



# Efficient and Robust Driver Fatigue Detection Framework Based on the Visual Analysis of Eye States

Yancheng LING<sup>1</sup>, Xiaoxiong WENG<sup>2</sup>

Original Scientific Paper  
Submitted: 23 Feb. 2023  
Accepted: 14 June 2023

<sup>1</sup> lingyancheng@126.com, School of Civil Engineering and Transportation, South China University of Technology

<sup>2</sup> cttwxweng@scut.edu.cn, School of Civil Engineering and Transportation, South China University of Technology



This work is licensed under a Creative Commons Attribution 4.0 International License

Publisher:  
Faculty of Transport and Traffic Sciences,  
University of Zagreb

## ABSTRACT

Fatigue detection based on vision is widely employed in vehicles due to its real-time and reliable detection results. With the coronavirus disease (COVID-19) outbreak, many proposed detection systems based on facial characteristics would be unreliable due to the face covering with the mask. In this paper, we propose a robust visual-based fatigue detection system for monitoring drivers, which is robust regarding the coverings of masks, changing illumination and head movement of drivers. Our system has three main modules: face key point alignment, fatigue feature extraction and fatigue measurement based on fused features. The innovative core techniques are described as follows: (1) a robust key point alignment algorithm by fusing global face information and regional eye information, (2) dynamic threshold methods to extract fatigue characteristics and (3) a stable fatigue measurement based on fusing percentage of eyelid closure (PERCLOS) and proportion of long closure duration blink (PLCDB). The excellent performance of our proposed algorithm and methods are verified in experiments. The experimental results show that our key point alignment algorithm is robust to different scenes, and the performance of our proposed fatigue measurement is more reliable due to the fusion of PERCLOS and PLCDB.

## KEYWORDS

fatigue detection; visual-based; fusion; PERCLOS; PLCDB.

## 1. INTRODUCTION

Fatigue driving is a significant cause of traffic accidents in daily life, and many traffic accidents have a high correlation with drowsy drivers [1–6]. Studies have shown that fatigue is a leading contributing factor in traffic accidents worldwide, especially for drivers who often have to work between midnight and early morning and in a dark environment [7].

Many studies of driver drowsiness have been performed; they can be classified into contact and noncontact approaches. Although physiological signals such as heart rate and brain activity can be more reliable and accurate [8–12], they disturb normal driving when intrusive methods are utilised; thus, their use is inconvenient in real environments. With the development of computer vision, driving behaviour analysis using drivers' facial features is becoming increasingly popular [13–18]. Many contactless detection methods that are based on monitoring the movement of the eyes and mouth have been proposed to detect driver fatigue [19–21]. Many of these methods use the open-source Dlib library [22] for face-critical alignment [14, 16, 18, 23–25], which would be accurate in an ideal environment and meet real-time requirements. However, conventional methods lack robust feature extraction capability for complicated settings. Imprecise positioning will increase when a driver has a head movement or is covered, as shown in *Figure 1*. The photo on the left (a) is taken in an ideal environment with a frontal face, and the photos in the middle (b) and right (c) are taken in natural environments with head movement or coverings.

During the outbreak of COVID-19, wearing masks has become necessary to stop the spread of the virus, whereas a large area coverage of the face would hinder the use of many fatigue systems that are based on mouth yawn detection [16, 23, 26, 27]. Under the circumstances, eye state-changing becomes the most crucial feature for vision-based fatigue detection systems. However, eye state monitoring also becomes less identifi-

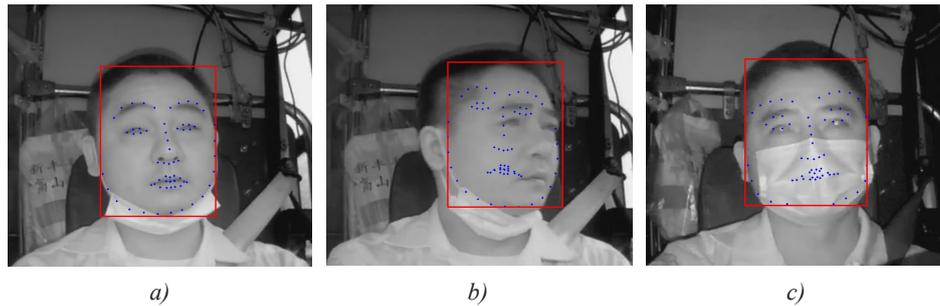


Figure 1 – Some examples of the most commonly employed face critical alignment algorithms in different scenarios

able due to the covering. So, developing a more robust framework for fatigue detection based on eye states is crucial. There are three main challenges for fatigue detection using the eye state changes of drivers:

- 1) Face critical alignment: Illumination, head poses and real-time detection speed are challenging tasks for eye key points location. Besides, during the COVID-19 outbreak, wearing masks was necessary to stop the spread of the virus. Masks cause a severe disturbance to face critical alignment. However, few studies consider this disadvantage.
- 2) Eyes state monitoring: The measurement of the eyes state is an essential foundation for fatigue parameters. Many proposed methods for eyes state monitoring are based on eye key points, and a fixed threshold is used for eyes state measurement. The fixed threshold is useful for state recognition when the eyes are fully open or closed. However, when the eyes are half-closed, it would be out of work due to the influence of individual eye size and illumination.
- 3) Fatigue measurement: Though many indicators, such as blink frequency, PERCLOS, yawn, etc. are used for fatigue measurement, PERCLOS has been verified as the most effective single indicator for vision-based fatigue measurement. However, single PERCLOS considers the total duration of prolonged blinks but disregards the influence of individual blinks.

This paper proposes an effective and reliable method for driver fatigue detection with eye state movement. First, we propose a light and robust face alignment algorithm for eye points location. Second, the dynamic threshold value is applied to recognise the state of the driver's eyes in real time. Third, a syncretic fatigue indicator that is based on the percentage of eye closure time per unit time [28] and the proportion of long closure duration blinks [29] is applied for fatigue measurement. The main contributions are highlighted as follows:

- 1) An effective and reliable eye landmark localisation alignment algorithm. Conventional eye alignment algorithms are unsatisfactory in exaggerated movement because they capture whole-face information while missing eye details. To fully utilise global face information and detailed eye information, we fuse them in an effective way.
- 2) The dynamic threshold value is applied to measure eye states. Individuals have different eye sizes and degrees of closure, whereas fixed threshold values disregard individual differences among drivers and often cannot effectively distinguish eye states.
- 3) A refined approach for driver fatigue measurement based on PERCLOS and PLCDB is proposed. The syncretic parameter considers the unit blink, which could cover the shortage of a single PERCLOS.

The following subsection discusses related work. Our methodology is presented in Section 3, and the results and discussion is included in Section 4. In Section 5, we present the conclusion.

## 2. RELATED WORK

Many vision-based fatigue detection systems have been proposed with the development of computing technology [24, 27, 30, 31]. The overall process can be summarised as follows: (1) eye detection: effective methods were employed to locate the eyes, then monitor their changes; and (2) fatigue detection: parameters such as blink frequency, yawns and PERCLOS were used to predict fatigue.

### 2.1 Eye detection

Conventional eye detection mainly relies on the eye's geometrical characteristics or grey features. In [32], an eye detection algorithm that is based on a horizontal projection histogram and edge map was proposed.

Tabrizi et al. used a chromatic algorithm to detect eyes [33]. With the development of face key point alignment algorithms, an increasing number of fatigue detection systems use landmarks of the face to locate eye positions and monitor eye changes [14, 16, 18, 23, 24, 25]. The main steps can be summarised as follows: (1) use of the face alignment algorithm to obtain facial feature points, (2) facial feature points around the eyes were extracted to obtain the eye regions, and (3) extracted eyes were fed to a trained SVM or another classifier to measure the statement of the eyes or key eyelid points were applied to monitor eye changes directly.

### 2.2 Fatigue detection

Many facial features were proven to be related to fatigue, including blink frequency, blink duration, PLCDB, PERCLOS and yawn. In [16], the duration of eye closure, blinking frequency and yawning were fused to predict fatigue. Mandal et al. proposed a development method based on PERCLOS to detect fatigue; it could monitor changes in fatigue well [34]. In [22], PERCLOS, blink rate, number of yawns, eyelid closing speed, blink duration statistics and eyelid reopening speed were extracted to measure fatigue and obtain satisfactory results. In addition, head movement was used to detect fatigue [24].

Single PERCLOS was proven to be the most related to fatigue for a long time [28, 35]. However, many studies have indicated that single PERCLOS considers the total duration of prolonged blinks but disregards the influence of individual blinks [36, 37]. Researchers have demonstrated new parameters, such as blinks’ fatigue-related amplitude/velocity ratio (AVRBs). In [37], it was proven that the proportion of long closure duration blinks (PLCDB) has a relationship with fatigue.

### 3. METHODOLOGY

The architecture of our system is shown in *Figure 2*. There are three main modules: PFLD-eyes, feature extraction and fatigue measurement. Firstly, the PFLD-eyes module acquires the eye key point based on the

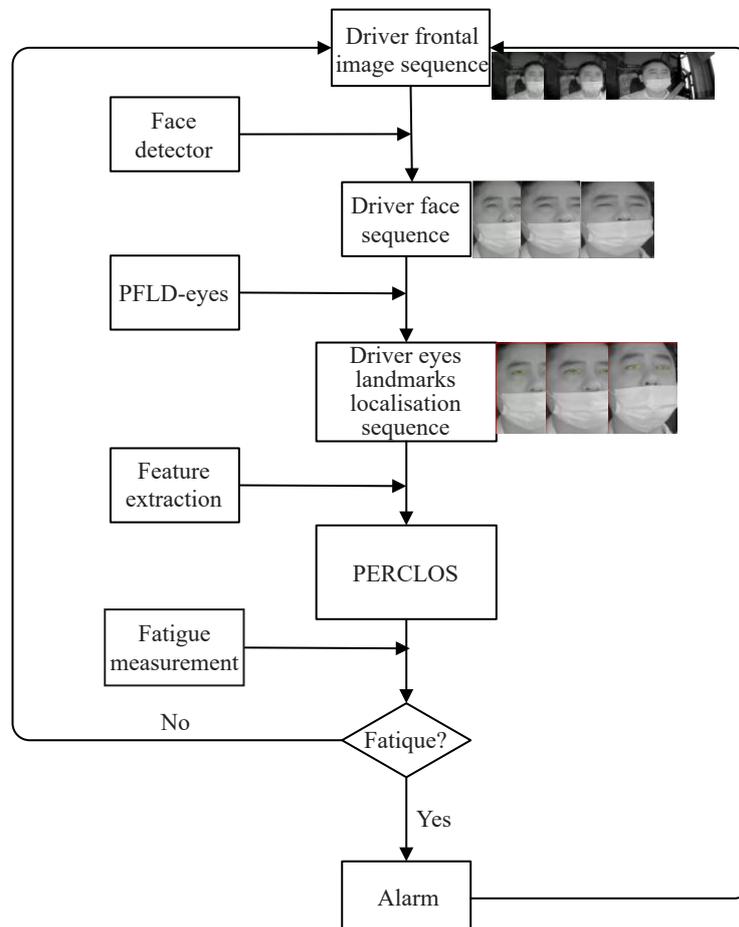


Figure 2 – The architecture of the fatigue detection system

detected face. Then, we apply the dynamic threshold value to measure the eye state in the feature extraction modules. Finally, we fuse a new parameter for fatigue measurement based on PERCLOS and PLCDB.

### 3.1 Eye detector PFLD-eyes

The eye state is an essential indication of fatigue. Many detectors would perform well to locate the eyes on the front side of the face. Still, they often fail because of movement of the head, exaggerated facial expressions and coverings, as shown in *Figure 3*. The defects of the existing model can be summarised as follows. (1) Feature extraction: a less powerful network to extract abundant information. (2) Face geometrical constraint: quadratic regression to obtain the face landmarks that do not consider face geometrical constraints.

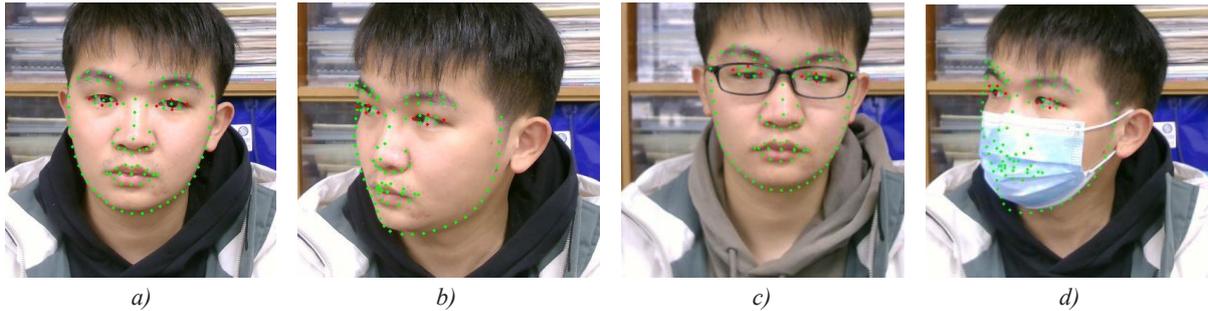


Figure 3 – A total of 106 face key point alignments are based on PFLD; 8 eye key points are labelled in red

Since the convolutional neural network (CNN) first landed in the ImageNet Challenge in 2012 and won first prize, the deep learning team developed by CNN has made rapid developments in recent years. CNN’s robust feature extraction ability has made great success in images and language. Many lightweight neural network models that balance speed and accuracy have been proposed.

The practical facial landmark detector (PFLD) [38] is a lightweight facial landmark alignment algorithm that has satisfactory performance with a balance of speed and accuracy with large poses. The PFLD achieves good performance due to its light structure and novel loss function. The regular quadratic regression loss is:

$$L = \frac{1}{M} \sum_{m=1}^M \sum_{n=1}^N \gamma_n \|d_n^m\|_2^2 \tag{1}$$

where  $\|\cdot\|$  designates a certain metric to measure the distance/error of the  $n$ -th landmark of the  $m$ -th input.  $N$  is the predefined number of landmarks per face to detect.  $M$  denotes the number of training images in each process.  $\gamma_n$  is a parameter that is always set to 1.

However, the novel loss function of PFLD could be defined as follows:

$$L = \frac{1}{M} \sum_{m=1}^M \sum_{n=1}^N \left( \sum_{k=1}^K (1 - \cos \theta_n^k) \right) \|d_n^m\|_2^2 \tag{2}$$

where  $\gamma_n$  was replaced with  $\sum_{k=1}^K (1 - \cos \theta_n^k)$  in the novel loss function, where  $\theta^k$  represents the head pose angle deviation value in pitch, yaw, and roll between the estimated value and the ground truth value. In the training phase, its branch structure could measure facial gestures  $\theta^k$  and supervise the backbone network to learn extra facial geometrical features in the novel loss function. Therefore, it would undergo substantial development with a larger pose.

PFLD achieves satisfactory performance for the global facial profile with exaggerated facial expressions due to geometrical features. However, it cannot always adequately locate eyelid key points, as shown in *Figure 3*. We obtained 106 facial landmark locations based on PFLD, and eight eye key points are labelled in red. There are eight accurate eye key point locations in the frontal pose and the large head movement in *Figures 3a and 3b*. However, the key points in the left eye do not align in *Figures 3c and 3d* when glasses and masks are worn. Because the eye region is a small part of the face, it is difficult to accurately locate the key points when glasses, masks, or other face-covering objects are worn, which causes a disturbance in recognising true eyelid contours. To obtain a more robust detector, we propose PFLD-eyes based on PFLD, which is shown in *Figure 4*. Five main

modules are shown: PFLD, eyes, key point extraction, eyes extraction, eyes-CNN and eyes key point fusion. Eyes-CNN and eyes key point fusion are the most important contributions. There are five main steps in the process. First, the PFLD detector is used to acquire 46 landmarks. Second, we obtain the preselected rectangle of the eyes by points  $p_0(x_0, y_0)$  and  $p_1(x_1, y_1)$ , which can be computed as:

$$x_0 = \frac{x_0}{2}, y_0 = \frac{\min(y_7, y_{17})}{2} \tag{3}$$

$$x_1 = \frac{(w + x_{44})}{2}, y_1 = y_{33} \tag{4}$$

where  $x_0, x_{44}, y_{17}, y_7,$  and  $y_{33}$  are the axes of the points in red, as shown in *Figure 4*, and  $w$  is the width of the input face image. Third, we obtain the eye region and eight key points  $\{e_0, e_1, \dots, e_7\}$  on the other branch, as shown in blue in *Figure 4*. Fourth, the preselected eye region in the second step was input into the eyes-CNN model, and eight key  $\{e'_0, e'_1, \dots, e'_7\}$  were obtained, as shown in yellow. Next, a fusion operation of eight key points is proposed, as shown in green:

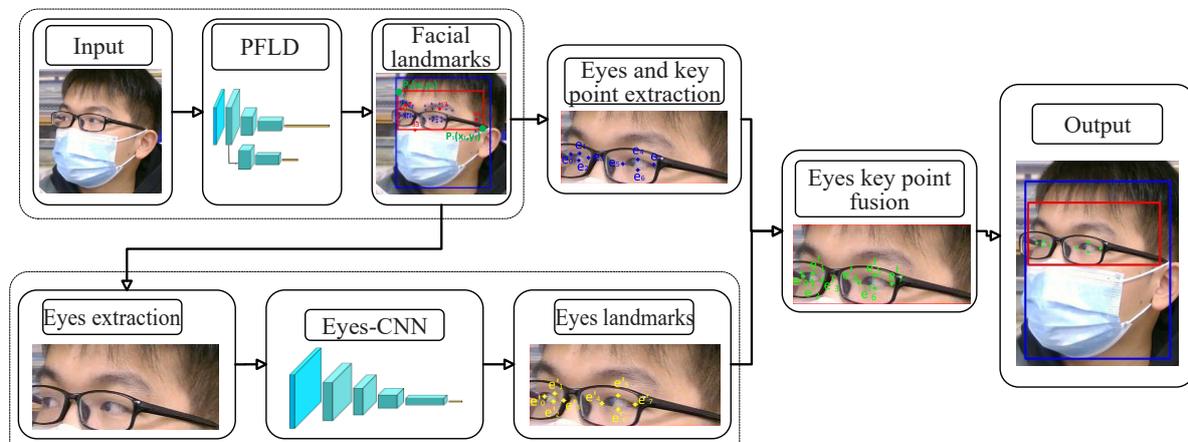


Figure 4 – Structure of PFLD-eyes

$$e_k(x) = e'_k(x), k \in \{0, 1, \dots, 7\} \tag{5}$$

$$e'_k(x) = e_k(y), k \in \{0, 1, \dots, 7\} \tag{6}$$

where  $e'_k(x)$  denotes the x-axis of the  $e'_k$  point;  $e'_k(y)$  denotes the y-axis of the  $e'_k$  point. It is necessary to note that we chose this novel fusion operation because of its excellent performance with many different permutations. It could be explained as the combination of full contour information of the face and details of the eyes.

We utilise the efficient convolutional neural networks for mobile vision applications framework, also known as MobileNetV1, as the eyes-CNN module to achieve a balance between speed and accuracy. The architecture of this framework is detailed in *Table 1*. To get a more accurate location of eye key points, we use the quadratic loss function:

$$loss = \frac{1}{M} \sum_{i=1}^N \|y'_i - y_i\| \text{ to } loss = \frac{1}{M} \sum_{i=1}^N \|y'_i - y_i\|^2 \tag{7}$$

where  $M$  represents the sample number in each training epoch,  $N$  denotes the double size of key points,  $y'_i$  is the prediction value of key points and  $y_i$  is the real value.

### 3.2 Eyes openness estimation

Continuous eye state changes are essential indicators of fatigue measures. The eye aspect ratio (EAR) is a commonly utilised parameter to distinguish eye states. Based on the eight key points that we obtained in Section 3.1, we could calculate the EAR as:

$$EAR = \frac{\|p_2 - p_3\|}{\|p_1 - p_4\|} \tag{8}$$

where  $\{p_1, p_2, p_3, p_4\}$  represent eyes key points as is shown in Figure 5.

Table 1 – Structure of MobileNetV1

Input	Operator	s	k	t	c
$112^2 \times 3$	Conv3 $\times$ 3	2	3	1	32
$56^2 \times 32$	Depthwise Conv3 $\times$ 3	1	3	1	64
$56^2 \times 64$	Depthwise Conv3 $\times$ 3	2	3	1	128
$28^2 \times 128$	Depthwise Conv3 $\times$ 3	1	3	1	128
$28^2 \times 128$	Depthwise Conv3 $\times$ 3	2	3	1	256
$14^2 \times 256$	Depthwise Conv3 $\times$ 3	1	3	1	256
$14^2 \times 256$	Depthwise Conv3 $\times$ 3	2	3	1	512
$7^2 \times 512$	Depthwise Conv3 $\times$ 3	1	3	5	512
$7^2 \times 512$	Depthwise Conv3 $\times$ 3	2	3	1	1024
$4^2 \times 1024$	Depthwise Conv3 $\times$ 3	1	3	1	1024
$4^2 \times 1024$	Full Connection	-	-	-	16

s represents stride, k represents kernel, t represents the times of repetition of operator, and c represents the dimensionality of outputs



Figure 5 – The EAR is based on eight key eye points when opening eyes or closing eyes

### 3.3 Fatigue parameters based on eyes state

Eye state serves as a crucial indicator for detecting fatigue. This section introduces PERCLOS and PLCDB, which are based on eye state, for the purpose of detection.

PERCLOS is an effective eye parameter correlated with fatigue [28]. The conventional PERCLOS calculation process is described as follows: (1) set the EAR threshold value to a value such as 0.2 and (2) calculate the percentage of closing time in a window of time; the window of time is usually 1 min or a multiple of 1 min.

However, its performance is limited by the notion that a fixed EAR threshold value is not valid for individuals, which has been conducted in [24]. In addition to the influence of individual eye size, illumination is a factor that cannot be disregarded, and people tend to narrow their eyes in bright sunlight. To obtain a more robust PERCLOS, we propose a dynamic and efficient method.

First, we obtain a dynamic EAR threshold value by the following steps.

- 1) Calculate EARs in different poses (as shown in Figure 6) and compute their eye state for approximately 5 seconds.
- 2) Collect the ten largest EARs as  $\{h_0, h_1, \dots, h_{10}\}$  and the ten smallest EARs as  $\{l_0, l_1, \dots, l_{10}\}$ .
- 3) Separately calculate the mean-variance of the sets in step 2 and delete the value beyond the threshold of triple standard deviation.
- 4) Average the processed sets in step 3 and obtain  $EAR_{max}$  and  $EAR_{min}$  as:

$$EAR_{max} = \frac{1}{N_1} \sum_{i=1}^{N_1} h_i \tag{9}$$

$$EAR_{min} = \frac{1}{N_2} \sum_{i=1}^{N_2} l_i \tag{10}$$

where  $N_1$  and  $N_2$  represent the element number of processed sets in step 3.

5) We can obtain the final threshold value  $EAR_{threshold}$  as follows:

$$EAR_{threshold} = EAR_{min} + (EAR_{max} - EAR_{min}) \cdot 0.2 \tag{11}$$



Figure 6 – Different poses to obtain the maximal EAR and minimal EAR

Second, let  $EAR_i$  denote the EAR at frame  $i$ , and let  $\Delta t$  be the interval time from frame  $i$  to frame  $i+1$ ,  $pe_i$  is the eye state in frame  $i$ . Compare  $EAR_i$  with  $EAR_{threshold}$  and set  $pe_i$  to 0, if  $EAR_i$  is greater than  $EAR_{threshold}$ ; otherwise, set it to 1. Given the period  $T$ , we obtain sets  $\{pe_0, pe_1, \dots, pe_{N_3}\}$ , where  $N_3$  is the minimum value to satisfy  $\sum_{i=0}^{N_3} \Delta t > T$ . We can calculate PERCLOS at frame  $t$  as:

$$PERCLOS_t = \frac{\sum_{i=0}^{N_3} pe_i}{N_3} \tag{12}$$

where  $PERCLOS_t$  is the PERCLOS in frame  $t$ .

Although PERCLOS is an effective indication of fatigue, it measures the total time of blinks and does not consider the unit eye blink [36, 39]. Some people naturally have a higher blink frequency, and only PERCLOS would be out of operation at times. Blinking duration is longer when drowsy drivers struggle to open their eyes to ensure safety, as shown in Figure 7; considering this finding would be effective. There is a prominent difference in closure time between normal and prolonged blinks. Many researchers have found that spontaneous eye-blink parameters, including blink duration, reopening time, closing time, amplitude-velocity ratio of blinks (AVRBs) and PLCDB, provide reliable information about fatigue. The proposed eye-blink parameters have intrinsic connections; for example, the blink duration consists of closing, closed and reopening, and PLCDB is highly correlated with blink duration. Thus, to compensate for the deficiency of PERCLOS and keep the monitor in real time, we choose PLCDB as a blink parameter.

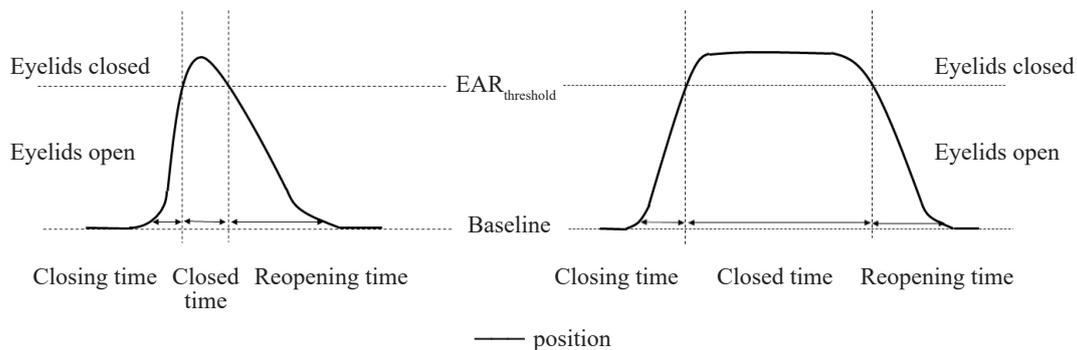


Figure 7 – Schematic of normal blinks and prolonged blinks

Let  $ECT$  denote the eye closure time, and  $\{ECT_0, \dots, ECT_{N_4}\}$  represent the sets of the most  $N_4$  number of eye closure times from present time. The  $pl$  represents the state of eye closure time. An average person has a blink frequency of approximately 16 per min. Given the period  $T$ , we could calculate  $\lceil (4T)/15 \rceil$  as  $N_4$ . The 500 msec has been proven as the maximal blink duration of an average blink [29, 36]. Set  $pl_i$  to 0, if  $ECT_i$  is less than 500 msec; otherwise, set it to 1. We would obtain the set  $\{pl_0, \dots, pl_{N_4}\}$ . The real-time PLCDB would be calculated as:

$$PLCDB_t = \frac{\sum_{i=0}^{N_A} pl_i}{n} \quad (13)$$

where  $PLCDB_t$  is the PLCDB at time  $t$ .

### 3.4 Parameters fusion to measure fatigue

PERCLOS is valid and reliable information for measuring the drowsiness level with a solid absolute correlation coefficient greater than 0.8 ( $p < 0.05$ ) [28, 34]. However, single PERCLOS considers the total duration of prolonged blinks but disregards the influence of individual blinks. PLCDB is an indicator that takes the unit blink time into account so that it would compensate for the defect of a single PERCLOS.

Compared to PERCLOS, PLCDB is a weaker parameter with an absolute correlation coefficient of approximately 0.369 ( $p < 0.05$ ) [29]. While fusing PERCLOS and PLCDB can enhance the effectiveness of fatigue measures, it is important to note that summing them with equal weight is suboptimal, as PERCLOS plays a more crucial role in measuring fatigue compared to PLCDB. To obtain a more reasonable measurement, we develop a novel method. Let  $f$  represent the fatigue level. We could calculate real-time fatigue risk as  $f$ :

$$f_t = \alpha \times PERCLOS_t + \beta \times PLCDB_t \quad (14)$$

where  $\alpha$  and  $\beta$  are coefficients of the fatigue equation, we could set  $\alpha$  and  $\beta$  according to the normalised correlation coefficient; we could set  $\alpha$  equal to 0.7 and  $\beta$  equal to 0.3. In this novel equation, we consider both PERCLOS and single blink time, which would compensate for the defect of a single PERCLOS.

## 4. RESULTS AND DISCUSSION

In this section, we experimentally demonstrate the validity of our model. Two main modules are verified, including PFLD-eyes and the fatigue measurement of our proposed method.

In the phase of PFLD-eyes, we use the open-source face aligning dataset, which includes 21,080 images used to train the PFLD model. Each face is annotated with 46 key points. We also obtained training sets with 21,080 images and eight key points in each face based on the open-source face aligning dataset to train our eyes-CNN. In the training phase of PFLD, a batch size of 128 is employed with an initial learning rate of  $10^{-4}$ . We tailor eye pictures based on the 46 face landmarks to train eyes-CNN. We train it using a batch size of 32 and an initial learning rate of  $10^{-4}$ .

In the fatigue detection method phase, few existing vision-based datasets are aimed at driver fatigue in complicated conditions, such as wearing masks and illumination changes. To evaluate the performance in a real scene, we invited 18 volunteers to drive in a simulated environment that had a camera with a similar position and distance, as well as changes in lighting conditions to real vehicles. Eighteen different volunteers perform 18 videos, and they contain two different scenes: daytime driving with masks and nighttime driving with masks. Each video lasted 1.5 hours, and each volunteer was asked to drive on the same roads with monotonous surroundings where people get drowsy easily. To obtain the real state of drivers, we used the NeuroSky TGAM



Figure 8 – Examples of our driving simulation videos

electroencephalogram (EEG) acquisition module to collect the brainwave information alpha ( $\alpha$ ) wave, theta ( $\theta$ ) wave and beta ( $\beta$ ) wave. We calculated the  $(\alpha+\theta)/\beta$  value as the ground true state of drivers, which has been proven to be highly correlated with fatigue in many types of research [40, 41]. The videos are captured by an ordinary infrared camera with the same view in the vehicle, as shown in *Figure 8*.

#### 4.1 Experiments on PFLD-eyes

In this section, we first introduce four performance indicators for evaluating our model. Next, we test our model on the public BioID dataset and compare its performance against the other three baseline models.

To evaluate the performance of PFLD-eyes in eye state, we tested it on the BioID, a public dataset with 1,521 grey images, and it is widely used in eyes state assessment for its challenge in illuminations and large poses. We use EAR as the evaluation indicator and set 0.2 as the fixed threshold value widely used in many types of research. It means that the eyes are open if the EAR is greater than 0.2 or the eyes are closed. Based on the predicted eyes state and the ground true eyes state, we calculate the evaluation indicators as follows:

$$Accuracy = \frac{TP + TN}{TP + FN + FP + TN} \% \quad (15)$$

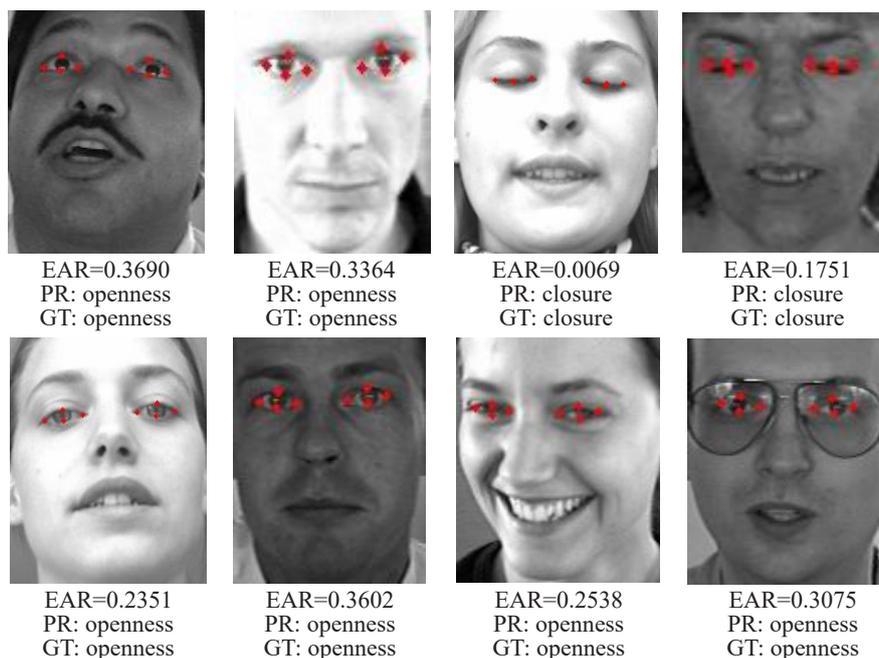
$$Precision = \frac{TP}{TP + FP} \% \quad (16)$$

$$Recall = \frac{TP}{TP + FN} \% \quad (17)$$

$$F1 = \frac{2 \cdot Precision \cdot Recall}{Precision + Recall} \quad (18)$$

where  $TP$  represents the true positive samples,  $TN$  represents the true negative samples,  $FP$  represents the false positive samples, and  $FN$  represents the false negative samples. Accuracy, precision and recall are parameters that evaluate the performance of a model in different aspects, and  $F1$  is a comprehensive indicator that is calculated by precision and recall.

To evaluate our model in eye state recognition, we compare PFLD-eyes with other different models in the BioID dataset. The result is shown in *Table 2*, and some samples of detected results are shown in *Figure 9*. PR



*Figure 9 – Some examples of the detection results on the BioID dataset*

Table 2 – The comparison of different models on the BioID database

Method	Recall (%)	Precision (%)	F1-Score (%)	Accuracy (%)
Ö. F. Soylemez	94.50	97.81	96.10	93.30
Cheng	88.58	98.00	93.05	94.00
Dlib	97.29	97.82	97.56	95.24
Ours	98.92	99.32	99.12	98.28

represents the predicted results based on the EAR, and GT represents the ground truth. Compared to the proposed methods, Ö.F.Soylezmez [42], Cheng [43] and Dlib [24, 27, 44], our model obtains the best performance in all indicators, which means that our model is robust in illuminations and large poses.

## 4.2 Experiments on fatigue measurement

In this phase, we evaluate the performance of our proposed model on fatigue detection. Before assessing the performance of fatigue measurement, we verify the robustness of the dynamic EAR threshold value.

The fixed threshold value is not robust enough for eye state monitoring, especially for half-closed eyes, due to the influence of individual eyes size and illumination. To verify the performance of dynamic EAR, we select five subjects with obvious differences in eye size. There were ten video clips in total, and every subject was asked to blink normally within 1 min during both day and night. We obtained the dynamic EAR threshold value at the first 5 s in each video clip; they represent five subjects' dynamic EAR threshold values during both day and night. We then calculated the average of ten dynamic EAR threshold values as the fixed EAR threshold value.

In *Figure 10*, we display some examples of the EARs for different subjects. The dynamic EAR threshold values for the EAR set are  $\{0.19, 0.12, 0.28, 0.26, 0.21, 0.22, 0.19, 0.22, 0.24, 0.16\}$ , and the average EAR threshold value is 0.209. The dynamic EAR threshold value ranges from 0.12 to 0.28, and even the same subject has a different EAR threshold value; for example, the fifth subject has a threshold value of 0.24 in the daytime and a threshold value of 0.16 at night. After analysing the video, we found that the EAR threshold value was susceptible to eye size, illumination and mental state. Some drivers closed their eyes unconsciously in the presence of strong light. Therefore, a fixed EAR threshold value is not robust for different drivers. To measure the effectiveness of the dynamic EAR threshold, we manually count the ground-truth blink and

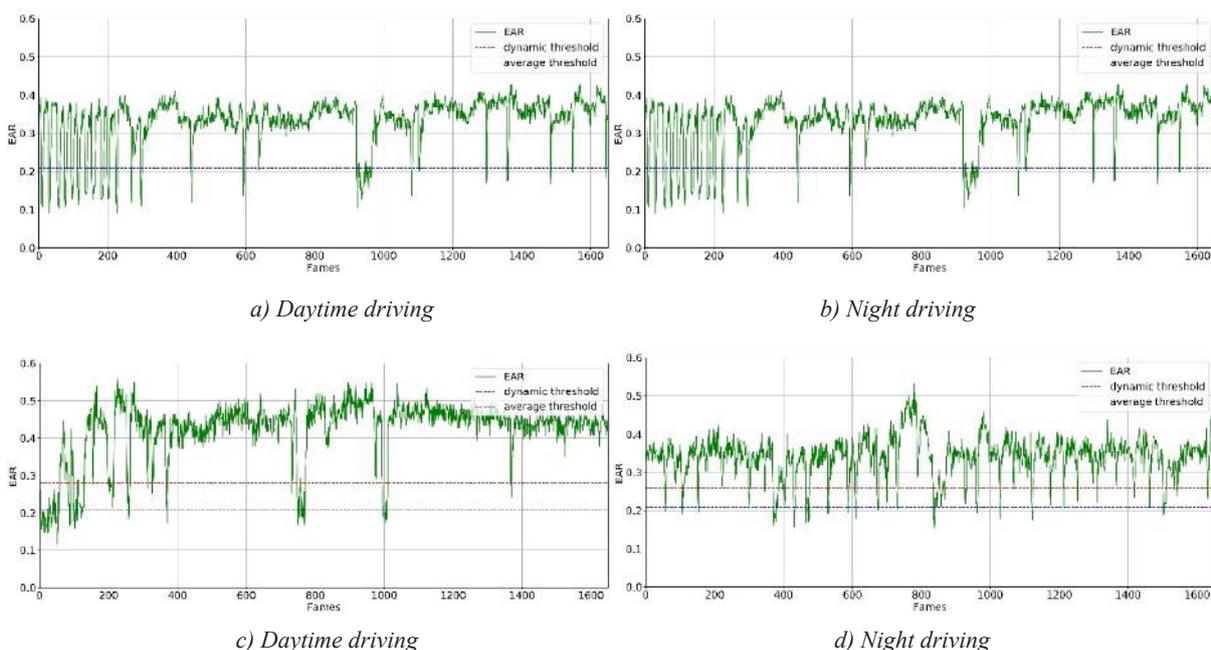


Figure 10 – Examples of changes in the EAR for different subjects; each subject took the initiative to blink for approximately 55 s (approximately 1,650 frames)

predict blinking based on the dynamic EAR threshold values and fixed dynamic EAR threshold values and calculate the accuracy as follows:

$$acc = \frac{\min(GT, PR)}{\max(GT, PR)} \% \quad (19)$$

where  $GT$  is the number of ground-truth blinks and  $PR$  is the number of predicted blinks based on the dynamic EAR threshold or fixed EAR threshold.

The result is shown in *Table 3*. It is observed that the accuracy based on the dynamic EAR threshold value (D-acc) is higher than the fixed EAR threshold value (F-acc) in most cases. We analysed the only case in which the F-acc is higher than the D-acc, as shown in Sub5 daytime, and found that our system misunderstands that the subject's eyes are closed when she looks down in a large pose; only the upper eyelid could be seen. The same reason applies to the Sub2, who put her head down in a larger pose to put her vehicle into gear during daytime driving. In most cases, the detected blink is larger by a few quantities than the ground-truth blink because it is ambiguous to manually estimate the number of blinks when the subject blinks in a small range. Generally, our dynamic EAR threshold value achieves satisfactory performance on blink detection.

*Table 3 – The accuracy based on different threshold values*

		GT	D-PR	F-PR	D-acc	F-acc
Sub1	Daytime	14	15	10	93.3	71.4
	Night	34	33	19	97.1	55.9
Sub2	Daytime	22	27	27	81.4	81.4
	Night	34	35	32	97.1	94.1
Sub3	Daytime	30	30	-	100	-
	Night	23	22	26	95.6	88.5
Sub4	Daytime	34	36	36	94.4	94.4
	Night	32	34	34	94.1	94.1
Sub5	Daytime	15	17	15	88.2	100
	Night	32	33	38	97.0	84.2

Furthermore, we conducted evaluations on test videos to further validate the accuracy of our fatigue measurement model. To obtain the real state of the subjects, we use the EEG data to calculate the indicators for fatigue measurement. The electrical waves are made up of various rhythmic waves, which could be divided into delta ( $\delta$ ),  $\theta$ ,  $\beta$  and  $\alpha$  according to their frequency. Each rhythmic wave appears with a higher frequency when people are in a different state. And the  $(\alpha+\theta)/\beta$  value has been proven to have a high correlation with fatigue [40, 41]. We obtained different brain waves' power spectral density (PSD) with varying rhythms from raw EEG signals. To reduce the influence of characteristic value fluctuation of EEG power spectrum, we calculated the mean value of power spectrum characteristic value within 5 min. For each subject, we removed the data for the first 10 min and the last 20 min, and we obtained 12 EEG characteristic values to describe the real state changes of a subject in 1 hour.

To evaluate the performance of our proposed system, we compared the single PERCLOS with improved parameters  $f$  (see *Equation 14*) for fatigue measures. We captured the video for each subject by the corresponding time of EEG data. Each video is divided into 12 clips and each clip contains 5 min. Based on the clips, we calculated the PERCLOS and our improved parameters by the PFLD-eyes model. The Spearman correlation coefficient was calculated to compare the predicted state with the real state, which is widely used in measuring the correlation of statistical variables.

The correlation coefficients of single PERCLOS and our method with EEG signals are shown in *Table 4*, and four typical results are presented in *Figure 11*. The results of the other 14 subjects are shown in *Appendix*. It is observed that the state predicted by both our method and single PERCLOS has a high correlation coefficient with the real state for most subjects. As shown in *Figure 11a*, the predicted fatigue degree of subject 2

increased from 0 min to 30 min gently and decreased from 40 min to 60 min, which is consistent with the real state. For subject 6, the predicted fatigue degree increased steeply from 0 to 40 minutes and decreased from 40 to 60 minutes, which has the same trend as the EEG curve. In this case, our method has the same changes as the PERCLOS method. However, for subject 7 and subject 17, the PERCLOS method could not keep up with the real state change, and the correlation coefficient between the PERCLOS and the real states is only 0.426 ( $p < 0.05$ ) (as shown in Table 4) for the subject 7 and 0.330 ( $p < 0.05$ ) for the subject 17. Compared to the PERCLOS method, our method has a significant improvement in correlation coefficient with 0.818 ( $p < 0.05$ ) for subject 7 and 0.809 ( $p < 0.05$ ) for subject 17. As shown in Table 4, the average correlation coefficient between the single PERCLOS and the EEG method is 0.847 ( $p < 0.05$ ), while the average correlation coefficient between our method and the EEG method is 0.900 ( $p < 0.05$ ), which has a significant improvement. It indicates that the proposed parameters, which fuse the PERCLOS and PLCDB are more robust for fatigue measures, and the PLCDB could add the single blink information, which could make up for the defect of PERCLOS.

Table 4 –The comparison of different parameters for fatigue measurement

Subject	R value ( $p < 0.05$ )		Subject	R value ( $p < 0.05$ )	
	PERCLOS	Ours		PERCLOS	Ours
1	0.697	0.709	10	0.958	0.958
2	0.860	0.888	11	0.970	0.964
3	0.888	0.888	12	0.927	0.927
4	0.999	0.999	13	0.930	0.906
5	0.940	0.943	14	0.855	0.782
6	0.906	0.906	15	0.999	0.964
7	0.426	0.918	16	0.758	0.758
8	0.976	0.976	17	0.330	0.815
9	0.916	0.925	18	0.915	0.964

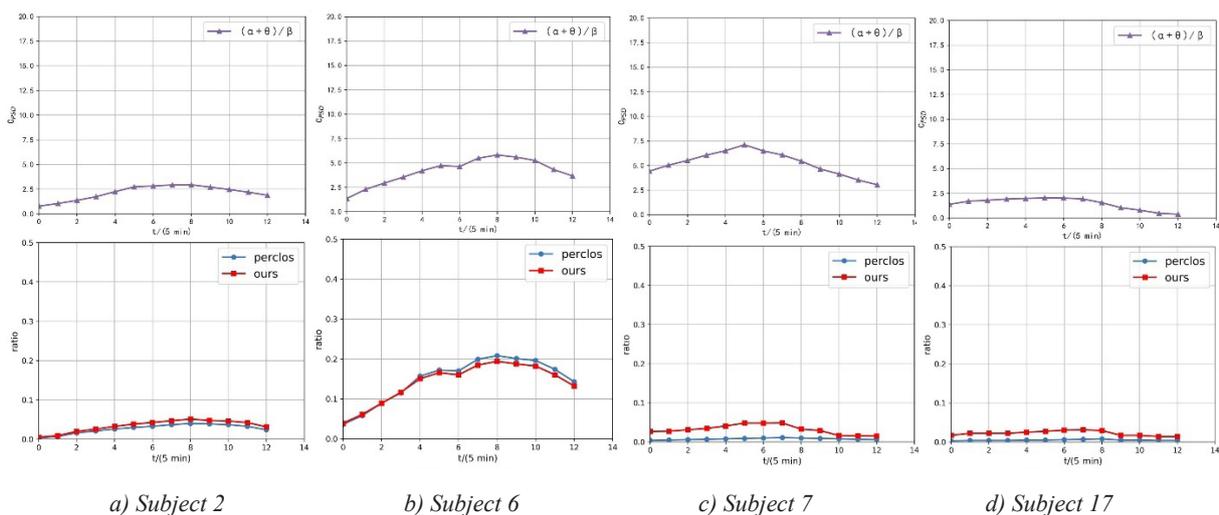


Figure 11 – Examples of state measurement by EEG signals, PERCLOS and improved methods

### 4.3 Computer environment

We train all our models, including PFLD-eyes, on a server with 2080 TI graphics. We test our system on a computer with 1050 TI graphics and an i5-8500 CPU with 3.0 GHz. The PFLD-eyes take approximately 0.011 s per frame. Our system takes approximately 0.026 s per frame. Our system runs approximately 38 frames per second (fps). Running our system on a single CPU without graphics takes around 0.1 s per frame and about ten fps. With the development of embedded modules, such as Jetson Xavier with 512 cores, which has pow-

erful reasoning capabilities for deep learning models, many deep learning models have been redeployed with real-time effects [45, 46]. Hence, our proposed system is robust and can be applied in the real world.

## 5. CONCLUSIONS

This paper introduces a robust fatigue detection system based on computer vision, capable of detecting fatigue in drivers even when they are wearing masks, or their facial expressions are exaggerated. The system first locates eight key points in the driver's eyes using PFLD-eyes, and then extracts fatigue characteristics using dynamic thresholds. A syncretic parameter is used to measure drivers' fatigue in real time. Experimental results demonstrate the system's robustness, as it achieves accurate results in different driving simulation scenes, captured from the same viewpoint as real vehicles. Our proposed syncretic parameter produces more reliable results, as it eliminates the impact of normal blinks on PERCLOS. Our method represents a significant improvement over existing fatigue detection frameworks in terms of both eye state detection accuracy and fatigue state measurement. Moreover, our model's lightweight structure allows for high processing speed, enabling it to meet real-time requirements.

In the future, we plan to increase the sample size of our study to enhance the reliability of the results and gain a deeper understanding of the different facets of driver fatigue. Additionally, we intend to deploy our fatigue detection system on Jetson Xavier and investigate ways to enhance driver safety using the insights from our detection results.

## REFERENCES

- [1] Toroyan T, Peden MM, Iaych K. WHO launches second global status report on road safety. *Injury Prevention*. 2013;19(2):150. DOI: 10.1136/injuryprev-2013-040775.
- [2] Zhang G, Yau KKW, Zhang X, Li Y. Traffic accidents involving fatigue driving and their extent of casualties. *Accident Analysis and Prevention*. 2016;87:34–42. DOI: 10.1016/j.aap.2015.10.033.
- [3] Williamson A, et al. The link between fatigue and safety. *Accident Analysis and Prevention*. 2016;43(2):498–515. DOI: 10.1016/j.aap.2009.11.011.
- [4] Dawson D, Searle AK, Paterson JL. Look before you (s)leep: Evaluating the use of fatigue detection technologies within a fatigue risk management system for the road transport industry. *Sleep Medicine Reviews*. 2016;18(2):141–152. DOI: 10.1016/j.smrv.2013.03.003.
- [5] National Highway Traffic Safety Administration. *Drowsy Driving 2015: A Brief Statistical Summary*. Oct. 2017. <https://crashstats.nhtsa.dot.gov/Api/Public/ViewPublication/812446>.
- [6] Amodio A, et al. Automatic detection of driver impairment based on pupillary light reflex. *IEEE Transactions on Intelligent Transportation Systems*. 2019;20(8):3038–3048. DOI: 10.1109/TITS.2018.2871262.
- [7] Seen KS, Mohd Tamrin SB, Meng GY. Driving fatigue and performance among occupational drivers in simulated prolonged driving. *Global Journal of Health Science*. 2010;2(1). DOI: 10.5539/gjhs.v2n1p167.
- [8] Kokonozi AK, Michail EM, Chouvarda IC, Maglaveras NM. A study of heart rate and brain system complexity and their interaction in sleep-deprived subjects. *Computers in Cardiology*. 2008;35:969–971. DOI: 10.1109/CIC.2008.4749205.
- [9] Kar S, Bhagat M, Routray A. EEG signal analysis for the assessment and quantification of driver's fatigue. *Transportation Research Part F: Traffic Psychology and Behaviour*. 2010;13(5):297–306. DOI: 10.1016/j.trf.2010.06.006.
- [10] Galley N. Blink Parameter as indicator of drivers sleepiness. *Nursing Standard*. 2003;23(35):26–27. <http://www.ncbi.nlm.nih.gov/pubmed/19749128>.
- [11] Patel M, Lal SKL, Kavanagh D, Rossiter P. Applying neural network analysis on heart rate variability data to assess driver fatigue. *Expert Systems with Applications*. 2011;38(6):7235–7242. DOI: 10.1016/j.eswa.2010.12.028.
- [12] Ahmed J, Li JP, Khan SA, Shaikh RA. Eye behaviour based drowsiness Detection System. *2015 12th International Computer Conference on Wavelet Active Media Technology and Information Processing, ICCWAMTIP*. 2015. p. 268–272. DOI: 10.1109/ICCWAMTIP.2015.7493990.
- [13] Chau L, Yap K. *Driver fatigue detection*.
- [14] Li K, Gong Y, Ren Z. A fatigue driving detection algorithm based on facial multi-feature fusion. *IEEE Access*. 2020;8:101244–101259. DOI: 10.1109/ACCESS.2020.2998363.
- [15] Luo XQ, Hu R, Fan TE. The driver fatigue monitoring system based on face recognition technology. *Proceedings of the 2013 International Conference on Intelligent Control and Information Processing, ICICIP*. 2013. p. 384–388. DOI: 10.1109/ICICIP.2013.6568102.
- [16] Deng W, Wu R. Real-time driver-drowsiness detection system using facial features. *IEEE Access*. 2019;7:118727–118738. DOI: 10.1109/access.2019.2936663.
- [17] Coetzer RC, Hancke GP. Eye detection for a real-time vehicle driver fatigue monitoring system. *IEEE Intelligent Vehicles Symposium, Proceedings*. 2011. p. 66–71. DOI: 10.1109/IVS.2011.5940406.

- [18] Savaş BK, Becerikli Y. Real time driver fatigue detection system based on multi-task ConNN. *IEEE Access*. 2020;8:12491–12498. DOI: 10.1109/ACCESS.2020.2963960.
- [19] Smith P, Shah M, da Vitoria Lobo N. Determining driver visual attention with one camera. *IEEE Transactions on Intelligent Transportation Systems*. 2003;4(4):205–218. DOI: 10.1109/TITS.2003.821342.
- [20] Picot A, Charbonnier S, Caplier A. On-line detection of drowsiness using brain and visual information. *IEEE Transactions on Systems, Man, and Cybernetics Part A: Systems and Humans*. 2012. DOI: 10.1109/TSMCA.2011.2164242.
- [21] Horng WB, Chen CY, Chang Y, Fan CH. Driver fatigue detection based on eye tracking and dynamic template matching. *IEEE International Conference on Networking, Sensing and Control*. 2004;1:7-12. DOI: 10.1109/ICNSC.2004.1297595.
- [22] King DE. Dlib-ml: A machine learning toolkit. *Journal of Machine Learning Research*. 2009;10:1755–1758.
- [23] Cheng Q, et al. Assessment of driver mental fatigue using facial landmarks. *IEEE Access*. 2019;7:150423–150434. DOI: 10.1109/ACCESS.2019.2947692.
- [24] You F, et al. A real-time driving drowsiness detection algorithm with individual differences consideration. *IEEE Access*. 2019;7:179396–179408. DOI: 10.1109/ACCESS.2019.2958667.
- [25] Sun Y, et al. Driver fatigue detection system based on colored and infrared eye features fusion. *Computers, Materials and Continua*. 2020;63(3):1563–1574. DOI: 10.32604/CMC.2020.09763.
- [26] Devi MS, Bajaj PR. Fuzzy based driver fatigue detection. *Conference Proceedings - IEEE International Conference on Systems, Man and Cybernetics*. 2010. p. 3139–3144. DOI: 10.1109/ICSMC.2010.5641788.
- [27] Savas BK, Becerikli Y. Real time driver fatigue detection based on SVM Algorithm. *2018 6th International Conference on Control Engineering and Information Technology, CEIT 2018*. 2018. DOI: 10.1109/CEIT.2018.8751886.
- [28] Wierwille WW, et al. *Research on vehicle-based driver status/performance monitoring: Development, validation, and refinement of algorithms for detection of driver drowsiness*. NTIS. Report No. DOT HS 808 247, 2019.
- [29] Caffier PP, Erdmann U, Ullsperger P. Experimental evaluation of eye-blink parameters as a drowsiness measure. *European Journal of Applied Physiology*. 2003;89(3–4):319–325. DOI: 10.1007/s00421-003-0807-5.
- [30] Bandara IB. Driver drowsiness detection based on eye blink. 2009. p. 1–4. [http://collections.crest.ac.uk/9782/1/Bandara Indra chapa - Thesis pdf.pdf](http://collections.crest.ac.uk/9782/1/Bandara%20Indrachapa%20-%20Thesis.pdf.pdf).
- [31] Huang R, Wang Y. P-FDCN based eye state analysis for fatigue detection. *2018 IEEE 18th International Conference on Communication Technology (ICCT)*. 2018. p. 1174–1178.
- [32] Devi MS, Bajaj PR. Driver fatigue detection based on eye tracking. *Proceedings - 1st International Conference on Emerging Trends in Engineering and Technology, ICETET 2008*. 2008. p. 649–652. DOI: 10.1109/ICETET.2008.17.
- [33] Tabrizi PR, Zoroofi RA. Open/closed eye analysis for drowsiness detection. *2008 1st International Workshops on Image Processing Theory, Tools and Applications, IPTA 2008*. 2008. DOI: 10.1109/IPTA.2008.4743785.
- [34] Mandal B, Li L, Wang GS, Lin J. Towards detection of bus driver fatigue based on robust visual analysis of eye state. *IEEE Transactions on Intelligent Transportation Systems*. 2017;18(3):545–557. DOI: 10.1109/TITS.2016.2582900.
- [35] Of O, Carriers M. PERCLOS: A valid psychophysiological measure of alertness as assessed by psychomotor vigilance. 1998;31(5):1237–1252. <http://scholar.google.com/scholar?hl=en&btnG=Search&q=intitle:PERCLOS+:+A+Valid+Psychophysiological+Measure+of+Alertness+As+Assessed+by+Psychomotor+Vigilance#0>.
- [36] Johns MW. The amplitude velocity ratio of blinks: A new method for monitoring drowsiness. *Sleep*. 2003;26:A51-52.
- [37] Trutschel U, et al. PERCLOS: An alertness measure of the past. *Driving Assessment Conference*. 2011. p. 172–179. DOI: 10.17077/drivingassessment.1394.
- [38] Guo X, et al. PFLD: A practical facial landmark detector. arXiv:1902.10859 [cs.CV]. DOI: 10.48550/arXiv.1902.10859.
- [39] Geng L, Liang X, Xiao Z, Li Y. Real-time driver fatigue detection based on morphology infrared features and deep learning. *Hongwai Yu Jiguang Gongcheng/Infrared and Laser Engineering*. 2018;47(2). DOI: 10.3788/IRLA201847.0203009.
- [40] Wang H, et al. A novel real-time driving fatigue detection system based on wireless dry EEG. *Cognitive Neurodynamics*. 2018;12(4):365–376. DOI: 10.1007/s11571-018-9481-5.
- [41] Jap BT, Lal S, Fischer P, Bekiaris E. Using EEG spectral components to assess algorithms for detecting fatigue. *Expert Systems with Applications*. 2009;36(2 PART 1):2352–2359. DOI: 10.1016/j.eswa.2007.12.043.
- [42] Söylemez ÖF, Ergen B. Eye location and eye state detection in facial images using circular Hough transform. *Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics), 8104 LNCS*. 2013. p. 141–147. DOI: 10.1007/978-3-642-40925-7\_14.
- [43] Cheng E, Kong B, Hu R, Zheng F. Eye state detection in facial image based on linear prediction error of wavelet coefficients. *2008 IEEE International Conference on Robotics and Biomimetics, ROBIO 2008*. 2009. p. 1388–1392. DOI: 10.1109/ROBIO.2009.4913203.
- [44] Miah AA, Ahmad M, Mim KZ. Drowsiness detection using eye-blink pattern and mean eye landmarks' distance. *Proceedings of International Joint Conference on Computational Intelligence. Algorithms for Intelligent Systems*. 2020. p. 111–121. DOI: 10.1007/978-981-13-7564-4\_10.

[45] Liu H, Soto RAR, Xiao F, Lee YJ. YolactEdge: Real-time instance segmentation on the edge (Jetson AGX Xavier: 30 FPS, RTX 2080 Ti: 170 FPS). 2020. <http://arxiv.org/abs/2012.12259>.  
 [46] Tsuyama M, et al. Embedded implementation of human detection using only color features on the NVIDIA Xavier. *Proceedings - 2019 International Symposium on Intelligent Signal Processing and Communication Systems, ISPACS 2019*. 2019. DOI: 10.1109/ISPACS48206.2019.8986281.

凌艳城，翁小雄

高效稳健的基于眼部状态视觉分析的驾驶员疲劳检测框架

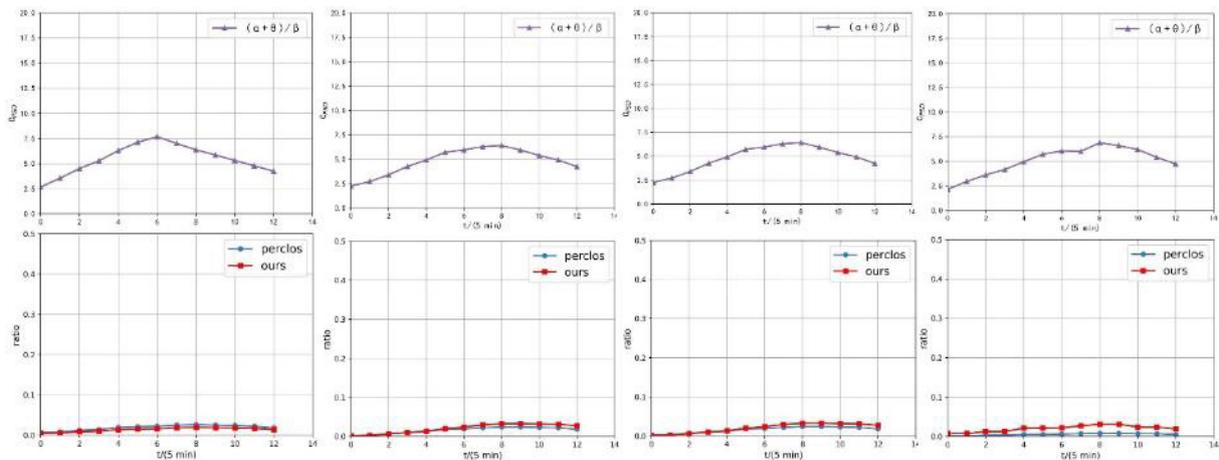
摘要

基于视觉的疲劳检测在车辆中被广泛应用，由于它具有实时和可靠的检测结果。随着2019年冠状病毒病（COVID-19）的爆发，许多基于面部特征的检测系统由于口罩的遮盖而变得不可靠。本文提出了一种基于视觉的疲劳检测系统，用于监测驾驶员，该系统在口罩遮盖、光照变化和驾驶员头部运动等条件下具有稳健性。我们的系统包括三个主要模块：面部关键点对齐、疲劳特征提取和基于融合特征的疲劳度量。创新的核心技术描述如下：1) 通过融合全局面部信息和局部眼部信息实现稳健的关键点对齐算法，2) 采用动态阈值方法提取疲劳特征，3) 基于眼睑闭合百分比（PERCLOS）和长时间闭合的眨眼比例（PLCDB）融合实现稳定的疲劳度量。实验证明了我们提出的算法和方法的出色性能。实验结果显示，我们的关键点对齐算法对不同场景具有稳健性，而我们提出的疲劳度量由于融合了PERCLOS和PLCDB而更可靠。

关键词：

疲劳检测；基于视觉的；融合；PERCLOS；PLCDB

APPENDIX – The state measurement by EEG signals, PERCLOS and improved methods for 14 subjects

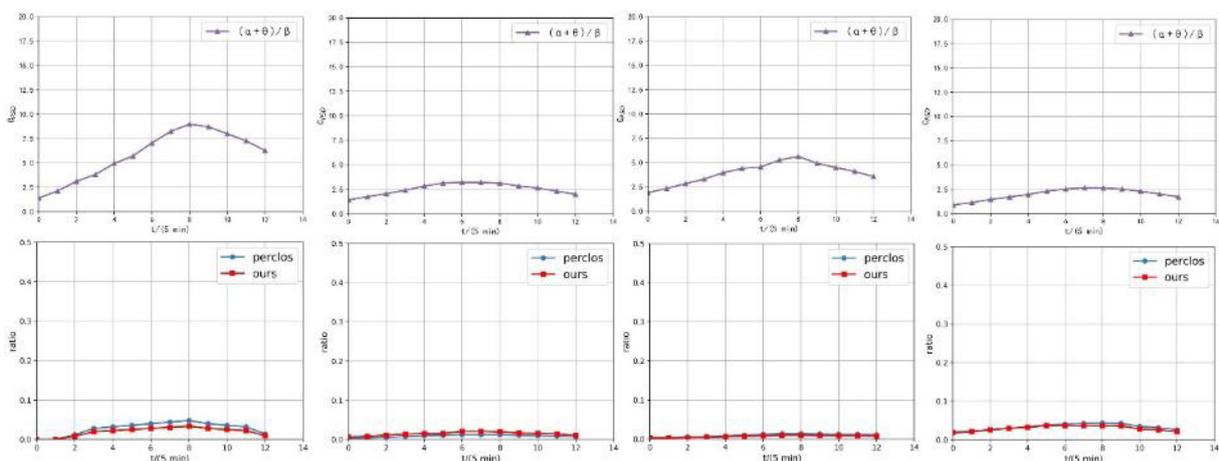


a) Subject 1

b) Subject 3

c) Subject 4

d) Subject 5

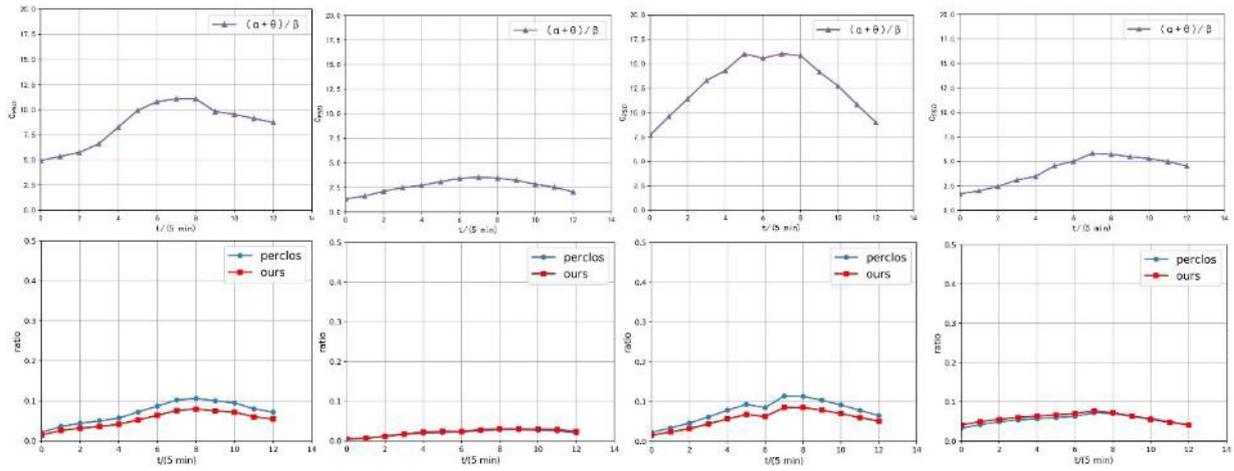


e) Subject 8

f) Subject 9

g) Subject 10

h) Subject 11

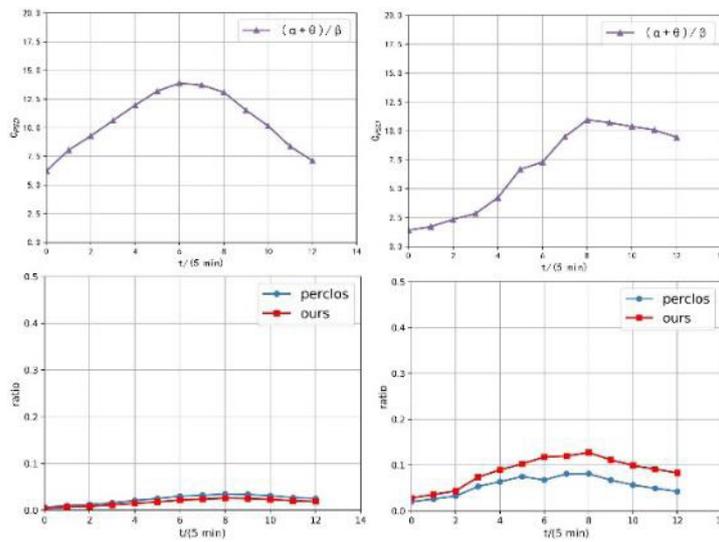


i) Subject 12

j) Subject 13

k) Subject 14

l) Subject 15



m) Subject 16

n) Subject 18