A study on how to generate fire data from video/image using the F-guess and ROI method

Jong-Sik Kim[1[0000-0002-1459-194]], Hye-Youn Lim[2[0000 0003 0186 2430]] and Dae-Seong Kang[3[0000-0003-0186-2430]]

[1]Dong-A University, Dept. of Electronics Engineering, 37 Nakdong-daero 550 beon-gil Saha-gu, Busan, Korea

kjsluke@dau.ac.kr, hylim@dau.ac.kr, dskang@dau.ac.kr

Abstract. This paper presents the best method for generating fire data and improving the fire recognition rate. At the same time, shorten the labelling time by using fire videos and images when there is a limit to collecting the data and hard to improve the recognition rate with a small label such as a fire. In order to improve the recognition rate and shorten the labelling time, the F-guessed method and the region of interest (ROI) expression method were used to process the data so that the predicted result labelling value and the correct answer value is similar. As a result, data generation increased by about 5.4 times from 5,565 data to 35,633 data compared to the initial labelling task, and mAP@0.5 improved by about 17.6% from 65.9% to 83.5%.

Keywords: semi-supervised learning, deep learning, fine-tuning, pseudo labeling

1 Introduction

In the field of computer vision, semi-supervised learning methods have developed rapidly over the past few years. Current state-of-the-art methods introduce hybrid methods by simplifying previous work in terms of architecture and loss function or by mixing different formulations [1]. However, the most representative methodology of deep learning is supervised learning. Supervised learning is a learning method that memorizes patterns in training data. Therefore, it is not easy to recognize data that has never been learned. For good generalization, more labeled data is required [2].
In cases where labelled data is limited, such as fire, applying the existing method is difficult. In other words, collected fire result labels cannot handle whole. When new test data not included in the training data came in, Mis-recognition was made often. Therefore, in this paper, the existing Pseudo-label method with fine-tune learning and classify to data with no correct answers. The study focuses on the shape and size of fires, particularly their small and similar shapes in the initial stages[3].
This paper proposes that set an ROI on a fire scene of fire ignition video and comparing it with ROI data that the initial learning model predicts. If differences are less than 50%, use Guessed Label method to increase the fire recognition rate and reduce labelling time. Fig. 1. is a conceptual diagram of initial fire data generation using the ROI comparison method of the proposed algorithm.

2 Understanding semi-supervised learning

Semi-supervised learning can be considered if there is a small amount of correct label data (Labeled) and a lot of uncorrected label data (Unlabeled) which can be used. Semi-supervised learning aims to improve performance by applying supervised learning to a small number of correct labels and unsupervised learning to many uncorrected labels. Various semi-supervised learning methodologies have appeared regarding how to use labels without correct answers for learning. Among the methodologies similar to the currently proposed are pseudo-labelling methods and MixMatch also FixMatch.

Pseudo labeling [4] uses a model primarily trained through supervised learning to make predictions on unlabeled data. Label it as Pseudo using the performed prediction result. Therefore, to perform Pseudo labelling, there must be a trained model and label data without correct answers. For labels without correct answers, performing secondary learning using extended data set after Pseudo labelling.

MixMatch[5][6] method generates processed correct label samples (X') and predicted labeled (U') when given correct label data (X) and uncorrected label data (U). This method applies entropy minimization, label consistency regularization, and mixup.

FixMatch [7] method uses cross-entropy loss to train a supervised learning model from correct label images. Two images are obtained by applying weak and strong enhancement methods to each image of the unanswered label. The weakly augmented image is passed to the model and the prediction for the class is obtained and the probability of the most confident class is compared against a threshold. The corresponding class uses a pseudo-label as the default label if it exceeds the threshold value.
The enormously augmented image passes through the model to make predictions for the class. This prediction is used as the cross-entropy loss compared to the correct pseudo-label. At this point, the two losses are combined, and the model is optimized.

3 Proposal method

“ A Study on Fire Data Generation and Recognition Rate Improvement using F-guessed and Semi-supervised Learning”[8] by the author proposed the Semi-supervised learning method that uses a Pseudo labelling method that extracts images per frame from a fire video and using a fire answer label. In this study, we changed the method to a method that can extract guessed label from a video without changing the video to images. In addition, for the pseudo-labelled guessed from image/video files, the guessed labels were determined by increasing the value approximately twice every step from 2000 files. Fig. 1. indicates the whole system fire data generating diagram that is proposed.

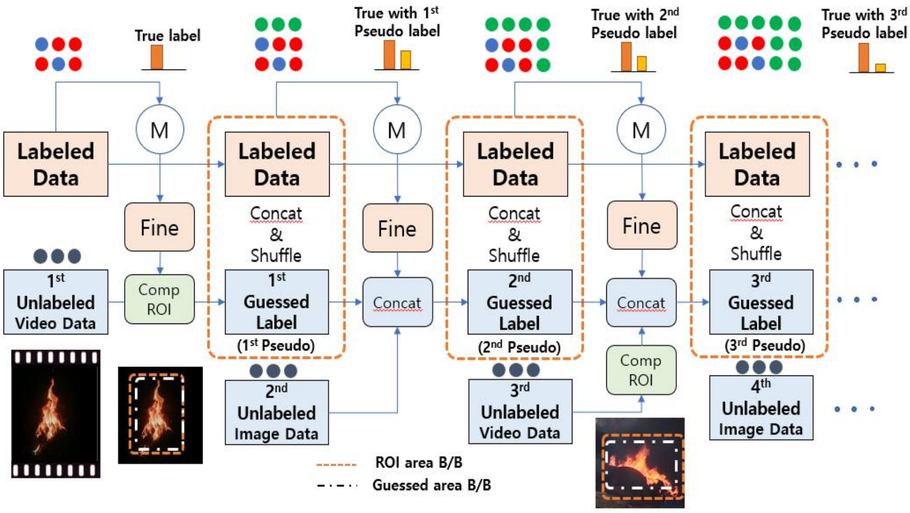

Fig. 1. Conceptual diagram of fire data generation using region of interest (ROI) comparison method.

Also, when the fire ignites, it gradually increases over a specific time, as in Fig 2. The fires have different shapes, but the sizes are similar before the fires increase as long as the camera is not moving. Focusing on this, once an ROI is set in a fire video which extracts pseudo labelling from a fire starting point in the video, the shape and size of the fire keep changing in the area of the ROI so that can collect good quality fire data without additional labelling.

Fig. 2. The shape and size of fire in the region of interest (ROI) in the video

The principle of operation is simple. Comparing fire data in ROI with fire B/B(Bounding Box) of ROIs recognized as the learning model of Labeled data. If there is no difference of more than 50%, Using Pseudo labelling of the B/B as Guessed labelled. As a result, in this study, the recognition rate was improved by using Pseudo labelling methods to minimize misrecognition when predicting fire image data by estimating the Guessed Labelled data, which can occur when there is little initial labelled data close to the correct label.

The reason for applying Fine-tuning was to transform the architecture for image data of new purpose based on previously learned models and to update the learned model before.

4 Experimental Results

The research experiment on the method of generating fire data using the F-guess approach was conducted on a computer system with the following specifications: CPU: AMD Ryzen 7 3700X 8-Core Processor 3.6 GHz, GPU: NVIDIA GeForce RTX 8000, RAM 32GB [8]. Darknet 53 was used for learning, and Yolov4 was used for fire object detection.

Table 1. displays the number of fire data generation in every step and the number of whole images by using F-guessed and Semi-supervised learning. The 5,565 Answer labels used in initial learning are manually labelled (Labeled data) by a person and result in Guessed answer labels (Pseudo labeled) which use videos and images of unlabeled data displayed as F-guessed quantity. The number of F-guessed continues to increase because it is designed to re-label with final Weight values by reaping the Pseudo Labelled Guessed Labelled every step.

Table 1. Pseudo labeled data set augmentation information.

Data	Labeled Q'ty	Unlabeled Q'ty	F-guessed Q'ty	Division
Basic labeled data	5,565	0	0	image
1st augmentation	5,565	2,587	2,587	video
2nd augmentation	5,565	2,778	5,365	image
3rd augmentation	5,565	6,940	12,305	video
4th augmentation	5,565	8,182	20,487	image
5th augmentation	5,565	9,581	30,068	image

As a result, Table 2. shows an experiment comparing the data of 37,259 F-guessed completed based on 35,633 correct answer labels manually labeled by humans and 5,565 initial correct answer labels. As a result of the experiment, Loss was improved by 0.98, mIOU by 8.9%, and mAP by 4.95% compared to manual labeling by humans.

Table 2. Labeled and guessed labeled data comparison experiment tables.

Data	Q'ty	Loss	mIOU	mAP

Basic labeled data	5,565	3.347	52.23	65.93
Manual labeled	36,749	2.38	69.42	78.34
Guessed labeled	35,633	1.40	78.33	83.29

5 Conclusions

The research in this paper has suggested a comparison of the F-guessed method and the ROI method that complements the existing Pseudo-labelling method of semi-supervised learning [8] to facilitate data collection and labelling in special cases where data collection is limited, such as in the field of fire and disaster. As a result, compared to the initial correct label data, Loss decreased by up to 1.95%, mIOU increased by 26.1%, and mAP@0.5 improved by 17.36%. In addition, the human resources and time to get the correct answer label took about a month, based on manual labelling. However, after the change, it was significantly reduced to 48 hours. And the number of secured data also secured 35,633 correct answer data, which increased by 5.4 times based on the initial correct answer label data of 5,565. In the future, additional research will be processed to get additional data so that the pseudo-labelling position automatically follows the camera movement in the fire video.

Acknowledgment

This work was supported by the National Research Foundation of Korea(NRF) grant funded by the Korea government(MSIT)(No.RS-2023-00247045)

References

1. Amit Chaudhary homepage, https://amitness.com/2020/07/ semi-supervised-learning, last accessed 2023/05/07.
2. Ouali, Y., Hudelot, C., Tami, M.: An Overview of Deep Semi-Supervised Learning. Machine Learn., arXiv:2006.05278 (2020).
3. Kim, J.S., Kang, D.S.: A Study on Automatic Data Generation and Object Recognition Rate Improvement using Heterogeneous Object Detection and Duplicate Labeling Method. KIIT, Vol. 20, No. 5, pp. 21 - 29 (2022).
4. Vinko Kodžoman homepage, https://dataw hatnow.com/pseudo-labeling-semi-supervised-learning, accessed 2023/05/07.
5. Berthelot, D., Carlini, N., Goodfellow, I., Papernot, N., Oliver, A., Raffel, C.: MixMatch: A Holistic Approach to Semi-Supervised Learning. NeurIPS, 32, (2019).
6. Berthelot, D., Carlini, N., Cubuk, E. D., Kurakin, A., Sohn, K., Zhang, H., Raffel, C.: ReMixMatch: Semi-Supervised Learning with Distribution Alignment and Augmentation Anchoring, Machine Learn, arXiv:1911.09785 (2020).
7. Sohn, K., Berthelot, D., Carlini, N., Zhang, Z., Zhang, H., Raffel, C. A., Cubuk, E., Kurakin, D., Li, C.: FixMatch: Simplifying Semi-Supervised Learning with Consistency and Confidence. NeurIPS, 33, pp.596-608. (2020).

8. Kim, J.S., Kang, D.S.: A Study on Fire Data Generation and Recognition Rate Improvement using F-guessed and Semi-supervised Learning. KIIT, Vol. 20, No. 12, pp.123-134 (2022).