Real-time smart driver sleepiness detection by eye aspect ratio using computer vision

Simon Chong Kai Yuen[1], Noor Hidayah Zakaria[1], Goh Eg Su[1], Rohayanti Hassan[1], Shahreen Kasim[2], Tole Sutikno[3,4]
[1]Faculty of Computing, Universiti Teknologi Malaysia, Johor Bahru, 81310, Johor, Malaysia
[2]Soft Computing and Data Mining Centre, Faculty of Computer Science and Information Technology, Universiti Tun Hussein Onn Malaysia, Parit Raja, Batu Pahat, 86400, Johor, Malaysia
[3]Master Program of Electrical Engineering, Faculty of Industrial Technology, Universitas Ahmad Dahlan, Yogyakarta, Indonesia
[4]Embedded Systems and Power Electronics Research Group, Yogyakarta, Indonesia

Article Info

Article history:

Received Oct 31, 2023
Revised Mar 4, 2024
Accepted Mar 15, 2024

Keywords:

Driver sleepiness
Eye aspect ratio
Facial detection
OpenCV
Real-time processing

ABSTRACT

The purpose of this study is to determine the optimal eye aspect ratio (EAR) for a prototype capable of using computer vision techniques to detect driver sleepiness based on eyelid size changes. The prototype, developed with Raspberry Pi and OpenCV, provides a real-time evaluation of the driver's level of alertness. The prototype can accurately determine the onset of sleepiness by monitoring and detecting instances of prolonged eyelid closure. Due to the fact that the eye aspect ratios of different individuals vary in size, the system's accuracy may be compromised. For the first experiment, the research focuses on determining the optimal EAR threshold of the proposed prototype using a sample of 20 participants ranging in age from 20 to 30, 31 to 40, and 41 to 50 years old. The study also examines the effects of various environmental conditions, such as dark or nighttime settings and the use of spectacle. The optimal EAR threshold value, as dedicated by the first experiment, is 0.225 after testing 20 participants with and without eyeglasses in low and bright lighting and 7 participants with a 0.225 EAR threshold in dark and bright lighting environments. The result shows 100% precision.

Corresponding Author:

Noor Hidayah Binti Zakaria
Faculty of Computing
Universiti Teknologi Malaysia
Johor Bahru, 81310, Johor, Malaysia
Email: noorhidayah.z@utm.my

1. INTRODUCTION

Sleepiness is one of the primary causes of road accidents, accounting for 21% [1] of car crashes in 2015. Road accidents may be caused by several factors, such as being drunk, reckless driving, speeding, distraction, a low lighting environment, and so on [2]. There are several methods to avoid or prevent this from happening with the aid of technology. For example, eyelid size detection, the number of yawnings, the number of eye blinks, and so on. Besides, yawning is one of the symptoms of sleepiness. It is possible to detect the number of yawnings and blinks to determine if someone is feeling sleepy. A study found that aged subjects yawn less frequently than young adults, particularly during morning and mid-afternoon hours [3].

The simplest way to determine if the driver is feeling sleepy is to detect the eyelid size by reading the eye aspect ratio, also known as EAR, in real-time with a camera. However, different people will have different eyelid sizes, especially among different age groups [4]. Therefore, this study is to determine the optimal eye aspect ratio for the majority of users with a developed prototype. The prototype was developed with a microcomputer (a

Raspberry Pi) attached to a night vision camera as its input and connected to a speaker as its output. When a driver closes their eye for approximately one second, the prototype assumes the driver is asleep and in a dangerous situation and plays an alarm sound accordingly to seek the attention of the driver. This invention is to prevent the driver from losing their eyesight on the road and falling asleep without awareness while driving.

OpenCV (Open-Source Computer Vision), an efficient open-source computer vision library [5], [6], and BSD (Berkeley Software Distribution)—licensed software [6], [7], were utilized in the prototype to achieve the research objectives. OpenCV allowed the computer to “see” and analyze the real-world environment by mimicking the human eye [8]. The main feature of the prototype is to read the eyelid size of the driver by obtaining and calculating the EAR. In this study, there were two experiments carried out to determine the optimal EAR threshold for most of the users and evaluate the sensitivity of the optimal EAR threshold. The rest of the paper will be discussed as follows: Previous works and technologies used were reviewed in Section 2.0. Section 3.0 explained the methodology of the prototype and how the experiment was being conducted. Section 4 explained and discussed the results of the experiments.

2. PREVIOUS WORKS

The previous study [9] proposes an image detection drowsiness system to detect the state of a car driver using the Eye Aspect Ratio (EAR) technique. The system is equipped with a Pi camera, a Raspberry Pi 4, and a GPS module to detect and analyze the state of eye closure in real-time. The system is able to recognize whether the driver is drowsy or not with an accuracy of 90% under various conditions such as wearing spectacles, dim light, and microsleep.

Another study [10] discusses detecting driver drowsiness based on adaptive thresholds using EAR and MAR. The system detects the face and determines the facial landmarks by which it can compute EAR and MAR to detect driver drowsiness based on an adaptive threshold. The system also uses head pose estimation to check the attention of the driver’s head with respect to the road. When the system detects that the driver is drowsy, it raises an alarm.

Chandiwala and Agarwal [11] proposed a system that can detect driver drowsiness in real time and issue a timely warning. The method is non-invasive and uses the Eye Aspect Ratio (EAR) and Mouth Aspect Ratio (MAR) to quantify drowsiness by locating, monitoring, and analyzing both the eyes and mouth, even in the dark. Comparing the real-time EAR to the initial EAR of the drive and the MAR to a threshold of 20. The system notifies the user via text and audio alerts. A smart alarm system that can be deactivated by detecting a hand gesture is also proposed. To test the system, data is collected in real-time from the user. This system accurately predicted nine out of ten cases in both dark and light conditions, with various facial features, including glasses.

The other study [12] proposed that, utilizing image processing techniques, the system captures video and identifies the driver's face in every frame. The system is able to detect facial landmarks, computes Eye Aspect Ratio (EAR) and Eye Closure Ratio (ECR), and uses adaptive thresholding to detect driver drowsiness. Using machine learning algorithms, the effectiveness of the proposed method has been evaluated. Using a random forest classifier, empirical results demonstrate that the proposed model can achieve an accuracy of 84%.

Additionally, this study [13] presents a method for detecting eye blinks in a series of real-time videos captured by a car dashboard camera. The suggested technique determines the facial landmark positions for each video frame and then extracts the vertical distance between the eyelids from the facial landmark positions. The algorithm that has been proposed estimates the facial landmark positions, extracts a single scalar quantity by making use of Eye Aspect Ratio (EAR), and identifies the eye closeness in each frame. In the end, blinks are recognized by employing the modified EAR threshold value in conjunction with a pattern of EAR values in a relatively short period of time. Experimental evidence indicates that the greater the EAR threshold, the worse the AUC’s accuracy and performance. Further, 0.18 was determined to be the optimum EAR threshold in their research.

2.1. Extraction of Face Feature

Only the driver’s eye information needed to be extracted instead of the entire driver’s face. After utilizing OpenCV to obtain frame-by-frame images containing human faces, dlib will be used to extract the characteristics of the faces. It includes a wide range of machine learning algorithms as well as numerous graphical model algorithms. It achieved an accuracy of 99.38% [14], [15] on the outdoor face detection database benchmark. The obtained facial information of humans will be converted into 68 points, as shown in Figure 1, and accessed with an array. In this study, only the eye data point is required, not the entire face. Therefore, the points between 37–42 (right eye) and 43–48 (left eye) were extracted for further processing. The rest of the points were omitted due to unnecessary data in this research.

2.2. Image pre-processing

Imutils is a collection of convenient features for performing fundamental image processing operations using OpenCV and Python. These operations include translation, edge detection, rotation, scaling, skeletonization, contour sorting, and displaying images using Matplotlib [7, 16]. Using Imutils makes these tasks significantly easier compared to working with OpenCV alone. In this research, once the driver's face is captured, it is used to convert the video into grayscale in terms of calculating the pixel value. The color does not make sense in this research. And also, it is used to convert the extracted landmark into a NumPy [17] array for the next step.

2.3. Eyes Aspect Ratio

As mentioned previously, only the eye information will be extracted in this research; these eye landmarks are retrieved by using an array to calculate the EAR value for identifying whether the driver falls asleep. The algorithm for getting the EAR value is as shown in Figure 2 [19]. The sum of the vertical distance will be calculated and divided by the horizontal distance multiplied by two. The EAR would drop to 0 when the eye closed, while rising to 3 when the eye open.

The equations (1) and (2) are the algorithms to obtain the EAR value from the human's eye. The sum of the vertical distances from the upper and lower eyelids will be calculated as (P2 + P6) + (P3 + P5). Once the vertical distance is obtained, divide it by two times the horizontal distance (P4 minus P1). So now the EAR value is obtained for one eye. In order to have an average EAR value for both eyes, just sum up two EARs, then divide by 2.

$$\text{EAR} = \frac{Sum\ of\ vertical\ distance}{2 * Horizontal\ distance} \tag{1}$$

$$\text{EAR} = \frac{(\text{P2}-\text{P6})+(P3-P5)}{2\ (P1-P4)} \tag{2}$$

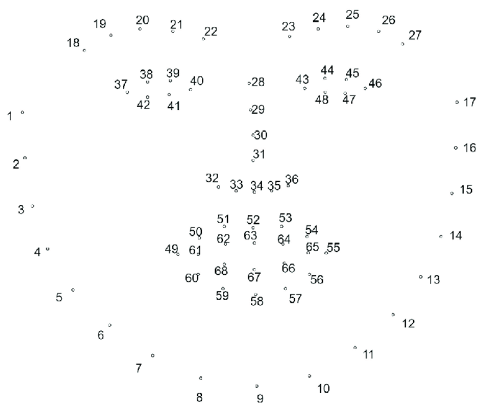


Figure 1. The 68 landmark points on the face [18]

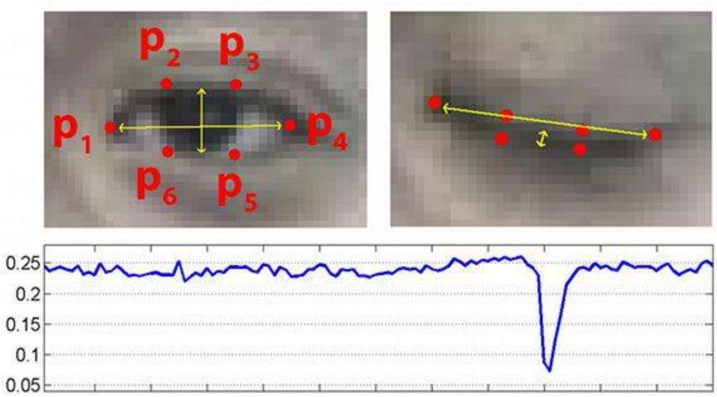
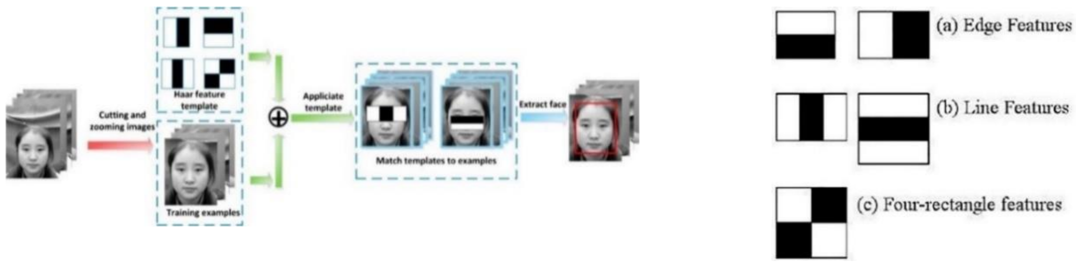


Figure 2. EAR of open and close eyes [19]

2.4. Algorithm of the pre-trained model

The Haar classifier algorithm detects objects. The characteristics of the image will be extracted in order to detect the item and determine what it is. Haar Cascading is a machine learning approach that involves drilling a classifier from a large number of positive and negative images [2, 19–23]. Object detection classifiers are implemented using Haar feature-based cascade classifiers, as shown in Figure 3. This classifier follows a machine learning process in which a cascade operation from the photographs is instilled to discover objects in further photos, as shown in Figure 3(a). Face recognition and facial expression identification in images are also successful. The session concludes with the classifier being shown positive and negative images. The properties are then extracted from the image, as shown in Figure 3(b). By dividing the sum of the pixels in the white and black rectangles, each characteristic has a distinct value. There are three features or masks in this recognition process, which are edge, line, and four rectangles [9, 23]. These various features have different functionality or results depending on the image pixel.



(a) (b)

Figure 3. Haar Cascade face recognition of (a) process of face recognition and (b) the kernel of Haar Cascade

3. METHOD

The methodology of development is shown in Figure 4, since all the functions or features have to be completed before we start testing. Communication or user involvement is unnecessary in this prototype development process. The current development status and progress in each phase of the research methodology are explained further in the following sub-sections. Also, two experiments have been carried out to determine the optimal EAR threshold for different age categories, with and without a spectacle, and in dark and bright lighting environments.

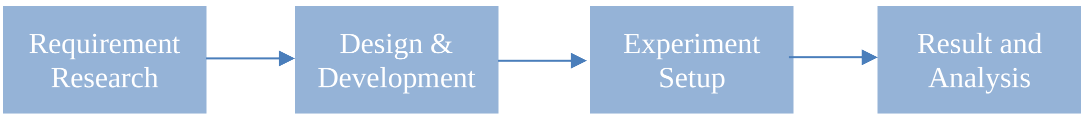


Figure 4. Methodology of this research

3.1. Requirement Research

Prior to commencing the experiment, limited research, studies, and discussions were conducted, focusing on the EAR algorithm, the technology required for the experiment, and cost-effective hardware and software. The Raspberry Pi is selected for its high portability and affordability. OpenCV is open-source, allowing for easy editing and sharing, along with numerous documentation and references. The programming language chosen was Python due to its abundance of high-functionality libraries and its widespread popularity, particularly for tasks related to artificial intelligence and computer vision.

3.2. Design and Development

Figure 5 shows the workflow of the prototype. Firstly, the camera was scanning and screening to see if a face was detected; if so, it proceeded to the next step. Once the human face is shown in the camera area, the data is converted into 68-point facial landmarks, and then only left eye and right eye data are extracted for further processing. Next, the size of both eyes was calculated based on the EAR algorithm in Figure 2, and an average value was obtained from both eyes. Lastly, if the result of the average value is below the threshold and persists for more than 1 second, the system plays an alarm sound, indicating sleepiness is detected.

A straightforward user interface was developed in order to facilitate the process of conducting our experiments and ensure that we have accurate records, as shown in Figure 6. On the user interface, a number of important and helpful pieces of information were displayed. Figure 6(a) shows that the eye is still open, and Figure 6(b) shows sleepiness detected. These included the EAR value, the number of times the eye was closed, the frame rate, and a warning message that appeared when the user was detected to be sleepy.

3.3 Experiment Setup

In this experiment, the duration of the closing eye is set as a constant because the human eye-blinking speed is 572 ± 25 ms [24]. Looking away from the forward roadway for more than two seconds can be dangerous [25]. As a result, this research prototype (shown in Figure 7) is set to one second, which means that if the driver closes their eyes for longer than one second, an alarm sound is played. The duration is not measured in seconds, but in video frames; the frame changes from time to time within the range of 8-12 FPS; thus, the conversion from frame to second is performed to consistently record the experimental data. For this purpose, at 8 FPS, the prototype's fastest time is processing 10 frames, indicating a duration of 0.833 seconds, which is still longer than the human maximum blinking time of 0.597 seconds and meets the needs and requirements of this experiment. Hence, the threshold of the frame is set to 10.

Eyelid size can vary depending on age and size, resulting in varying EAR thresholds among users. During the first experiment, 20 participants were tested with three variables to determine the optimal EAR threshold. The initial variable includes three age groups: 20-30 (12 participants), 31-40 (6 participants), and 41-50 (6 participants). The second variable contains four different EAR threshold parameters: 0.18, 0.2, 0.225, and 0.25. Some of them looked good wearing spectacles, while others did not. The third variable entails testing the prototype in both bright and dark environments to determine its performance under low-light conditions. When a participant closed their eyes, the prototype manually measured the time until the alarm sound played. The assessment was carried out to ensure that the research prototype functioned properly within the specified timeframe of about one second, under the various conditions previously mentioned. These conditions include a variety of age groups and lighting settings. Three different testing parameters are used to calculate the average time for all participants at each EAR threshold: 1) EAR threshold, 2) age category, and 3) lighting conditions. When the prototype detects the participant's eyes closing, it sounds an alarm that lasts longer than the longest human blink, confirming its proper functionality. The shortest time will be recorded.

Seven participants had been tested in the second experiment; the determined optimal EAR threshold from experiment 1 would be used to test its sensitivity in this experiment. Participants are required to open and close their eyes for ten seconds in both dark and bright light environments and without wearing spectacles to see if the built prototype can consistently be detected under the selected optimal EAR threshold. Figure 7 (a) and (b) show the experiment setup in a discussion room with a participant.

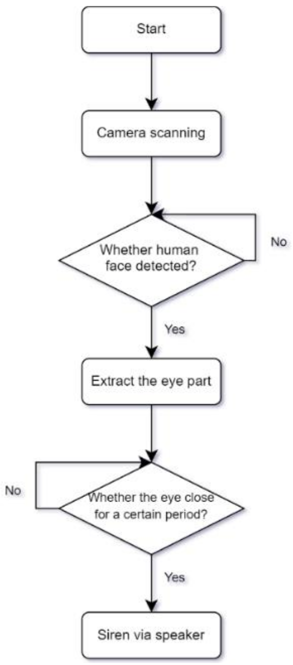
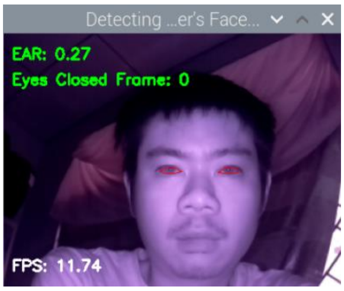


Figure 5. Workflow of the developed prototype



(a)

(b)

Figure 6. The user interface of (a) open eyes and (b) closed eyes

(a) (b)

Figure 7. Experiment setup, (a) experiment, and (b) lighting setup

4. RESULTS AND DISCUSSION

Figure 8 shows the average time taken by the research prototype for the age 20–30 category. Figure 8 (a) shows the average time taken by the research prototype for the age 20–30 category in a dark or low-light environment. There were 12 participants in this category. From the result, the EAR threshold with 0.25 has the fastest reaction time among all 4 thresholds (0.18, 0.2, 0.225, and 0.25), but it has only 7 out of 10 valid data points (false negative and false positive data were counted as invalid), while 0.225 has the second fastest time consumed by the research prototype to react, but it has 9 out of 10 valid data points. Therefore, 0.225 was chosen with consideration of the time and number of valid data points. In Figure 8 (b), the same participant, but tested in a bright environment, had the fastest reaction by the research prototype with 8 valid data points. While 0.18 had the second highest among all thresholds, it only had 5 out of 10 valid tested data. While the 0.225 EAR threshold had the highest valid data, which was 9 out of 10, Therefore, the 0.225 threshold has been chosen for this category.

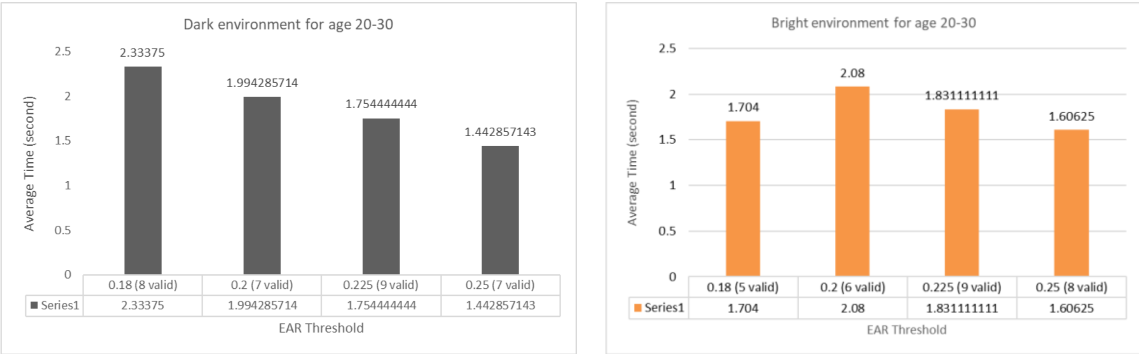
Figure 9 displays the typical time required by the research prototype for the age group of 31 to 40. Figure 9 (a) displays the result of the dark environment for the age 31–40 category, with a total of 6 participants. Unlike the result in the previous experiment, the EAR threshold with 0.225 had the shortest time among the four thresholds. But only with 4 valid data points out of 6, while for 0.25, it had 5 out of 6 valid tested data points. Therefore, 0.25 was chosen for this category, which was the age between 21 and 40 in a dark environment. Figure 9 (b) explains the experiment result for the age 31–40 category, but in a dark or low-light environment, the same result as when tested in a bright environment. 0.225 had the fastest time taken by the research prototype to detect the condition that users were closing their eyes for more than 1 second, and it had 5 out of 6 valid data points. Apparently, the 0.225 EAR threshold won the prize here.

The next experiment had a total of 6 participants in the age 41–50 category. Figure 10 displays the typical time required by the research prototype for the age group of 31 to 40. In Figure 10 (a), the 0.25 threshold had the fastest reaction time tested by the research prototype, but it only had 4 valid data points out of 6. While 0.225 and 0.2 had 5 valid data points out of 6, in this case, 0.225 had a faster reaction time compared to 0.2. Therefore, 0.225 was chosen from this experiment. Figure 10 (b) displays the result of 0.25, which is the best EAR threshold value among all 4 thresholds, but it only has 3 valid data points out of 6, which is only 50% correct. While 0.2 and 0.225 had 5 out of 6 valid data points, which is 83.33% valid data in this experiment, 0.225 has a faster reaction time taken by the system. Therefore, 0.225 has been chosen in this experiment.

From the experiment, there is an interesting result: when a participant was wearing spectacles, it had a serious effect on the research result, especially the thick glasses. That was one of the reasons that made the result invalid, as shown in Table 1. The optimal value for each age category for dark and bright lighting has been chosen as shown in Table 1. In conclusion, the EAR threshold of 0.225 seems to be the optimal value, as 5 out of 6 results show this value, while only the 31–40 age category has a better result in a dark environment.

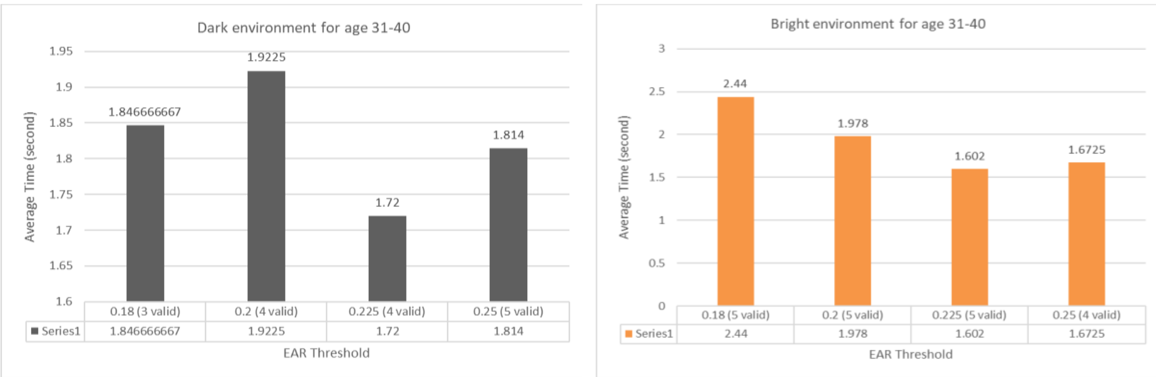
As shown in Table 2, the selected EAR threshold of 0.225 from the first experiment appears to be working well for these seven participants. The research prototype detected 7 out of 7 cases correctly when participants closed and opened their eyes for longer than one second. This is likely due to the participants' larger eyes, which prevented the research prototype from sounding even when they kept opening them. In contrast, in the first experiment, with an EAR threshold of 0.25, the research prototype sounded even when the participants did not close their eyes because the detected EAR value was still below the threshold.

Figure 11 explains the performance of the 0.225 EAR threshold in the confusion matrix. The true positive and true negative matrix had 100% precision with the selected EAR threshold in the second experiment. Therefore, 0.225 is the optimal threshold based on this research.



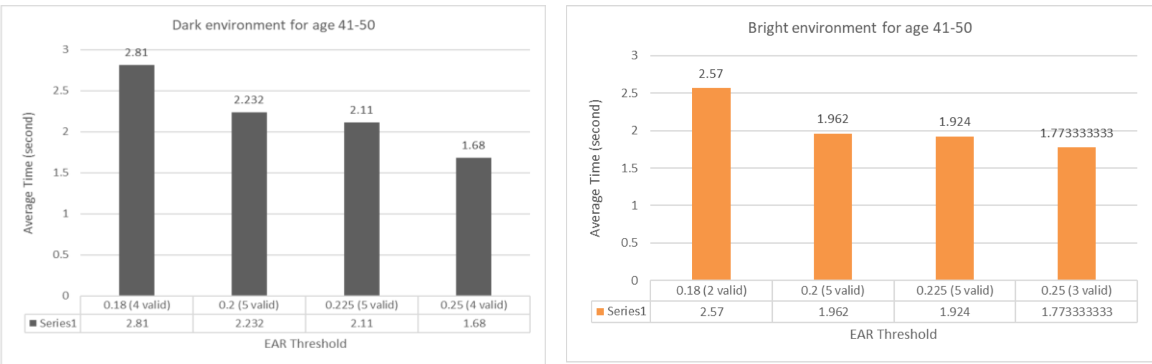
(a) (b)

Figure 8. Average time taken for age 20-30 category, (a) dark environment, and (b). bright environment



(a) (b)

Figure 9. Average time taken for age 31-40 category, (a) dark environment, and (b). bright environment



(a) (b)

Figure 10. Average time taken for age 41-50 category, (a) dark environment, and (b). bright environment

Table 1. The performance in different lighting environment

Age (Years old)	Dark (EAR)	Bright (EAR)
20-30	0.225	0.225
31-40	0.25	0.225
41-50	0.225	0.225

Table 2. The performance of 0.25 EAR

Participants	Eyes Open	Eyes Closed
Participant 1	Alarm not sound	Alarm sounds
Participant 2	Alarm not sound	Alarm sounds
Participant 3	Alarm not sound	Alarm sounds
Participant 4	Alarm not sound	Alarm sounds
Participant 5	Alarm not sound	Alarm sounds
Participant 6	Alarm not sound	Alarm sounds
Participant 7	Alarm not sound	Alarm sounds

		True Class	
		Positive	Negative
Predicted Class	Positive	7	0
	Negative	0	7

Figure 11. Confusion matrix of the experiment result.

5. CONCLUSION

The optimal EAR threshold was determined to be 0.225 after conducting two experiments: one to identify the optimal threshold and another to assess its sensitivity. This threshold appears to be effective across various age groups and lighting conditions. The second experiment yielded a favorable result of 100% with 0.225, indicating that the research prototype functioned logically without any detection errors or faults among the 7 participants. The prototype performed effectively in low-light conditions due to the inclusion of two infrared lights integrated into the camera, eliminating the need for wearing glasses. Additionally, the tested prototype has limitations, such as the inability to detect faces when a user is wearing a mask due to the requirement for the prototype algorithm to capture the human face before obtaining their eye information. Wearing thick glasses can impact the accuracy of the tested prototype or lead to detection errors. Position the prototype being tested in front of the driver in a way that does not block their vision. It is crucial because the prototype depends on detecting both the left and right eyes. Users need to customize the prototype's height to align with their eye level and adjust its angle horizontally for optimal performance. There are limited prospective studies that can be conducted in the future. The accuracy of the prototype determines the ideal mouth size for yawning. The ideal frequency of yawning for various age groups of participants. Additionally, researching distraction is also feasible. For instance, using a phone, moving one's head, or observing the surroundings while the vehicle is in motion. These secondary tasks while driving are also contributing factors to car accidents on the road.

ACKNOWLEDGEMENTS

This work was supported by the UTM Fundamental Research Grant (UTMFR) from Universiti Teknologi Malaysia under Grant vot No. PY/2022/01312. The authors are grateful to Universiti Teknologi Malaysia, Universiti Tun Hussein Onn Malaysia, and Universitas Ahmad Dahlan (Yogyakarta, Indonesia) for their contributions to this collaborative research. Also, thank you to the Embedded System and Power Electronics Research Group in Yogyakarta, Indonesia.

REFERENCES

[1] N. F. Watson et al., "Confronting drowsy driving: The American Academy of Sleep Medicine Perspective," *J. Clin. Sleep Med.*, vol. 11, no. 11, pp. 1335–1336, 2015, doi: 10.5664/jcsm.5200.

[2] A. Shetty, Bhoomika, Deeksha, J. Rebeiro, and Ramyashree, "Facial recognition using Haar cascade and LBP classifiers," *Global Transitions Proc.*, vol. 2, no. 2, pp. 330-335, 2021, doi: 10.1016/j.gltp.2021.08.044.

[3] I. Zilli, F. Giganti, and V. Uga, "Yawning and subjective sleepiness in the elderly," *J. Sleep Res.*, vol. 17, no. 3, pp. 303-308, 2008, doi: 10.1111/j.1365-2869.2008.00659.x.

[4] C. A. Flores et al., "Quantitative 3-dimensional geometry of the aging eyelids," *Plast. Reconstr. Surg. - Glob. Open*, vol. 7, no. 11, 2019, doi: 10.1097/gox.0000000000000251.

[5] OpenCV, "About," OpenCV, Nov. 4, 2020. [Online]. Available: https://opencv.org/about/.
[6] OpenCV, "License," OpenCV, Oct. 20, 2020. [Online]. Available: https://opencv.org/license/.
[7] PyImageSearch, "PyImageSearch/imutils: A series of convenience functions to make basic image processing operations such as translation, rotation, resizing, skeletonization, and displaying matplotlib images easier with opencv and python," GitHub. [Online]. Available: https://github.com/PyImageSearch/imutils.
[8] G. Bradski and A. Kaehler, Learning OpenCV, America: O'Reilly, 2009.
[9] S. Sathasivam et al., "*Drowsiness detection system using eye aspect ratio technique*," in Proc. 2020 IEEE Student Conf. on Research and Development (SCOReD), 2020, doi: 10.1109/scored50371.2020.9251035.
[10] T. V. Sri Mounika et al., "*Driver drowsiness detection using eye aspect ratio (EAR), mouth aspect ratio (MAR), and driver distraction using head pose estimation*," in ICT Systems and Sustainability, pp. 619–627, 2022, doi: 10.1007/978-981-16-5987-4_63.
[11] J. Chandiwala and S. Agarwal, "*Driver's real-time drowsiness detection using adaptable eye aspect ratio and Smart Alarm System*," in Proc. 2021 7th Int. Conf. on Advanced Computing and Communication Systems (ICACCS), 2021, doi: 10.1109/icaccs51430.2021.9441756.
[12] S. Mehta et al., "Real-time driver drowsiness detection system using eye aspect ratio and eye closure ratio," SSRN Electron. J., 2019, doi: 10.2139/ssrn.3356401.
[13] C. Dewi et al., "Eye aspect ratio for real-time drowsiness detection to improve driver safety," *Electronics*, vol. 11, no. 19, p. 3183, 2022, doi: 10.3390/electronics11193183.
[14] D. Zhang, J. Li, and Z. Shan, "*Implementation of dlib deep learning face recognition technology*," in Proc. 2020 Int. Conf. on Robots & Intelligent System (ICRIS), 2020, doi: 10.1109/icris52159.2020.00030.
[15] F. Schroff, D. Kalenichenko, and J. Philbin, "*Facenet: A unified embedding for face recognition and clustering*," in Proc. IEEE Conf. on Computer Vision and Pattern Recognition, 2015, pp. 815-823.
[16] A. Rosebrock, "My imutils package: A series of opencv convenience functions," PyImageSearch, Jun. 9, 2023. [Online]. Available: https://pyimagesearch.com/2015/02/02/just-open-sourced-personal-imutils-package-series-opencv-convenience-functions/.
[17] "Eye blink detection with OpenCV, Python, and dlib," GeeksforGeeks, Feb. 26, 2022. [Online]. Available: https://www.geeksforgeeks.org/eye-blink-detection-with-opencv-python-and-dlib/
[18] M. S. Patil et al., "LSTM based lip reading approach for Devanagiri script," *ADCAIJ: Adv. in Distrib. Comput. and Artif. Intell. J.*, vol. 8, no. 3, pp. 13–26, 2019, doi: 10.14201/adcaij2019831326.
[19] T. Soukupova and J. Cech, "*Real-Time eye blink detection using facial landmarks*," in Proc. 21st Computer Vision Winter Workshop, Rimske Toplice, Slovenia, Feb. 2016, p. 2.
[20] Q. Chen and L. Sang, "Face-mask recognition for fraud prevention using Gaussian mixture model," *J. Vis. Commun. Image Represent*., vol. 55, pp. 795-801, 2018. [Online]. Available: https://doi.org/10.1016/j.jvcir.2018.08.016
[21] S. Naido and R. R. Porle, "*Face detection using colour and Haar features for indoor surveillance*," in Proc. 2020 IEEE 2nd Int. Conf. Artificial Intelligence in Engineering and Technology (IICAIET), 2020. [Online]. Available: https://doi.org/10.1109/iicaiet49801.2020.9257813
[22] P. Viola and M. Jones, "*Rapid object detection using a boosted cascade of simple features*," in Proc. IEEE Computer Society Conf. on Computer Vision and Pattern Recognition (CVPR 2001), 2001. [Online]. Available: https://doi.org/10.1109/cvpr.2001.990517
[23] "Cascade classifier," OpenCV. [Online]. Available: https://docs.opencv.org/4.x/db/d28/tutorial_cascade_classifier.html
[24] K.-A. Kwon et al., "High-speed camera characterization of voluntary eye blinking kinematics," *J. R. Soc. Interface*, vol. 10, no. 85, 20130227, 2013. [Online]. Available: https://doi.org/10.1098/rsif.2013.0227
[25] B. G. Simons-Morton et al., "Keep your eyes on the road: Young driver crash risk increases according to duration of distraction," *J. Adolesc. Health*, vol. 54, no. 5, 2014. [Online]. Available: https://doi.org/10.1016/j.jadohealth.2013.11.021
[26] Carsomeadministrator, "8 common causes of car accidents in Malaysia," Carsome Malaysia, Jun. 4, 2020. [Online]. Available: https://www.carsome.my/news/item/8-Common-Causes-of-Car-Accidents-in-Malaysia

BIOGRAPHIES OF AUTHORS

Simon Chong Kai Yuen received his B. Sc. Degree of computer science from Universiti Teknologi Malaysia, Johor, Malaysia. He waived his master degree and persuing his Ph.D. in Univerting Teknologi Malaysia. He is doing his research in computer vision field. He can be contacted at email: simon@graduate.utm.my.

Noor Hidayah Binti Zakaria Noor Hidayah Zakaria holds the position of Senior Lecturer in the Faculty of Computing at Universiti Teknologi Malaysia. She earned her B.Sc., M.Sc., and Ph.D. degrees in Computer Science from Universiti Teknologi Malaysia (UTM) in 2011, 2014, and 2018, respectively. Dr. Zakaria's research focuses on artificial intelligence, natural language processing, bioinformatics, and software engineering testing. Presently, she is serving as the Academic Coordinator at the Faculty of Computing, UTM. In addition to her academic accomplishments, Dr. Zakaria possesses professional certifications, including Microsoft Certified: Azure AI Fundamental and Certified Python Programmer for Data Science. She can be contacted at email: noorhidayah.z@utm.my.

Ts. Dr. Goh Eg Su received the B.S. degree in information technology and the M.S. degree in technical and vocational education from Universiti Tun Hussein Onn (UTHM), Malaysia, in 2006 and 2008, respectively. She was a Lecturer with the Kolej Komuniti Jempol, Negeri Sembilan, Malaysia within 2009-2011 and Kolej Komuniti Tanjung Piai, Pontian, Johore, Malaysia, within 2011-2016. She further studied within 2016-2019 and received her Ph.D. degree in Computer Science from Universiti Teknologi Malaysia UTM, in 2019. She was a former Scholar and a Researcher with the Media and Game Innovation Centre of Excellence (MaGICX), UTM. She has published some articles in national and international conferences including publication in high impact journals. Her research interests including computer graphics and information technology (especially in augmented reality and human-computer interaction). After the study leaves, she had been posted as a senior lecturer with Politeknik Balik Pulau, teaching Android Application Development and Integrative Programming and Technologies from August 2019 until Jun 2020. Currently, she is a senior lecturer with Faculty of Computing, UTM. Email: eg.su@utm.my.

Assoc. Prof. Dr. Rohayanti Hassan is an Associate Professor in the Faculty of Computing at Universiti Teknologi Malaysia. She obtained her B.Sc., M.Sc., and Ph.D. degrees in Computer Science from Universiti Teknologi Malaysia (UTM) in 2003, 2006, and 2011, respectively. Dr. Hassan's research interests encompass artificial intelligence, natural language processing, bioinformatics, and agile software engineering. Currently, she serves as the Head of the Software Engineering Research Group at the Faculty of Computing, UTM. In addition to her academic achievements, Dr. Hassan holds professional certifications, including Microsoft Certified: Azure AI Fundamental, Certified Tester Foundation Level, Certified Professional Requirements Engineering, and AgilePM Foundation & Practitioner Certificate. She can be contacted at email: rohayanti@utm.my.

Shahreen Kasim is currently an Associate Professor in Department of Security Information and Web Technology, Faculty of Computer Science and Information Technology, Universiti Tun Hussein Onn Malaysia. Her areas of interest include bioinformatics, soft computing, data mining, web and mobile application. She can be contacted at email: shahreen@uthm.edu.my.

Tole Sutikno is a lecturer and the head of the Master Program of Electrical Engineering at the Faculty of Industrial Technology at Universitas Ahmad Dahlan (UAD) in Yogyakarta, Indonesia. He received his Bachelor of Engineering from Universitas Diponegoro in 1999, Master of Engineering from Universitas Gadjah Mada in 2004, and Doctor of Philosophy in Electrical Engineering from Universiti Teknologi Malaysia in 2016. All three degrees are in electrical engineering. He has been a Professor at UAD in Yogyakarta, Indonesia, since July 2023, following his tenure as an Associate Professor in June 2008. He is the current Editor-in-Chief of TELKOMNIKA and Head of the Embedded Systems and Power Electronics Research Group (ESPERG). He is one of the top 2% of researchers worldwide, according to Stanford University and Elsevier BV's list of the most influential scientists from 2021 to the present. His research interests cover digital design, industrial applications, industrial electronics, industrial informatics, power electronics, motor drives, renewable energy, FPGA applications, embedded systems, artificial intelligence, intelligent control, digital libraries, and information technology. He can be reached via email at tole@te.uad.ac.id.