Exploring the Potential of Webcam-Based Eye-Tracking for Traditional Eye-Tracking Analysis

Cheng-Hui Chang[1] , Jason C. Hung[1], Jia-Wei Chang[1,*]

[1] Department of Computer Science and Information Engineering, National Taichung University of Science and Technology, Taichung City, Taiwan
w5566w8488@gmail.com, {jhung, jwchang*}@nutc.edu.tw

Abstract. Traditional eye-tracking systems can be costly and may pose a barrier to entry for researchers interested in studying gaze behavior. In recent years, there have been significant developments in simulating eye-tracking using webcams. However, little research has explored the use of webcam-based eye-tracking data for traditional eye-tracking analysis. In this paper, we propose a webcam-based eye-tracking system that utilizes an dilated convolutional neural networks to detect point of gaze and calculate a range of analysis indicators, such as duration of first fixation and latency of first fixation. By integrating these indicators, we aim to explore the potential of webcam-based eye-tracking for traditional eye-tracking analysis. This approach could significantly reduce the barrier to entry for researchers in the field of gaze behavior research and open up new avenues for studying gaze behavior.

Keywords: Eye tracking, Analysis indicators, Gaze estimation , Convolutional neural network (CNN).

1 Introduction

The saying "eyes are the windows to the soul" highlights the significant relationship between eye gaze and human cognitive processes. Eye tracking is the entry point to exploring this relationship and the field of eye tracking has made significant progress in recent years. The development of eye-tracking devices or "eye trackers" has played a critical role in improving the accuracy of eye movement tracking. Eye trackers have garnered widespread attention for analyzing and quantifying data and for related applications.

However, the eye-tracking system cannot be widely researched and applied because obtaining high-accuracy data requires powerful eye trackers, which are often expensive. This has made it difficult for novice scholars to enter this field. In the past, some researchers have proposed low-cost solutions to reduce the cost of eye-tracking equipment [3], laying the foundation for the development of eye-tracking devices. On the other hand, the proliferation of mobile devices has made camera lenses ubiquitous, attracting more people to this field, hoping that this technology can become a reality. The

advancement of computer vision in deep learning has also made it possible to use these everyday visible devices for eye tracking.

2 Related Work

In recent years, with the significant advancements in deep learning techniques, some researchers have proposed using deep learning to track eye movements through webcams and have developed related datasets [2]. These methods have been able to achieve high accuracy while running smoothly on personal computers or mobile devices, greatly increasing the accessibility of eye-tracking applications that use webcams.

Most deep learning-based eye-tracking methods have focused on appearance-based methods. For instance, in [1], researchers utilized a neural network to process facial images, allowing them to capture features even when the head moves slightly, thereby reducing errors. In [2], the authors used a similar concept but input both eye features and facial features, achieving high accuracy even without calibration.

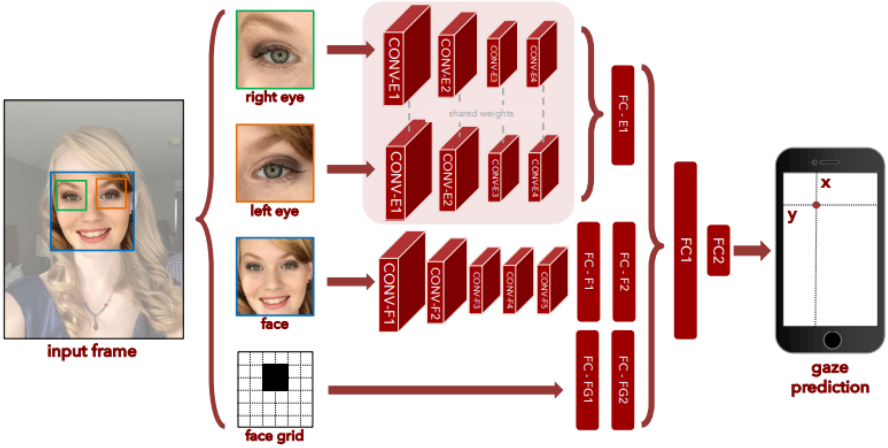

Fig. 1. The model architecture of the dilated convolutional neural network.

In the field of eye-tracking, deep learning methods have mainly been implemented using convolutional neural networks (CNNs), as in [4] and [5]. In [3], the authors proposed using dilated convolutions to reduce the loss of image information when capturing eye features. This method achieved 20.8% higher accuracy than models without dilated convolutions.

3 Design of the System

Although many studies have focused on using webcams to achieve tasks such as gaze estimation, there are few that use this data to perform traditional eye-tracking analysis metrics. This paper aims to investigate whether webcam-based eye-tracking can achieve such metrics.

We will employ the method of extended convolutional neural networks [3] to develop a webcam eye-tracking system that can detect the point of gaze and calculate various analysis metrics, such as Duration of First Fixation (DFF) and Latency of First Fixation (LFF).

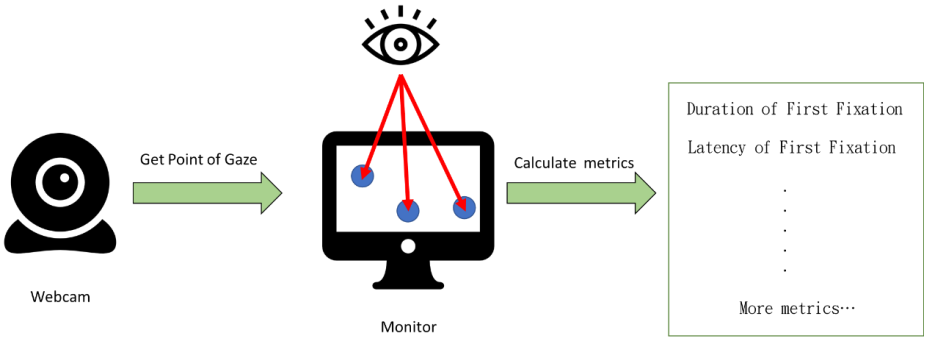

Fig. 2. The architecture and flow of the system design.

4 Conclusion

The high cost of traditional eye-tracking systems has been a barrier for many researchers to enter the field. In recent years, there has been significant development in simulating eye-tracking systems using webcams. This paper aims to explore the future of webcam-based eye-tracking analysis by combining various analysis indicators. By focusing on traditional eye-tracking indicators, we hope to discover new analysis methods and lower the barrier of entry to eye-tracking research using webcams.

5 Future work

Our next step in research will focus on advancing emotion recognition using webcam eye-tracking technology. Other studies [7] have analyzed various methods for using eye-tracking technology to analyze emotional states, and combining it with other physiological states can lead to even higher accuracy in evaluating gaze patterns and emotions. This will expand the application range of webcam eye-tracking technology.

Acknowledgement

This work was partially supported by the National Science and Technology Council, Taiwan, R.O.C. [grand number MOST 110-2221-E-025 -005].

References

1. Zhang, Xucong, et al. "It's written all over your face: Full-face appearance-based gaze estimation." Proceedings of the IEEE conference on computer vision and pattern recognition workshops. 2017.
2. Krafka, Kyle & Khosla, Aditya & Kellnhofer, Petr & Kannan, Harini & Bhandarkar, Suchendra & Matusik, Wojciech & Torralba, Antonio. (2016). Eye Tracking for Everyone.
3. Dongheng Li, Jason Babcock, Derrick Parkhurst, "openEyes: a low-cost head-mounted eye-tracking solution", https://doi.org/10.1145/1117309.1117350, pp. 95-100, Eye Tracking Research & Application, (2006)
4. Chen, Zhaokang, and Bertram E. Shi. "Appearance-based gaze estimation using dilated-convolutions." Asian Conference on Computer Vision. Springer, Cham, 2018.
5. C. Meng and X. Zhao, "Webcam-Based Eye Movement Analysis Using CNN," in IEEE Access, vol. 5, pp. 19581-19587, 2017, doi: 10.1109/ACCESS.2017.2754299.
6. Y. Yin, C. Juan, J. Chakraborty and M. P. McGuire, "Classification of Eye Tracking Data Using a Convolutional Neural Network," 2018 17th IEEE International
7. Zhan Z, Zhang L, Mei H, Fong PS. Online Learners' Reading Ability Detection Based on Eye-Tracking Sensors. Sensors (Basel). 2016 Sep 10;16(9):1457. doi: 10.3390/s16091457. PMID: 27626418; PMCID: PMC5038735.