Journal of Soft Computing and Artificial Intelligence

Journal homepage: https://dergipark.org.tr/en/pub/jscai

International Open Access

Volume 03
Issue 02

December, 2022

Research Article

Real-time Iris Center Detection Based on Convolutional Neural Networks

Kenan DONUK[1], Davut HANBAY[2]

[1] *Cizre Vocational School, Computer Programming Department. Sirnak University, 73200, Sirnak, Türkiye*
[2] *Department of Computer Engineering, Inonu University, 44280, Malatya, Türkiye*

ARTICLE INFO

Article history:
Received December 8, 2022
Revised December 15, 2022
Accepted December 21, 2022

Keywords:
GI4E
CNN
Pupil center detection
Iris center detection

ABSTRACT

It is an active field of study in studies where the iris center is referenced, such as iris center detection, gaze tracking, driver fatigue detection. In this study, an approach for real-time detection of iris centers based on convolutional neural networks is presented. The GI4E dataset was used as the dataset for the proposed approach. Experimental results estimated the test data of the proposed convolutional neural network model with an accuracy of 97.2% based on the 0.025 error corresponding to the closest position to the iris center according to the maximum normalized error criteria. The study was also tested in real time with a webcam built into the computer. While the test accuracy is satisfactory, real-time speed performance needs to be improved.

1. Introduction

The eyes, one of the complex organs of the human body, are an organ located in the eye socket, providing vision with its spherical structure. The visible part of the eye consists of three parts, the sclera, iris, and pupil. The sclera forms the white area of the eye. The iris acts as a diaphragm, which forms the middle layer of the eye, whose color varies from person to person, and which determines the size of the pupil in the middle. The muscle fibers, which are positioned in a circle around the iris, cause the pupil to contract when they contract, adjusting the amount of light entering the eye according to the changing environmental conditions. Vision occurs when light is refracted from the cornea and transmitted to the lens/pupil and from there to the retina (fovea).

Detection of the iris center or pupil is an important field of study in many different sectors such as gaze tracking [1], [2], driver fatigue [3], medical research [4]. To be able to detect the iris center, first of all, face and eye detection should be done on the images. For face and eye detection, there are ready-made algorithms such as DLIB [5], [6], MTCNN [7], OpenCv [8] in the literature. Some studies in the literature for pupil/iris center detection after eye detection are as follows.

Yu et al. proposed a robust iris center detection method for eye gaze tracking in their study. The method they propose is based on the geometric relationship between the iris center and the eye corners. It also presents an algorithm that predicts iris edge points to overcome the occlusions in the iris region because of eye movements [1]. Donuk and Hanbay in their study, they proposed a pupil center detection method based on U-Net architecture. They trained the proposed Mini-Unet architecture with the GI4E dataset. The test performance of the architecture detected a position close to the center of the pupil with an accuracy of 98.40% based on an

[1]Corresponding author
e-mail: kenandonuk@sirnak.edu.tr
DOI: 10.55195/jscai.1216384

error value of 0.025 [9]. Lee et al. presented a fast-acting method of pupil center detection in robust and real-time applications. They used non-local block and self-attention block mechanisms in their proposed method. Non-local block mechanisms have been applied to reduce the delay in face detection, and self-attention block mechanisms have been applied to provide good image quality for glasses removal in images with glasses. The test success of the method on the GI4E dataset is 99.84% [10].

2. Material and Method

2.1. Dataset: GI4E

GI4E is a publicly available dataset of 103 user images taken via webcam for Iris center detection. These images were obtained from 12 different images with a resolution of 800x600 pixels for each user and looking at different points on the screen. Alongside the users' images in the dataset, there is a manually labeled guide containing the coordinates of the iris centers and eye corners of the eyes in the images in pixels [11].

2.2. Dataset preprocessing

All images of the GI4E dataset have been converted to gray format and normalized for ease of data manipulation and Convolutional Neural Network (CNN) training. The eyes in the images were separated from the rest of the image by using the eye corner coordinates in the data set. The new iris center coordinates (x, y) of the separated eye images were recalculated using the eye corner coordinates. A total of 2472 eye images, 2 of each image, and the new iris center coordinates of the eye images are divided into 80% training and 20% test data for the training of the CNN network. In Figure 1, the process on the image in the data set is visualized.

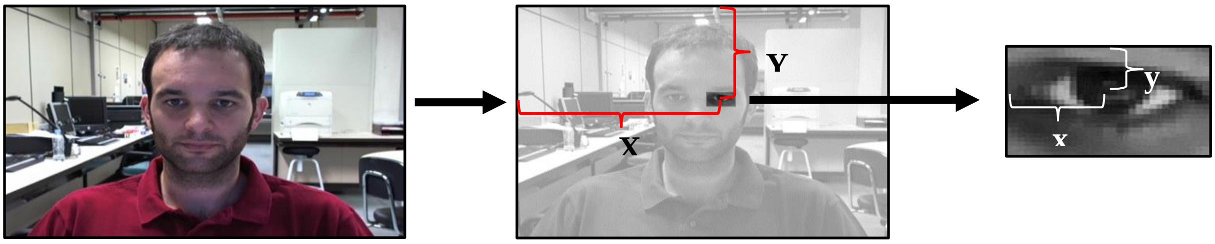

Figure 1 Dataset preprocessing

2.3. Proposed Method and Training

With data preprocessing, 2472 right and left eye images obtained from the GI4E data set were converted to 32x32 spatial dimensions in accordance with the input of the CNN model. In the CNN model, firstly, three convolutions (3x3) filters were applied to the input data, respectively. After the second and third convolution layers, the size is reduced by applying maxpooling. Then, the obtained features are flattened and transmitted to the full connected layer of the CNN model. The full connected layer consists of 128, 64 neurons and classification layers, respectively. It has been tried to increase the generalization ability of the network by applying 2 dropouts (0.4) in these layers, respectively. “BatchNormalization” and “ReLU” were applied as normalization and activation functions in all layers, respectively. Finally, as the learning optimization of the network, the “Adam” optimization algorithm and for loss detection the euclidean_distance function, which expresses the distance between the real iris center coordinates and the iris center coordinate values estimated by the CNN network, are used. In this function, p and q represent the actual and estimated iris center coordinates, respectively. The distance between p and q is given by Equation 1.

$$d(p,q) = \sqrt{\sum_{i=1}^{n}(q_i - p_i)^2} \qquad (1)$$

The proposed network reached the minimum error value in 400 iterations. The proposed CNN structure is shown in Figure 2. The accuracy and loss graphs of the proposed network are shown in Figures 3 and 4, respectively. When the loss and accuracy graphs are examined, we can see that the network is learning properly.

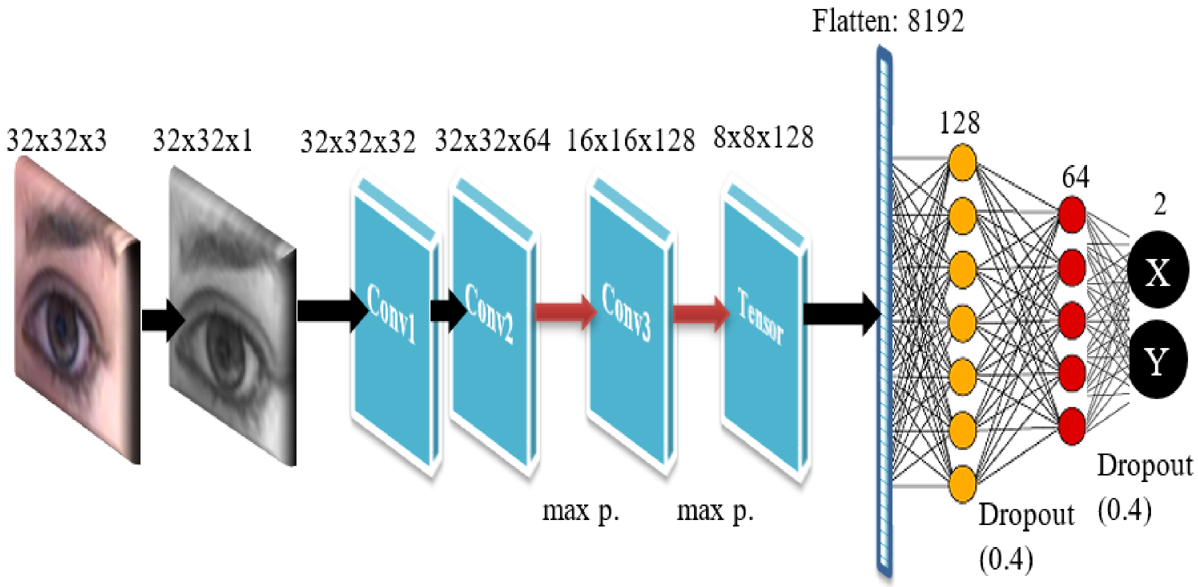

Figure 2 Proposed CNN structure

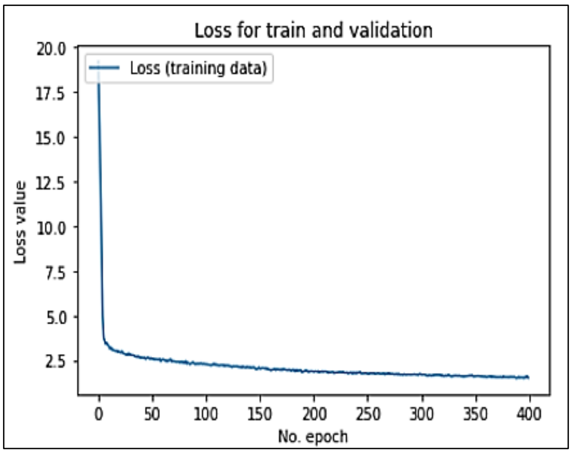

Figure 3 Loss graph

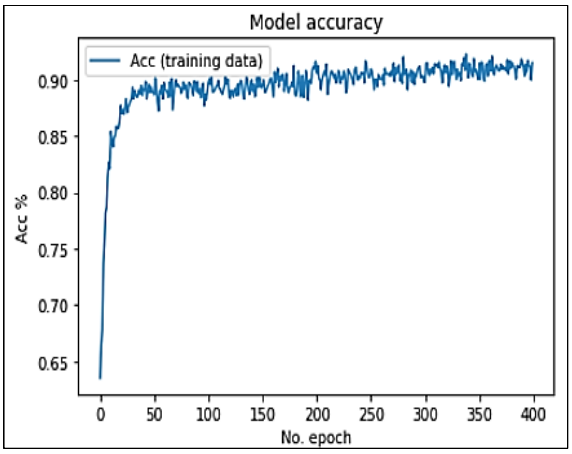

Figure 4 Accuracy graph

3. Experimental Results

To evaluate the performance of the trained network with the right and left eye images obtained from the data set and the coordinate labels corresponding to these images, an estimation was made with the 20% test data of the data set that was not used in the training. For the analysis of the estimation performance of the network, the maximum normalized error value criterion, which is widely used to measure the pupil center estimation performance, is used [12]. The formula for this error criterion is given in Equation 2.

$$d_{eye(error)} = \frac{max\,(d_l,d_r)}{\|C_l - C_r\|} \quad (2)$$

Given in the equation, d_l is the distance between the true iris center of the left eye and the iris center predicted by the CNN model, and d_r is the distance between the true iris center of the right eye and the iris center predicted by the CNN model. Error detection is obtained by the ratio $\|C_l - C_r\|$ of the value greater than these two values to the actual distance between the iris centers of the right and left eyes. If the obtained $d_{eye(error)}$ value is less than 0.025 error value, the estimated iris center coordinate represents the closest value to the real iris center coordinate. If the error value is less than 0.05, the estimated iris center coordinate obtained represents a coordinate approximately within the actual boundaries of the pupil [9], [13]. In Figure 5, the detected iris center detected manually in the data set and the estimated iris center are marked on the eye image in white and red, respectively.

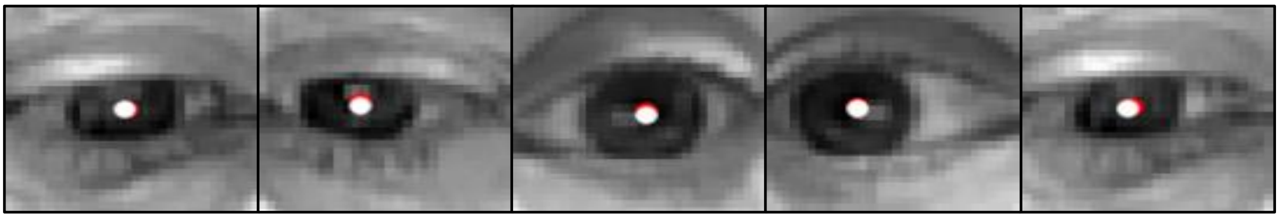

Figure 5 Actual and predicted iris center positions

The prediction results of the proposed CNN model on 500 test data obtained an accuracy of 97.2% according to 0.025 error and 99.8% according to 0.05 error from the maximum normalized error criterion. The proposed system has also been tested in real time with a webcam. In Figure 6, images of the real-time iris center detection application are given. For real-time application, first the face and then the eyes are detected from the webcam via OpenCv. After gray formatting and normalization, the dimensions of the eye regions are converted to 32x32 spatial dimensions. The coordinates of the iris center are determined by giving the obtained instant eye images as input to the trained CNN model. However, these coordinates do not reflect the actual coordinate values because the spatial dimensions of the real eye image have been changed for CNN. Therefore, new iris center coordinate values are obtained by applying the inverse of the size change for the CNN input of the eye images to the detected coordinate values. The actual coordinate values obtained are marked on the face.

Figure 6 Real-time iris center detection

4. Conclusion

The proposed model revealed an accuracy of 97.2% for the GI4E dataset. Although the model gives good results in forward-facing eye positions, iris center detection results in higher errors in different face and eye poses (right, left, down and up). For more robust iris center detection, the generalization ability of the model can be increased by using datasets with different eye positions. Also, the proposed system is slow for real-time applications. This is because the system first detects face and eye in real-time applications. The resulting eyes are then sent to the CNN model for iris center detection. This process causes the system to experience lag. Therefore, this delay can be overcome with face and eye detection algorithms that promise high speed in the literature and deep learning architectures that will reduce the cost for the CNN model.

References

[1]. M. Yu *et al.*, “An easy iris center detection method for eye gaze tracking system,” *J Eye Mov Res*, vol. 8, no. 3, 2015, doi: 10.16910/JEMR.8.3.5.

[2]. K. Donuk and D. Hanbay, “Video Based Real-Time Eye Tracking,” in *In: 28th IEEE Signal Processing and Communications Applications (SIU-2020)*, 2020, pp. 21–24.

[3]. Q. Zhuang, Z. Kehua, J. Wang, and Q. Chen, “Driver fatigue detection method based on eye states with pupil and iris segmentation,” *IEEE Access*, vol. 8, pp. 173440–173449, 2020, doi: 10.1109/ACCESS.2020.3025818.

[4]. H. Yan and Y. Zhang, “Detection of the pupil eigenvalues in medicine,” *Proceedings - 2010 International Conference on Computational and Information Sciences, ICCIS 2010*, pp. 989–992, 2010, doi: 10.1109/ICCIS.2010.244.

[5]. “dlib C++ Library.” http://dlib.net/ (accessed Dec. 08, 2022).

[6]. V. Kazemi and J. Sullivan, “One millisecond face alignment with an ensemble of regression trees,” in *2014 IEEE Conference on Computer Vision and Pattern Recognition*, Jun. 2014, pp. 1867–1874. doi: 10.1109/CVPR.2014.241.

[7]. K. Zhang, Z. Zhang, Z. Li, and Y. Qiao, “Joint Face Detection and Alignment Using Multitask Cascaded Convolutional Networks,” *IEEE Signal Process Lett*, vol. 23, no. 10, pp. 1499–1503, Oct. 2016, doi: 10.1109/LSP.2016.2603342.

[8]. “Home - OpenCV.” https://opencv.org/ (accessed Dec. 08, 2022).

[9]. K. Donuk and D. Hanbay, “Pupil Center Localization Based on Mini U-Net,” *Computer Science*, pp. 185–191, Oct. 2022, doi: 10.53070/BBD.1173482.

[10]. K. il Lee, J. H. Jeon, and B. C. Song, “Deep Learning-Based Pupil Center Detection for Fast and Accurate Eye Tracking System,” *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, vol. 12364 LNCS, pp. 36–52, 2020, doi: 10.1007/978-3-030-58529-7_3/COVER.

[11]. A. Villanueva, V. Ponz, L. Sesma-Sanchez, M. Ariz, S. Porta, and R. Cabeza, “Hybrid method based on topography for robust detection of iris center and eye corners,” *ACM Transactions on Multimedia Computing, Communications, and Applications (TOMM)*, vol. 9, no. 4, Aug. 2013, doi: 10.1145/2501643.2501647.

[12]. O. Jesorsky, K. J. Kirchberg, and R. W. Frischholz, “Robust Face Detection Using the Hausdorff Distance,” 2001, pp. 90–95. doi: 10.1007/3-540-45344-X_14.

[13]. A. Larumbe-Bergera, G. Garde, S. Porta, R. Cabeza, and A. Villanueva, “Accurate Pupil Center Detection in Off-the-Shelf Eye Tracking Systems Using Convolutional Neural Networks,” *Sensors*, vol. 21, no. 20, p. 6847, Oct. 2021, doi: 10.3390/s21206847.