APAMI2020 Poster Presentation Sessions | APAMI 2020 | Poster Presentation Sessions

Artificial Intelligence

Sun. Nov 22, 2020 3:00 PM - 4:00 PM Room E-2 (Congress center 5F - Conference Room 53)

[AP2-E2-4-01] Effects of the Sampling Frequency Change in Eye Movement Analysis Using Deep Convolutional Neural Network: Comparison with Other Analyses by Open Annotated Gaze Data

*Takayoshi Terashita[1], Tetsuo Sato[1], Shoko Tsutsumi[1], Mitsuru Sato[1], Toshihiro Ogura[1], Kunio Doi[1,2] （1. Graduate School of Radiological Technology, Gunma Prefectural College of Health Sciences, Japan, 2. Department of Radiology, University of Chicago, USA）

Keywords: Eye Movement Analysis, Sampling Frequency of Gaze Data, Deep Convolutional Neural Network

In our previous work, an eye movement analysis using a deep convolutional neural network was proposed. However, the low sampling frequency decreased the detection accuracy of eye events. The purpose of this study is to evaluate the effects of the sampling frequency change in our eye movement analysis. The gaze data of Lund University were used as open annotated data. It was then clarified that our method had high accuracy rates of approximately 90%. Moreover, the change in the accuracy was continuously high regardless of the sampling frequency in comparison to other methods.

©Japan Association for Medical Informatics

Effects of the Sampling Frequency Change in Eye Movement Analysis Using Deep Convolutional Neural Network: Comparison with Other Analyses by Open Annotated Gaze Data

Takayoshi Terashita[a], **Tetsuo Sato**[a], **Shoko Tsutsumi**[a], **Mitsuru Sato**[a], **Toshihiro Ogawa**[a] and **Kunio Doi**[a, b]

[a] *Graduate School of Radiological Technology, Gunma Prefectural College of Health Sciences, Japan*
[b] *Department of Radiology, University of Chicago, USA*

Abstract

In our previous work, an eye movement analysis using a deep convolutional neural network was proposed. However, the low sampling frequency decreased the detection accuracy of eye events. The purpose of this study is to evaluate the effects of the sampling frequency change in our eye movement analysis. The gaze data of Lund University were used as open annotated data. It was then clarified that our method had high accuracy rates of approximately 90%. Moreover, the change in the accuracy was continuously high regardless of the sampling frequency in comparison to other methods.

Keywords:

Eye Movement Analysis, Sampling Frequency of Gaze Data, Deep Convolutional Neural Network

Introduction

It is beneficial to elucidate the difference of features between the beginner and the expert in the training of an advanced skill. Many examples use eye movement features for advanced skill training, especially in medical image interpretation in healthcare education [1].

An eye movement analysis detects eye events from the gaze data acquired by an eye-tracker [2]. Fixation, a vital eye event, is defined as the state where the gaze is relatively stable at or around a single point. There are many detection techniques, such as manual and statistical detection. In addition, there are advanced techniques that use deep learning technology [3].

In our previous study, we suggested the eye movement analysis using a deep convolutional neural network (DCNN) [4]. It had a high detection accuracy and allowed eye events to be detected by classifying the images with a drawn path line of eye movements.

The effect of the sampling frequency change in detection techniques are a critical issue because it was known that the low sampling frequency decreased the detection accuracy of eye events generally [5]. The popularized eye-trackers, such as gaming input devices, are inexpensive and have low sampling frequencies under 100Hz. Thus, it was necessary to investigate the effects of the sampling frequency change in our method.

The purpose of this study is to evaluate the effects of the sampling frequency change in our eye movement analysis.

Materials and Methods

Materials

In this study, we adopted the open annotated gaze data from the Humanities Lab, Lund University [6]. This was focused on eye movements when viewing still images, videos, and moving dots, using a hi-speed eye-tracker with a sampling frequency of 500 Hz. Events had been segmented into five eye categories by two experts manually. These were fixation, saccade, post-saccadic oscillation, smooth pursuit, and blink. We selected the gaze data of still images because our target was medical images such as radiography, additionally limited to the annotations of one expert (RA) who had processed a large amount of sample. Eventually, 79,805 gaze points comprised the total gaze data from 16 participants. The gaze data of ten participants were separated for training the DCNN that was being developed, and residue data were used for the validation. The validation data of each sampling frequency were created by resampling at intervals for target frequency from raw data (Table 1). Annotations to created gaze data were reclassified into the fixations and the non-fixations because it was challenging to detect detailed eye events from data of the low sampling frequency.

Table 1- Properties of each sampling frequency in the validation data

Sampling frequency [Hz]	Intervals [ms]	Number of gaze points	Number of fixations
500	-	29,332	20,518
250	2	14,693	10,267
125	6	7,359	5,138
100	8	5,905	4,112
50	18	2,966	2,065
25	38	1,488	1,043

Eye movement analysis using the DCNN

Path line images were created by drawing path lines that connected short consecutive gaze points [4]. Those gaze points were extracted by a window size of 100 ms, because the fixation was defined as maintaining a gaze for ≥ 100 ms [2]. Other settings of path line images were as follows; image size, 28 × 28 pixels; compressed the coordinate of gaze points to 1/8;

allocated the center of gravity of those gaze points to the center of the image; three pixels width and the grayscale color of path line. The supervised labels were determined to the annotations at the central point of gaze points. Those procedures were conducted in the training data and the validation data of each sampling frequency. However, the number of fixations were three times larger than that of non-fixations, and this bias was inappropriate with training a neural network. Thus, the number of fixations and non-fixations in the training data were equalized by selecting randomly from fixations. The parameters of our DCNN were learned by this training data. The model of our DCNN consisted of four convolution layers, two max-pooling layers, three dropout layers, a global average pooling layer, and a full-connected layer [4]. *TensorFlow* 1.12.0 was used as the deep learning framework and *Python* 3.5.6 as the programming language.

Comparison with other eye movement analyses

Our eye movement analysis using DCNN was compared with other methods. The "*emove*" for R package was adopted as the dispersion technique [7]. A dispersion technique is a conventional method for eye movement detection whereby distances are calculated from x and y coordinates of visual points when in the threshold range. The "*gazepath*" for R package was adopted as the velocity technique [8]. A velocity technique is a conventional method for eye movement detection when distances per unit of time exceed the threshold. The Startsev's method (hereafter called "*BLSTM*") was adopted as a method using deep learning published in the current journal [3]. It is a highly accurate method and is that the eye movements are detected by qualifying the alterations of x and y coordinates of gaze points as time-series data, using a technique of long short-term memory in a recurrent neural network.

The F score of fixation was used as an evaluation index and was calculated from the confusion matrix of the predictions from each eye movement analysis and the ground truth, which was the annotations by a Lund University's expert.

Results and Discussion

Figure 1 shows a relationship between the F score of fixation and the sampling frequency in each method. The scores of the emove and the gazepath remained constant up to the medium sampling frequency. However, it did decrease at the low sampling frequency, and a theoretical result was obtained. The BLSTM showed the highest score at 93% at a sampling frequency of around 200 – 100 Hz. Although it had the lowest score at 53% at the low sampling frequency. Furthermore, because the publicly available parameters of the BLSTM were learned by 250 Hz gaze data, the score could be improved by training the parameters according to the sampling frequency.

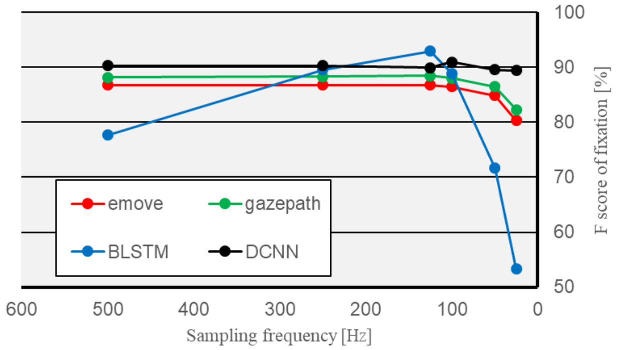

Figure 1- The relationship between the F score of fixation and the sampling frequency in each method

In contrast, the scores of DCNN were stable at around 90% regardless of the sampling frequency and were higher than other methods at the high and low sampling frequency. In our eye movement analysis using DCNN, the eye events were detected by classifying the images in drawn path lines of gaze points. That is identical to the data interpolation because the drawing of the path line filled intervals between gaze points. Therefore, the F score of fixation in our method could not be decreased, even in the sparse data.

Conclusion

In this study, the effects of the sampling frequency change were clarified for our eye movement analysis using DCNN. Compared with the other methods, it had stable and high accuracy, as well as the best performance, regardless of the sampling frequency. Although, the accuracy was even more profound, especially at the low sampling frequency.

Acknowledgments

This work was supported by JSPS KAKENHI grant number 18K09951.

References

[1] Kundel HL, Nodine CF, Carmody D. Visual scanning, pattern recognition and decision-making in pulmonary nodule detection. *Invest Radiol.* 1978; 13:175-81.

[2] Kenneth H. Eye Tracking: a comprehensive guide to methods and measures. Oxford University Press. 2011.

[3] Startsev M, Agtzidis I, Dorr M. 1D CNN with BLSTM for automated classification of fixations, saccades, and smooth pursuits. *Behav Res Methods.* 2019; 51:556-72.

[4] Terashita T, Sato T, Tsutsumi M, *et al.* Eye movement analysis using deep learning for medical image interpretation. *ECR.* 2019; C-0155: DOI: 10.26044/ecr2019/C-0155.

[5] Zemblys R, Niehorster DC, Komogortsev O, Holmqvist K. Using machine learning to detect events in eye-tracking data. *Behav Res.* 2018; 50:160-81.

[6] Larsson L, Nystrom M, Stridh M. Detection of saccades and post saccadic oscillations in the presence of smooth pursuit. *IEEE Trans Biomed Eng.* 2013; 60(9):2484-93.

[7] Salvucci DD, Goldberg JH. Identifying fixations and saccades in eye-tracking protocols. Proceedings of the 2000 symposium on Eye-Tracking research & applications 2000: pp.71-8.

[8] van Renswoude DR, Raijmakers MEJ, Koornneef A, *et al.* Gazepath: An eye-tracking analysis tool that accounts for individual differences and data quality. *Behav Res.* 2017; DOI: 10.3758/s13428-017-0909-3.

Address for correspondence

Takayoshi Terashita
Graduate school of radiological technology
Gunma prefectural college of health sciences, Japan
E-mail: terapist@gchs.ac.jp