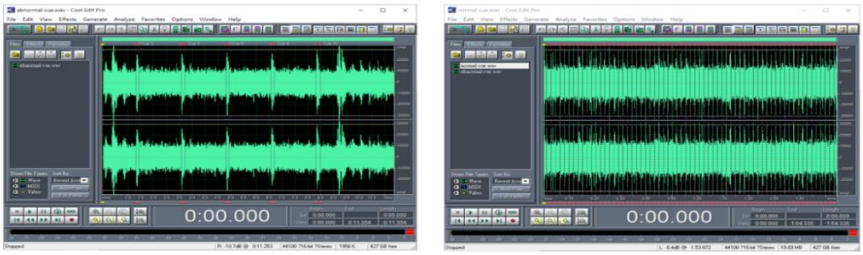


Figure 2. Example of normal and abnormal sound (buckling) voiceprints

3 Methodology and Experimental Results

Feature engineering is performed on the AE signals. The method most commonly applied to acoustic data is MFCC, which can transform collected AE signals into frames for processing. The framing criterion of the duration is 32ms. Moreover, to make sure the change between two adjacent frames is not too large, there is an overlapping area of 706 sample points for every two adjacent frames. After finishing the framing for the AE signals, a total of 70,233 normal sounds, 1,677 abnormal sounds (raised burr), 2,464 abnormal sounds (slug marks), 779 abnormal sounds (buckling), and 512 environmental sounds are produced. Table 5 shows the counts of various types of AE signals.

Table 5. The sample size of various types for AE signals

Items	Normal	Abnormal (raised burr)	Abnormal (slug marks)	Abnormal (buckling)	Environmental
Counts	70,233	1,677	2,464	779	512

Then, pre-emphasis and Fast Fourier transform (FFT) are performed on the processed samples. With MFCC, the Mel Frequency for the spectrum of each frame can be obtained. Then, the logarithmic function and the discrete cosine transform are applied to each Mel Frequency and the results are transformed into MFCCs. Ultimately, 39-

dimensional AE signal features are obtained for each framing. Figure 3 shows the abnormal sounds after the MFCC processing.

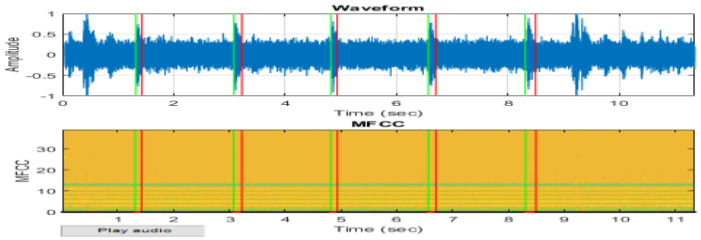


Figure 3. Example of the abnormal sound (buckling) after the MFCC processing

After the feature extraction is completed, the processed AE signals will be sent to HMM algorithm to build an analysis model. The HMM algorithm is viewed as the extension of Markov chains. Considering the fact that the actual problem of engine fault diagnosis is more complex than that can be described by Markov chains, in an HMM algorithm, each state does not correspond to an observable event, but is connected to a group of probability distributions of the state. The HMM algorithm is characterized by the following features:

1. N represents the number of states in the model. The individual states are denoted by $S = S_1, S_2,\ldots,S_N$, and the state at time t is indicated by q_t and obviously $q_t \in S$.
2. M denotes the number of distinct observation symbols per state (i.e., discrete observations). We represent the observation symbols as $V=V_1,V_2,\ldots,S_M$, and the observable symbols at time $t \in S$.
3. $\pi = \pi_i$, the initial state probability distribution, where $\pi_i = P(q_1=S_1)$, $1 \le i \le N$.
4. $A=a_{ij}$, the state transition probability distribution matrix, where $a_{ij} = P(q_{t+1} = \frac{S_j}{q_{t+1}}=S_i)$, $1 \le i, j \le N$.
5. The observation symbol probability distribution in state j, $B = b_j(k)$, where $b_j(k) = P(\frac{v_{k(t)}}{q_t}=S_i)$, $1 \le j \le N$, $1 \le k \le M$

Therefore, HMMalgorithm may be denoted as λ (N,M, π , A, B) . Given appropriate values ofN,M, π , A, B, the HMM can be used as a generator to give an observationsequence as follows:

$$O=O_1,O_2,\ldots,O_T;\ O_T \in V$$

1. Choose the initial state $q_1 = S_i$, based on the initial state distribution π .
2. For t=1 to T
   Choose $O_t = V_k$, according to the symbol probability distribution instate $S_i, b_i(k)$.
   Transition to a new state $q_{t+1} = S_j$ according to the state transition probability $S_i, b_i(k)$.
3. Increment t by 1, return to step 2, if $t \le T$; else terminate the procedure.

In HMM algorithm, the parameters are set as follows: The maximum likelihood estimate based on the Gaussian mixture modelis performed for state probability. Bigram statistics via frame counts are used for transition probability. And then Diagonal covariance matrix is applied for each Gaussian in 5 mixtures. Finally, Leave-one-file-out cross validation is adopted to evaluate the performance of anomaly sound classifier.

The analysis results of employing the anomaly sound classifier are shown in Figure 4. 64,473 normal sounds are correctly classified, and the accuracy of classifying normal sounds is 91.80%. 1,562 abnormal sounds (slug marks) are correctly classified, and the accuracy of classifying abnormal sounds (slug marks) is 93.14%. 1,383 abnormal sounds (raised burr) are correctly classified, and the accuracy of classifying abnormal sounds (raised burr) is 56.13%. 714 abnormal sounds (buckling) are correctly classified, and the accuracy of classifying abnormal sounds (buckling) is 91.66%. 505 environmental sounds are correctly classified, and the accuracy of classifying environmental sounds is 98.63%. Furthermore, we further exam the misclassified abnormal sounds. It is found that 854 abnormal sounds (raised burr) are misclassified into normal sounds, with the proportion being the highest (34.66%). A primary cause of classifier misjudgment may be that abnormal sounds (raised burr) of die aging are similar to normal sounds. Apart from abnormal sounds (raised burr), the classification effects of the classifier based on HMM for abnormal sounds (slug marks) and abnormal sounds (buckling) are good. Only 71 abnormal sounds (raised burr) are misclassified into normal sounds (4.20%), and 28abnormal sounds (buckling) into normal sounds (3.59%).

	Normal sound	Abnormal sound - slug marks	Abnormal sound - raised burr	Abnormal sound - buckling	Environmental sound
Normal sound	91.80% (64,473)	4.25% (2,986)	2.37% (1,668)	1.20% (846)	0.04% (284)
Abnormal sound - slug marks	4.20% (71)	93.14% (1,562)	1.50% (25)	0.95% (16)	0.23% (4)
Abnormal sound - raised burr	34.66% (854)	5.93% (146)	56.13% (1,383)	3% (74)	0.28% (7)
Abnormal sound - buckling	3.59% (28)	2.70% (21)	1.54% (12)	91.66% (714)	0.51% (4)
Environmental sound	0.39% (2)	0.20% (1)	0.20% (1)	0.59% (3)	98.63% (505)

Figure 4. Analysis results for all AE signals

4 Conclusion

This study proposes a quality monitoring technique which can be applied to AE signals. The anomaly sound classifier based on the HMM algorithm is applied to predict the major failure causes for the stamping machine. First, AE signals are collected through the hardware sensors which are installed on the stamping machine of the relay terminals, and then AE signals are manually classified into five types. Then, 39-dimensional AE signal features are extracted using MFCC. This study applies these features to train an HMM anomaly sound classifier. According to the experiment results, the classification accuracy for normal sounds, abnormal sounds (raised burr), abnormal sounds (buckling) and environmental sounds are all over 91.66% except for abnormal sounds (raised burr). It means that the proposed classifier can be used to precisely identify AE signals to help engineers determine whether a stamping machine is in a normal or abnormal state in most cases.

References

1. Wollschlaeger, M., Sauter, T., &Jasperneite, J.: The future ofindustrial communication: Automation networks in the era of the internet of things and industry 4.0. IEEE industrial electronics magazine, 11(1), 17-27(2017).
2. Hagino, N., Komiya, S., Endou, J., &Ishihama, M.: In-process Monitoring during Stamping of Thin Plate: Propagation Characteristics of Ultrasonic Waves in Cylindrical Stamping Dies. Sensors and Materials, 31(10), 3129-3140 (2019).
3. Shanbhag, V. V., Pereira, P. M., Rolfe, F. B., &Arunachalam, N.: Time series analysis of tool wear in sheet metal stamping using acoustic emission. In Journal of Physics: Conference Series, 896(1), 12-30,(2017).
4. Tsuruya, T., Danseko, M., Sasaki, K., Honda, S., & Takeda, R.: Process monitoring of deep drawing using machine learning. In: 2019 IEEE/ASME International Conference on Advanced Intelligent Mechatronics, pp. 1227-1232, Hong Kong, China(2019).
5. Skåre, T., &Krantz, F.: Wear and frictional behaviour of high strength steel in stamping monitored by acoustic emission technique. Wear, 255(7-12), 1471-1479(2003).
6. Shanbhag, V. V., Rolfe, B. F., Griffin, J. M., Arunachalam, N., & Pereira, M. P.: Understanding galling wear initiation and progression using force and acoustic emissions sensors. Wear, 436-437 (2019).
7. Ubhayaratne, I., Pereira, M. P., Xiang, Y., & Rolfe, B. F.: Audio signal analysis for tool wear monitoring in sheet metal stamping. Mechanical Systems and Signal Processing, 85, 809-826(2017).
8. Seemuang, N., Panich, S., &Slatter, T.: Bendability evaluation of sheet metals in three-point bending test by using acoustic emission features. The Journal of Applied Science, 16(2), 15-22(2017).
9. Kothuru, A., Nooka, S. P., & Liu, R.: Application of audible sound signals for tool wear monitoring using machine learning techniques in end milling. The International Journal of Advanced Manufacturing Technology, 95(9), 3797-3808(2018).
10. Wang, L., Mehrabi, M. G., &Kannatey-Asibu Jr, E.: Hidden Markov model-based tool wear monitoring in turning. J. Manuf. Sci. Eng., 124(3), 651-658 (2002).
11. Ertunc, H. M., Loparo, K. A., &Ocak, H.: Tool wear condition monitoring in drilling operations using hidden Markov models (HMMs). International Journal of Machine Tools and Manufacture, 41(9), 1363-1384(2001).

Depth Image Based Rendering System by Kinect Sensor

Beom-Seok Oh[1[0000-0002-7521-7197]], Eungyeol Song[2[0000-0002-5065-2969]] and Sunjin Yu[3[0000-0001-9292-4099]]

[1] Dept. of Computer Science, Gyeongsang National University, South Korea
[2] Research and development department, Codevision Inc., South Korea
[3] Dept. of Culture Techno, Changwon National University, South Korea
sjyu@changwon.ac.kr

Abstract. Various technologies are being studied for realistic video conferencing. Research on stereoscopic images is active during the study of realistic video conferencing. In this paper, we propose a Depth Image Based Rendering(DIBR) system with KINECT. Kinect is a device that can obtain 3D information with a depth camera and is used to generate 3D images. In the proposed system, we create 3D images from depth images obtained from KINECT through DIBR technology. Also, we control the three dimensionality through hand motion recognition.

Keywords: DIBR, Kinect, Depth Image

1 Introduction

In recent years, with the development of communication technologies, there has been an increasing interest in communication methods through remote video calls or video conferences[1-5]. However, in reality, the demand for using a video call or a video conference is not so much compared to the interest, and a more realistic image is demanded as the user-provided environment is improved.

For realistic video conferencing, 3D stereoscopic imaging technology is required, but specialized software technology is required to calibrate specific equipment and cameras such as 3D stereo camera rigs. In addition, 3D stereoscopic image acquisition equipment is expensive and bulky. The Depth Image Based Rendering (DIBR) is used to solve many of these problems. A representative real-time depth image acquisition device used in DIBR technology is KINECT. The Kinect provides depth information between the camera and the object[6].

In this paper, we propose a Depth Image Based Rendering(DIBR) system with KINECT. In the proposed system, DIBR technology[7,8] is applied to generate 3D stereoscopic images. This paper describes the proposed system in Chapter 2 and explains the experimental results in Chapter 3. Finally, we conclude the discussion

2 Proposed System

In the depth image obtained from KINECT, the depth map represented by the X-coordinate system and the Y-coordinate system is expressed by the X-coordinate system and the Z-coordinate system image for segmenting only the human figure of the foreground. Next, the connected component labeling method is used to set the label of each component and the bounding box to set the depth range and X-coordinate direction of each object. In the depth image acquired by Kinect, a phenomenon occurs that is lost in reflective areas or black areas where light is absorbed well. To solve this problem, objects are separated precisely by using depth information and segmentation information so that depth values lost from real objects can be restored. We apply DIBR technology to generate 3D stereoscopic images from depth images[7,8].

We use 3D hand tracking to locate the hand and adjust the three-dimensional effect through the Z-axis information value of the hand position[9]. First, the location of the hand is extracted using kinect's 3D hand tracking library. We assume that the Interocular Distance(IOD) control environment has been established when the hand position is paused for a certain period of time.

3 Experimental Results

From the proposed system, we generate stereoscopic images. Fig. 1. shows pre-processed depth image acquired from a Kinect camera and RGB image.

Fig. 1. Pre-processed depth and RGB images

We build a system so that users can control the three-dimensionality through hand movements. Also, we configure the stereoscopic image to be output as active type and passive type. Fig. 2. shows the three-dimensional effect of IOD transformation as passive type. Fig. 3. is shown as active type through left image and right image for three-dimensional expression.

Fig. 2. Experimental Results for Passive Type

Fig. 3. Experimental Results for Active Type

4 Conclusion

Kinect camera provides both depth information and RGB information at the same time. It is possible to create stereoscopic images through the convergence of depth information and RGB information. In this paper, we propose a 3D stereoscopic image generation system using KINECT. The proposed system is designed to allow the user to adjust the binocular disparity so that the user can provide different stereoscopic effects. In the future, we will apply 3D hand tracking and augmented reality technologies to the proposed system.

Acknowledge

This research was supported by Changwon National University in 2019~2020. And it was supported by the Technology Innovation Program (20004088, Development of

smart underwater robot and underwater environment monitoring system based on multisensor artificial intelligence) funded By the Ministry of Trade, Industry & Energy(MOTIE, Korea)

References

1. Julian Jang-Jaccard, Surya Nepal, Branko Celler, Bo Yan: WebRTC-based video conferencing service for telehealth, Computing, 98, 169–193 (2016).
2. Z. Liao, L. Zhang: Scheduling Dynamic Multicast Requests in Advance Reservation Environment for Enterprise Video Conferencing Systems, in IEEE Access, vol. 8, pp. 76913-76928 (2020)
3. Stefano D'Aronco, Sergio Mena, Pascal Frossard: Distributed Rate Allocation in Switch-Based Multiparty Videoconferencing System, ACM Trans. Multimedia Comput. Commun. Appl., Vol. 13, No. 3s, Article 41 (2017)
4. Y. Bai and F. Kao: Design of the intuitive touch sèreen and the camera control of a remote video conference system, 2018 IEEE International Conference on Consumer Electronics (ICCE), Las Vegas, NV (2018)
5. M. H. Hajiesmaili, L. T. Mak, Z. Wang, C. Wu, M. Chen, A. Khonsari: Cost-Effective Low-Delay Design for Multiparty Cloud Video Conferencing, in IEEE Transactions on Multimedia, vol. 19, no. 12, pp. 2760-2774 (2017)
6. https://developer.microsoft.com/en-us/windows/kinect/
7. Information Technology—MPEG Video Technologies, Part 3: Representation of Auxiliary Video and Supplemental Information ISO/IEC 23002–3 (2007)
8. C. Fehn: A 3D-TV approach using depth-image-based rendering (DIBR), *Proc. VIIP 03* (2003).
9. Joongrock Kim, Sunjin Yu, Dongchul Kim, Kar-Ann Toh, Sangyoun Lee: An adaptive local binary pattern for 3D hand tracking, Pattern Recognition, 61, 139-152 (2017).

LAD Analysis of Policy Reports for Low Birth Patterns

Mi-Jin Kim[1] and Byung-Man Kim[2]

[1] Inje University, Injero 197, Gimhae 50834, South Korea
[2] Kyungnam University, Kyungnamdaehakro 7, Changwon 51767, South Korea
bmkim@kyungnam.ac.kr

Abstract. The purpose of this study was to analyze the main topics of policy reports related to low birth status during the last 15 years (from 2005 to 2019). To this end, the LAD (Late Dirichlet Allocation) technique was applied to the policy report that was reported for 15 years to extract topics of low birth phenomena and to analyze major issues every five years. The results of this study are as follows. First, related topics often featured words such as 'Society', 'Aging society', 'Population', 'Our country', 'Economy', 'Family', 'Education' and 'Environment'. Second, the Topic Modeling Analysis of Policy Reports has been used to extract the main topics according to the time of low birth phenomena and to identify meaningful differences in the composition ratio by time. This study was able to draw major topics related to low birth phenomena in policy reports, identify major issues of low birth phenomena by comparing them with timing, and explore the changes to broaden the understanding of low birth rate phenomena. It can also provide meaningful implications for subsequent policymaking and research implementation.

Keywords: Low Birth Status, Policy Reports, Big Data, Topic Modeling, Latent Dirichlet Allocation

1 Introduction

The ultra-low birth rate, which began in the 1990s, is a problem faced not only by Korea but also by many countries, but among them, the birth rate of Korea is decreasing at an unprecedented rate worldwide. Judgment on low fertility is generally based on the total fertility rate(TFR), and the total fertility rate of 1.5 or higher is considered as the minimum safety line, and when it is 1.3 or less, it is regarded as the lowest-low fertility [1]. Korea's total fertility rate was 6.0 in 1960, but it fell sharply down to 4.53 in 1970 and 3.57 in 1975 due to the population restraint policy implemented to successfully achieve the economic development plan since 1961. The score was 2.08, which was below the population replacement level. Since then, the trend of decline in fertility rate has been intensified, and in 2005, it reached the lowest level in the world at 1.07 and entered the Lowest-fertility State. In 2015, the fertility rate rose slightly to 1.23 people, but in 2019, 0.92 people remained low in fertility for more than 30 years, and to this day, it still shows a very low level of birth rate, far below the population replacement rate [2].

It is predicted that the current lowest level of fertility will continue in the long term, not short-term, and it is difficult to be optimistic about the recovery of fertility. The reason for the long-term persistence of low fertility is that the problems caused by the decrease in population size and the imbalance of the demographic structure go beyond just the cost, and the negative ripple effect exerts influence across society, such as the loss of national growth engines and imbalance between generations. Therefore, it can ultimately cause a very wide range of social problems [3-5].

The government is seeking to find a variety of ways to cope with low birth rates, but its effectiveness and performance are insufficient. Policy related to low birth rate is a representative policy that the government supports and expands to cope with low birth status, and is being implemented for the purpose of revitalizing women's economic activities, supporting work-family compatibility, and healthy development of children as well as raising the birthrate. The government defines the low birth status as a major national phenomenon or crisis, but it approaches the focus of the solution in line with individual factorial and open-ended responses from a short-term perspective rather than from a long-term perspective. Although various policies to cope with low birth rates have been under way for a long time, the phenomenon of low birth rate is getting worse.

This raises the need to look at low birth as a social phenomenon that reveals complexity, uncertainty and instability [6-9]. In addition, attempts to enhance understanding of the phenomenon are needed by comprehensively examining the numerous issues that arise in relation to the low birth phenomenon and the concepts, targets and languages that are produced accordingly. The purpose of this study is to derive the topical features of low-birth phenomena within the policy study report regarding low-birth phenomena, and to examine the overall trends of policies, discourse and social issues. In addition, the purpose of analyzing the characteristics and structure of the major objects or concepts of low birth rate was to compare and analyze them according to historical trends. To this end, the Latent Dirichlet Allocation (LAD) technique was applied to policy reports on low-birth phenomena reported from 2005 to May 2019, and the main issues were analyzed every five years.

In this study, the main issues of awareness, status, and major issues of low birth status in our society were identified by eliciting topics that were noted in the policy study report on low birth phenomena and comparing them by timing. This has allowed us to extend our understanding of low-birth phenomena, and we intend to provide meaningful implications for policy-making and subsequent research.

The research questions for the purpose of this study are as follows.

1. What is the main topic of the policy report on low birth status?
2. What changes do under-births represent depending on the timing of the policy report?

2 Research Method

2.1 Data collection

For this study, data collection was done by crawling the results of searching policy reports to 'low birth' using Selenium. Selenium is a testing framework for web applications and is commonly used for automated testing. Starting in 2005, the five-year cycle was set to be one period (from 2005 to 2009), the second period (2010 to 2014), and the third period (2015 to 2019). The information in the policy report was downloaded in the form of pdf after searching the keyword ‘low birth’ and ‘low birth’ at http://policy.nl.go.kr/. Policy reports that can only be found through the agreed libraries have been excluded because they are not accessible. For the collected data, the policy report is pdf type and contains images other than text, so only text types were imported into the python3.6 program and saved as a file. Through the above process, 242 policy reports on low birth status reported from 2005 to 2019 were chosen as the final analysis target for this study.

2.2 Data processing

In this study, text preprocessing was carried out before the topical analysis of the collected policy repots. For text preprocessing, performed an analysis of the Korean morphology using a function of the Mecab dictionary in Python 3.6 Program KoNLPy Package. Afterwards, the technical terms and proper nouns were organized and added to the dictionary of the users so that the nouns could be properly dealt with in the dictionary of words. For words with similar meanings, the conversion was done to prevent overlapping words with the same meaning as the result of a morphological analysis. Words with unnecessary meanings such as ‘low birth’, ‘policy’, and ‘birth’ were added to the list of words to be removed from the analysis.

2.3 Data analysis

In this study, word frequency analysis and topical analysis were carried out using the R 3.6.0 program after text processing to identify major issues related to low birth status of policy reports collected. Specifically, the tm, topicmodels, lda, slam, ldating, and dplyr packages were used to analyze the data in this study.

First, we analyzed the frequency of the words and checked the frequently used words, and then excluded the meaningless words from the analysis by registering them as unnecessary words. In order to implement the topic modeling of big data, this study used the Latent Dirichlet Allocation (LDA) to discover the core topics of vast texts [10] and to understand the interrelationships among subjects by inferring the parameters of the probability distribution from which the original data was generated from the observed data. LDA is a probability model that estimates the distribution of words and distribution of documents to see which topics exist in each document for a given document. As a sampling method, Gibbs sampling method, which is a technique of changing

only one variable, fixing the remaining variables, and excluding some unnecessary variables from sampling, was used. Iteration was set as 5000 times. In the policy report, we searched for the words with the highest appearance of low birth phenomena, derived topics that were semantically consistent, derived key words for each topic, and analyzed the difference in topic composition and composition ratio for each cycle every 5 years. In setting the number of topics, the researcher used a method to specify an arbitrary number [11] in consideration of the interpretability and suitability of the field.

3 Results

3.1 Top appearance words in policy reports on low birth status

The top 30 words wewre detected throughout 15 years and by timing are presented in Table 1 to look at the key keywords in the policy report on low birth phenomena.

Table 1. Top Appearance Words in Policy Reports on Low Births.

Ranking	2005-2009 (1st Period)		2010-2014 (2nd Period)		2015-2019 (3rd Period)		Total	
	Word	Frequency	Word	Frequency	Word	Frequency	Word	Frequency
1	Nurture	25	Society	56	Education	35	Society	115
2	Society	25	Environment	56	Problem	35	Aging society	114
3	Plan	24	Economy	56	Family	34	Population	113
4	Population	24	Aging society	56	Increase	34	Our country	113
5	Environment	24	Population	55	Economy	34	Economy	112
6	Our country	24	Increase	55	Response	34	Family	112
7	Aging society	24	Change	55	Diverse	34	Education	112
8	Descendant	23	Country	55	Country	34	Environment	112
9	Family	23	Target	55	Target	34	Increase	111
10	Support	23	The public	55	Direction	34	Support	111
11	Burden	23	Government	55	Expansion	34	Age	110

12	Growth	23	System	55	Popula-tion	34	Country	110
13	Propulsion	23	Female	55	Change	34	Expan-sion	110
14	Govern-ment	22	Analysis	55	Society	34	Female	110
15	Increase	22	Diverse	55	Research	34	Problem	110
16	Economy	22	Goal	55	Re-searcher	34	Plan	110
17	Establish-ment	22	Burden	55	Our country	34	Reduc-tion	109
18	Child care	22	Labor	55	Aging society	34	Survey	109
19	Survey	22	Expan-sion	55	Support	33	Average	109
20	Way	22	Family	55	Female	33	Change	109
21	Counter-plan	22	Genera-tion	55	Reduc-tion	33	Target	109
22	Phenome-non	22	Income	55	Enhance	33	Govern-ment	109
23	Welfare	22	Support	55	Data	33	Response	108
24	Female	22	Provision	55	Change	33	Phenom-enon	108
25	Child	22	Welfare	55	Age	33	Research	108
26	Education	22	Introduc-tion	55	Structure	33	Data	108
27	Employ-ment	22	Applica-tion	55	Plan	33	System	108
28	Average	22	Reduc-tion	55	Survey	33	Analysis	107
29	Age	22	Age	55	Key	33	Direction	107
30	World	22	Develop-ment	55	Next	33	Nurture	107

Word frequency analysis shows that 'Society' (total 115 times, first 25 times, second 56 times and third 34 times) are the most common. In addition, the following words appeared at a high frequency throughout the period such as 'Aging society' (Total 114 times, 1st 24 times, 2nd, 2nd 56 times, 3rd 34 times), 'Population' (Total 113 times, 1st, 25 times, 3rd 34 times), 'Our country' (Total 113 times, 1st, 24 times, 3rd 34 times), and 'Economy' (Total 112 times, 34 times, 3rd, 3rd, 32 times, 36 times, 3rd, 3rd, 3rd, 3rd, 3rd, 32 times). Looking at each period, from 2005 to 2009 words, such as 'Nurture', 'Society', 'Plan', 'Population', 'Environment', 'Our country', 'Aging society', 'Descendant', 'Family' and 'Support' were found much frequently. In particular, unlike other times, the words 'Nurture' and 'Plan' were there. From 2010 to 2014, 'Society', 'Environment', 'Economy', 'Aging society', 'Population', 'Increase', 'Change', 'Country', 'Target' and 'The public'. In particular, unlike other times, the words 'Change' and 'The public' were found to have high frequency. Next, from 2015 to 2019, articles reported in the following order were 'Education', 'Problem', 'Family', 'Increase', 'Economy', 'Response', 'Diverse', 'Country', 'Target' and 'Direction'. In particular, unlike other times, the words 'Education', 'Problem', and 'Family' were found to have been shown to be highly frequent in the third period.

3.2 Composition of the Low Birth Phenomenon Policy Report Topics by Period

The main words for each topic extracted from policy reports published from 2005 to 2009 are shown in Table 2.

Table 2. Key words by topic in the 2005-2009 Policy Reports

	Topic 1	Topic 2	Topic 3	Topic 4	Topic 5
1	Facility	Child	The birth rate	Society	Right
2	Service	Survey	Aging society	Descendant	Broadcasting
3	Welfare	An annual pension	Evaluation	Upbringing	News
4	Target	Participation	Propulsion	Education	Upbringing
5	Protection	Culture	Plan	Our country	Female
6	A newborn baby	Function	Government	Life	Mom
7	Local government	Revision	Institution	Child care	Contraception
8	Pregnant women	A target person	Allowance	Environment	A single mother
9	A checkup	Infant	Market	School	Son
10	Infertility	Respondents	Way	Work	Gender
Topic name	Welfare perspective	Supporting target	Policy intervention	Support for education and child care	Woman's birth perspective

Topic names were given after repeated reviews of the composition of the main words by topic and their differentiation from other topics. As a result, the Topic Name was selected as 'Welfare perspective', 'Supporting target', 'Policy intervention', 'Support for education and child care' and ‘Women's birth perspective'. Policy reports from 2005 to 2009 highlighted the characteristics of the low-birth phenomenon era such as analysis of birth rate trends, related statistical analysis surveys, and a decrease in the population of school-age. In particular, the social reality in which the scope of responsibility for child care and education is excessively focused on parents and issues concerning the national welfare and policy intervention to address them were identified as the main topics.

Next, the main words for each topic extracted from the policy report published from 2010 to 2014 are shown in Table 3.

Table 3. Key words by topic in the 2010-2014 Policy Report

	Topic 1	Topic 2	Topic 3	Topic 4	Topic 5
1	Descendant	Childcare	Family	Society	Female
2	Increase	Expansion	System	Allowance	Service
3	Nurture	Target	Welfare	Economy	An annual pension
4	Life	Budget	Country	Time	Marriage
5	People	Provision	Employment	Change	Parental leave
6	Response	School	Law	Plan	Effort
7	Cost	Operation	Public	The birth rate	Institution
8	Market	Assurance	Spending	Government	Participation
9	Experience	Work	Work	Parents	Recognition
10	Production	Counterplan	Salary	The masculine gender	Retirement
Topic name	Cost of child rearing	Financial support	Home welfare support	Major limitations of childbirth	Plan to restore the birth rate

Topics derived from media coverage during this period were ‘Cost of child rearing’, ‘Financial support’, ‘Home welfare support’, ‘Major limitations of childbirth’ and ‘Plan to restore the birth rate’. The 2010-2014 policy reports highlighted marriage and childcare. These results allow us to infer that social members in their marriageable age pay for childcare, have a pessimistic view of childbirth and demand high financial and home welfare support. In other words, we can confirm that in this period people had great pressure on marriage, and awareness review and policy intervention have emerged as the main topics in this period.

Finally, from 2015 to 2019, the main words for each topic extracted from a policy reports are shown in Table 4.

Table 4. Key words by topic in the 2015-2019 Policy Reports

	Topic 1	Topic 2	Topic 3	Topic 4	Topic 5
1	Population	Descendant	Budget	Service	Nurture
2	Task	Goal	Time	Target	Plan
3	Labor	Health	Means	Inspection	Expansion
4	Educational expenses	Parents	Parental leave	Operation	Protection
5	School	Allowance	Health	Internalization	Safety
6	Response	Newly wed	Institution	Management	Responsibility
7	Student	Public	Job	Dissatisfaction	Day care center
8	Reform	Dual-income	Development	After school	People
9	Capabilities	Provision	Resource	Local government	Happiness
10	Function	Encourage	Private sector	Insufficient	Equality
Topic name	Future social response perspective	National support	Practical support	Internalization of low birth response policy	Expectation of birth rate recovery

Topics derived from media coverage during this period were selected as 'Future social Response perspective', 'National support', 'Realistic support', 'Internalization of low birth response policy' and 'Expectation of birth rate recovery'. The policy reports from 2010 to 2014 highlighted marriage and childcare. In this period, the changes in social and structural structure through low birth rate and the actualization and internalization of support to address the problem of low birth rate were emphasized. In addition, the main topics of the policy reports were expanding in view of future prospects, direction and expectations in preparation for future society. In other words, through the topics on future social responses, national support, and internalization of policies, we were able to confirm that the view of low birth status was beyond analysis.

4 Conclusions

This study analyzes the main topics of policy reports related to the low birth rate during the last 15 years(2005-2019), draws attention to topics in policy research reports on low birth rate, and compares them by period. The purpose was to identify the current status of birth phenomena and major issues, and to provide implications for establishing policies related to low birth. Discussion is based on the results presented in this study and is as follows.

First, the main topics extracted from the policy reports published from 2005 to 2009 are 'Welfare Perspective', 'Supporting Target', 'Policitation of Policy Interventions', 'Support of Education and Child Care' and 'View of Women's Birth'. In this period, topics such as 'women', 'mom', 'unwed mother', 'pregnant women', and 'infertility' were considered as the top, and the subject of low fertility problems was recognized as women, and the nation's public responsibility for raising was strengthened. It can be seen that the government tried to solve the low birth rate problem by increasing the government-led fiscal input. In other words, during this period, while emphasizing the arrival of the low birth rate, the low birth rate was defined as a national crisis, and the improvement of the nurturing environment led to the improvement of marriage and childbirth desires, but still the birth of the 'family' rather than the 'women' decision-making problem. It was presumed to be a matter of 'or' individual' [12]. Therefore, there is a limitation that the government's low birth rate policy is approaching the economic paradigm of women, welfare and investment.

Second, the main topics extracted from the policy reports released from 2010 to 2014 are "costs of child rearing," "financial support," "home welfare support," "major limitations of childbirth" and "rehabilitation of fertility." During this period, topics such as 'parental leave', 'male', and 'parent' were drawn to the top, and beyond the support for women's pregnancy and childbirth, education and childcare, the government's low fertility policy led to housing stability, health, and routine. Assumptions were expanded to equilibrium. In addition, the target of the low birth rate policy was expanded from low-income families to working families at this time, and it was possible to confirm the change in the government's policy direction by reducing the burden of childbirth and parenting to support work-family balance. At this time, the government's low birth rate policy expanded the focus of the policy by balancing work and family, but it still has limitations in recognizing childbirth and care as women's problems.

Third, the main topics extracted from the policy reports released from 2015 to 2019 are 'Future Social Response Perspective', 'National Support', 'Realistic Support', 'Internalization of Low Birth Response Policy' and 'Expectation of Birth Rate Recovery'. In the policy report published at this time, policy support based on improving quality of life and establishing gender equality stands out. This reflects the criticism that the existing low birth rate policies did not consider the gender aspect, and is evaluated as a new reconstruction of the low birth rate policy from the perspective of gender equality [13]. During this period, the scope of the policy was further expanded by encompassing diverse types of families, from married women to unmarried men and women, and it was confirmed that policies were implemented to establish a gender equality culture such as parental leave for men. In addition, through topics on future society response, national support, and policy internalization, it can be seen that the perspective of viewing low birth rates goes beyond the level of actual analysis and includes prospects and directions for the future society.

For 15 years, the government's low birth rate policy has gradually expanded its targets, and the paradigm of low birth rate problems and support measures has changed. However, the government still lacks the view of the people who raise children because the national view is dominant rather than the 'people'-centered approach [9], and does not include various aspects of life and social structure. Is pointed out. Therefore, in

order to solve the low birth rate problem, as deviating from the long-established assumption that all things could be explained from a mechanical interaction point of view, it is necessary to pay attention to the complexity of the low birth rate, which is a dynamic system itself, and to listen to the discourse raised differently in the low birth rate policy.

Acknowledgments

This work was supported by the Ministry of Education of the Republic of Korea and National Research Foundation of Korea (NRF-2018S1A5A2A03033511)

References

1. McDonald, P.: Low fertility and the state: The efficacy of policy. Population and Development Review 32(3), 485-510 (2006).
2. Statistics: Statistics Korea Population Trend Survey. http://kosis.kr/, last accessed 2020/03/15.
3. Kwak, Y. C.: The effects of economic factors on household formation: Focused on relative income and housing expenses. Seoul National University Doctoral Dissertation (2017).
4. Park, H. Y., Yeo, Y. G.: An analysis of the causes of low-fertility rate and the directions of future child-care policies in Korea. Korea Journal of Child Care and Education 85, 69-95 (2014).
5. Saracho, O., Spodek, B.(eds.). International perspectives on research in early childhood education. IAP. (2005).
6. Park, S. H.: Problems in and future improvement directions for childbirth policy. Hankuk University of Foreign Studies Master's Dissertation (2015).
7. Lee, I. S.: A Study on the Factors Affecting Low Fertility and the Implication of Socal Welfare. Korean Journal of Social Welfare 57(4), 67-90 (2015).
8. Jang, H. K., Lee, M. J., Kim, K. M., Kim, Y. R.: Women's and National Response Strategy in the Low Birth Age. Korea Women's Policy Institute Research Report 240-15, 1-343 (2004).
9. Jung, H. Y., Kang, S. J., Lee, Y. S.: Trends in low-fertility policy by keyword network analysis: Childcare policy recommendations. Journal of Korean Council for Children & Rights 23(4), 821-844 (2019).
10. Blei, D. M.: Probabilistic topic models. Communications of the ACM 55(4), 77-84 (2012).
11. Kim, H. J., Kim, S. W., Kim, S. T.: Topic Modeling of Media Reports on Smartphone Addiction - A Study on the Comparison of Government Policies between 2010 and 2018. Journal of Broadcasting and Telecommunications Research 104, 38-69 (2018).
12. Shin, K. A.: Issues and dilemmas of population policy in Korea: Gender policy without gender perspective. Journal of Feminism Research 10(1), 89-120 (2010).
13. Kim, E. J., Song, H. J., Bae, H. J, Sun, B., Y., Choi, J. H., Hwang, J., M.: Paradigm shift in policy response to low birthrates(I): An analysis of the gendered life prospective and policy validation of among youth. Seoul: Korea Women ’s Development Institute (2019).

A Semantic Network Analysis of Research Trends on Media Education of Elementary School Students

Su-Jeong Jeong[1]

[1] Tongmyong University, Nam-gu Pusan 48570, Republic of Korea
jsjs@tu.ac.kr

Abstract. In the era of the Fourth Industrial Revolution, one of the core competencies required of students is media literacy. In particular, elementary school students' media education is important because elementary school students have high adaptability and capacity.

In this regard, this study aims to examine the discourse about media education in Korean society for effective media education in the era of smart media. The purpose of this study is to find out the research trend of media education for elementary school students. For this study, online big data collected over the past 5 years has been used on Internet portal sites. Keywords used in this study are "elementary School Students media education" and "elementary School Students media literacy". In the last 5 years, 2,755KB of data was collected through the data mining process from the big data retrieved. In this study, a semantic network analysis method was used.

Based on these research results, elementary school students are using a lot of smartphones or other digital devices, and it seems that there are many media education approaches especially for English education and coding education. It is the time that more effort is needed for better education utilizing media in media friendly environment of elementary school.

Keywords: Semantic Network Analysis, Media Education, Elementary School Students.

1 Introduction

With advent of fourth industrial revolution, the educational field of today has changed a lot. In the knowledge-informatized society, there is a growing awareness that no one can escape the influence of the media, and there is an increasing need to develop the ability to critically accept the contents of media. As like reading, writing and understanding text and grammar, media education should be used in order to gain ability to adapt media literacy and technology of newly emerging media, as well as it is necessary to facilitate content analysis. In particular, the use of media for children has continued to increase, and in previous studies, the research for how the use of old media has had an impact on children's health has been conducted. Reported that media use has an impact on negative development and potential media harm [1].

This is the time when the importance of education for media use and media literacy education has emerged for students in elementary school, when the use of media is rapidly increasing. In particular, elementary school students' media education is important because they are highly adaptive and competent. [2,3].

Therefore, in this study, in order to investigate the social debate about media education of elementary school students in Korea, the data for the last 5 years on perceptions, preferences, attitudes, etc. of elementary school students on media was analyzed.

However, according to the results of our previous studies, In order to analyze how the concept of the Internet medium is defined and recognized in Internet-related education at elementary schools, an in-depth interview was conducted[4]. The results that they do not understand the media from the perspective of a simple tool, and that they must proceed to Internet media education for critical literacy that is understood by a single educational and cultural environment were suggested[4].

And, as a result of a questionnaire survey conducted for elementary school students, the type of media that uses excessively more than 7 hours a day has already been reported, there is a research that emphasizes a multi-dimensional instruction is needed to form a lifestyle for media use from early elementary school period. [5].

Therefore, it was suggested that media education is a cultural environment now should be understood and it was concerned that elementary school students spend too much time on using media. In addition, it was also suggested that the quality of media used by elementary school students was considered

Yet, since the results of previous studies are from most surveys and interviews, so there are limitations to generalizing the results of these studies. In order to be able to bring this generalized conclusion, the Korean society's perception of elementary school media education through big data analysis was examined in this study.

The following are the research questions used to achieve the goal of this research study.

1. What is the social discussion about media education of elementary school students in Korea with big data?

2 Materials and Methods

The purpose of this study is to examine social trends in media education of elementary school students. For this, big data was collected using Internet portal sites such as Google, Naver and Daum. The collected big data was used in this study to analyze the semantic network through the process of data mining. In addition, programs such as R-program, Ucinet 6.0 and Netdraw were used for big data analysis.

In this study, 2,755KB data is collected through big data searched from 2014.08.07 ~ 2019.08.07 with keywords “Elementary School Student Media Education” and “Elementary School Student Media Literacy”.

3 Results and Discussion

In order to find out the research trend of elementary school students' media education, the statistical significance of elementary school students' media education was examined in this study. As a result of the research, the average distribution value of the network data was 24.3090, and the standard error was 2.3838. The Z-score value was Z = 9.9578, and the probability that the test statistic was observed to be higher than the absolute value of the Z value was 0.0002 (p <.01).

Table 1. Frequency and TF-IDF of "Elementary Media Education" & "Elementary Media Literacy"

No	WORD	COUNT	TF-IDF	No	WORD	COUNT	TF-IDF
1	Child	1414	2533.548	31	Practice	258	902.099
2	Education	921	2023.644	32	Coding	256	931.2221
3	Teacher	901	2074.629	33	Need	251	869.8997
4	Object	755	1810.411	34	Understand	251	877.6234
5	Elementary School	730	1750.464	35	Mathematics	244	944.573
6	Class	661	1744.417	36	Training	244	893.909
7	Student	649	1664.652	37	English tutoring	240	929.0882
8	Utility	598	1533.84	38	Video	233	828.3961
9	School	590	1597.75	39	Study	232	831.3764
10	Program	536	1518.602	40	Posts	225	793.4311
11	Media	510	1414.02	41	Textbook	225	840.9757
12	Digital Devices	476	1471.336	42	Era	224	802.7082
13	Operation	467	1375.053	43	Multimedia	223	792.8426
14	Progress	460	1329.571	44	Activity	223	799.1247
15	Middle School	422	1242.553	45	Development	219	790.791
16	Book	421	1281.744	46	Lecture	216	812.4192
17	English	418	1310.637	47	Parents	205	771.046
18	Learning	409	1245.21	48	Future	203	743.703

19	Teenager	402	1223.898	49	Contents	198	740.0586
20	Start	401	1201.289	50	Games	197	762.6266
21	High school	401	1290.769				
22	News	348	1171.819				
23	Smartphone	340	1144.881				
24	Work	322	1023.333				
25	You-Tube	313	1103.751				
26	Think	295	972.2719				
27	Information	278	954.6484				
28	Middle Schooler	276	929.3736				
29	Experience	265	918.42				
30	Summer Vacation	259	913.3274				

As a result of analyzing the word frequency according to the simultaneous appearance by collecting online data of "Elementary Media Education" & "Elementary Media Literacy", 'Child' showed the highest frequency in 1414 times. Subsequently, 'Education' appeared 921 times, 'Teacher' 901 times, and 'Object' 755 times.

In addition, the frequency was shown in the order of Elementary School(730), Class(661), Student(649), Utility(598), School(590), Program(536),Media(510), Digital Devices(476), Operation(467), Progress(460), Middle School(422),Book(421), English(418), Learning409, Teenage(402), Start(401), High school(401), News(348), Smartphone(340), Work(322), You-Tube(313), Think(295), Information(278), Middle Schooler(276), Experience(265), Summer Vacation(259), Practice(258), Coding(256), Need(251), Understand(251),Mathematics(244), Training(244), English tutoring(240), Video(233), Study(232), Posts(225), Textbook(225), Era(224), Multimedia(223), Activity(223), Development(219), Lecture(216), Parents(205), Future(203), Contents(198), and Games(197).

It can be seen that the words Child, Education, Teacher, Object, Elementary School, Class and Student, etc. are derived as key words in elementary media education.

As such, the fact that many words such as Child, Education, and Teacher are mentioned can be interpreted as the importance of education based on the interaction between elementary school students and teachers in media education.

That is, when media is used as an educational tool, interaction between students and teachers may not be active. Therefore, in education using media, education that the interaction between people is thoroughly considered must be designed.

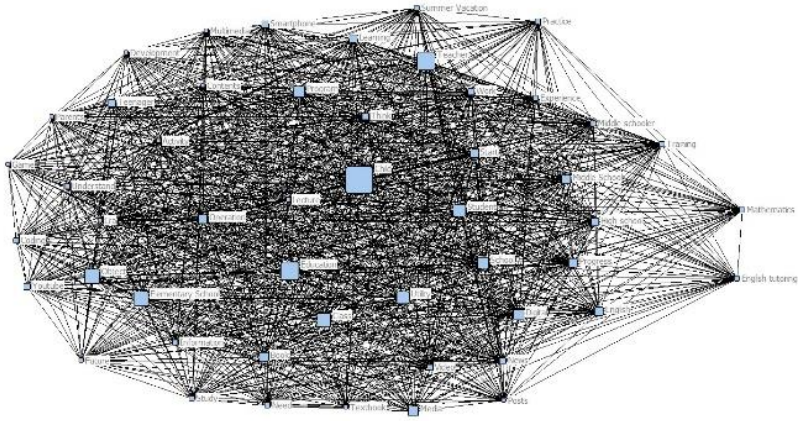


Fig. 1. 50 Words Graph of "Elementary Media Education" & "Elementary Media Literacy"

Figure 1 shows the Network analysis of "Elementary Media Education" & "Elementary Media Literacy" through Netdraw.

Words such as Child, Education, Teacher, Object, Elementary School, Class and Student are more prominent. In other words, Child, Education, Teacher, Object, Elementary School, Class and Student can be seen as key words in this study.

In Elementary Media Education, words such as Child, Education, Teacher, Object, Elementary School, Class and Student are shown as key words in this study. As mentioned above, many words such as Child, Education, and Teacher can be interpreted to mean that the interaction between elementary school students and teachers, who are the subjects of education, is important in the classroom.

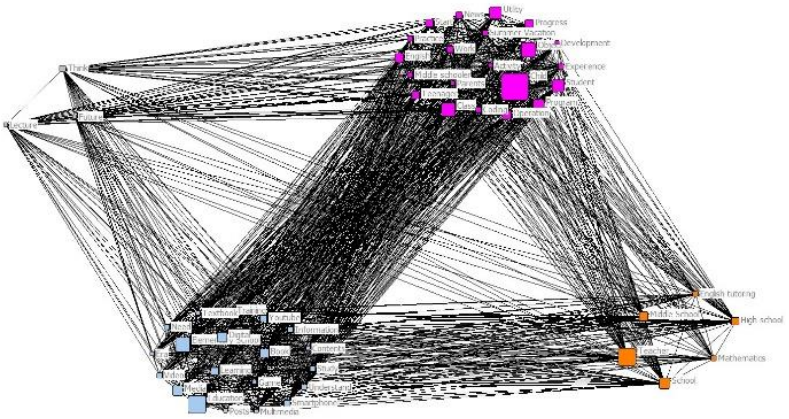


Fig. 2. "Elementary Media Education" & "Elementary Media Literacy" 50 Words CONCOR Graph

In particular, through Concor analysis, which visually groups and displays 50 words, it is possible to confirm what words are highly relevant to the Elementary Media Education. Looking at four representative groups, it is as follows.

First, Concor analysis of this study shows that words such as Child, Object, Student, Utility, Program, and Coding are combined. These results indicate that Elementary Media Education is being applied through programs such as coding education that children experience.

Second, it consists of words such as Education, Elementary School, Smartphone, YouTube, Digital Devices, Media, and Information. Through this, media and digital devices are actively used to educate elementary school students. Especially, the use of media has become more common as smartphones are widely distributed to elementary school students.

In particular, it can be inferred various ways are being tried in the educational field to acquire various information and improve problem solving ability in order to enhance learning ability using media[6,7].

Third, it consists of words such as Teacher, School, Middle School, High School, English Tutoring, and Mathematics education. In other words, it can be inferred that the teacher uses media education to teach English[8] and mathematics[9, 10] to middle and high school students.

Finally, it consists of words such as future, lecture, and thought. This means that deep thinking about lectures is necessary for future students.

Table 2. Analysis of "Elementary Media Education" & "Elementary Media Literacy" using Ucinet"

No	Word	Degree	Closeness	Eigen-vector	Between
1	Child	3170	49	0.265	1.418
2	Education	2508	49	0.118	1.418
3	Teacher	2779	49	-0.379	1.418
4	Object	2422	51	0.324	0.618
5	Elementary School	1904	51	0.082	0.618
6	Class	1680	49	0.069	1.418
7	Student	1751	49	0.042	1.418
8	Utility	2048	49	0.183	1.418
9	School	1804	49	-0.441	1.418
10	Program	1870	49	0.273	1.418
11	Media	1553	51	0.119	1.254
12	Digital Devices	1677	49	0.137	1.418
13	Operation	1554	50	0.23	0.998
14	Progress	1479	49	0.205	1.418
15	Middle School	1243	49	-0.189	1.418
16	Book	1032	50	0.093	1.213

17	English	1248	50	0.116	1.367
18	Learning	1312	49	0.103	1.418
19	Teenager	1214	51	0.157	0.618
20	Start	884	49	0.101	1.418
21	High school	1562	49	-0.978	1.418
22	News	909	50	0.09	1.348
23	Smartphone	873	52	0.111	0.518
24	Work	935	49	0.099	1.418
25	You-Tube	815	52	0.091	0.597
26	Think	717	49	-0.01	1.418
27	Information	800	50	0.082	1.038
28	Middle Schooler	948	50	0.032	1.336
29	Experience	899	49	0.121	1.418
30	Summer Vaca-tion	900	53	0.174	1.066
31	Practice	816	57	0.087	0.835
32	Coding	663	52	0.079	0.414
33	Need	829	51	0.061	1.119
34	Understand	726	51	0.079	0.618
35	Mathematics	1168	62	-0.998	0.413
36	Training	956	54	0.015	0.958
37	English tutoring	1145	63	-1	0.362
38	Video	593	50	0.063	1.336
39	Study	535	52	0.02	0.985
40	Posts	412	52	0.018	1.059
41	Textbook	550	51	0.042	1.104
42	Era	675	51	0.062	0.618
43	Multimedia	660	52	0.046	1.146
44	Activity	632	51	0.055	0.976
45	Development	569	53	0.046	0.474
46	Lecture	614	49	0.017	1.418
47	Parents	479	52	0.014	0.597
48	Future	554	55	0.007	0.2
49	Contents	547	51	0.062	1.226
50	Game	549	53	0.056	0.45

The results of the analysis of the centrality of the 50 major words related to Elementary Media Education are as follows. First of all, when looking at the Degree centrality, the values were higher in the order of 'Child', 'Education', 'Teacher', and 'Object'. This can be confirmed that it is directly connected with various words among the key words related to 'Elementary Media Education'.

Second, in Closeness centrality, 'Object', 'Elementary School', 'Media', 'book', 'Teenager', 'Smart Phone', 'You-tube' appeared in the order. Closeness centrality words mean that all words in the data network can be easily influenced. That is, words such as 'Object', 'Elementary School', 'Media', 'Book', 'Teenager', 'Smart phone', 'You-tube' etc. can be seen as playing a central role in the network structure.

Finally, if the Between centrality values was analyzed, 'Child', 'Education', 'Teacher', 'Class', 'Student', 'Utility', 'School', 'Program', 'Media', 'Digital Devices ', 'Progress' and 'Middle School' are derived. As a result, keywords such as 'Child', 'Education', 'Teacher', 'Class', 'Student', 'Utility', 'School', 'Program', 'Media', 'Digital Devices', 'Progress' and 'Middle School' are influential because they act as mediators in linking words. This means that media education programs experienced in elementary school play an important role for both students and teachers, and are even affecting middle schools.

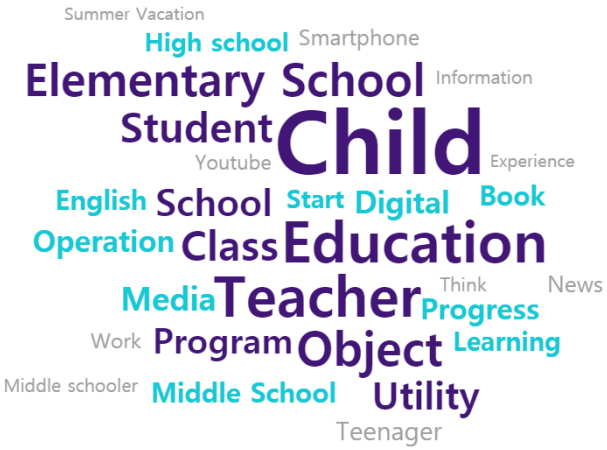


Fig. 3. Word Cloud for "Elementary Media Education" & "Elementary Media Literacy"

4 Conclusion

As a result of analyzing about 7,577KB of online data to analyze social perceptions and trends of elementary media education, significant information about elementary media education was extracted.

First of all, the keywords with the highest frequency of appearance were 'Children, Education, Teachers, Object, Elementary Schools, Classes, Students'.

And, as a result of analyzing the semantic network, the keywords with the degree centrality were 'Children, Education, Teachers, Object, Elementary School, Class, English Tutoring'. Keywords with high betweenness centrality were 'Object, Elementary School, Media, Operations, Teenagers, Smartphones, English, Coding', etc. to confirm the influence of smartphones. The keywords with closeness centrality were ‘Children, Education, Teachers, Classes, Students, Utilization, Programs, Digital Devices, and English Tutoring.′

Through these research results, elementary school students are using a lot of smart phones and digital devices, and in particular, it seems that a lot of media education approaches for English education or coding education are occurring. In order to understand the media in an educational and cultural environment, elementary media literacy education may be needed more often.

Students will enhance their ability to learn new knowledge by collecting, editing, and generating a variety of information, using the important factors of media education such as immediate access, knowledge extensibility, and collaborative interactivity. Education using media is becoming more common in recent years, and media has been actively utilized as a method of self-directed learning for students as well as elementary school classrooms. When considering these situations, it is expected that learning ability using media will be noted as more important core ability in the future.

Therefore, we should continue to strengthen our educational efforts to elementary school students in their early elementary school stage on the right attitudes and ethical norms about the use of media as well as activities in the world of smart media.

We should increase our ability to have more positive attitudes and beliefs. In addition, when media education is carried out in the elementary school classroom, it is necessary to focus on the interaction between the teacher and the student, rather than focusing only on media utilization as an educational medium. In particular, it is suggested that teachers provide meticulous guidance in the field for students to improve their ability to choose good contents of media and to control their ability to facilitate media so that students can improve their ability to solve problems through deep thinking.

5 Acknowledgment

This work was supported by the Ministry of Education of the Republic of Korea and the National Research Foundation of Korea (NRF-2019S1A5B5A02034983)

References

1. Strasburger, V. C., Jordan, A. B., Donnerstein, E.: Health Effects of Media on Children and Adolescents. Pediatrics 125(4), 756-767 (2010).
2. Ryu, H. K.: The Nature of Elementary Education: A Proposal for the Research Orientation. The Journal of Elementary Education 19(1), 1-16(2006).
3. Park, J. K.: The Constructivist Interpretation on Child-centeredness. The Journal of Elementary Education 19(1), 17-36(2006).
4. Lee, J. S.: Defining the Internet for Internet Media Education In Elementary School. Media, Gender & Culture 15, 195-238(2010).
5. Jeong, S. J.: The Factors and Development Outcomes of Latent Types in Children's Time Use. Doctoral dissertation. Ewha Womans University, Seoul (2018).
6. Sung, E. M.: The Relationship of Smart Media Literacy`s Factors for Primary School Student on Subject Attitude and Achievement. Journal of Korean Association for Educational Information and Media 21(2), 215-243 (2015).
7. Jang, S. P., Kim, J. I.: Developing an Resource-based Learning Model for Media Center in Elementary School. The Journal of Elementary Education 19(1), 129-152 (2006).
8. Cho, H. R.: Use of Digital Textbooks as an Alternative to Translation-based Elementary. Doctoral dissertation. Sejong University, Seoul(2020).
9. Eom, W. Y., Lee, H. M.: The Effects of Flipped Learning Using Educational SNS: Focusing on Mathematics in Elementary School. Journal of Korean Association for Educational Information and Media 23(1), 109-138 (2017).
10. Lee, S. S., Jin, C. H., Kang, J. C.: Development of STAD-Based Elementary Mathematics Online Cooperative Learning Game Model. Journal of Korean Association for Educational Information and Media 20(2), 217-246 (2014).

Implementaion for Used Trading Management System Based BlockChain(Case: Used Car)

Byeongtae Ahn[0000-0003-3431-9493]

Liberal & Arts College, Anyang University, 22, 37-Beongil, Samdeok-Ro, Manan-Gu, Anyang 430-714, South Korea.
ahnbt @anyang. ac. kr

Abstract. Internet on e-commerce management system is gradually increasing and various items in the transaction made it. In particular, the used car market scale continuously growing there. But online on sale transactions and the car as high the transaction the seller and purchaser between reliability is very important. Several of the areas used transactions, but the amount is big owned the car for transaction management system by developing reliable improvement was increased. This system uses smart contracts to ensure reliability based ethereum without third party intervention. In addition, this system mitigated information asymmetry between buyers and sellers, and reduced and prevented brokerage fees in the distribution process without third parties.

Keywords: *Trading system, BlockChain, Ethereum, Smart Contract, Used car*.

1. Introduction

The Internet and advances consumer products to buy more and more ways to diversify and has , more cheaper products to buy for online sale transactions for sites using their buyers steadily increasing it . However, in online transactions, there is a problem in that it is difficult to guarantee the reliability of the transaction since the buyer can not directly check the goods [1]. Most of the online sale trading platform a scam to avoid for safety trading sites or sellers of fraudulent transaction history to be queried may in fraud damages lookup sites available that recommends it. This also precautions despite a hassle due to the prevention of site utilization is high, that is a part. Moreover, safe trading site itself phishing (phishing) to be used also because it completely safe that way sometimes difficult . The purchaser third party intervention without a deal trustworthy can you order the seller and selling the goods for the certificates is needed. If the seller's sales history or reliability is provided safely, the reliability of the transaction can be further enhanced. blockchain technology, the information center server instead of the P2P network to the nodes by distributing the information reliability which guarantees a technique [2]. Present this paper these issues to resolve to block the chain , based on a reliable vehicle information provided by trading participants the information asymmetry sex less smart a contract with a stable trading environment that provides used car transaction management system was developed . This paper follows as it is composed.

2 Section System development for related research presented and the three chapter provider of Leeum with used car trade management system to designed. 4 chapter public block chains based provider of Leeum using used car deals Systems designed for based system was implemented. Finally, In 5 chapter, it proposed conclusions and the future works.

2. Related Studies

This chapter in the introduction mentioned one block chains based on pre trade management system to develop for relevant studies were described. Block chain technology transaction information ledger (Ledger) to all nodes (users) are equally distributed to manage the technology. Of the ledger the transaction information as a block, and predetermined time unit (bit coin 10 minute, Ether Solarium 12 second) to the new transaction information by collecting the blocks produces. The resulting block before the block are connect, such a process all associated blocks (block chain) is made to [3]. We deal also can P2P system built. Expert blocks the chain forged or modulation can't be very secure as technology called because the hash value [4]. Transfer to occur before the block a unique one hash value is generated. Block head, the previous hash value and the current hash value is. If someone attempts to forge or change the data value of a particular block for malicious purposes, the hash value of the block changes [5]. Blockchain are working to prove (PoW) algorithm is used. PoW is a bit of coin to get you used an algorithm [6]. Computer hard math problems by creating a chain all the nodes have the problem solved and the random finds. A random locate nodes in the transmission success and Fig. 1, and like the previous block and a hash to a value associated with the structure have. In addition to node -bit coin service. This is called mining. In addition, transfer the two kinds of a way is done. The first second is above mentioned one is mining the two second is the bit coin other users sent to be when raised. In this case, the output section sends information about how much money is being sent to which users.

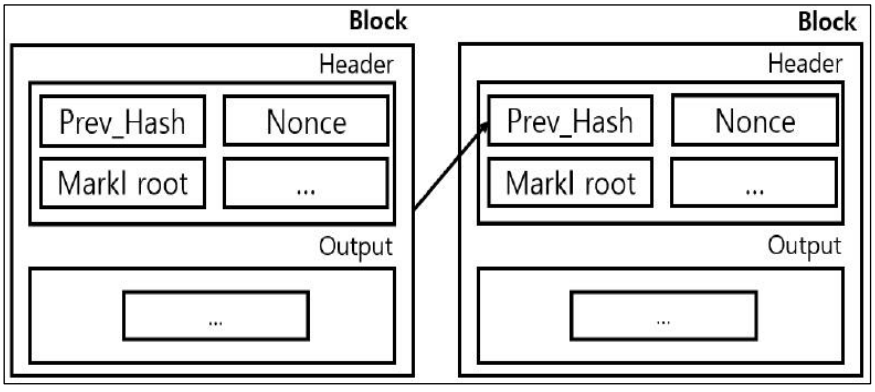


Fig. 1. Block Head Architecture in BlockChain

Bitcoin, first proposed by Satoshi Nakamoto, is the first attempt to change the central bank method used by existing banks to a decentralized method using blockchain technology[7]. The blockchain is not a method of storing users 'data on a central server but rather all data on a user's computer, so that all users participating in the blockchain can access not only my data but also all users' data. The blockchain uses a hash function to establish a very secure data storage method that cannot forge or tamper with data. The inside of the block of the blockchain used by Bitcoin is shown in Fig. 2. Inside the block, the old hash value is stored so you can see it as if it were chained together. If a malicious user attempts to forge or tamper with the data in a particular block, the currently stored hash value is changed and accordingly, the hash value connected to the front and back must be changed accordingly[8]. It is not possible to change the hash value of every block in the blockchain. Even though it is a decentralized system, users can use it without worrying about data tampering.

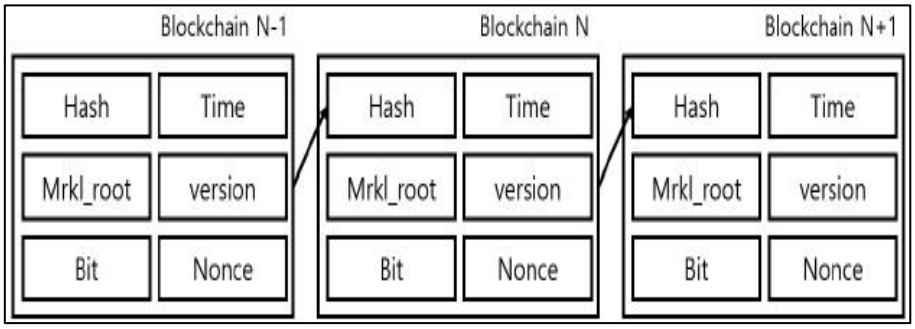


Fig. 2. Block Inside of Bitcoin

Smart contract is in Ether Solarium decentralized applications (DApp, Decentralized Application) for developing for the program. Although other cryptocurrencies mainly provide the function as a currency, Ethereum can be applied to various application fields using smart contracts.

Smart contract is Ethernet Solarium dependent language solidity language used to develop can and Fig. 2 and as a block in the Smart the contract are included. The block transaction is history that is shared as such contracts also account among the shared storage distribution is.

Blockchain multiple node data check and verify the data based between nodes of data sharing can be by allowing digital to the data of the trust to create . Reliable capable of data based on smart this contract blocks and chains together have been published [9].

3. Blockchain based used car trading system design

In this paper, blockchain is used to solve the asymmetry of used car information. Block chain is distributed as a director system in a number of nodes all of the information has. In addition, the block chain is an agreement algorithms and encryption, hash (Hash) the used data structures, etc. because of the following, such a characteristic having [10].

Integrity: the block chain hash (Hash) the use of data structures and many of the nodes in the agreement by management being done because once entered data above / modulation is quite that difficult to attribute to have it.
Reliability: the chain block by using the data for integrity is assured users block chain system in the data trust and use can have. We used the car trust that the information provide above for such characteristics have in the block chain as suitable was determined. Fig. 3 shows the configuration of this system. Using the Node.js server-based database Mongo data establish that smart a contract stored hash of IPFS was used to return a function. The server used Web3.JS to interact with the functions of the smart contract and Tx for the smart contract call . And blockchain transactions were sent to the Ethereum blockchain through geth .

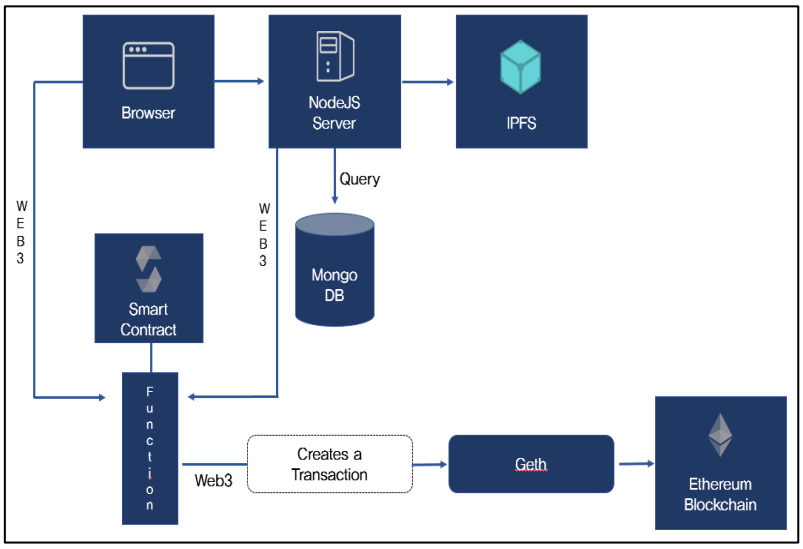


Fig. 3. Flowchart Diagram of System

In this paper, we use Ethereum-based smart contracts to provide contractual vehicle reliability without providing transparent vehicle data and third-party assurance. Representative characteristics of blockchain are integrity and transparency. Each block header hash-encodes the information of the previous block. Therefore, the transaction history recorded in the linked blockchain cannot be forged or altered unless the previous transaction history is modified. In addition, anyone can view the transaction history recorded in the block and participate as a node. Smart contracts automate the process of contract execution and verification into the network, enabling secure contract execution without the need for third-party involvement, based on the trust of implications.

4. Implementation of used car transaction management system

This system is a Windows 10- based system and uses Visual Studio Code 1.26 version, which supports code intersense function and Typescript. Visual Studio is IDE of expansion feature the most outstanding. The server is full stack is enabled and asynchronous programming that supports Node.js 8.11 the used. As a language for writing smart contracts, Solidity version 0.4.21 was used. HTML5, CSS3, and Javascript languages were used as programming languages for web and app support. Server in a database of memory DB of MongoDB 3.4.1 for the use transactional processing have improved Contr ares CT compilation for the real-time efficiency Good Remix for used.

In order for a seller to register a used car, he must access the site and register the vehicle he / she wants to sell. In order to register a used car for sale, you must access the relevant site, enter the seller's basic information, and then enter the relevant information in the vehicle registration.

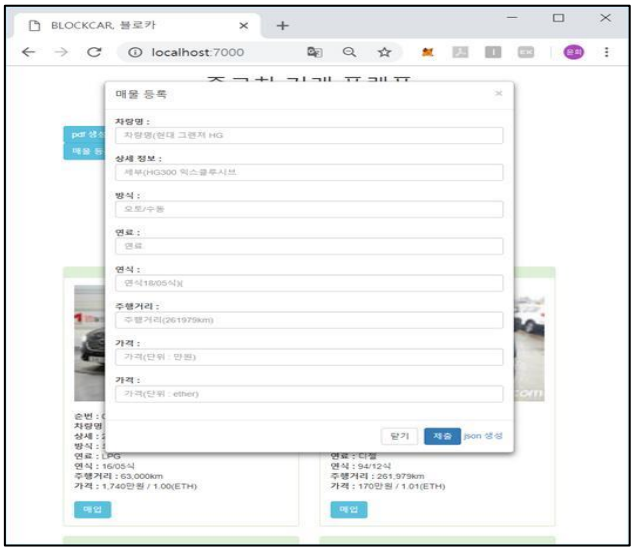


Fig. 4. Register List of Used car

The Fig. 4 shows the initial screen for registering a vehicle for sale. The most basic is to input the vehicle name and detailed information about the vehicle. In addition, the gear shift method, fuel type, year of delivery, mileage, and won price are entered. And rather than All prices in Ethernet Solarium deal to be able Ethereum is that prices are registered. After inputting the detailed information and clicking the Submit button, it is registered in the vehicle registration list.

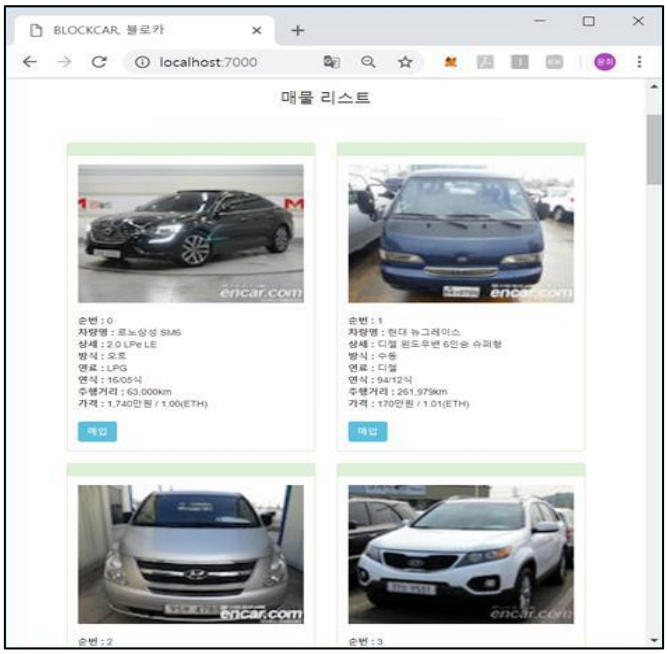


Fig. 5. Used car Information List

Fig. 5 shows the vehicle information list according to the property registration. Vehicle information is listed according to the input order and may be newly sorted according to the year and price of delivery. If the buyer wants to view by price, click the price in the relevant listing to automatically sort the vehicle price in the lowest order.

5. Conclusion and future works

Present the paper used car trading market problems to solve for Ethernet Solarium based smart a contract to use the used car trading system was implemented. The system is reliable, it can be data that the input is assumed. In other words, the data input after the integrity guarantee, but entered before validation This verification must be accompanied. This also limits though, this study is the last issue, the block chain technology using the existing used car market problems solved by reliability is guaranteed used car to trade can be had. Future Graduate School of zero used automatically car deals as well as the entire automotive data transparently trading managed to be so that the system must be extended to.

References

1. Ching-Hsue Cheng, Chung-Hsi Chen, You-Shyang Chen, Ho-Long Guo and Chien-Ku Lin(2019), "Exploring Taiwanese's smartphone user intention: an integrated model of technology acceptance model and information system successful model", International

Journal of Social and Humanistic Computing, Vol.3 No.2, 97-107, DOI: 10.1504/IJSHC.101591.

2. YutingPanaXiaosongZhangaYiWangaJunhuiYanaShu-onvZhouaGuanghuaLiaJiexi.(2019) Application of Blockchain in Carbon Trading. Energy Procedia, 158(2), 4286-4291. DOI: 10.1109/TIFS.2017.2725820
3. Zhang C, Wang Q and Shi D(2016), "Scenario-based potential effects of carbon trading in China: An integrated approach", Applied Energy, 182, 177-190
4. Noor S, Yang W and Guo M(2018), “Energy Demand Side Management within micro-grid networks enhanced by blockchain", Applied Energy, 228, 1385-1398.
5. Zhou Y, Wu J and Long C(2018), Evaluation of peer-to-peer energy sharing mechanisms based on a multiagent simulation framework", Applied Energy, 222, 993-1022.
6. Sikorski JJ, Haughton J and Kraft M(2017), "Blockchain technology in the chemical industry: Machine-to-machine electricity market", Applied Energy, 195, 234-246
7. X. Xu, I. Weber, M. Staples, L. Zhu, J. Bosch, L. Bass, C. Pautasso, P. Rimba(2017), "A taxonomy of blockchain-based systems for architecture design", IEEE International Conference on Software Architecture (ICSA), 243-252.
8. Zibin Zheng, Shaoan Xie, Hong-Ning Dai and Weili Chen(2020), "An overview on smart contracts: Challenges, advances and platforms", Future Generation Computer Systems, Vol. 105, 475-491.
9. Ahmet Bugday, Adnan Ozsoy, Serdar Murat Öztaner and Hayri Sever(2019), "Creating consensus group using online learning based reputation in blockchain networks", Pervasive and Mobile Computing, Vol. 59, 111-125
10. Matevž Pustišek and Andrej Kos(2018), "Approaches to Front-End IoT Application Development for the Ethereum Blockchain", Procedia Computer Science, Vol. 129, 410-419.

Design of Real Estate Contract Management System based Blockchain

SoonHyeong Jeong[1][0000-0001-8596-0382] and Byeongtae Ahn[2][0000-0003-3431-9493]

[1] Onther Inc., 527, Gangnam-daero, Seocho-gu, Seoul, Republic of Korea
kevi n. j @nt her . i o
[2] Liberal & Arts College, Anyang University, 22, 37-Beongil, Samdeok-Ro, Manan-Gu, Anyang 430-714, South Korea.
ahnbt @ anyang. ac. kr

Abstract. The existing real estate transaction is highly interested in a contract that is guaranteed to be reliable due to anxiety about the threat of fraud. In particular, it is timely and economically efficient to manage contract information online and prevent forgery and duplication of contract information. Blockchain technology is emerging as an alternative to the enormous cost and reliability issues associated with offline real estate contracts. If blockchain technology is applied to online contract management, reliability and scalability are efficient and confidentiality can be ensured. Therefore, in this paper, we designed an Ethereum-based online real estate contract management system among blockchain technologies. The real estate contract management system enables online contract management and discrimination of contract forgery through blockchain. In particular, it was possible to prevent fraud until the contract was concluded and the contract was terminated.

Keywords: *Real Estate Trading system, BlockChain, Ethereum, Smart Contract*

1. Introduction

The domestic real estate market trading volume year deal number is increasing, but real estate deals damage and fraud continued to increase it. Contract fraud prevention for the procedure in accordance with the agreements in progress, but the process of loopholes to exploit various fraud incident occurs there. Real Estate Contracts four groups to avoid for the third -party intermediaries via a contract that signed this approach is common. However, fraud occurs due to the lack of credibility of a third- party broker. And the third party is lacking, not duplicate contract fraud is also the case frequently occurs there[1].

In order to solve these cases, it is necessary to have a technical system that prevents forgery of contracts and guarantees trust in advance. Therefore, the this paper reliability is 100% with guarantee block chain technology using P2P trust -based networks by building brokers without a secure real estate The agreement is established can be in the system was designed . This system was able to forge and falsify by creating a contract using smart contracts on the basis of Ethereum, a distributed computing platform. Pre-

sent in paper contracts for which management API for Open API to haejumyeon provide real estate Brokerages, Stock Brokerages such as contracts related to service contracts for management for the functions that will provide be it. The platform API for utilizing Brokerages are present on the system of small, only fees you pay online on safe and reliable in the contract to manage a low cost is possible. In Chapter 2 of this paper, a related study was prepared, and in Chapter 3, a real estate transaction management system was designed. Finally, the fourth chapter, conclusions and the future work is proposed.

2. Related Studies

The paper in the in the introduction mentioned block chain -based systems to develop for the block chain related studies out look.
The block chain Ether Solarium is proof-of-work(Proof of Work PoW) algorithm is used [2]. PoW is Bitcoin (bitcoin) the to acquire used algorithms in computer hard math problems inside the chain participation by all nodes that the issue released any number (nonce) to discover the behavior you are. Any of a number found node to the transmission is successful, and Fig. 1, and as the previous block and the hash to the value by which connecting structure has is.
Also applicable for node schedules the bit coin service and this drilling (mining) referred to. This only as transfer 2 kinds on the way by there is performed the first second is above mentioned is mined, the other users of Bitcoin to transfer when the transfer takes place. At this time, the output section (output section) in which the user for some time as to whether transfer of my dragon platter sent is [3].

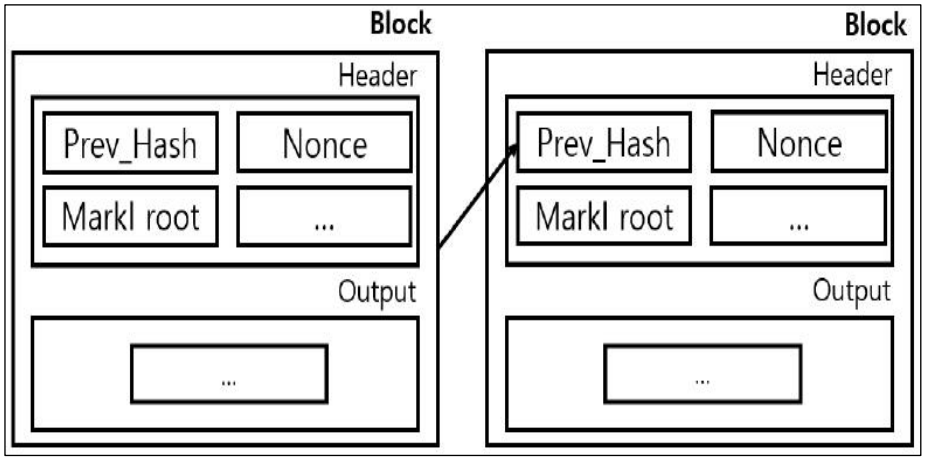


Fig. 1. Block Head Architecture in BlockChain

Smart Contract is Nick Szabo is 1994 years first proposed the concept existing contract (Contract) is written there is agreement the conditions to implement this person directly

be done , but the digital contract (digital contract) is the condition according automatically the vote that will . Block chain is a number of the node data validation and, proven ones based between the nodes sharing the way through digital data trust generates . Believe can in a data -based smart contract is to block the chain together was at issue [4].
Smart Contract techniques apply Block Chaining operation structure follows the same. Assuming that there are buyers and sellers who want to purchase the goods, the seller registers the goods to be sold on the blockchain. The smart contract the transaction in accordance with that in the block chain new node registration and real-time updating i. Buyers search (Query) to use that block the chain to query can have the goods if you buy smart contract is that of the node database updating it.

Ether Leeum's programming available to block chaining various blocks chain services to develop to be added can be there. Ethereum is a distributed network platform that can be operated as a client program [5].
Leeum provider of currency units ether as a deal for a fee are used. In Ethereum, programming is not required in Gold Dragon, but it is supported for smart contracts. The sender and the receiver between a transaction (the financial transaction) or contract non-financial transactions) in the form and constitutes , passed when the system to protect from dimensional charge of ether to together passes . Numismatic type plain deal for ether, a small payment for finney, trading upon fee payment for szabo, wei is there [6].
Fig. 2 is a provider Solarium -based smart contract that is driven the process shown would. The contract to be developed is included in the block and transferred to other blocks in the blockchain to be executed during verification. Smart contract is to be executed in order Ether Solarium virtual machine (EVM) to over the drive is 7].

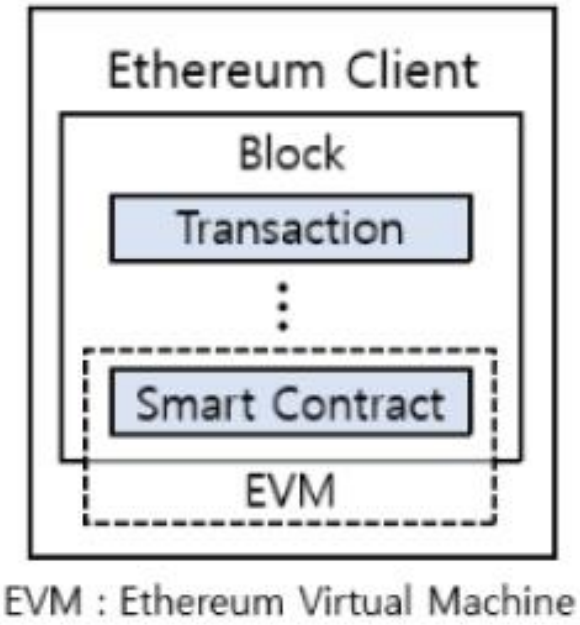


Figure 2. Smart Contract Runtime Environment in Ethereum

3. Real estate transaction management system design

Real estate transaction management system is offline procedures included can only free trading contracts online on which management will. This system is the two kinds of advantages are. First, it prevents forgery and alteration of the contract . Present the thesis contract information and contracts storing and managing to block chains leverage. The block chain contract storage and to manage if trust -based networks on a contract to save because the contract is forgery impossible. Second, it prevents duplicate contracts. Already the contract is entered into the item for duplicate an agreement to proceed if their contract in advance allows filtering. However contract target item present on the system registered the identity has been granted to be state and [8].

The system functions by five different modules (USER, CONTRACT, FILE, TRANSACTION, MAILING) to be configured. USER module createUser, getUser, login, modifyUser, deleteUser to be classified. CONTRACT module addContract, findContract to be classified. The FILE module and the MAILING module are configured independently. And TRANSACTION module createTransaction, getSenderTransact ion, getRecipietnTransaction, acceptTransaction to be classified. Each of the modules in accordance with API of the design and the model was constructed [9].

Present in the system configuration of the block chains of based IPFS storage how to use so as to be driven was. Fig. 3 is present on the system configuration also shown to.

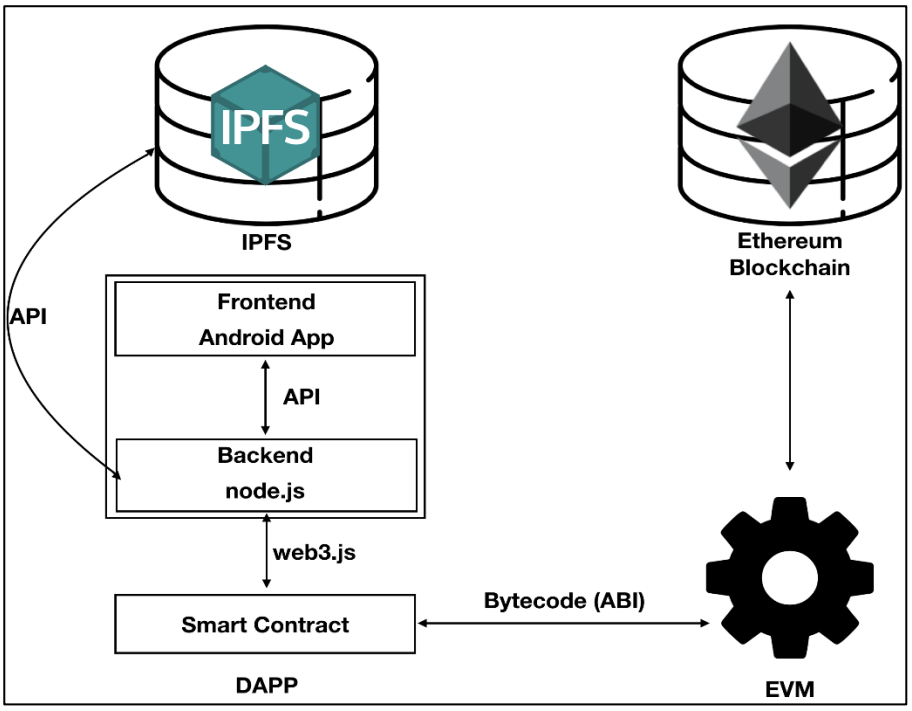


Fig. 3. System Architecture

This system is the system performance, a maximum for node.js the server were used block chains of based distributed computing the network supporting the Solarium Ether was used . Smart contract document storage to block chain -only storage system IPFS was used IPFS in providing that API for using a document accessible to enable were [10]. IPFS in documents Move to web3.Js the used. Ether Solarium virtual machine, byte code smart contract with each other is moved to was. The system will block the chain through the agreement now manage, because online on safety and reliability makes the contract makes management . The flat on the form of transactions registered by the contract to manage the agency to report that the same legal effect obtained is not present on the platform that managed just also notarized the effect obtained may have . Therefore this system to take advantage of trading of the company if the deal the companies you want to use to customers I deal the company a contract to clear that that can prove be there. And customers in the admission of this in the system contract to manage the supplier contracts that simply fraud prevention can have.

4. Conclusion and future tasks

Present in this paper is forgery prevention and duplicate prevention for real estate transaction management system was designed. The system's design to real estate contracts, equity contracts, including offline procedures include forced to free trading contract

credibility in contract management services provided are P2P trust -based network on a contract to be managed because of the contract, the transparency was ensured. Future challenges include these system designs based on substantially real estate contract deal be established can allow a system to develop it.

References

1. Ching-Hsue Cheng, Chung-Hsi Chen, You-Shyang Chen, Ho-Long Guo and Chien-Ku Lin(2019), "Exploring Taiwanese's smartphone user intention: an integrated model of technology acceptance model and information system successful model", International Journal of Social and Humanistic Computing, Vol.3 No.2, 97-107, DOI: 10.1504/IJSHC.101591.
2. H. S. Park, J. W. Chung, and U. M. Kim, "A Study On Shared EMR(Electronic Medical Record By BlockChain(Ethereum)", Proceedings of KIIT Summer Conference, 436-437, December 2017.
3. Zhang C, Wang Q and Shi D(2016), "Scenario-based potential effects of carbon trading in China: An integrated approach", Applied Energy, 182, 177-190
4. Noor S, Yang W and Guo M(2018), “Energy Demand Side Management within micro-grid networks enhanced by blockchain", Applied Energy, 228, 1385-1398.
5. Zhou Y, Wu J and Long C(2018), Evaluation of peer-to-peer energy sharing mechanisms based on a multiagent simulation framework", Applied Energy, 222, 993-1022.
6. Sikorski JJ, Haughton J and Kraft M(2017), "Blockchain technology in the chemical industry: Machine-to-machine electricity market", Applied Energy, 195, 234-246
7. X. Xu, I. Weber, M. Staples, L. Zhu, J. Bosch, L. Bass, C. Pautasso, P. Rimba(2017), "A taxonomy of blockchain-based systems for architecture design", IEEE International Conference on Software Architecture (ICSA), 243-252.
8. Zibin Zheng, Shaoan Xie, Hong-Ning Dai and Weili Chen(2020), "An overview on smart contracts: Challenges, advances and platforms", Future Generation Computer Systems, Vol. 105, 475-491.
9. Ahmet Bugday, Adnan Ozsoy, Serdar Murat Öztaner and Hayri Sever(2019), "Creating consensus group using online learning based reputation in blockchain networks", Pervasive and Mobile Computing, Vol. 59, 111-125
10. R. James, "A Next-Generation Smart Contract and Decentralized Application Platform", "https://github.com/ethereum/wiki/wiki/White-Paper", 2017.

A Study for Analysis of Stock Price Information through Extraction of News Articles

Seok-Woo Jang[1][0000-0001-5580-4098] and Byeongtae Ahn[2][0000-0003-3431-9493]

[1] Dept. Of Software, Anyang University, 22, 37-Beongil, Samdeok-Ro, Manan-Gu, Anyang 430-714, South Korea.
swjang7285@gmail.com
[2] Liberal & Arts College, Anyang University, 22, 37-Beongil, Samdeok-Ro, Manan-Gu, Anyang 430-714, South Korea.
ahnbt@anyang.ac.kr

Abstract. For stock investors, news articles serve as an important basis for stock trading signals. In particular, the user who receives the information verifies whether the information is correct, and then buys and sells. At this time, the purchase and the sale often fail due to the delay of information verification. Therefore, in this paper, we conducted a news event extraction and researched a system that provides meaningful information through learning by extracting relevant stocks for 18 years of news data and stock price data, and classifying them into five levels. The system provides information using past stock analysis and real-time news and stock price capture data. In addition, it shortens investors' judgment by providing real-time news on the stocks in real time.

Keywords: *Stock Price, Stock Information, Stock Analysis, Deep Learning, Big Data*

1. Introduction

Domestic stock investment includes value investment, chart trading, and theme stock trading [1]. Among them, trading stock through news is called theme stock trading. In Korea, there are many cases of stock trading through the sale of theme stocks. However, these stock transactions are defined in just one sentence, making it difficult for investors to buy and sell at the right time. In particular, buying and selling through news is difficult to obtain accurate information in a timely manner, and even when accurate information is obtained, the timing of buying and selling is often missed due to lack of judgment by investors [2].

In the stock market, at least timing and quickness are directly related to the investor's profits. Therefore, in this paper, we designed a system through a verification machine in which information delivered to investors can be provided more quickly than other investors and few advanced information can be recommended in real time [3]. This system provides verification information more quickly than other investors so that investors can concentrate and invest.

About 3,000 to 5000 domestic news stories occur daily, and it is easy to miss timing to sell and buy stocks by analyzing these vast articles individually [4]. Therefore, the system extracts only meaningful information and provides a vast amount of news articles quickly. And it filters out spam news to provide pure stock-related news information.

The composition of this paper is as follows. In Chapter 2, the research trend for grasping the stock price information through the extraction of news articles is examined, and in Chapter 3, the system for stock information analysis is designed. Chapter 4 suggests conclusions and future tasks.

2. Related Studies

In this chapter, BIGKinds, KakaoStuck, and Google Alerts are systems that extract meaningful information using the news articles mentioned in the introduction. BIGKinds, which is the most used in Korea, is a new news analysis service made by integrating big data analysis technology into the largest article database including comprehensive daily newspapers, economic journals, regional daily newspapers, and broadcasters [5,6]. This service is provided free of charge and provides various services such as analysis data download and personalization service. The system consists of three categories: news collection system, analysis system, and storage system. News provided by domestic newspapers is classified into text, photo, and electronic document files, and stored in a news collection DB, a news collection system. Stored data is automatically classified by news category and extracted and standardized as key keywords in the news [7]. The standardized data is categorized as meaningful information by morpheme analyzer, entity name analyzer, and network analyzer in the article analysis module of the news analysis system and stored as metadata. This metadata is provided where users need it [8].

KakaoStuck is an app that provides stock information on your smartphone. This system allows you to quickly and easily order stocks of tradeable securities companies, and check stock quotes and hot issue information in real time. Basically, the news of domestic securities companies is crawled and only meaningful information is extracted and provided. To extract meaningful information, only specialized stock information is extracted and provided using deep learning of artificial intelligence. As the main service, real-time information is provided for the items of interest designated by the user, and the transaction details of participants are provided in real time. In addition, domestic and foreign stock index and exchange rate information is provided in real time, and chart-based supply and demand diagnosis and fundamental-based item diagnosis service are provided in real time [9,10].

3. Design of Electronic Voting System

In stock trading, meaningful information from news articles is very important to investors. Therefore, in this system, a system with four main functions was designed. First, it filters advertisements and advertising news from news articles. The filtering technique was performed using artificial intelligence to extract only meaningful information [11]. Second, it provides article analysis results in order to provide information on the items, industries, and themes related to news articles [12]. The analysis technology uses big data to divide categories into categories and extract only meaningful data for the data. Third, the results of seed analysis are provided to provide basic information on the item and other items showing similar patterns [13]. Fourth, it provides additional utilities such as stock calendar, news push of interest stock, and stock discussion community [14].

Fig. 1 shows the system configuration.

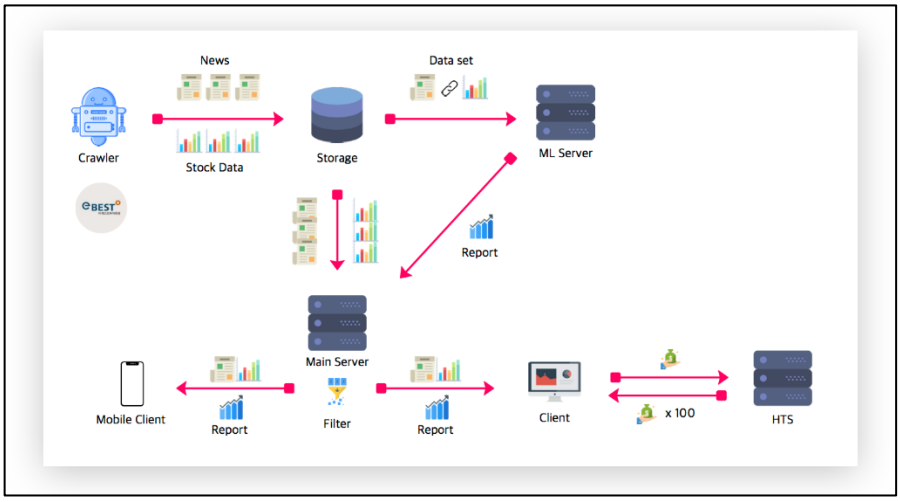


Fig. 1. Structure of System

This system crawls news and stock data in real time on the web and stores it in a storage system by a call back method. The accumulated news and data are provided to ML Server as a training data set, and an information model is generated through the stock information analysis module. The generated analysis result provides only the desired information to the client's request from the main server. Clustering that extracts only news content is classified using KoNLP and TF-IDF. In the classified news, stocks, industries, and themes are extracted using Elastic.

The development environment of this system was mainly developed in MacOS, and it can be accessed using Xing Api. The editor used Xcode, Phcharm, and the Python library used KoNlpy, NLTK, Annie, Gensim, etc. Elastic Search was used as a search

engine. The development language was developed with Cross, plaatform, desktop, and app using JS / Flow. The mobile app is written in native language and the server uses django / python / mypy, matching the core division and language.

Through this study, users can judge the importance and reliability of information faster than other investors through the verification machine and intuitive UI. And non-experts can easily buy and sell theme stocks. In addition, it provided various information that the existing products, Big Kinz and Kakao Stark, were not provided.

4. Conclusion and future works

The system designed in this paper increased the reliability and verification of information than the existing service, providing investors with quick judgment. In addition, the existing system measured the fluctuation of news and stock prices on a daily basis, but this system captures the news in real time and provides the stock price fluctuation of the news stock in real time. Future research projects should be practically developed based on these designs.

References

1. An Automated Framework for Incorporating News into Stock Trading Strategies. IEEE TRANSACTIONS ON KNOWLEDGE AND DATA ENGINEERING, VOL. 26, NO. 4, APRIL 2014
2. X. Tang, C. Yang, and J. Zhou. "Stock price forecasting by combining news mining and time series analysis," Proc. of the 2009 IEEE/WIC/ACM Int. Joint Conf. on Web Intelligence and Intelligent Agent Technology, 2009.
3. T. Fu, et al. "Discovering the correlation between stock time series and financial news," Proc. of Web Intelligence and Intelligent Agent Technology, 2008.
4. M.A. Mittermayer, and G. Knolmayer, "Text mining systems for predicting market response to NEWS," Proc. of IADIS European Conf. on Data Mining, 2007.
5. R. Schumaker, and H. Chen. "Textual analysis of stock market prediction using breaking financial news: The AZFinText system," ACM Transactions on Information Systems, vol. 27, no. 2, 2009.
6. Machado JAT , Lopes AM . Relative fractional dynamics of stock markets. Non- linear Dyn 2016;86(3):1–7 .
7. Chen YM , Wei YQ , Liu DY , Boutat D , Chen XK . Variable-order fractional nu- merical differentiation for noisy signals by wavelet denoising. J Comput Phys 2016;311(C):338–47 .
8. Mahato K . On the boundedness result of wavelet transform associated with fractional hankel transform. Integr Transf Special Functions 2017;28(2):1–12 .
9. Rahimkhani P , Ordokhani Y , Babolian E . A new operational matrix based on bernoulli wavelets for solving fractional delay differential equations. Numer Al- gorithms 2017;74(1):223–45 .
10. Nandal A , Gamboa-Rosales H , Dhaka A , Celaya-Padilla JM , Galvan-Tejada JI , Galvan-Tejada CE , et al. Image edge detection using fractional calculus with fea- ture and contrast enhancement. Circuits Syst Signal Process 2018;37(18):1–27 .

11. Boubaker H . A comparative study of the performance of estimating long-mem- ory parameter using wavelet-based entropies. Comput Econ 2016;48(4):1–39 .
12. Liu Q , Wang Y , Wang J , Wang QH . Optical image encryption using chaos-based compressed sensing and phase-shifting interference in fractional wavelet do- main. Opt Rev 2018;25(1):46–55 .
13. Gang X , Yu W , Zhang S , Xi C . Fractional domain singularity power spectrum. Nonlinear Dyn 2016;85(3):1–13 .
14. Srivastava HM , Shah FA , Abass R . An application of the Gegenbauer wavelet method for the numerical solution of the fractional Bagley-Torvik equation. Russ J Math Phys 2019;26(1):77–93 .

Solid particle dynamics by using Moving Particle semi-implicit Method

Kyung Sung Kim[0000-0001-7409-0895]

Tongmyong University, Busan Korea 48512
keiuskim@tu.ac.kr

Abstract. The particle method for computational fluid dynamics was dramatically increased its applicability on the both of engineering and sciences. The most important factor of its expansion is that less restriction of physical assumption because each particle carries own physical quantities including velocity, density and pressure. Due to its advantage, its applicability was expanded not only to fluid dynamics, but also solid particle territory. In this paper, one of the most well-known particle method, Moving Particle Semi-implicit (MPS) method was used to demonstrate solid particle behavior. In order to simulate solid particle phenomena, new models for particle interaction, which is invented and validated by author, was adopted. The soil penetration problem was simulated by newly developed MPS method for solid particle, and then its numerical results were compared to corresponding experiments. Through the comparison, it is shown that the results and tendency were well agreed between numerical and experimental results.

Keywords: Particle Method, Moving particle semi-implicit, solid particle, penetration

1 Introduction

Computational fluid dynamics (CFD) has been developed and expanded its application field dramatically based on the increased performance of computing technology and methodologies on the CFD method. The most impressive development on CFD is the establishing of fully Lagrangian approach such as particle method.
The particle method is following Lagrangian, thus all the physical quantities carries on the each particle. Among the particle method, smoothed particle hydrodynamics (SPH) and moving particle semi-implicit (MPS) method are well known. The SPH method firstly announced by Monaghan et al. (1994). According to Monaghan et al (1994), the density and pressure were measured by kernel function and equation of state. The SPH was very precise to simulate massive fluid movement, however, several pressure fluctuations were also observed. Xu et al. (2009) was improved pressure problems on SPH by using Poisson pressure equation (PPE). The MPS method was originally proposed by Koshizuka et al. (1996). Tanaka et al. (2010) improved MPS method by using multi-source terms in the PPE to suppress pressure fluctuations. Lee et al. (2011) developed MPS method by enhancing free surface searching method. Kim et al. (2014) expanded

MPS method to multi-phase flows and then it was developed for the solid particles at Kim et al. (20XX).

In this study, the newly developed MPS method for solid particle uses to simulate soil bed penetration problem which regard spud-can installation in the offshore industry. The numerical results were compared to the corresponding experimental results.

2 MPS for solid object

The governing equations of MPS method are the Continuity and Navier-Stokes' Equations as follows:

$$\frac{D\rho}{Dt} = 0 \tag{1}$$

$$\rho \frac{D\vec{u}}{Dt} = -\nabla p + \rho \nu \nabla^2 \vec{u} + \sigma \kappa \vec{n} + \vec{F} \tag{2}$$

where ρ is density, t is time, $\vec{u}$ is fluid velocity, p is pressure, σ is surface tension coefficient, κ is curvature for surface tension, $\vec{n}$ is normal vector of interface, $\vec{F}$ is external force including gravitational force, ∇ is gradient, and ∇^2 is Laplacian.

Although MPS is designed and widely used in fluid dynamics problems, its characteristic can be well matched with solid particle problems such as land sliding, dumping problems. However, for the MPS for solid particle requires to new model to alternate viscous and diffusion term which are not applied in the solid phenomena such as friction, drag force, collision and etc., which were not identified in fluid particle as follows:

$$\rho \frac{D\vec{u}}{Dt} = -\nabla p + \vec{F} + \vec{F}_f + \vec{F}_D \tag{3}$$

where $\vec{F}_f$ is friction force, $\vec{F}_D$ is drag force. Since drag force can be applied if only solid particle is positioned in the flow at least is influenced by fluid, else $\vec{F}_D$ is always zero.

Since solid particle does not have viscosity effect, the second term of Eq.(2) has been removed and replaced with friction term. Moreover, drag force can be considered when solid particle is on flow, thus drag force term were added into modified Naver-stokes' equation. Detailed information about each model will be announced.

3 Particle Interaction Models

Since the MPS method follows fully Lagrangian Approach, all differential terms require to be replaced with form of particle interaction. In this regard, the kernel function which measures effectiveness of neighboring particles to center particle was introduced. The kernel function used in this study as follows:

$$w(\mathrm{r}_{ij}) = \begin{cases} \left(1-\frac{r_{ij}}{r_e}\right)^3\left(1+\frac{r_{ij}}{r_e}\right)^3 & \left(0 \le r_{ij} < r_e\right) \\ 0 & \left(r_e < r_{ij}\right) \end{cases} \quad (4)$$

where subscription ij denotes $\left(r_j - r_i\right)$, r_i is position of particle i, and r_e represents effective range. In this study, r_e is set to 2.1 which is chosen by Lee at al. (2010). According to Eq. (4), the neighboring particle cannot affect to center particle which means the effectiveness is zero when the neighboring particles are located far from center particle.

Another key factor of MPS method to keep system continuum is particle number density. It is sort of density instead of fluid density for particle method. The particle number density can be calculated by using the kernel function as follows:

$$n_i = \sum_{j \neq i} w\left(|\,\vec{\mathrm{r}}_{ij}\,|\right) \quad (5)$$

For the solid particle simulation, the Navier-Stokes' equation was modified as Eq. (3). The third-term of right-hand-side of Eq. (3) denotes friction force on center particle. The friction force is kind of resistance force; thus, it matches to the viscous force. The friction force can be measured by acceleration including gravitational acceleration and friction coefficient.

For the acceleration, relative velocity between center and neighboring particle was required. And then acceleration of center particle for friction can be calculated as follows:

$$F_f^{(1)} = \sum_{j \neq i} \frac{1}{2} {l_0}^2 \left(\mu_i + \mu_j\right)\left[\rho_j a_{n,j} - \rho_i a_{n,i}\right]\vec{t}_{ij} \quad (6)$$

where l_0 is particle distance at initial arrangement, μ denotes the friction coefficient, a_n is acceleration of particle to normal direction of opposing particle, $\vec{t}_{ij}$ denotes the unit vector of friction force. The friction force direction can be decided according to the arrangement of particle with respect to the direction of movement as shown in Fig. 1(a).

Another component of the friction force is the accumulated weight of particle. As shown in Fig. 1(b), the particles over the center particle are involved in the calculation of accumulated weight. To avoid overestimation, the specific range of x-direction is introduced. When the neighboring particle over the center particle is out of the range of x-direction, it is excluded from accumulation effects. The accumulated weight can be calculated as follows:

$$p_{h,i} = \sum_{j \neq i} \rho_j g \quad if\ only\ z_j > z_i \quad and \quad x_i - r_f l_0 < x_j < x_i = r_f l_0 \quad (7)$$

$$p_{s,ij} = \frac{p_{h,i} + p_{h,j}}{2} \quad (8)$$

where $p_{h,i}$ is accumulated weight of stacked particles, $p_{s,ij}$ is static weight by averaging of accumulated weight between center and neighboring particle, r_f is effective range of center particle for friction, x *and* z is position of particle for x- and z-direction, respectively. With calculated accumulated weight, the friction force due to static weight can be calculated as follows:

$$F_f^{(2)} = \sum_{j \neq i} \frac{1}{2}\left(\mu_i + \mu_j\right) p_{s,ij} l_0^{\,2} \cos\theta_{ij} \vec{t}_{ij} \qquad (9)$$

where θ_{ij} is angle between center and neighboring particles.

By summation of Eq. (6) and Eq. (9), total friction force on center particle can be measured as follows:

$$F_f = \frac{l_0^{\,2}}{2}\left(\mu_i + \mu_j\right) \sum_{j \neq i} \left[\left(\rho_j a_{n,j} - \rho_i a_{n,i}\right) + p_{s,ij} \cos\theta_{ij} \right] \vec{t}_{ij} \qquad (10)$$

The friction force is resistance force of movement; it can only reduce or stop the object, not trigger the movement of particle.

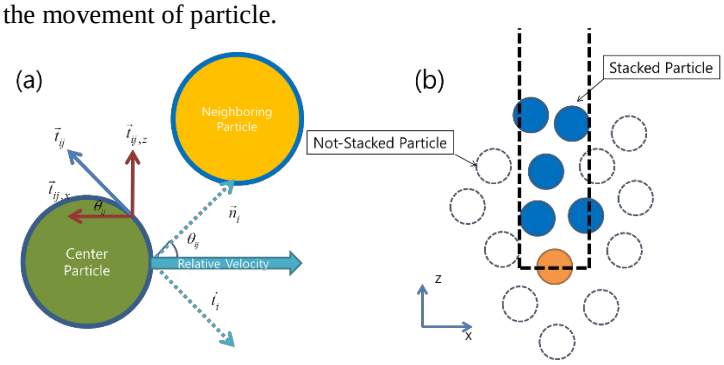


Fig. 1 (a) Schematic of contact model and (b) Accumulated weight of center particle

4 Application

In this study, the soil penetration problems was considered to simulate using MPS method for solid particle. Randolph et al. (2005) was selected as the corresponding experiment. The soil particles were captured container, and the spud-can moves to downward to penetrate soil particle. The numerical model is shown in Fig. (2). The object(Spud-can) moves to downward with 0.5m/s speed, which is same as experiment. The soil has physical phenomena as clay.

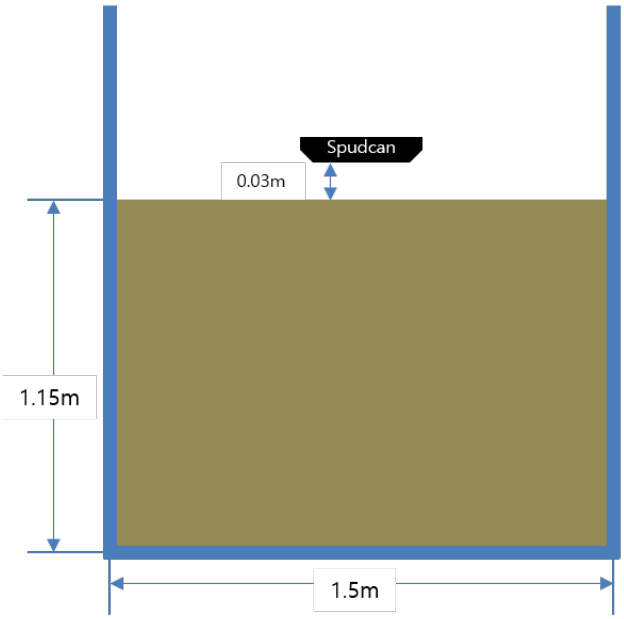
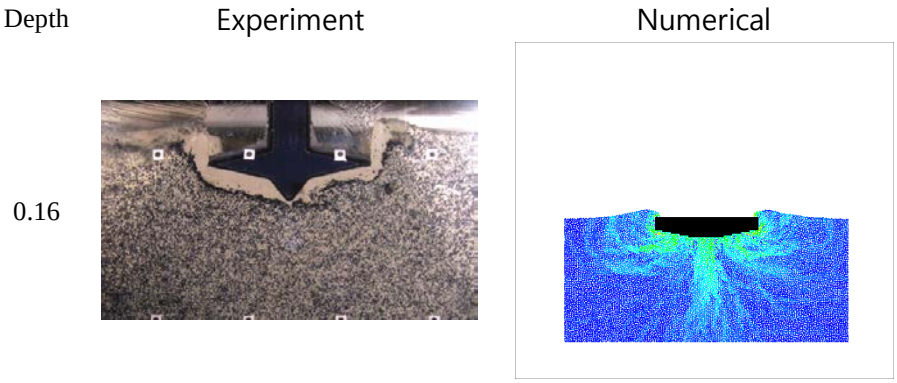


Fig. 2. Schematic model for numerical simulation

In the Fig. (4), the comparison of numerical and experimental results by using snap-shots. There were slightly difference between them, however it can be regard as 3-deimensional effects and soil consistency problems which can be ignored in the engineering sense.

The vertical loads on the spud-can measured and compared in fig. (5). Since the case of uninstallation (upward movement) was not considered at this time, only loads under installation was compared and it shows good-agreement between numerical and experimental results.



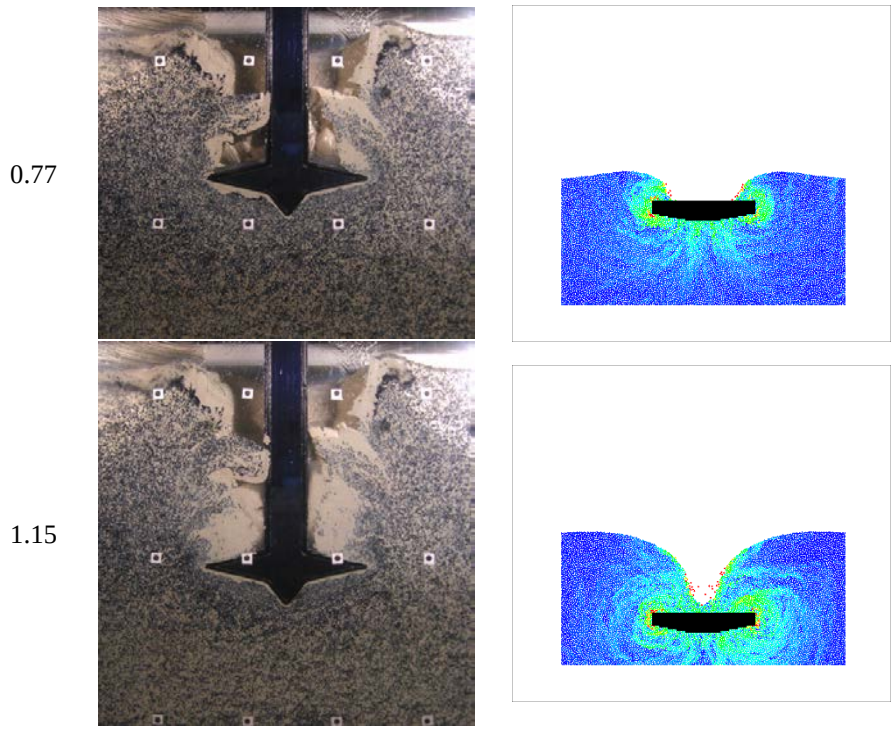


Fig. 3. Snashot comparison

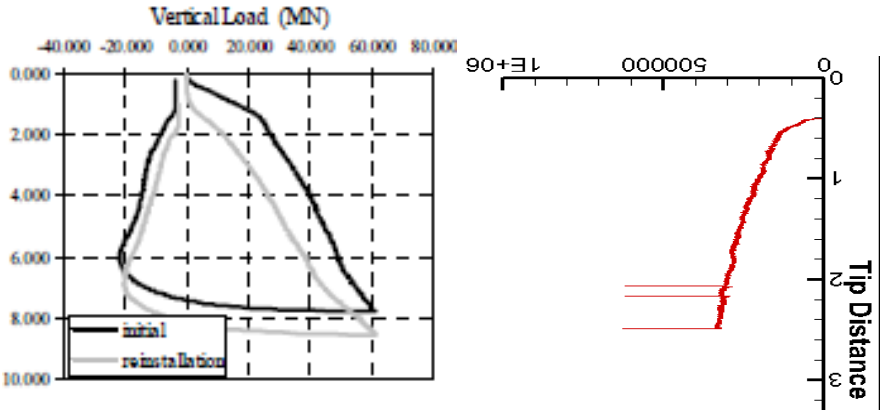


Fig. 4. Vertical load comparison

5 Conclusion

In general, Moving Particle Semi-implicit method is suitable to simulate fluid dynamics problem. An applicability of particle system, however, can extend its capacity of object not only fluid particle, but also solid particle due to its characteristics which is that particles can be located anywhere having physical properties including velocity, acceleration and pressure. Compare to fluid particle, solid particle have different characteristics, e.g. friction force instead of viscous effect. In this regard, governing equations of MPS have to be modified considering solid particle properties.
The newly developed MPS method program demonstrated soil penetration problem and then the numerical results were compared to the corresponding experimental results. There were slightly difference between numerical and experimental results including snapshot and vertical load, however, they were acceptable by considering soil particle consistency and 3-dimensional effects.
Through simulations and comparisons, it was seen that the applicability of particle system to not only fluid particle, but also solid particle such as gravel, sand and etc. The existence of angle of repose is significant difference which distinguishes particle whether solid or fluid particles, and it is clearly shown in this study. Comparison between two considered cases, broken dam and cargo dump problems, shows that angle of repose can be changed due to its kinetic energy and it is well-agreed with theory. The applicability of MPS for solid particle has been investigated and validated. Furthermore, this study can be a milestone to develop system to simulate solid particle behavior in the fluid flow which can be applied for sedimentation, bridge scour problem and etc.

Acknowledgement

This research was supported by Basic Science Research Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Education (NRF-2018R1D1A1B07048254)

References

1. Monaghan, J.J.: Simulating Free Surface Flows with SPH. J. Comput Phys, 110(2), 399–406 (1994). http://dx.doi.org/10.1006/jcph.1994.1034.
2. Xu, R., Peter S., Dominique L.: Accuracy and stability in incompressible SPH (ISPH) based on the projection method and a new approach. Journal of computational Physics 228.18 6703-6725 (2009).
3. Koshizuka, S., Oka, Y.: Moving-particle semi-implicit method for fragmentation of incompressible fluid. Nuclear science and engineering, 123(3), 421-434, (1996).
4. Kim, K. S., Kim, M. H., Park, J. C.: Simulation of Multiliquid-Layer Sloshing With Vessel Motion by Using Moving Particle Semi-Implicit Method.Journal of Offshore Mechanics and Arctic Engineering, 137(5), 051602. (2015).

5. Tanaka, M., Masunaga, T.: Stabilization and smoothing of pressure in MPS method by quasi-compressibility. J. Comput. Phy. 229(11), 4279–4290. (2010).
6. Lee, B.H., Park, J.C., Kim, M.H., Hwang, S.C.: Step-by-step improvement of MPS method in simulating violent free-surface motions and impact-loads. Comput. Methods Appl Mech Eng. 200(9), 1113–1125. (2011).
7. Kim, K.S., Kim, M.H., Park, J.C.: Development of Moving Particle Simulation Method for Multiliquid-Layer Sloshing. Math. Probl. Eng. 2014. dx.doi.org/10.1155/2014/350165. (2014).
8. Randolph, M. F., Cassidy, M. J., Gourvenec, S. M., Erbrich, C.: Challenges of offshore geotechnical engineering, Proc. 16th Int. Conf. Soil Mech. Geotech. Engng, Osaka, 1, pp 123–176. (2005).

Research on Remote Control of Self-propelled Lawn Mower

Wan-Rong Chen[1*], Tung-Shou Chen[1], Jeanne Chen[1], Fang Rong Hsu[2], Yu-Hsun Kuo[2]

[1] National Taichung University of Science and Technology, Taichung 404, Taiwan
[2] Feng Chia University, Taichung 407 Taiwan
*s1810732002@nutc.edu.tw

Abstract. Current lawn mower in the market are mainly driven by fuel engines, and users need to operate them at close range. Problems such as noise, dust and foreign object splashes generated when the machine is running are not fully protected. This poses a huge threat to people's health and well-being. In order to avoid the above-mentioned injuries, this study modified the self-propelled lawn mower. Initially, personnel were required to follow the machine for control and switch to remote control operation for weeding. In this study, the walking mechanism was changed to DC wheel motor drive. Using the RC remote control for control, the low-cost microcontroller reads its signal and controls the motor to act according to the signal. The time from receiving to reaction must be processed in real time, so use GPIO interrupt to achieve.

Research measurements have shown that the average noise of the lawn mower in operation is 108.84 dB. The amount of dust generated during mowing can reach $6.292 * 10^{-2}$ g/ cm^2. At a distance of 25 meters from the lawn mower, the noise is reduced to 65 dB, the amount of dust is almost to 0 g/ cm^2, and the threat of foreign objects flying over long distances is greatly reduced. In this study, the user's remote control distance can be at least 80 meters or more, so it can be proved that this method can effectively reduce the damage caused to users by noise, dust and foreign matter splashing during the mowing process.

Keywords: Remote Control farm Mower, health hazard, microcontroller, weed control.

1 Introduction

In the past, agricultural practices often used pesticides to remove weeds, but to regain environmental awareness, pesticides are considered to be one of the main factors that pollute the environment [1]. In addition to environmental pollution, many studies have found that pesticides may cause human health problems [2-7]. Taiwan ’s Taiwan regulations restrict the use of pesticides, so a variety of mowing machines appear on the surface of the city to remove weeds. Nowadays, the power source of the lawn mowers that cut long weeds is mainly fuel engines. Therefore, when the lawn mower is running, it will generate a lot of noise, dust, and foreign matter splashes. Users also need to operate the machine at close range. Although there are related wearable protective gears

available in the market, the mowing operation is often under the sun, and the protective gears are easy to cause sultry and uncomfortable, so most users will not wear protective gears perfectly. Therefore, it poses a great threat to the body and health of users who do not wear comprehensive protection. For example, excessive intake of dust may cause pneumoconiosis [8], noise may easily lead to permanent hearing loss [9], and foreign objects hit at close range The body will cause immediate damage.
Therefore, in order to avoid the above problems, the target of this study is to transform the self-propelled lawn mower, using compact electronic parts to transform it, to open the distance between the user and the machine, noise, dust and foreign matter splash, etc., so that users can Operate in a fairly safe environment.

This study observed the operation of self-propelled lawn mowers and found that users mainly control the direction and stop of self-propelled lawn mowers. These actions are mainly reflected on the steering wheel. In this study, the walking mechanism of the original self-propelled lawn mower was replaced with a hub motor. Because the wheel motor can achieve relatively small movements, and can simplify the control method and mechanism, it is very conducive to on-site control. And through the remote control RC (Radio Communication) 2.4G remote control to send the signal to the receiver, and then use the microcontroller Arduino to analyze the signal, and then let the H-bridge circuit control the motor according to the signal. In addition to improving safety, this method also makes the operation of the lawn mower easier.

2 Materials and Methods

2.1 Mower Specifications

In this experiment, the self-propelled hammer cutter (SF-660-A2, TASOON, Taiwan) was used for transformation. The mowing mechanism uses a 10-horsepower fuel engine as the power source. Because the torque of the fuel engine is larger than that of the motor, it can break thick branches or long weeds.

Table 1. Self-propelled lawn mower specifications.

Lawn Mower	TASOON SF-660-A2
Outline Dimension	L*W*H: 1650*840*950 (mm)
Job Width	660 (mm)
Number Of Claw Hammers	38 (pcs)
Weight	145 (kg)
Knife Shaft Speed	2600 – 3000 (rpm)
Working Height	0 – 70 (mm)
Engine	10 horsepower
Maximum Horsepower	13.0 (ps) / 2000 (rpm)
Weight	38 (kg)

2.2 System Structure

In this study, the remote control lawnmower keeps the engine as the power source, so it is easily interfered by the engine when the signal is transmitted. Therefore, it is necessary to select the RC (Radio Communication) 2.4 GHz remote control receiver with high interference resistance. The RC receiver uses a set of 8 channel 2.4 GHz DMSS (XG8, JR PROPO, Japan) for use.

The RC controller sends signals to the receiver, such as throttle, front, rear, left and right. The receiver will output the corresponding PWM signal corresponding to the joystick of the remote control. The control part uses the microcontroller Arduino UNO to analyze the PWM signal sent by the RC receiver. Then output to the H bridge according to the control signal, so that the drive motor rotates in its specific direction.

The microcontroller Arduino UNO board is based on the ATmega328P chip and has a 16MHZ quartz oscillator [10]. Responsible for controlling the signal of the remote control receiver and the actuation of the motor. The motor of two wheels system is powered by 24V which is connected in series by two 12V lead-acid batteries (55B24L, YUASA, Taiwan). And then converts 24V to 9V through the buck module for the power of microcontroller Arduino board. The system diagram is shown in **Fig. 1** below.

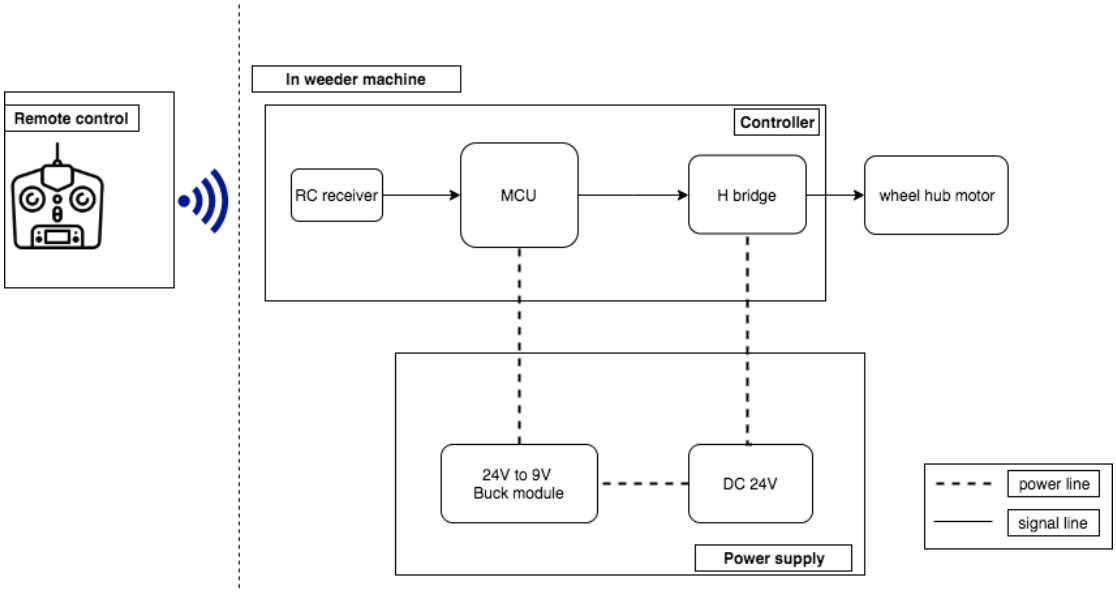


Fig. 1. Architecture of the remote control system.

2.3 Real-Time System

The remote control receiver used is a multi-channel signal output, and the signal sent from the remote control must be analyzed to determine the current remote control sta-

tus. However, the multi-channel cannot be processed using the polling method, because it will cause the judgment time to be too long, causing the machine to respond too slow, and the machine may fall into the ditch and hit obstacles.

The operating frequency of the RC receiver is 50Hz. In order for the system to get an immediate response, the signal must be processed before sending the next signal. The signal transmission frequency of the receiver is very high. If you poll the processing speed, it is too late to reflect the operation in time. Therefore, the program uses GPIO interrupt plus timer interrupt to process the signal, as shown in **Fig. 2**. Switching to interrupt processing for one second will process 100 signals, ensuring that all receiver signals can be processed for real-time control.

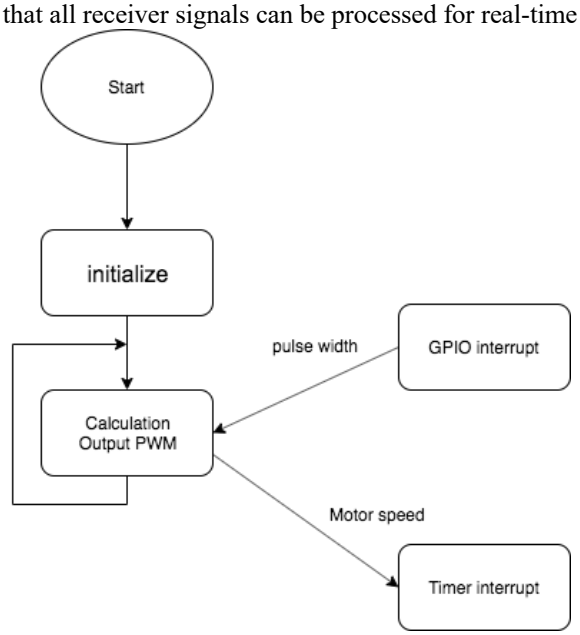


Fig. 2. Flowchart for the remote control farm mower

3 Results and discussion

3.1 Noise and dust measurement data

In order to know that the distance between the remote control mower and the user is safe, this chapter measures the noise and dust when using the mower to understand its actual situation.

In the noise experiment, the distance between the machine and the user is designed to be measured at a distance of 0, 5, 10, 15, 20, and 25 meters. Record 10 noise levels and average the data results. The measurement data results are shown in **Table 2.** Distance and noise measurement results (dB) below.

Through multiple measurement data to analyze the obtained results, the average measurement distance between the user and the machine at 0 meters is 108.8 dB. When

the noise exceeds 85 dB, it will already cause hearing loss [11]. According to the NIOSH safety guidelines, exposure to 95.0 dB should not exceed 47 minutes without any hearing protection equipment [12].

As shown in **Fig. 3.** Machine noise and distance as the distance increases, the decibel number of the noise will be greatly reduced. When the distance reaches 25 meters, the amount of noise has been reduced to 65 dB, allowing users to obtain a more comfortable operating environment. It can be confirmed that the farther the user is from the machine, the smaller the noise problem when the engine is running.

Table 2. Distance and noise measurement results (dB)

Distance (m)	Measurement Data										Average
0	99.5	117.5	108	100	113	99	115.4	114.3	105.4	116.3	108.84
5	80.3	80.1	80.6	85.1	81.4	82.3	81.6	82.9	81.7	81.1	81.7.1
10	78.6	78	78.1	78.9	78.4	78.2	78	79	77.7	78.6	78.35
15	74.1	74.3	74.8	74	74.6	73.9	78.3	73.4	74	73.7	74.51
20	69.7	68.7	68.2	68.3	68.4	71.6	70.7	68.5	68.1	69.1	69.13
25	65.2	65.4	65	65.3	65.9	65	65.1	65.4	65.5	65.9	65.37

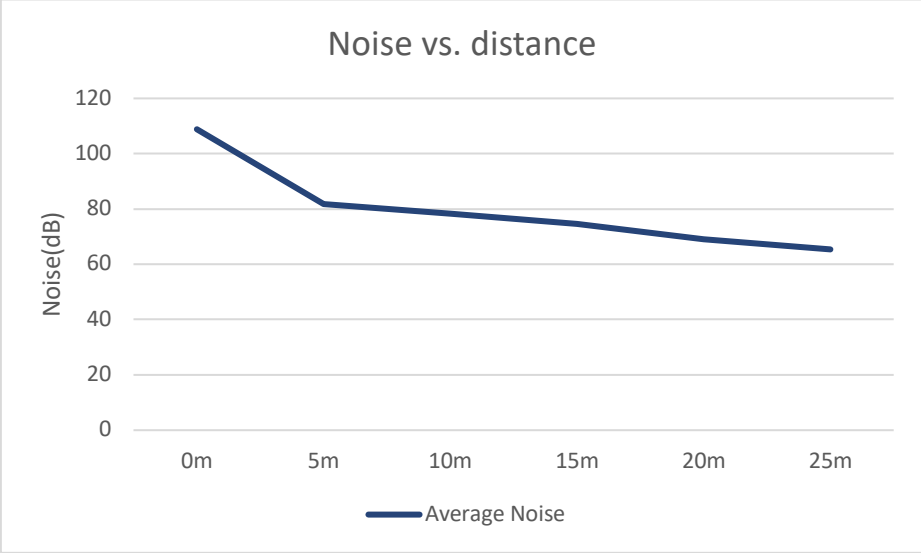


Fig. 3. Machine noise and distance

Dust reduction experiment, select different distances from the machine to measure dust reduction. Place 10 × 10 cm aluminum foil paper on the soil surface respectively at 0, 1, 2, 3, 4 and 5 meters. Put the cutting depth of the machine into the soil table about 1 cm, and then run the machine in place for 5 minutes. The measurement results obtained are shown in Table 3.

When measuring, no wind is flowing. Therefore, dust is concentrated near the machine, and the minimum measurement result at a distance of 3 meters is close to 0 grams. The maximum amount of dust dropped at 0 meters is 6.292 * 10^{-2} g/cm^2, which is a considerable amount. It usually takes more than 5 minutes to cut the grass, so the amount of dust must be more terrible. Inhaling too much dust may also cause lung disease.

Table 3. Dust measurement results(g/cm^2).

Distance (m)	Measurement Data			Average
0	1.999	6.292	0.536	2.942
1	0.02	0.011	0.012	0.014
2	0.001	0.004	0.003	0.002
3	0	0.003	0.006	0.003
4	0	0.003	0.01	0.004
5	0	0.004	0.004	0.002

When the distance exceeds 3 meters, the dust is almost 0, and the noise must reach 25 meters to be within the safe range. Experiments show that the longer the distance, the less noise and dust damage. This study used an RC remote control receiver. The remote control distance of the measured equipment can reach more than 80 meters. Since human vision can no longer clearly see the state of the lawnmower, it is easy to cause danger. The mower should be kept within a distance of 40 meters to ensure the current mowing conditions of the machine are safe.

3.2 Remote mower

The control board that controls the walking mechanism is installed behind the engine. The overall view of the remote control mower is shown in **Fig. 4**.

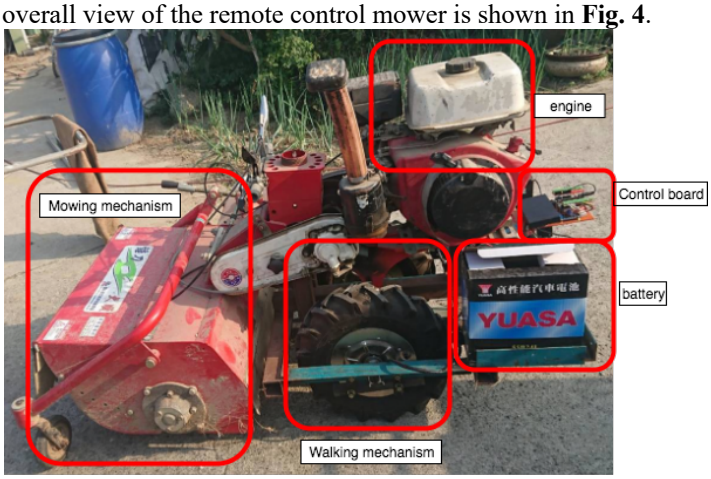


Fig. 4. Complete drawing of remote control mower.

4 Conclusions

This study solves the problems of noise, dust and foreign objects caused by self-propelled lawn mowers that must be operated at close range. Use the remote control to keep users away from the machine to reduce risk. Remote control is achieved through low-cost electronic components. The experimental results prove that the lawn mower can be operated remotely, which greatly reduces the risk of weeding, such as noise and dust decrease with distance.

The RC 2.4G receiver can be controlled remotely, allowing you to maintain a safe range when operating the lawn mower. It not only increases the safety of the user when weeding, but also reduces the fatigue when weeding.

References

1. Mostafalou , S., Abdollahi, M.: Pesticides and human chronic diseases: Evidences, mechanisms, and perspectives. Toxicology and Applied Pharmacology 268(2), 157–177 (2013).
2. Saad-Hussein, A., Noshy, M., Taha, M., El-Shorbagy, H., Shahy, E., Abdel-Shafy, E. A.: GSTP1 and XRCC1 polymorphisms and DNA damage in agricultural workers exposed to pesticides. Mutation Research/Genetic Toxicology and Environmental Mutagenesis 819, 20–25 (2017).
3. Bolognesi, C., Merlo, F.D.: Pesticides: Human Health Effects. Encyclopedia of Environmental Health. 2nd edn. Elsevier (2019).
4. Benedetti, D., Nunes, E., Sarmento, M., Porto, C., Santos, Eliete Iochims dos Santos, C., J. F., da Silva, J.: Genetic damage in soybean workers exposed to pesticides: Evaluation with the comet and buccal micronucleus cytome assays. Mutation Research/Genetic Toxicology and Environmental Mutagenesis 752(1-2), 28–33 (2013).
5. Smit, L.: Respiratory Diseases in Farmers. Reference Module in Biomedical Sciences. Elsevier (2019).
6. Singh, S., Kumar, V., Thakur, S., Dev Banerjee, B., Chandna, S., Singh Rautela, R., Sunder Grover, S., Singh Rawat, D., Tazeen Pasha, S., Kumar Jain, S., Lal Ichhpujani, R., Rai, A.: DNA damage and cholinesterase activity in occupational workers exposed to pesticides. Environmental Toxicology and Pharmacology 31(2). 278–285 (2011).
7. Li, Z.: The use of a disability-adjusted life-year (DALY) metric to measure human health damage resulting from pesticide maximum legal exposures. Science of The Total Environment 639, 438–456 (2018).
8. Baur, X., Sanyal, S., Abraham, J.: Mixed-dust pneumoconiosis: Review of diagnostic and classification problems with presentation of a work-related case. Science of The Total Environment 652, 413–421 (2019).
9. Aung K.Zaw, Aung M.Myat, MyaThandar, Ye M.Htun, Than H.Aung, Kyaw M.Tun, Zaw M.Han,: Assessment of Noise Exposure and Hearing Loss Among Workers in Textile Mill (Thamine), Myanmar: A Cross-Sectional Study. Safety and Health at Work (2020).
10. Arduino UNO, https://datasheet.octopart.com/A000066-Arduino-datasheet-38879526.pdf, last accessed 2020/03/16.

11. O.Ahmed, H., H.Dennis, J., G.Ballal, S.: The accuracy of self-reported high noise exposure level and hearing loss in a working population in Eastern Saudi Arabia. International Journal of Hygiene and Environmental Health 207(3), 227–234 (2004).
12. Ibraheem Jaafar, N., Khairi Md Daud, M., Mohammad, Normastura Abd Rahman, I.: Noise-induced hearing loss in grass-trimming workers. Egyptian Journal of Ear, Nose Throat and Allied Sciences 18(3), 227–229 (2017).

Application model research of visual focus on exhibition item in museum.

Min-Feng Lee[1]

[1] National Museum of Natural Science, 1, Guancian Rd., Taichung, ROC
National Museum of Natural Science
antonio@nmns.edu.tw

Abstract. Most people who want to understand the visit process and visual experience of museum visitors can only be done through interviews and questionnaires. In this model research, the visual trajectory data of the visitors will be used to collect and analyze the structure research and planning of the cognitive process of the museum visitors' exhibits. This research will use a mobile eye tracker to record and analyze the history of visitors viewing exhibits, and provide an application model for evaluating the perception of visitors in museums or various exhibits in the future.

Keywords: Visual experience, eye tracker, affective computing.

1 Introduction

A large part of the information received from people participating in museum visits is based on vision. The movement track of the visitor's sight can reflect the cognitive process of people's visit. Through these visual processes, it integrates affective calculation analysis and expands eye tracker research. In the realm of museums. In order to explore cognitive features such as information selection patterns. With the rapid development of smart computing, more and more researchers realize that the use of various sensing devices can facilitate the acquisition of monitoring information, and use eye trackers to record the various eye movement indicators of visitors, understand their visual search methods, and understand It responds to differences in actions. In recent years, the continuous development of Eye Tracker technology has made it convenient to obtain real-time data of visual overall information in various research fields, and use the collected data to explore the knowledge exploration and cognitive behavior of visitors. With the continuous advancement and maturity of the technology of the eye tracker, various manufacturers have begun to improve the eye tracking device in all aspects, such as: price, eye comfort after wearing, and accuracy of scanning eyeballs, etc. All of them have made great progress, and in response to various research needs, they have developed matching packages, computing software and visualization software to provide researchers with data analysis and use, and also created a new field of eye tracker research[1].

2 Related Work

2.1 Background and Motivation

Eye tracker is a kind of eye tracking technology equipment that collects eye movement and other related information for eye tracking, and is used to measure the moving direction and staying position of the observer's gaze. When people observe objects, they will keep their eyes Move to achieve three modes of gaze behavior: Fixation, Saccade and Pursuit[2].

At present, eye trackers have gradually been applied in different research fields, as follows:

A. Interface interaction: The interactive design of the human-machine interface allows people to operate the computer mouse and keyboard through the movement of the eyeball into another type of input device, which is also provided to more groups who need multiple input methods.
B. Visual Cognitive Science: understand the cognitive model of learners or readers through eye movement, understand the cognitive style of the observer through the visual reading process, and use eye movement to understand the emotional state of the observer [3], and There are related studies that apply related visual cognition technology to the museum field[4].
C. Limb movement measurement: Through the eye movement index measurement during limb movement, researcher can understand the relevant visual search information and related reaction actions when the athlete is performing limb movement, so as to understand the movement mode and perform subsequent sports training improvement modes.
D. Interface information browsing: record the viewing of commercial information and advertisements by shoppers or pedestrians on the road, study whether the display method of commercial information attracts the attention of consumers, and analyze the displayed media and presentations What are the impacts of classification and performance on consumers?

An introduction to the three-common eye-tracking behavior modes:

A. Fixation is the behavioral purpose of making the fovea of the eye align with the observed object. When gazing is observed, the eye is not completely motionless. From the research of various eye trackers, three kinds of weak movements are found, including irregular ones, Slow visual axis changes-drift, high frequency, low amplitude visual axis vibration-nystagmus (tremor), and when staring at a certain area of the still life for more than a certain time, or because the starting point is not When the sight information is received by autonomous drifting, tiny, involuntary saccades will be produced.
B. The visual observation mode of saccade means that when viewing an object, the gaze of the eye will stay at a certain position in the visual observation object for a while, wait for the gaze mentioned in the previous paragraph, and then switch to another position to perform the new object position. Fixation, this process, the process of eyeball jumping is called saccade.

C. Pursuit, this eye behavior mode is when people keep their gaze on a moving object and follow the direction of the object to move, such as staring while watching the train moving in a car. This kind of eye observation the process of moving is called Pursuit.

3 Visual focus on exhibition item in museum

3.1 Research design

In this research plan shown in Figure 1, it is mainly divided into three steps, as follows:
Stage I is to set the observation objects to be selected according to the exhibition objects of the Science Museum, provide visitors with a visit design, conduct a visit experiment for the content of the exhibits viewed by the visitors, and record the Fixation, pupil size and other related information to facilitate subsequent feature selection. According to the defined observation mode, evaluate whether the trajectory of visitors under different conditions is suitable for the experiment.
Stage II is data collection. After confirming all the contents of the visit, a large amount of eye movement data program is collected. The program selects testers and collects eye movement data according to different observation objects. After selecting effective gaze point features, training data of different dimensions is established through data normalization.
Stage III is result verification and analysis. Classifier training and training result evaluation are carried out on different dimensions of data. The cross-validation analysis program finds out the appropriate data dimensions, and finally verifies the classification results of the test data and conducts in-depth discussion and analysis of the results. .

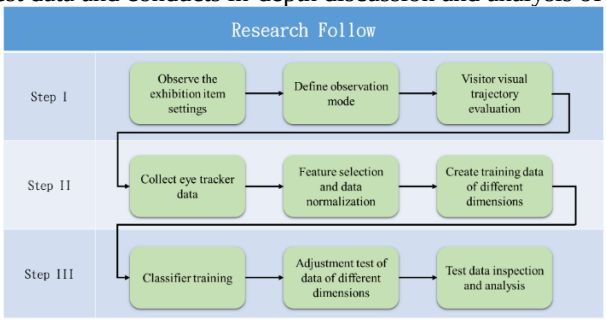


Fig. 1. Visual focus on exhibition item in museum research follow.

4 Result & Discuss

In the framework of this research, data collection and analysis are still being continuously expanded. The current analysis data is as follows. Figure 2 for a single exhibition item is shown in Figure 2. Figure 2 shows the collected visitor's gaze trajectory focus analysis and calculation results. The visual trajectory of different testers is analyzed and

predicted by the RNN improved mode LSTM (Long short-term Memory) in artificial intelligence to predict the visual focus of visitors. As shown in Figure 3, after 60 drops in the calculation mode, it has reached 97.8 accuracy. Accuracy rate. It is also expected to expand the application of visitors' visit patterns in follow-up research, record the complete visit process, and conduct predictive analysis.

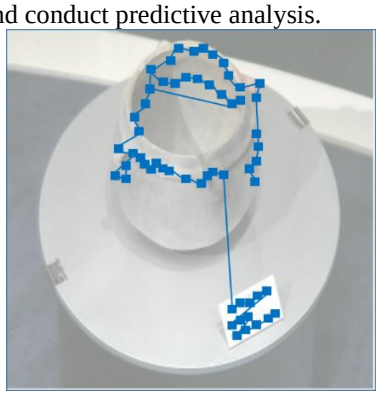


Fig. 2. Focus map of visitors' visual trajectory.

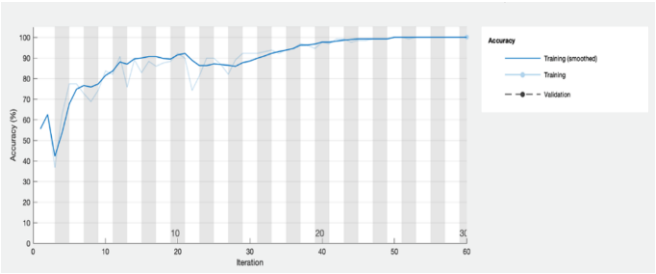


Fig. 3. Visual trajectory recognition training chart with LSTM.

References

[1] R. Kothari, Z. Yang, C. Kanan, R. Bailey, J. Pelz, and G. Diaz, "Gaze-in-wild: A dataset for studying eye and head coordination in everyday activities," *arXiv preprint arXiv:1905.13146*, 2019.

[2] B. B. Velichkovsky, N. Khromov, A. Korotin, E. Burnaev, and A. Somov, "Visual Fixations Duration as an Indicator of Skill Level in eSports," *arXiv preprint arXiv:1906.01699*, 2019.

[3] J. De Lemos, G. R. Sadeghnia, Í. Ólafsdóttir, and O. Jensen, "Measuring emotions using eye tracking," in *Proceedings of measuring behavior*, 2008, vol. 226, pp. 225-226.

[4] T. Toyama, T. Kieninger, F. Shafait, and A. Dengel, "Museum guide 2.0-an eye-tracking based personal assistant for museums and exhibits," in *Proc. of Int. Conf. on Re-Thinking Technology in Museums*, 2011, vol. 1, p. 24.

Determinants of Competitiveness in Immersive Technologies with Sensor Networks

Byun Jeongeun[1][0000-0002-2563-4443], Hong Jae-pyo[2], Byun Jeoung-woo[3]

[1] Technology Commercialization Center, Korea Institute of Science and Technology Information (KISTI), 66 Hoegi-ro, Dongdaemun-gu, Seoul, 02456, Korea
[2] Science, Technology and Innovation Office, Ministry of Science and ICT, 194, Gareum-ro, Sejong-si, 30121, Korea
[3] Department of Hotel Management, Kyung Hee University, 26 Kyungheedae-ro, Dongdaemun-gu, Seoul, 02447, Korea
jwbyun@khu.ac.kr

Abstract. In this paper, we analyze the importance weights of the factors that influence the national competitiveness of immersive technologies with sensor networks in Korea, based on Porter's diamond model as well as through analysis of opinions from experts of related fields. We find that the factor influencing the national competitiveness of immersive technologies is technological competitiveness of information and communications technology (ICT), followed by domestic market and business performance. In order to improve the national competitiveness of immersive technologies, it is required to establish the mechanism to expand technology commercialization that embraces firms as the player of innovation, technologies as the object of innovation, and environment including market.

Keywords: Immersive Technologies, Sensor Networks, National Competitiveness, Diamond Model

1 Introduction

Various immersive technologies with sensor networks, representative examples of which include augmented reality (AR), virtual reality (VR), and mixed reality (MR), have developed beyond the research and development (R&D) stage and many are now positioned far closer to the market. The purpose of this study is to generate strategic priorities to secure national competitiveness in the immersive technologies with sensor networks, which will drive future growth. Based on the diamond model developed by M. Porter (1990) [1], this study collects views of experts to analyze the relative importance of the various factors that determine the national competitiveness of immersive technologies in Korea.

2 Analytical Method

This study analyzes relative importance of the factors that determine the national competitiveness in immersive technologies, based on Porter's diamond model as well as with considering competitive advantages of ICT, which underlie the immersive technologies, as the nation ranked 1st in ICT adoption pillar in the 2018 Global Competitiveness evaluated by World Economic Forum (WEF). The analytic hierarchy process (AHP) designed questionnaire survey was conducted for 17 experts of technology, market, and policy from academia, industry and government funded institutes that are related to immersive technology. As a result of consistency ratio test, answers from 15 respondents excluding two who lacked in consistency were geometrically averaged and applied to the analysis model.

3 Results

Our main results are reported as follows: First, related and supporting industries is the most important factor among four major determinants. It suggests that strategic investment not only in immersive technologies, but also in ICT, is important in order to reinforce the national competitiveness of immersive technologies. Second, the fact that among the 12 sub-factors, technological competitiveness of ICT, domestic market, and business performance were led to three major factors that influence the national competitiveness of immersive technologies. Third, it is presumed that, among the 12 sub-factors, that labor was determined to be relatively less important factor than anything else is derived from the technology-intensive characteristics of high technologies.

4 Conclusion

In this study, we suggest the following implications. First, to improve the national competitiveness of immersive technologies, it is regarded that the development of ICT based technologies is more important than anything else. Second, since immersive technologies are characterized by end-user friendliness, it is necessary to establish the mechanism that will lead technological innovation to consumer markets to facilitate the national competitiveness. Finally, it is presumed that the reason why the factor labor got the lowest importance reflects experts' expectations that the growth of technology and market based on ICT infrastructure would drive workforce to growth in turn. However, it is not desirable to interpret based on the results that labor is not important in facilitating national competitiveness of immersive technologies.

References

Porter, M.: The Competitive Advantage of Nations. The Free Press, New York (1990).

Using Computer-based Feedback Help Student Nurses Improving Accuracy of Performing Clinical Operations

Ying Geng[1] and Po-Sen Huang[2]

[1] Library and Information Center, Tajen University, Pingtung 90741, TW
[2] Department of Nursing, Tajen University, Pingtung 90741, TW
phuang@tajen.edu.tw

Abstract. Nurses are the ones who activate the clinical care to patients, and are the key factor influencing the implementation of following clinical interventions. However, novice nurses often feel not confident on providing clinical judgments and making decisions, which may prevent novice nurses on providing effective and high quality clinical care that a patient expects. As a powerful tool assess student nurses' clinical skills and prepare them to fit into future workplace, OSCE is approved to be essential but not applicable in nursing classroom. Moreover, student nurses are not able to receive timely feedback to improve their clinical performance in the traditional OSCE. The present study would be a pilot study explores whether student nurses could utilize the electronic feedback to improve their clinical performances after they complete the OSCE. The results showed a positive effect on student nurses' clinical performance, which suggested that student nurses who received immediate electronic feedback could perform clinical performance more accurately at the end of the semester.

Keywords: Computer-based Feedback, EOSCE, clinical performance.

1 Introduction

Nurses are the ones who activate the clinical care to patients, and are the key factor influencing the implementation of following clinical interventions (Pantazopoulos, et al., 2012). The cases that a nurse encounters in the workplace are often complex, varied and situated, which requires a nurse to evaluate patients' clinical condition and apply accurate clinical judgments and to articulate the judgments clearly to other clinicians in order to provide appropriated interventions to patients (Benner, 2012). The society posts increasingly high demand toward nurses on expecting a novice nurse to be quite an expert on their first day to serve the patient, to engage in nursing practice and to offer effective health care. Such a demand, in the meantime, requires the nursing education to help student nurses be fully equipped with high-level of clinical skills in before they encounter any comprehensive and complex situation in the future workplace (Turner & Dankoski, 2008; Setyonugroho et al., 2015). And hence expecting those student nurses be an active nurse who could engage in the real cases effectively on their first day to work. Novice nurses, however, often feel not confident on providing clinical judgments and making decisions, which may prevent novice nurses on providing effective and high quality clinical care that a patient expects (Gorman, 2018).

2 Literature Review

2.1 A call for nursing curriculum reform

While the society posts increasingly high-expectations toward nurses, the nurse education in reality faces great challenges that student nurses who receive knowledge from the traditional nursing pedagogies may have difficulties to connect their knowledge to the real cases in their workplace (Giddens & Brady, 2007), and hence, there is a call for developing and implementing alternative pedagogies to help student nurses be better equipped with high-level of skills in school (Setyonugroho et al., 2015). The Objective Structured Clinical Skills Examination (OSCE) was introduced into nursing curriculum as an essential part to help student nurses to integrate all functional knowledge they have learned in classroom and put it into practice with simulated patients under supervision of nurse experts. OSCE, first established in 1975 and developed over decades, is now a well-established method of offering clinical learning environment and assessing clinical competence for pre-registration nursing professionals. More importantly, OSCE provides an objective feedback for students to see their actual performance before they are moving to the practice setting. Such feedback is valuable to pre-service nurses to overcome the weaknesses and to reinforce the strengths (McWilliam & Botwinski, 2010), which eventually will make those nursing students be more well-prepared and competitive in future.

Nursing OSCE offers a controlled stimulated working environment and consists a number of stations. The design of OSCE allows students to perform a broad range of clinical skills in a controllable simulated environment and get evaluation and feedback from professional nurses who have to deal the real cases in the workplace, but as students have limited time to perform clinical skills in each station, in return, the professional examiner in that station, has limited time to observe the procedure and to score the performance. When student nurses move to another station, they have to face a different set of professional examiner and trained patients. Even with a proper training and a rubric, there might be slightly difference and variations due to human errors (Tredwell, 2006). There is a strong need to transfer the traditional OSCE to an affordable method, such as EOSCE, that could be implemented into daily classroom.

2.2 Developing a computer-based timely feedback

Despite the scarcity of existing literature addressing the use and the effectiveness of electronic OSCE or similar software as a replacement to the traditional paper-based OSCE, researchers tend to agree that examination efficiency and timely individualized feedback can be achieved by employing a computerized system (Snodgrass et al., 2014). Studies showed EOSCE can effectively reduce errors and post-exam workload following by transferring paper-based scores into electronic files (Treadwell, 2006, Schmitz et al., 2011). In an electronic OSCE, with an electronic scoring system, it enables examiners to upload feedback automatically after scoring student nurses' performance, and hence it is possible for student nurses to receive immediate feedback toward their performance which could possibly reinforce their memories on performing clinical procedures and help they correct the mistakes they made during the practices (Snodgrass et al., 2014). In a study conducted across two consecutive years on comparing

quantity and quality of feedback between paper-pencil checklist and iPads scoring system, the electronic scoring system lead to an increase on the quantity of feedback (Munro et al, 2018). If student nurses could receive timely and comprehensive feedback, there would be a large chance that the feedback would influence student nurses on performing clinical skills. However, there is not much identified evidences regarding to the quality of the computer-based feedback to improving student nurses' skill practice and the effectiveness of using a computerized system feedback for student nurses to improve the accuracy of performing clinical skills (Snodgrass et al., 2014; Ashby et al, 2016).

The present study would be a pilot study explores whether student nurses could utilize the electronic feedback to improve their clinical performances after they complete the OSCE.

3 Methods

3.1 Participants

Participants from this study were 42 college nursing students. All of them were in college junior class and recruited from a Nursing Skills & Implementation course at one university located in southern Taiwan. Age ranged from 20-22. All of them were required to take nursing OSCE by the end of the course.

3.2 Experimental Procedure

The study conducted in Nursing Skills & Implementation course lasted 18 weeks. In each week, participants received a 2-hour class for clinical skills training. During the first 6 weeks, students were asked to be familiar with the content knowledge. Started from week 7 to week 14, students were asked to practice OSCE training sections including special mouth care, bed bath, bedpan use, heat and cold therapy, perineal care, nasal gastric tube feeding & care, finger sugar test, oxygen therapy, steam inhalation, and wound care. All students were required to take OSCE during week 15-16. During the OSCE weeks, 20 students were randomly assigned to the traditional OSCE group (Group T), in which examiners used paper and pencil scoring checklist, whereas the other 22 students were in the EOSCE group (Group E), in which examiners used an iPad app to score students' performance and the score will be automatically stored once the examiner complete scoring. After the OSCE, students in Group T were asked to self-reflect their performance during the OSCE, whereas students in Group E were allowed to login and check their scores and feedback on the apps. In week 18, all 42 students have to perform 5 randomly selected clinical procedures as their finals.

4 Results and Discussion

The present study interested in exploring whether student nurses could utilize the electronic feedback to improve their clinical performances in following practices. The independent t-test was conducted to assess students' prior knowledge on to-be-learned

clinical performances (Table 1) and showed show no significant difference between the Group E and Group T (t= .59, p=.56>.05) existing in prior knowledge between two groups. Therefore, students in both group do not have difference before they receive the semester-long clinical skills training.

Table 1. t-test of the two group on the pre-test results

Group	N	Mean	S.D.	t(42)
Group E	22	6.27	10.64	-1.20
Group T	20	4.50	8.72	

*p<.05

The independent sample of analysis of covariance was conducted to evaluate students' clinical performance after they receive different types of feedback (Table 2), immediate electronic feedback (Group E) versus self-reflection (Group T).

Table 2. ANCOVA results of the two group on the post-test results

Group	N	Mean	S.D.	F(1,40)	p-value
Group E	22	91.22	5.48	6.08	0.02
Group T	20	74.83	5.91		

*p<.05

The result showed a significant difference after excluding the influence of the pre-test results (F= 6.08, P = 0.02<0.05), which indicate a positive effect on students' learning achievement. The results suggested that student nurses who received immediate electronic feedback could perform clinical performance more accurately at the end of the semester.

5 Conclusion and Future Suggestion

The present study investigated the possibility of employing immediate electronic feedback to improve nursing students' clinical performance. The timely feedback students have received after OCSE has a positive effect on helping student nurses to confirm the correctly performed clinical procedures and to correct wrongly performed ones, and hence the accuracy of clinical performance has been improved in their following practices. However, as a pilot study, the present study was constrained by relatively small sample size and the limited test trails. The future study could include a larger sample size and more test trails. It would be interesting to further explore the improvements student nurses could make through multiple test trails. It would be also interesting to investigate on student nurses' perceptions of receiving electronic feedback versus receiving face-to-face feedback during OSCE.

References

1. Ashby, E. S., Sondgrass, H. S., Rivett, A. D., Russell, T.: Factors shaping e-feedback utilization following electronic Objective Structured Clinical Examinations. Nursing and Health Science 18, 362-369
2. Benner, P. Educating nurses: A call for radical transformation –how far we come? Journal of Nursing Education. 51(4), 183-184, 2012.
3. Giddens, F. J., Brady, P. D. Rescuing nursing education from content saturation: The case for a concept-based curriculum. Journal of Nursing Education 46(2), 65-69 (2007).
4. Gorman, V. Future emergency nursing workforce: What the evidence is telling us. Journal of Emergency Nursing, 45(2), 1-5 (2018).
5. McWilliam, P., Botwinski, C.: Developing a successful nursing Objective Structured Clinical Examination. J Nurs Educ 49 (1), 36–41 (2010).
6. Munro, J. A., Cumming, K., Clelan d, J., Denison, R. A., Currie, R. G.: Paper versus electronic feedback in high stakes assessment.
7. Pantazopoulos, I., Tsoni, A., Kouskoini, E., Papadimitriou, L., Johnson, E., Xanthos, T: Factors influencing nurses' decisions to activate medical emergency teams. Journal of Clinical Nursing 21(18), Issue 18, 2668-2678 (2012).
8. Preheim, G. Armstrong, G, Barton, A. The new fundamentals in nursing: Introducing beginning quality and safety education for nurses' competencies. Journal of Nursing Education. 48(12), 694-697.
9. Schmitz, F. M., Zimmermann, P. G., Gaunt, K.: Electronic rating of Objective Structured Clinical Examinations: Mobil digital forms beat paper and pencil checklist in a comparative study. Springer, Berlin, 501-513 (2011)
10. Segall, N., Doolen, T. L., Porter, J. D.: A usability comparison of PDA-based quizzes and paper-and-pencil quizzes. Computer & Education 45(4), 417-432 (2005).
11. Setyonugroho, W., Jennedy, K. M., Kropmans, T. J. B.: Reliability and validity of OSCE checklists used to assess the communication skills of undergraduate medical students: A systemic review. Patient Education and Counseling, 98, 1482-1491(2015).
12. Snodgrass, J. S., Ashby, E. S., Rivett, A. D.: Implementation of an electronic Objective Structured Clinical Exam for assessing practical skills in pre-professional physiotherapy and occupational therapy programs: examiner and course coordinator perspectives. Australasian Journal of Educational Technology, 30(2), 152-166 (2014).
13. Tanner, C. A. Transforming prelicensure nursing education: Preparing the new nurse to meet emerging health care needs. Nursing Education Perspectives. 31(6), 347-353 (2010).
14. Treadwell, T.: The usability of personal digital assistants (PDAs) for assessment of practical performance. Medical Education, 40 (9), 855-861 (2006).

A Study of 2D multi-person Pose Estimation using Distance Scaling on Images

Boney Labinghisa[1][0000-0001-8241-0377] and Dong Myung Lee[2][0000-0002-1314-8763]

[1,2] Tongmyong University, Nam-gu BUSAN 48520, Republic of Korea
[1,]blabinghisa@yahoo.com, [2]dmlee@tu.ac.kr

Abstract. The innovation of convolutional networks in artificial intelligence led to many studies generally on image processing and object detection. This study aims to make use of human detection in 2D images to correctly estimate human pose. In addition, due recent advances in pose estimation, we aim to apply multi-person human detection with distance scaling in 2D images. The proposed system makes use of multi-person 2D human detection with pose detection and convert the image into a 3D model which estimates the distance between human poses to create a 3D scale. The system is divided into three modules: 2D image multi-person human detection, 2D to 3D pose estimation, and 3D pose distance scaling.

Keywords: 2D Human Pose Estimation, 3D Human Pose Estimation, Deep Learning, Multi-person Detection, Computer Vision, Distance Scaling.

1 Introduction

Image processing led to many studies on object detection from 2D RGB images, video footages and real-time camera streams. Many methods utilized deep convolutional neural networks (CNNs) and have improved in accuracy while also publishing their data and findings in the internet [1-2]. This also led to recent large datasets available publicly for further studies in the field. The 3D representation of a 2D images is a difficult task with many researches focusing on this field [3-5]. A single-person 3D human pose estimation usually crops human figure in images by bounding boxes as predicted by pre-trained human detection model. This cropped part of the image is used as input for a 3D pose estimation module where ground truth is based on the key point such as the pelvis that represents the center joint [6].

The aim of this study is to use a multi-person approach in human detection module before feeding the cropped part of the image into the 3D human pose estimation module and finally make a scaled estimation of distances between the key points from the 3D human pose.

2 Related Studies

In order to understand the advances of human pose estimation, several topics of interest are discussed in the section and how they differ in feature and application. These will serve as the focus of our research.

2.1 OpenPose Model

OpenPose is an open source project developed in Carnegie Mellon University which applied real-time multi-person system to detect human body, hand, facial, and foot key points. The identified a total of 135 key points in an image with a single person [3]. This increases the performance accuracy of human pose estimation with lower variation as compared to earlier human pose detections in a 2D image. Another advantage of this model gave the possibility of real-time estimation of human pose without the need to represent the image in a 3D environment. The model also publicly provides their own dataset for further training and validation.

2.2 Multi-Person 2D Pose Estimation

In comparison to OpenPose, previous multi-person 2D pose estimation system trained using object detection models [7]. This approach was called multi-person pose estimation in the wild, and this top-down approach consists of two stages. The first stage is to predict the location and scale of the bounding boxes likely to contain human figures, while the second stage, estimates the key points of the person contain in the bounding boxes. They used COCO dataset with 17 identified key points, of which 12 were the body joints and 5 were the face landmarks.

2.3 2D to 3D Pose Estimation

One paper used a fusion neural network of two images of 2D human pose by using heatmaps to estimate their key points. The two images are from different angles of camera and the fusion layer is where the heatmaps of the two images are fused together. The system used Recursive Pictorial Structure Model (RPSM) to represent 3D pose from the estimated multi-view of 2D heat maps [8].

3 Multi-person 2D to 3D Pose Estimation and Scaling

3.1 Proposed System Architecture

The overall flow of the proposed system is to use OpenPose Model as the core component in processing 2D images into 3D pose. The system will utilize the key points in OpenPose to acquire a higher accuracy of human pose in 2D images. Then we will use the suggested method in related studies on how to represent 2D pose into 3D pose estimation. After obtaining the 3D estimated pose with key points, the final module is the scaling of distances between key points of multiple persons in the image. The

difference in distance is then used to scale the image. The flow of the proposed system is as seen in Fig. 1.

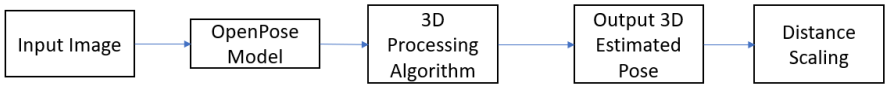


Fig. 1. Basic flow of the proposed system of multi-person 2D pose estimation with distance scaling.

3.2 Distance Scaling Algorithm

The main contribution of the proposed system is the distance scaling applied after the estimated 3D pose from OpenPose Model and 2D to 3D multi-person pose estimation.

To achieve distance scaling, human pose estimation is essential since key points produced from OpenPose can be used to anchor specific parts of the human body. OpenPose first identifies the human figure then applies 17 key points that corresponds to mostly our joints and face. As see in Fig. 2, 2D RGD image is marked with OpenPose key points. From the 17 key points, two key points are used in order to determine the distance from the camera to the test subject, these are the Right Ear and the Torso as seen in the figure.

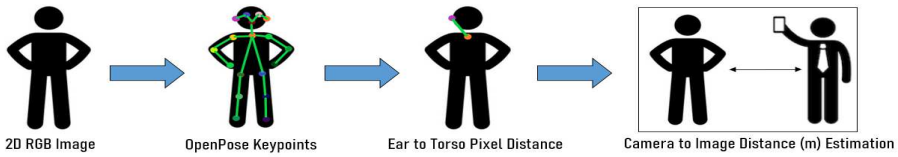


Fig. 2. Application of OpenPose to 2D RGB image and the use of ear-to-torso distance estimation to scale camera to image distance.

A distance scaling is then calculated as seen in equation (1), where d_{kp} is the length in in image scale (pixels) between the key points of the ear and the torso as defined in OpenPose. Whereas, d is the distance in real space (m) of the camera from the person in the 2D image [8]. For a and b, they will be valued at zero as the correlation coefficient between the x and y axis. Euclidean distance is used to compute the image scale d_{kp} as seen in equation (2).

$$d_{kp} = a + \frac{k}{d-b} \; ; \quad \frac{k}{d} \tag{1}$$

$$d_{kp} = \sqrt{(x_{ear} - x_{torso})^2 + (y_{ear} - y_{torso})^2} \tag{2}$$

The reason for using the distance between the ear and the torso for determining the scale is because the face will always appear larger when near the camera and will also appear smaller when far away from the camera. Another reason is the detection of the key points even when the camera is facing towards the face or at the back of the person

[9]. To complete the calculation of the real space distance *d*, the correlation coefficient between pixels and meter *k*, must first be identified. To find *k*, images were gathered with human subjects having fixed distances between the camera. Using the relationship in equation (1) and linear regression, *k* is calculated at 147.56. Based on equations (1) and (2), the real space distance *d* can be easily computed by direct substitution.

4 Experimentation and Analysis

4.1 Results Analysis

The main objective of this paper is to scale the distance in images with multi-person pose estimation. To first test the validity of the scale, three distances were set as the standards as seen in Table 1. As seen in the table the scale made on real space distance is compared with the actual distance from the camera and the human subject.

Table 1. Average differences of the proposed distance scaling compared to the actual distance of the test subject.

Actual distance of test subject	1.0m	1.5m	2.0m	2.5m	3.0m
Average scaled distance	1.11m	1.43m	2.19m	2.23m	3.26m

The results show that there is very little difference between the proposed distance scaling and the actual distance of the test subject. This validates the correlation of distance between ear and torso key points to the real space distance of the test subject. The average error distance of this paper is about 3.87% by root mean square error (RMSE), which is represents very little error that can be caused by the noise on the OpenPose system when determining the key points for the ears and the torso. Even though the results showed a high accuracy, it should also be considered that the test was done on very short distances and might change depending on the linear regression coefficient of *k*, when longer distances are used to estimate the distances.

In Fig. 3, the relation of Euclidean distance and estimated distance of the proposed system can be seen. The estimated distance are represented in blue spheres from the tests with actual distances of test subjects 1.0, 1.5, 2.0, 2.5, and 3.0 meters. While the red line represents the linear regression of Euclidean distance (pixel) and estimated distance (m). As seen on the figure, the greater the Euclidean distance the nearer the test subject is from the camera, while as the scaled distance increases, the Euclidean distance also decreases. And at certain points near the actual distance of 3.0 meters, the Euclidean distance does not decrease further because there can be no negative distance between in the y-axis.

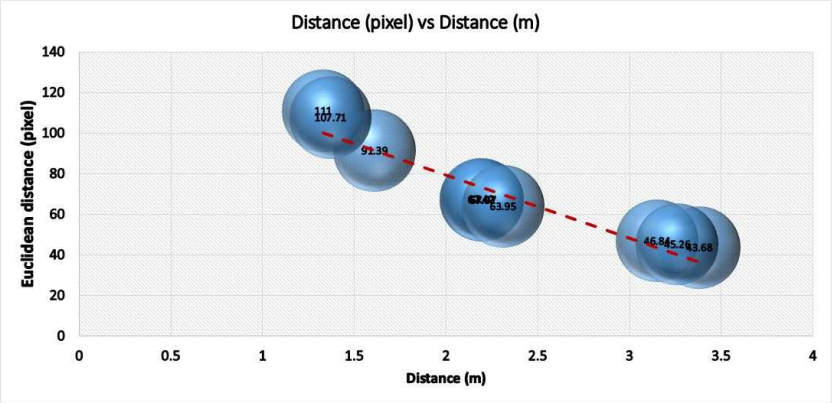


Fig. 3. Euclidean distance (pixel) in the y-axis with scaled distance (m) in the x-axis. The red line represents the linear regression of ear-to-torso pixel distance and the scaled distance (camera to test subject).

4.2 Design Limitations and Considerations

Since the paper is still an ongoing study and experimentation is not yet done, many limitations and considerations are applied. The following are the known limitations and considerations:

- Only applicable in 2D images (as of now, no known real-time pose estimation to 3D using camera).
- Requires knowledge and implementation of human body 3D pose dataset (Human3.6M, MPII, MuCo-3DHP, MuPoTS-3D datasets, in the wild COCO). In this paper, COCO dataset was used in the experiments and 17 key points were used to represent human skeleton.
- Distance scaling is based on the estimation of the Euclidean distance between the key points of the ear and the torso of the human skeleton derived from OpenPose. Knowing the distance between the camera source and the human figure will give the scale of the 2D image.

5 Conclusion

With the development of OpenPose, it made a great contribution is computer vision technology. OpenPose showed a very accurate identification of skeletal key points to represent a human figure and it also correctly identified pose estimation of multi-persons in 2D images. This paper's goal is the utilize the identified key points and transform the 2D image to real space by calculating the distance of the human figure

from the camera giving it a scale in 3D space. The results showed a high performance with only 3.87% error distance.

Despite the high accuracy of the proposed system, more test should be conducted using longer distances and further study should be made on the application of OpenPose and the distance scaling in real-time.

Acknowledgment

This work was supported by the National Research Foundation of Korea (NRF) grant funded by the Korea government (MSIT) (No. 2019R1F1A1062670).

References

1. Kuznetsova A, Rom H, Alldrin N, Uijlings J, Krasin I, Pont-Tuset J, Kamali S, Popov S, Malloci M, Duerig T, and Ferrari V. The open images dataset v4: Unified image classification, object detection, and visual relationship detection at scale. arXiv preprint arXiv:1811.00982 (2018).
2. Han J, Zhang D, Cheng G, Liu N, and Xu D. Advanced deep-learning techniques for salient and category-specific object detection: a survey. IEEE Signal Processing Magazine.10; 35(1): 84-100 (2018).
3. Martinez J, Hossain R, Romero J, and Little JJ. A simple yet effective baseline for 3D human pose estimation. In ICCV, pp. 2640-2649 (2017).
4. Pavlakos G, Zhou X, Derpanis KG, and Daniilidis K. Coarse-to-fine volumetric pre-diction for single-image 3D human pose. In CVPR, pp. 1263-1272 (2017).
5. Rhodin H, Salzmann M, and Fua P. Unsupervised geometry-aware representation for 3D human pose estimation. In ECCV, pp. 750-767 (2018).
6. Sun X, Xiao B, Wei F, Liang S, and Wei Y. Integral human pose regression. In ECCV, pp. 529-545 (2018).
7. Jiang B, Luo R, Mao J, Xiao T, and Jiang Y. Acquisition of localization confidence for accurate object detection. In ECCV, pp. 784-799 (2018).
8. Yang W, Ouyang W, Wang X, Ren J, Li H, and Wang X. 3D human pose estimation in the wild by adversarial learning. In CVPR, pp. 5255-5264 (2018).
9. Pathi SK, Kiselev A, Kristoffersson A, Repsilber D, and Loutfi A. A novel method for estimating distances from a robot to humans using egocentric RGB camera. Sensors. Jan; 19: 3142, pp. 1-13 (2019).
10. Cao Z, Hidalgo G, Simon T, Wei SE, and Sheikh Y. OpenPose: real-time multi-person 2D pose estimation using part affinity fields. arXiv preprint arXiv:1812.08008 (2018).

A Study of Zero-Knowledge Circuit-based for Reduce Transaction Storage Cost

SoonHyeong Jeong[1][0000-0001-8596-0382] and Byeongtae Ahn[2][0000-0003-3431-9493]

[1] Onther Inc., 527, Gangnam-daero, Seocho-gu, Seoul, Republic of Korea
kevi n. j @nt her . i o
[2] Liberal & Arts College, Anyang University, 22, 37-Beongil, Samdeok-Ro, Manan-Gu, Anyang 430-714, South Korea.
ahnbt @ anyang. ac. kr

Abstract. Recently, blockchain technology accumulates and stores all transactions. In addition, existing blockchains cannot participate in the network because they do not become a block sink by using a low-spec server. Due to this problem, as the time passes, the data size of the blockchain network becomes larger and it becomes impossible to participate in the network except for users with abundant resources. Therefore, in this paper, by designing a zero-knowledge circuit that reduces transaction storage space, anyone can participate in the blockchain network. In addition, a blockchain with a zero-knowledge proof-based virtual machine capable of general operation verification can be created to improve transaction processing efficiency.

Keywords: *Zero-Knowledge, validation, transaction, BlockChain, Ethereum*

1. Introduction

The blockchain-based distributed application market is expected to grow from about $ 3.2 billion in 2019 to more than $ 60 billion in 2024. Among them, the market with 'transaction processing' as a profit model is expected to reach 55% of the total. This means that blockchain-based distributed applications are generally provided on the basis of open source, so transaction fees rather than content usage fees are inevitably accepted by users. Therefore, the economic value of the technology to efficiently process transactions is very positive[1]. Blockchain is a decentralized digital ledger that secures the integrity of transaction details and shares the details without the involvement of trusted third parties in the peer-to-peer (P2P) network. A typical example of applying blockchain is cryptocurrency such as Bitcoin and Ethereum. Ethereum introduced the Ethereum virtual machine (hereafter EVM). With EVM, users can program their own way, rather than performing a predefined set of tasks[2]. However, EVM is very inefficient compared to existing virtual machines such as Java Virtual Machine (JVM). And it is difficult to support a complex application environment. In the past decade, numerous blockchain implementations have appeared, but there has been no significant innovation in terms of accumulating and storing transactions. In this case, there is a limit to scalability because all transaction contents must be verified. In addition, since a sepa-

rate verification algorithm is used for each transaction type, a verification burden increases as the size of the transaction increases. Also, due to the nature of the blockchain that stores all transaction data, the data on the blockchain continues to increase over time. When zero-knowledge proof technology is applied to storage of transaction data, data storage space can be saved by compressing the data by pruning actual data and leaving only proof of data. As time goes by, the data of the blockchain will gradually accumulate, and accordingly, the computing resources required to operate the full node are gradually increasing. In the case of Ethereum, it is already difficult for an individual to operate a full node, and in the future, only a large company or large hands that can have sufficient computing resources can operate the full node. These factors will lead to the centralization of the blockchain, and this problem can be solved by reducing the resources required for data storage and verification through a virtual machine with zero knowledge proof technology[3]. In this paper, we designed a zero-knowledge circuit that reduces storage space to improve the transaction efficiency of the blockchain. Section 2 introduces related research and Section 3 introduces domestic and foreign cases. Section 4 proposes a zero-knowledge circuit design, and Section 5 presents a conclusion and future tasks.

2. Related Studies

Blockchain technology can be divided into a simple type of blockchain made of UTXO (Unspent Transaction Output) and a complex type of blockchain that deals with the State Tree. Currently, in the simple form of blockchain, zero-knowledge proof is used at the protocol level only in some transaction processing. However, although some complex forms of blockchain use smart contracts using smart contracts, there are limitations in terms of performance and utilization because they are implemented in the upper layer. Proof size of a single operation created through the proposed SNARKs algorithm is about 1,500 bytes (1.5 kbytes) [4].

* Bullet Proof algorithm -Transaction size of UTXO-based blockchain is measured in (in * 254 * 146 + out * 254 * 33 + 10) bytes, and increases arithmetically according to the number of * in, out used. * It occupies about 45,000 bytes (45kbytes) based on 1 in and 1 out. -Regardless of the type of transaction, the transaction size can be fixed to 1.5 kbytes, and even the simplest transaction standard is more than 70% economical. -The blockchain-based distributed application market is expected to grow from about $ 3.2 billion in 2019 to more than $ 60 billion in 2024 (Blockchain Market Shares, Market Strategies, and Market Forecasts, 2018 to 2024, IBM, 2018). Among them, the market with 'transaction processing' as a profit model is expected to reach 55% of the total.Since these blockchain-based distributed applications are generally provided on the basis of open source, transaction fees rather than content usage fees are inevitably accepted by users[5]. Therefore, the economic value of the technology to efficiently process transactions is very positive. Even if all verification nodes do not participate in block verification, the general operation is verified with the same security strength as all nodes participated and verified using zero-knowledge proof technology, thereby

providing the same effect as storing the entire transaction without saving all transaction data[6].

Currently, as the value of using personal information increases, discussions on how to provide personal information have been actively conducted. Currently, one of the most common methods of providing personal information is a group that uses personal information to obtain personal consent and use personal information. However, the above method has two problems. First, information that is more than the information required by the institution for the use of personal information is exposed. Second, whenever a company requests personal information, there is a problem that a trusted party must provide authentication information for the information to the company[7]. In order to solve the above problems, this paper proposes a privacy-protected personal information management method using zk-SNARK (zero-knowledge Succinct Non-interactive ARgument of Knowledge) technique and blockchain. The privacy-protected personal information management technique can guarantee the privacy while guaranteeing privacy when providing personal information through zk-SNARK. In addition, it is possible to manage personal information data while ensuring the integrity of the data through the blockchain, and sharing personal information can be performed more easily than the existing authentication method.

3. Domestic & International cases

There are several companies with blockchain virtual machine technology. The most representative virtual machine is EVM, Ethereum's virtual machine. EVM is the first blockchain virtual machine and based on EVM, Ethereum has grown into a basic platform for smart contracts, tokens, and decentralized applications (Dapps). And many blockchain projects are using Ethereum's EVM when creating the mainnet. Currently, Ethereum is planning to upgrade to Ethereum 2.0, and when Ethereum 2.0 is introduced, the current virtual machine EVM will be converted to eWASM[8].

EOS-VM is a virtual machine created by EOS.IO and is not limited to the blockchain industry, and is expected to be used in traditional software development fields such as game engines, databases, and web frameworks. EOS-VM is a virtual machine dedicated to the blockchain system, and it can be expected to save development resources (CPU), improve blockchain scalability, and improve development efficiency compared to the first blockchain virtual machine, EVM.Tron's virtual machine TVM is developed based on Ethereum's EVM and is characterized by being compatible with Ethereum[9]. By designing a unique virtual memory mechanism, the amount of memory actually used can be greatly reduced, and the operation cost of a decentralized application can be greatly reduced by providing developers with almost unlimited memory capacity. And you can save resources by optimizing the compiler.Table 1 shows domestic and International cases as a table.

Table 1. Domestic & International cases

Coin name	Consensus method	Characteristic	Market cap	Remark
Ethereum	EVM	Turing completeness as the first blockchain virtual machine	$ 13 billion	Focusing on decentralization and security
EOS	EOS-VM	Consensus algorithm similar to indirect democracy	$ 2.4 billion	Value for scalability
Tron	TVM	EVM-enhanced virtual machine featuring Ethereum compatibility	$ 1.8 billion	

Currently, the domestic blockchain technology is mainly biased to the underlying technologies related to the main net, such as distributed ledgers and consensus algorithms. Due to the nature of the domestic technology ecosystem, the area that can lead in the global market is the distributed application area rather than the mainnet area. And there is currently no zero-knowledge proof-based virtual machine that can efficiently verify complex operations required for distributed applications, not just bookkeeping.Therefore, by designing a system for improving the amount of code verification based on zero-knowledge proof applicable to various distributed applications and smart contract execution environments, it will become a distributed application-based technology with great growth potential in the future[10].

4. Design of Zero-knowledge circuit

Due to the nature of the blockchain that stores all transaction data, the data on the blockchain continues to increase over time. When zero-knowledge proof technology is applied to storage of transaction data, data storage space can be saved by compressing the data by pruning actual data and leaving only proof of data. As time goes by, the data of the blockchain will gradually accumulate, and accordingly, the computing resources required to operate the full node will gradually increase.

In the case of Ethereum, it is already difficult for an individual to operate a full node, and in the future, it is expected that only large companies or large hands that can have sufficient computing resources can operate the full node. This will cause the centralization of the blockchain, and this problem can be solved by reducing the resources required for data storage and verification through a virtual machine with zero knowledge authentication technology.

Figure 1 is designed to apply the zero-knowledge proof algorithm to the virtual machine. It shows the flow of the operation method of the Ethereum virtual machine for applying zero knowledge proof technology. Since Solidity, the smart contract language of Ethereum, is a language created for human understanding, it needs to be changed to a machine language understandable by a virtual machine in order to operate in a virtual machine. Code written in Solidity is converted to Ethereum bytecode by the compiler. This bytecode is executed by EVM, Ethereum's virtual machine. When a specific

bytecode is executed, all nodes in the Ethereum network execute the same bytecode respectively to verify the transaction. At this time, if zero-knowledge proof technology that can perform general operation verification is applied to the virtual machine, even if the virtual machine does not execute the transaction, it is possible to know whether the corresponding transaction is the correct transaction by performing verification on the zero-knowledge evidence.

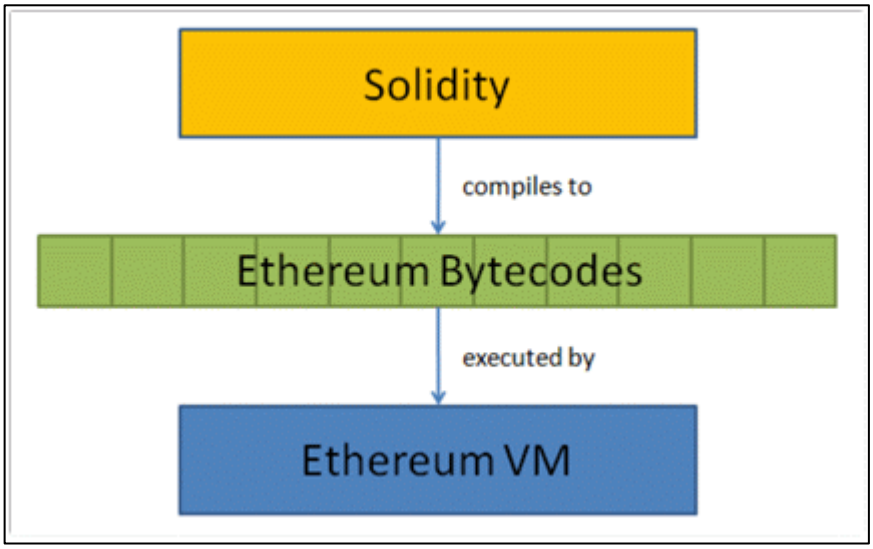


Figure 1. execution process of Ethereum virtual machine

In order to modify the virtual machine, it is necessary to understand the structure. Therefore, Figure 2 shows the architecture and execution flow diagram of the Ethereum virtual machine. Once you understand how the virtual machine is running, you need to figure out what parts of the virtual machine need to be modified to apply zero-knowledge techniques.

6

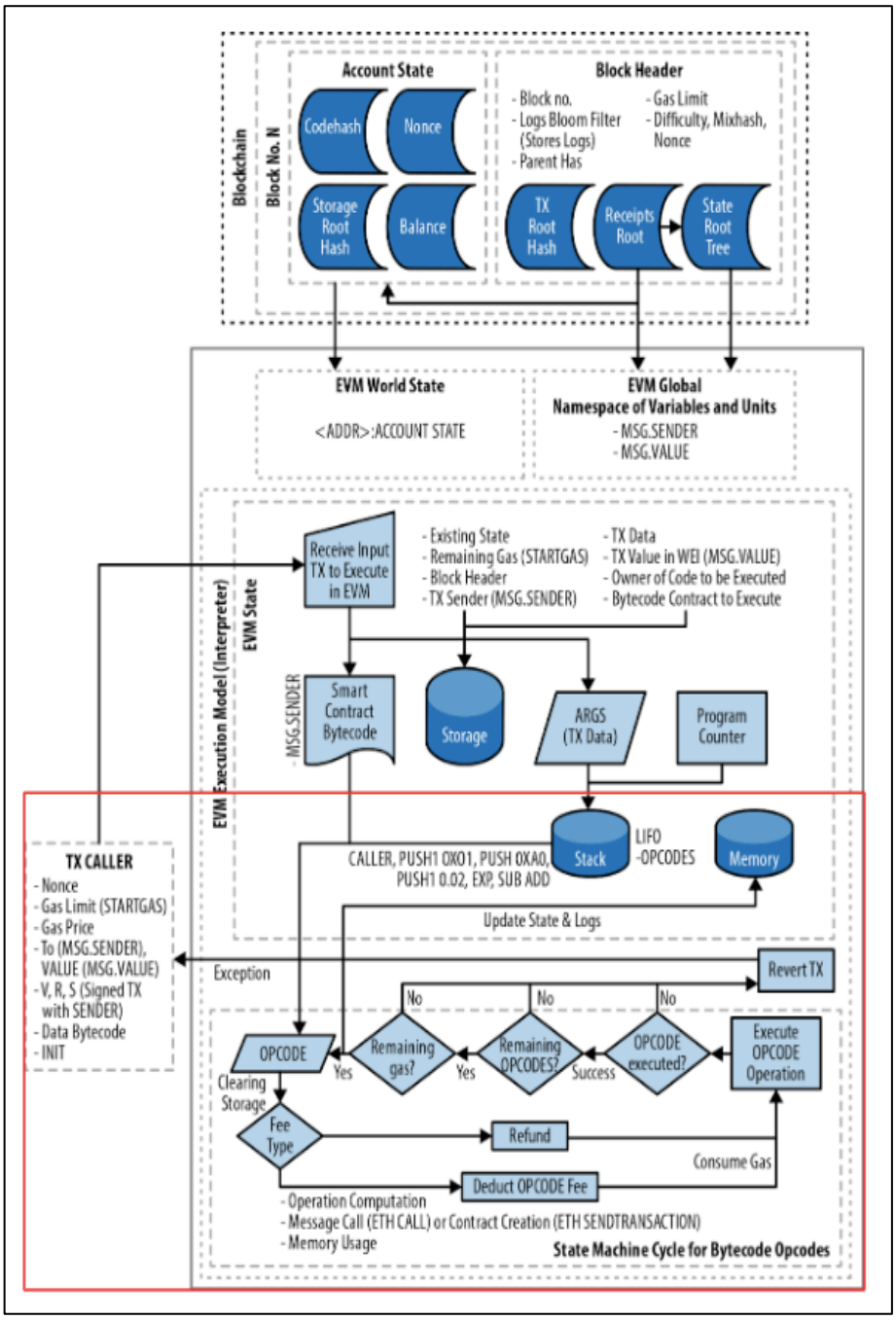


Figure 2. Architecture & Excution Flow Chart of Ethereum Virtual Machine

In order to execute the transaction, it needs to be changed to Ethereum bytecode as mentioned. These bytecodes are decomposed into what are called opcodes, stacked on the stack and executed one by one. You must subtract the gas cost for running the virtual machine before the opcode runs. The opcodes are now executed if the gas cost is not insufficient. Figure 3 shows the parts that need to be changed in the virtual machine when the opcodes are executed.

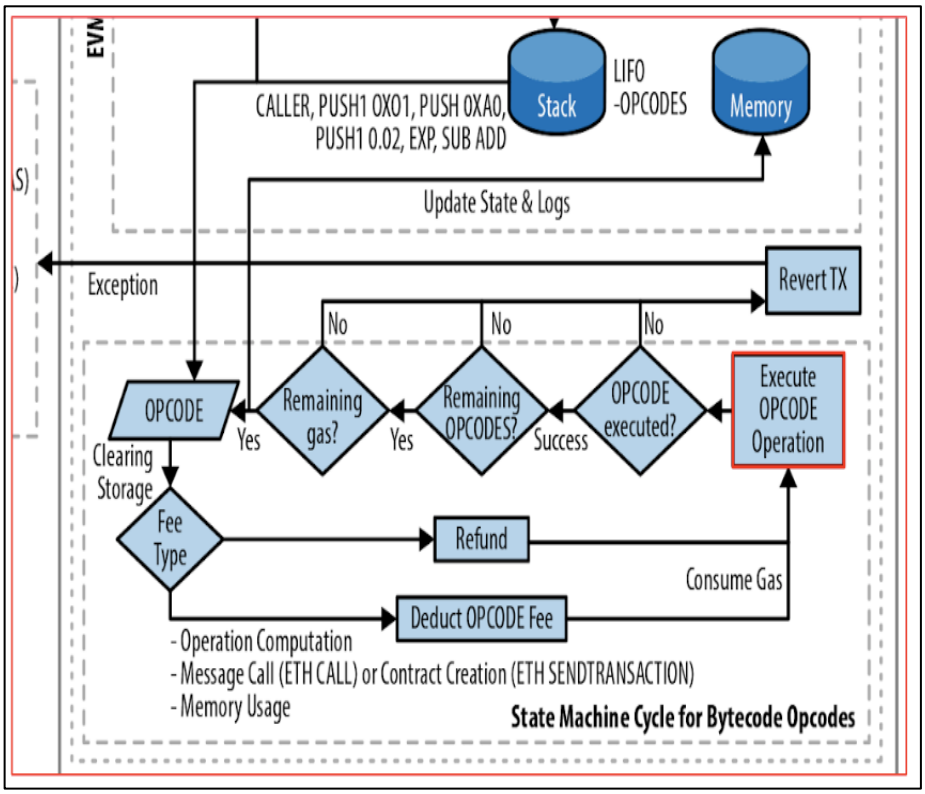


Figure 3. Change Part in Virtual Machine

In the figure above, the part marked with a red box is the part to which zero-knowledge proof technology should be applied, and the part to create a universal circuit that can execute the opcode. Figure 4 shows the change in data stored after the zero-knowledge proof technology is applied. After the zero-knowledge proof technology is applied to the part that performs the opcode, TX data among the data stored in the existing storage is replaced with the proof of the zero-knowledge proof. And we will create and test a virtual machine with zero knowledge proof.

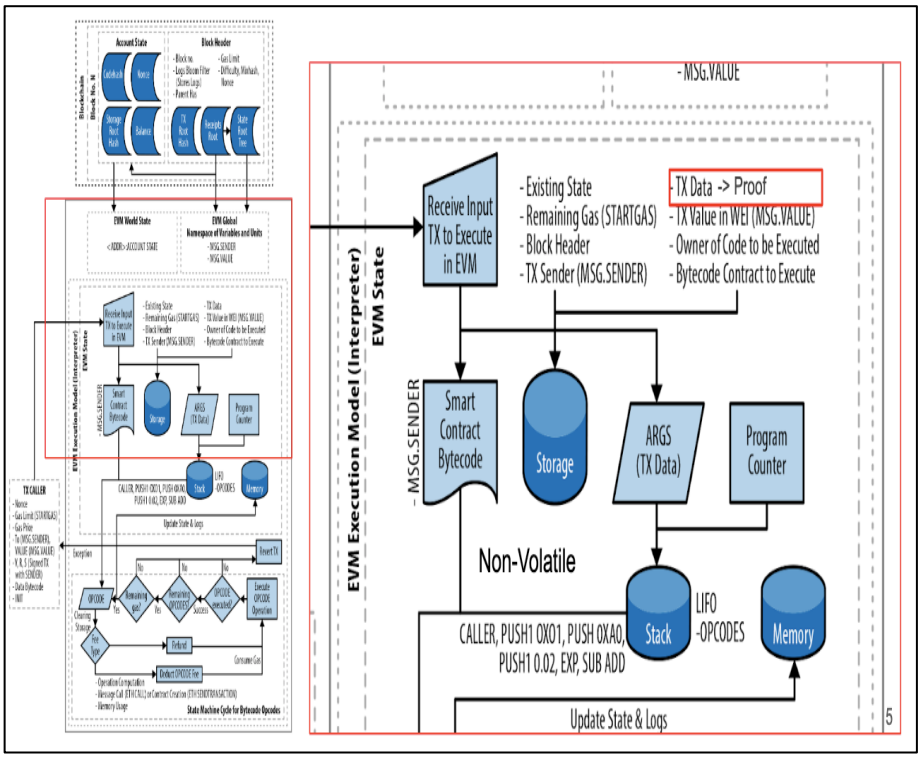


Figure 4. Data Changes after zero knowledge proof technology

5. Conclusion and future works

In this paper, we designed a zero-knowledge circuit to reduce transaction storage space. It was designed by applying the existing verification amount optimization to the virtual machine. This study is a practical example of applying the zero-knowledge proof algorithm capable of semi-operational verification, and can develop two different blockchains in the future. In addition, Crypto Currency implementation with zero-knowledge proof algorithm for general operation verification can be developed. It can also create new blockchain business opportunities, such as platform services, where DApps linked to public blockchains can be integrated with each other. In the future, we will develop our own platform through practical implementation based on this design.

References

1. Sikorski JJ, Haughton J and Kraft M(2017), "Blockchain technology in the chemical industry: Machine-to-machine electricity market", Applied Energy, 195, 234-246
2. X. Xu, I. Weber, M. Staples, L. Zhu, J. Bosch, L. Bass, C. Pautasso, P. Rimba(2017), "A taxonomy of blockchain-based systems for architecture design", IEEE International Conference on Software Architecture (ICSA), 243-252.
3. Zibin Zheng, Shaoan Xie, Hong-Ning Dai and Weili Chen(2020), "An overview on smart contracts: Challenges, advances and platforms", Future Generation Computer Systems, Vol. 105, 475-491.
4. Ahmet Bugday, Adnan Ozsoy, Serdar Murat Öztaner and Hayri Sever(2019), "Creating consensus group using online learning based reputation in blockchain networks", Pervasive and Mobile Computing, Vol. 59, 111-125
5. Matevž Pustišek and Andrej Kos(2018), "Approaches to Front-End IoT Application Development for the Ethereum Blockchain", Procedia Computer Science, Vol. 129, 410-419.
6. Eli Ben-Sasson, Alessandro Chiesa, Eran Tromer and Madars Virza. “Succinct Non-Interactive Zero Knowledgefor a von Neumann Architecture”, https://eprint.iacr.org/2013/879.pdf
7. [BCGT13a] Eli Ben-Sasson, Alessandro Chiesa, Daniel Genkin, and Eran Tromer. Fast reductions from RAMs to delegatablesuccinct constraint satisfaction problems. In Proceedings of the 4th Innovations in Theoretical Computer ScienceConference, ITCS ’13, pages 401–414, 2013.
8. [Val08] P. Valiant. “Incrementally verifiable computation or proof of knowledge imply time/space efficiency”, In: Theory of Cryptography. Ed. by Ran Canetti. Berlin, Heidelberg: Springer, 2008, pages 1–18. isbn: 978-3-540-78524-8
9. [Sasson14] E. B-Sasson, A. Chiesa, E. Tromer 그리고 M. Virza. “Scalable zero knowledge via cycles of elliptic curves (extended version)”. In: Advances in Cryptology - CRYPTO 2014. Vol. 8617.
10. [Bowe19] S. Bowe, J. Grigg, D. Hopwood, “Halo: Recursive Proof Composition without a trusted setup”, https://eprint.iacr.org/2019/1021.pdf

A Study on the Factors of Performance Improvement in Manufacturing SMEs on Overseas Market

Seung Gyun Yoo

Department of Global Economics and Commerce, Dongguk University-Gyeongju, 123, Dongdae-ro, Gyeongju-si, Gyeongsangbuk-do, 38066, Republic of Korea

bluetrade@dongguk.ac.kr

Abstract. This study analyzed the effect of marketing strategy on the performance of SMEs in overseas market by setting up domestic and overseas resources, marketing mix as parameters. The result shows that the domestic environment of SMEs has a significant negative effect on product strategy and a positive effect on distribution strategy. Utilization of information has a significant positive effect on the marketing mix strategy. Managerial competence has a significant positive effect on price strategy and promotion strategy. The marketing mix strategy has a significant positive effect on non-financial and financial performance. The following is how to use the marketing mix strategy to improve the performance of SMEs. First, information and resources that SMEs do not have must utilize external networks. Second, management needs to build a team that can share trust and vision. The results of this study are also meaningful in that they can re-examine the internal competency and marketing capabilities of SMEs, and provide the basis for restructuring the capabilities to improve the performance of the company. In the present situation where economic growth is stagnating, only job creation through activation of manufacturing companies can vitalize the national economy. However, this study was conducted for SMEs and has limited results. Nevertheless, I hope that it can be a foundation for the revitalization of the entire manufacturing company.

Keyword : SMEs, Marketing, Manufacturing, Overseas, Performance

1. Introduction

The growth of SMEs plays an important role in economic growth. In particular, the size and sales of manufacturing companies among small and medium-sized enterprises are significant because they are linked to employment in domestic and foreign markets.

Recently, various studies related to manufacturing SMEs have been conducted, and are also used in marketing strategies to improve the performance of companies. There are also strategies that utilize fashion such as Hallyu.

SMEs mostly focus on the domestic market. However, the overseas market can increase the size of a company with an important role to expand performance. In addition, SMEs can quickly respond to market changes through an efficient network of internal and external resources. In this regard, it is necessary to confirm which factors play an important role for manufacturing SMEs to improve their

performance in overseas markets.

This study explored the success factors of a company's overseas market in terms of marketing. To this end, empirical research was conducted on the company's domestic marketing activities, managers' capabilities, information utilization capabilities, and performance. Through research, it was suggested through various strategic measures to improve the performance of SMEs.

2. Literature Reviews

Connant, Mokwa, and Varadarajan (1990) also analyzed the relationship between the strategy type and marketing competency suggested in Miles and Snow (1978), and found that companies pursuing an enterprising strategy were more than companies pursuing other types of strategies. It turns out that he has a lot of marketing capabilities.

Park(2017) also if the domestic market is large enough, the company will pursue an expansion strategy in the domestic market until the cost of expanding additional business in the domestic market is higher than entering the overseas market. Barney(1991) presented that from a resource-based point of view, resources are regarded as an important influencing factor affecting strategy and performance as a source of corporate competitive advantage.

Kim(2005), Lee(2012), Kim(2010) and Kim(2008) said the Korea Wave is a popular trend of Korean culture and a term that means the Korean phenomenon spread throughout the global region. It was initially used as a simple approach to content disseminated in China, but gradually expanded its domain and meaning, becoming the term for Korean pop culture that spread beyond East Asia and Central Asia to Europe. Hoeffler & Keller(2018) presented that the fourth industrial revolution, which is spreading recently, focuses more on software than hardware, and there are so many phenomena that we call the age of content. In particular, companies competing in the global market are engaged in marketing activities with a variety of local content. And Lee(2019) have studied the improve performance.

These studies can be interpreted as a favorable shift in interest in SMEs products, and the higher the interest in SMEs, the better quality of SMEs products, and attitudes toward SMEs products. Taking advantage of trend who are leading the recent wave, Korea Wave marketing is a good example of this, not only of businesses but also of the government's policy approach. Tourism marketing at the filming site of the drama, traditional marketing focusing on food and culture, and marketing of Korea Wave products, where celebrities are concentrated, are good approaches for companies to actively utilize.

3. Model and Result

3.1 Model

The model of this study is as follows.

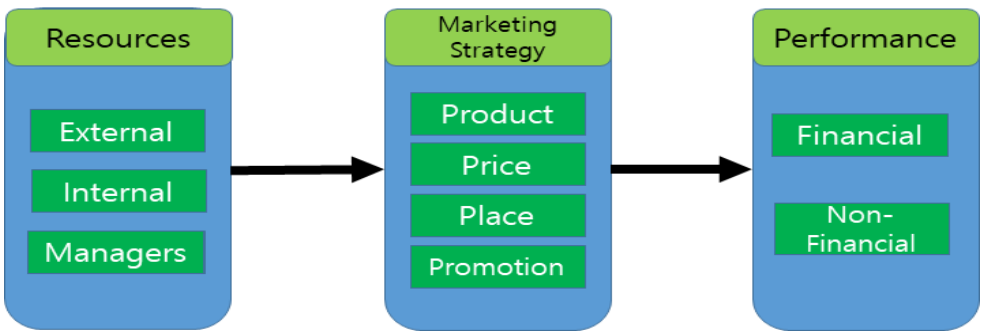


Fig.1. Reserch model

3.2 3.2 Results

Demographic characteristic of respondents by survey is as follows.

Table 1. Demographic Characteristic of Respondents

Classification	Detail	Response(No.)	Persentage(%)
Size	under 50 Labors	142	33.5
	51-100	112	26.4
	101-200	80	18.9
	201-300	62	14.6
	Over 301	28	6.6
Build	under 10 Years	75	17.7
	10-20	78	18.4
	over 20	271	63.9
Marketing Manager	None	0	0.0
	under 5%	193	45.5
	under 20%	100	23.6
	under 50%	72	17.0
	over 51%	59	13.9
Total		424	100.0

Criteria for goodness of fit is as follows.

Table 2. Criteria for Goodness of Fit

x2/DF	GFI	AGFI	NFI	CFI	PGFI	RMR	RMSEA
≤3.000	≥0.900	≥0.800	≥0.900	≥0.900	≥0.600	≤0.080	≤0.080

The correlation by factors is as follows.

Table 3. Correlation by Factors

Classify	A	B	C	D	E	F	G	H	I

External(A)	1								
Internal(B)	.444***	1							
Managers(C)	.593***	.569***	1						
Product(D)	.460***	.458***	.513***	1					
Price(E)	.390***	.514***	.581***	.486***	1				
Place(F)	.414***	.452***	.623***	.439***	.671***	1			
Promotion(G)	.182***	.290***	.302***	.458***	.362***	.297***	1		
Financil(H)	.287***	.297***	.322***	.435***	.287***	.377***	.261***	1	
Non-Financial(I)	.381***	.311***	.403***	.364***	.375***	.379***	.265***	.362***	1

* p<.05, ** p<.01, *** p<.001

The route analysis results are as follows.

Table 4. Route Analysis Results

Path			COE	E.M	T-V	P-V	Results
External	→	Marketing Strategy	0.107	0.143	0.072	1.979*	Accept
External	→	Financial	0.252	0.258	0.064	4.050***	Accept
External	→	Non Financial	0.169	0.163	0.052	3.103**	Accept
Internal	→	Marketing Strategy	0.094	0.074	0.034	2.192*	Accept
Internal	→	Financial	0.163	0.225	0.075	3.012**	Accept
Internal	→	Non Financial	0.139	0.113	0.036	3.183**	Accept
Managers	→	Marketing Strategy	0.127	0.126	0.053	2.388*	Accept
Managers	→	Financial	0.120	0.098	0.036	2.706**	Accept
Managers	→	Non Financial	0.155	0.148	0.060	2.492*	Accept
Marketing Strategy	→	Financial	0.378	0.518	0.073	7.138***	Accept
Marketing Strategy	→	Non Financial	0.320	0.437	0.073	6.020***	Accept

* p<.05, ** p<.01, *** p<.001

4. Conclusion

The various problems that have arisen in global, coupled with the changing perception of the social environment, make us re-recognize that the global market moves like one. This shows that the reaction of companies to the market and consumer perception of companies are moving together. In addition, SMEs' marketing in overseas is based on their initial character, and consumers in global are also helping companies improve their performance by responding to the trend and marketing.

In the external environment, the factor that significantly influences the marketing mix strategy of small and medium-sized

manufacturers is the domestic environment, which shows that the overseas competitiveness of SMEs is relatively inadequate. Moreover, in the domestic environment, product strategy was found to have a negative effect. This shows that the marketing activities focused on products by SMEs are reducing the performance.

Given the size and sales of manufacturing companies, improving the distribution structure of overseas markets was not able to obtain meaningful results through empirical analysis of this study. It is necessary to focus more on the domestic market, which shows that improving relations with trading partners may be more efficient than product development or low-priced policies.

In addition, it was found empirically that management had an influence on product price and promotion activities. This shows that responding to the price of own products and raising awareness of customers and customers has a more significant effect on performance.

Marketing mixes have a significant impact on financial performance. This means marketing activities are needed to improve corporate performance. It is not known what factors, such as characteristics of domestic and overseas markets and timing of transactions, will lead to more significant results, but this study has identified the need for marketing activities.

In particular, factors such as sales, operating profit ratio, and value-added ratio, which are financial performance variables, are variables that can evaluate the performance of the marketing mix, which is one of the major contributors to this study.

Marketing mixes have a significant impact on non-financial performance. This means marketing activities are needed to improve corporate performance. In particular, factors such as non-financial performance variables such as customer satisfaction, customer retention rate, and employee ability are variables that can evaluate the performance of the marketing mix, which is one of the main contributors to this study.

Measures to enhance business performance through research results are as follows. First, small and medium-sized manufacturing companies should establish and utilize an enterprise external network for information and resources that are not in the enterprise. If the network is formed, it has the advantage of sharing resources needed for marketing activities and also having the scale to create new opportunities. It is also easy to supply and demand funds, such as capital liquidity and the use of guarantee insurance. Second, we need to build a marketing team. This reduces internal conflicts and transaction costs, and can lead to sharing of corporate vision and realization of objectives.

It is also meaningful that the results of this study can be re-examined the internal capabilities and marketing capabilities of small and medium-sized manufacturing companies, and provide the basis for restructuring to enhance the performance of companies.

The limitations of this study are as follows. First, since the data were collected as a measurement method for Internet questionnaires, the survey results may depend on the attitude of the survey respondents. Second, because the positions of the survey respondents vary, it is difficult to accurately reflect the circumstances of the entity.

References

1. Allen, N. J. and J. P. Meyer, "The Measurement and Antecedents of Affective, Continuance and Normative Commitment to the Organization", Journal of Occupational Psychology, Vol.63(1990).
2. Andolsek, D. M. and Stebe, J., "Multinational Perspectives on Work Values and Commitment," International Journal of Cross Cultural Management, Vol. 4(2)(2004).
3. Barczak G. New product strategy, structure, process, and performance in the telecommunications industry. J Prod Innov Manage, 12(3), pp.224–34(1995).

4. Catherine. K., "Organizational Citizenship and Withdrawal Behaviors in USA and India:Does Commitment Make a Difference?", International Journal of Gross Cultural Management, Vol.3(1)(2003).
5. Chen, Chung-Jen, "The Effects of Knowledge Attribute, Alliance Characteristics, and Absorptive Capacity on Knowledge Transper Performance," R&D Management, Vol.34(3)(2004).
6. Davies H, Walters P. Emergent patterns of strategy, environment and performance in a transition economy. Strateg Manage J, pp.25(4):347− 64(2004).
7. Eun Young Kim, "A SWOT Analysis of Korean Wave and Marketing Strategy of Korean Movie in Asian Markets", The Journal of Korean Movies, 27(0), 89-110(2005).
8. Goeldner, C. R. and J. R .B. Ritchie, "Tourism: Principles, Practices, Philosophies (9th edition)", John Wiley & Sons, New Jersey(2003).
9. Han E. K., "A Study on the Indexation of Corporate Social Responsibility", Korean Journal of Boarding & Telecommunication Studies, 17(3), 274-303(2003).
10. Hong S. T, Kang D. K, Ohya Yoshihiko., "The Effects of Korean Wave on Evaluation and Purchase Intention for Products Made in Korea", The Journal of Marketing Managment, 12(1), 71-90(2007).
11. Hyung-Shik Jung, "The Effects of Consumer's Perception of Korean Wave(Hallyu) on Korean Product Purchase and Country Image in Chinese Market", Research of Consumer, 17(3), 79-101(2006).
12. Kim, Dong Kyun. & Koh, I. K., "A Study on the Influence of Brand Image and Product Attributes upon Purchase Intention: Focusing on the Analysis of Full-Sized Sedan", Korean Enterprise Management Studies, 17(4), 215-231(2010).
13. Kim, M. J., "The Influence of Korean Image Media on the Korean Wave Visitors", Master Thesis, Sejong University(2010).
14. Kim, M. J., "The Influence of Korean Image Media on the Korean Wave Visitors", Master Thesis, Sejong University(2004).
15. Kim, S. S &, Kim, M. J., "The Effects of Preference for Korean Cuisine Culture on Image of Korea and Intention to Visit Korea as a Potential Tourism Destination as Perceived by Residents of Asian Countries Affected by Hallyu", The Journal of Foodservice Managment Society of Korea, 11(3), 117- 145(2008).
16. Lee, B. S., "A Study on the Impact of Country Image on Selection of Goods and the Moderating Effect of Brand Personality", Korea Trade Review, 44(2), 73-90(2019).
17. Lee, J. K., Oh, S. K. & Xu, D. P., "A Study on the Influence of Korean Wave and CSR on Company Image and Intentions to buy; Focused on the Chinese University Student", The Journal of Korea Research Society for Customs, 13(4), 323-340(2012).
18. Leonidas Leonidou C., Constantine S. Katsikeas and Saeed Samiee, “Marketing Strategy Determinants of Export Performance : A Meta-Analysis”, 「Journal of Business Research」, Vol.55, pp.51-67(2003)
19. Menon, S., & Khan, B., "Corporate sponsorship of philanthropic activities: When do they impact perception of sponsor brand?", Journal of Consumer Psychology, 13(3), 316-327(2003).
20. Porter, M. E. & Kramer, M. R., “Strategy and Society: The Link between Competitive Advantage and Corporate Social Responsibility,” Harvard Business Review, 84(12), 56-68(2006).
21. Pothukuchi,. F., Damanpour, J. Choi, C. C. Chen and S. H. Park, "National and Organizational Culture Differences and International Joint Venture Performance," Journal of International Business Studies, Vol. 33(2)(2002.)
22. Suh, C. H. & Yang, J. Y., "The Effects of Korean Wave Contents on Country Image and Behavioral Intentions",

Korea Management Association, 25(4)(2012).

23. Yoon, Kak & Suh, S. H., "The Influence of Corporate Advertising and Social Responsibility Activities on Corporate Image and Brand Attitudes", Communication Research, 61, 47-72(2003).

The Design and Implementation of Blockchain-based Supply Chain System with Traceability

Chien-Ying Chen[1], Yu-Wei Chan[2], Chih-Hung Chang[2], Tsan-Ching Kang[2], Chun-Hong Huang[3]*, Yin-Te Tsai[4]

[1] Computer Science and Information Engineering, Providence University
[2] College of Computing and Informatics, Providence University, Taiwan
[3]Lunghwa University of Science and Technology
[4]Department of Computer Science and Communication Engineering, Providence University
ryomil2@gmail.com, ywchan@gm.pu.edu.tw, ch.chang@gm.pu.edu.tw, ch.huang@mail.lhu.edu.tw, tckang@pu.edu.tw, yttsai@pu.edu.tw

Abstract. In the past decade, consumers have paid more and more attention to the source and safety of commodities and food, and the traceability of the supply chain has become very important. Most of the current supply chain adopts centralized management. If multiple parties are involved (production, processing, transportation, etc.), the existing centralized management method requires a lot of manpower costs to help achieve the supply due to the isolation of data and lack of trust. In recent years, blockchain technology has been widely used due to its inherent characteristics, such as distributed ledger and cryptography, to overcome the trust problem among entities which means that entities do not rely on the central organization (e.g. central authority) or a third-party intermediary unit (e.g. Intermediary). Thus, the blockchain technology is applied to many applications and industries, especially the logistics industry. Currently, the logistics company needs to help buyers and sellers to advance and import goods by using the phone to negotiate for the money. This process is very inefficient. In this work, we propose a multi-party supply chain system with traceability based on the blockchain technologies.

Keywords: Supply Chain, Blockchain, Smart Contract, Traceability.

1 Introduction

In recent years, the demand for traceability from raw materials to retailers to consumer goods has become higher and higher [1]. The global supply chain is large and complex. If the need for traceability needs to be met, in the traditional supply chain, supplementary companies should cooperate. To trace the source of the product accompanies the transformation process [2]. However, each participant in the traditional supply chain system is isolated and requires a unified format, and the trust mechanism is very difficult [3]. We use the blockchain technology to implement a supply chain system for

achieving warehousing and traceability across multiple parties. The characteristics of blockchain are: Decentralization, immutability and standardization [4], which can solve the shortcomings of traditional blockchain.

In this work, we propose the smart contract mechanism, which create goods with the smart contract to record information such as characteristics and sources [5], and then transfer, exchange the information through the smart contract for recording the composition of goods. Thus, our system can not only trace the origin of the processed ingredients and raw materials, but also manage the purchase and sales of each role. In addition, we solve the problem of the traditional supply chain system, which is that the traditional supply chain system is not able to trace the source, and make food inspection very easy. You can directly view the ingredients and sources of products through this system [6].

2 Literature Review

2.1 Traditional supply chain

The traditional supply chain is to centrally store data, such as resource supply, producers, and customers. These data are isolated or independent from each other, so as to realize each small segment of monitoring. In the context of supply chain management, traceability has promoted good development. The traceability refers to the process from raw materials to products and customers [7]. The traditional supply chain achieves short-distance traceability. If you need to retrieve the source of the upper layer, you must provide it through the supplier and a trusted third party. Although traditional supply chain systems can completely manage every commodity, traceability is very limited.

2.2 The blockchain technology

Recently, the blockchain technology provides a decentralized database to improve the shortcomings of centralized management of traditional supply chains. The blockchain technology has a privacy protection framework that can provide data security, and an encrypted private key is deployed to ensure data privacy.

In addition, the blockchain technology eliminates third parties and intermediaries. Relatively, it has a consensus mechanism to ensure that only effective transactions are executed. And all actions are visible to all participants in the chain, so it can prevent malicious attacks.

2.3 Smart contract

In order to apply warehousing, trading and other functions in the blockchain, the relevant parties must reach the consensus established in the smart contract, and the transaction will not be conducted until both parties agree. Originally proposed by Nick Szabo in 1994 [8], smart contracts execute set scripts without third-party intervention to complete functions such as transactions and ownership transfer.

2.4 The distributed file system

In order to access files in the blockchain, we use a distributed file system (InterPlanetary File System, IPFS) as a file access platform. The IPFS is a distributed file system [9], which is allocated for each access file a unique address, and you can set this file to be public, private, or to a specific person.

3 System Design

In this work, we propose the architecture of the blockchain-based supply chain system, which is shown in Fig. 1. The architecture includes five roles: Producer, processing plant, logistics, retailer, and customer. In this work, the design of smart contract is used to implement the blockchain-based traceability supply chain system.

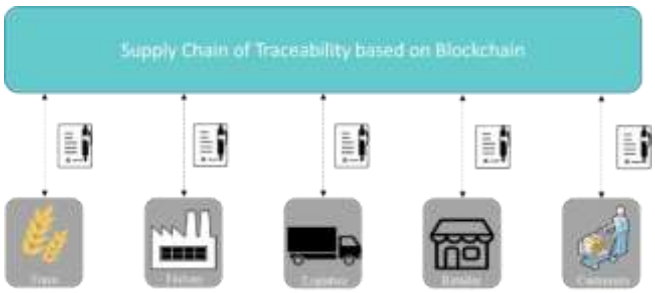


Figure 1 The five roles designed in the proposed system.

In the following, we will present the functions of the five roles.

(1) **Producers**: The functions are designed for producers are build, forsale, accept, send and other functions. Producers use build to build the goods and record the characteristics of the goods (origin, variety, price, etc.) when producing the goods. After confirming that you can sell, use forsale to find buyers. When a buyer places an order, the producer needs the accept function to accept the transaction, and the producer decides whether logistics is needed to help deliver the goods. If you do not need logistics assistance, you can directly use send to send the goods to the buyer; when you need logistics assistance, you need to find through the ship.

(2) **Processing plant**: The processing plant needs buy, received, process, forsale, accept, ship, send and other functions. The processing plant must use buy function to make a transaction with the seller when it sees that there is good goods. After the seller accepts the transaction, whether the seller delivers the goods in person or through logistics, the goods need to be received to accept the goods when they arrive at the processing plant. When the processing plant needs to process the goods, it needs to process the goods a and b to process the goods c. After finishing the processing, you need to use forsale to find buyers. When a buyer places an order, the processing plant needs the accept function to accept the transaction, and the processing plant decides whether it needs logistics to help deliver the goods. The processing plant decides whether logistics is needed to assist in the delivery

of the goods. If you do not need logistics assistance, you can directly use send to send the goods to the buyer; when you need logistics assistance, you need to find through the ship.

(3) Logistics: The logistics needs delivery, ship, accept, received, arrived and other functions. Logistics must use deliver to establish routes and record remaining space and other information. Logistics can use the ship to find the seller who needs logistics assistance, or use accept to accept the logistics assistance that the seller has proposed. After confirming the need for logistics assistance, receiving goods from the seller requires that the received representative has received and started shipping. When shipping to the buyer, the goods need to be delivered to the buyer through arrived.

(4) Retailers: Retailers need buy, received, forsale, soldout and other functions. The retailer must use buy to make a transaction with the seller when it sees that there is good goods. After the seller accepts the transaction, whether the seller delivers the goods in person or through logistics, the goods need to be received to accept the goods when they arrive at the retailer. Retailers need to use forsale to start selling goods. When the goods are sold, a sellout is required to record that the goods have been sold.

(5) Customer: Customers need getstate to check the source of the goods.

4 Application Scenarios

In order to realize the traceability of the supply chain and the system of process automation, Figure 2 uses smart contracts to record the characteristics and actions of all raw materials and goods to achieve a clear source and destination for all goods. Each role is communicated through a smart contract. The production and transaction process in the system:

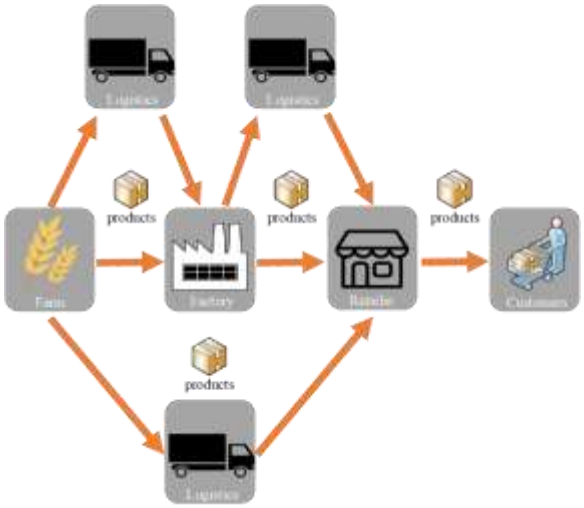


Figure 2 An example of supply chain system based on blockchain.

Figure 3 shows the process of the producer. After the producer produces the goods, he can find the buyer. When the buyer places the order, the seller can deliver it to the buyer by himself or through logistics.

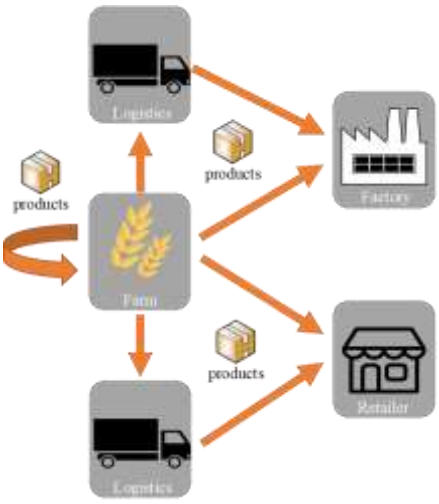


Figure 3 The process of the producer.

Figure 4 shows the process of the processing plant. After receiving the goods from other roles, the processing plant processes it internally, processes the goods A and B into the goods C, and then looks for the buyer. When the buyer places the order, the seller can do it by himself or through Logistics delivery to buyers.

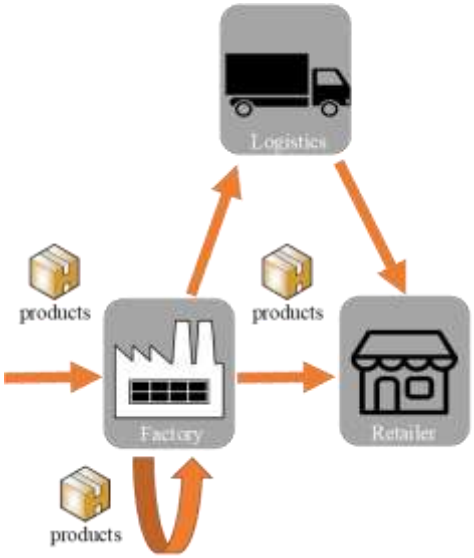


Figure 4 The process of the processing plant.

Figure 5 shows the process of the retailer. After receiving the goods from other roles, the retailer puts the goods on the shelf and sells them. Customers who bought the goods can query the traceability of the goods through this system.

Figure 5 The process of the retailer.

5 System Implementation

Figure 7 shows that how the producers produce rice, how to go through the logistics and retail process, and finally reach the customer and trace the products. After the farmers produce rice, they build the goods through the system (as shown in Figure 6 above), and they start looking for buyers of rice for sale.

Figure 6 The process of producing goods.

Figure 7 shows that how the retailer needs to buy rice, finds the required rice through the system, and presses to buy.

After the retailer places the order, the producer will receive the transaction and choose whether to accept the transaction (as shown in Figure 8).

Figure 7 The process of that how the retailer buys goods.

Figure 8 The process of how the producer accepts the transaction and checks goods.

6 Conclusions

In this work, we proposed a supply chain system based on blockchain technologies with traceability. In this system, all the processes of commodities could be recorded in production. In addition we used the smart contract technology to record every raw material and processed product. Through the system, information about the products sold were presented, so that consumers could understand the source of the product more clearly. By using smart contracts as a bridge between the roles of the supply chain, whether it

was demand, input or output were exposed on the chain. Through our proposed system, more safe and convenient supply chain system could be used.

References

1. F. Dabbene, P. Gay, and C. Tortia, "Traceability issues in food supplychain management: A review," *Biosystems Engineering*, vol. 120, pp. 65– 80, 2014.
2. J. Gualandris, R. D. Klassen, S. Vachon, and M. Kalchschmidt, "Sustainable evaluation and verification in supply chains: Aligning and leveraging accountability to stakeholders," *Journal of Operations Management*, vol. 38, pp. 1 – 13, 2015.
3. S. A. Abeyratne and R. P. Monfared, "Blockchain ready manufacturing supply chain using distributed ledger," *International Journal of Research in Engineering and Technology*, vol. 05, September 2016.
4. S. Nakamoto, "Bitcoin:A Peer-to-Peer Electronic Cash System," Manubot, 2019.
5. K. Toyoda, P. T. Mathiopoulos, I. Sasase, T. Ohtsuki, "A novel blockchain-based product ownership management system (POMS) for anti-counterfeits in the post supply chain," *IEEE Access*, vol. 5. pp. 17465-17477, 2017.
6. F. Tian, "An agri-food supply chain traceability system for China based on RFID & blockchain technology," *Proc. of the IEEE 13th international conference on service systems and service management (ICSSSM)*, 2016.
7. A. Bechini, M. G. Cimino, F. Marcelloni, and A. Tomasi, "Patterns and technolo-gies for enabling supply chain traceability through collaborative e-business," *Information and Software Technology, vol.* 50, no. 4, pp. 342– 359, 2008.
8. M. Kolvart, M. Poola, A. Rull, "Smart contracts," The Future of Law and etechnologies. *Springer*, Cham, 133-147. 2016.
9. J. Benet, "Ipfs-content addressed, versioned, p2p file system," *arXiv preprint arXiv:1407.3561* 2014.
10. S. S. Kamble, A. Gunasekaran, R. Sharma, "Modeling the blockchain enabled traceability in agriculture supply chain," *International Journal of Information Management* (2020): 101967.
11. Z. Wang, T. Wang, H. Hu, J. Gong, X. Ren, Q. Xiao, "Blockchain-based framework for improving supply chain traceability and information sharing in precast construction," *Automation in Construction* (2020): 103063.
12. K. Behnke, M. Janssen, "Boundary conditions for traceability in food supply chains using blockchain technology," *International Journal of Information Management* (2020): 101969.
13. S. Saberi, M. Kouhizadeh, J. Sarkis, L. Shen "Blockchain technology and its relationships to sustainable supply chain management," *International Journal of Production Research* (2019): 2117-2135.

WiFi Location-based 3D Map for Device Connections

Chao-Tung Yang[1]*, Chen-Kun Tsung[2], Wei-Chen Chen[1], Jia-Hao Zhang[1], Shih-Kuang Chang[3], and Ming-Shang Hsu[3]

[1] Department of Computer Science,
Tunghai University, Taichung City 407302, Taiwan (R.O.C.)
[2] Department of Computer Science and Information Engineering,
National Chin-Yi University of Technology, Taichung City 41170, Taiwan (R.O.C.)
[3] Krystal Technology Co., Ltd, Taiwan (R.O.C.)

Abstract. Wireless network stability is critical for organizations. For at least part of its day-to-day activities, most companies rely on a strong internet connection. It is highly necessary to show the fast and high capacity of Wi-Fi is. Wireless network management also plays a significant part in supplying the user with quality service and helping the system administrator manage and track the network infrastructure. The user must be able to stay connected. In this paper, the Wi-Fi connection engine was applied using the Analysis and Locations Engine (ALE). ALE provides the information, including MAC address, location, floor information, the building where the device belongs. From this data, we presented the web visualization of the Wi-Fi network monitoring system on the map using Cesium.

Keywords: Wireless Network, Analysis and Positioning Engine, Mobile Device, 3D Model Map, WiFi Positioning

1 Introduction

Wireless monitoring has the power to change the way organizations conduct business. Everyone has a mobile phone in the current era, so this has become a significant source of data. What kind of data should be collected becomes a topic [1]. For commercial usage, from selling goods and analyze the customer group. If we can more directly understand the passenger flow, Wi-Fi positioning is an effective method, because we can detect the user's location, and then make a heat map of the masses, and master the location of the masses at the first time, which helps in-depth analysis of the relationship between time and customer groups [2][3].

In today's society, it can be said that everyone has a mobile phone, and ALE can retrieve the data of its existence. If we can capture the current mobile device's data, we can know how many people there are [4]. If we can use this

* Corresponding author: ctyang@thu.edu.tw

information by visualizing, we can know the status of population movement in various buildings. It can be used to observe the status of population movement in future org activities, and then it can then further improve the movement line [5]. In addition, when a disaster occurs, the golden rescue time is very urgent. Immediately grasping the victims who hold the mobile device can effectively help the rescue route planning, reduce the rescue time, improve the survival rate of the wounded, reduce the loss of personnel to the minimum, and let the disaster relief Personnel can perform rescue tasks more efficiently [6].

2 Background Review and Related Work

In this section, we discuss the background review of this work, including ALE, ALE positioning system and Cesium.

2.1 ALE

ALE is a set of analysis and positioning engine provided by the company Aruba. To understand ALE, we start with the entire architecture and then look at what services ALE provides.The bottom layer in Figure 1 is the WiFi user's machine and AP. Smartphones and laptops are usually WiFi user's machines. The Access Point (AP) is a wireless access point that detects the mobile phone. It sends information to the Local Controller, which shows that the Local Controller is used to collect the detected information detected by a cluster of AP machines. Then integrate the information and send it to the ALE engine. It is worth mentioning that there is a line between ALE and AirWave. It says VisualRF Maps. According to the instructions on the AirWave document The module provides a picture that shows The wireless signal strength and coverage of the machine.

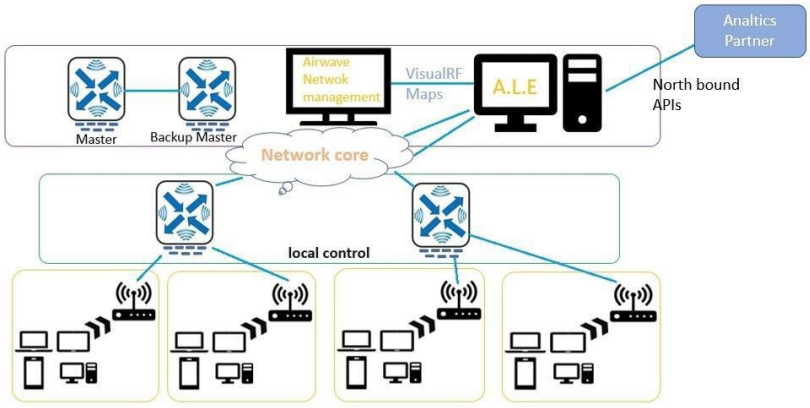


Fig. 1. Diagram of Relationship Between ALE and Other Devices

2.2 ALE positioning system

There are two ALE positioning systems, namely Context with Device Location (Estimated) and Context with Device Location (Calibration), referred to as Estimated and Calibration [7]. The Estimated set is based on the information provided by AirWave. According to the ALE document, you can get the map and AP machine location information through AirWave. Use AP-AP RSSI 5 message to create a path-loss model. As mentioned on Wikipedia, path-loss mainly refers to that the wall will affect the propagation of wireless signals, and RSSI is the strength of the signal, and the full name is Received Signal Strength Indicator. The AP can receive the signal strength [8][9]. After understanding the information provided by Airwave, let's look at the algorithm of ALE. The Estimated set uses RSSI and lateration based. In simple terms, first find the relationship between signal strength and distance, and then refer to the information of different nearby machines to get the location

However, this has a disadvantage. If there is only one AP around, the accuracy will be greatly reduced. If there are three or more APs, the accuracy will be more accurate. Another system is Calibration. According to the ALE document, the positioning accuracy is more accurate because of the use of data from fingerprinting. Later, the web service Aruba Nao Campus and the mobile application Aruba Nao Logger were developed for practice fingerprinting [10]. The general operation method is to first define our target building and route in Aruba Nao Campus, and then use Aruba Nao Logger to actually go to the location to give accurate location information and test positioning, and check whether the planned route is wrong or missing. So, we need to take the mobile phone to this point to provide information to create an actual comparison data table to help us locate [11].

2.3 Cesium

Cesium is a JavaScript database that can directly create 3D maps on a web browser without the need to install additional plug-ins. It uses WebGL to achieve hardware accelerated graphics, adjust for dynamic data visualization, and cross-platform and cross-browser [12]. Cesium can greatly improve the readability of data, by rendering layers, importing models, and applying terrain modules. Cesium is based on the Apache open source agreement and supports commercial and non-commercial free use.

3 System Design and Implementation

3.1 System Architecture

Our system is based on the Linux operating system. The Linux operating system is more flexible than Windows, and is more suitable for setting up a web server.The architecture diagram is shown in Figure 2. The system we use can be divided into two parts. On the left is ALE with script to send data to MySQL.

MySQL interacts with the website through PHP. Use Apache, and will use HTML, CSS, JS, and finally present our website.

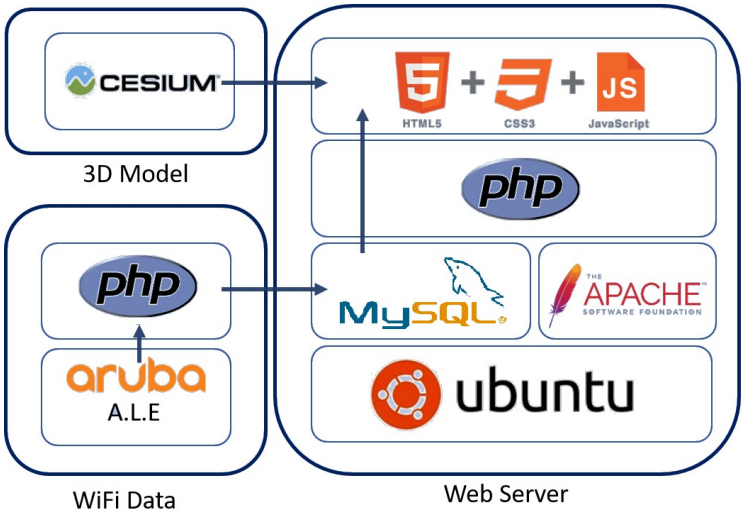


Fig. 2. Software Architecture

3.2 Model Building

Assuming that only the points of the user's device are on the map, it does not look so beautiful, so we chose to use SketchUP to build our model software. Through SketchUP's Extensiom Warehouse, there are many plug-ins with special shapes, which make the East China Sea rich. Artistic buildings can be made, as shown in Figure 3.

4 Experimental Results

Through the process of the previous chapter, we have actually made it possible to receive the ALE API information system, and successfully presented it on the map. We also received the location information and made a heat map, so that the user can more easily understand the population distribution s position. The ALE API information is presented on the 3D map and contains its detailed information, as shown in Figure 4.

5 Conclusions and Future Work

When designing the system architecture, most of them are biased towards functional requirements. However, today's requirements for non-functional require-

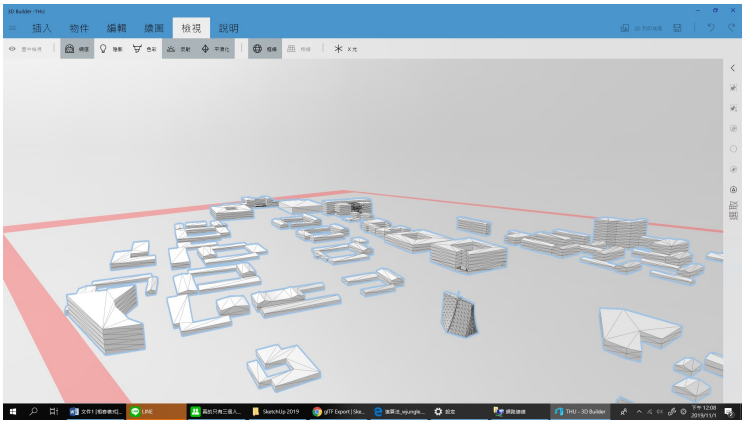


Fig. 3. SketchUP Implementation Campuse

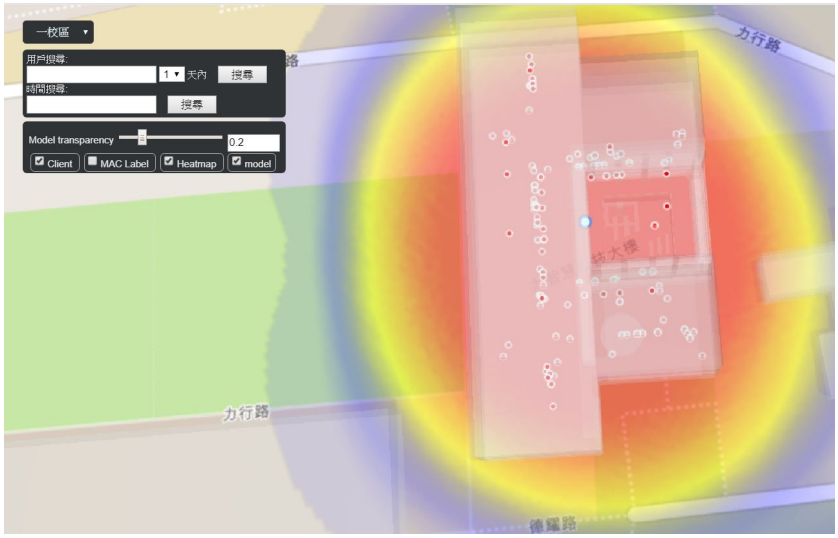


Fig. 4. Crowd Density Heat Map Presentation

ments are also very high. Therefore this implementation can also allow us to practice non-functional requirements. Since this is our first large-scale implementation, it took some time to research and modify the use of the database or the writing of the web page to achieve the goal gradually. In the process of implementing this system, we understand how to obtain the data, convert and organize it. Finally, present the sorted data.

Acknowledgement

This work was sponsored by the Ministry of Science and Technology (MOST), Taiwan, under Grant No. 108-2622-E-029-007-CC3, 108-2221-E-029-010-, and 108-2745-8-029-007-.

References

1. Junhai Luo and Liang Fu. A smartphone indoor localization algorithm based on wlan location fingerprinting with feature extraction and clustering. *Sensors*, 17(6):1339, 2017.
2. Han Zou, Zhenghua Chen, Hao Jiang, Lihua Xie, and Costas Spanos. Accurate indoor localization and tracking using mobile phone inertial sensors, wifi and ibeacon. In *2017 IEEE International Symposium on Inertial Sensors and Systems (INERTIAL)*, pages 1–4. IEEE, 2017.
3. Alexandra Zayets and Eckehard Steinbach. Robust wifi-based indoor localization using multipath component analysis. In *2017 International Conference on Indoor Positioning and Indoor Navigation (IPIN)*, pages 1–8. IEEE, 2017.
4. Ye Tao and Long Zhao. A novel system for wifi radio map automatic adaptation and indoor positioning. *IEEE Transactions on Vehicular Technology*, 67(11):10683–10692, 2018.
5. Nguyen Dinh-Van, Fawzi Nashashibi, Nguyen Thanh-Huong, and Eric Castelli. Indoor intelligent vehicle localization using wifi received signal strength indicator. In *2017 IEEE MTT-S international conference on microwaves for intelligent mobility (ICMIM)*, pages 33–36. IEEE, 2017.
6. Afaz Uddin Ahmed, Neil W Bergmann, Reza Arablouei, Branislav Kusy, Frank De Hoog, and Raja Jurdak. Fast indoor localization using wifi channel state information. In *2018 17th ACM/IEEE International Conference on Information Processing in Sensor Networks (IPSN)*, pages 120–121. IEEE, 2018.
7. JC Aguilar Herrera, Paul-Gerhard Plöger, André Hinkenjann, Jens Maiero, M Flores, and A Ramos. Pedestrian indoor positioning using smartphone multi-sensing, radio beacons, user positions probability map and indoorosm floor plan representation. In *2014 international conference on indoor positioning and indoor navigation (IPIN)*, pages 636–645. IEEE, 2014.
8. Julie Yixuan Zhu, Jialing Xu, Anny Xijia Zheng, Jiaju He, Chaoyi Wu, and Victor OK Li. Wifi fingerprinting indoor localization system based on spatio-temporal (st) metrics. In *2014 International Conference on Indoor Positioning and Indoor Navigation (IPIN)*, pages 611–614. IEEE, 2014.
9. Hu Xujian and Wang Hao. Wifi indoor positioning algorithm based on improved kalman filtering. In *2016 International Conference on Intelligent Transportation, Big Data & Smart City (ICITBS)*, pages 349–352. IEEE, 2016.
10. Hu Xu Jian and Wang Hao. Wifi indoor location optimization method based on position fingerprint algorithm. In *2017 International Conference on Smart Grid and Electrical Automation (ICSGEA)*, pages 585–588. IEEE, 2017.
11. Pratik Palaskar OnkarPathak, Rajesh Palkar, and Mayur Tawari. Wi-fi indoor positioning system based on rssi measurements from wi-fi access points–a tri-lateration approach. *International Journal of Scientific & Engineering Research*, 5(4), 2014.
12. Cesium. Changing how the world views 3d: Build world-class 3d geospatial applications.

Using Spark Distributed Deep Learning to Analyze NetFlow in Data Lake System

Cheng-Tian Jiang[1], Chao-Tung Yang[1*], Yu-Wei Chan[2], Endah Kristiani[34], and Jung-Chun Liu[1]

[1] Department of Computer Science,
Tunghai University, Taichung City 407302, Taiwan (R.O.C.)
[2] College of Computing and Informatics,
Providence University, Taichung City, Taiwan, (R.O.C.)
[3] Department of Industrial Engineering and Enterprise Information,
Tunghai University, Taichung City 407302, Taiwan (R.O.C.)
[4] Department of Informatics, Faculty of Engineering and Computer Science,
Krida Wacana Christian University, Jakarta 11470, Indonesia

Abstract. The usage of data to improve information security has become the focus of research in the development of big data and artificial intelligence. However, the traditional single database technology is unable to handle the current complex and massive data processing. While network attacks have also occurred frequently and produce massive data. This paper proposes a framework for storing and analyzing web log data based on the concept of data lakes. The need for processing, accessing, querying, analyzing, and visualizing large amounts of data in a short time is also critical. Therefore, the Grafana tool was used to provide the network administrator a real-time monitoring and visualization system management.

Keywords: Data Lake, Distributed Deep Learning, NetFlow Analysis, Cloudera Cluster

1 Introduction

According to a security report published by Netscout: In 2019, a total of 8.4 million Distributed Denial of Service (DDoS) were detected, equivalent to 16 occurrences per minute [1]. As the number of cyber attack incidents has increased year by year, the attack methods have become more and more diversified. How to prepare the system architecture that automatically analyze whether there is an attack in the NetFlow log under big data has become a very important concern [2][3].

In the current state of artificial intelligence development, a large amount of training data and computing power are extremely important. The concept of data lakes and parallel computing are also important in the future development direction [4][5]. This paper uses the Cloudera platform to run Spark on the

* Corresponding author: ctyang@thu.edu.tw

resource scheduling platform Yarn for distributed deep learning purpose. It provides relevant experience and practical data for a high-performance distributed deep learning system [6]. The specific goals are as follows:

1. Use Cloudera to build a data lake, the system can access a large amount of real-time data, and the files are in various formats such as text, pictures and videos.
2. Use Kafka to collect streaming data and store in HBase.
3. Use Spark for the data preprocessing, and cut the dataset of the model training process by cross-validation.

2 Background Review and Related Work

In this section, we discuss the background review of this work, including Data Lake, Spark Deep Learning, and other related works.

2.1 Data Lake

This paper stores Data Materials in a pool implement by Hbase. The processed data is temporarily stored in MySQL. The main purpose is to accelerate the real-time query data processing and reduce the server loading [7]. When the raw data is needed for historical record queries, neural network training, or other purposes, Hive can be used to query big data through MapReduce [8][9].

2.2 Spark Deep Learning

Spark is a parallel computing framework developed by Algorithms Machines and People Lab, which focuses on SQL query, stream processing, machine learning, and deep learning [10]. Although several scholars also use Spark to analyze NetFlow, but mostly for machine learning methods [11].

2.3 Related Works

Hongyu Liu et al. [12] proposed a detection method for instant port-to-port, using PL-CNN (a convolutional neural network-based payload classification method) and PL-RNN (Neural Network-Based Payload Classification Method) performs attack detection. The two methods learn feature representation from the original payload without feature engineering and support end-to-end detection.

Duygu Sinanc Terzi et al.[13] proposed a new unsupervised anomaly detection method. Its purpose is to determine the anomaly caused by a UDP flood attack on a specific IP. This approach is implemented on public NetFlow data in case studies.

Rick Hofstede et al. [14]explain all phases of NetFlow's traffic output and typical traffic monitoring settings, covering the full range from packet capture to data analysis.

3 System Design and Implementation

3.1 System Architecture

The hardware architecture of this paper is shown in Figure 1. The data source is from multiple network switches. NetFlow data is accumulated into the Cloudera cluster through the NetFlow Collector. Each node in the cluster is in 1Gb Ethernet connection specification, Intel Core i7- 4770 CPU, 16GB memory and 1TB HDD storage space.

This system is built on the Cloudera big data platform. The operating systems are all running 64-bit CentOS7. Real-time data is written to HBase by means of Kafka using Spark Streaming. The old system data is manually imported into HBase from MySQL using Sqoop. The data query part uses HBase to integrate Hive and Impala as search engines. The collected data will be identified by the trained deep learning module and then stored in MySQL to facilitate Grafana for visualization purpose. The MySQL database set up here only stores the latest real-time data. The intention is to improve the query speed of real-time data visualization. The architecture of this system is shown in Figure 2.

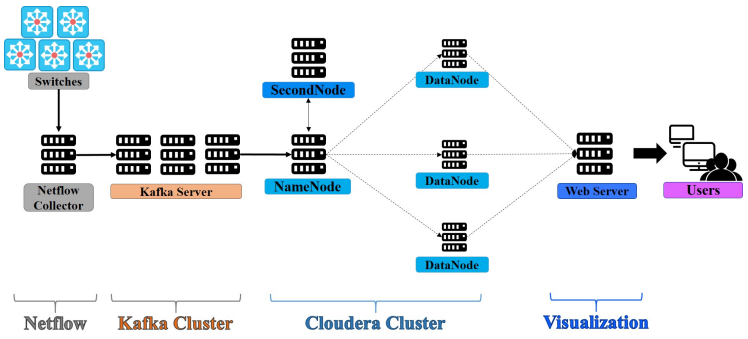


Fig. 1. Hardware Architecture

3.2 Data Collection

The data collected in this paper is a campus wired network. The average data volume is about 100,000 (16MB) per minute. The number of data in a day is more than 100 million. There are three ways to write data:

1. Real-time data: use Kafka to collect real-time streaming data.
2. Historical data: use Sqoop to import from related database.
3. Exception: Use ImportTsv to import text files. Confusion Matrix.

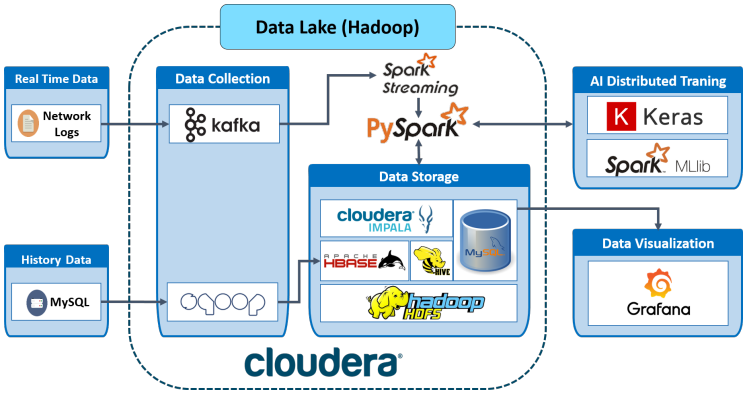


Fig. 2. Software Architecture

3.3 Data Storage

This system uses Hive and HBase based on HDFS as the storage database. The data is mainly stored in HBase, because HBase has the characteristics of high transmission and low latency, which is very suitable for reading and writing operations of big data. Hive responsible for converting SQL syntax to MapReduce Jobs to search the data on HDFS, this usually used for offline big data queries.

4 Experimental Results

4.1 Experimental Environment

This section introduces our hardware experiment environment. A total of eight hosts are used in this experiment, and the data lake is composed of five hosts. Ideally, an independent Kafka Cluster should be built. Because of resource constraints, this experiment builds Kafka Cluster with NetFlow Collector and the data lake. Detailed equipment is shown in Table 1.

4.2 Data Visualization Result

The grafana provides usage information of each machine's CPU, network, and hard disk. The chart allows the cluster administrator to quickly grasp the status of the cluster without having to query the working status separately through commands or the background of each service. Figure 3 is the cluster CPU status, HDFS read and write, hard disk read and write, and network traffic visualization in sequence.

5 Conclusions and Future Work

In this study, we used Cloudera to build a data lake system, which is a completely open source management platform. It can be easily built by Cloudera users. The

Table 1. Computing Environment

Host	CPU	RAM	Disk	Roles
NetFlow Collector	Intel(R) Core(TM) i7-4790 CPU @ 3.40GHz	64G	1TB	NetFlow Collector Kafka Producer Kafka Broker
Broker	Intel(R) Core(TM) i7-4790 CPU @ 3.40GHz	16G	1TB	Kafka Broker
master	Intel(R) Core(TM) i7-4770 CPU @ 3.40GHz	64G	1TB	Namenode Kafka Consumer Kafka Broker
node01	Intel(R) Core(TM) i7-4770 CPU @ 3.40GHz	16G	1TB	Datanode
node02	Intel(R) Core(TM) i7-4770 CPU @ 3.40GHz	16G	1TB	Datanode
node03	Intel(R) Core(TM) i7-4770 CPU @ 3.40GHz	16G	1TB	Datanode
node04	Intel(R) Core(TM) i7-4770 CPU @ 3.40GHz	16G	1TB	Secondary Namenode
Web Server	Intel(R) Core(TM) i9-9980XE CPU @ 3.00GHz	48G	500GB	Grafana Server MySQL

Fig. 3. Cluster Hardware Usage

system supports the collection, storage, and query of large amounts of data. Also, it provides a series of functions such as analysis, distributed neural network training and visualization. Grafana is used to visualize the data, draw a variety of cluster environments and network usage behavior charts, and provide network managers with a real-time monitoring platform.

Acknowledgement

This work was sponsored by the Ministry of Science and Technology (MOST), Taiwan, under Grant No. 108-2622-E-029-007-CC3, 108-2221-E-029-010-, and 108-2745-8-029-007-.

References

1. Xiaoyong Yuan, Chuanhuang Li, and Xiaolin Li. Deepdefense: identifying ddos attack via deep learning. In *2017 IEEE International Conference on Smart Computing (SMARTCOMP)*, pages 1–8. IEEE, 2017.
2. Abebe Abeshu Diro and Naveen Chilamkurti. Distributed attack detection scheme using deep learning approach for internet of things. *Future Generation Computer Systems*, 82:761–768, 2018.
3. Rafał Kozik. Distributing extreme learning machines with apache spark for netflow-based malware activity detection. *Pattern Recognition Letters*, 101:14–20, 2018.
4. Xiaoyi Lu, Haiyang Shi, Rajarshi Biswas, M Haseeb Javed, and Dhabaleswar K Panda. Dlobd: A comprehensive study of deep learning over big data stacks on hpc clusters. *IEEE Transactions on Multi-Scale Computing Systems*, 4(4):635–648, 2018.
5. Markus Ring, Daniel Schlör, Dieter Landes, and Andreas Hotho. Flow-based network traffic generation using generative adversarial networks. *Computers & Security*, 82:156–172, 2019.
6. Chao-Tung Yang, Jung-Chun Liu, Endah Kristiani, Ming-Lun Liu, Ilsun You, and Giovanni Pau. Netflow monitoring and cyberattack detection using deep learning with ceph. *IEEE Access*, 8:7842–7850, 2020.
7. Amr A Munshi and Yasser Abdel-Rady I Mohamed. Data lake lambda architecture for smart grids big data analytics. *IEEE Access*, 6:40463–40471, 2018.
8. Chao-Tung Yang, Shuo-Tsung Chen, Wei-Hsun Cheng, Yu-Wei Chan, and Endah Kristiani. A heterogeneous cloud storage platform with uniform data distribution by software-defined storage technologies. *IEEE Access*, 7:147672–147682, 2019.
9. Chen-Kun Tsung, Hsiang-Yi Hsieh, and Chao-Tung Yang. An implementation of scalable high throughput data platform for logging semiconductor testing results. *IEEE Access*, 7:26497–26506, 2019.
10. Fabrizio Carcillo, Andrea Dal Pozzolo, Yann-Aël Le Borgne, Olivier Caelen, Yannis Mazzer, and Gianluca Bontempi. Scarff: a scalable framework for streaming credit card fraud detection with spark. *Information fusion*, 41:182–194, 2018.
11. Leonel Cruz, Ruben Tous, and Beatriz Otero. Distributed training of deep neural networks with spark: The marenostrum experience. *Pattern Recognition Letters*, 125:174–178, 2019.
12. Hongyu Liu, Bo Lang, Ming Liu, and Hanbing Yan. Cnn and rnn based payload classification methods for attack detection. *Knowledge-Based Systems*, 163:332–341, 2019.
13. Duygu Sinanc Terzi, Ramazan Terzi, and Seref Sagiroglu. Big data analytics for network anomaly detection from netflow data. In *2017 International Conference on Computer Science and Engineering (UBMK)*, pages 592–597. IEEE, 2017.
14. Rick Hofstede, Pavel Čeleda, Brian Trammell, Idilio Drago, Ramin Sadre, Anna Sperotto, and Aiko Pras. Flow monitoring explained: From packet capture to data analysis with netflow and ipfix. *IEEE Communications Surveys & Tutorials*, 16(4):2037–2064, 2014.

Using Long Short-Term Memory Deep Learning for Short-Term PM2.5 Prediction in Taiwan

Hao Lin[1], Wen-Yen Lin[2], Chao-Tung Yang[1]*, and Jwu-Rong Lin[3]

[1] Department of Computer Science,
Tunghai University, Taichung City 407302, Taiwan (R.O.C.)
[2] Department of Information Management,
National Taichung University of Science and Technology, Taichung City 40401, Taiwan (R.O.C.)
[3] Department of International Business,
Tunghai University, Taichung City 407302, Taiwan (R.O.C.)

Abstract. Predicting air pollution will enable the people and administrative units to have a clear and powerful quantitative basis for their lives and government orders. In this study, we will use deep learning techniques, recurrent neural networks, and long- and short-term memory models to use serialized data to make predictions about the future level of air pollution. And the deep learning model prediction visualization results are presented on the web page, to assist developers to conduct experiments and to facilitate user to use.

Keywords: Air Pollution, Deep Learning, Long Short-Term Memory

1 Introduction

In recent years, due to climate change, the shortage of natural resources and the destruction of the living environment, coupled with more and more national modernization and other factors, the air pollution in various regions has become more and more serious [1]. At the same time, air pollution has caused many environmental disasters and affected the living space and the quality of life. With the improvement of environmental protection awareness, people are paying more and more attention to having good air quality. Many people realize that the impact of air pollution on the environment and human health is very great [2]. Therefore, air pollution has become the most important issue. Time series data forecasting is a challenging forecasting model study, because time series need to consider the time dependence between input variables. Recurrent neural networks (RNNs) can deal with time series problems because they can use their output as the next input to maintain the state from the previous iteration to the next iteration. His cyclic ability has proved that RNN is specifically used for Powerful engine for time series data [3][4].

* Corresponding author: ctyang@thu.edu.tw

2 Background Review and Related Work

2.1 CNN

Convolutional neural network architectures often include single or multiple convolution layers, pooling layers, and are connected to a fully-connected layer at the output. The function of the convolution layer is to capture pictures. The action of taking features, after finding the best features, and then classifying, and the pooling layer is often added between the convolutional layers. It is a way to compress pictures and retain important information. Finally, at the end of the model A fully connected layer will be added to the face, and the role of the fully connected layer is to be used to achieve classification [5][6].

2.2 RNN

RNN, the purpose of recurrent neural network is to process sequence data. The reason why RNN is called a recurrent neural network is because the output in the sequence is related to the previous output. Specifically, the network will remember the previous information and apply it to the calculation of the current output, but hide it. Neurons between layers are no longer connected, but are connected to each other, and the input of the hidden layer also includes not only the output of the input layer, but also the output of the hidden layer of the previous time [7].

2.3 LSTM

LSTM, Long Short-Term Memory is a time recursive neural network. LSTM solves the problem of gradient vanishing, and the key to the problem of gradient disappearance is that it has more gates. 'S function is to filter the information flow obtained by the model [8][9][5].

2.4 GRU

GRU replaces the forget gate and input gate in the LSTM with an update gate, and hides the unit status and The status is merged. Due to the simplified network structure of GRU, its training speed is faster, and it has similar capabilities to LSTM. When processing data sets with fewer data parameters, GRU can converge faster to LSTM [10].

2.5 Related Works

Ma et al. [11] proposes a transfer learning-based stacked long-term bidirectional memory (TLS-BLSTM) network to forecast air pollution for the new stations which is lacking in data. The proposed method combines advanced deep learning techniques and transfer learning strategies to enhance prediction by transferring the knowledge learned from existing air quality stations into new stations.

Zhang et al. [12] discusses the cross-correlations between concentration of PM2,5 and meteorological factors (such as, temperature, air pressure, relative humidity and wind speed) in Beijing and Hong Kong. They use the multifractal detrended cross-correlation analysis (MF-DCCA) to examine cross-correlations, and then research the asymmetric features of cross-correlations by multifractal asymmetric detrended cross-correlation analysis (MF-ADCCA).

Zao et al. [13] presented a data-driven model, called as long-term memory-fully connected neural network (LSTM-FC). Their forecasting model consists of two parts: First, by using a long-term memory (LSTM)-based temporal simulator to model local PM2.5 contamination variation; and second, by using a neural network-based spatial combinator to capture spatial relations between central station PM2.5 contamination and adjacent station contamination. The results show that their model of neural network LSTM-FC gives better performance in predictive terms.

3 System Design and Implementation

3.1 System Architecture

In this study, we used Linux as the operating system of this study through Ubuntu 16.04, and conducted machine learning and deep learning experiments through Tensorflow and Keras to establish a prediction model that targets air pollution prediction and uses PM2.5 as the main forecasting objects. In the establishment of the model, we set up several models to make predictions in counties and cities, in order to expect that different models can predict PM2.5 according to different geographical conditions in different regions As a reference, people can quickly understand the level of air pollution.The architecture of this system is shown in Figure 1.

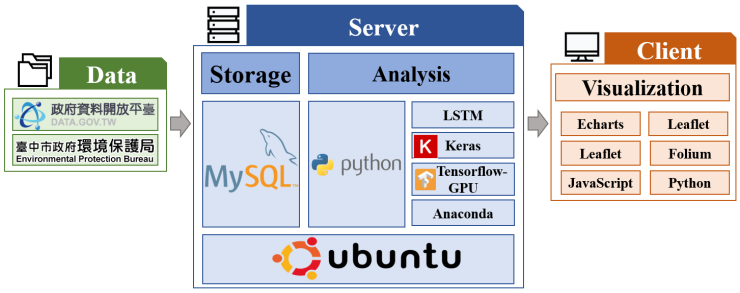


Fig. 1. System Architecture Diagram

3.2 Data correlation

In this study, each air pollution factor in the air pollution quality data set and PM2.5 were tested for Pearson correlation coefficient and Spearman correlation coefficient, and some air pollution factors that showed high correlation were screened out. At the same time, we also conducted a principal component analysis of each air pollution factor in the air pollution quality data set, observed the correlation between each factor, after discussion and experiment, found the appropriate factor as a training machine learning prediction model Input data in order to expect better prediction results,

3.3 Building and training prediction models

After setting the activation function, loss function and optimizer, you can start to adjust the number of neural layers of the machine learning model and the neurons of each neural network layer. Generally speaking, the more neural layers and nerves the number of meta-models for model training can learn more features, but it is also more likely to cause over fitting. At this time, it is necessary to use the conventional processing mentioned later, including the discarding method, The L1 and L2 conventional methods, etc., which often happen, need special attention. Finally, after stacking this machine learning model architecture, we need to select hyper parameters, each hyper parameter represents a different meaning, and epoch represents how many total training data sets have been trained during this training process in Keras, the parameter update is carried out in batches, which is the so-called small batch reduction algorithm. It divides the data into several groups. This is called batch, and batch size refers to the number of samples in each group of data.

4 Experimental Results

4.1 Correlation

The choice of input variables will affect the accuracy of this air pollution premodel. Choosing data with low correlation to enter the model for prediction will make it difficult for the machine to grasp the relationship between the data. In order to avoid such problems as much as possible In this experiment, the correlation of air pollution factors will be measured first, and several pollution factors will be selected as the input of this prediction model, in order to expect better prediction results. Table 2 shows the input variable correlation.
It can be found from Table 1 above that the correlation factors obtained from the experimental results are actually the same. The main difference is that Spearman's related experiments have more NOx. However, Spearman is the main evaluation standard for measuring nonlinear values. , So this experiment will include NOx in the air factor reference.

Table 1. Input variable correlation

	Pearson correlation	Spearman correlation
σ(PM2.5,SO2)	0.369	0.337
σ(PM2.5,CO)	0.453	0.449
σ(PM2.5,O3)	0.059	0.063
σ(PM2.5,PM10)	0.765	0.673
σ(PM2.5,NOx)	0.299	0.320
σ(PM2.5,NO)	0.174	0.186
σ(PM2.5,NO2)	0.349	0.338
σ(PM2.5,THC)	0.379	0.380
σ(PM2.5,NMHC)	0.379	0.373
σ(PM2.5,TEMP)	0.197	0.174
σ(PM2.5,RAINFALL)	-0.143	0.221
σ(PM2.5,CH4)	0.324	0.314
σ(PM2.5,UVB)	0.067	0.082
σ(PM2.5,RH)	0.056	-0.006

4.2 Comparison of neural network prediction models

In the experiment of this study, the model was selected as 6 kinds of neural networks such as CNN, RNN, LSTM, GRU, bidirectional LSTM and bidirectional GRU, adding multiple variables into the model, training, and obtaining training results and evaluation values. The following is a comparison of the prediction results of the PM2.5 concentration of CNN, RNN, LSTM, GRU, bidirectional LSTM and bidirectional GRU in the next 8 hours, including the true value, predicted value, and error value of each hour, etc. In Table 3, we can see that the average error value of the prediction result of LSTM is smaller than the average error value of the prediction result of other neural networks. Therefore, we can think that LSTM is more suitable for the type of air pollution data used in this study than other neural networks. Therefore, in the subsequent experiments, we will choose LSTM as the neural network of the prediction model.

Table 2. Comparison of different model's prediction results

Time	Real	CNN	RNN	LSTM	GRU	Bi-LSTM	Bi-GRU
1 hr	12	12	12	12	12	11	12
2 hr	17	19	19	17	16	15	16
3 hr	20	21	20	20	20	17	19
4 hr	27	23	27	25	22	23	24
5 hr	24	22	27	25	25	21	25
6 hr	28	25	30	27	27	27	24
7 hr	29	26	30	30	29	26	27
8 hr	31	26	32	29	29	29	29
Average Error		2.5	1.1	1.0	1.25	2.3	1.75

4.3 Visualization

Use JavaScript, Leaflet and E-charts together to mark the location of each station in Taiwan in dotted mode, and present the actual value and predicted value in a line chart. The following Fgure 2 is the distribution of each station in Taiwan. Dotted presentation can help users quickly and easily find the area they need to query. The following Figure 3 shows the actual value of PM2.5 at each station and the predicted value of the next 8 hours.

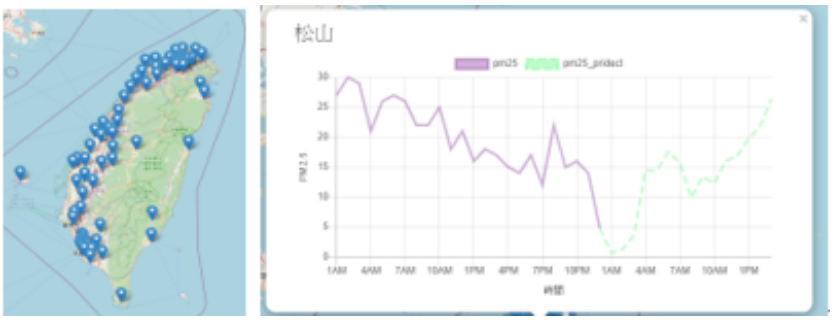


Fig. 2. Map of All Stations and Visualization of Predict Result

5 Conclusions

This study uses correlation calculation methods to provide a clear quantitative standard, and combines the correlation calculation method and the Lag time mechanism to find the time points with high correlation between pollution factors. In this study, we chose to use a variety of neural networks to build deep learning models to effectively predict short-term future air quality conditions, and use MAPE as a model evaluation method to compare the prediction accuracy and error values of different neural network prediction models. Through the calculation of the inverse distance weighting method, the PM2.5 concentration in the surrounding areas of each air quality monitoring station is estimated, and the predicted future PM2.5 concentration data is presented on the web page for users by means of maps and line charts reference.

Acknowledgement

This work was sponsored by the Ministry of Science and Technology (MOST), Taiwan, under Grant No. 108-2221-E-029-010 and 109-2119-M-029-001-A. This study was also sponsored by the Higher Education Sprout Project, Ministry of Education, Taiwan.

References

1. Chao-Tung Yang, Lung-Ying Lin, Yu-Tse Tsan, Po-Yu Liu, and Wei-Cheng Chan. The implementation of a real-time monitoring and prediction system of pm2. 5 and influenza-like illness using deep learning. *Journal of Internet Technology*, 20(7):2237–2245, 2019.
2. Chao-Tung Yang, Yuan-An Chen, Yu-Wei Chan, Chia-Lin Lee, Yu-Tse Tsan, Wei-Cheng Chan, and Po-Yu Liu. Influenza-like illness prediction using a long short-term memory deep learning model with multiple open data sources. *The Journal of Supercomputing*, pages 1–27, 2020.
3. Yagmur Gizem Cinar, Hamid Mirisaee, Parantapa Goswami, Eric Gaussier, and Ali Aït-Bachir. Period-aware content attention rnns for time series forecasting with missing values. *Neurocomputing*, 312:177–186, 2018.
4. Bo Zhang, Hanwen Zhang, Gengming Zhao, and Jie Lian. Constructing a pm2. 5 concentration prediction model by combining auto-encoder with bi-lstm neural networks. *Environmental Modelling & Software*, 124:104600, 2020.
5. SHI Xingjian, Zhourong Chen, Hao Wang, Dit-Yan Yeung, Wai-Kin Wong, and Wang-chun Woo. Convolutional lstm network: A machine learning approach for precipitation nowcasting. In *Advances in neural information processing systems*, pages 802–810, 2015.
6. Yanlin Qi, Qi Li, Hamed Karimian, and Di Liu. A hybrid model for spatiotemporal forecasting of pm2. 5 based on graph convolutional neural network and long short-term memory. *Science of the Total Environment*, 664:1–10, 2019.
7. Ching-Fang Lee, Chao-Tung Yang, Endah Kristiani, Yu-Tse Tsan, Wei-Cheng Chan, and Chin-Yin Huang. Recurrent neural networks for analysis and automated air pollution forecasting. In *International Conference on Frontier Computing*, pages 50–59. Springer, 2018.
8. Qunli Wu and Huaxing Lin. Daily urban air quality index forecasting based on variational mode decomposition, sample entropy and lstm neural network. *Sustainable Cities and Society*, 50:101657, 2019.
9. Ricardo Navares and José L Aznarte. Predicting air quality with deep learning lstm: Towards comprehensive models. *Ecological Informatics*, 55:101019, 2020.
10. Youru Li, Zhenfeng Zhu, Deqiang Kong, Hua Han, and Yao Zhao. Ea-lstm: Evolutionary attention-based lstm for time series prediction. *Knowledge-Based Systems*, 181:104785, 2019.
11. Jun Ma, Zheng Li, Jack CP Cheng, Yuexiong Ding, Changqing Lin, and Zherui Xu. Air quality prediction at new stations using spatially transferred bi-directional long short-term memory network. *Science of The Total Environment*, 705:135771, 2020.
12. Chen Zhang, Zhiwei Ni, and Liping Ni. Multifractal detrended cross-correlation analysis between pm2. 5 and meteorological factors. *Physica A: Statistical Mechanics and its Applications*, 438:114–123, 2015.
13. Jiachen Zhao, Fang Deng, Yeyun Cai, and Jie Chen. Long short-term memory-fully connected (lstm-fc) neural network for pm2. 5 concentration prediction. *Chemosphere*, 220:486–492, 2019.

A Container-Based of Edge Device Monitoring on Kubernetes

Halim Fathoni[1,2], Hao-Yi Yen[3], Chao-Tung Yang*[3], Chin-Yin Huang[1], and Endah Kristiani[1,4]

[1] Department Industrial Engineering and Enterprise Information, Tunghai University, Taichung City 407302, Taiwan (R.O.C.)
[2] Departement of Ekonomi dan Bisnis, Politeknik Negeri Lampung, Indonesia
[3] Department Computer Science, Tunghai University, Taichung City 407302, Taiwan (R.O.C.)
Corresponding author: ctyang@thu.edu.tw
[4] Departement of Informatics, Krida Wacana Christian University, Jakarta 11470, Indonesia

Abstract. The concept of the Internet of Things has widely used in various applications in life, from home to big cities that have a variety of edge devices to facilitate edge computing. This paper attempts to use Kubernetes to establish an environment for collecting the internal resource usage status of various node devices and implementing monitoring and visualization. The deep learning models are deployed on heterogeneous devices for performance evaluation and verification. In this experiment, Kubernetes used to maintain a Docker container that used to deploy all applications mentioned above. All the data then visualize through Grafana.

Keywords: Kubernetes, Docker, Monitoring, Prometheus, Grafana, Edge Computing

1 Introduction

The rise of the Internet of Things in recent years has made extensive use of various edge computing devices in various fields to assist in data collection and pre-processing. However, due to the differences in architecture caused by heterogeneous edge computing devices, the devices have particular difficulties in performance evaluation and monitoring. With the increase of various edge computing devices, we try to develop a performance monitoring environment that is instantly monitored and easy to maintain. This environment will use containerized tool management methods by collecting and visualizing various resources and system indicators in the device [1]. In terms of management and deployment environment, Docker's containers are easy to deploy and have features that suitable for computing devices with lower efficiency. Kubernetes is used to manage and deploy the efficiency and life cycle of Docker containers, maintaining the

entire environment in a highly available state [3]. In terms of device data collection, this paper evaluates the deployment of Prometheus as a tool for data collection and storage of heterogeneous devices [4][6]. Prometheus collects data by Node exporter as an agent and installed in each edge device [5]. Then use Grafana to obtain the data source on Prometheus to visualize the collected data [11].

2 Related Works

2.1 Kubernetes

Kubernetes has a valuable feature that effectively used to manages microservices [7]. Scalable Linux containers automatically keep maintaining the container always in the best performance. Compared to manually deploying containers to multiple machines, Kubernetes can use the *Master* as a containerized service distribution and management for the *Slaves* in the entire cluster by establishing a cluster relationship with multiple nodes[8].

2.2 Docker

Docker is open-source software, an open platform for developing, deploying, and executing applications. Docker allows users to separate applications in the system environment to form smaller containers, thereby increasing the speed of software deployment. Docker containers are similar to virtual machines, but in principle, containers virtualize operating systems, and virtual machines virtualize hardware, so containers are more portable than virtual machines and consume fewer system resources [2].

2.3 Prometheus

Prometheus is open source software for environmental monitoring and alerting. It records real-time indicators in a time-series database (TSDB) built using the HTTP pull model and has flexible query and instant alert functions.

2.4 Node exporter

Node exporter is the data source of Prometheus. It captures different index data in real-time and constructs an HTTP pull model to provide Prometheus for collection and storage.

2.5 Grafana

Grafana is open source software for multi-platform analysis and interactive visualization. When it connects to a supported data source, it will provide charts, graphs, and alerts for the Web. It can be extended through the plug-in system. Users can use the interactive query builder to create sophisticated monitoring dashboards.

2.6 DeepStream

DeepStream is a tool based on NVIDIA, which is mainly used in the solution of the entire visual process. It is different from other visual libraries (such as OpenCV) in that it establishes a complete end-to-end support solution. In other words, whether the data source is a camera, movie, or video on a cloud server, from video codec, it can handle all the details of the background image judgment to the complete process of the displayed screen, including the setting of parameters.

3 System Design and Implementation

This paper mainly uses a server as the host for management and distribution services. Its specifications are Intel (R) Core (TM) i7-5960X CPU @ 4.00GHz CPU, 128G memory, and 2T HDD hard drive. Four edge computing devices are used as working nodes for performance monitoring, namely: Raspberry Pi 3, Raspberry Pi 4 (4GB), NVIDIA Jetson Nano, NVIDIA Jetson TX2. Node exporter will collect internal system indicators of each device, and then monitor the node exporter of each device through Prometheus to return the indicator data and store it in TSDB. Then access the Prometheus database through Grafana to get the data Fig. 1 describes the system architecture.

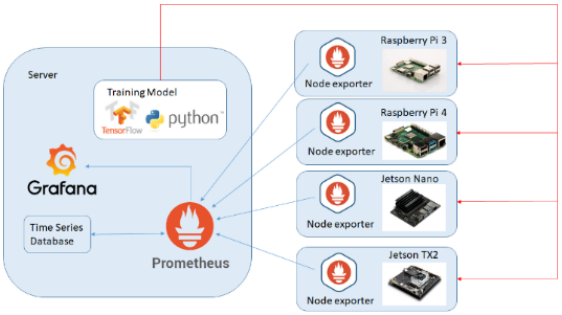


Fig. 1: System Architecture Diagram

3.1 Environmental Construction and Service Deployment

Prometheus, Node exporter, and Grafana are all deployed by Docker's container deployment service. Kubernetes used to establish a cluster system. The server host is used as the master node. The remaining edge computing devices are used as worker nodes for service allocation and scheduling—the smallest unit for deploying services in a Pod. Figure 2 shows the architecture of Kubernetes cluster.

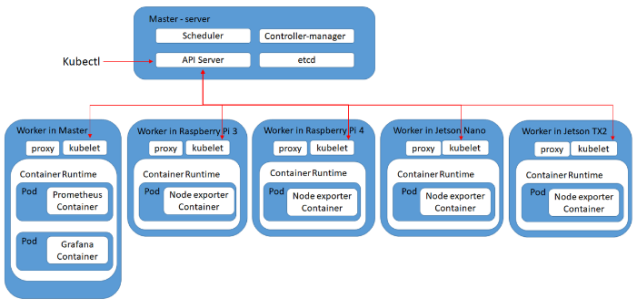


Fig. 2: Kubernetes Cluster Architecture

3.2 Collecting Resource Indicators

Each node device uses the Node exporter to collect internal resource indicators. Different devices must use a separate node exporter due to different architectures[9]. After the installation is completed, each resource index retrieved according to internal settings will be exported through port (9100). The derived resource indicators are: CPU usage, memory usage, system load, etc. After collecting and storing data. After establishing the entire monitoring environment, in order to verify whether the purpose of monitoring resource usage can be achieved, this experiments runs deep learning models on two heterogeneous edge device nodes, Raspberry Pi 4 and NVIDIA Jetson Nano.This experiment uses Grafana to observes and executes the changes in resource usage, before and after when running deep learning models, the results of this experiment are presented.

4 Experimental Result

In this experiment, the deployment of environmental tools has significantly improved efficiency through Docker. Service-based containerization eliminates the need for complicated steps like manual installation, and the virtualization of the container operating system makes the architecture between heterogeneous devices. In addition, through Kubernetes excellent management of containers and clusters, service errors can be quickly redeployed, allowing the entire system to monitor the status of all devices without interruption, and effectively expand and update service resources.

4.1 Data Collection and Visualization

In terms of data index collection, Prometheus has added four heterogeneous devices for monitoring, including NVIDIA Jetson Nano, NVIDIA Jetson TX2, Raspberry Pi 3, and Raspberry Pi 4 as resource monitoring devices, as shown in Fig. 3.

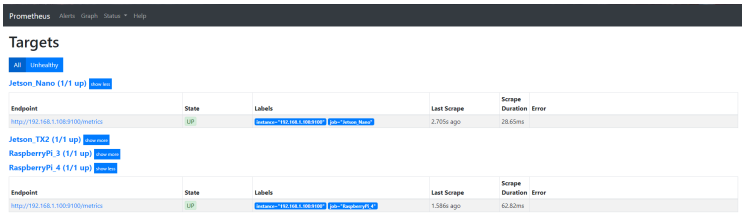


Fig. 3: Heterogeneous Devices Monitored on Prometheus

Among them, Raspberry Pi 4 and NVIDIA Jetson Nano are the monitoring targets of this resource change. After setting up the data source through Grafana, draw out various resource charts, such as: CPU usage, memory usage, system average load, etc. Experimental verification of resources in terms of changes in Raspberry Pi 4 performed by the Tensorflow trained Inception v3 model, which is the identification of the model vehicle, respectively, performed FP16. Figure 4a and Fig. 4b show the execution FP16 longitudinal CPU resources and memory of a change.

(a) CPU

(b) Memory

Fig. 4: running on Inception v3 FP 16 model

The experimental results show that the internal resource usage information of heterogeneous devices collected through Prometheus can be presented in real time through Grafana.

5 Conclusion

In this paper, Kubernetes used to build a high-quality service monitoring environment of edge heterogeneous equipment performance through Docker container, which has a features rapid deployment container services. In this study, the proposed focus on a high quality of service monitoring deployment plan. In the future, this research hopes to create a set of self-monitoring cluster services with the ability to train deep learning, with the goal of including this deep learning service.

Acknowledgement

This work was sponsored by the Ministry of Science and Technology (MOST), Taiwan, under Grant No. 108-2622-E-029-007-CC3, 108-2221-E-029-010-, and 108-2745-8-029-007-.

References

1. Sukhija, Nitin, and Elizabeth Bautista. "Towards a Framework for Monitoring and Analyzing High Performance Computing Environments Using Kubernetes and Prometheus." 2019 IEEE SmartWorld, 2019.doi:10.1109/SmartWorld-UIC-ATC-SCALCOM-IOP-SCI.2019.00087
2. Lingayat, Ashish, Ranjana R. Badre, and Anil Kumar Gupta. "Performance evaluation for deploying docker containers on baremetal and virtual machine." 2018 3rd International Conference on Communication and Electronics Systems (ICCES). IEEE, 2018. DOI:10.1109/CESYS.2018.8723998
3. Marathe, Nikhil, Ankita Gandhi, and Jaimeel M. Shah. "Docker Swarm and Kubernetes in Cloud Computing Environment." 2019 3rd International Conference on Trends in Electronics and Informatics (ICOEI). IEEE, 2019
4. Chen, Wenyan, Kejiang Ye, and Cheng-Zhong Xu. "Co-Locating Online Workload and Offline Workload in the Cloud: An Interference Analysis." 2019 IEEE 21st International Conference on High Performance Computing and Communications; IEEE 17th International Conference on Smart City; IEEE 5th International Conference on Data Science and Systems (HPCC/SmartCity/DSS). IEEE, 2019. DOI:10.1109/HPCC/SmartCity/DSS.2019.00316
5. Slhoub, Khaled, Marco Carvalho, and Fitzroy Nembhard. "Evaluation and Comparison of Agent-Oriented Methodologies: A Software Engineering Viewpoint.", 2019 IEEE International Systems Conference (SysCon). IEEE, 2019.DOI:10.1109/SYSCON.2019.8836962
6. Slhoub, Khaled, Marco Carvalho, and Fitzroy Nembhard. "Evaluation and Comparison of Agent-Oriented Methodologies: A Software Engineering Viewpoint." 2019 IEEE International Systems Conference (SysCon). IEEE, 2019.DOI:10.1109/SYSCON.2019.8836962
7. SMedel, Víctor, et al. "Characterising resource management performance in Kubernetes."Computers & Electrical Engineering 68 (2018): 286-297.DOI:10.1016/j.compeleceng.2018.03.041
8. Jiang, Congfeng, et al. "Energy aware edge computing: A survey."Computer Communications (2020) DOI:10.1109/SEC.2018.00038
9. Ning, Huansheng, et al. "Heterogeneous edge computing open platforms and tools for internet of things."Future Generation Computer Systems (2020)DOI:10.1016/j.future.2019.12.036
10. Dey, Swarnava, and Arijit Mukherjee. "Implementing deep learning and inferencing on fog and edge computing systems."2018 IEEE International Conference on Pervasive Computing and Communications Workshops (PerCom Workshops). IEEE, 2018.DOI:10.1109/PERCOMW.2018.8480168
11. Kristiani, Endah, Chao-Tung Yang, and Chin-Yin Huang. "iSEC: An Optimized Deep Learning Model for Image Classification on Edge Computing." IEEE Access 8 (2020): 27267-27276.DOI:10.1109/ACCESS.2020.2971566

Flame Recognition System Using YoLo

Chao-Tung Yang[1]*, Wen-Yen Lin[2], Yi-Chun Chen[1], Zheng-Yao Wang[1], and Chia-Hsin Lee[3]

[1] Department of Computer Science,
Tunghai University, Taichung City 407302, Taiwan (R.O.C.)
[2] Department of Information Management,
National Taichung University of Science and Technology, Taichung City 40401, Taiwan (R.O.C.)
[3] IOEZ INC., New Taipei City, Taiwan (R.O.C.)

Abstract. This research is to design a management system that uses drone-mounted cameras combined with object identification and live broadcast systems to assist disaster relief. In order to avoid violations of flight regulations and weather factors, the system will add restricted area restrictions and weather information. YoLo was used as the model of recognition system. When a disaster occurs, each rescuer can use multiple drones to observe the scene of the disaster and immediately grasp the status of the scene through the live broadcast system and object identification to achieve the shortest time to rescue and improve the efficiency of rescue.

Keywords: Object Detection, Image Detection, Video Streaming, UAV, Deep Learning

1 Introduction

The history of the drones development have ranged from the military use that the U.S. military used to fight thousands of kilometers to the enemy to the civilians today [1]. It has far exceeded the larger and wider range of applications. Because the cost of various electronic parts has been greatly reduced, the performance is far from Beyond the past, the threshold for consumer-grade drones is increasingly lowered [2][3]. In addition to the common handheld remote-controlled aircraft of the general public, it is also commonplace to take pictures of unmanned aerial cameras in major gathering occasions and TV programs. In view of the ability of unmanned aerial cameras [4][5]. Due to the limitation of terrain [6], many government units have used unmanned aerial cameras to cooperate with firefighting units to carry out various rescue operations [7][8].

This paper combines unmanned aerial cameras and object detection to assist disaster relief through artificial intelligence object recognition technology. Microsoft Azure platform is used as a development system platform. OpenCV and YoLo are implemented to provide image processing and model recognition.

* Corresponding author: ctyang@thu.edu.tw

The system will perform image recognition on the real-time transmitted images, detection and classified objects.

2 Background Review and Related Work

In this section, we discuss the background review of this work, including RTMP, HLS, OpenCV, Pytorch, YOLO, and other related works.

2.1 RTMP

RTMP (Real Time Messaging Protocol) originally developed by Macromedia to transfer streaming audio, video and data between Flash players and a server via the Internet. Macromedia was later acquired by Adobe Systems. This agreement was established on the basis of the TCP protocol or Polling on top of the HTTP protocol.

2.2 HLS Live Broadcast Agreement

HTTP Live Streaming (Abbreviated as HLS) is an HTTP-based streaming media network transmission protocol proposed by Apple. It works by first cutting the entire media stream into small HTTP files, and downloading only a small part each time. When the media stream is playing, the client wants to download the same resource from many different alternative sources, and download at different rates, allowing the streaming media session to adapt to different data rates. HLS only requests basic HTTP messages. Unlike RTMP, HLS can pass through any firewall or proxy server that allows HTTP data to pass through. It is also very easy to use the content distribution network to transmit media streams as shown in Figure **??**.

2.3 Related Work

Rivera et al. [9] proposed a modular system for human detection and geolocation. The study implemented a structure that can find possible human survivors in areas affected by disasters, including two shots And a geolocation module installed on the drone, according to the results mentioned in the article, the thermal detection effect is the best when used in night operation. On the other hand, the data implies that the optical detection can also operate effectively during the day, and with Nighttime deployment has higher accuracy than daytime deployment.

Gaszczak at al. [10] discuss about the method for real-time detection of people and vehicles for unmanned aerial vehicles (UAV) for aerial reconnaissance and surveillance. This paper proposes a method to automatically detect vehicles based on the use of multiple trained Haar classifiers and secondary confirmation in thermal imaging. In addition, a related method is proposed for human detection in thermal images combined with additional multivariate Gaussian shape

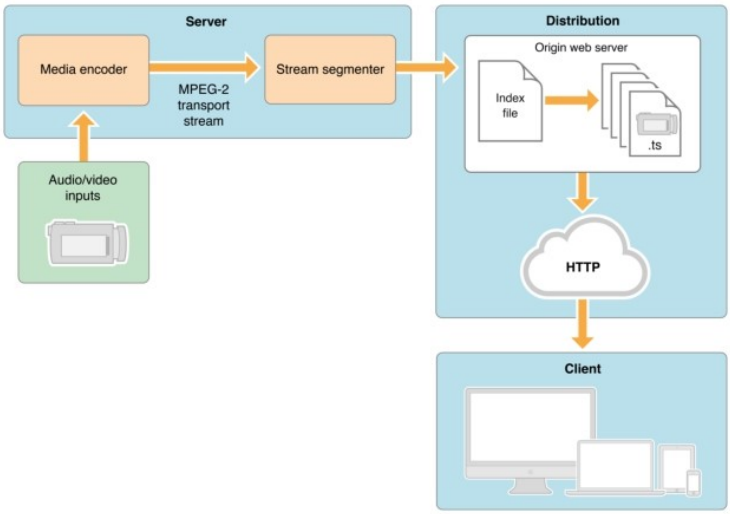


Fig. 1. Architecture diagram of UAV object identification system

matching. The results given indicate that vehicles and people are successfully detected in chaotic rural and urban environments under different conditions.

Giitsidis et al. [11] proposed the use of unmanned aerial vehicles equipped with optical and thermal sensors, as well as complex image processing and computer vision algorithms to detect suspicious activities or prevent the spread of fire. This paper is mainly special, the reason is that the human detection process is suitable for high-altitude and vertical shooting images, in contrast to most other similar works that consider lower altitudes and different shooting angles.

2.4 System Architecture

The architecture diagram of the UAV object identification system applied in the disaster relief site is shown in Figure 2. In order to be able to watch the drone image in real time, the environment part will build an RTMP streaming server, and the identification part will diversify the analyzed items. Various dangerous items are used as analysis items to facilitate future rescuers to use this system, which can improve the safety of rescuers and facilitate the rapid planning and discussion of rescue lines.

By setting Nginx as a web server, although Nginx is a lightweight server, the RTMP module written by others can also host RTMP streams, and then configure RTMP and HTTP HLS live broadcast protocols under Nginx Configuration.

2.5 Development environment

The environment part uses the Ubuntu 18.04 operating system as the development environment, uses Python3.6, Tensorflow1.13.1 and Keras2.2.4.

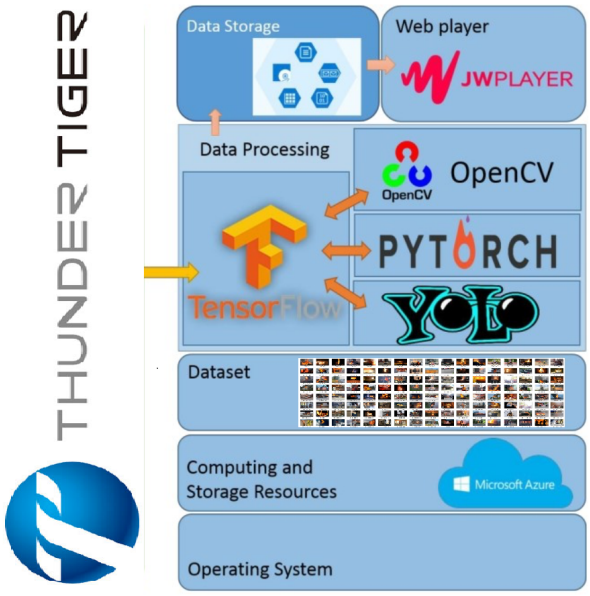


Fig. 2. Software Architecture

3 Experimental Results

In this section, we present HLS Stream, YOLO Model, and video test.

3.1 HLS Stream

After passing Nginx With RTMP Module, we can convert the RTSP signal output by Python into HLS and display it on the webpage. Because HLS can pass through any firewall or proxy server that allows HTTP data to pass, we don't need to install a special plug-in and You can watch the live broadcast.

3.2 YOLO Model

YoLo is a real-time object detection which classify or localizers the images. High scoring regions of the image are considered detections. YoLo model has several advantages compared to other classifier model. The prediction system is based on looking at the whole image to predict picture by global meaning. This also allows predictions unlike systems like R-CNN that allow thousands for a single picture for a single network evaluation. That makes it incredibly fast, more than 1000x quicker than R-CNN and 100x quicker than Short R-CNN [12]. The dataset used is from 2500 flame and smoke images, with 1250 flame and 1250 smoke images. The batch size is set to 64. Figure 3 shows the classification loss of the training model. Figure 4 describes the overall loss.

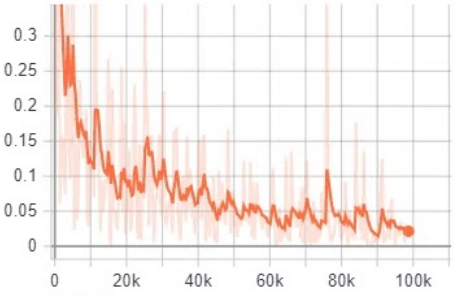


Fig. 3. Classification loss

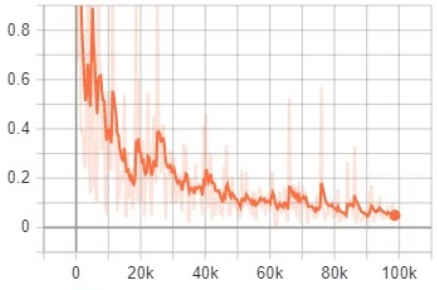


Fig. 4. Overall loss

Figure 5 shows the crowd density of heat map presentation. Figure 6 describes the application testing on flame recognition.

Fig. 5. Crowd Density Heat Map Presentation

Fig. 6. Test1

4 Conclusions and Future Work

This paper designs a UAV object recognition system that is applied to disaster relief sites. It can watch drone images in real time, helping disaster relief personnel to track the spread of fire and fire sources, and assisting in disaster relief.

In the future, we will try to migrate the recognition module to the Nvidia Jetson Xavier NX development board. Set up the development board on the aerial camera to achieve the function of real-time recognition. The recognized images will be returned to the server for future convenience. When the rescuers use this system, they can improve the safety of the rescuers and facilitate the rescuers to quickly plan and discuss the rescue line.

Acknowledgement

This work was sponsored by the Ministry of Science and Technology (MOST), Taiwan, under Grant No. 108-2622-E-029-007-CC3, 108-2221-E-029-010-, and 108-2745-8-029-007-.

References

1. Xiufang Shi, Chaoqun Yang, Weige Xie, Chao Liang, Zhiguo Shi, and Jiming Chen. Anti-drone system with multiple surveillance technologies: Architecture, implementation, and challenges. *IEEE Communications Magazine*, 56(4):68–74, 2018.
2. Dong Hyun Kim, Jae Ho Shin, and Jong Deok Kim. Design and implementation of wi-fi based drone to save people in maritime. *Journal of the Korea Institute of Information and Communication Engineering*, 21(1):53–60, 2017.
3. Juan A Besada, Luca Bergesio, Iván Campaña, Diego Vaquero-Melchor, Jaime López-Araquistain, Ana M Bernardos, and Jose R Casar. Drone mission definition and implementation for automated infrastructure inspection using airborne sensors. *Sensors*, 18(4):1170, 2018.
4. Amarjot Singh, Devendra Patil, and SN Omkar. Eye in the sky: Real-time drone surveillance system (dss) for violent individuals identification using scatternet hybrid deep learning network. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops*, pages 1629–1637, 2018.
5. Florian Schmidt and Stefan Hinz. A scheme for the detection and tracking of people tuned for aerial image sequences. In *ISPRS conference on Photogrammetric image analysis*, pages 257–270. Springer, 2011.
6. Li Deng, Yuanhua He, and Quanyi Liu. Research on application of fire uumanned aerial vehicles in emergency rescue. In *2019 9th International Conference on Fire Science and Fire Protection Engineering (ICFSFPE)*, pages 1–5. IEEE, 2019.
7. Bodnar Laszlo, Restas Agoston, and Qiang Xu. Conceptual approach of measuring the professional and economic effectiveness of drone applications supporting forest fire management. *Procedia engineering*, 211:8–17, 2018.
8. Paweł Smyczyński, Łukasz Starzec, and Grzegorz Granosik. Autonomous drone control system for object tracking: Flexible system design with implementation example. In *2017 22nd International Conference on Methods and Models in Automation and Robotics (MMAR)*, pages 734–738. IEEE, 2017.
9. AJA Rivera, ADC Villalobos, JCN Monje, JAG Mariñas, and CM Oppus. Post-disaster rescue facility: Human detection and geolocation using aerial drones. In *2016 IEEE Region 10 Conference (TENCON)*, pages 384–386. IEEE, 2016.
10. Anna Gaszczak, Toby P Breckon, and Jiwan Han. Real-time people and vehicle detection from uav imagery. In *Intelligent Robots and Computer Vision XXVIII: Algorithms and Techniques*, volume 7878, page 78780B. International Society for Optics and Photonics, 2011.
11. Themistoklis Giitsidis, Evangelos G Karakasis, Antonios Gasteratos, and G Ch Sirakoulis. Human and fire detection from high altitude uav images. In *2015 23rd Euromicro International Conference on Parallel, Distributed, and Network-Based Processing*, pages 309–315. IEEE, 2015.
12. Joseph Redmon and Ali Farhadi. Yolov3: An incremental improvement. *arXiv*, 2018.

Comparison of Influenza Disease Prediction Using ARIMA and LSTM models for Central Taiwan

Kieu Lan Phuong Nguyen[1,2][0000-0002-4487-0185], Ho-Wen Chen[*1][0000-0002-7677-1122], Chao-Tung Yang[3][0000-0002-9579-4426], and Endah Kristiani[4,5][0000-0003-2925-2992]

[1] Department of Environmental Science and Engineering,
Tunghai University, Taichung City 407302, Taiwan (R.O.C)
[2] Nguyen Tat Thanh University, Ho Chi Minh City 70000, Vietnam
[3] Department of Computer Science,
Tunghai University, Taichung City 407302, Taiwan (R.O.C)
[4] Department of Industrial Engineering and Enterprise Information,
Tunghai University, Taichung City 407302, Taiwan (R.O.C.)
[5] Department of Informatics, Faculty of Engineering and Computer Science.
Krida Wacana Christian University, Jakarta 11470, Indonesia
Corresponding Author: hwchen@thu.edu.tw

Abstract. The spreading infectious diseases could be a serious problem for the society, such as COVID-19 pandemic recently. The need to monitor the trend of the disease is crucial to prevent the disease transmission. In this paper, we propose prediction models for influenza trends using ARIMA and LSTM based on air quality dataset training operation. The study seeks to compare the prediction between ARIMA and LSTM in three areas of central Taiwan, i.e. Changhua County, Taichung City, and Nantou County between 2014 and 2018. The prediction of ARIMA relies only on five-year time series dataset of air pollutants and meteorological conditions, meanwhile there are three types of LSTM models related to feature selection technique. One of the LSTM model does not apply feature selection, and the two other LSTM models using matrix correlation and extra trees classifier to select the top five most important factors to conduct training operation. As a result, the experiment demonstrated that the LSTM models outperforms ARIMA in term of model performance in all three areas. Among three LSTM models, LSTM ETC achieved the lowest RMSE value indicating the highest accuracy. The RMSE value of influenza prediction in Taichung is substantially higher than that in Changhua and Nantou. However, LSTM ETC result obtained the highest accuracy improvement (73%) when comparing to ARIMA outcome. Therefore, implementing the LSTM models for prediction of influenza and another disease could be adopted for early detection as well as reduction of disease transmission.

Keywords: Influenza disease, ARIMA, LSTM, air pollution datasets.

* Corresponding author: hwchen@thu.edu.tw

1 Introduction

Ambient air pollution is the cause of about 4.2 million deaths annually, according to World Health Organization (2016). Of which, it is estimated that around 26% of respiratory infection deaths. There is synergistically potential interaction of air pollutants and influenza viruses on disease pathogenesis due to mutual effect on respiratory system (Wong et al., 2009). The distribution of air pollutants and influenza viruses is in space and time. For instance, meteorological conditions affect dispersion of air pollutants meanwhile fine particular matters may facilitate long-range transmission of influenza viruses. The finding prediction model using ARIMA and LSTM could be used for early detection of influenza as well as proposing emission reduction strategy of specific air pollutants which highly correlated with the disease (Yazdi et al., 2019).

Therefore, in this study, we aim to compare the finding prediction model of ARIMA and LSTM, which are models used widely for time-series data. Moreover, the variation of influenza cases in spatial dimension will be also surveyed in Changhua County, Taichung City, and Nantou County located in central Taiwan. The specific objectives of the study are presented as follows:

1. Designing the training model for ARIMA based on weekly data and for LSTM in combination with feature selection.
2. Comparing the accuracy of all models through an evaluation metric.

2 Background Review

2.1 Definition of the disease

Influenza is a contagious, acute respiratory illness caused by influenza viruses. The disease can cause mild to severe illness, particularly for those who are at higher risk for serious complications of influenza such as pregnant women, older adults, children under 5 years of age, and people with certain chronic health conditions (World Health Organization, 2015). In tropical and subtropical regions influenza circulation often during rainy season or year-round with some peaks during rainy season (Tamerius et al., 2013; World Health Organization, 2015).

2.2 Concepts of models

ARIMA. ARIMA is an acronym of AutoRegressive Integrated Moving Average which is denoted as ARIMA (p,d,q): p (autoregression)- a regression model based on the relationship between a sample and lagged sample; q (integrated)- the number of times that original observations of time series data are differenced to be stationary; q (moving average)- the model uses the relationship between each sample and each residual error from moving average model for lagged sample.

LSTM. LSTM is an acronym of Long-Short Term Memory network which is a special

kind of Recurrent Neural Network. The advantage of LSTM deep learning model is capability of keep information in persistent, thereby it might improve model performance in time-series prediction.

Evaluation metric. The study uses the square root of the mean squared error (RMSE) to evaluate the accuracy of models with its equation as follows:

$$RMSE = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(y_i - \hat{y}_i)^2} \quad (1)$$

where y_i is the actual value of observations; $\hat{y}_i$ is the predicted value of observations from the models; N is the total number of observations. The higher accuracy of model prediction, the closer value to 0.

2.3 Related Works

Both influenza viruses and ambient air quality have been documented well as one of the causes of health issues (Chen et al., 2017; Liang et al., 2014; Thach et al., 2010; Wong et al., 2009). For instance, Chen et al. (2017) claim that 10.7% of influenza cases across 47 cities in China result from PM2.5 exposure in ambient atmosphere. Wong et al. (2009) figure out significant effect modification of influenza for ozone effects. In order to predict influenza cases, the traditional statistical approach is applied in a study of Rao et al. (2020). Meanwhile, deep learning using LSTM model is also applied to predict influenza-like illness (Yang et al., 2019) and respiratory cases (Nguyen et al., 2020). There is, however, little study use feature selection technique in combination with LSTM to improve the model performance. Thus, in the current study, we employ some feature selection tools before training LSTM model, finally the results will be compared with ARIMA outcome.

3 System Design and Methodology

3.1 System Design

To proceed data analysis, system architecture is in need for specific purpose, see Figure 1. The study used dataset of air pollutants and meteorological conditions collected from Taiwan Environmental Protection Administration (EPA) (shown in Table 1) and influenza cases collected from Taiwan Centers for Disease Control (CDC) from 2014 to 2018 to train and predict using ARIMA and LSTM. Numpy and Pandas libraries were used for data manipulation and preprocessing, while Matplotlib library was used to visualize actual and predicted observations. For LSTM model, air pollutants and meteorological conditions dataset was trained using TensorFlow and Keras libraries. Scikit-learn library was adopted for evaluating the models through RMSE metric as well as measuring feature importance of parameter using extra trees classifier estimator.

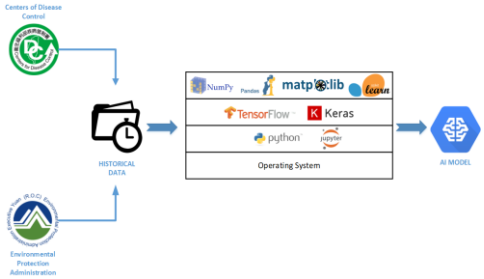


Fig. 1. System Architecture

Table 1. Air pollutants and meteorological dataset

No.	Parameter	Unit	Name
1	CO	ppm	Carbon monoxide
2	NO	ppb	Nitrogen monoxide
3	NO_2	ppb	Nitrogen dioxide
4	NO_x	ppb	Nitrogen oxide
5	O_3	ppb	Ozone
6	$PM_{2.5}$	$\mu g/m^3$	Fine aerosol
7	PM_{10}	$\mu g/m^3$	Aerosol
8	SO_2	ppb	Sulfur dioxide
9	AMB_TEMP	°C	Ambient temperature
10	RAINFALL	mm	Rainfall
11	RH	%	Relative humidity
12	WIND_SPEED	m/s	Wind speed
13	WIND_DIREC	degree	Wind direction

3.2 System Workflow

The system analysis flow is displayed in Fig. 2. After collecting the datasets, we executed data preprocessing which includes replacing special characters and missing data by 0, removing outliers. Because the original EPA's dataset in hourly form, the study converted it to weekly form to unify with CDC's dataset. After that, for ARIMA prediction model, the datasets were splitted into training and validation parts with the ratio of 80:20, respectively. In the meanwhile, there are three types of LSTM prediction model in combination with feature selection techniques. The first type-ALL- indicate no feature selection step was applied, thus all 13 parameters of EPA's dataset were utilized to train for the next step. The other two types include MC and ETC indicating operation of matrix correlation and extra trees classifier, respectively for feature selection before conducting training data. The ratio of evaluation partitions in LSTM model is as same as ARIMA model (80:20). Finally, we compare the accuracy of prediction models between ARIMA, and LSTM ALL, and LSTM MC, and LSTM ETC in three areas.

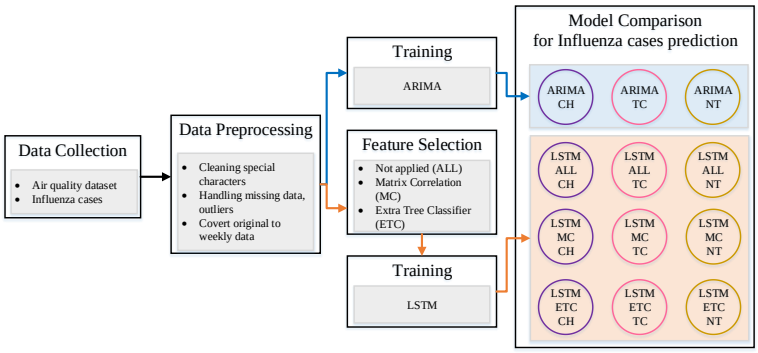


Fig. 2. System Analysis Flow

3.3 Experimental Procedure

The experimental procedure involves major contents: data preprocessing, feature selection, ARIMA model, and LSTM model.

Data Preprocessing. The stage consists of: replacing special characters and missing data by zero value; removing outliers; converting the air quality dataset from hourly to weekly by making average; and extracting influenza cases in Changhua, Taichung, and Nantou from the original CDC data.

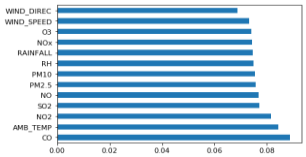
Feature selection.
All parameters (ALL). There is no application of any feature selection for all 13 air quality parameters. Thus, all parameters were adopted for training procedure.

Matrix Correlation (MC). The MC tool evaluates the correlation between the number of influenza and 13 air quality data. As a result, top five parameters have highest relationship will be chosen, shown in Table 2.

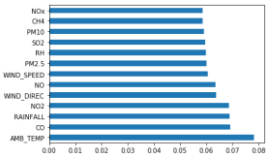
Table 2. Top 5 parameters selected by matrix correlation technique

Ranking	Changhua	Taichung	Nantou
1	AMB_TEMP	AMB_TEMP	NO
2	NO_x	NO_x	WIND_DIREC
3	NO_2	NO	WIND_SPEED
4	CO	NO_2	NO_x
5	NO	CO	SO_2

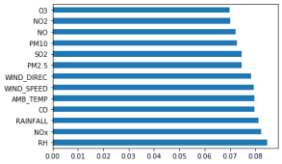
Extra Trees Classifier (ETC). The ETC estimator evaluates feature importance of parameters of the dataset using randomized decision trees.



(a) ETC for Changhua County's dataset



(b) ETC for Taichung City's dataset



(c) ETC for Nantou County's dataset

Fig. 3. Feature importance of air quality parameters in central Taiwan

Training parameters.

ARIMA Training Model. Before conducting ARIMA model, we determine if the time series data is stationary through ADF test. When the dataset is stationary, it will be portioned into training and validation parts with the ratio of 80:20, respectively.

LSTM Training Models. The three LSTM models including LSTM ALL, LSTM MC, and LSTM ETC were also splitted into training and validation portions with the same ratio as ARIMA model for easier comparison.

4 Experimental Results and Discussion

The finding results of predicted values by ARIMA and different types of LSTM models comparing with the actual values in Changhua, Taichung, and Nantou are presented in Table 3, 4, and 5, respectively.

Table 3. Actual and predicted last 10 values using different models in Changhua County

YW	Actual	Prediction			
		ARIMA	LSTM ALL	LSTM MC	LSTM ETC
201842	1550.0	1836.7	1570.9	1620.7	1606.9
201843	1603.0	1919.0	1500.7	1557.1	1550.7
201844	1658.0	2090.5	1556.2	1583.7	1573.9
201845	1531.0	2246.8	1668.8	1677.6	1672.4
201846	1618.0	2029.3	1597.0	1643.4	1629.5
201847	1674.0	2191.0	1620.6	1657.7	1654.6
201848	1906.0	2246.5	1688.8	1729.9	1713.4

YW	Actual	Prediction			
		ARIMA	LSTM ALL	LSTM MC	LSTM ETC
201849	2161.0	2324.8	1933.2	1943.8	1931.8
201850	2188.0	2482.9	2263.5	2294.4	2259.7
201851	2520.0	2480.2	2349.1	2384.3	2342.7

Table 4. Actual and predicted last 10 values using different models in Taichung City

YW	Actual	Prediction			
		ARIMA	LSTM ALL	LSTM MC	LSTM ETC
201842	2967.0	5353.0	3169.7	3014.2	3119.5
201843	3260.0	5355.6	2960.6	2795.1	2861.4
201844	3537.0	5694.0	3149.0	3044.9	3138.1
201845	3516.0	5699.3	3554.9	3456.1	3592.1
201846	3504.0	5618.6	3825.3	3695.3	3866.2
201847	3349.0	5683.7	3814.2	3661.5	3811.0
201848	3532.0	5992.1	3583.0	3433.7	3566.0
201849	4161.0	6178.0	3587.5	3446.7	3556.1
201850	4326.0	6342.4	4150.5	4052.2	4186.6
201851	5094.0	6559.2	4617.9	4516.0	4708.5

Table 5. Actual and predicted last 10 values using different models in Nantou County

YW	Actual	Prediction			
		ARIMA	LSTM ALL	LSTM MC	LSTM ETC
201842	240.0	579.5	294.5	244.2	255.9
201843	264.0	580.3	297.4	248.0	261.8
201844	288.0	596.9	315.5	266.3	282.7
201845	255.0	614.4	336.6	285.7	304.3
201846	201.0	660.6	324.4	271.1	286.8
201847	170.0	669.0	282.3	229.8	240.7
201848	209.0	705.8	248.7	198.9	208.6
201849	173.0	729.0	260.5	214.2	229.0
201850	230.0	773.5	237.8	191.4	209.0
201851	225.0	724.4	274.2	224.5	247.1

The result of RMSE in the four models is shown in Fig. 4. Because the original influenza cases in Taichung is much larger than the other counties, the RMSE value of influenza prediction model is substantially higher than that in Changhua and Nantou Counties. Moreover, it can be seen that the three LSTM models have the lower RMSE than that of ARIMA model. In other words, the prediction of influenza cases using LSTM models are more accuracy than the outcome of ARIMA model. Among three LSTM models, LSTM ETC achieved the lowest RMSE value indicating the highest accuracy. With respect to the accuracy improvement in the prediction results of LSTM

models when comparing to ARIMA outcome, LSTM ETC result obtained the highest value which improved about 73% of the accuracy.

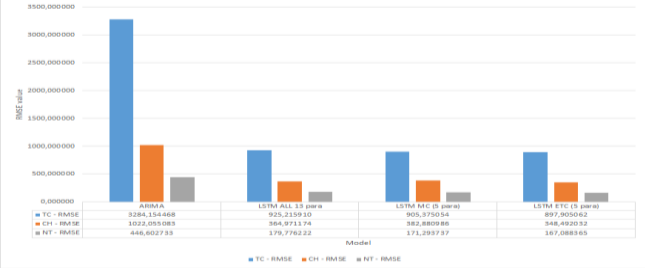
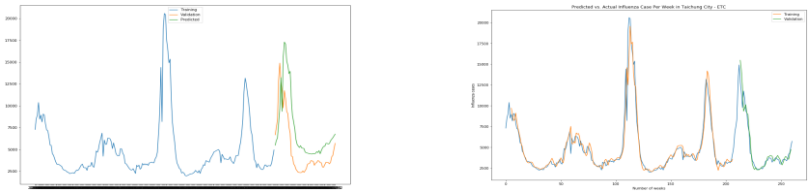


Fig. 4. RMSE comparison of influenza prediction model in central Taiwan

In Fig. 5, we present only the influenza time-series plot in Taichung in which has the highest accuracy enhancement, and the largest difference is the result of ARIMA and LSTM ETC models.



(a) ARIMA (b) LSTM ETC

Fig. 5. Comparative prediction results using different models in Taichung City between 2014 and 2018 (a) ARIMA; (b) LSTM ETC

5 Conclusions and Further Works

In this study, influenza cases in Changhua, Taichung, and Nantou were predicted using by ARIMA and LSTM models. While ARIMA is based on typical temporal structures which change over time, LSTM relies on multivariate air pollutants and meteorological conditions datasets from 2014 to 2018 combined with feature selection. Thereby, four prediction models were executed including ARIMA, LSTM ALL, LSTM MC, LSTM ETC. The experiment demonstrated that the LSTM models outperforms ARIMA in term of model performance in all three areas. When comparing three LSTM models, LSTM ETC achieved the lowest RMSE value indicating the highest accuracy. Although the RMSE value of Taichung models is substantially larger than the other counties, LSTM ETC result in Taichung obtained the highest accuracy improvement (about 73%) compared to ARIMA results.

Therefore, implementation of the LSTM models for prediction of influenza and another disease could be useful for real-time detection and disease transmission prevention.

6 Acknowledgement

The authors are grateful for the financial support from the Ministry of Science and Technology in Taiwan under Contract No. 108-2221-E-029-017-MY3.

References

1. Chen, G., Zhang, W., Li, S., Zhang, Y., Williams, G., Huxley, R., . . . Guo, Y. (2017). The impact of ambient fine particles on influenza transmission and the modification effects of temperature in China: A multi-city study. *Environment international, 98*, 82-88.
2. Liang, Y., Fang, L., Pan, H., Zhang, K., Kan, H., Brook, J. R., & Sun, Q. (2014). PM 2.5 in Beijing–temporal pattern and its association with influenza. *Environmental Health, 13*(1), 102.
3. Nguyen, K. L. P., Chen, H.-W., Yang, C.-T., & Kristiani, E. (2020). Implementation of a Respiratory Disease Forecasting Model Using LSTM for Central Taiwan. In *Information Science and Applications* (pp. 441-450): Springer.
4. Rao, X., Chen, Z., Dong, H., Zhu, C., & Yan, Y. (2020). Epidemiology of Influenza in hospitalized children with respiratory tract infection in Suzhou area from 2016 to 2019. *Journal of Medical Virology*.
5. Tamerius, J. D., Shaman, J., Alonso, W. J., Bloom-Feshbach, K., Uejio, C. K., Comrie, A., & Viboud, C. (2013). Environmental predictors of seasonal influenza epidemics across temperate and tropical climates. *PLoS pathogens, 9*(3).
6. Thach, T.-Q., Wong, C.-M., Chan, K.-P., Chau, Y.-K., Thomas, G. N., Ou, C.-Q., . . . Hedley, A. J. (2010). Air pollutants and health outcomes: assessment of confounding by influenza. *Atmospheric environment, 44*(11), 1437-1442.
7. Wong, C. M., Yang, L., Thach, T. Q., Chau, P. Y. K., Chan, K. P., Thomas, G. N., . . . Peiris, J. M. (2009). Modification by influenza on health effects of air pollution in Hong Kong. *Environmental health perspectives, 117*(2), 248-253.
8. World Health Organization. (2015). Influenza. https://www.who.int/immunization/diseases/influenza/en/
9. World Health Organization. (2016). Ambient air pollution. https://www.who.int/gho/phe/outdoor_air_pollution/en/
10. Yang, C.-T., Lin, L.-Y., Tsan, Y.-T., Liu, P.-Y., & Chan, W.-C. (2019). The Implementation of a Real-time Monitoring and Prediction System of PM2. 5 and Influenza-Like Illness Using Deep Learning. *Journal of Internet Technology, 20*(7), 2237-2245.
11. Yazdi, M. D., Wang, Y., Di, Q., Zanobetti, A., & Schwartz, J. (2019). Long-term exposure to PM2. 5 and ozone and hospital admissions of Medicare participants in the Southeast USA. *Environment international, 130*, 104879.

The Design and Implementation of Dynamic Costume Projection System

Chuan-Feng Chiu[1], Han-Yun Hsieh[1], Wei-Chuan Chung[1] and Shwu-Huey Yen[2]

[1] Minghsin University of Science and Technology, Department of Information Management, ROC
[2] Tamkang University, Department of Computer Science and Information Engineering, ROC
cfchiu@must.edu.tw, mr8407266@gmail.com, 105390@mail.tku.edu.tw

Abstract. With the development of machine vision theory and the advancement of hardware technology, the application of 3D models and 3D reconstruction, medical treatment, and life is becoming more and more widespread. Projection mapping is one of the special applications in this field. Therefore, in this paper we proposed a dynamic virtual costume system that use the Kinect V2 somatosensory device to acquire the human body grid structure based on the 3D depth image, preprocesses the depth 3D image to obtain a simplified grid, and use Iterative Closest Point(ICP) algorithm to calculate the projecting virtual costume. We align the position of the points between physical human body 3D mesh and virtual costume 3D mesh. After aligning the virtual costume 3D mesh to physical human body 3D mesh, the final virtual costume is suitable for physical human body via rotation and translation. The projector would project the generated virtual costume onto human body. And the proposed system also showed the ICP algorithm could be used with different object registration.

Keywords: Projection Mapping, Kinect, Iterative Closest Point

1 Introduction

With the advance of 3D technology and hardware, this led the popular development in 3D application area. Augment reality and mixed reality are such kind of applications. The one of popular usage in augment reality is the projection mapping application. Projection mapping has been used in many entertainment and advertisement activity including projection on Cinderella Castle in Disney Park by Disney[7] and the Fête des Lumières festival in Lyon by P. Warrener. The application of projection mapping would have potential for future usage and demonstration in other fields. Projection mapping is a technique which is mapping the video or image on the surface of selected object which resulting in special effect for human eyes. Most projection mapping demonstration focus on mapping the multimedia resource onto static objects including building etc.. So, we design and implement a system that project the costume onto the human body in a performance in this paper. We use Kinect to capture and track human body, calculate the projecting costume and project it onto. The rest of this paper is organized as following, Section 2 describe the related literatures regarding the projection mapping

research. Section 3 reveals the design of the dynamic costume projection system. Finally, the brief conclusion describes the contribution and future works.

2 Related Works

Projection Mapping had gained a lot of attention in recent year, it could project video, image etc. onto the surface of the objects by aligning the points between objects which could be used for advertisement or entertainment etc.. Rigid projection and non-rigid projection are two basic approach in this field. In rigid projection, the projection object is rigid, we could display rich texture onto the object surface by using projection mapping technology. Zhou *et. al.*[9] proposed a dynamic projection mapping system called Pmomo which could track the real object in 6DoF degree and generate the image that would project onto the tracking object. The proposed method comprised the usage of low-density point cloud to register the projection point and could against the occlusion issue. Siegl *et. al.*[3] proposed a novel method for projection mapping using multiple projectors. One projector could cover the unclear part of another projector and produce the high-quality projection image. The multiple projector could project onto objects by different luminance to generate interesting projection result. Non-rigid projection mapping focus on the objects which would have stretching or shrinking behavior. Because of the deformation of the surface of the object, the registration point or position would be more difficult to discovered. Bermano *et. al.*[17] proposed a method for projecting facial image onto real-time human faces. The real-time capturing the configuration if human faces is difficult. The majority of the proposed paper is to reduce the computation cost for projecting image onto human faces. The paper projected the captured human faces expression into low dimensional feature space to have non-rigid deformation and use adaptive Kalman filtering to have better result. Holman *et. al.*[10] proposed the projection mapping onto papers, the system could have more interesting interaction for users. The system captures the status of the paper changing and project material onto the surface changed papers. Lee *et. al.*[11] extend the idea to another foldable object including umbrella, reel, newspaper etc.. They add some markers to the foldable object and project images onto the surface to have more visual effect. Harrison *et. al.*[12] proposed a OmniTocuh system that track the surface of objects to project interactive 3D virtual user interface onto the object and track user's fingers to know the action pointing onto the virtual user interface. Besides the projecting mapping technology described above, with the popular usage in electronic commerce the virtual fitting room system is another application related with projection mapping. Traumann *et. al.*[13] proposed a virtual fitting room system for users to fitting the clothes in the virtual room. The system captures the human body and generate the virtual clothes image with texture and draw onto the captured human body image. The proposed system is based on the image processing technology and use Kinect to capture the physical human body. Therefore, in this paper we proposed a ICP-based dynamic virtual costume system which generate the virtual costume and project onto human body.

3 Dynamic Costume Projection Mapping System

In this section, we would describe the design of the proposed dynamic costume projection system showed in Fig. 1. The system comprises the Kinect RGB-D camera to capture the physical object and a projector which would project virtual costume onto human body. The virtual costume is estimated by the backend Dynamic Costume Projecting System and project it onto by the projector. In order to make projector to have correct projection position, we would estimate the offset of the object of the projecting image. Because of Kinect and projector are two hardware and placed in different position. First, we place Kinect and projector in the same line of horizontal axis. The distance D_p and D_k are the distance from Kinect and project to the horizontal plane of the position physical human body. The angle between Kinect, projector and physical human body is θ. The distance between Kinect and projector is K. Therefore, we could calculate the offset F or image that would be showed by projector with ε which is an offset with respect to the Kinect and physical human body in Equation(1).

$$F = D_p \times \tan\theta + \epsilon, where\ \tan\theta = \frac{K}{D_k} \tag{1}$$

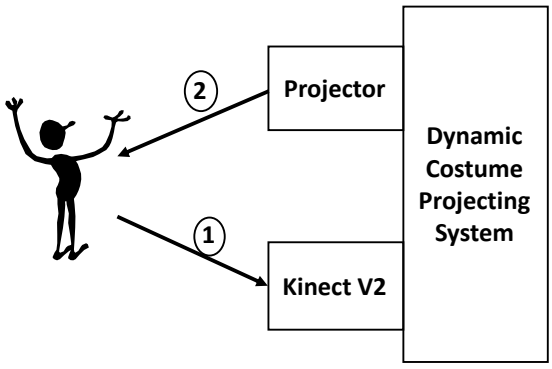


Fig. 1. Dynamic Costume Projecting System Architecture

In order to tracking and capturing the human body, we use Kinect to capture the moving human body to acquire 3D mesh information by using the Kinect SDK. On the other hand, the system had prepared the virtual costume which will be deformed as same as possible as the captured human body. In order to deform the virtual costume as same as the human body, we use Iterative Closest Point(ICP) algorithm to calculate the necessary rotation and translation to get approximative 3D mesh of virtual costume to physical human body. ICP is an algorithm proposed by P.J.Besl *et. al.*[15] to match points between different view of object. By finding the matching points, the overall 3D

model could be reconstructed from these different views of object. However, ICP focus on same object for the original usage. In this paper, we extend ICP for matching points between different objects, that is to project one object onto another object.
At first, we have human body 3D mesh H and virtual costume 3D mesh C. In order to reduce the calculation cost, we reduce the H and C by eliminate the redundant point, reduce the 3D triangle faces and have same number of points in reduced 3D mesh called H' and C'. Each iteration of ICP, we would have new C' called C_i' that is closed to H'. The process would be repeated until the distance between H' and C_i' of each iteration is lower a threshold that means we have rotation R and translation t which make the H' and C_i' closest. After getting C_i', we re-scale C_i' to have same volume between H' and C_i'. The overall flow is showed in Fig. 2. Fig. 3 shows our result that project clothes onto human body. The result also shows the clothes is close to the human body visually.

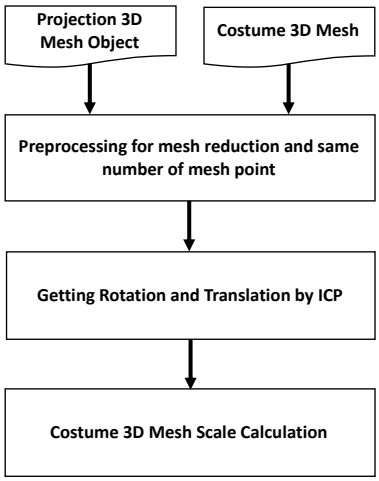


Fig. 2. Virtual Costume Calculation

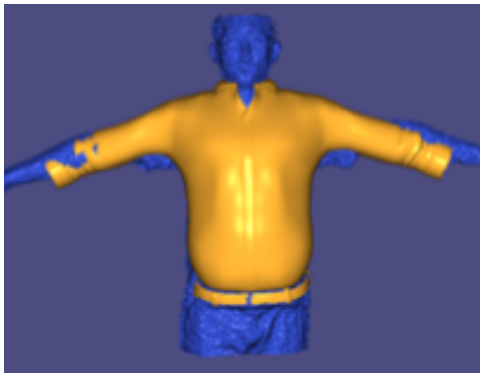


Fig. 3. Result of Our Design System

4 Conclusion

In this paper we proposed dynamic virtual costume system to project virtual costume onto human body which is an actor in a performance. The system makes the performance more flexible and exciting. The proposed mechanism is based on ICP algorithm. Traditionally, ICP algorithm is used for reconstructing 3D model which based on different view of the object. However, in this paper we extend the ICP to align one object with respect to another object and prove the possibility of the ICP algorithm in the projection mapping. But there still exists some issue that we could enhance in the future. The computation cost is still high and this make real-time processing not smooth. We would assist with GPU for enhancing the computation speed in the future.

Acknowledgment
The authors would like to thank the Ministry of Science and Technology (MOST) of R.O.C. financially supporting in part of the research under the grant number MOST 107-2221-E-032-035-MY2.

References

1. R. B. Rusu, N. Blodow and M. Beetz.: Fast Point Feature Histograms (FPFH) for 3D registration. In: 2009 IEEE International Conference on Robotics and Automation Proceedings, pp. 3212-3217. Kobe(2009).
2. C. Siegl, V. Lange, M. Stamminger, F. Bauer and J. Thies.: FaceForge: Markerless Non-Rigid Face Multi-Projection Mapping. IEEE Transactions on Visualization and Computer Graphics 23(11), 2440-2446(2017).
3. Christian Siegl, Matteo Colaianni, Lucas Thies, Justus Thies, Michael Zollhöfer, Shahram Izadi, Marc Stamminger, and Frank Bauer.: Real-time pixel luminance optimization for dynamic multi-projection mapping. ACM Transactions on Graphics (TOG), 34(6), 1-11(2015)
4. Bimber, O., & Raskar, R.: Spatial augmented reality: merging real and virtual worlds. AK Peters/CRC Press (2005).
5. R. Koch.: 3-D surface reconstruction from stereoscopic image sequences. In: IEEE International Conference on Computer Vision, Cambridge Proceedings, pp. 109-114. MA, USA(1995).
6. J. Weingarten and R. Siegwart.: EKF-based 3D SLAM for structured environment reconstruction. In: 2005 IEEE/RSJ International Conference on Intelligent Robots and Systems Proceedings, pp. 3834-3839. Edmonton, Alta.(2005).
7. M. R. Mine, J. van Baar, A. Grundhofer, D. Rose and B. Yang.: Projection-Based Augmented Reality in Disney Theme Parks. Computer 45(7), 32-40(2012).
8. Lange V., Siegl C., Colaianni M., Kurth P., Stamminger M., Bauer F.: Interactive Painting and Lighting in Dynamic Multi-Projection Mapping. In: De Paolis L., Mongelli A. (eds) Augmented Reality, Virtual Reality, and Computer Graphics. AVR 2016. Lecture Notes in Computer Science, vol. 9769. Springer, Cham(2016)
9. Y. Zhou, S. Xiao, N. Tang, Z. Wei, and X. Chen.: Pmomo: projection mapping on movable 3D object. In: the CHI Conference on Human Factors in Computing Systems Proceedings, pp. 781–790. Santa Clara, Calif, USA(2016).

10. David Holman, Roel Vertegaal, Mark Altosaar, Nikolaus Troje, and Derek Johns.: Paper windows: interaction techniques for digital paper. In: the SIGCHI Conference on Human Factors in Computing Systems (CHI '05) Proceedings, pp. 591–599. New York, NY, USA(2005).
11. Johnny C. Lee, Scott E. Hudson, and Edward Tse.: Foldable interactive displays. In: the 21st annual ACM symposium on User interface software and technology (UIST '08) Proceedings, pp. 287–290. New York, NY, USA(2008).
12. Chris Harrison, Hrvoje Benko, and Andrew D Wilson.: OmniTouch: wearable multitouch interaction everywhere. In: the 24th annual ACM symposium on User interface software and technology Proceedings, pp. 441--450. Santa Barbara, California, USA(2011).
13. A. Traumann, G. Anbarjafari and S. Escalera.: A new retexturing method for virtual fitting room using Kinect 2 camera. In: IEEE Conference on Computer Vision and Pattern Recognition Workshops (CVPRW) Proceedings, pp. 75-79. Boston, MA(2015).
14. M. Pauly, M. Gross and L. P. Kobbelt.: Efficient simplification of point-sampled surfaces. In: IEEE Visualization Proceedings, pp. 163-170. Boston, MA, USA(2002).
15. P. J. Besl and H. D. Mckay.: A method for registration of 3-D shapes. IEEE Transactions on Pattern Analysis and Machine Intelligence14(2), 239-256(1992).
16. Li Biqing and Cai Yong.: An Efficient Variant of the ICP Algorithm for Multiview Registration. Mechanical Engineer 2, 73-75(2009).
17. Amit H. Bermano, Markus Billeter, Daisuke Iwai, and Anselm Grundhöfer.: Makeup Lamps: Live Augmentation of Human Faces via Projection. Comput. Graph. Forum 36(2) , 311–323(2017).
18. G. Narita, Y. Watanabe and M. Ishikawa.: Dynamic Projection Mapping onto Deforming Non-Rigid Surface Using Deformable Dot Cluster Marker. IEEE Transactions on Visualization and Computer Graphics23(3), 1235-1248(2017).
19. Hashimoto, Naoki & Koizumi, Ryo & Kobayashi, Daisuke.: Dynamic Projection Mapping with a Single IR Camera. International Journal of Computer Games Technology, 1-10(2017).

Unidentified PII detection with a k-Nearest Neighbors Approach

Tzu-Yin Liao[1], Yu-Chih Wei [2] and Wei-Chen Wu[3]

[1] National Taipei University of Technology, Taiwan
t108ab8019@ntut.org.tw
[2] National Taipei University of Technology, Taiwan
vickrey@mail.ntut.edu.tw
[3]Feng Chia University, Taiwan
weicwu@fcu.edu.tw

Abstract. Since the implementation of the EU General Data Protection Regulation ("GDPR") and similar legislation on personal data protection in Taiwan, enterprises must now provide adequate protection for their customers' personal data. Many enterprises use automated personally identifiable information ("PII") scanning systems to process PII to ensure full compliance with the law. However, personal data saved in non-electronic form cannot be detected by these automated scanning systems, resulting in PII not being able to be accurately identified. We propose a k-nearest neighbors ("k-NN") approach to detect unidentified PII to close the loopholes. Relevant peripheral information attributes of PII were identified and used in our study for machine learning and modeling to establish a model for detecting PII that otherwise could not be detected by automated scanners. Our study showed the F1-Measure of our proposed model achieved at least 80%, higher accuracy rate than that of automated scanners in detecting PII in an enterprise's inventory of information assets.

Keywords: Personally Identifiable Information, k-NN, Privacy

1 Introduction

In order for an enterprise to meet its customer expectations of service, its customers are often required to provide certain information for or feedback on the services provided. The information provided by the customers may include their personal and confidential information. In light of the EU GDPR and the legislation on personal data protection in Taiwan, enterprises will be at risk of law violation if they do not have adequate information management infrastructure in place.

For the purposes of information management, an inventory of information assets, including PII, must first be conducted. Information security and risk assessments are then carried out to detect anomaly against items in an inventory. Due to the sheer volume of information assets in an inventory, risk assessments carried out manually often incur huge amounts of professional fees, and human errors, such as omissions and mistakes in manual detection of PII, are ultimately inevitable.

A k-NN approach of PII detection is discussed in this paper to examine information assets (including PII) contained in an information asset inventory. In our study, we focused on the detection of non-electronic personal data that automated personal asset scanning systems cannot detect or accurately identify. We sought to set up an effective detection mechanism with machine learning to improve the accuracy in detecting unidentified PII, so that the cost of consultants and auditors could be reduced, and the quality of risk assessment data improved. Attributes of information assets that were related to certain business processes of an enterprise or contained key peripheral information were identified and analysed for our k-NN approach.

The following sections of this paper are arranged as follows: Section 2 describes relevant researches on risk assessments and machine learning. Section 3 presents our research method. Section 4 describes the results of this study. Conclusions and recommendations are given in Section 5.

2 Related Work

Very few academic works have been carried out on this topic. Some studies on risk assessments, risk predictions and anomaly detection by machine learning are mentioned here for informational purposes. Eminagaoglu et al. [1] studied the information security risk of a human resources department of a logistics company. 11 machine learning classifiers were used, and 342 samples of HR data were collected for their study. However, the number of samples used in their study was too low to be sufficient for effective training and testing for machine learning. Zhao et al. [2] used wavelet neural networks ("WNNs") for information risk assessments. Wavelet proposed the WNNs to avoid shortcomings of BPNN. The WNNs were known for their strong learning ability and high accuracy. The WNNs' network architecture was also relatively simple, and they can achieve convergence of speed more quickly. The WNN models are more suitable for quantitative than qualitative analyses.

Paltrinieri et al. [3] used deep neural networks (DNNs) to predict risk associated with oil wells. Kaplan et al[4] defined risk (R) as a combination of answers to these three questions: (1) what could go wrong (scenario s), (2) the likelihood of that happening (probability p) , and (3) the severity of the consequences (consequence c).

$$R = f(s, p, c) \tag{1}$$

In Paltrinieri's study, they examined each layer of the risk factors, e.g. how equipment at the oil wells was operated, what accidents had occurred and how experts could have prevented the accidents. These factors were then used as indicators for their risk assessment models. The output of their models represented the probability of accidents occurring.

Many unsupervised machine learning algorithms for detecting outliers, e.g. the k-NN, LOF, and the anti-k-NN algorithms, etc., had been developed based on the distance-based outlier detection technique developed by Knorr et al.[5]. Knorr defined outliers as objects that deviated substantially from a fixed distance threshold. The

distance threshold was further extended to k-NN distance, i.e. the distance of a point to its kth nearest neighbors. Later studies proposed dealing with outliers in a different way by assigning each object an anomaly score representing its outlierness, instead of basing on its k-NN distance. [6].

Malini et al. [7] analyzed credit card fraud identification techniques using the k-NN outlier detection. Machine learning, genetic programming, fuzzy logic, sequence alignment, etc. were used to detect credit card fraudulent transactions. Their results proved that the k-NN method was accurate and efficient.

Local anomaly detection algorithms, such as the LOF, were developed by measuring the local deviation of a given point to its k-NN. The local method was more flexible than the global method and better reflected the outlierness of each group of data. LOF compared the local reachability density (lrd) of an object to the local densities of its neighbors. Points that had a substantially lower density than their neighbors (with higher LOF scores) were considered to be outliers [6].

3 Methodology

This section analyzes anomaly processing and screening with automated systems for risk assessments and the setup of k-NN models for this study. In our study, Chinese text was first pre-processed, segmented into words or phrases and converted into vectors in a vector space model ("VSM"). The segmentation process involved Chinese text processing, the selection of features, the calculation of weighted features and the selection of classification algorithms. The output of the VSM model represented the category a segmented Chinese character or phrase belongs to.

The k-NN algorithm is one of the most popular unsupervised machine learning algorithms for text classification [8]. It is also known as a lazy algorithm. It is a local algorithm. It calculates the distance of an object to its neighboring points though Euclidean distance and sets parameter k to decide the number of nearest neighbors to be in the voting process of the algorithms. The classification of a target object is decided by the majority of votes from the voting neighbors. If k=1, the object is assigned to its nearest neighboring object to determine its classification.

In this study, a grid search method was used to find the optimal K value for the k-NN. We used F1-Measure as a metric for measuring the performance of our k-NN model. Generally, the larger the value of k is, the less sensitive to noise variations a k-NN model will be. However, a large k value is not necessarily suitable for specific tasks that a k-NN model is designated to tackle.[9] A combined method of the gird search and cross validation was therefore adopted in this study to find the optimal value of K for our model.

In our study, independent data was also used to verify and determine k value in cross-validation. The data was divided into a N number of subsets, one set of which was treated as the testing data and the remaining (N-1) sets of which training data. This process were repeated N times, and each subset was verified once.

3.1 Word segmentation

Jieba is a widely used PHP Chinese word segmentation module[10], as it allows users to add new words and build their own customized dictionaries. This function gives Jieba better accuracy in word segmentation than other traditional word segmentation modules. However, there are other word segmentation tools available these days that offer even better accuracy. The Academia Sinica in Taiwan has released an open source library, CkipTagger, which implements NLP tools, such as word segmentation (WS), part-of-speech tagging (POS), and named entity recognition (NER). It also offers features, such as (1) performance improvements, (2) no word limits on sentences, and (3) supporting user-defined recommended and must-have word lists. CkipTagger was tested with the ASBC 4.0 Test Split (50,000 sentences). The results for accuracy, recall and F1-score of CkipTagger were all about 97%, higher than those of Jieba. Therefore, CkipTagger was chosen as a tool for word segmentation in this study[11].

Table 1. An extract of CkipTagger, CKIPWS (classic), Jieba-zh_TW comparison [11]

Tool	(WS) prec	(WS) rec	(WS) f1	(POS) acc
CkipTagger	97.49%	97.17%	97.33%	94.59%
Jieba-zh_TW	90.51%	89.10%	89.80%	--

3.2 A comparison of word frequency and word similarity methods

The TF-IDF evolved from the IDF proposed by Sparck Jones (1972, 2004) [12]. The TF-IDF uses statistical methods to evaluate the importance of a word in a document in a file or a corpus based on the frequency of it appearing in the document. The more important the word is, the higher its TF-IDF score is. The TF is used to calculate the frequency of a word in a document, whereas the IDF the frequency of a word in a corpus. The TF-IDF is obtained by multiplying the two results.

The latent semantic analysis (LSA) is used to minimize global reconstruction errors by finding the best subspace approximation to the original document space. It is based on the singular value decomposition (SVD) and projects the document vectors into an approximated subspace to calculate the cosine value to represent similarity[13].

Multi-words are mainly segmented by two methods: (1) statistical methods, and (2) linguistic methods through grammar and sentence rules.

Zhang, Yoshida, and Tang[14] used statistical methods in their study to compare the performance of the TF-IDF, the LSA and Multi-words in text classification. They analyzed their semantic quality and statistical quality and discussed their application to the Chinese and English text retrieval and classification. Their results showed that the LSI and the multi-words method produced better semantic quality than the TF-IDF, and the TF-IDF produced better statistical quality than the other two methods[15]. A comparison table of the three methods is shown in Table 2. The smaller the value of a method in the table was, the better the method was. As scores would be assigned to words or phrases in accordance with their number of occurrences (i.e. the frequency)

in a document or a corpus in this study, the TF-IDF, which offered better statistical quality, was chosen to calculate word frequency for our study.

Table 2. TF-IDF, LSI, Multi-word comparison [15]

Task	TF-IDF	LSI	Multi-words
Chinese information retrieval	1	2	3
Chinese text classification	2	1	2

4 Evaluation

4.1 Dataset

Our study data was collected from a listed company in Taiwan. The dataset contained 4096 information assets each of which has three variables, including asset names, asset attributes, and PII. The asset names and assets attributes converted into vectors for the VSM. The CkipTagger was used in our study for word segmentation, the TF-IDF and one-hot encoding for the text vectorization , and the k-NN for word classification.

4.2 Implementation process

There were 4096 raw information assets in total, all of which were pre-processed. Asset names of the raw information assets were segmented into words or phrases. As our main task was to identify PII of the company's customers, uncertain data or data with missing values were removed from our study data. Data relating to employees of company or internal correspondence was also removed, as it had negative impact on our machining training and was not considered PII for the purposes of this study. Duplicate words, punctuation marks, stopped words, building floor numbers, numbers, and words with only one Chinese character were also removed. There were 1104 words or phrases segmented from the asset names remaining, each of which was then treated as a feature of a column in the inventory, totaling 1104 features/columns. There were 47 categories of asset attributes in the raw information data. All of the 47 categories of the asset attributes were treated as 1 (one) feature. Each category was given a categorical number, representing the numerical value of the entry in the dataset.

After the word segmentation and classification of the asset names and the asset attributes of the raw information data, there were a total of 1105 features for our model, containing both PII and non-PII information assets. Our model was then used to process the features for PII detection (see **Fig. 1.**).

After the data was pre-processed, it was then arranged into 1225 rows of information assets in our inventory. One-hot encoding and the TF-IDF were carried out to train our model for feature extraction. One-hot encoding was carried out on the features of $s_{i1},...,s_{i1104}$ which are the result in the word segmentation of asset names. The category

of asset attributes is s_{i1105}. After one-hot encoding, "1" was entered in the inventory indicating a feature existent in an asset name, whereas "0" non-existent. After the TF-IDF, if the feature existed in an asset name, it was given a TF-IDF score, and if not, "0".

When validating our model, we used the train_test_split in scikit-learn to split the testing and training data. There were 980 rows of information assets in the inventory as training data and 245 as testing data for validation. KNeighbors Classifier was used to classify all training data. Cross-validation was used to find the best k value. We used GridSearchCV to perform cross-validation. The F1-Measure, the average and standard deviation were obtained through 15-fold cross-validation, the results of which were compared to find the best k.(see Table 3.)

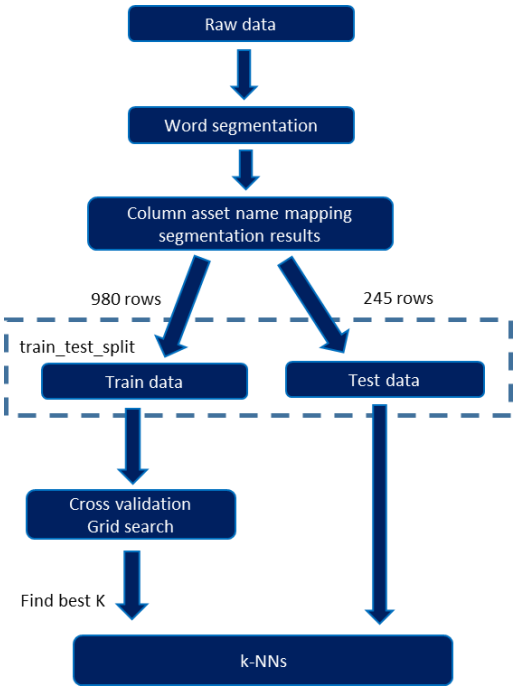


Fig. 1. The process of unidentified PII detection

Table 3. 15-fold cross-validation

k	CV1	CV2	CV3	CV4	CV5	CV6	CV7	CV8	CV9	CV10	Mean	StdDev
1	0.845	0.843	0.857	0.701	0.792	0.776	0.763	0.811	0.731	0.791	0.791	0.048
2	0.759	0.769	0.804	0.651	0.651	0.725	0.727	0.750	0.699	0.716	0.725	0.047
3	0.792	0.812	0.808	0.704	0.747	0.828	0.824	0.781	0.776	0.813	0.788	0.037
4	0.742	0.809	0.792	0.667	0.737	0.787	0.733	0.750	0.756	0.795	0.757	0.040
5	0.824	0.784	0.857	0.673	0.752	0.796	0.788	0.771	0.768	0.809	0.782	0.046
6	0.771	0.755	0.787	0.707	0.708	0.825	0.763	0.747	0.750	0.813	0.763	0.037

7	0.774	0.752	0.765	0.654	0.769	0.784	0.780	0.774	0.772	0.821	0.764	0.041
8	0.755	0.742	0.796	0.634	0.740	0.816	0.768	0.780	0.703	0.839	0.757	0.056
9	0.757	0.714	0.769	0.685	0.796	0.800	0.772	0.766	0.702	0.816	0.758	0.042
10	0.725	0.716	0.780	0.673	0.781	0.796	0.768	0.738	0.681	0.830	0.749	0.048
11	0.755	0.762	0.757	0.686	0.807	0.800	0.772	0.752	0.708	0.813	0.761	0.039
12	0.757	0.752	0.755	0.699	0.765	0.816	0.763	0.743	0.717	0.796	0.756	0.032
13	0.759	0.731	0.738	0.718	0.815	0.777	0.725	0.766	0.688	0.760	0.748	0.034
14	0.750	0.755	0.707	0.700	0.769	0.784	0.737	0.772	0.667	0.784	0.742	0.037
15	0.725	0.757	0.733	0.704	0.804	0.784	0.720	0.722	0.695	0.796	0.744	0.037

245 rows of data in the inventory were tested with TF-IDF and one-hot encoding in our model. The F1-Measure of test data was 82.93% and 83.42% with one-hot encoding and TF-IDF respectively.22 rows of data in the inventory were tested with TF-IDF and one-hot encoding in our model. The F1-Measure of evaluation was 81.82% and 63.33% with one-hot encoding and TF-IDF respectively. TF-IDF scored higher than one-hot encoding in test data, but there was not much difference between TF-IDF and one-hot encoding. However, in terms of the overall result of the model, one-hot encoding scored higher than TF-IDF. One-hot encoding was therefore chosen for our study due to its stability. The result of two methods are shown in Table 4. We also did an experiment for hierarchical k-NN model compared with one-hot encoding k-NN of which result is shown in next section.

Table 4. The result of F1-Measure in TF-IDF and one-hot encoding

F1-Measure	**TF-IDF**	**One-hot encoding**
Test data (245 rows data)	83.42%	82.93%
Evaluation (22 rows data)	63.33%	81.82%

4.3 The hierarchical k-NN model

A hierarchical k-NN model, like a single k-NN model, processes data though word segmentation and assigning categorical values to asset attributes. However, a hierarchical k-NN model consists of 3 k-NN models, namely (1) an asset name single k-NN model, (2) an asset attribute single k-NN model, and (3) a combined k-NN model of the two single k-NN models.

In our study, asset names were segmented into 1104 features after the word segmentation for the asset name single k-NN model. The target variables of asset name single k-NN model were same as the one-hot encoding k-NN model. Categories of the asset attributes were treated as features in the asset attribute single k-NN model. The target variables of asset attribute single k-NN model were same as the one-hot encoding k-NN model. Features of the combined k-NN model were features of the other two

models combined. The target variables of combined k-NN model were also the same as the one-hot encoding k-NN model .

Our results showed that the one-hot encoding k-NN model was better than the hierarchical k-NN model, as the F1-Measures of the single k-NN for both the cross-validation and evaluation were higher (see Table 5). Consequently, the single k-NN model was used in this study.

Table 5. F1-Measure of Hierarchical k-NN model

F1-Measure	Asset name	Asset attribute	Combined
Test data (245 rows data)	98.58%	78.53%	82.18%
Evaluation (22 rows data)	77.25%	73.21%	77.25%

5 Conclusion

In our study, we used k-NN models to detect unidentified PII. Three models, namely the TF-IDF with single k-NN, one-hot encoding with single k-NN and the hierarchical k-NN models, were tested in our study. Our test results showed that the F1-Measure of the TF-IDF single k-NN model was lower than the one-hot encoding single k-NN model, possibly due to the fact that the data used for our study was not populations, that the frequency of a word appearing in different samples was different and that it had no positive impact on the target variables.

The reason why the F1-Measure of the hierarchical k-NN model was lower than the one-hot encoding k-NN model was that the number of asset attribute features used for training were insufficient and that assets that shared the same asset attributes might include both PII and non-PII assets. The use of asset names for training, however, resulted in better detection of PII. As both asset names and asset attributes were used in the combined k-NN model of the hierarchical k-NN model for training, the F1-Measure of hierarchical k-NN model was lower than the one-hot encoding single k-NN model.

The F1-Measure of our single one-hot encoding k-NN model achieved more than 80% after it was tested by cross-validation and additional data. Traditionally, auditors would manually review each asset name and asset attribute in the inventory. The k-NN model proposed in our study utilized simple machine learning to expeditiously determine whether any row of data might contain PII, and it could therefore help reduce the time required for audits and professional cost.

Our k-NN model will be offered to auditors to be used in their auditing for feedback on the predictions given by our model. The model will be used to obtain more training data. Once we have more understanding of the relationship between the use of an asset name in the business process of an enterprise, our model can also be improved by

adding columns for business processes for training for more effective detection of PII in an information asset inventory.

6 References

1. Eminagaoglu, M. and S. Eren. *Implementation and comparison of machine learning classifiers for information security risk analysis of a human resources department*. in *2010 International Conference on Computer Information Systems and Industrial Management Applications (CISIM)*. 2010. IEEE.
2. Zhao, D.-M., J.-X. Liu, and Z.-H. Zhang. *Method of risk evaluation of information security based on neural networks*. in *2009 International Conference on Machine Learning and Cybernetics*. 2009. IEEE.
3. Paltrinieri, N., L. Comfort, and G.J.S.s. Reniers, *Learning about risk: Machine learning for risk assessment.* 2019. **118**: p. 475-486.
4. Kaplan, S. and B.J.J.R.a. Garrick, *On the quantitative definition of risk.* 1981. **1**(1): p. 11-27.
5. Knorr, E.M. and R.T. Ng. *A unified approach for mining outliers*. in *Proceedings of the 1997 conference of the Centre for Advanced Studies on Collaborative research*. 1997. IBM Press.
6. Campos, G.O., et al., *On the evaluation of unsupervised outlier detection: measures, datasets, and an empirical study.* 2016. **30**(4): p. 891-927.
7. Malini, N. and M. Pushpa. *Analysis on credit card fraud identification techniques based on KNN and outlier detection*. in *2017 Third International Conference on Advances in Electrical, Electronics, Information, Communication and Bio-Informatics (AEEICB)*. 2017. IEEE.
8. Manning, C.D., C.D. Manning, and H. Schütze, *Foundations of statistical natural language processing*. 1999: MIT press.
9. Ma, C.-M., W.-S. Yang, and B.-W.J.J.o.A.S. Cheng, *How the parameters of k-nearest neighbors algorithm impact on the best classification accuracy: In case of parkinson dataset.* 2014. **14**(2): p. 171-176.
10. Huiqin, W. and L. Weiguo. *Analysis of the Art of War of Sun Tzu by Text Mining Technology*. in *2018 IEEE/ACIS 17th International Conference on Computer and Information Science (ICIS)*. 2018. IEEE.
11. Li, P.-H. and W.-Y. Ma. *CkipTagger*. 2019; Available from: https://github.com/ckiplab/ckiptagger.
12. Jones, K.S.J.J.o.d., *A statistical interpretation of term specificity and its application in retrieval.* 1972.
13. Berry, M.W., S.T. Dumais, and G.W.J.S.r. O'Brien, *Using linear algebra for intelligent information retrieval.* 1995. **37**(4): p. 573-595.
14. Justeson, J.S. and S.M.J.N.l.e. Katz, *Technical terminology: some linguistic properties and an algorithm for identification in text.* 1995. **1**(1): p. 9-27.
15. Zhang, W., T. Yoshida, and X.J.E.S.w.A. Tang, *A comparative study of TF* IDF, LSI and multi-words for text classification.* 2011. **38**(3): p. 2758-2765.

Artificial Intelligence Identification Model for Chronic Kidney Disease

Ya-Fang Cheng[a], Hsiu-An Lee[b], Chien-Yeh Hsu[c,*]

[a] Department of Information Management, National Taipei University of Nursing and Health Sciences, Taipei, Taiwan

[b] Department of Computer Science and Information Engineering, Tamkang University, New Taipei, Taiwan

[c] Department of Information Management, National Taipei University of Nursing and Health Sciences, Taipei, Taiwan

[*] Master Program in Global Health and Development, Taipei Medical University, Taipei, Taiwan

Abstract:

Many people suffer from Chronic Kidney Disease (CKD). Nowadays, CKD is one of the top ten causes of death. CKD should via invasive examination to understand participants health status. If a non-invasive identification model can be established, it can provide users with self-assessment which let users quickly understand their physical condition. This study used machine learning method to establish an identification model of Chronic Kidney Disease.

This study found the associated factors with kidney failure from the literature. Selected MJ database as information resources. Used two different factor selection methods to training model. Compared the performance with K-Nearest Neighbor, Support Vector Machine, Logistic Regression, Artificial Neural Network, Decision Tree, Random Forest, eXtreme Gradient Boosting and Vote Algorithms, used the better one to establish the model.

In this study, the best model used Vote algorithm to establish the model, and can only use 13 non-invasive factors. The accuracy is 88%, the precision is 73%, the sensitivity is 69%, the specificity is 93%, and the AUC is 0.92. The contribution of this study is to use non-invasive factors to identify Chronic Kidney Disease, but it is a preliminary evaluation and ultimately requires doctors to diagnose.

Keywords: Chronic Kidney Disease, Machine Learning, MJ Database

1. Introduction

Many people suffer from Chronic Kidney Disease (CKD). According to Central Health Insurance Agency Ministry of Health and Welfare's reported, the top ten diseases for universal health insurance and outpatient medical expenses are acute renal failure and chronic kidney disease in 2018[1]. United States Renal Data System show that the prevalence of dialysis is the highest in the world[2]. According to the statistics on the causes of death in 2018 Taiwan people released by the Statistics Department of the Ministry of Health and Welfare, kidney disease is the ninth of the top ten causes of death. In summary, kidney disease not only requires huge medical expenses, but it is also one of the main causes of death for Taiwanese.

Machine learning can analyze the hidden parameters between the associations that humans cannot find, and assist humans in the development of various fields, such as assisting doctors in the medical field to make diagnosis, or predicting the probability of disease and finding out the relationship between disease and disease. In the past, many studies have used machine learning techniques to predict the occurrence or probability of kidney disease [3-5], stage prediction of chronic kidney disease [6], and find out biochemical values related to renal function or lifestyle habits or diseases [7, 8]. Current prediction models and analyses related to kidney disease mostly use data from patients and visits in different countries and hospitals [3, 4], or use publicly available data sets [5-8]. At present, the examination of kidney disease requires invasive examination to obtain biochemical values, and the biochemical values are used to assess the condition of the kidney. If a non-invasive self-screening is established and provided to users, it can avoid the discomfort of invasive examinations. Self-screening methods Fast and free.

Therefore, this study used different machine learning method to established an identification model of Chronic Kidney Disease, and model can help participants and doctor to understand the participant's kidney status and reduce end-stage renal disease the risks that occur.

2. Method

2.1 Tool

We applied the SPSS26, SQL Server2014 Management and Python3.6.5. SPSS was used to process data, count the population and analyze the related factors. SQL Server was used for data access, and Python was used for training, establishing and evaluation models.

2.2 Data resource

We utilized the MJ health database from 2001-2015. In this study, we have 190,200 participants and 377,124 records. The data included examination data and questionnaire data. The questionnaire data included diseases and living habits that the participants have suffered.

2.3 Research Process

In this study, collected non-invasive factors related to chronic kidney disease from literature, such as examination data and questionnaire data, as shown in Table 1. The standard for the identification of CKD is to refer to the estimation of the glomerular filtration rate(eGFR) proposed by the American National Kidney Foundation. According to the standard, if eGRF < 60 ml/min/1.73 m^2 and more than three months[9], which means that the participant may already have CKD. In MJ database we only have one record one year, so we need to make sure participant situation, the diagnostic criteria for CKD were as follows:

1. If eGFR < 60 ml/min/1.73 m^2 for two consecutive years, the participant is judged to be chronic kidney disease.
2. If eGFR >=60 ml/min/1.73 m^2 for two consecutive years, the participant is judged to be without chronic kidney disease.

This study was divided into experimental group and control group. People who did not meet the conditions were deleted, such as 114,279 participants who had less than two health check-ups, and 52,995 participants who did not have health check-ups for two consecutive years. A total of 22,926 participants were included in the study. After data processing, there were 573 participants with chronic kidney disease and 22,353 participants without chronic kidney disease. Because the number of people in the two groups was very different, we used random sampling method to match participants who without chronic kidney disease. The matched ratio of the design of this study was 3:1 between no chronic kidney disease and chronic kidney disease, and the data is divided into training and testing sets 8:2, as shown in Figure 1.

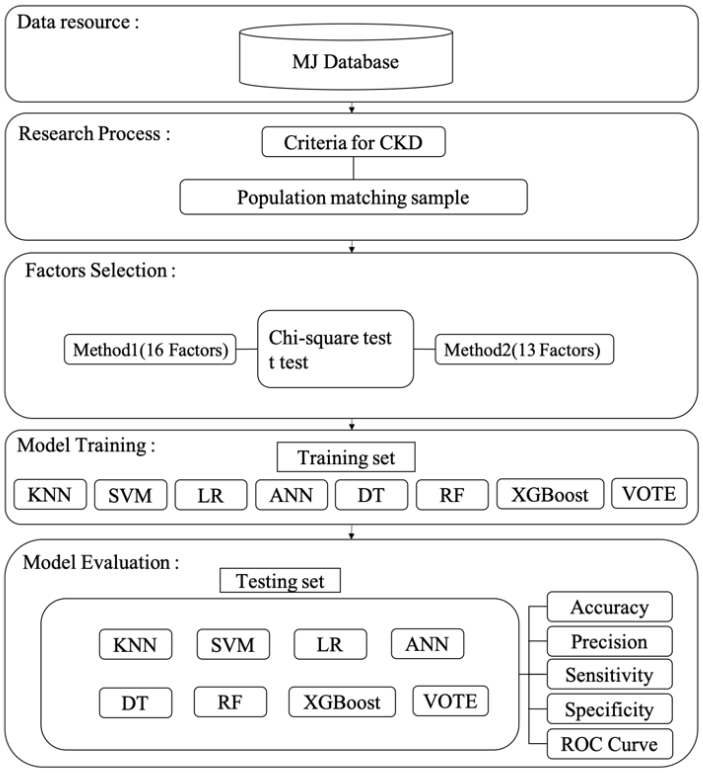


Figure 1 Research architecture

2.4 Factors Selection

We had method1 and method2, method1 was used all factors, method2 was used significant factors. We divided into categories and numerical factors according to the MJ Health database. The factors belong to the category types use the Chi-square test to investigate whether they are related to chronic kidney disease, and the numerical factors use the t test to investigate whether the average value between the two samples is different. In this study, the P value set in < 0.01.

2.5 Model Training

We used eight different machine learning methods and compared accuracy, including K-Nearest Neighbor (KNN), Support Vector Machine (SVM), Logistic Regression (LR), Artificial Neural Network (ANN), Decision Tree (DT), Random Forest (RF), extreme Gradient Boosting (XGBoost) and voting algorithm (Vote).

1. KNN: The basic idea of the KNN algorithm is to calculate the distance between the data to be classified and all the training samples, and compare the distance similarity of the training samples, and classify the data to be classified into the samples with the closest distance.
2. SVM: The concept of support vector machine is to find a hyperplane and effectively distinguish data. Hyperplane means a multi-dimensional plane. In terms of two dimensions, the support vector will find a line to distinguish the two

types of data. It is expected that the distance between the two types of lines (Margin) is as large as possible, which can effectively distinguish the data.

3. LR: LR is mostly used in binary classification, related directions, or explaining the relationship between variables and variables. The goal is to find a function that best represents all observations and distinguish the two types of data.
4. ANN: It can perform a large number of calculations, storage, and processing tasks, and through automatic learning, find out the correlation between samples, assist human analysis and effectively solve problems. Its advantage is that it can tolerate larger data errors and larger data adaptability.
5. DT: A decision tree classifies problems in a tree structure, and can also solve regression problems. A decision tree mainly contains root nodes, internal nodes, and leaf nodes. The root node is the data sample and is the starting point of the decision tree. Internal nodes will be trained to generate classification rules. The leaf node is the result of the final data sample that can be classified. Each piece of data will start from the root node, determine the direction of the sample through the rules of the internal nodes, and finally know which category it belongs to the leaf node.
6. RF: Random forest is derived from the concept of decision tree. It is first predicted by multiple individual decision trees, and the results are jointly determined by the predicted results.
7. XGBoost: The concept of XGBoost is to continuously increase the tree and split continuously. The reason is to achieve the effect of improving the model. When the growth of the tree is completed, the predicted result is the score corresponding to the same feature of each tree, and the total is the predicted value.
8. Vote: Use the above seven classifiers as voting methods, including KNN, SVM, LR, ANN, DT, RF, XGBoost, and use different weighted values to vote.

2.6 Model Evaluation

In this study, we used confusion matrix and receiver operating characteristic curve (ROC curve). Confusion matrix used accuracy, sensitivity and specificity to evaluate the model.

3. Results

3.1 Demography

The random matching sample is based on the participants with the disease to match the participants without the disease. disease: without disease = 1:3, if there are 573 participants, and the matched participants are 1,719 without the disease. In this sample, the male participants with the disease were 30-69 years old, and the male participants without the disease were 30-49 years old. The female participants who were sick were 50-79 years old, and the female participants who were not sick were 30-59 years old, as shown in Table 1 and Table 2.

Table 1 Demographic of categories factors and p-value

Factors **SUM**	**Description**	**Experimental Group (CKD)** **573**	**Control Group (Non-CKD)** **1,719**	**p-value**
Gender	male	38	430	.000
	female	535	1,289	
Age	<30	4	240	.000
	30-39	21	618	
	40-49	67	487	
	50-59	101	243	
	60-69	227	99	
	70-79	126	30	
	>80	27	2	
Pedal Edema	no	551	1,710	.000
	yes	22	9	
Hypertension	no	365	1,644	.000
	yes	208	75	
Diabetic	no	491	1,688	.000
	yes	82	31	
Cardiovascular Diseases	no	474	1,682	.000
	yes	99	37	
anemia	no	478	1,456	.465
	yes	95	263	
smoke	no	20	112	.000
	second-hand smoking	515	1,362	
	smoking cessation	10	51	
	1-3(times/week)	10	55	

Factors **SUM**	**Description**	**Experimental Group (CKD)** **573**	**Control Group (Non-CKD)** **1,719**	**p-value**
	4-5(times/week)	1	33	
	everyday	17	106	
Drink	no	541	1,562	.006
	abstinence	8	13	
	1-2(times/week)	15	100	
	3-4(times/week)	5	34	
	everyday	4	10	
Exercise	low-intensity exercise	288	828	.000
	moderate-intensity exercise	30	699	
	high-intensity exercise	2	149	
	strenuous exercise	0	29	
	no exercise	253	5	
Sleep Time	<4(hours/day)	21	12	.000
	4-6(hours/day)	177	360	
	6-7(hours/day)	280	1,053	
	7-8(hours/day)	81	255	
	8-9(hours/day)	10	35	
	>9(hours/day)	4	4	

Table 2 Numerical factors and p-value

Factors	Experimental Group (CKD)		Control Group (Non-CKD)		p-value
	Mean	SD	Mean	SD	
BMI	23.9	3.5	28.3	3.6	.000
Body Fat Percentage	30.8	6.8	27.4	6.9	.484
Waistline	77.3	9.0	73.5	9.2	.843
Systolic Blood Pressure	129	21.5	112	15.9	.000
Diastolic Blood Pressure	73	11.3	68	10.1	.001

3.2 Factors selection

We had two different method in factors selection, respectively are Method1 and Methld2.

1. Method1: a total of 16 factors, namely gender, age, hypertension, diabetes, cardiovascular disease, anemia, smoking, drinking, chewing betel nut, sleep time,

exercise, waist circumference, systolic blood pressure, diastolic blood pressure, Lower limb edema and body mass index.

2. Method2：Used significant factors(P value < 0.01). A total of 13 factors, namely gender, age, hypertension, diabetes, cardiovascular disease, smoking, drinking, sleep time, exercise, systolic blood pressure, diastolic blood pressure, Lower limb edema and body mass index.

3.3 Model Evaluation

In Method1 the best classifier was used Vote algorithm and need to 16 factors. the accuracy is 88%, the precision is 71%, the sensitivity is 71%, the specificity is 92%, and the AUC is 0.92. However, In Method2 the best classifier was used Vote algorithm and only need to 13 factors. the accuracy is 88%, the precision is 73%, the sensitivity is 69%, the specificity is 93%, and the AUC is 0.92, as shown in Table 3 .Confusion matrix as shown in Figure 2 and Figure 3.

Table 3 Method1 and Method2's result.

	Method1					Method2				
	Accuracy	Precision	Sensitivity	Specificity	AUC	Accuracy	Precision	Sensitivity	Specificity	AUC
KNN	0.84	0.64	0.53	0.92	0.79	0.85	0.64	0.61	0.91	0.83
SVM	0.8	0.5	0.91	0.77	0.92	0.81	0.51	0.91	0.78	0.92
LR	0.79	0.49	0.91	0.75	0.92	0.78	0.48	0.93	0.74	0.91
ANN	0.81	0.77	0.11	0.99	0.81	0.52	0.29	0.93	0.42	0.85
DT	0.81	0.52	0.61	0.86	0.73	0.76	0.44	0.55	0.82	0.69
RF	0.86	0.7	0.57	0.94	0.9	0.86	0.7	0.61	0.93	0.9
XGBOOST	0.88	0.71	0.69	0.93	0.91	0.88	0.72	0.68	0.93	0.91
VOTE	**0.88**	**0.71**	**0.71**	**0.92**	**0.92**	**0.88**	**0.73**	**0.69**	**0.93**	**0.92**

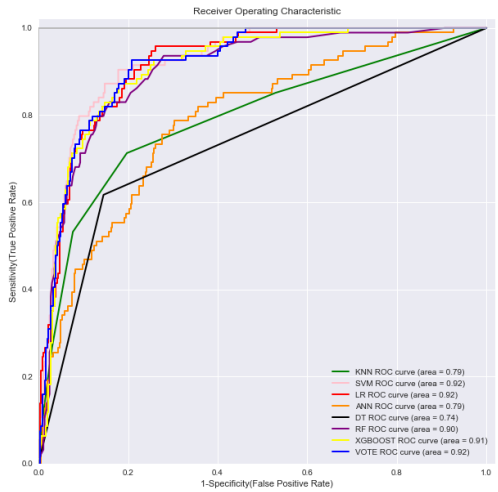


Figure 2 Method1's confusion matrix.

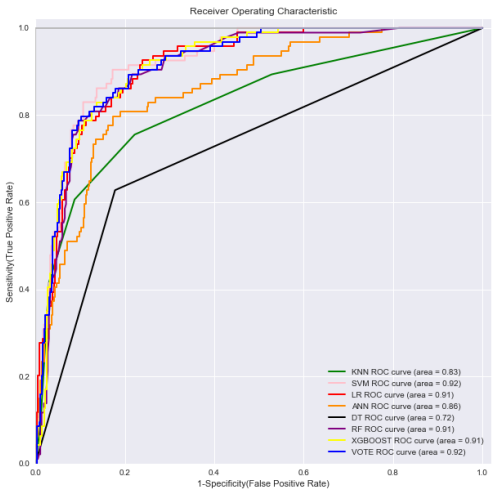


Figure 3 Method2's confusion matrix.

4. Discussion

The results show that Vote is the best method for this research. The Vote method has the best results for multiple different selection factors. It can be verified by this research that combining multiple weak classifiers into a powerful classifier is better than a single classifier. The result of the evaluation is good.

In the past, it is found that most of the current kidney disease models use blood, urine, physical examination and disease history to evaluate together. For example, some scholars use patients' blood, urine, and disease history data to establish chronic kidney disease[3, 5, 6]. In this study, the model established is the use of significant factors collected in the literature. Non-invasive factors can be used as a way to evaluate the disease. The factors are gender, age, hypertension, diabetes, cardiovascular disease, smoking, drinking, sleep time, Exercise, systolic blood pressure, diastolic blood pressure, lower extremity edema, body mass index total 13 factors to assess the disease, and the accuracy can reach 0.88.

From the database of this study, among the 190,200 participants, 133,883 were women and only 56,317 were men. Therefore, when this study judged whether the participants had chronic kidney disease, the number of men suffering from chronic kidney disease was lower than that of women, with 38 men and 535 women, resulting in a large gender gap in the study population. The MJ Health Database has more female data than male. The reason may be that women pay more attention to personal health and more women actively participate in health examinations. Therefore, this study has more data on female than male.

5. Conclusion

We established the model only used non-invasive factors. The factors had gender, age, hypertension, diabetes, cardiovascular disease, smoking, drinking, sleep time, Exercise, systolic blood pressure, diastolic blood pressure, lower extremity edema, body mass index total 13 factors to assess the disease, and the accuracy can reach 0.88.

Using non-invasive factors to identify chronic kidney disease, you only need to acquire the participant's general examination data and questionnaire data, such as disease information, weight or height. In this study, the model can help participants to quickly understand whether they have chronic kidney disease without invasive examinations. However, the identification in this study is a preliminary assessment, and ultimately requires professional medical personnel to diagnose. In the future, it can add such as medication history, family medical history and eating habits, maybe there are better evaluation results.

Reference

1. 中央健康保險署醫務管理組, *2018 年國人全民健康保險就醫疾病資訊*. 2019.
2. *Annual Data Report.* UNITED STATES RENAL DATA SYSTEM, 2018.
3. Diciolla, M., G. Binetti, T. Di Noia, F. Pesce, F.P. Schena, A.M. Vågane, R. Bjørneklett, H. Suzuki, Y. Tomino, and D. Naso, *Patient classification and outcome prediction in IgA nephropathy.* Computers in Biology and Medicine, 2015. **66**: p. 278-286.
4. Di Noia, T., V.C. Ostuni, F. Pesce, G. Binetti, D. Naso, F.P. Schena, and E. Di Sciascio, *An end stage kidney disease predictor based on an artificial neural networks ensemble.* Expert Systems with Applications, 2013. **40**(11): p. 4438-4445.
5. Almansour, N.A., H.F. Syed, N.R. Khayat, R.K. Altheeb, R.E. Juri, J. Alhiyafi, S. Alrashed, and S.O. Olatunji, *Neural network and support vector machine for the prediction of chronic kidney disease: A comparative study.* Computers in Biology and Medicine, 2019. **109**: p. 101-111.
6. Rady, E.-H.A. and A.S. Anwar, *Prediction of kidney disease stages using data mining algorithms.* Informatics in Medicine Unlocked, 2019. **15**: p. 100178.
7. Akben, S.B., *Early Stage Chronic Kidney Disease Diagnosis by Applying Data Mining Methods to Urinalysis, Blood Analysis and Disease History.* IRBM, 2018. **39**(5): p. 353-358.
8. Sheng, Y.-H., J.-H. Chen, J.-M. Chiou, K.-S. Tsai, Y.-Y. Lee, C.-K. Tsao, and Y.-C. Chen, *Association of Renal Function and Menopausal Status with Bone Mineral Density in Middle-aged Women.* Scientific reports, 2015. **5**: p. 14956-14956.
9. *Clinical Practice Guidelines For Chronic Kidney Disease: Evaluation, Classification and Stratification.* National Kidney Foundation K/DOQI, 2002.
10. Fawcett, T., *An introduction to ROC analysis.* Pattern Recognition Letters, 2006. **27**(8): p. 861-874.

Using AI algorithm to Establish the CVD Risk Assessment Model

Yin-Chen Chen[1], Hsiu-An Lee[2] and Chien-Yeh Hsu[1,3*]

[1] National Taipei University of Nursing and Health Sciences Department of Information Management, Taipei, Taiwan
[2] Tamkang University Department of Computer Science and Information Engineering, New Taipei City, Taiwan
[3] Taipei Medical University Master Program in Global Health and Development, Taipei, Taiwan

Abstract. In Taiwan, diseases with cardiovascular include heart disease and cerebrovascular disease among the top ten causes of death. With the development of data mining in the medical field, it can be used to establish the risk prediction models of disease as a tool to assist physicians in decision-making in diagnosis.

The purpose of this study is taking Taiwanese health examination data as an example, to identify the significant risk factors of cardiovascular disease and to establish the risk assessment model of cardiovascular disease using data mining. Using Chi-square test and Information Gain identify the correlation between various factors and cardiovascular diseases. Using eight algorithms such as decision tree, random forest, XGBoost, neural network, logistic regression, support vector machine, K nearest neighbor algorithm and voting algorithm to establish the risk assessment model of cardiovascular disease and using confusion matrix and AUC as model evaluation. Compare model performance with different factor combinations.

Experiment result shows that it finds 22 questionnaires and biochemical variable risk factors affecting cardiovascular disease and 10 questionnaire factors. ANN and VOTE with 22 factors of Information Gain (threshold>0.01) are the best models. Both models have the same accuracy (0.88) and AUC (0.90). The best model of questionnaire variable is ANN with the accuracy rate is 0.89, and the AUC is 0.91.

Keywords: Cardiovascular disease, Data mining, Feature selection, Risk assessment model

1 Introduction

1.1 Background

Cardiovascular diseases are the most common non-communicable diseases in the world, and it is also the first in the global mortality ranking. According to the research, about 17.7 million people died of cardiovascular diseases globally in 2017,

and more than three-quarters of them occurred in low-income and middle-income countries [1].

In Taiwan, cardiovascular diseases contain heart disease and cerebrovascular disease, and both of them are among the top ten causes of death. Table 1 shows the statistics of the top ten causes of death in Taiwan in the most recent three years from the Ministry of Health and Welfare. Therefore, the prevention of cardiovascular diseases is paying more and more attention to their health [2].

According to the Global atlas on cardiovascular disease prevention and control guideline, many risk factors, such as age, gender, personal lifestyle and disease history, are the causes of atherosclerosis [3]. Therefore, early control or reduction of the impact of risk factors on cardiovascular disease can reduce the incidence and mortality of cardiovascular diseases.

In order to reduce the incidence and mortality of cardiovascular diseases, the World Health Organization actively promotes the prevention of cardiovascular diseases in the world. More and more counties build the risk prediction models of cardiovascular diseases. However, many studies have shown that the famous Framingham risk prediction model, when the population data of different countries are used for evaluation, it is observed that the risk values are overestimated and underestimated, which may be caused by regional and ethnic differences, and different risk factor combinations [4, 5].

Table 1. The statistics of the top ten causes of death in Taiwan[2]

Years	105		106		107	
Causes of death	Death toll	Death toll ratio(%)	Death toll	Death toll ratio(%)	Death toll	Death toll ratio(%)
Malignant neoplasms	47,760	27.7	48,037	27.8	48,784	28.2
Diseases of heart	20,812	12.1	20,644	11.9	21,569	12.5
Pneumonia	12,212	7.1	12,480	7.2	13,421	7.8
Cerebrovascular diseases	11,846	6.9	11,755	6.8	11,520	6.7
Diabetes mellitus	9,960	5.8	9,845	5.7	9,374	5.4
Accidents and adverse effects	7,206	4.2	6,965	4.0	6,846	4.0
Chronic lower respiratory diseases	6,787	4.0	6,260	3.6	6,146	3.6
Hypertensive diseases	5,881	3.4	6,072	3.5	5,991	3.5
Nephritis, nephrotic syndrome and nephrosis	5,226	3.0	5,381	3.1	5,523	3.2
Chronic liver disease and cirrhosis	4,738	2.7	4,554	2.6	4,315	2.5

In the medical field, many researches have begun to use data mining and combine the clinical data as variables for analysis to apply the prediction model of diseases, such as breast cancer [6, 7], cardiovascular disease [8, 9]. Some research also combined with feature selection technology, used to confirm the risk factors of the disease, and can increase the accuracy of the model. Through the data mining technique can help physicians as one of the tools to assist diagnosis and decision-making, and achieve the goals of reducing medical costs and effectively improving medical quality services.

1.2 Object

In this study, we use Taiwan health examination data as an example, and combine data mining technique to build a risk assessment model of cardiovascular disease in Taiwan. The following are the goals of this study:

(1) Using data mining to identify relevant risk factors affecting cardiovascular diseases.
(2) Using risk factors related to cardiovascular diseases, use data mining to establish a cardiovascular disease risk assessment model.

2 Literature Review

According to the report of the American College of Cardiology/American Heart Association Task Force on Clinical Practice Guidelines [10], risk factors such as age, gender, type 2 diabetes, hypertension, hyperlipidemia, obesity, smoking, diet and exercise affect cardiovascular diseases. Another research shows that kidney disease, sleep and drinking are also one of the risk factors for cardiovascular diseases. And also give advice on prevention and treatment [11].

Many researches indicate that data mining techniques can be used as a tool for assisting diagnosis and decision-making in the medical field, and also apply in the cardiovascular diseases [12, 13]. In the field of cardiovascular diseases, feature selection can not only be used as an important risk factor for finding cardiovascular diseases, but also improve the efficiency of classification models. At present, many literatures propose different methods for selecting features, including combination(2n-1), Information gain and Gini Index used in decision trees, principal component analysis(PCA), mRMR algorithm, etc., which can be used as algorithms for finding features.

Amin scholar proposed a comprehensive search for the risk factors of cardiovascular diseases recognition features, the method used combination (2n-1), combined all the factors in sequence, and compared the accuracy of the combination of features in different classification models. The result shows that the highest accuracy rate after feature selection is 87.41%, and the accuracy rate of without feature selection is 86.30%. Therefore, the model combined with feature selection technology has a higher accuracy rate [14].

Mustaqeem scholar uses information gain as the feature selection method. This method sorts according to the importance of factors on cardiovascular diseases, and

puts the selected important factors into the model. In comparison, the results show that the systolic blood pressure, fasting blood glucose and cholesterol rank in the top three, while in the evaluation model, the accuracy rate of the model without feature selection is 86.49%, and the accuracy rate of the model with feature selection is 91.89% [15].

Hsu Scholar uses the improved mRMR algorithm to identify the risk factors for cardiovascular diseases. The original 21 variables are selected, and finally 9 items are selected as important factors. The factors include age, exercise, weight, height, waist circumference, and kidney the skein filtration rate, low-density cholesterol, body mass index, and pre-meal blood glucose value. The research results show that using the improved mRMR method can effectively increase the accuracy of the model. The highest accuracy rate of the model is 57.80%. And the accuracy rate of the model without using feature selection is 51.79% [12].

Based on the above literature, different feature selection methods are proposed. After comparing the accuracy of each method, it is shown that Information gain has better accuracy. Therefore, the study chooses this method as the feature selection method.

Data mining technique is to find hidden knowledge from a large amount of data. At present, there are many literatures that use different data mining technique to compare model performance, select the algorithm with the best model performance, and combine the use of clinical data.

A critical study published by Alizadehsani scholar, collecting literature related to cardiovascular disease and machine learning methods from 1992 to 2019, and found that most studies used algorithms for machine learning methods, and included categories in the top ten rankings. Neural network, decision tree, support vector machine, K nearest neighbor algorithm, random forest, logistic regression. It is proposed that due to ease of use and low computational load, most studies adopt the above algorithm as a build prediction model approach [16].

According to Amin scholars, the prediction model of cardiovascular disease established with 7 classification models, including K nearest neighbor algorithm, decision tree, naive Bayes, logistic regression, support vector machine, neural network and maximum the voting algorithm. The voting algorithm, uses the combination of Naive Bayes and Logistic regression as the method of building a prediction model, and compares the evaluation indicators of the seven models. The results show that compared with all models, the accuracy of the voting algorithm is 86.30%, which has better model performance [14].

Liu scholar proposed to use XGBoost algorithm for predicting heart disease, comparing random forest and support vector machine. The results show that the accuracy of XGBoost and random forest is higher than support vector machine, and the running time of XGBoost model is better effectiveness [17].

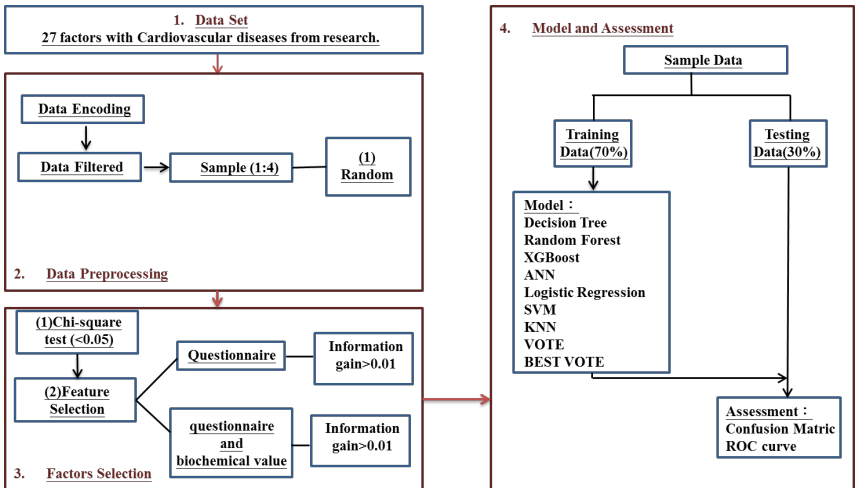
Based on the compilation of the above literature, this study use six algorithms that most research used, including neural network, decision tree, support vector machine, K nearest neighbor algorithm, random forest, logistic regression, and adding voting algorithm of the method and XGBoost of the optimized decision tree are used as the method of building the model, and compare the effectiveness of the model.

3 Method

3.1 Research Framework

Figure 1 shows the structure of the research data analysis process in this study, which is mainly divided into four parts. The first part is organize the literature review to identify the risk factors that have been proven to affect cardiovascular disease ; the second part is preprocesses the data, including data encoding, Data Filtered ; the third part is the feature selection, using chi-square test and feature selection methods to explore the correlation between each risk factor and cardiovascular disease, and identify the best risk factor ; the fourth part is establish model and assessment, using the 7:3 ratio as the training and testing data, using eight classification algorithm like decision tree, random forest, XGBoost, etc. to establish a risk assessment model, the model assessment uses confusion matrix and ROC curve to compare performance of the model.

Fig. 1. The structure of the research data analysis process in this study



3.2 Data Resource and tools

MJ Health Resource Center offers three databases. In this study, MJ Health Survey Data and MJ Biodata were used as the data set.

This study uses SPSS 22 statistical software and Python 3.5 program as research tools.

3.3 Data Preprocessing

The data that the study used were identified according to exclusion and inclusion criteria. The inclusion criteria were as follows：(1) Cases with the number of health

examinations more than two times and have cardiovascular after second examinations.

Due to the disparity between the number of cases and the control group, the study chose sampling data at a ratio of 1:4 can provide better verification power [18]. The data included a total of 8,120 samples, of which 6,496 cases did not have cardiovascular disease, and 1,624 were the first cases of cardiovascular disease during the follow-up period.

3.4 Feature Selection

The chi-square test is used to explore the correlation between the two groups of category variables and determine whether the two groups are independent or related to each other. The p value is less than 0.05, rejecting H0 as two groups. The variables are related and indicate that the two groups of variables have significant differences.

Information gain is another feature selection method.

Another feature selection method is Information gain algorithm that mainly composed of Entropy, and is used as the basis for variable segmentation. Its characteristic is used to measure the expected value of a random variable, indicating the measurement of uncertainty. The Entropy value is between 0~ 1. When the value is higher, it means that there are many different classifications in the data, which need to be classified; when the value is lower, it means that the classification in the data tends to be more consistent, so the variable node will be selected as the classification node with a smaller value, And continue to calculate Entropy for other variables until the decision tree is completely generated.

3.5 Classification model

Decision tree is composed of three parts, including root nodes, branch nodes, and final nodes or target variables. Random forests are composed of multiple decision trees. XGBoost is also composed of trees, and it mainly optimized decision tree.

The structure of neural network includes an input layer, a hidden layer and an output layer. Logistic regression is a multivariate analysis that can handle multiple variables. The characteristics of the support vector machine mainly convert the feature vector of the data from the low-dimensional space to the high-dimensional space, and find a boundary that can maximize the two classifications to achieve the purpose of separation. The characteristic of KNN is to find the k most similar samples in the feature space of new samples, and classify the new samples into the same classification. The weighted average voting algorithm, the calculation method is based on the weight value β of each model, and the probability value of the result, and the weighted average is calculated. The model finally selects the classification result with the larger probability value. The maximum voting algorithm based on the majority votes as the result of model.

4 Result

4.1 Demography

Table 2 shows the comparison of baseline characteristics between cardiovascular diseases and no cardiovascular diseases. All the variables which were significantly different between cardiovascular diseases and no cardiovascular diseases were considered as input variables.

Table 2. Comparison of baseline characteristics

	Cardiovascular Diseases (N=1,624)	No Cardiovascular Diseases (N=6,496)	
Category variable	N (%)	N (%)	p
Gender			
Male	15(0.9)	1,743(26.8)	<0.001
Female	1,609(99.1)	4,753(73.2)	
Personal medical history			
Hyperlipidemia	161(9.9)	34(0.5)	<0.001
Hypertension	686(42.2)	195(3.0)	<0.001
Diabetes	164(10.1)	60(0.9)	<0.001
Chronic nephritis	54(3.3)	82(1.3)	<0.001
Obesity	326(20.1)	452(7.0)	<0.001
Female			
Menopause	979(60.3)	818(12.6)	<0.001
Living habit			
Smoke	47(2.9)	707(10.9)	<0.001
Drink	67(4.1)	524(8.1)	<0.001
Sleep Time (>=six hours)	1,029(63.4)	5,131(79.0)	<0.001
Sport	689(42.4)	895(13.8)	<0.001
Vegetable intake (>=1bowl)	688(42.4)	2,138(32.9)	<0.001
Fruit intake (>=2 servings)	263(16.2)	923(14.2)	<0.05
Continuous variable	Mean(SD)	Mean(SD)	p
Physiological Information			
Age	56.8(14.0)	37.5(12.3)	<0.001
Height (cm)	155.3(6.2)	161.2(8.7)	<0.001
Weight (kg)	58.2(10.0)	57.9(11.7)	0.287
Body Mass Index (kg/m^2)	24.2(4.0)	22.2(3.5)	<0.001

Waist circumference (cm)	77.3(9.4)	72.9(9.3)	<0.001
Systolic blood pressure (mmHg)	127.3(22.2)	113.9(15.9)	<0.001
Diastolic blood pressure (mmHg)	73.2(12.1)	67.9(10.4)	<0.001
Biochemical value			
Fasting blood glucose (mg/dl)	107.4(25.6)	96.4(15.1)	<0.001
Urea Nitrogen (mg/dl)	15.0(5.7)	12.8(3.3)	<0.001
Triglycerides (mg/dl)	118.9(62.9)	95.0(55.7)	<0.001
Total cholesterol (mg/dl)	201.0(36.2)	188.4(34.3)	<0.001
High density lipoprotein (mg/dl)	60.1(14.8)	59.8(15.2)	0.535
Low-density lipoprotein (mg/dl)	117.8(32.3)	109.7(30.6)	<0.001
C reactive protein (mg/dl)	0.3(0.5)	0.2(0.3)	<0.001
High sensitivity C-reactive protein (mg/L)	0.08(0.66)	0.02(0.41)	<0.01

4.2 Correlation of factors with Cardiovascular diseases

Table 3 presents ranking of factors based on information Gain. The criteria of split based on research, therefore the chosen threshold is >0.01 as the basis for ranking.

Table 3. Ranking of factors based on information Gain

Ranking	Factor types	Factors	Information gain
1	Questionnaire	Age	0.204880917
2	Biochemical value	Body Mass Index	0.193043087
3	Questionnaire	Menopause	0.161049999
4	Questionnaire	Hypertension	0.143267651
5	Biochemical value	Height	0.119927131
6	Biochemical value	Fasting blood glucose	0.081007618
7	Biochemical value	Systolic blood pressure	0.078887038

8	Questionnaire	Gender	0.067358258
9	Biochemical value	Triglycerides	0.061613503
10	Biochemical value	Urea Nitrogen	0.058447825
11	Questionnaire	Sport	0.052578923
12	Biochemical value	Total cholesterol	0.035539271
13	Biochemical value	Waist circumference	0.034726138
14	Questionnaire	Hyperlipidemia	0.032418137
15	Biochemical value	Diastolic blood pressure	0.031622301
16	Biochemical value	Low-density lipoprotein	0.027408115
17	Questionnaire	Diabetes	0.027166106
18	Questionnaire	Obesity	0.019412209
19	Biochemical value	High sensitivity C-reactive protein	0.01692414
20	Questionnaire	Sleep(more than six hours)	0.014372357
21	Biochemical value	C reactive protein	0.014016526
22	Questionnaire	Smoke	0.01100842

4.3 Performance of models

Table 4 shows the performance of models based on Chi-square test with questionnaire and biochemical value variables. The best model was ANN and VOTE. Table 5 the performance of models based on Information Gain (threshold>0.01) with questionnaire and biochemical value variables. The best model was ANN and VOTE. Table 6 the performance of models based on Information Gain (threshold>0.01) with questionnaire variable. The best model was ANN.

Table 4. The performance of models based on Chi-square test with questionnaire and biochemical value variables

Model	Cut Point	Accuracy	Sensitivity	Specificity	Precision	AUC
Decision Tree	0.5	0.79	0.86	0.78	0.5	0.89
	0.5	0.79	0.86	0.78	0.5	0.89
Random Forest	0.5	0.85	0.74	0.88	0.62	**0.9**
	0.46	0.84	0.77	0.86	0.6	**0.9**
XGBoost	0.5	0.87	0.66	0.92	0.68	0.88
	0.31	0.83	0.73	0.86	0.58	0.88

ANN	0.5	**0.88**	0.51	0.98	0.85	**0.9**
	0.17	0.85	0.78	0.87	0.6	**0.9**
Logistic Regression	0.5	0.82	0.72	0.85	0.56	0.86
	0.36	0.77	0.83	0.75	0.47	0.86
SVM	0.5	0.82	0.82	0.82	0.54	**0.9**
	0.28	**0.87**	0.74	0.91	0.68	**0.9**
KNN	0.5	0.8	0.3	0.93	0.52	0.69
	0.26	0.68	0.65	0.68	0.35	0.69
VOTE	0.5	**0.88**	0.67	0.94	0.74	**0.9**
	0.39	0.85	0.76	0.88	0.61	**0.9**
BEST VOTE	—	0.84	0.78	0.86	0.59	—

Table 5. The performance of models based on Information Gain (threshold>0.01) with questionnaire and biochemical value variables

Model	Cut Point	Accuracy	Sensitivity	Specificity	Precision	AUC
Decision Tree	0.5	0.79	0.86	0.78	0.5	0.89
	0.5	0.79	0.86	0.78	0.5	0.89
Random Forest	0.5	0.85	0.75	0.88	0.62	**0.9**
	0.45	0.84	0.79	0.85	0.58	**0.9**
XGBoost	0.5	0.86	0.62	0.92	0.66	0.87
	0.2	0.79	0.78	0.79	0.49	0.87
ANN	0.5	**0.88**	0.6	0.96	0.8	**0.9**
	0.28	**0.86**	0.76	0.89	0.64	**0.9**
Logistic Regression	0.5	0.82	0.72	0.56	0.85	0.86
	0.4	0.8	0.78	0.51	0.8	0.86
SVM	0.5	0.81	0.82	0.81	0.53	**0.9**
	0.21	0.83	0.8	0.84	0.57	**0.9**
KNN	0.5	0.8	0.3	0.93	0.52	0.69
	0.26	0.68	0.65	0.68	0.35	0.69
VOTE	0.5	**0.88**	0.68	0.93	0.72	**0.9**
	0.39	**0.86**	0.75	0.89	0.63	**0.9**
BEST VOTE	—	0.83	0.8	0.84	0.57	—

Table 6. The performance of models based on Information Gain (threshold>0.01) with questionnaire variable

Model	Cut Point	Accuracy	Sensitivity	Specificity	Precision	AUC
Decision Tree	0.5	0.79	0.86	0.78	0.5	0.89
	0.62	0.83	0.81	0.83	0.56	0.89
Random Forest	0.5	0.85	0.79	0.86	0.6	0.9
	0.51	0.85	0.78	0.86	0.6	0.9

XGBoost	0.5	0.84	0.79	0.85	0.59	0.9
	0.53	0.86	0.78	0.87	0.62	0.9
ANN	0.5	**0.89**	0.59	0.96	0.81	**0.91**
	0.27	**0.87**	0.76	0.9	0.67	**0.91**
Logistic Regression	0.5	0.87	0.77	0.89	0.65	0.9
	0.5	**0.87**	0.77	0.89	0.65	0.9
SVM	0.5	0.83	0.8	0.84	0.57	0.9
	0.26	0.84	0.78	0.86	0.59	0.9
KNN	0.5	0.86	0.55	0.94	0.72	0.82
	0.29	0.79	0.76	0.8	0.5	0.82
VOTE	0.5	0.88	0.68	0.93	0.73	**0.91**
	0.44	0.86	0.78	0.88	0.64	**0.91**
BEST VOTE	—	0.86	0.79	0.88	0.63	—

5 Conclusion

In this study, we use Taiwan health examination data as an example, and combine data mining technique to build a risk assessment model of cardiovascular disease in Taiwan.

The best models with 22 questionnaires and biochemical value variables were ANN and VOTE based on Information Gain (threshold value>0.01). The best models with 10 questionnaires variables were ANN based on Information Gain (threshold value>0.01).

References

1. Roth, G.A., et al., *Global, regional, and national age-sex-specific mortality for 282 causes of death in 195 countries and territories, 1980–2017: a systematic analysis for the Global Burden of Disease Study 2017.* The Lancet, 2018. **392**(10159): p. 1736-1788.
2. 行政院衛生福利部統計處, *國人死因統計*. 2018.
3. Mendis, S., et al., *Global atlas on cardiovascular disease prevention and control*. 2011: Geneva: World Health Organization.
4. Wallisch, C., et al., *External validation of two Framingham cardiovascular risk equations and the Pooled Cohort equations: A nationwide registry analysis.* International Journal of Cardiology, 2019. **283**: p. 165-170.
5. Cook, N.R. and P.M. Ridker, *Calibration of the Pooled Cohort Equations for Atherosclerotic Cardiovascular Disease: An Update.* Ann Intern Med, 2016. **165**(11): p. 786-794.

6. Sarvestani, A.S., et al. *Predicting breast cancer survivability using data mining techniques*. in *2010 2nd International Conference on Software Technology and Engineering*. 2010. IEEE.
7. Jonsdottir, T., et al., *The feasibility of constructing a Predictive Outcome Model for breast cancer using the tools of data mining.* Expert Systems with Applications, 2008. **34**(1): p. 108-118.
8. Shouman, M., T. Turner, and R. Stocker. *Using data mining techniques in heart disease diagnosis and treatment*. in *2012 Japan-Egypt Conference on Electronics, Communications and Computers*. 2012. IEEE.
9. Yeh, D.-Y., C.-H. Cheng, and Y.-W. Chen, *A predictive model for cerebrovascular disease using data mining.* Expert Systems with Applications, 2011. **38**(7): p. 8970-8977.
10. Arnett, D.K., et al., *2019 ACC/AHA Guideline on the Primary Prevention of Cardiovascular Disease: Executive Summary: A Report of the American College of Cardiology/American Heart Association Task Force on Clinical Practice Guidelines.* Journal of the American College of Cardiology, 2019. **74**(10): p. 1376-1414.
11. Benjamin, E.J., et al., *Heart Disease and Stroke Statistics—2019 Update: A Report From the American Heart Association.* Circulation, 2019. **139**(10): p. e56-e528.
12. Hsu, W.-Y., *A decision-making mechanism for assessing risk factor significance in cardiovascular diseases.* Decision Support Systems, 2018. **115**: p. 64-77.
13. Shinde, S. and A. Priyadarshi, *Diagnosis of Heart Disease Using Data Mining.*
14. Amin, M.S., Y.K. Chiam, and K.D. Varathan, *Identification of significant features and data mining techniques in predicting heart disease.* Telematics and Informatics, 2019. **36**: p. 82-93.
15. Mustaqeem, A., et al., *A statistical analysis based recommender model for heart disease patients.* International Journal of Medical Informatics, 2017. **108**: p. 134-145.
16. Alizadehsani, R., et al., *Machine learning-based coronary artery disease diagnosis: A comprehensive review.* Computers in Biology and Medicine, 2019. **111**: p. 103346.
17. 刘宇 and 乔木, *基于聚类和 XGboost 算法的心脏病预测.* 计算机系统应用, 2019. **28**(1): p. 228-232.
18. 白其卉, *打一場公平的拳賽-淺談病例對照研究之設計原則.* 腦中風會訊, 2010. **17**(1): p. 10-11.

An Architecture of Real-World Data Database

Hsiu-An Lee[1] and Chien-Yeh Hsu[2,3*]

[1] Tamkang University Department of Computer Science and Information Engineering, New Taipei City, Taiwan
[2] National Taipei University of Nursing and Health Sciences Department of Information Management, Taipei, Taiwan
[3] Taipei Medical University Master Program in Global Health and Development, Taipei, Taiwan

Abstract.
In the analysis of medical data, it is now important to use "Real-World Data (RWD)". In the past, clinical research in the medical field was often completed by clinical trials. Clinical trials are a kind of use of listed drugs and comfort according to research objectives. The process of comparing drugs or other medical treatments on subjects as a control group.
The Real-world data research database is essential to precision health. In order to develop the value of clinical data, improve the medical care service, and promote precision medicine, the goal of this research is to design a Real-World Data Research Database, aiming at the wide range and standardization of data, the detail component including:
1.International standard Health Level 7-Fast Health Interoperability Resources (HL-7 FHIR)
2.Data sorting and cleaning mechanism.
3.Subject-based database research applications.
4.Blockchain-based technology for identity verification and database authorization recognition mechanism.
Keywords:
Real World Data, Blockchain, FHIR, DOI implementation, Clinical Research database

1 Introduction

1.1 Background

In recent years, big data analysis and artificial intelligence have developed vigorously. Different research fields have incorporated "big data" as a basis for improving the current situation. In the domain of clinical medical data research, big data has become the basis of precision medicine. Clinical research sources usually from a single database or a specific research data set, the format and content of clinical data have not been standardized.

The data preprocessing of each research is completed based on the researcher's experience, there is no common medical research data sharing platform.

Many of these studies are completed by randomized clinical trials (RCTs), a clinical trial is a way to use marketed drugs, placebos or therapies as a control group according to the research objectives, and conduct drugs or other medical treatments on subjects. In the process of clinical trial comparative testing, the researcher first decides the therapy to be tested (such as a drug or device). After it, decides which therapy to compare with it, and which type of patient must be found as the test subject.

Therapeutic drugs need to prove that they can effectively prolong the life of the patient, reduce specific symptoms or reduce the occurrence of adverse events to improve the quality of life of the patient. However, RCTs must establish many limitations in advance, which may not be effective in achieving the results of the trial in practice in the reality. Therefore, in recent years, the use of real-world data (RWD) has been gradually pursued. RWD in medicine is data from different patient groups and sources in the real environment, such as patient surveys, clinical trials, observational cohort studies, and patient report outcome. RWD refers to observational data, not data collected in experimental environments such as randomized controlled trials. RWD may come from electronic health records (EHR), insurance claims (health insurance data), medical products, IoT sensors, and health examination. The goal of RWD is that the data used for analysis should not be affected by factors such as environmental variables, experimental control, and research context setting.

The research results obtained through RWD are called Real-World Evidence (RWE), which can be derived from retrospective or prospective observational research. In 2018, the United States also promulgated the "21st Century Cures Act"[1] FDA expands the role of RWD and research evidence. RWE is mainly used to assist in the analysis of clinical trials that cannot truly target all patient groups with specific diseases. It can answer questions about specific treatment effects and analyze the effects of drugs over a longer period. In clinical practice, RWE is used to understand different patients to provide appropriate care measures for different patients.

RWE analysis requires that RWD data must have sufficient quality. Data quality is defined as consisting of three parts: (1) Consistency (data should follow the specified standard format); (2) Completeness (completeness of actual data); (3) Reasonability (reliability, accuracy, and timeliness of data value) [2]. To comply with RWE, the clinical data database should standardize the data to have sufficient quality.

RWD data comes from four sources-clinical data, management/claims (health insurance) data, patient report outcome data and other data sources (social media and IoT data). The data sources are diverse. Big data analysis continues to develop and mature rapidly. There has some clinical research database for medical research, the relevant descriptions are as follows:

Table 1.

Country	Database	Category	Data volume (million)	Authorized unit
Japan	MHLW	Country Level Database	126	Academic Research
USA	CMS	Country	120	Academic

		Level Database		Research
France	SNIIRAM	Country Level Database	60	Academic Research and Policy maker
	PMSI	Public Hospital Database	60	Academic Research
England	CPRD	EMR Database	53	Part of data for Commercial
	HES	EMR Database	15	Academic Research
Germany	AOK, WldO	Regional level database	24	Academic Research
	Barmer GEK		9	Academic Research
	TK, Wineg		7	Academic Research
Danmark	Sundhed.dk	Country Level Database	6	Academic Research

Many large companies have also begun to provide their own data sets. For example, Microsoft predicts pancreatic cancer from the analysis of search records. Its research has found that retrospective analysis of user searches (such as symptoms) can be performed several months before the formal diagnosis. Clearly identify 5-15% of the undiagnosed population[3]. For this disease with a high fatality rate, a few months of advance prediction can bring about huge changes.

RWE research needs to ensure that key results of effectiveness can be observed and established in RWD sources, and clinical RWD database can provide correct data to affect the overall research design and its clinical and cost burden.

For example, it may be necessary for a data repository to incorporate new data into the EHR or to collect and store biomarkers for subsequent analysis. In clinical research, studies such as Salford Lung Study[4] have proven that the use of RWD data can make potential progress in generating RWE, but it has also found many challenges related to infrastructure and research funding resources. Therefore, a universal RWD database is essential to promote clinical medical research and AI model development.

1.2 Object

The Real-world data research database is essential to precision health. In order to develop the value of clinical data, improve the medical care service, and promote precision medi-cine, the goal of this research is to design a Real-World Data Research

Database, aim-ing at the wide range and standardization of data, the detail component including:

1.International standard - Health Level 7-Fast Health Interoperability Resources (HL-7 FHIR)

2.Data sorting and cleaning mechanism.

3.Subject-based database research applications.

4.Blockchain-based technology for identity verification and database authorization recognition mechanism.

2 Literature Review

In the medical domain, the application of big data has played a very important role in clinical decision support. According to the American Heart Association's report on the primary prevention of cardiovascular disease[5], the significant factors which affect cardiovascular disease including: age, gender, Type 2 diabetes, high blood pressure, hyperlipidemia, obesity, and smoking. The American Heart Association also proposed a study of heart disease and stroke [6]. The result showed that the kidney disease and sleep are the risk factors for cardiovascular disease.

According to one prospective study explored the correlation between gender, age and other cardiovascular risk factors and coronary arteries[7]. The cardiovascular risk factors proposed in the study included smoking, total cholesterol, blood pressure, body mass index and diabetes. The study pointed out when the level of these risk factors increases, the incidence and mortality of coronary arteries are correlated with age.

As the age increases, the proportion of coronary arteries is higher. Among genders, women are more likely to suffer from them than men. According to the American Heart Association's statistics on the prevalence of coronary artery disease in 2010, when the age group is older than 65 years old, the proportion is 20%, 45 to 64 years old is 7%, 18 to 45 years old is 1.3%, and in the same age group, The incidence of men is higher than that of women[6], so gender are related to coronary artery disease. Another study uses clinical data to establish hypertensive prediction models and evaluate the importance risk factor variables in a nationally representative sample of adults over 20 years old (n = 19,799) in the United States and their relationship with the prevalence of hypertension[8]. Arnett et al. present a MDRD formula to calculate the stage of chronic kidney disease using big data analysis[9].

There also have a lot of research based on RWD. Thomas et al.[10] found that sex and age are specific distribution of small cell lung cancer among never smokers and suggests that have a distinct biology of SCLC in never smokers compared with smokers. Emily et al. through the RWD to evaluated the impact of oncology care model reporting in biomarker testing and treatment[11]. Jeffery et al. presented a high accuracy machine learning model to predicted metastatic status from unstructured EHR data. The model can highly determine the status of metastasis in more than 75% of cases without additional manual inspection, which can more effectively identify clini-

cal trial candidates and perform clinical trial matching, thereby reducing the main barriers to clinical trial participation in community clinics[12].

With so many types and multi-faceted applications, big data analysis can find that data analysis and application can effectively help the development of various domains. Especially in the medical domain which can assist physicians in making more accurate judgments and diagnoses.

3 Method and Result

To build an RWD database, there are several necessary elements. This study mainly provides an application structure of the RWD database, and explains the usage, advantages, and design methods.

The main points included are 1. Standardized data architecture, 2. Thematic database construction, and 3. Blockchain architecture security protection.

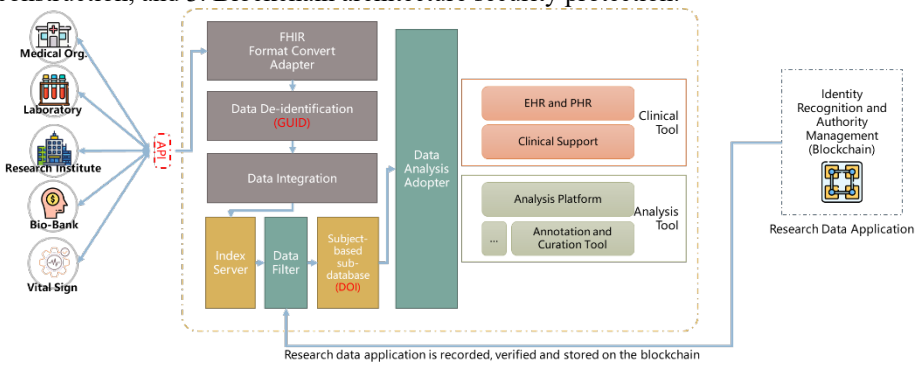


Fig. 1. The architecture of RWD Database

3.1 Standardized data architecture

In the standardized data architecture, the data is first de-identified by Globally Unique Identifier (GUID), which is a unique identifier generated by an algorithm, and can integrate subject information, but will not expose subject identification information.

The purpose of GUID is to allow all elements in a distributed system to have unique identification information. The identification information needs to be specified through the central control terminal. In this way, everyone can create a GUID that does not conflict with others. In this case, there is no need to consider the problem of duplication of names when the database is created. Currently the most widely used GUID, and other important applications, there are Linux ext2/ext3 file system, LUKS encrypted partition, GNOME, KDE, Mac OS X and so on. In addition, we can also find the implementation in the GUID library in the e2fsprogs package. GUID is composed of a string of 16-byte (also called 128-bit) hexadecimal numbers, so the theoretical total of GUID is 216 x 8 = 2128, which is approximately 3.4 x 1038. In other words, if 1 trillion GUIDs are generated every nanosecond, it will take 10 billion years to use up all GUIDs. The standard form of GUID contains 32 hexadecimal

numbers, divided into five segments by hyphens, in the form of 32 characters of 8-4-4-4-12.

After data de-identifier, an international standard should be used for data management. Fast Healthcare Interoperable Resources (FHIR) as fast medical interoperable resources [13] was launched by the HL7 organization in 2011. It is based on the HL7 v2 and HL7 v3 medical information standards, but it is easier to use. The purpose of FHIR is to promote more effective communication and sharing of medical information between medical units [14]. Medical information is widely used on different and diverse devices, such as computers, tablets and smartphones, so that medical units and everyone can more easily Provides and uses medical service information, and provides systematic information to third-party application developers. By labeling different data with exclusive tags, different systems can intercept and combine data more quickly in data exchange.

The FHIR concept is to decompose all health data into many data elements for storage, and is a network-based technology. Compared with HL7 v3, it is more professional and comprehensive. These features make it easier for developers to develop telemedicine applications. Model [15]. In addition, FHIR supports a RESTful web page architecture, which is suitable for applications on mobile devices. It is an HTTP-based programming syntax [16]. If a system has complied with FHIR and uses HTTP protocol, you can get a connection with other systems. Interoperability. In addition, FHIR can be converted into other HL7 standards, which helps to narrow the gap between systems based on previous standards. For example, there is a conversion structure between FHIR and HL7 standard clinical document architecture (CDA), but other standards or unique modules. The system needs a special module to convert to conform to the FHIR format.

In the FHIR data standard, every data element has a tag. Just like on the Internet, every web page has its own URL—Uniform Resource Locator (URL). The FHIR tag is make a unique identifier (Universally Unique Identifier, UUID), in order to make any FHIR browser application can read, FHIR format can also choose JSON or XML for data expression, also because of the unified format, so developers only need to follow Format to read the data content, add the data to the existing system, you can effectively and quickly share data, integrate care, decision-making and analysis in different medical units [16].

This research develop a data conversion encoder in the format of FHIR in C# programming language, and convert the clinical data and pathology reports into the data type of FHIR standard format. In the future, RWD will be provided for research and FDA review at the same time, it has the effect of standardization, formatting and consistency, and can be used as external verification data for others to research and use.

The establishment process is: 1. Inventory the content and structure of the data, 2. Corresponding FHIR resources, 3. Design the corresponding structure and format, 4. Establish the format conversion formula, it is expected to use the composition resources of FHIR to create the content of the entire data as The main framework connected to each other can link all individual resources together to provide an overall data package, and has clinically proven applications.

3.2 Thematic database construction

Researchers filter through specific data, such as age, gender, disease, therapy, medication, etc., to establish a thematic database, which can help data sharing and research reproduction. Thematic database should be established according to common diseases such as: hypertension, diabetes, cardiovascular disease, cancer, etc., and can provide researchers with quick access and analysis.

After thematic database established, a digital object identifier should be given for database identifier. Digital Object Identifier (DOI) is a mechanism used to identify digital resources, including videos, documents, reports, Books or any electronic resources, etc. It also includes a mechanism for naming different resources, as well as a protocol for parsing identification numbers into specific addresses. This research will import the DOI object identification number as the identification number of the research thematic database. In the future, the research data analysis, application, model design, and result publication can be quickly connected to the corresponding database by DOI to obtain relevant information. And continue to promote data sharing and meaningful use. Importing the DOI application can supplement the deficiencies of the uniform resource symbol. Through the DOI, the information of the personally developed model can be located and shared permanently with other authorized users.

After the DOI code is identified and analyzed, it can be connected to more than one data. However, the identification number cannot obtain the digital resource body, and only guides and connects it to ensure the security of the digital resource. The DOI resolution agreement includes RFC 3652 and RFC 3651 describing the naming mechanism, and RFC 3650 describes its architecture. DOI will analyze the identification code through the Handle system developed by it. For example, access the URL https://doi.org/10.1000/2342 and you can see the information corresponding to the digital resource with the identification number 10.1000/2342.

3.3 Blockchain architecture security protection.

Japanese scholar Satoshi Nakamoto proposed the concept of "blockchain" in the 2008 study "Bitcoin White Paper[17]", and founded the Bitcoin network in 2009, establishing the first block, namely "Creation Block". Blockchain is an architecture that does not rely on third parties and uses its own distributed nodes to store, verify, transmit, and communicate network data. It uses cryptography [18, 19] to concatenate and protect content. Even transaction records (also known as blocks), each block contains the encrypted hash of the previous block, the corresponding time stamp and transaction data, which allows the two parties to effectively record the transaction and permanently verify the transaction. Blockchain records combined with the concept of distributed ledger and the design of encrypted hash calculations make the blockchain difficult to tamper with.

Blockchain technology has the following advantages:

1. Decentralization

The structure does not rely on additional third-party management agencies or hardware facilities, nor does it require central control. The blockchain itself, through the

distributed calculation, verification, and consensus mechanism, is completed by each node to complete data verification, transmission and management.

2. Verification

The foundation of blockchain technology is a public ledger. The data contained in the blockchain is open to everyone. Anyone can query the encrypted data of the blockchain through the public interface, except for the transaction party with the decryption private key. To the encrypted data.

3. Independence

Based on consensus specifications and agreements, the entire blockchain system does not need to rely on other third parties to operate, and all nodes can automatically and securely verify and exchange data within the architecture.

4. Security

With the architecture design of the private chain, users cannot manipulate and modify block data arbitrarily, which makes the block chain itself relatively safe and avoids subjective and artificial data tampering.

5. Anonymity

Unless required by laws and regulations, due to blockchain technology, the identity information of each block node does not need to be disclosed or verified, and the information transmission can be done anonymously.

The blockchain architecture is first applied in the financial cryptocurrency market, with algorithm changes, improvements, and workload proof methods to complete the use of different advantages and disadvantages. Subsequently, the blockchain architecture has continued to evolve after its launch. In recent years, the first virtual currency sale ICO, the smart contract blockchain of Ethereum, and the asset token sharing economy of "light ownership and heavy usage rights" have emerged. At present, the blockchain architecture develops decentralized applications (Dapp) in different fields [20] , and builds decentralized autonomous organizations and decentralized autonomous communities (Decentralized autonomous society, DAS)

To establish a RWD research database, the basic information of data screening and application should be made public, such as sample number, sex ratio, gender ratio, screening criteria, various basic statistics, etc., for reference and verification by other researchers. In order to meet the characteristics of RWE, the analysis data must have the characteristics that can be verified. This research proposes to use the characteristics of the blockchain to assist in data verification and authentication. When the RWD research needs to be reviewed, use the blockchain to provide information to confirm that there is no error.

4 Conclusion and Discussion

In the past, clinical research related to the medical field was often completed in clinical trials. Clinical trials are a way to use listed drugs, placebos or therapies as a control group according to the research objectives. The process of comparing drugs or other medical treatments on subjects. In a clinical trial, the researcher must first decide the therapy to be tested, such as a drug or device, and then decide which therapy to compare with it, and which type of patient must be found as the test subject. For therapeutic drugs, it must be proved that it can effectively prolong the life of the patient, relieve specific symptoms or reduce the occurrence of adverse events to improve the quality of life of the patient. However, many restrictions must be established in advance for clinical trials, which may not be effective in achieving the results of the trial in practice.

The real-world evidence is critical to clinical research and medical information application development. RWD research database should develop in different country based on our design. Data can efficiency exchange and management through the architecture we presented. And the data requirement record by blockchain which allows the two parties to record the transaction and permanently verify the transaction.

In the blockchain security design, the records combined with the concept of distributed ledger and the design of encrypted hash calculations make the blockchain difficult to tamper with.

References

1. Hudson, K.L. and F.S. Collins, *The 21st Century Cures Act—a view from the NIH.* New England Journal of Medicine, 2017. **376**(2): p. 111-113.
2. Kahn, M.G., et al., *A harmonized data quality assessment terminology and framework for the secondary use of electronic health record data.* Egems, 2016. **4**(1).
3. Paparrizos, J., R.W. White, and E. Horvitz, *Screening for pancreatic adenocarcinoma using signals from web search logs: Feasibility study and results.* Journal of Oncology Practice, 2016. **12**(8): p. 737-744.
4. Bakerly, N.D., et al., *The Salford Lung Study protocol: a pragmatic, randomised phase III real-world effectiveness trial in chronic obstructive pulmonary disease.* Respiratory research, 2015. **16**(1): p. 101.
5. Arnett, D.K., et al., *2019 ACC/AHA Guideline on the Primary Prevention of Cardiovascular Disease: Executive Summary: A Report of the American College of Cardiology/American Heart Association Task Force on Clinical Practice Guidelines.* Journal of the American College of Cardiology, 2019. **74**(10): p. 1376-1414.
6. Benjamin, E.J., et al., *Heart Disease and Stroke Statistics—2019 Update: A Report From the American Heart Association.* Circulation, 2019. **139**(10): p. e56-e528.
7. Jousilahti, P., et al., *Sex, Age, Cardiovascular Risk Factors, and Coronary Heart Disease.* Circulation, 1999. **99**(9): p. 1165-1172.
8. López-Martínez, F., et al., *Machine learning classification analysis for a hypertensive population as a function of several risk factors.* Expert Systems with Applications, 2018. **110**: p. 206-215.
9. Matsuo, S., et al., *Current status of estimated glomerular filtration rate (eGFR) equations for Asians and an approach to create a common eGFR equation.* Nephrology, 2010. **15**(s2): p. 45-48.
10. Thomas, A., et al., *Clinical and Genomic Characteristics of Small Cell Lung Cancer in Never Smokers: Results From a Retrospective Multicenter Cohort Study.* Chest, 2020.
11. Castellanos, E.H., et al., *Evaluating the Impact of Oncology Care Model Reporting Requirements on Biomarker Testing and Treatment.* JCO Oncology Practice, 2020: p. JOP.19.00747.
12. Kirshner, J.J., et al., *An automated EHR-based tool for identification of patients (pts) with metastatic disease to facilitate clinical trial pt ascertainment*. 2020, American Society of Clinical Oncology.
13. International, H.L.S. *FHIR*. 2019 March 22, 2019; Available from: http://wiki.hl7.org/index.php?title=FHIR.
14. Hatt, N. *Why FHIR is "Better"*. 2015 August 3, 2017; Available from: https://www.redoxengine.com/blog/fhir-better/.
15. International, H.L.S. *FHIR Homepage*. 2018; Available from: https://www.hl7.org/fhir/.
16. Franz, B., A. Schuler, and O. Krauss, *Applying FHIR in an integrated health monitoring system.* EJBI, 2015. **11**(2): p. 51-56.
17. Nakamoto, S., *Bitcoin: A peer-to-peer electronic cash system.*

18. Staff, E.J.T.E.R., *Blockchains: The great chain of being sure about things.* 2016. **18**.
19. Narayanan, A., et al., *Bitcoin and cryptocurrency technologies: a comprehensive introduction*. 2016: Princeton University Press.
20. Raval, S., *Decentralized Applications: Harnessing Bitcoin's Blockchain Technology*. 2016: " O'Reilly Media, Inc.".

An Innovation Study on Applying Deep Learning to Recognize Gesture in Sign Language

Chen-Hu Chou [1], Yu-Yu Yen [2], Yu-Chen Su[3], Horng-Twu Liaw [3], Wei-Chen Wu[4]

[1] Department of Public Policy and Management, Shih Hsin University, Taipei, Taiwan
[2] Department of Finance, Shih Hsin University, Taipei, Taiwan
[3] Department of Information Management, Shih Hsin University, Taipei, Taiwan
[4 *] Department of Finance, Feng Chia University, Taipei, Taiwan, Taichung

Abstract. Deaf people because they cannot hear the outside voice or cannot make a voice, is relying on sign language gestures as a bridge of communication. Not everyone can read sign language to convey the meaning. Deep learning about Convolutional Neural Networks (CNN) in image recognition has a fairly excellent performance. Entering a large number of video data to train models to do identification in recent years. In order to expand the social to recognize 37 Mandarin Phonetic Symbols with deep learning.

Using today's most popular smartphones as an image-shooting tool, a database of 10,598 images of right-handed gestures is built with an average of about 250 to 300 images of right-handed gestures of 720x402 pixels per gesture. A moving image with 96x96x3 input to convolutional neural network, with LeNet and SmallVGGNet Model to do training. Adjust the batch size of training and reduce the learning rate function to do experiments and compare the results. The highest training data accuracy rate is more than 96%, and then with different background gesture images to test its model identification rate, test LeNet up to 51.35 percent and SmallVGGNet up to 29.73 percent.

Built 37 gesture images of Mandarin Phonetic Symbols with single background to train CNN model. The resulting model can be tested with up to 51.35% recognition accuracy with different background gesture images. The experimental process and results can be applied to further studies, and move forward on the language translations between sign and spoken language.

Keywords: . Recognizing Gesture in Sign Language, Deep Learning, Convolutional Neural Network, Mandarin Phonetic Symbols.

1 Introduction

Each country has its own different sign language system and pattern. Taiwan sign language can be broadly divided into phonetic finger, word finger, meaning sign language[1, 2]. Narrow sense of communication function can be divided into natural sign language and grammar sign language[1, 3]. The phonetic symbol refers to the text is a kind of statement, by means of finger forms indicating Pinyin text to communicate with each other.

Mandarin Phonetic Symbols are still the basis of learning the pronunciation of words can be used well when natural sign language fails to express its intention clearly, and the general public can understand better. Utilize deep learning methods to recognize gestures in sign language images for the following purposes:

Above all is using the deep learning method to identify the Pinyin of the gesture in Sign Language refers to the text. With the research of Mandarin Phonetic Symbols which applied in gesture in Sign Language at today's school for the Hearing Impaired, 37 gesture database was established. Secondly, the results of the image data training of different gestures with a single background were experimented with different deep learning models. Lastly, the result of the recognition rate of the model completed by testing the training with another background image data.

2 Materials and methods

The types of sign language and the development artificial intelligence, machine learning and deep learning, as well as the source and model content of the deep learning CNN model used.

2.1 Deep learning and CNN

Neurons are made up of cell body, Dendrite and Axon, as shown in Figure 2-1. After comparing the input signal from the Dendrite to the Threshold, within the nerve cell's Threshold value. When the input signal exceeds the Threshold, the message flows to the Axon output signal.

The structure of the neural network is imitated by the Dendrite input (x), multiplying the weight (w). After adding this value, which is triggered when it exceeds the Threshold (h) to the axon output (y), as shown in Figure 2-2[4].

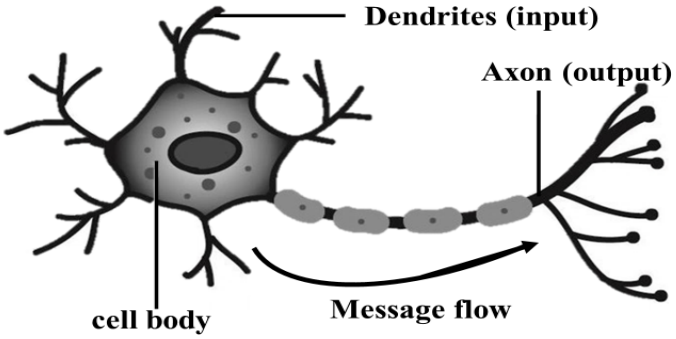


Fig. 2-1. Neural structure

Figure 2-2 is the simplest neural model, which is called Perceptron by adjusting the mathematical formula of weight w or threshold h to the artificial neural network model. The Perceptron, which has a single-layer and multi-layers for dealing with more complex problems [5, 6].

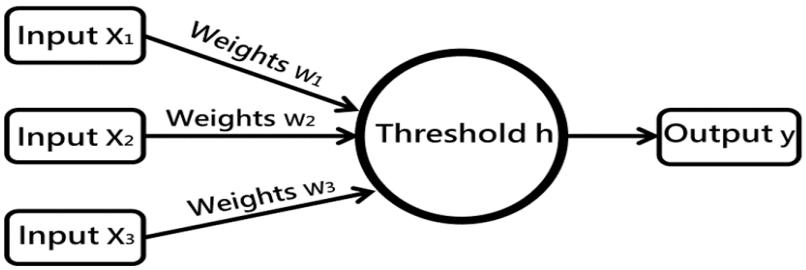


Fig. 2-2. Form of Neural Structure

2.2 Deep learning and CNN

The results of the histogram classification were good, but the histogram method could only be classified for a small number of letter hand shapes or gestures to produce good results. The effect of all 26 English glyph classifications was not good, used the method of Canny edge detection, which resulted in a good identification rate of 92.33%.

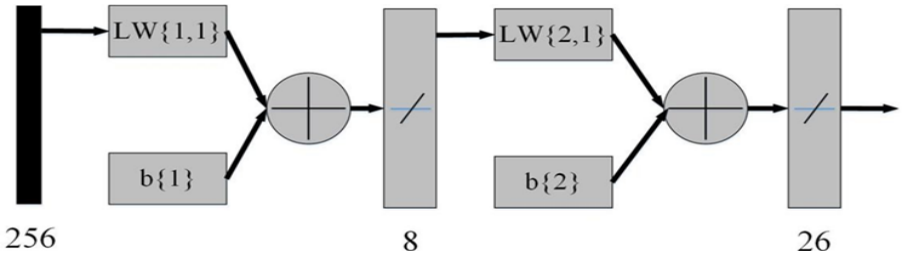


Fig. 2-3. Classification Network

According to Fleck and Forsyth's [7]1996 paper on ECCV (European Conference On Computer Vision), human skin at proper saturation is made up of red (blood) and yellow (melanin), and these feature information can be used in gesture tracking algorithms. Skin tone filtering is converted to the corresponding value with I, R_g and B_y at every point of RGB pixels, the formula is as follows

$$L(x) = 105 * log10(x + 1 + n)$$
$$I = L(G)$$
$$R_g = L(R) - L(G)$$
$$B_y = L(B) - \big(L(G) + L(R)\big)/2$$

Displayed equations are centered and set on a separate line.

$$x + y = z$$

I, R_g and B_y correspond to the color channels of green, red, and blue. Green is the best to represent image intensity cause red and blue have poor spatial resolution in some photographs. Constant 105 is to allow the output of the log function to fall between 0 and 254, n is a random noise value of 0 to 1, is to prevent the black area of the picture banding artifacts produced. Constant 1 is intended to prevent over-optimization in very dark colors. Using the log function conversion to calculate the excellent phase and saturation, the skin tone area can be marked with the following properties[8]:

$$110 \leqq Hue\ 150\ and\ 20 \leqq Saturation \leqq 60$$
$$130 \leqq Hue\ 170\ and\ 30 \leqq Saturation \leqq 130$$

Use the pre-trained CNN model to do edge and corner feature extraction at the bottom. With the ImageNet Large-Scale Visual Recognition Challenge (ILSVRC) 2014 training model, 22-layer deep network architecture in the ILSVRC has the first accuracy rate of 68.7% and 5th accuracy 88.9%. The decline 70% of rate to avoid overfitting with the CNN model also used the Stochastic Gradient Descent (SGD) and the Cross-Entropy based loss E's Softmax function, as Formula follow [9].

$$\mathrm{E} = -\frac{1}{N}\sum_{n=1}^{N}\log p(k|x_n)$$

At the last layer of CNN, a 1024-size feature extraction convolution layer was set and the size was reduced to 200 using Principal Component Analysis (PCA) formalization, Measured the Sign Language Collection Word Error Rate (WER) with RWTH-PHOENIX-Weather 2014 Multisigner and SIGNUM, which are based on the following formula[9]:

$$\mathrm{WER} = \frac{\#\text{deletions} + \#\text{insertions} + \#\text{substitutions}}{\#\text{number of reference observations}}$$

Although sign language in the sign language film has been evenly expressed sign language gestures, but the language of the oral expression has different length and arrangement. Sign language film translation into the discourse has sequence-to-sequence problems, the goal is to have a T-frame gesture film as a U-word and statement y's probability p, which can be represented by the following formula [10]:

$$\mathrm{p} = \frac{y}{x} = \frac{(y_1, y_2, \ldots, y_U)}{(x_1, x_2, \ldots, x_T)}$$

If the T >> U indicates that the gesture film frame is much larger than the translation word of oral, it is usually not aligned between sign language to the speaking word. Unlike other text translations, the input material used here is a film, which makes the model construction of RNN quite difficult, so combines the focus on Attention-based Encoder-Decoders CNNs network calculates the p-value. The full picture of the research targets of Camgoz et al. are shown of picture 2-4[10].

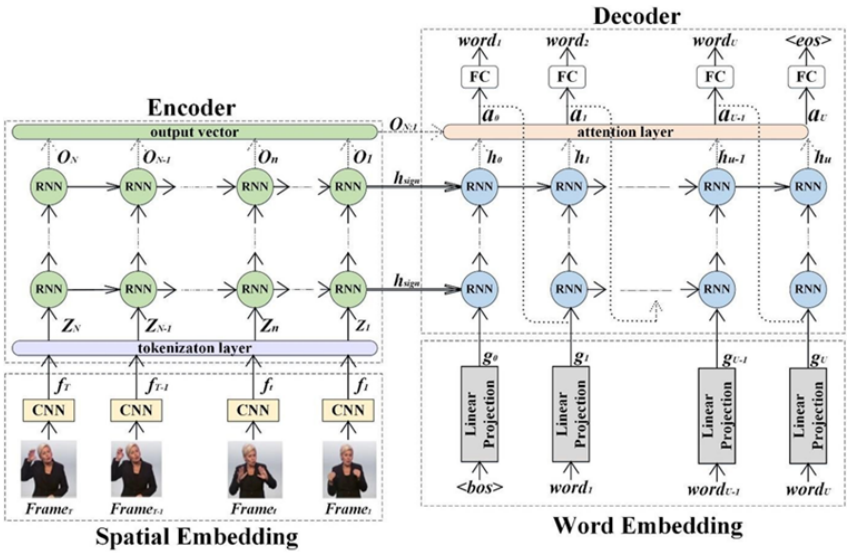


Fig. 2-4. Camgoz et al. Studied the full picture of SLT's translation of sign language film into a spoken language target.

3 Materials and methods

3.1 Establish Gesture image database

Contains shooting gesture images, pre-processing of image materials, and creating gesture image database. After shooting 37 sign language Mandarin Phonetic Symbols images, tearing the image into still gesture images. Remove images that include blacking edges, large arm's proportions and excessive angle rotation. Establish a database of gesture images of about 250 to 300 gestures per gesture.

3.2 CNN model setup with parameters

The CNN model used: LeNet and SmallVGGNet have layer-number architectures, described with the relevant parameters at training time, including the categorical cross entropy, Adam Optimization Algorithm, Batch Size and Epoch, ReLU Activation function, Softmax Activation Function and Learning Rate Function.

3.3 Image data entry and data augmentation kit

Use Scikit-Learn's Import Train_Test_Split module to differentiate gesture image data into training data and validation data. The input gesture image use data augmentation kit to flip, rotate, crop, zoom in and out and so on, to promote or replace the training data with specified condition. Make training data more variable, reduce over adaptations caused by re-training of original image and improve learning outcomes.

3.4 Establish Gesture image database

Because Mandarin Phonetic Symbols' database refers to the related gesture database research is less, so this research is self-established. Chinese sign language Mandarin Phonetic Symbols gestures such as Table 3-1. This study used a smartphone as a shooting tool to record images of each gesture, which are mp4 format files with image sizes of 1920x1080 pixels and 29.99 fps (Frames Per Second).

Table 3-1. Spoken and sign language of Mandarin Phonetic Symbols gesture

NO	Phonetic Mark	Signal	NO	Phonetic Mark	Signal	NO	Phonetic Mark	Signal
0	ㄅ		11	ㄐ		24	ㄦ	
1	ㄆ		12	ㄑ		25	ㄚ	
2	ㄇ		13	ㄒ		26	ㄛ	
3	ㄈ		14	ㄓ		27	ㄜ	
4	ㄉ		15	ㄔ		28	ㄝ	
5	ㄊ		16	ㄕ		29	ㄞ	
6	ㄋ		17	ㄖ		30	ㄟ	
7	ㄌ		18	ㄗ		31	ㄠ	
8	ㄍ		19	ㄘ		32	ㄡ	
9	ㄎ		20	ㄙ		33	ㄢ	
10	ㄏ		21	ㄧ		34	ㄣ	
			22	ㄨ		35	ㄤ	
			23	ㄩ		36	ㄥ	

Fig. 3-1. The captured gesture image

Fig. 3-2. Take gesture images from the camera

The recorded image material is then output to 720x480 images at 25 fps, and since the number of Python syntax starts at 0, the number 0 to 36 corresponding to 37 kinds of Mandarin Phonetic Symbols which are used as the folder name. After removing images with too large proportion of arms, or too much angle rotation, the converted images

were taken with about 250 to 300 images each for a total of 10,598 images. Images with corresponding of Mandarin Phonetic Symbols as in Table 3-2.

Table 3-2 The database number and number of image sheets corresponding to the Mandarin Phonetic Symbols.

Table 3-2. The database number and number of image sheets corresponding to the Mandarin Phonetic Symbols

NO	Mandarin Phonetic Symbols	Images	NO	Mandarin Phonetic Symbols	Images	NO	Mandarin Phonetic Symbols	Images
0	ㄅ	292	11	ㄐ	300	24	ㄦ	300
1	ㄆ	250	12	ㄑ	260	25	ㄚ	280
2	ㄇ	260	13	ㄒ	260	26	ㄛ	280
3	ㄈ	290	14	ㄓ	300	27	ㄜ	300
4	ㄉ	290	15	ㄔ	300	28	ㄝ	290
5	ㄊ	280	16	ㄕ	300	29	ㄞ	280
6	ㄋ	270	17	ㄖ	260	30	ㄟ	290
7	ㄌ	280	18	ㄗ	290	31	ㄠ	300
8	ㄍ	300	19	ㄘ	300	32	ㄡ	290
9	ㄎ	300	20	ㄙ	290	33	ㄢ	288
10	ㄏ	280	21	ㄧ	280	34	ㄣ	296
			22	ㄨ	280	35	ㄤ	297
			23	ㄩ	300	36	ㄥ	295

3.5 Model parameters

1. Categorical cross entropy

In the field of machine learning, the error that compares the result to the output data is called Lose. The function that represents the degree of machine learning whose error is called Loss Function. Loss function will give back its value to the neural network. It updates the parameters and then continue to train to reduce the loss and improve the learning of the neural network. Loss functions are used Cross Entropy Error, which is how much difference there is between group A and group B. Its formula represents, as in the [11]:

y is the output data that the machine learning algorithm determines the result, and t is the correctly answered data, k is a category. Cross Entropy Error is from the probability of containing a variety of labels to extract the probability of the correct answer. The probability of taking the number of pluses, and then its value feedback class neural network to do parameter correction to improve learning rate. Feedback correction of model parameters is used during model training to find the most appropriate combination of parameters, Activation Function refers to the input is output after activation. ReLU (Rectified Linear Unit), Activation Function is when the input is greater than 0, the output is equal to the input; If the input is less than 0, the output is equal to 0, as shown in Figure 3-3.

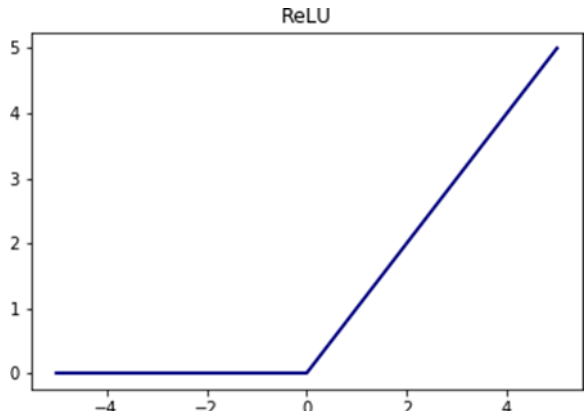


Fig. 3-3. ReLU Activation Function

2.Softmax Activation Function

The value is between 0 and 1, and its purpose is to increase the gap with multiple input values in a formula. The highest input value to make it approximate to 1 as output, and other input values to approximate its output to 0. It can be multi-classified output, the output of approximately 1 is its classification result. The formula represents as in the s[11]:

$$\sigma(y)_j = \frac{e^{y_j}}{\sum_{k=1}^{K} e^{y_k}}$$

y_j for an input value, K represents the number of output categories, $\sigma(y)_j$ which represents the natural logarithm of y_j values taking K natural logarithm plus. The range of 0 to 1 with the total number of the possibility, and if one input y_j is greater than the other y values, the output y_j is approximately 1, and the other output is approximately 0, thus representing the possibility of classified output.

3. Learning Rate Function : ReduceLROnPlateau

The goal of Gradient Descent method is to optimize to the lowest point. When the model is not ideal, add Learning Rate Function (LR Function): ReduceLROnPlateau to do the training learning rate adjustment, setting the val_acc to halve the learning Rate

when three consecutive Epochs have not risen. The minimum learning rate limited to 0.00001 to optimize the training parameters.

3.6 CNN model building

1. LeNet

Use (96, 96, 3) as input layer data, i.e. 96x96x3 add convolution layer, activation function ReLU, pooling layer for the first feature extraction then add convolution layer. Activation function ReLU, pooling layer for the second feature extraction and then establish a flat layer with 500 neurons and activation function ReLU. Finally establish 37 neurons and the Softmax activation function as a classification to identify of the output, 37 neurons are used as the output layer due to 37 gestures, as shown in Figure 3-4.

```
_________________________________________________________________
Layer (type)                 Output Shape              Param #
=================================================================
conv2d_1 (Conv2D)            (None, 96, 96, 20)        1520
_________________________________________________________________
activation_1 (Activation)    (None, 96, 96, 20)        0
_________________________________________________________________
max_pooling2d_1 (MaxPooling2 (None, 48, 48, 20)        0
_________________________________________________________________
conv2d_2 (Conv2D)            (None, 48, 48, 50)        25050
_________________________________________________________________
activation_2 (Activation)    (None, 48, 48, 50)        0
_________________________________________________________________
max_pooling2d_2 (MaxPooling2 (None, 24, 24, 50)        0
_________________________________________________________________
flatten_1 (Flatten)          (None, 28800)             0
_________________________________________________________________
dense_1 (Dense)              (None, 500)               14400500
_________________________________________________________________
activation_3 (Activation)    (None, 500)               0
_________________________________________________________________
dense_2 (Dense)              (None, 37)                18537
_________________________________________________________________
activation_4 (Activation)    (None, 37)                0
=================================================================
Total params: 14,445,607
Trainable params: 14,445,607
Non-trainable params: 0
_________________________________________________________________
None
```

Fig. 3-4. LeNet model structuring with its number of parameters

A total of 14,445,607 parameters are used for learning to zoom scalable vector graphics (SVG) which format is displayed as Figure 3-5.

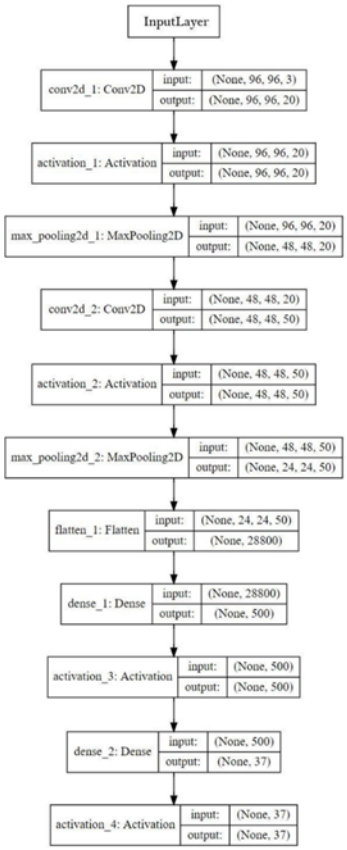
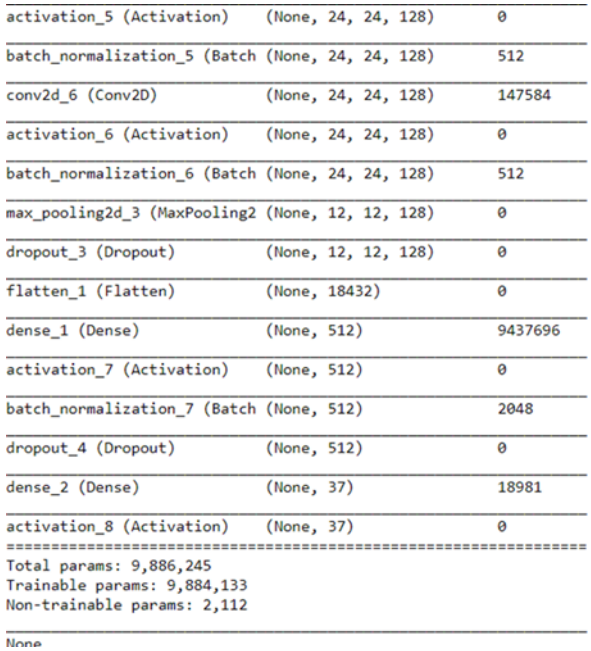


Fig. 3-5. LeNet model in SVG format

The LeNet model can be clearly seen in the SVG format. Entered by neurons in the input layer, extracted from the convolution layer, pooling layer features and finally to the flat layer or fully connected layer for the classification to identify 37outputs.

2. SmallVGGNet

In this study Pyimagesearch's [12]Small VGGNet model, based on the VGG model, but with a shallow number of layers, was another training model for this study. Adding a layer of batch normalization, which the effect is batch data outed feature values are standardized and then entered to the next layer. To ensure that the feature values are fixed in order to speed up the training, prevent the gradient from disappearing. Improve the accuracy of the training, SmallVGGNet's model construction and parameters are shown in Figure 3-12, and the SVG format is shown in Figure 3-6.



activation_5 (Activation)	(None, 24, 24, 128)	0
batch_normalization_5 (Batch	(None, 24, 24, 128)	512
conv2d_6 (Conv2D)	(None, 24, 24, 128)	147584
activation_6 (Activation)	(None, 24, 24, 128)	0
batch_normalization_6 (Batch	(None, 24, 24, 128)	512
max_pooling2d_3 (MaxPooling2	(None, 12, 12, 128)	0
dropout_3 (Dropout)	(None, 12, 12, 128)	0
flatten_1 (Flatten)	(None, 18432)	0
dense_1 (Dense)	(None, 512)	9437696
activation_7 (Activation)	(None, 512)	0
batch_normalization_7 (Batch	(None, 512)	2048
dropout_4 (Dropout)	(None, 512)	0
dense_2 (Dense)	(None, 37)	18981
activation_8 (Activation)	(None, 37)	0

Total params: 9,886,245
Trainable params: 9,884,133
Non-trainable params: 2,112

None

Fig. 3-6. SmallVGGNet model construction with its number of parameters

3.7 Read the input data

The training of the Deep Learning System that is made with Images and Real Values (Labels) as the system learning data on the basic identification Pyimagesearch [13] of CNN. System is modified and adjusted to meet the needs of this study, and the deep learning and identification of the sign language with Mandarin Phonetic Symbols is established. The results of its classification experiments are obtained.

Reading the input data, the picture taken by the database is processed in 720 pixels wide, 402 pixels high, and color RGB mode, which can be represented as 720x402x3, first rewriting its picture size to 96x96x3. Image read-in data then get its classification of the digital folder name, as the label of this kind of image picture, read it in sequence into the data matrix. Completed the deep learning system input layer required image values and real value input data.

Using the Sklearn Suite Import Train_Test_Split module to distinguish between 80% training data and 20% validation data. The number of gesture image data in this study was 10,598, which was divided into 8,478 pieces of training material and 2,120 validation data using the ImportTrain_Test_Split module.

4 Data Analysis and Discussion

Self-established gesture image database, input it into the completed LeNet and SmallVGGNet models. Use data segmentation and data amplification kits for gesture image recognition training, adjust its Batch Size parameters one by one.

Observe the training accuracy and error rate of each training results. Enter other background gesture image data into the model completed by the model to evaluate its gesture recognition rate.

4.1 LeNet Training Results

This study trained the equipment specifications as CPU 3.5GHz, 8G RAM memory, BS represents the batch Size batch training quantity size, and uses Epoch as 20 times, Categorical cross entropy, Adam optimization algorithm, 37 classification outputs, and Softmax activation functions for the training of deep learning convolution network models.

First, Batch Size=64

Enter 96 pixels x96 pixels x3 (3 for RGB color images), first set Bathc Size to batch 64 samples to learn, to see how well they learn, parameters such as Table 4-1.

Table 4-1. LeNet enters 96x96x3 training parameters

Model	LeNet
Input	96 x 96 x 3
Batch size	64
Epochs	20
Loss	Categorical crossentropy
Optimizer	Adam
Output	37
Activation	Softmax

The Epoch recognition rate for the first training session was 3.2%, after which the accuracy of each training increased gradually, as shown in Figure 4-1. The accuracy of the training data (Accuratey, train_acc) and the error rate (Loss, train_loss), as well as the accuracy of (val_acc) and error rate (val_loss) of the validation data, can be expressed in Figure 4-2, the 20th Epoch training data accuracy acc is 92.75 percent, the

error rate is 22.47%; The accuracy rate of the validation data was 99.15% val_acc and the error rate of the verification data was val_loss 3.94%.

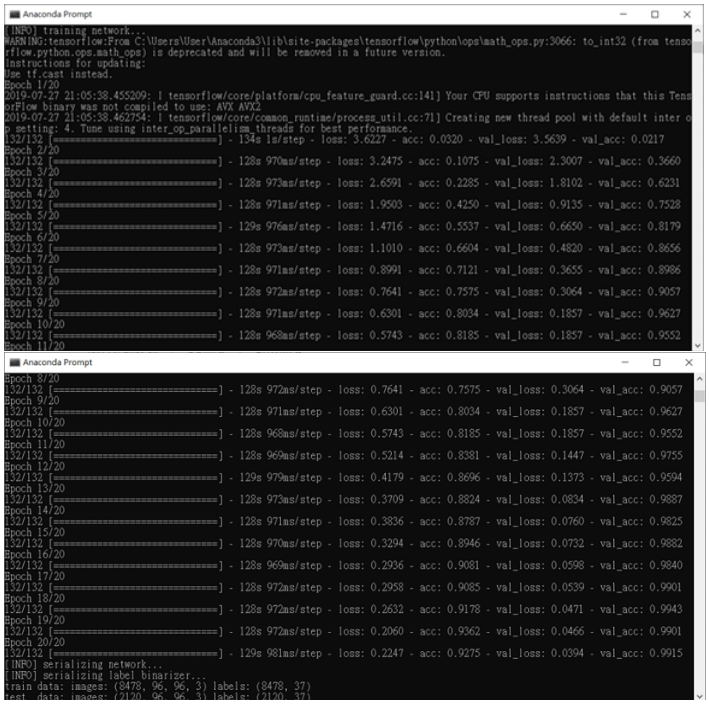


Fig. 4-1. LeNet BS=64 training process

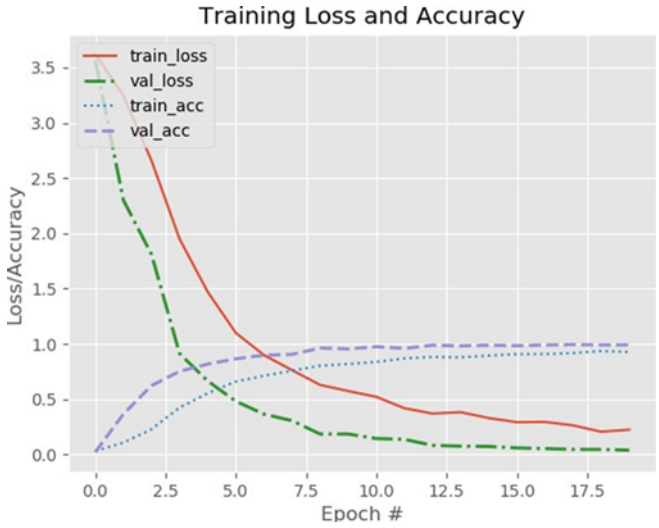


Fig. 4-2. LeNet BS=64 training diagram

In the confusion matrix results, the largest misjudgments are predicting the gesture of label 18 (ㄗ) as label 5 (ㄊ), and predicting the gesture of label 23 (ㄩ) as label 12 (ㄑ), as shown in Figure 4-3. There are six pictures that are incorrectly predicted. It is speculated that the prediction is misjudged due to the high similarity of gestures.

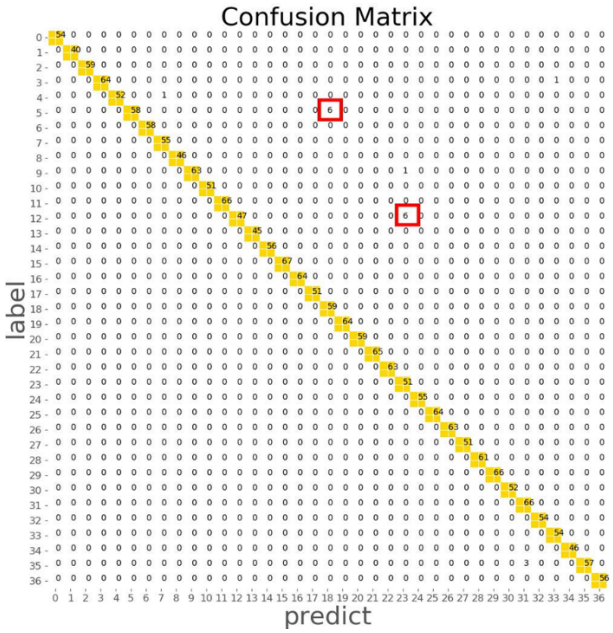


Fig. 4-3. Confusion matrix of training result of LeNet BS=64

Six, LeNet model classification test

The training results of LeNet models are shown in Table 4-2. The best performance is with Batch Size of 128. After adding the learning rate reduction function, the training accuracy is increased by 1.15% to 96.64%, and the verification accuracy is improved. 0.38% reaches 99.67%, as number five in Table 4-2. The result is a better model trained in LeNet in this study.

Table 4-2. LeNet model training results

Model: LeNet				Epochs=20 training results (unit :%)				Training time (Unit: minutes)
NO.	Input (RGB)	Batch Size	LR func-tion	train_acc	train_loss	val_acc	val_loss	
1	96 x 96	64	None	92.75	22.47	99.15	3.94	43
2	96 x 96	96	None	93.91	19.5	97.59	6.72	41
3	96 x 96	128	None	95.49	14.43	99.29	3.03	41
4	96 x 96	256	None	93.55	22.7	99.2	5.43	39
5	96 x 96	128	Yes	96.64	11.57	99.67	1.79	43

SmallVGGNet training results of various models, with BS=128 and adding LR Function as the best training model, as shown in number 4 in Table 4-3, use this model to test images and pictures.

Table 4-3. SmallVGGNet model training results

Model: SmallVGGNet				Epochs=20 training results (unit)				Train-ing time (Unit: minutes)
NO.	Input (RGB)	Batch Size	LR func-tion	train_acc	train_loss	val_acc	val_loss	
1	96 x 96	64	None	99.13	2.82	12.17	1080.96	97
2	96 x 96	64	Yes	99.8	0.77	94.01	15.18	92
3	96 x 96	96	Yes	99.93	0.71	88.87	28.03	91
4	96 x 96	128	Yes	99.93	0.74	98.02	4.54	92
5	96 x 96	256	Yes	99.66	3.05	75.42	91.63	88

The experimental results of SmallVGGNet have the highest recognition rate of 29.73% in the recognition results of monochrome background and increased background complexity because its non-input training gesture images, unlike training data images, have a high recognition rate.

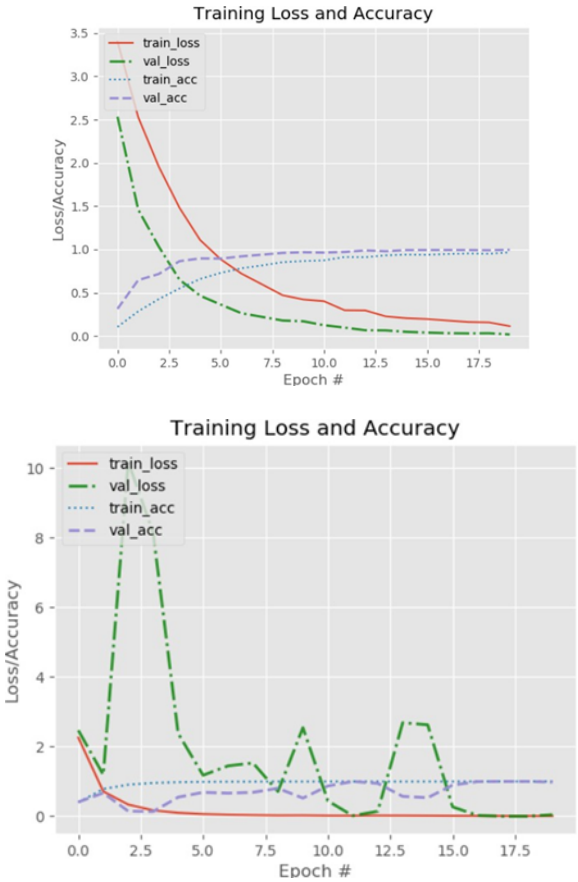


Fig. 4-4. LeNet and SmallVGGNet model training process

In terms of test data performance, LeNet training data image recognition rate is 100%, SmallVGGNet recognition error one image, recognition rate 97.3%. In terms of the recognition rate of right-hand and left-hand on monochrome background and complex background, LeNet has the highest 51.35% and the lowest 13.51%, while SmallVGGNet has the highest 29.73% and the lowest 13.51%, as shown in Table 4-4.

Table 4-4. LeNet and SmallVGGNet model recognition rate

Model	Epochs=20 training results (unit)				Model recognition rate (Unit: %)			
	train_acc	train_loss	val_acc	val_loss	Training data	Monochrome background	Complex background (Right hand)	Complex background (Left hand)
LeNet	96.64	11.57	99.67	1.79	100	13.51	51.35	40.54
SmallVGGNet	99.93	0.74	98.02	4.54	97.3	13.51	29.73	18.92

The recognition results for monochrome background and increased background complexity have a recognition rate of 51.35%. The recognition rate is affected by the diversification of the training data, and different gesture angles, different gestures but overlapping gestures at a certain angle and different light and dark conditions will affect the recognition rate of the model.

5 Conclusion

Shooting 37 gesture images of Mandarin sign language phonetic symbols, each of which includes three combinations of front, clockwise and counterclockwise rotation, combined with distance and near shooting. The recorded image data is output as a picture at 25 fps, with the numbers 0~36 corresponding to 37 phonetic symbols, about 250-300 pictures of each gesture are obtained, a total of 10,598 pictures, and a database of 37 right-hand gestures of Chinese sign language Mandarin Phonetic Symbols with a single background is established.

The completed Chinese sign language Mandarin phonetic symbols and gesture database was established and trained with two different convolution neural network models, LeNet and SmallVGGNet. Set the gesture database image picture to be divided into 80% training data and 20% verification data, and use input image size 96x96x3, loss function cross entropy error, Adam optimization algorithm, ReLU activation function and Softmax activation function, data amplification kit, Epoch Experiment with 20 times as the training parameters of the convolution neural network model. Adjust the Batch Size during training and add the Learning Rate Function to reduce the learning rate at the right time. Finally, compare the results and get the Batch Size as 128, enabling the learning rate reduction function is the best result of training, and the accuracy of training data and verification data are both above 96%.

In addition to the diversification of data with different gesture angles, gestures but overlapping gestures at a certain angle or gesture image pictures with different light and dark conditions, the selection of parameters will affect the recognition rate of the model.

Use the self-created gesture image database as a deep learning convolution neural network model training, use a single background image and image training model to test its recognition rate results. Future research can add other complex backgrounds and different angles. The gesture images of light and dark and age or skin color can enrich the database of gestures will be closer to real life scenes.

The recognition in this research is static image recognition. Future research suggests that you can do further real-time dynamic gesture recognition, and convert the dynamic recognition of sign language gestures into spoken sentences, or develop and design the relevant interface of the sign language recognition real-time comparison system will be more helpful for gesture recognition.

References

[1] (2020). *Discussion on the types of physical gestures communication*. Available: http://mailstu.nutn.edu.tw/sign/2005/mastergroup7.htm

[2] H.-Y. Chang, "The role of syntax in hearing-impaired students reading written language," Graduate Institute of Learning and Instruction, National Central University, 2008.

[3] H. T. Liu, "Taiwanese Sign Language Receptive Skill Test for Children as An Evaluation for Elementary Level Students at Deaf Schools," *Bulletin of Special Education,* vol. 44, no. 1, pp. pp90-116, 2019.

[4] L. K. Hansen and P. Salamon, "Neural network ensembles," *IEEE transactions on pattern analysis and machine intelligence,* vol. 12, no. 10, pp. 993-1001, 1990.

[5] S. Lawrence, C. L. Giles, A. C. Tsoi, and A. D. Back, "Face recognition: A convolutional neural-network approach," *IEEE transactions on neural networks,* vol. 8, no. 1, pp. 98-113, 1997.

[6] X. Zhang, X. Zhou, M. Lin, and J. Sun, "Shufflenet: An extremely efficient convolutional neural network for mobile devices," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2018, pp. 6848-6856.

[7] M. M. Fleck, D. A. Forsyth, and C. Bregler, "Finding naked people," in *European conference on computer vision*, 1996, pp. 593-602: Springer.

[8] T.-N. Nguyen, H.-H. Huynh, and J. Meunier, "Static hand gesture recognition using artificial neural network," *Journal of Image and Graphics,* vol. 1, no. 1, pp. 34-38, 2013.

[9] O. Koller, H. Ney, and R. Bowden, "Deep hand: How to train a cnn on 1 million hand images when your data is continuous and weakly labelled," in *Proceedings of the IEEE conference on computer vision and pattern*

recognition, 2016, pp. 3793-3802.

[10] N. Cihan Camgoz, S. Hadfield, O. Koller, H. Ney, and R. Bowden, "Neural sign language translation," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2018, pp. 7784-7793.

[11] T.-C. Lee, "Filtering Spam Mails Using Convolutional Neural Networks and Parallel Computing," The Department of Computer Science, Tunghai University, 2016.

[12] P. Adrian Rosebrock. (2018). *Keras Tutorial: How to Get Started With Keras, Deep Learning, and Python*. Available: https://www.pyimagesearch.com/2018/09/10/keras-tutorial-how-to-get-started-with-keras-deep-learning-and-python/

[13] P. Adrian Rosebrock. (2018). *Keras and Convolutional Neural Networks (CNNs)*. Available: https://www.pyimagesearch.com/2018/04/16/keras-and-convolutional-neural-networks-cnns/

Research on Medical Information Exchange Mechanism of Blockchain Combined with Health Passbook

Ying-Che Huang [1], Yu-Yu Yen [2], Jui-Hung Kao[3], Horng-Twu Liaw [3], Wei-Chen Wu[4]

[1 *] Center for General Education, Shih Hsin University, Taipei, Taiwan

[2 *] Department of Finance, Shih Hsin University, Taipei, Taiwan

[3 *] Department of Information Management , Shih Hsin University, Taipei, Taiwan

[4 *] Department of Finance, Feng Chia University, Taipei, Taiwan, Taichung

Abstract. With the gradual development of the global information technology and mobile communication networks, the integration of the medical system with information and network technologies has become an inevitable trend. All countries are also accelerating their efforts to replace traditional paper books with Electronic Medical Record (EMR). Medical records and electronic medical record implementation can save costs annually, increase economic efficiency by improving the efficiency and safety of medical systems, and accelerate innovation in the medical industry. The implementation of EHR will undoubtedly bring many conveniences, stakeholders who are concerned with EHR at this stage are still not satisfied with the privacy disclosure and information security regarding the possible implementation of electronic medical records. And the information between them still lacks good interoperability.

Ministry of Health and Welfare promote health passbooks in recent years which allows patients to choose their own medical information by themselves. How to securely and efficiently integrate information from different platform systems and make it easier for people to get complete personal information thus allows individuals and organizations to securely access and exchange is the key requirement for thinking about smart, precise and health care.

Aims at designing a personal health passbook exchange mechanism with high safety strength specifications to protect the special funds, two-factor identity card technology with smart cards, an authorization mechanism for applying real-name block chain system and digital signature to medical records, and the overall system integrate the existing electronic medical record signing mechanism with the overall personal health passbook exchange mechanism to identify users from the very beginning as empowered to grant users the right to personal health passbook exchange. Such medical record exchange is both safe and trustworthy and easy to use. According to the development vision of the Ministry of Health and Welfare, electronic medical records are stored through mobile carriers(health passbooks) to achieve a

win-win situation of promoting universal health and medical progress from an information security point of view.

Keywords: Real-name, Blockchain, Health Passbook, Electronic Medical Record Exchanges

1 Introduction

Today's personalized health information needs have changed the way we use medical system information. Establishing a health care data-specific big data network and integrating personal health information is a nationally driven goal and an inevitable trend in global IT advancement and action with the gradual maturity of the era of communication networks, countries have also accelerated the replacement of traditional paper based medical record with Electronic Medical Record (EMR). The implementation of electronic medical records can save costs every year and increase economic efficiency by improving the efficiency and safety of medical systems.

And accelerate innovation in the medical industry[1-3]. The implementation of electronic medical records will undoubtedly bring a lot of convenience, but at present, the stakeholders related to electronic medical records, whether they are patients, medical institutions, medical staff or information experts, are doubts about safety caused by the implementation of electronic medical records may still disclose the patient's privacy leakage and information. [4].

2 Materials and methods

2.1 Electronic medical record exchange and preservation mechanism based on blockchain

Based on the development of the electronic medical records and under the premise of the compliance procedures guarantee with sensitive personal information for patients (electronic medical records), in order to ensure that the right to use electronic medical records is in the holder of medical records, new security maintenance measures for expanding the electronic medical record exchange authorization are proposed. The electronic medical record exchange and preservation mechanism based on blockchain, in addition to protecting the confidentiality, integrity and access security of the transmission data, ensuring the non-repudiation exchange of electronic medical records between the two parties after the authorization is completed. In order to meet the needs of practical operations, mainly using the technology based on blockchain and group signature to applied to the hospital's medical record access control process. The following briefly introduces the relevant technologies of this mechanism:

Application of smart card identity identification mechanism: Since the information about the user on the network is represented by data, in order to ensure that the user's

actual identity corresponds to the data identity on the network, there is a need to use identity identification authentication. There are three main ways to identify user authentication. [2, 5, 6]：

(1). Identify your identity (Something You Know): the account number, password, picture and order set by the user.
(2). Identify your identity by your owner (Something You Have): Binding identity RFID card, smart card, etc.
(3). Identify your identity according to your characteristics: such as the fingerprint, voiceprint and profile of the user.

The above identity identification authentication methods are most widely used for account number and password. In identity identification authentication the specific account password must be memorized; while the smart card is used as the identity identification authentication the account password is no need to remember account password, the risk of fraudulent use duplication is existing when the card is lost. Although the use of fingerprints, voiceprints and other characteristics for identity identification authentication has the highest safety factor, the cost is also higher than the other two ways. In order to avoid the risk of identifying the identity through a single factor, there is also a mechanism to use two dual identity authentication methods to enhance system security, called two-factor authentication.

Digital currency Bitcoin has become an important currency for today's investment and private transactions. The concept is derived from the blockchain theory proposed by Haber and Stornetta in 1991 [7], using timestamps and hash values to form a link, each block must contain its own timestamp hash value and the timestamp hash value of the previous block (as shown in Figure 1), which forms the base of the Bitcoin-Ledger.

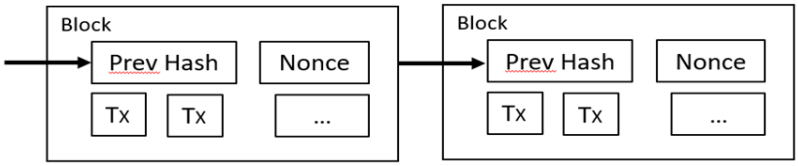


Fig. 1: Schematic diagram of blockchain architecture for Bitcoin

When Bitcoin is carrying out actual transaction, it needs to be able to ensure that the recipient takes ownership, and the next recipient must be able to verify the ownership of the previous owner. To achieve the above objectives, the bitcoin transaction model diagram is shown in Figure 3. The users participating in the transaction must hold their own Public Key and Private Key, and the block design the signature of the owner, the recipient can be verified by the owner's public key, and after the transaction is completed the old Bitcoin is recycled. The recipient obtains a new Bitcoin and then signature to complete the transfer of ownership.

The health passbook storage method of this will adopt the concept of blockchain. Except that each health passbook record becomes a block, in order to achieve its own health passbook that can be mastered function by oneself, each health passbook record

writing will become a transaction; and in order to maintain the transaction record of each health passbook to achieve the integrity and non-repudiation of the transaction record, each block of the health passbook block chain will be signed by medical personnel, and because the medical staff is plural, the mechanism that is suitable for adopting Digital Multisignature Scheme.

2.2 Blockchain transaction agreement

Scholars Harn and Kresler proposed a multi-signature mechanism in 1989，This mechanism is to solve the situation where the same document requires multiple people to sign together. The multiple signatures of the same document are the definition of Multisignature. The two scholars completed the multi-signature mechanism by signing each member's private key one by one in the RSA cryptosystem; in past studies, the Threshold Signature and the group signature are the implementation of the RSA signature by the participants using the private key[6, 8].

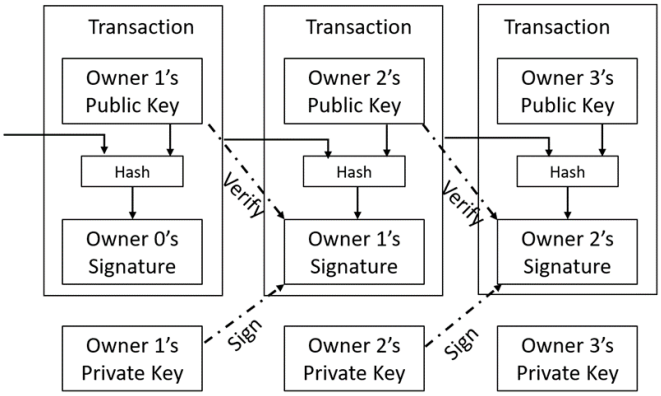


Fig. 2: Schematic diagram of Bitcoin blockchain transaction agreement [9]

1. Initial phase of the system
   (1) Generation of the trust center is required RSA related parameters and public key e_o and private key d_o.
   (2) Publish the public parameters n_o、e_o、$g_o$and save the private parameters p_o、q_o、d_o
2. Key generation phase
   (1) Each signer F_l participating in the signature generates the relevant RSA parameters:
   public key e_i, private key $d_i$and n_i
   (2) The member of the group G participating in the signature is $G = \{F_1, F_2, \ldots F_t\}$。

3. Signature stage
   (1) Suppose that the message m is to be signed, and the signatures are sequentially signed by the signersF_1, F_2,…F_t , and F_i calculates s_i= m^{di}mod n and transmits it to the signer F_{i+1}
   (2) Each signer participating in group G continues the calculation of the preceding item, that the signature is transmitted to the next signer until F_{t-1} ，F_i expression of the calculation is also $s_i = s_{i-1}^{d_i} \bmod n_i$ 。
   (3) The signature of the final signature holder F_twill complete the group signature of the entire group, that is, the group signature s $= s_t = s_{t-1}^{d_t} \bmod n_t$ ， after the signature, the examiner's public key e_v is encrypted, that is, $S = s_t^{e_v}$, and is transmitted to the examiner for confirmation.
4. Verification phase
   (1) The inspector needs to decrypt with his own private key d_v, and then use each the signere F_i of the public signature(n_i, e_i) to calculate signature to reply to the message, according to the number of signatures to calculate $s_{i-1} = s_i^{e_i} \bmod n_i$,where $t \geq i \geq 1$ 。
   (2) If $s_0 = m$ the message completion check, the number of multiple signatures is established, the message m is acceptable.

In addition to the blockchain storage concept, the health passbook will be signed and protected through the digital multi-signature of the public key encryption system, and used for transmission and preservation. In addition to ensuring that the health passbook can be mastered on its own, it also ensures the integrity and non-repudiation of health passbooks.

3 Methods

The current electronic medical record storage methods are centralized control. The electronic medical record storage and operation process are all passed through the hospital's medical information system, leaving only operational records and failing to meet the basic requirements of information security management. Once the system described above finds a data leakage problem, it will be difficult to lock down the user who leaked the electronic medical record, and the centralized control method does not grant authority to the electronic medical record holder who should hold all the rights of the electronic medical record; failure to include signatures from all health care providers. Therefore, which will design smart card identification, blockchain transaction process and multi-signature storage method, and refer to relevant domestic and foreign literature, design various simulation situations for actual medical conditions, in order to design better personal health passbook.

3.1 Health passbook design access and disclosure process

The current electronic medical record storage methods are centralized control. The electronic medical record storage and operation process are all passed through the hospital's medical information system, leaving only operational records and failing to meet the basic requirements of information security management. Once the system described above finds a data leakage problem, it will be difficult to lock down the user who leaked the electronic medical record, and the centralized control method does not grant authority to the electronic medical record holder who should hold all the rights of the electronic medical record; failure to include signatures from all health care providers. Therefore, this will design smart card identification, blockchain transaction process and multi-signature storage method, and refer to relevant domestic and foreign literature, design various simulation situations for actual medical conditions, in order to design better personal health passbook.

To design access and open processes for health passbooks. The storage structure and process are divided into the following sections:

1. Health passbook storage structure: Introduce the healthy passbook storage structure set by the blockchain.
2. Initial stage of the system: The multi-signature of the operation requires parameters and behaviors that are processed at the initial stage of the system, that is, the parameters required by the third-party trust center to calculate the public key cryptosystem.
3. Personnel registration stage: The role level involved in needs to be registered with a third-party trust center to obtain the parameters required by the public key cryptosystem to enable the system to operate.
4. Health passbook writing mechanism: After the visit, the user directly writes a record to the user's health passbook according to the storage structure of the blockchain. To ensure the integrity and non-repudiation of the stored data, this mechanism will be signcrypt the data through a multi-signature.
5. Health Passbook Access Mechanism: The act of accessing the user's health passbook with the consent of the user, to ensure user consent and to handle the signcryption information mechanism.
6. Open the Health Passbook Storage Structure: In order to contribute to the owner's contribution to the removal of the identified medical materials for the purpose of medical content sharing and research, which is based on the public health passbook storage structure set up by the blockchain public account.
7. Health Passbook Disclosure Mechanism: For the smooth development of medical related research, limited data sharing is necessary. Therefore, also proposes a public account book structure for health passbooks for the research units to maintain and synchronize with each other.

The following describes the roles that will be involved in, the mechanism’s storage architecture and system flow: the participating actors include the attending physician, the emergency physician, the general medical staff and the health passbook’s user; the storage structure of the mechanism has a healthy passbook storage structure. And the public health passbook storage structure; Mechanism system process division system

initial stage, health passbook writing stage, health passbook access stage and health passbook open stage. In the above-mentioned system stage, how to effectively use the blockchain's private account book and the public account book and the multi-signature signcryption mechanism to achieve the purpose of the research - the user masters his own health passbook, is both private and effectively disclose healthy passbooks [10].

The symbol table to be used by this mechanism is shown in Table 1:

Table 1. Symbol Table of the Mechanism

Symbol	Description
D_i	Medical staff i
DID_i	Medical staff i's account
G	Medical staff group
U_i	User i
UID_i	User i's account
P_{ID}	Health insurance card information
$\oplus$	XOR operation
$h(.)$	Hash function

3.2 Blockchain storage architecture and process description

Various medical institutions and people exchange with the personal health passbook at exchange center, they will first authenticate with the personal health passbook exchange center, complete the establishment of a safe passage for safe registration and related Parameter exchange. After the registration of the personal health passbook exchange center is completed, the relevant parameters of the multi-signature can be obtained for multi-signature and identity identification. The electronic medical record of the holder of the personal health passbook is stored in its own mobile device, and can be synchronized to the personal health passbook exchange center by means of a decentralized database through the relevant mechanism of the blockchain. However, under normal circumstances, the holder is personal health passbook holders themselves [7, 11].

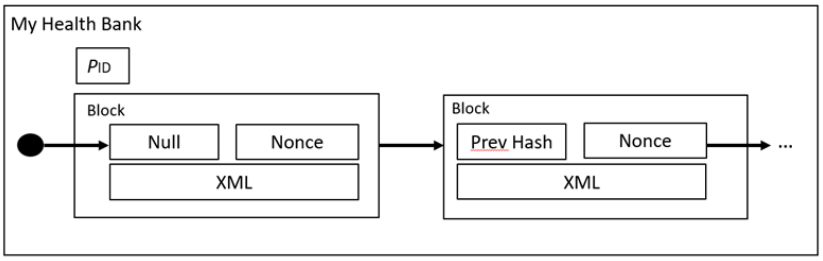


Fig. 3: Schematic diagram of the blockchain architecture of the health passbook

1. Initial stage of the system: Because uses the public key cryptosystem for digital multi-signature, in the initial stage of the system parameters related to the published system will be generated, and the system will operate smoothly. The system must be operated by a third-party trust center. The relevant parameters of the public key encryption system are as follows:
   (1) The third-party trust center takes two large prime numbers p and$q$ and calculates$N = pq$
   (2) The third-party trust center then selects the generator g from the multiplication group Z_n^*.
   (3) The third-party trust center selects a set of key pairs(e, d) , where e is the public key and d is the private key, and needs to satisfy $e \times d \equiv 1\ (mod\ \emptyset(N))$.
   (4) The third-party trust center discloses its own public keys e and N.
2. Personnel registration stage: The medical staff $D_i$who starting a job in a medical institution, the following steps must be taken to complete the registration process and obtain the parameters of the public key cryptosystem, and generate the precalculation content of the digital multi-signature.
   (1) The medical staff D_i registers with the third-party trust center. After registration, the medical personnel card and other vehicles will be issued. The registration information will be stored in the vehicle of the medical staff D_i , including the identification code 〖ID〗_i of the medical staff D_i, and the necessary disclosure parameters are handled by a third-party trust center; in reality, the card will be protected with a PIN after it is issued.
   (2) The medical staff $D_i$then registers with the third-party trust center on the basis of the medical staff team group G, with its own DID_i. Add a random integer $x_i \in Z_n^*$as the private key in the medical staff team group G, and then calculate $v_i = g^{-x_i}\ mod\ N$.
   (3) Transfer $DID_i$and v_i to the third-party trust center.
   The third-party trust center calculates the public key of the medical staff D_i in the medical staff team group G .
   $$y_i = (v_i^{h(v_i)} - ID_i)^d\ mod\ N$$
   (4) When the users of the medical staff team group $G = \{D_1, D_2, \dots, D_n\}$ are registered,

the group identification code GID is determined, and calculate the group public key $Y = ((\prod_{i=1}^{n} v_i^{h(v_i)}) - GID)^d \ mod \ N$, whose private key needs to be calculated jointly by the medical staff participating in the team $X = \sum_{i=1}^{n} x_i^{h(v_i)} \ (\mathrm{mod}\ \emptyset(N))$. That is, all the medical staff in the team can provide $v_i = g^{-x_i} \ mod \ N$ and then use the formula $Y^e + GID \equiv \prod_{i=1}^{n} v_i^{h(v_i)} \ mod \ N \equiv g^{-X} mod \ N$ to ensure the correctness of Y.

3. Health passbook writing mechanism: This stage describes the process of writing the medical passbook to the medical staff team; although the normal situation will first access the data and then write the behavior, in order to facilitate the introduction of the digital multi-signature mechanism, the Introduction to the process of writing mechanism first. When the user completes a new medical record, the medical staff team is required to perform a multi-signature on the medical record M for the purpose of ensuring the integrity and non-repudiation function, and writing the medical record in the user's health passbook blockchain, the steps are as follows:

(1) Medical staff D_i first selects a random number $w_i \in Z_n$ and calculates $r_i = g^{w_i} \ mod \ N$, and then broadcasts r_i to the medical staff in all teams($G = \{D_1, D_2, \dots, D_n\}$) for Group G.

(2) Upon completion of the broadcast exchange message, the agent $D_x$of the medical staff team will be selected, and the medicalD_i staff ($G = \{D_1, D_2, \dots, D_n\}$)

$$R = M \prod_{j=1}^{n} r_j \ mod \ N$$

$$K = h(M, R, GID)$$

$$\mathrm{s}_i = \mathrm{w}_i + \mathrm{x}_i \times h(\mathrm{v}_i) \times K$$

After the calculation, send s_i to the agent of the medical staff team D_x

(3) After the agent D_x of the medical staff team receives the signature (r_i, s_i)transmitted by the medical staff, all the (r_i, s_i)will be tested by the following mathematical formula:

$$g^{s_i} \ ? = (\mathrm{y}_i^e + ID_i)^K \times r_i \ mod \ N, \text{for } i = 1, \dots, n$$

When all (r_i, s_i) are checked by the agent D_x of the medical staff team, the agent D_x of the medical staff team performs the calculation of the multi-signature as follows:

$$S = \sum_{i=1}^{n} s_i$$

(4) After calculating the multi-signature (R, S, K) of the medical record team group G for the medical record M, it will be stored in the new block of the healthy passbook blockchain. The data of the blockchain changes as shown in Fig.4. And when there are multiple pieces of data, the schematic diagram of the multi-signature of each block is shown in Figure 4:

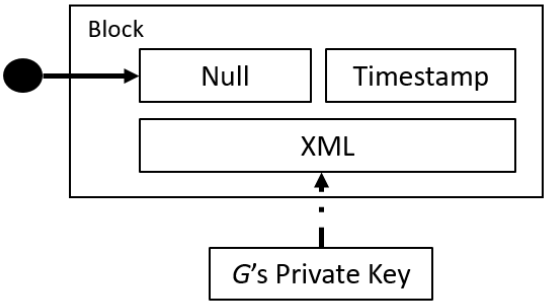


Fig. 4: Schematic diagram of the change of single data in the health passbook blockchain

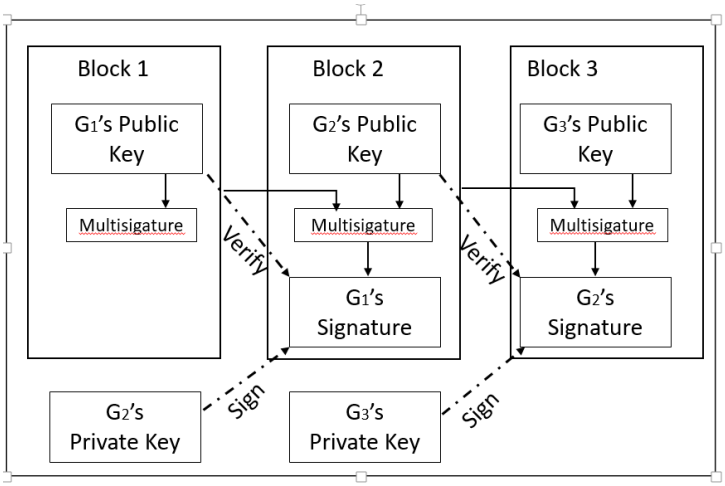


Fig. 5: Schematic diagram of multiple data changes in the blockchain of the health Passbook

4. Health passbook access mechanism: When the medical staff team G or the user $U_i$wants to access the medical record M of a specific block of the health passbook blockchain, the recovery and verification of the medical record M must be performed by a number of multiple signatures (R, S, K) to ensure that the medical record has not been tampered with. The steps are as follows:

(1) The medical record M of a specific block of the healthy passbook block chain is restored by the following mathematical formula.

$$M' = R \times g^S \times ((Y^e + GID)^K)^{-1} \, mod \, N$$

In order to complete the above mathematical formula, the relevant parameters and mathematical expressions involved are as follows:

$$S = \sum_{i=1}^{n} s_i = \sum_{i=1}^{n} x_i \times h(\mathrm{v}_i) \times K$$

$$X = \sum_{i=1}^{n} x_i^{h(v_i)} \left(\text{mod } \emptyset(N)\right)$$

$$Mg^{\Sigma_{i=1}^{n} w_i} \equiv M \prod_{j=1}^{n} r_j \equiv R \; mod \; N$$

After completing the calculation of the relevant parameters and mathematical formulas involved in it, the mathematical formula can be calculated as follows:

$Mg^{S} \equiv Mg^{\Sigma_{i=1}^{n} w_i + \Sigma_{i=1}^{n} x_i \times h(v_i) \times K} \equiv Mg^{\Sigma_{i=1}^{n} w_i} \times g^{X \times K}$

$\equiv R \times g^{X \times K} = R \times (g^{-X})^{-K}$

$\equiv R \times (Y^{e} + GID)^{-K} \; mod \; N$

get $M = R \times g^{S} \times ((Y^{e} + GID)^{K})^{-1} \; mod \; N$。

(2) Check the digital multi-signature with the following mathematical formula:

$$K =? h(M', R, GID)$$

When this mathematical expression is immediately expressed, the digital multi-signature (R, S, K) passes the test, medical staff team group G signed a multi-signature for the medical record M of a specific block of the health passbook blockchain, the medical staff team can trust and receive the digital multi-signature and medical record.

5. Public health bank passbook storage structure: For the smooth progress of medical research, limited data sharing is necessary. Therefore, it is also proposes Public health bank structure for health passbooks for the research units to maintain and synchronize with each other. The schematic diagram is shown in Figure 6:

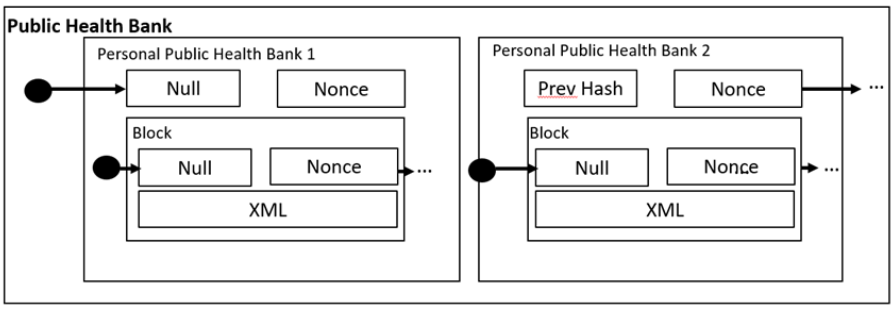


Fig. 6: Schematic diagram of the public health bank blockchain chain structure of the health passbook

6. Health Passbook Public Mechanism: When the user $U_i$wants to disclose the medical record of his health passbook to relevant organizations for sharing, to assist in the promotion of medical related research, the medical institution that can choose a system will be processed to update the health passbook to the public bank blockchain, the schematic diagram is shown in Figure 7. The steps are as follows:

(1) The medical records that the user has de-identified are taken out as shown in Figure 7 and uploaded to the medical institution.

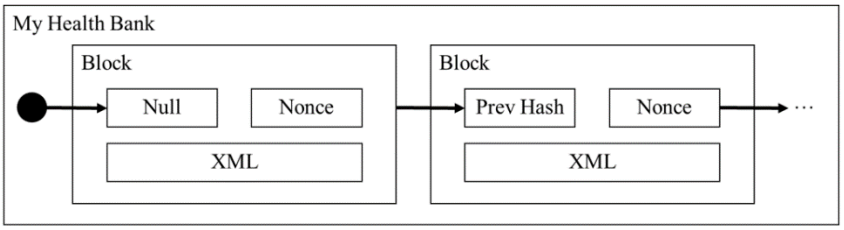


Fig. 7: De-identified medical records

(1) In order to update the information to the public account, the medical institution must select a medical institution as a fair third party to calculate the new hash value in the open account blockchain, so the third party is required to update.
(2) The third party receives the request demand and calculates the new hash value for the new key to be provided to the medical institution.
(3) The two medical institutions will disclose their respective information to the members participating in the public, and complete the information update and synchronization of the public account.

4 Conclusion

Through the above process, which can be used to ensure the safety of the health passbook and the protection of the patient's sensitive personal information by using the smart card identity identification mechanism and the ABAC authorization mechanism, so that the medical personnel and the people all over the country can use personal health passbooks with both efficiency and security. Make the information that physicians can refer to more complete. In the large-scale environment, the traditional RBAC authorization mechanism will increase the related management time with the increase of the implementation time, the increase of the data volume and the complexity. Uses the ABAC to determine the authorization and weaken the role binding authority relationships directly. Effectively remove the problem of responsibilities division of powers through indirect relationships between attributes. In terms of setting, it also sets access the content related to the patient's health passbook only for the different attributes of the person. For example, the pharmacist cannot read the patient's medical record originally, but after the attribute is converted, The information about the part of the patient and the pharmacist is allowed to access, so after the modularization of the health passbook and the block authorization, in addition to allowing the medical staff to have access to relevant information, it also has a higher specification guarantee of sensitive personal information for the patient. In the case of cross-hospital medical record exchange, and because of the group (each hospital is a group) attribute, a limited authorization and medical record search can be performed under the group status. In addition, because the personal health passbook exchange center has only the de-identification of

the personal health passbook index, even if it is attacked, it will not flow out sensitive personal information. At the same time, the control server is authorized to perform group authorization certification for hospitals that have joined the personal health passbook exchange program, and exchanged using the Health Leavel 7 (HL7) standard. The advantage is that the authorization status can be customized according to the hospital's status format, and is also compatible with the HL7 Clinical Document Architecture, Relese 2.0 (HL7 CDA R2) standard format. Therefore, which is expected to add an authorization mechanism to the existing personal health passbook platform, and combine with the original personal health passbook signature mechanism to complete an electronic medical record information system with higher safety specifications.

References

[1] O. Kemkarl and D. Dahikar, "Can electronic medical record systems transform health care? potential health benefits, savings, and cost using latest advancements in ict for better interactive healthcare learning," *International Journal of Computer Science & Communication Networks,* vol. 2, no. 3/6, pp. 453-455, 2012.

[2] B. K. P. Groves, D. Knott, and S. V. Kuiken. "The 'big data' revolution in healthcare: Accelerating value and innovation." http://200.25.59.71:8081/jspui/bitstream/11146/465/1/1661-The_big_data_revolution_in_healthcare.pdf (accessed.

[3] R. Hillestad *et al.*, "Can electronic medical record systems transform health care? Potential health benefits, savings, and costs," *Health affairs,* vol. 24, no. 5, pp. 1103-1117, 2005.

[4] M.-F. Tsai, S.-Y. Hung, W.-J. Yu, C. C. Chen, and D. C. Yen, "Understanding physicians' adoption of electronic medical records: Healthcare technology self-efficacy, service level and risk perspectives," *Computer Standards & Interfaces,* vol. 66, p. 103342, 2019.

[5] A. S. Radwan, A. A. Abdel-Hamid, and Y. Hanafy, "Cloud-based service for secure electronic medical record exchange," in *2012 22nd International Conference on Computer Theory and Applications (ICCTA)*, 2012: IEEE, pp. 94-103.

[6] K. Itakura and K. Nakamura, "A public-key cryptosystem suitable for digital multisignatures," *NEC Research & Development,* no. 71, pp. 1-8, 1983.

[7] S. Nakamoto, "Bitcoin: A peer-to-peer electronic cash system," Manubot, 2019.

[8] Z. Li, A. V. Barenji, and G. Q. Huang, "Toward a blockchain cloud manufacturing system as a peer to peer distributed network platform,"

Robotics and Computer-Integrated Manufacturing, vol. 54, pp. 133-144, 2018.

[9] S. Nakamoto. "Bitcoin: A Peer-to-Peer Electronic Cash System." https://bitcoin.org/bitcoin.pdf (accessed Dec.27, 2017).

[10] N. Nakamoto, "Centralised Bitcoin: A Secure and High Performance Electronic Cash System," *Available at SSRN 3065723,* 2017.

[11] E. Heilman, A. Kendler, A. Zohar, and S. Goldberg, "Eclipse attacks on bitcoin's peer-to-peer network," in *24th {USENIX} Security Symposium ({USENIX} Security 15)*, 2015, pp. 129-144.

Using RoBERTa and Linguistic Features to Detect Fake News

Jhu-Jyun Huang [1], Yen-Heng Tsao[1], Zi-Ying Chen[1], You-Chuan Yang[1]

[1] Innovative Digi-Tech-Enabled Applicaiton and Service Institute,
Institute for Information Industry, Taiwan
{Ginnyhuang, isatsao, ziyingchen , ycyang }@iii.org.tw

Abstract..
Auto-detecting fack news has been a important part in the recent years. In this study, we define 5 linguistic features, which may have a correlation of various grammatical categories to fake news. To classify the linguistic features in Fake news, we collected 1000 fake article and tagged 3310 sentence which split by each article. Then we used RoBERTa model to classify the linguistic features. After alculating and testing the model, we found that using RoBERTa model is feasible to to classify the linguistic features of fack news. By using the model, we hope that will help the Fact-checking reduce their checking time.

Keywords: Fake News, RoBERTa, linguistic features.

1 Introduction

Rumer has existed for a very long time. But the term"Fake News" became popular during the 2016 US election campaign. At that time, President Donald Trump usually used the term "fake news" to critical mainstream media (CNN Cable News Network, ABC News, Washington Post, etc.) which report the negative press coverage of himself.

We are defining "Fake news" as those writings are not true: It might means that writing is fabricated, with no any verifiable facts, sources or line. Sometimes these writing designed to mislead the reader, for specific purpose. It may be designed as "clickbait" which use for online fraud. Another possibility is that someone deliberately write disinformation to change other people's mind. It is also known as media framing.

"Fake news" usually represents a formal news format. Sometimes it also be appear "Fake news" usually represents a formal news format. Sometimes it also be appeared in the Informal frame of news, such as message, picture, video, or only one sentence.

In recent years, fake news have increased rapidly via social media, since they are so easily and quickly shared online. Therefore, How to detect quickly fake news have become an important part in our daily life. Moreover, providing a basic guidelines for fact-checking in various types of content also need to do it.

2 linguistic features

There are some past researches showed that “Fake News” maybe including some specific linguistic features for the purpose of increase the confidence of article. Facebook (2018) [1], IFLA (2020) [2], Veronica P´erez-Rosas1, Bennett Kleinberg2, Alexandra Lefevre1 Rada Miracle (2018).

By reading these researches and our found of article, we summarize the relationship of various grammatical categories to fake news. There are 5 linguistic features, their definitions, and what we asked of annotators.

1. Contain emotional, radical feature
   Fake news usually contains the following language features that may inspire readers’ emotions, thinking, and action. It may using “Clickbait” Title, shocking, satirical twists on current affairs Facebook (2018), Veronica P´erez-Rosas1, Bennett Kleinberg2, Alexandra Lefevre1 Rada Miracle (2018)
2. Contain fabricate expert/ Anonymous source
   Fake news sometime fabricate expert that endorsements content. Anonymous source (eg. Did you hear about……) is another way using in fake news to Deceive readers. On top of that, Fake news often use the third person pronouns, like “he” and “she” rather than full names.
3. Contain Arbitrary mood
   Fake news reports usually contain more extreme, or using arbitrary expressed, such as “absolutely”, “it must be”. The researcher of Fatimah Taraba As at Simon Fraser University in Canada(2017), found that some sentence of fake news which can be used to exaggerate are all found more often in deliberately misleading sources. These included superlatives, such as “the most” and “the worst”.
4. Contain the social call
   Fake news have a lot of sharing call that try to spread more other people sometime it expects to use social network sharing to earn traffic, personal information or audience fans, so there are Articles with similar content may be false
5. Contain the Inferential tone:
   False news often use hypothetical situations inference related terms, trying to influence readers’ judgments. The implications can be dramatic, especially in medical trials, Veronica P´erez-Rosas1, Bennett Kleinberg2, and Alexandra Lefevre1 Rada Miracle (2018) Data

[1] https://www.cna.com.tw/news/afe/201810220368.aspx

[2] https://www.ifla.org/publications/node/11174

3 Fact Checking Organizations of Taiwan

All of Fact-checking agency/platform/app in Taiwan fellow the Code of IFCN-Principles (international fact-checking network) scrupulously: There are 5 principle,

1. A commitment to nonpartisanship and fairness
2. A commitment to transparency of sources
3. A commitment to transparency of funding & organization
4. A commitment to transparency of methodology
5. A commitment to open and honest corrections

In short, all of fact-checking agency/platform/app should be "OPEN" and "RANSPARENCY". There are 2 type of checking below:

1. Civilian Collaboration
   First, the suspicious information is reported by the public online, and then the data will be imputed in database. Next, another group of online user can view the information on app/website, check it, and make it public the checked-report to others. The user is not necessarily an expert, such as Cofacts.
2. FactCheck Center agency
   A group that make up from media literacy professionals. The fact-checking report published aperiodically. Like MyGoPan, Taiwan Fact Checking Center.

No matter which Fact-check method or not. People need to check the news item by item, that really require professionals participation.

4 Methodology

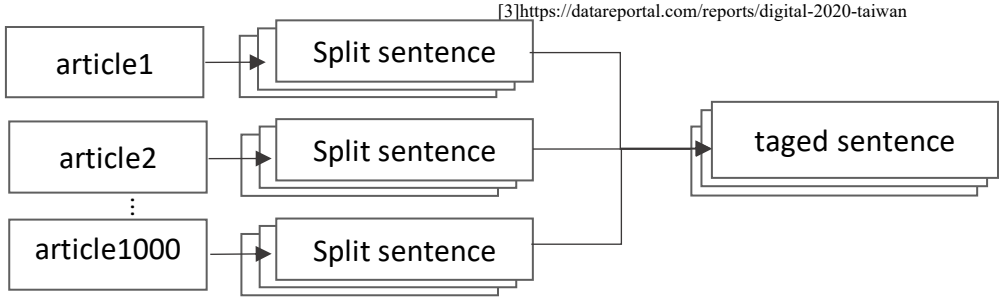
In this section, we try to find an automatic way to detect the linguistic features. And we hope to help media literacy professionals reducing their checking times is.

4.1 Data preprocessing

Internet have become one of indispensable part in our daily life. A recent survey, digital 2020, showed that there are 88% social media users in Taiwan. Ranking of top 3 app of user are LINE, Facebook, and Facebook message [3]

Thus, Fake News usually be sharing in social media, especially using in private messages.

[3]https://datareportal.com/reports/digital-2020-taiwan



a total of 3307

Fig. 1 split fake news process

In the study, we use the fake news that has been proved by three fact checking organizations of Taiwan (FactCheck Center agency, Mygopen, Cofact)
We selected a 1,000 Fake news, And split each article into petty sentence according to '\n'," ◦ ". If the number of character more then 150, divide the number of words by two, and go forward to find the last punctuation mark of each sentence. Then tag each sentence with different linguistic facture, shown in Figure 1.

4.2 Sentence tag

Our sentence tag use "Contain emotional, radical feature", "Contain fabricate expert/ Anonymous source", "Contain Arbitrary mood", "Contain the social call", "Contain the Inferential tone".

We use 3 group of training set to compare the differences between each model.

The first round of training set we talked 551 sentence of emotional, radical feature; 460 sentence of arbitrary mood; 145 sentences of social call; 145 sentences of inferential tone; 357 sentence of fabricate expert/ Anonymous source; and 460 sentences without any linguistic features. Totaling 2215 sentences. In this part, we mainly want to know the practicality of linguistic features.

The second round of training set we talked 583 sentence of emotional, radical feature; 484 sentence of arbitrary mood; 289 sentences of social call; 330 sentences of inferential tone; 373 sentence of fabricate expert/ Anonymous source; and 711 sentences without any linguistic features, totaling 2770 sentences for training;

In this part, we want to solve the problem of the imbalance in the number of linguistic features in each sentence. The ratio of the number of feature sentences is adjusted to an average. The ratio of 0.2 for emotion; 0.2 for arbitrary; 0.1 for social call; 0.1 for inference; 0.1 for fabricate expert/ Anonymous source; and 0.3 for not any linguistic features.

You may confuse that why do not separated into ratio equally among the linguistic features. Because we think that by adjusting unbalance ratio, it is possible to simulate the actual proportion of fake news with various language features.

The third round of training set. We want to increase the number of sentences with various language features, and expect to improve accuracy, the ratio of 0.2 for emotion; 0.2 for arbitrary; 0.1 for social call; 0.1 for inference; 0.1 for fabricate expert/ Anonymous source; and 0.3 for not any linguistic features. We talked 600 sentence of emotional, radical feature; 500 sentence of arbitrary mood; 440 sentences of social call; 450 sentences of inferential tone; 450 sentence of fabricate expert/ Anonymous source;

and 871 sentences without any linguistic features. A total of 3310 sentences were trained.

Finally, we also prepared a set of testing set, a total of 356 sentences, to verify the accuracy of linguistic feature classification.

4.3 model

For classify the linguistic feature, we use RoBERTa, which improve the BERT (Bidirectional Encoder Representations from Transformers):

BERT is a recent learning method which published by researchers at Google AI Language. This model used a new technique called Masked Language Model (Masked LM). Masked LM allow bidirectional training in models which can looked at a sequence from left to right and right-to-left training.

The RoBETa model uses a dynamic mask, which is different from the pre-mask processing of the corpus in the original BERT. RoBERTa uses 10 different masks during 40 rounds of training. The model removed the NSP task and found that it can slightly improve the performance of downstream tasks. The model uses larger training data and larger batch size. The original BERT uses a 30K of words, than RoBERTa uses 50K words.

4.4 Evaluation index

We use Precision, Recall, and accuracy to evaluating classification models.

1. Precision is the fraction of information retrieved that are relevant.
2. Recall is the fraction of relevant information that are retrieved.
3. Precision is the fraction of information retrieved that are relevant.

Accuracy is the fraction of predictions our model got right.
When the index arrive to 1, the better model we have.

5 Results

In these section, we shown that test results of fake news linguistic features classification. (Table 1) We used a total of three sets of training sets for analysis, and the test data used the same set.

From the comparison of one and two groups, it can be seen that adjusting the proportion of each category can increase the accuracy, from 78.37% to 82.87% From the precision and recall of each category, the overall situation is improved, except the fabricate expert/ Anonymous source. Although the precision and recall of fabricate expert/ Anonymous source was reduced, the reduction was not large. Precision decreased from 0.913 to 0.870; reccall decreased from 0.750 to 0.714.

Comparing the second and third groups, it can be seen that adjusting the proportion of each category can increase the accuracy, from 82.87% to 85.67%

From the precision and recall of each category, the arbitrary tone decreased from 0.609 to 0.580, but the reduction was not large.

To sum up, it can be seen that the precision of inquiries and arbitrary terms is poorer than other linguistic features, and the others are pretty good. The prediction accuracy of the third classification model is 86.67%.

Table 1. The result from each group

➢ First group			
	precision	recall	accuracy
Contain emotional, radical feature	0.667	0.794	78.37%
Contain Arbitrary mood	0.538	0.875	
Contain the social call	0.833	0.714	
Contain the social call	0.364	0.667	
Without any linguistic features.	0.907	0.779	
Contain fabricate expert/ anonymous source	0.913	0.750	
➢ Second group			
	precision	recall	accuracy
Contain emotional, radical feature	0.743	0.873	82.87%
Contain Arbitrary mood	0.609	0.875	
Contain the social call	0.909	0.714	
Contain the social call	0.462	1.000	
Without any linguistic features.	0.931	0.826	
Contain fabricate expert/ anonymous source	0.870	0.714	
➢ Third group			
	precision	recall	accuracy
Contain emotional, radical feature	0.791	0.877	86.67%
Contain Arbitrary mood	0.580	0.906	
Contain the social call	0.867	0.929	
Contain the social call	0.500	1.000	
Without any linguistic features.	0.962	0.836	
Contain fabricate expert/ anonymous source	0.962	0.893	

6 Conclusion

The linguistic features of fake news can be seen from the analysis results, which can be distinguished by using RoBERTa model. However, the accuracy of arbitrary tone and

inferential tone are relatively insufficient then other, which may be related to the definition of linguistic features.. Any sentence of “maybe” or “if...than” will be regarded as inference terms, but sometimes it is incorrect. There is no sense of inference from the sentence. I think we can discuss the indicators with the people who work in the Fact-checking center, and modify the definition of feature. We hope linguistic features of Fack news detection model can be more and more accurate in the future.

References

1. Pérez-Rosas, V., Kleinberg, B., Lefevre, A., Mihalcea, R.: Automatic detection of fake news. CoRR abs/1708.07104 (2017)
2. Amy X. Zhang, Aditya Ranganathan, Sarah Emlen Metz, Scott Appling, Connie Moon Sehat, Norman Gilmore, Nick B. Adams, Emmanuel Vincent, Jennifer Lee, Martin Robbins, et al.. A structured response to misinformation: Defining and annotating credibility indicators in news articles. In Proceedings of the Web Conference. International World Wide Web Conferences Steering Committee, 603--612. (2018)
3. Hannah Rashkin, Eunsol Choi, Jin Yea Jang, Svitlana Volkova, and Yejin Choi. b. Truth of varying shades: Analyzing language in fake news and political fact-checking Proceedings of EMNLP. 2921—2927(2017)
4. Devlin, J., Chang, M. W., Lee, K., & Toutanova, K. Bert: Pre-training of deep bidirectional transformers for language understanding. arXiv preprint arXiv:1810.04805(2018).
5. Liu, Y., Ott, M., Goyal, N., Du, J., Joshi, M., Chen, D., … & Stoyanov, V. Roberta: A robustly optimized bert pretraining approach. arXiv preprint arXiv:1907.11692. (2019).

Developing an Accessibility Assessment System for Mobile Devices

Ting-Fang Wu[1], Chi-Nung Chu[2], Hui-Shan Lo[3]

[1]Graduate Institute of Rehabilitation Counseling, National Taiwan Normal University
162, Sec. 1 Hoping East Rd., Taipei, Taiwan
tfwu@ntnu.edu.tw
[2]Department of Information Management, China University of Technology
56, Sec. 3, Xinglong Rd., Taipei , Taiwan
nung@cute.edu.tw
[3]Department of Special Education, National Taiwan Normal University
162, Sec. 1 Hoping East Rd., Taipei, Taiwan
80209003E@ntnu.edu.tw

Through the rapid evolution of the technology, more and more mobile devices have been used in a learning environment. The intuitional operational layout and less physical efforts in operating touchscreen, most people find it is easier for use compared to traditional input devices, such as keyboards and mice. However, persons with physical disabilities may encounter difficulties in accessing or operating those mobile devices due to poor control of movements and dexterities. The purpose of this study is to develop an application to help the research team to build, test, maintain, update, and scale the accessibility assessments of mobile phones. The Accessibility Assessment System was designed as a cloud-based application, and its design is task-oriented, including tapping, sliding and dragging. The time of completion of all tasks, error rate of manipulations, correctness of usage and the size of the finger press area were recorded. In addition, the results of the assessments would be automatically uploaded to the cloud database for further analysis. Through the three functional evaluation tasks of the Accessibility Assessment System, we would be able to understand the proper icon sizes, and the controllable area of touch screen on their mobile devices for the individuals with physical disabilities. For further studies, we will test AAS within different diagnoses of persons with physical disabilities, including cerebral palsy, spinal cord injury, and muscular dystrophy to look at their operational performance on the mobile devices.

KEYWORD: Accessibility Assessment, Mobile Devices, Persons with Physical Disabilities

1 Introduction

According to a survey of e-Market, the population of mobile phone users in Taiwan is about 73.4% [1]. With the popularity of mobile devices, it has become a necessity of daily life; more and more people use it for education. The intuitive operating mode of iPod or iPad can provide a more effective and efficient teaching and learning method compared to the tradition ones [2]. Although there are more and more studies using mobile devices to teach, it still needs more discussion about whether it is suitable for persons with disabilities.

According to a survey by Rehabilitation Engineering Research Center for Wireless Technologies from 2012 to 2013, it was found that approximately 40% of people with physical disabilities only have a simple mobile phone instead of a touch-screen smart phone. This reason may be that traditional mobile phones have large buttons, which are easier for persons with physical disabilities to access. However, smart phones with a touch-screen layout usually require more dexterous finger movements to operate, thus this may cause distress for persons with physical disabilities [3]. Duff et al. (2010) also found that persons with physical disabilities have a higher error rate when entering numbers on the on-screen keyboard than those without disabilities [4]. Another study also demonstrated that persons with physical disabilities have higher error rates, slower speeds and longer dwell times when operating the touch screens than people without disabilities [5]. From the literature, we can see that people with physical disabilities have many difficulties using smartphones, but smartphones are mainstream products and have their advantages in use.

Newton & Dell (2011) had mentioned that there were two major obstacles for the use of assistive devices for persons with disabilities; they were high costs and high abandonment rates [6]. Since the mobile devices are mainstream products, their prices are much cheaper than that of customized assistive devices, therefore, persons with disabilities might accept a mobile device as its low cost and as a mainstream trend. Therapists servicing persons with physical disabilities would like to provide accessibilities suggestions for them to use mobile phones barrier-free. The purpose of this study was to develop an application for mobile devices, which can simulate the functional tasks when operating the mobile devices' touch-screen layout. The APP developed by this study can assist the therapist to evaluate the difficulties of using mobile devices for persons with physical disabilities.

2 Method

The Accessibility Assessment System (AAS) was designed as a cloud-based application accessed through the web to help the research team to build, test, maintain, update, and scale the accessibility assessments of mobile phones much more quickly and less expensively. The web server architecture combined the Apache web server with PHP, Perl, and Maria DB, allowing users to easily connect the web server with their mobile phones. The design of AAS is task-oriented, including tapping, sliding and dragging. The time of completion of all tasks, error rate of manipulations, correctness of usage and the size of the finger press area were recorded. The AAS used the game style scenario to induce the subject's motivation, and the task orientation design to collect the related times and tracks of tapping, sliding or dragging through user's operation on the multi-touch screen. In addition, the results of the assessments would be automatically uploaded to the cloud database for further analysis.

3 Results

The AAS consisted of three functional tasks, which were tapping, sliding and dragging.

3.1 Tap Assessment

“Tap” means to point and select the icon on the touch screen of the mobile phone. In the assessment of tap, the sizes of icons and the areas the person is able to access were assessed through the continuous tapping tasks. Four different modules (5*4, 4*3,3*2, and 2*1, fig 1a-1d) were designed according to the different sizes of the icons. For example, the module 5*4 means that the layout of the icons on the touch screen are arranged in 4 columns and 5 rows (Fig 1a). The subjects will be asked to tap on the flashing icon on the screen continuously; and the time of completion of tapping tasks, accuracy and the area of correct tapping were recorded (Fig 1e-1h). For instance, the figure 1e indicated the correct tapping area of the 5*4 module.

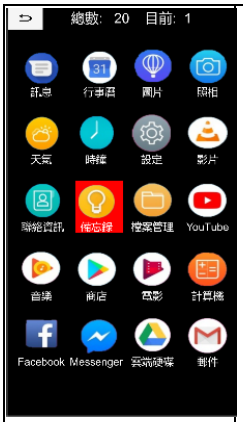


Fig 1a: touch screen layout of 5x4 module

Fig 1b: touch screen layout of 4x3 module

Fig 1c: touch screen layout of the 3x2 module

Fig 1d: touch screen layout of 2x1 module

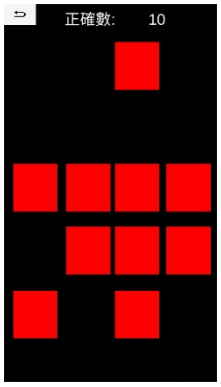


Fig 1e: correct tapping area of 5*4 module.

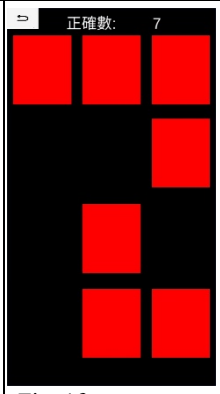


Fig 1f: correct tapping area of 4*3 module.

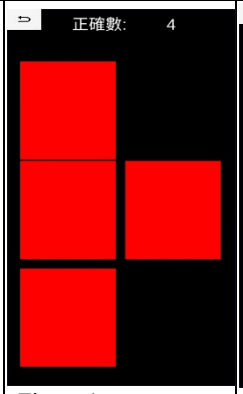


Fig 1g: correct tapping area of 3*2 module.

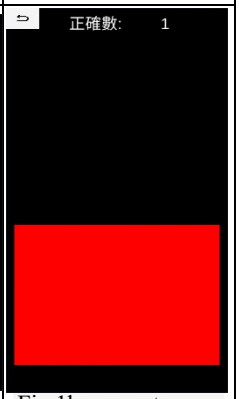


Fig 1h: correct tapping area of 2*1 module.

3.2 Slide

This sliding test simulates the action of the person turning the page left or right or moving the page up and down on the mobile devices. In this test, the screen shows four arrows in different directions: up, down, left, and right. The subject will be asked to slide from button to top (up, Fig 2a), from top to button (down, Fig2b), from left to right, (right, Fig 2c) and from right to left (left, Fig 2d) on the screen. The accuracy, the time to complete the tasks, and trajectory of sliding will be recorded.

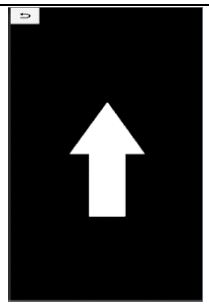


Fig 2a: slide from button to top

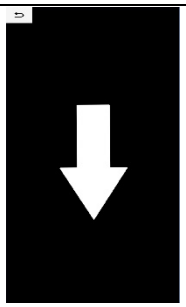


Fig 2b: slide from top to button

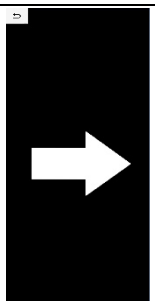


Fig 2c: slide from left to right

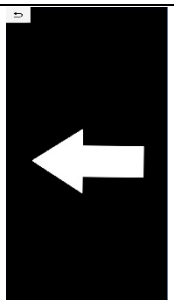


Fig 2d: slide from right to left

3.3 Drag

"Drag" refers to holding the target on the screen and dragging the target to the specified location. In the drag test, the subject will be asked to drag from the blue line on the bottom to the red line. Four tasks, dragging from button to top (Fig 3a), from top to button (Fig 3b), from left to right (Fig 3c), and from right to left (Fig 3d) on the screen were tested. The location where the subject first touched the screen, and the distance and trajectory of dragging, as well as the time to completion will be recorded for further analysis.

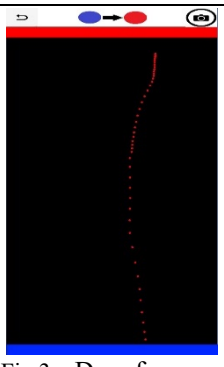


Fig 3a: Drag from button to top

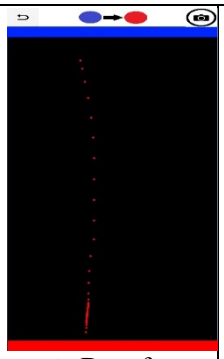


Fig 3b: Drag from top to button

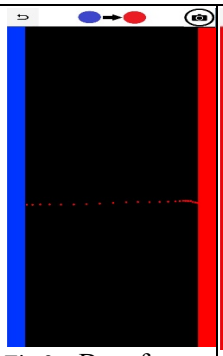


Fig 3c: Drag from left to right

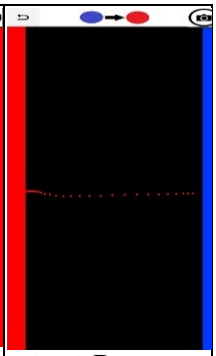


Fig 3d: Drag from right to left

4 Discussion

The AAS has been developed, and it seems suitable to evaluate the operation performance of mobile devices for persons with physical disabilities. Through the three functional evaluation tasks of the AAS, we would be able to understand the proper icon sizes, and the controllable area of touch screen on their mobile devices for the individuals with physical disabilities. For further studies, we will test AAS within different diagnoses of persons with physical disabilities, including cerebral palsy, spinal cord injury, and muscular dystrophy to look at their operational performance on the mobile devices.

Acknowledgment

The authors would like to thank the National Taiwan Normal University for financially supporting this research under Grant No. T10807000107.

References

1. Sodano, S. (2016). Mobile Taiwan: A look at a highly mobile market. Retrieved form-https://www.emarketer.com/Article/Mobile-Taiwan-Look-Highly-Mobile-Market/1014877
2. Manuguerra, M. & Petocz, P. (2011). Promoting student engagement by integrating new technology into tertiary education: The role of the iPad. *Asian Social Science, 7*(11), 61-65.
3. Morris, J., Mueller, J., Jones, M. L., & Lippincott, B. (2014) . Wireless technology use and disability: Results from a national survey. *Journal on Technology and Persons with Disabilities, 1,* 70-80.
4. Duff, S. N., Irwin, C. B., Skye, J. L., Sesto, M. E., & Wiegmann, D.A. (2010). The effect of disability and approach on touch screen performance during a number entry task. *Proc. HFES'10,* 566–570.
5. Irwin, C. B. & Sesto, M. E. (2012). Performance and touch characteristics of disabled and non-disabled participants during a reciprocal tapping task using touch screen technology. *Applied Ergonomics, 43*(6)*,* 1038–1043
6. Newton, D. A., & Dell, A. G. (2011). Mobile device and students with disabilities: What do best practices tell us? *Journal of Special Education Technology, 26*(3), 47-49.

A Mobile Module Design of Rhythm in Music Based on Practicing Strategy

Yu Ting Huang[1], Chi Nung Chu[2]

[1] Department of Music, Shih Chien University,
No.70 Ta-Chih Street, Chung-Shan District,Taipei, Taiwan, R.O.C.
yuting11@mail.usc.edu.tw
[2] Department of Management of Information System, China University of Technology,
No. 56, Sec. 3, Shinglung Rd., Wenshan Chiu, Taipei, Taiwan 116, R.O.C.
nung@cute.edu.tw

Abstract. This research proposed a prototype design module of smart mobile phone apps for the difficulty of music rhythm learning. The learning effect of traditional music rhythm is mainly achieved through one-on-one instruction by the teacher, and the learner's continuous self-practice with the metronome through hearing. Therefore, in order to improve the efficiency and cost of music rhythm learning, this study proposes a music rhythm learning model in which the smart mobile phone is implemented as a self-practice tool for the learners. Learners can choose one of two modules in touch screen tapping or screen auto-rotation induction through the mobile module design of rhythm in music based on practicing strategy. It could carry out music rhythm practice for the learners at any time and any place to achieve the learning effect. This kind of practice strategy using the smartphone apps can achieve the correctness of the traditional rhythm of the tapping with the hands monitored by teacher.

Keywords: smart mobile phone, music rhythm, learning effect

1 Introduction

The elements of music include pitch, rhythm, and timbre [5, 6]. Dealing with the music rhythm, what the many beginners often feel difficult to master is to understand the notes on the music staff, and then how to clearly grasp the tempos [12, 16, 21]. Traditionally the solutions for these problems are not easy for the beginners. In addition to requiring a lot of time to practice repeatedly, a supervisor to supervise is also needed. With the development of information technology, the acquisition of knowledge has become easier [10, 11, 13, 18]. The gap between music and learners has become closer, and the learning opportunities have become more diverse in the coming era of information technology [7, 15, 20]. Even so, the practice of music rhythm cannot be completely separated from the way that requires human guidance. For beginners, the correctness of the music rhythm is often difficult to detect because of the abstract hearing [9, 17]. Even if the wrong rhythm is pointed out, they could not identify pre-

cisely where the difference is. In addition, as the instructors cannot be consulted right away, the frustrating learning problems will cause beginners to wander in the learning place, and even gradually retreat from music learning [8, 14].
Mobile learning facilitating instructions with portable technology provides learners with the maneuverability [3, 19]. Beyond the traditional instruction limitations in classroom, the utilization of handheld devices allows comprehensive learning. The convenience of mobile learning happens as long as there are learning materials that can be retrieved by a mobile phone [1, 2]. Learning process can be carried out directly. Therefore, with the diffusion of hand-held technology, almost everyone could access the Internet at any time to obtain information. Many learning tasks would be gradually changed to the mode of distance teaching. Music rhythm learning could thus keep up with the trend in mobile learning to facilitate beginners to solve problems with flexibility.

2 Design of Mobile Module in Music Rhythm

Music rhythm which is presented in time pattern is the key element of all music. The design of Mobile Module in Music Rhythm is a learning system facilitating the novice learners to practice the music rhythm ubiquitously. The APP could listen to the learner's tapping or swinging with tempo estimation to evaluate how well the learner achieves the tempo pattern, and the learning system could also give immediate feedback to the learner.

2.1 System Development Strategy

The mobile module design of rhythm in music is based on the learning theory in behaviorism which claims that knowledge or skill can be acquired with practice through the mechanism of stimulus and feedbacks [4]. Learning music rhythm has more to do with the perception of time. Whatever a given chunk of music notes present with pitch or timbre patterns, music rhythm could exist without melody such as the drumbeats, but melody could not exist without rhythm. Therefore it becomes more important for the learner to get the skill of music rhythm through practicing repeatedly. The learning interface is shown as Fig.1.

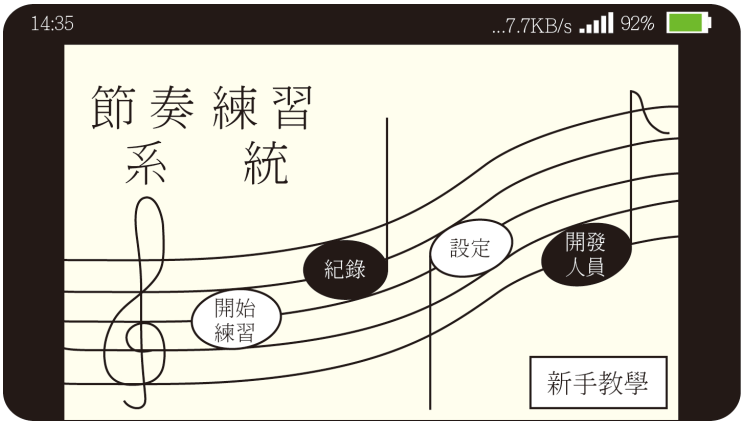


Fig. 1. System of Mobile Module in Music Rhythm

2.2 Design of Mobile Module in Music Rhythm

Music rhythm refers to the sequence of music notes and rests in time. A practicing rhythmic pattern consists of a series of notes and rests. (Fig.2)

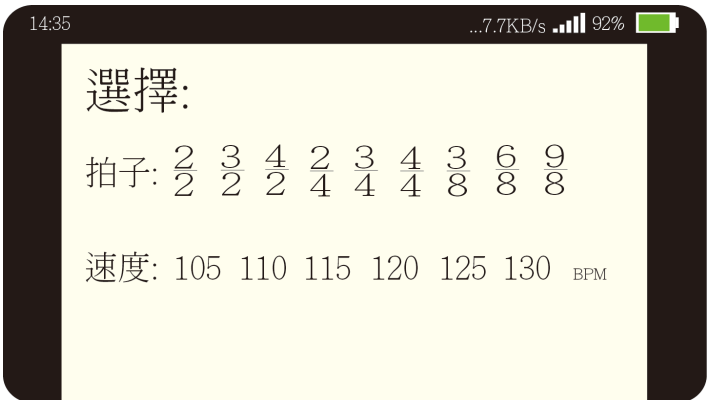


Fig. 2. Selection interface for praticing items

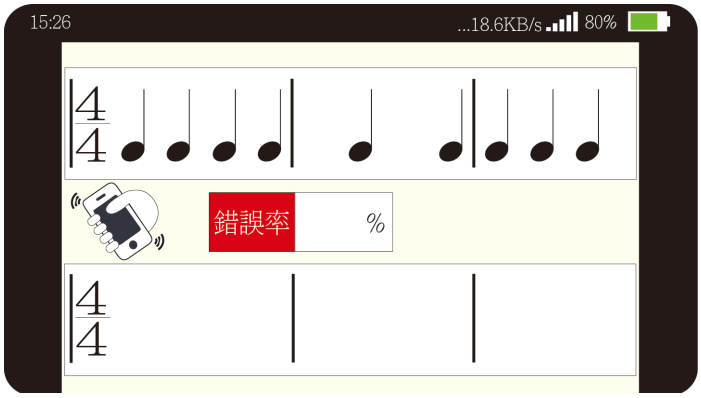


Fig. 3. Tapping mode

In addition to indicating which practicing item is set, the mobile module design of music rhythm provides alternative tapping mode (Fig.3) and screen orientation mode (Fig.4) for better support to the novice at music rhythm learning.

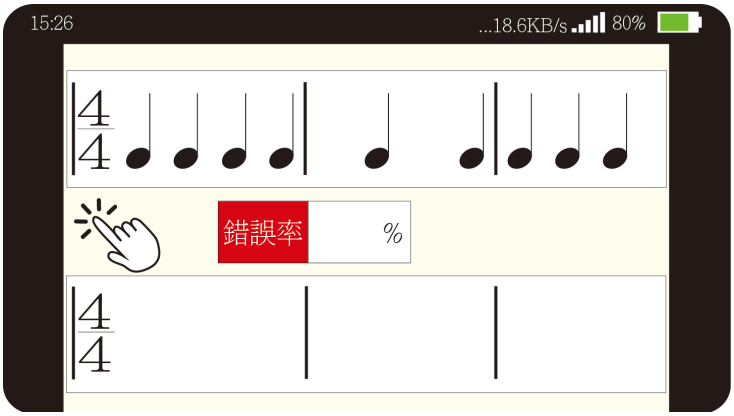


Fig. 4. Screen orientation mode

3 Conclusion

The elements of practicing strategy in the mobile module design of music rhythm with alternative tapping mode and screen orientation mode intend to better support the novice at music rhythm learning. A further assessment for the mechanisms in learning effects and retention with learners will be conducted in order to verify what the proposed strategy expects. Furthermore, music rhythm learning through mobile services

that are inherited by intermission time and location properties needs to further evaluate in the future.

References

1. Ahmed, M. S. (2016). Technology acceptance of smartphones as mobile learning tools: A contextual comparative study of engineering and education colleges.
2. Al-Harrasi, H., Al-Khanjari, Z., & Sarrab, M. (2015). Proposing a new design approach for m-learning applications. International Journal of Software Engineering and its Applications, 9(11), 11-24.
3. Berking, P., Haag, J., Archibald, T., & Birtwhistle, M. (2012). Mobile learning: Not just another delivery method. In Interservice/Industry Training, Simulation, and Education Conference (I/ITSEC) (pp. 1-30).
4. Bryant, L. C., Vincent, R., Shaqlaih, A., & Moss, G. (2013). Behaviorism and behavioral learning theory. In B. J. Irby, G. Brown, R. Lara-Alecio, & S. Jackson (Eds.), The handbook of educational theories (p. 91–103). IAP Information Age Publishing.
5. Burton, R. (2015). The elements of music: what are they, and who cares?. In Music: Educating for Life. ASME XXth National Conference Proceedings (p. 22). Australian Society for Music Education.
6. Cutietta, R. A. (1993). The Musical Elements Who Said They're Right? Robert A. Cutietta suggests that we have been teaching the wrong things for many years, focusing on musical elements rather than musical perception. Music Educators Journal, 79(9), 48-53.
7. Debevc, M., Weiss, J., Šorgo, A., & Kožuh, I. (2020). Solfeggio learning and the influence of a mobile application based on visual, auditory and tactile modalities. British Journal of Educational Technology, 51(1), 177-193.
8. Duran, D., Ribosa, J., & Sánchez, G. (2020). Peer tutoring for improvement in rhythm reading fluency and comprehension. International Journal of Music Education, 0255761419898313.
9. Furby, V. J. (2016). The effects of peer tutoring on the aural skills performance of undergraduate music majors. Update: Applications of Research in Music Education, 34(3), 33-39.
10. Guillén-Gámez, F. D., Mayorga-Fernández, M. J., Bravo-Agapito, J., & Escribano-Ortiz, D. (2020). Analysis of Teachers' Pedagogical Digital Competence: Identification of Factors Predicting Their Acquisition. Technology, Knowledge and Learning, 1-18.
11. Leidner, D. E., & Jarvenpaa, S. L. (1995). The use of information technology to enhance management school education: A theoretical view. MIS quarterly, 265-291.
12. Levitin, D. J., Grahn, J. A., & London, J. (2018). The psychology of music: Rhythm and movement. Annual review of psychology, 69, 51-75.
13. Mohammadi, H. (2015). Investigating users' perspectives on e-learning: An integration of TAM and IS success model. Computers in Human Behavior, 45, 359-374.
14. Phillips-Silver, J., & Trainor, L. J. (2007). Hearing what the body feels: Auditory encoding of rhythmic movement. Cognition, 105(3), 533-546.
15. Portowitz, A., Peppler, K. A., & Downton, M. (2014). In Harmony: A technology-based music education model to enhance musical understanding and general learning skills. International Journal of Music Education, 32(2), 242-260.
16. Ravignani, A. (2017). Visualizing and interpreting rhythmic patterns using phase space plots. Music Perception: An Interdisciplinary Journal, 34(5), 557-568.

17. Sloboda, J. A. (2000). Individual differences in music performance. Trends in cognitive sciences, 4(10), 397-403.
18. Sun, P. C., Tsai, R. J., Finger, G., Chen, Y. Y., & Yeh, D. (2008). What drives a successful e-Learning? An empirical investigation of the critical factors influencing learner satisfaction. Computers & education, 50(4), 1183-1202.
19. Uden, L. (2007). Activity theory for designing mobile learning. International Journal of Mobile Learning and Organisation, 1(1), 81-102.
20. Wise, S., Greenwood, J., & Davis, N. (2011). Teachers' use of digital technology in secondary music education: illustrations of changing classrooms. British Journal of Music Education, 28(2), 117-134.
21. Wolf, D. L. (2004). A hierarchy of rhythm performance patterns for kindergarten children. Visions of Research in Music Education, 4.


The Institution of
Engineering and Technology
國立臺中科技大學
NATIONAL TAICHUNG UNIVERSITY OF SCIENCE AND TECHNOLOGY