2022 International Conference on Future Trends in Smart Communities (ICFTSC)
December 1-2, 2022, Borneo Conventional Centre Kuching Sarawak, Malaysia

Enhancements to PERCLOS Algorithm for Determining Eye Closures

Muhammad Ammar Zulkarnanie
Department of Electrical and Electronics Engineering
Universiti Teknologi PETRONAS
Perak, Malaysia
muhammad_17003999@utp.edu.my

Kiranraj Siva Shanmugam
Institute of Health Analytics
Department of Electrical and Electronics Engineering
Universiti Teknologi PETRONAS
Perak, Malaysia
kiranraj_20001956@utp.edu.my

Nasreen Badruddin
Institute of Health Analytics
Department of Electrical and Electronics Engineering
Universiti Teknologi PETRONAS
Perak, Malaysia
nasreen.b@utp.edu.my

M Naufal M Saad
Institute of Health Analytics
Department of Electrical and Electronics Engineering
Universiti Teknologi PETRONAS
Perak, Malaysia
naufal_saad@utp.edu.my

2022 International Conference on Future Trends in Smart Communities (ICFTSC) | 979-8-3503-3454-8/22/$31.00 ©2022 IEEE | DOI: 10.1109/ICFTSC57269.2022.10039811

Abstract— **This study presents an algorithm that can detect people's facial features being studied and then applied mainly on daily basis activities, as an example in driving which is detection of driver drowsiness. In this study, the algorithm named 'PERCLOS' which stands for 'percentage of eye closure' was tested to detect face by using two face landmark detectors, that are pre-trained model and library Dlib's 68-points facial landmark and 468 3D face landmarks detector from MediaPipe by Google as an alternative and detects the condition of a person's eye based on Eye Aspect Ratio (EAR). Initial assessment of the Dlib's solution on 151,537 frames (about 84 minutes) of one of tested subjects revealed that 98.66% of eye states were properly identified, resulting in 378 blinks to be recorded. Despite having rather good accuracy, the algorithm produced 166 more blinks than the 212 blinks that were expected. As for MediaPipe, with 264 blinks and only 52 additional blinks, the MediaPipe Face Mesh solution was able to categorize the identical subject with a classification accuracy of 99.87%. Additionally, adaptive thresholds for different subjects were applied in order to investigate a way to improve the studied algorithm. Surprisingly, the adaptive threshold method being studied resulted in decreasing accuracy and precision for some of the subjects. For one of tested subject, the resulted precision of studied algorithm somehow drops from 100% to 98.60%.**

Keywords—PERCLOS, Eye Aspect Ratio (EAR), Dlib, MediaPipe, modified EAR

I. INTRODUCTION

A drowsy driver puts everyone and everything around them in danger, including themselves. Fundamentally, there is no way to totally avoid the problem of driver fatigue or drowsiness, instead, drivers with such issues should be counseled early in their driving term to avoid keeping their hands on the steering wheel wherever and whenever they are tired and unable to focus fully on their driving. The Malaysian Institute of Road Safety Research (MIROS) reported about 477,204 cases with 6,915 deaths due to road crashes in 2013, with an average of 19 deaths occurring per day [1]. From the report, it was determined that the main crash risk factor was identified to be driver fatigue. Further study showed that fatigue contributed to about 7.7% of crashes involving passenger cars, 9.9% of crashes involving lorries, and 7.9% of crashes involving buses. The increasing prevalence of traffic accidents from the year 2007 to 2010 indicates that driving while fatigued is one of the primary contributing factors. According to MIROS, being fatigued while driving is riskier than previously anticipated. The growing number of car accidents caused by fatigue worldwide indicates that inattentive driving is a serious problem that requires greater research and inquiry to prevent many more accidents from happening in the future. Based on [2], despite dozens of studies done, the dreadful number of 1.17 million deaths per year due to vehicle accidents remains pretty much unchanged.

Eye closure detection has been used in other research work as a technique to detect fatigue or drowsiness. Based on [3], the most accurate and valid method of determining a driver's attentiveness level was discovered to be PERCLOS. PERCLOS is defined as the percentage of eyelid closure over the pupil during a period of time. Based on [4], technically, PERCLOS is defined as the percentage of time that the eyes are more than 80% closed.

This study's attentive driver eye recognition prototype system is made feasible by using two facial landmark libraries, one is a pre-trained model and library Dlib's 68-point facial landmark by [5]. This model is based on a face landmark annotation technique developed by [6] that has processed 755,370 photos from 337 distinct people. This introduces the 68-point mark-up facial landmark that we will utilize for our annotations. An alternative provided by [9] is a 3D face landmark detector from MediaPipe by Google which estimates a total of 468 3D facial landmarks in real-time. In this study, we also investigated if the PERCLOS algorithm can be enhanced by adopting modified eye aspect ratio (EAR) threshold, as proposed in [7].

II. LITERATURE REVIEW

A. Facial Landmarks

This project employs two pre-trained and certified facial landmark models which are Dlib and MediaPipe to identify and interpret facial features such as the eyes, brows, nose, mouth, and face shape. For Dlib, the tested coding for this study was able to identify and recognize any face in a short period of time and efficiently by importing any video or footage that includes a face by applying the pre-trained model and library by [8] for facial landmarks.

When Dlib is provided with an input image a shape predictor algorithm is triggered to locate critical points of interest along the structure given in the image. The objective in the context of facial landmarks is to use shape prediction algorithms to identify essential facial features like those shown in Fig. 1. The pre-trained facial landmark detector in the Dlib library is used to identify the position of 68 (x, y) coordinates associated with face structures.

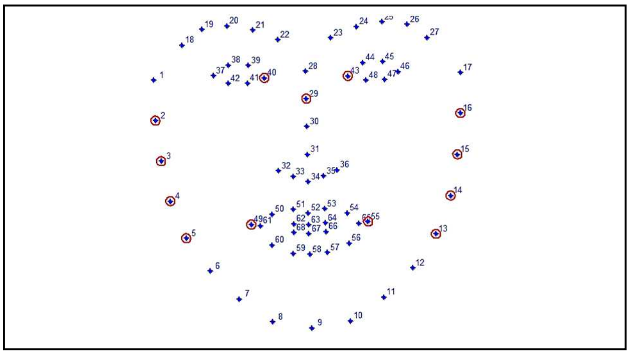

Fig. 1. Dlib 68 points facial landmarks

This project also utilized MediaPipe Face Mesh, a 468 3D facial landmark detector as an alternative to the Dlib facial landmark detector. MediaPipe Face Mesh which is developed by Google predicts a total of 468 3D face landmarks as shown in Fig. 2 in real-time. It is a simple approach designed specifically for mobile GPU inference. The MediaPipe Face Mesh solution already has a face identification model and does not require pairing with additional face detectors, in contrast to Dlib's 68 facial landmarks detector. MediaPipe's face detection model named BlazeFace, when compared to the MobileNetV2-SSD model, achieved a lower inference time of 0.6 milliseconds on the GPU of the Apple iPhone XS. When compared to MobileNetV2-SSD, which received a precision of 97.95, the BlazeFace model likewise obtained superior precision of 98.61. These characteristics make MediaPipe Face Mesh an effective substitute for DLib's solution.

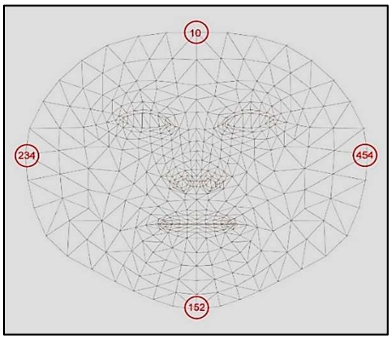

Fig. 2. MediaPipe 468 points facial landmark

B. PERCLOS Tracking

PERCLOS tracking, or also known as eye closure or eye droop tracking is tracking the closure or droopiness of the eyes of a driver when he or she is driving. This includes the opening, closing, and blinking of the eyes. This study focuses on the eyes droopiness because whilst driving, a driver should always be focused and stay in a conscious state and should not be drowsy or fatigued. To determine if one is drowsy or not, their eye closures must comply with PERCLOS standard threshold values based on [9], which are P70; the proportion of time when the eyes are closed at least 70 percent, and P80; the proportion of time when the eyes are closed at least 80 percent. Any eye closure value taken that is below P70 threshold value will not be considered as drowsy. PERCLOS can be calculated by using the formula shown in (1):

$$f = \frac{t_3 - t_2}{t_4 - t_1} \tag{1}$$

where f is the value of PERCLOS, t_1 and t_2 are the time that the eyes close from largest to 80% open and from 80% to 20%, respectively, t_3 is the time from 20% closed to 20% open, and t_4 is the time of eyes opening from 20% to 80%.

C. Eye Aspect Ratio (EAR)

The eye aspect ratio (EAR) is a measure of how wide the eyes are open. Based on [5], the EAR was calculated using the eye landmarks shown in Fig. 3 for pictures with faces. For every video frame where eye landmarks are recognized, the authors employ an SVM classifier to train the dataset. Because each person's eyelid movement is different, they used a Hidden Markov Model to construct a blinking model that addresses the varying EAR values over

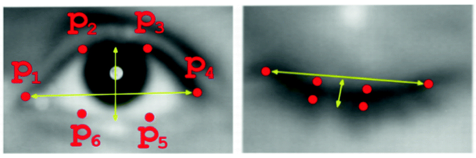

Fig. 3 EAR landmarks represented by p_i

time. As a result, the prototype system was able to detect the driver's eyes and perform the PERCLOS computation to identify the status of the driver's drowsiness.

$$EAR = \frac{\|p_2 - p_6\| + \|p_3 - p_5\|}{2\|p_1 - p_4\|} \tag{2}$$

From (2), the distances between the vertical points (p_2, p_3, p_5 and p_6) are calculated for the numerator and the distance between the horizontal points (p_1 and p_4) is determined for the denominator. Whenever a person blinks or closes their eyes, through the EAR method, we can observe that the distance between p_1 and p_4 horizontal points remain the same, but the vertical points value will decrease which will trigger EAR calculation and return results to indicate the person is either drowsy or not.

D. Modified Eye Aspect Ratio

It is known that different people have different eye sizes with different eye widths and heights. In this study, based on [7], the EAR value for each tested subject was recalculated and then applied, with each subject having their own modified EAR threshold value. The proposed modified eye aspect ratio (modified EAR) is calculated based on the following equations:

$$\mathrm{EAR_{closed}} = \frac{\mathrm{EAR_{l,\,min}} + \mathrm{EAR_{r,\,min}}}{2} \tag{3}$$

$$\mathrm{EAR_{open}} = \frac{\mathrm{EAR_{l,\,max}} + \mathrm{EAR_{r,\,max}}}{2} \tag{4}$$

$$\text{Modified } \mathrm{EAR_{Threshold}} = \frac{\mathrm{EAR_{closed}} + \mathrm{EAR_{open}}}{2} \tag{5}$$

where $\mathrm{EAR_{l,\,min}}$ and $\mathrm{EAR_{r,\,min}}$ are the minimum EAR for the left and right eyes, respectively, while $\mathrm{EAR_{l,\,max}}$ and $\mathrm{EAR_{r,\,max}}$ are the maximum EAR for the left and right eyes, respectively.

Equation (3) is to obtain the minimum average value of EAR for closed eyes (left and right eyes) and equation (4) is the equation to obtain the maximum average value of EAR for both eyes open. The result of these 2 equations will then be used in (5) to calculate the Modified EAR value of the subjects which will be the threshold value of each subject accordingly.

The eye status is determined using the Modified EAR threshold, where

$$\text{EAR} \leq \text{Modified EAR}_{\text{Threshold}} = \text{Eyes Closed} \quad (6)$$

$$\text{EAR} \geq \text{Modified EAR}_{\text{Threshold}} = \text{Eyes Open.} \quad (7)$$

The EAR output range with the eyes open and closed is shown in eqns. (6) and (7), respectively. When the eyes are closed, the EAR value will be close to 0, but when the eyes are open, it can be any integer greater than 0.

III. METHODOLOGY

A. Eye Closure Detection Process

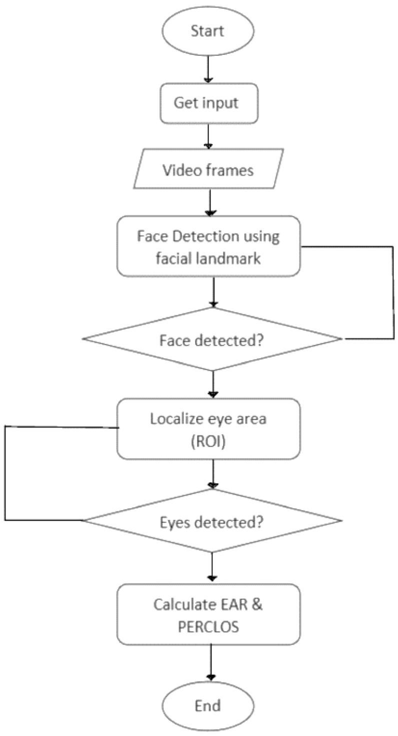

Fig. 4. Flowchart of proposed algorithm

Based on Fig. 4, the algorithm starts with getting input to be fed into the system's coding, which is a video dataset of people's faces from a Higher Institution Center of Excellence (HiCoE) project on driver drowsiness detection. After that, the algorithm divides the video into frames and analyzes each of the frames, whereby in this project the videos were analyzed for the first 10 minutes. The longer the input's duration, the longer the time needed for the algorithm to finish analyzing the input. The purpose of analyzing through each frame of input is to detect face in the input using facial landmark detection implemented in the algorithm's coding. As a result, if the algorithm does not detect any face throughout the input, it will turn back to the analyze frames section again to re-analyze the input. If the algorithm does detect any face in the input, then it will automatically localize eye areas which is the region of interest (ROI) that are contained to further calculate the eye aspect ratio (EAR) and percentage of eye closure (PERCLOS) analyzed in the input of the algorithm. By following the study done by [10], the subject in the video dataset will be considered drowsy only if the EAR threshold value of the subject is lower than the subject's individual modified EAR threshold value for a duration of at least 3 seconds.

B. Modified EAR Identification

The modified EAR threshold values were calculated for all tested subjects in this project. The purpose of obtaining the modified EAR threshold value for each subject was to investigate this method's reliability to be used in future PERCLOS algorithm implementation. Equations (3), (4), and (5) are the formulas to calculate the modified EAR threshold values, and (5) and (6) are the conditions that are to be set to determine the condition of the eyes of subjects. The modified EAR was calculated using the MediaPipe Face Mesh landmarks of both left and right eyes, which are point index 33, 159. 133 and 145 for the left eye and 362, 386, 263, and 374 for the right eye.

C. Performance Evaluation Metrics

After processing the dataset videos using the PERCLOS algorithm, manual visual inspection was conducted on the frames to determine if the frames were classified correctly by PERCLOS. The results of the checked frames were evaluated by dividing them into 4 groups as shown in Fig. 5, where

i. **True Positive (TP)**: PERCLOS detects eyes closed when eyes are actually closed

ii. **True Negative (TN)**: PERCLOS detects eyes as open when eyes are actually open

iii. **False Positive (FP)**: PERCLOS detects eyes as closed when eyes are actually open

iv. **False Negative (FN)**: PERCLOS detects eyes as open when eyes are actually closed

		Predicted	
		Negative	**Positive**
Actual	**Negative**	True Negative	False Positive
	Positive	False Negative	True Positive

Fig. 5. Evaluation matrix

These groups of frames then were used as variables to calculate the accuracy, precision, recall, and F-measure of the performance of Dlib and MediaPipe landmark detectors. The definitions of these performance metrics are given in eqns. (7) – (10).

$$\text{Accuracy} = \frac{\text{TP+TN}}{\text{TP+FP+TN+FN}} \quad (7)$$

$$\text{Precision} = \frac{\text{TP}}{\text{TP+FP}} \quad (8)$$

$$\text{Recall} = \frac{\text{TP}}{\text{TP+FN}} \quad (9)$$

$$\text{F1-measure} = \frac{2\times\text{Precision}\times\text{Recall}}{\text{Precision+Recall}} \quad (10)$$

Accuracy refers to how close a measurement is to the true value. Precision is used to determine how good the model is in predicting positive results, while recall refers to the result of TP divided by TP and FP, recall is a useful measure when the cost of false negatives is high. A good F-measure (closer to 1) indicates that there are low false positives and low false negatives.

IV. RESULTS AND DISCUSSION

A. Dataset

The PERCLOS algorithm was tested on a dataset of 10 people consisting of 5 males and 5 females with ages ranging from 20 to 37 years old. The dataset used for this study is provided by the Centre of Intelligent Signal and Imaging Research (CISIR) of Universiti Teknologi PETRONAS, Malaysia. The subjects were required to perform a driving task on a driving simulator for 80 minutes. The driving scenario was a lane-keeping task at a constant speed of 100 kilometers per hour. The subjects were recorded using a Logitech HD Pro Webcam C920 and its data was synchronized with physiological signal data. The webcam was placed one meter apart, facing the subjects' frontal faces. The basic description of all 10 subjects is shown in TABLE I.

TABLE I
Demographics of the 10 selected subjects

Subject	Gender	Wearing Glasses?	Frames Eye Closed	Drowsiness Status
1	Female	Yes	53	Alert
2	Male	No	166	Alert
3	Male	No	40	Alert
4	Female	No	520	Alert
5	Male	No	371	Alert
6	Male	Yes	2811	Drowsy
7	Male	No	444	Alert
8	Female	No	493	Alert
9	Female	Yes	2226	Drowsy
10	Female	No	3867	Drowsy

Only the first 10 minutes (18,000 frames) of the 80-minute video for each subject were examined for the comparison study. Three subjects in total expressed drowsiness signs, such as prolonged blink length, even though the study had only been conducted for 10 minutes.

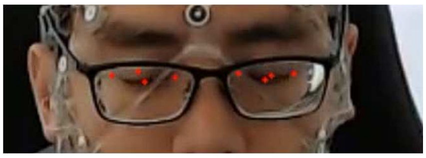

Fig. 6. Error locating Subject 6's eye

Fig. 7. Incorrect labeling of landmarks for Subject 7

A total of three subjects who wore glasses were chosen for the variability of the study

B. Comparison of Dlib and MediaPipe

The result of both Dlib and MediaPipe can be compared in TABLE II below. The landmarks are being compared by using the accuracy, precision, recall, and f-measure of each landmark to provide a better analysis of the detectors. This is because, from the initial evaluation that was done earlier in this project, it was noticed that comparing these landmark detectors using accuracy alone is not enough. The video dataset was labeled with PERCLOS values using the Dlib solution. However, the initial assessment of Dlib on 151,537 frames (about 84 minutes) of Subject 3 revealed that 98.66% of eye states were properly identified, resulting in 378 blinks being recorded. Despite having a rather good accuracy, the algorithm produced 166 more blinks than the 212 blinks that were expected because several crucial frames were incorrectly classified. As a result, the MediaPipe Face Mesh solution was put out as a substitute based on Google's stated real-time performance and precision statistics. With 264 blinks and only 52 additional blinks, the MediaPipe Face Mesh solution was able to categorize the identical subject with a classification accuracy of 99.87 percent.

From TABLE II, it can be observed that MediaPipe Face Mesh gives better results in terms of achieving higher values of accuracy, precision, recall, and F-measure for all 10 subjects. Thus, MediaPipe Face Mesh is clearly the better solution for detecting facial landmarks. However, not all results from MediaPipe landmarks detector are accurate enough since there was a case where the facial landmark did not detect the subject's face points correctly. This can be seen in subject 6. While manually checking the video of subject 6 using the MediaPipe landmark, there were some

TABLE II
DATA COMPARISON OF DLIB AND MEDIAPIPE

Subject	DLib 68 Facial Landmarks				MediaPipe Face Mesh			
	Accuracy	Precision	Recall	F-measure	Accuracy	Precision	Recall	F-measure
1	94.09	4.58	96.23	8.75	99.97	100.00	88.68	94.00
2	99.34	65.16	60.84	62.93	99.97	97.08	100.00	98.52
3	99.75	35.29	15.00	21.05	100.00	100.00	100.00	100.00
4	66.34	6.30	76.73	11.64	99.74	99.79	91.35	95.38
5	96.67	27.91	38.81	32.47	100.00	100.00	100.00	100.00
6	84.46	61.81	1.21	2.37	97.98	98.27	88.65	93.21
7	97.44	0.00	0.00	0.00	98.92	**69.65**	99.77	**82.04**
8	97.29	100.00	1.01	2.01	99.97	99.39	99.39	99.39
9	85.85	14.91	3.05	5.07	99.23	97.15	96.59	96.89
10	75.31	43.98	54.62	48.73	99.49	97.75	99.95	98.84

frames where the right eye landmark did not detect the right eye area of the subject accurately, shown in Fig. 6. This kind of error will lead to producing false positive or false negative frames, which will then affect the accuracy of the landmark detector. It should be noted that subject 6 head was tilted a little to the left of the camera's point-of-view which might lead to the mentioned error.

It can also be observed that the precision of Subject 6 using MediaPipe is 69.65% even though the subject was alert during the simulation driving. By taking a closer look at Subject 7's eye in Fig. 7, the eye landmark on the right eye of the subject was labeled incorrectly due to the landmark error itself. Specifically for Subject 7, this error happened in many frames which might result in a lower accuracy value.

C. Modified Eye Aspect Ratio Threshold Value

The modified EAR threshold values for the 10 subjects were calculated for further study in this project and listed in TABLE. III. The respective modified EAR threshold value was then applied to each subject. It can be observed that the modified EAR threshold values obtained for all subjects are slightly above the normal standardized EAR threshold value which is 20.

TABLE. III.
PARAMETERS OF MODIFIED EAR THRESHOLD VALUE

Subject	EAR_{closed}	EAR_{open}	Modified EAR
1	5.13	41.355	23.2425
2	6.08	40.045	23.0625
3	7.25	41.22	24.235
4	5.87	40.85	23.36
5	7.69	41.4	24.55
6	7.725	39.175	23.45
7	3.175	44.645	23.91
8	8.86	45.515	27.1875
9	8.14	39.165	23.6525
10	3.435	48.08	25.7575

However, this study will only discuss the result of applying the modified EAR threshold value on only the first 5 subjects due to the time constraints in repeating the classification process. The videos of the 5 subjects were processed using the calculated modified EAR threshold values by using MediaPipe Face Mesh.

TABLE IV shows the performance comparison between using the normal threshold value and the modified EAR threshold value. The videos of the 5 subjects with modified EAR values were checked thoroughly frame by frame to validate the method proposed by [10]. Two of the subjects showed similar or improved performance when using the modified EAR threshold. For Subject 1, the accuracy, recall, and f-measure values increased a little compared to normal EAR threshold value's result, but the precision somehow drop from 100% to 98.60% due to some false positive or false negative. Subject 3 resulted in the same behavior as before, which is 100% on all the recorded variables.

From this study, it can be concluded that unlike the study in [7], the method using modified EAR may improve the performance in some cases but not all. It can be seen that for Subjects 2, 4, and 5 the performance using the modified EAR in fact became worse. This might be due to some of the mistaken values of maximum or minimum EAR for both right and left eyes that were to be used in calculating the modified EAR threshold value. This result also might be due to the eye landmarks fluctuating/jittering just between higher and lower than the modified EAR threshold values, such as in the case of Subject 2 and Subject 4 which had 53 and 96 false positive (FP) frames respectively as shown in Figs. 8 and 9.

Fig. 8. FP frame of Subject 2

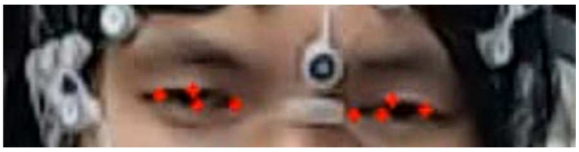

Fig. 9. FP frame of Subject 4

V. CONCLUSION

For years, researchers have been interested in detecting a person's drowsiness by implementing PERCLOS algorithm, and it will continue to be an attractive topic of research in the future. This is mostly due to the fact that changes in algorithms always occur and thus researchers will always try their best to develop and investigate the best way or method to implement PERCLOS for other application too besides in determining driver drowsiness. By stating that, throughout this research, new approach for PERCLOS were studied and tested, that are by comparing Dlib and MediaPipe facial landmark detectors performance and applying the concept of modified EAR threshold value for all subjects in order to reach the objective of this project, that is to develop an algorithm for detecting percentage of eye closure in determining eye closure. From the results discussed in Results and Discussion earlier, it is best to conclude that MediaPipe Face Mesh serves better solution and performance in detecting the opening and closing of the eyes.

TABLE IV.
NORMAL EAR AND MODIFIED EAR COMPARISON

Subject	Normal EAR				Modified EAR			
	Accuracy	Precision	Recall	F-measure	Accuracy	Precision	Recall	F-measure
1	99.97	100.00	88.68	94.00	99.99	98.60	98.60	98.60
2	99.97	97.08	100.00	98.52	99.70	79.38	99.51	88.31
3	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00
4	99.74	99.79	91.35	95.38	99.41	88.94	98.72	93.58
5	100.00	100.00	100.00	100.00	99.93	97.11	100.00	98.54

However, in this study, it also can be seen that some of the results are not accurate enough due to many factors such as the jittering and misplacement of eye landmarks. Meanwhile for the modified EAR threshold value, this study somehow did not achieve the same result as [10] might be due to inaccurateness in analyzing of data, inconsistent head position of tested subjects which affected the values obtained from dataset videos. These types of limitations should be addressed in future studies or work to create a better PERCLOS algorithm that can detect the eye closures accurately. Some recommendations include adaptive thresholding with better data analyzation, recording of tested subjects to be standardize in terms of the head angels and position, eye level, and distance of the subject with the camera.

ACKNOWLEDGMENT

The authors would like to thank the Yayasan UTP for supporting this research through YUTP-Fundamental Research Grants, Ref: 015LC0-241 and ICRF-International Collaborative Research Fund (Universitas Muhammadiyah, Sukarta, Indonesia), Ref: 015ME0-255. The authors wish to also thank Raja Nur Hamizah Raja Khairuddin, Sreeza Tarafder and Dr Rodney Petrus Balandong, for helpful discussions regarding the paper and their time spent in manually labelling the video dataset.

REFERENCES

[1] N. Othman, S. N. Jamaludin, N. M. Johari, A. Shabadi, "Accessing Teacher Readiness in Implementing the Scoring of Pedestrian Facilities by Schools in Selangor," Malaysian Institute of Road Safety Research, MRR No. 377, 2020

[2] A. Kareem, "Review of Global Menace of Road Accidents with Special Reference to Malaysia- A Social Perspective," The Malaysian Journal of Medical Sciences: MJMS. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3561885 [accessed February 21, 2022].

[3] U.S. Department of Transportation (Ed.) PERCLOS: A Valid Psychophysiological Measure of Alertness As Assessed By Psychomotor Vigilance. PsycEXTRA Dataset. 1998.https://doi.org/10.1037/e509282006-001. [accessed February 21, 2022].

[4] Wierwille, W. W., & Ellsworth, L. A. "EVALUATION OF DRIVER DROWSINESS BY TRAINED RATERS," Accident Analysis & Prevention, vol. 26(5), pp 571–581, 1994.

[5] Soukupová, T., & Cech, J., "Real-Time Eye-Blink Detection Using Facial Landmarks," 21st Computer Vision Workshop, 2016.

[6] Gross, R., Matthews, I., Cohn, J., Kanade, T., & Baker, S. "Multi-PIE," 2009. https://www.sciencedirect.com/science/article/pii/S0262885609001711. [accessed February 21, 2022].

[7] C. Dewi, R.-C. Chen, X. Jiang, and H. Yu, "Adjusting eye aspect ratio for strong eye blink detection based on facial landmarks," PeerJ Computer Science, vol. 8, 2022.

[8] C. Sagonas, E. Antonakos, G. Tzimiropoulus, S. Zafeiriou, M. Pantic, "300 Faces In-The-Wild Challenge: Database and Results," Image and Vision Computing, 2016.

[9] Y. Peng, Y. Dong, & D. Cheng, "Design and Implementation of a Driver's Eye State Recognition Algorithm Based on PERCLOS," Chinese Journal of Electronics, vol. 23(4), 2014.

[10] C. Lugaresi, J. Tang, H. Nash, C. McClanahan, E. Uboweja, M. Hays, F. Zhang, C.-L. Chang, M. G. Yong, J. Lee, W.-T. Chang, W. Hua, M. George and M. Grundmann, "Mediapipe: A Framework for Building Perception Pipelines," June 2019.