Perception-Free Calibration of Eye Opening and Closing Threshold for Driver Fatigue Monitoring

CHENG MING1,2 AND YAN YUNBING1

1School of Automotive and Transportation Engineering, Wuhan University of Science and Technology, Wuhan, Hubei 430200, China
2Intelligent Software Center, Dongfeng Motor Group Technology Center, Wuhan, Hubei 430056, China

Corresponding author: Yan Yunbing (yanyunbing@wust.edu.cn)

This work was supported in part by the National Natural Science Foundation of China under Grant 51975428 and Grant 52002298; and in part by the Key Research and Development Project of Hubei Province, Research and Application of Key Technologies of Advanced Autonomous Driving in Complex Driving Environment, under Grant 2020AAA001.

ABSTRACT Analyzing the opening and closing states of eyes and mouths by detecting the driver’s face feature points is an effective method for judging driver fatigue. However, in practical engineering applications, with the expansion of user groups, the false identification problem caused by the differentiation of individual facial features of drivers is prominent, especially for people with small eyes. To solve this problem, this paper uses the Mediapipe Facemesh module to detect face feature points and designs a perception-free calibration method for setting personalized eye opening and closing threshold combined with head postures. Compared with the traditional method of setting a fixed threshold, the precision of eye state recognition is improved by 36.4%. Finally, the model deployment and post-processing compilation are completed on the Xavier vehicle chip, achieving a running speed of 34 frames per second at most, and the subjective evaluation experience of the fatigue monitoring system is significantly improved.

INDEX TERMS Small eyes, fatigue monitoring, Mediapipe, face feature points, perception-free calibration.

I. INTRODUCTION Through the analysis of the predisposing factors of China’s road traffic accidents, it is shown that the main causes of traffic accidents lie in the drivers. According to the report of the American Automobile Association, accidents caused by driver fatigue account for the largest proportion, which is about 7% of total accidents and 21% of fatal traffic accidents [1]. Therefore, the monitoring function of driver fatigue degree needs to be applied in engineering. Currently, there are three types of fatigue monitoring methods, namely the methods based on visual detection, physiological parameters monitoring, and the behavioral states of drivers with onboard sensor monitoring. The mainstream idea of the first type of method is to first detect the driver’s face feature points, judge the opening and closing degree of eyes and mouth and the opening and closing time, and comprehensively determine the fatigue degree. Some methods adopt the end-to-end technology route, such as the CNN+LSTM method, to analyze image features and temporal features to judge the dynamic behavior of drivers in a certain period [2], [3]. Additionally, a large number of optimization studies have been conducted based on the C3D model [4], in which drivers’ behavioral characteristics are learned by 3D convolution. For example, Zhuang and Qi [5] proposed a method of combining the pseudo-3D (P3D) convolutional neural network and attention mechanism to improve the reasoning speed of the C3D model and compress parameters. The fatigue detection precision on the YawDD dataset was as high as 98%. However, after the compression of the model, the inference time is 660 ms, which does not meet the requirements of onboard real-time monitoring. The second type of method is based on physiological parameters, such as brain-computer sensing, skin sensing, and other methods, which require physical contact between the driver and the sensor [6], and the application scope is limited by the high hardware cost. The third type of method is based on behavioral states of drivers with onboard sensor monitoring [7]. This is a low-cost solution because no additional sensors are required. However, this method is
fatigue state is easy to occur. Thus, reasonably setting the threshold of eye opening and closing has become a new difficulty. At present, few scholars have studied the facial differentiation of individual drivers. The Summer [8] fatigue detection scheme distinguishes the driver’s individual facial information and uses an adaptive algorithm to compensate for feature point position errors, thus effectively reducing the distortion of the face and improving the accuracy of face recognition. However, this scheme does not solve the problem of facial differentiation. This study will focus on this issue, and the main contributions are as follows: 1. The top-down structure is adopted. Firstly, the driver’s head in the original image is set as the ROI region of interest to reduce the size of the image input to the network. Then, the clipped image is sent to the Mediapipe Facemesh model to detect the face frame and face feature points. 2. Based on the detection results of face feature points, a perception-free calibration scheme is innovatively designed for setting personalized eye opening and closing threshold to solve the problem of false recognition caused by drivers’ facial differences. Combined with the head posture, this method can reasonably limit the differentiation threshold. 3. By using the idea of the Perclos algorithm, the post-processing judgment logic is designed based on the opening and closing degrees of eyes and mouth, and the judgment criteria of first-degree fatigue and second-degree fatigue are proposed, which helps to enrich the strategies and approaches of warning drivers. 4. The model and post-processing logic code was compiled by BAZEL and successfully deployed in the Xavier vehicle controller, which can run normally at 26-34 frames per second. Twenty-three workers of genders and ages were invited to participate in the test. The test results showed that the recognition precision of the eye state was improved by 36.4% and the subjective evaluation result of the fatigue monitoring system was significantly improved after the addition of the no-perception calibration scheme.

Hence, in practical engineering applications, with the expansion of all users, the differentiation of individual drivers’ facial features becomes more obvious. Especially, for people with small eyes, the misidentification of a fatigue state is easy to occur. Thus, reasonably setting the threshold of eye opening and closing has become a new difficulty. At present, few scholars have studied the facial differentiation of individual drivers. The Summer [8] fatigue detection scheme distinguishes the driver’s individual facial information and uses an adaptive algorithm to compensate for feature point position errors, thus effectively reducing the distortion of the face and improving the accuracy of face recognition. However, this scheme does not solve the problem of facial differentiation. This study will focus on this issue, and the main contributions are as follows: 1. The top-down structure is adopted. Firstly, the driver’s head in the original image is set as the ROI region of interest to reduce the size of the image input to the network. Then, the clipped image is sent to the Mediapipe Facemesh model to detect the face frame and face feature points. 2. Based on the detection results of face feature points, a perception-free calibration scheme is innovatively designed for setting personalized eye opening and closing threshold to solve the problem of false recognition caused by drivers’ facial differences. Combined with the head posture, this method can reasonably limit the differentiation threshold. 3. By using the idea of the Perclos algorithm, the post-processing judgment logic is designed based on the opening and closing degrees of eyes and mouth, and the judgment criteria of first-degree fatigue and second-degree fatigue are proposed, which helps to enrich the strategies and approaches of warning drivers. 4. The model and post-processing logic code was compiled by BAZEL and successfully deployed in the Xavier vehicle controller, which can run normally at 26-34 frames per second. Twenty-three workers of genders and ages were invited to participate in the test. The test results showed that the recognition precision of the eye state was improved by 36.4% and the subjective evaluation result of the fatigue monitoring system was significantly improved after the addition of the no-perception calibration scheme.

II. FACE FEATURE POINT DETECTION BASED ON MEDIAPIPE

Mediapipe has attracted much attention since it was publicly released by Google. It includes rich frameworks and directly callable solution modules, of which the Face Mesh module and the iris recognition module are used in this study. The Face Mesh module is developed based on Blazeface [9], which is optimized for mobile GPU reasoning and improved based on the backbone of MobilenetV2 and SSD detector. In this study, a total of 468 feature points are detected by the Facemesh module, and 10 feature points are detected by the iris detection module. The details of mouths and eyes are marked, and the detection ability of large head postures is strong, as shown in Figure 1.

III. PERCEPTION-FREE CALIBRATION SCHEME FOR SETTING THE EYE OPENING AND CLOSING THRESHOLD

The perception-free calibration scheme for setting the eye opening and closing threshold is the main method to solve the problem of fatigue misidentification caused by individualized differences in drivers’ faces in this paper. Firstly, the definition of eye opening and closing threshold is clarified. Eye threshold refers to the ratio of width and height of two eyes, which can be divided into left eye threshold Ratio_left (RL) and right eye threshold Ratio_right (RR). The formula is shown in (1).

\[
\begin{align*}
RL &= \frac{W_l}{L_l} \\
RR &= \frac{W_r}{L_r}
\end{align*}
\]  

(1)

where \(W_l, L_l\) represent the width and height of the left eye; \(W_r, L_r\) represent the width and height of the right eye. As shown in Figure 2. In the traditional scheme, RL and RR are set as fixed values [10], [11]. When the value is less than this threshold, the eyes are considered to be closed; otherwise, the eyes are considered to be open. However, the fixed eye opening and closing threshold may cause misrecognition for people with small eyes, which does not consider drivers’ facial individualization differences.
Our method aims to make the fatigue detection system acquire drivers’ eye opening and closing threshold without sensing, and distinguish drivers’ facial differences. This method has been upgraded and changed many times in the research process. First, the simplest pre-entry method was used, and the user needed to manually trigger the entry instruction. When the input program is started, the user makes corresponding actions according to the prompts, such as opening and closing eyes. The camera collects face feature points and analyzes the left and right eye thresholds according to the geometric relationship of feature points in a single frame. However, the main problem of this method is that the collected threshold is not bounded to the driver, and the feature points collected in a single frame are unstable. Then, the method was optimized, and the FaceNet [12] face recognition module was adopted to collect face feature points and bind them with drivers to establish a driver information base. Meanwhile, single-frame acquisition was changed to multi-frame acquisition, and the eye opening and closing threshold was calculated from the average value of multi-frame data, which improved the credibility of the eye threshold. However, there are still two main problems with this method: (1) users have to manually trigger the face entry and recognition module, which leads to an unsatisfactory user experience because users prefer to adopt the perception-free threshold calibration method; (2) although multi-frame acquisition is used, the eye threshold changes when the user turns his/her head. Therefore, the head angle and eye opening and closing threshold must be integrated to formulate the standard for fatigue judgment. The following sections will introduce how to fuse the head angle information into the eye opening and closing threshold, as well as “perception-free calibration”, and the methods of face recognition input and driver information base establishment will not be repeated.

After the face key point is detected by Mediapipe, the head pose angle can be solved by the classic PNP algorithm [13], and it is described here. The six points (represented by C1~C6) of the nose tip, chin, left corner of eye, right corner of eye, left corner of mouth, and right corner of mouth are selected as the key points to construct the head coordinate system. Then, the key 3D template coordinates of face are denoted as $C = \{C_1, C_2, C_3, C_4, C_5, C_6\}$, as shown in Equation 2.

$$C = \{C_1, C_2, C_3, C_4, C_5, C_6\}$$

$$C_1 = (0, 0, 0)$$

$$C_2 = (0, -330, -65)$$

$$C_3 = (-225, 170, -135)$$

$$C_4 = (225, 170, -135)$$

$$C_5 = (-150, -150, -125)$$

$$C_6 = (150, -150, -125)$$

(2)

The 2D image coordinates corresponding to these 6 points detected by face feature points are denoted as shown in Equation 3.

$$U = \{U_1, U_2, U_3, U_4, U_5, U_6\}$$

(3)

In addition, the internal parameters and distortion parameters of the camera were obtained by the calibration method proposed by Zhang [14], as shown in Equations 4.

$$\text{Camera matrix} = \begin{bmatrix} f_x & 0 & C_x \\ 0 & f_y & C_y \\ 0 & 0 & 1 \end{bmatrix}$$

(4)

where $(f_x, f_y)$ is the focal length of the camera, and $(C_x, C_y)$ is the optical center.

$$U = S \ast \text{Camera matrix} \ast \begin{bmatrix} R & T \\ 0 & 1 \end{bmatrix} \ast C$$

(5)

where $S$ denotes the scaler of the scaling factor. $C$ is the 3D template coordinate of the feature points shown in Equation(2), $U$ is the 2D coordinate corresponding to these feature points shown in Equation(3), and the Camera matrix is the internal parameter of the camera shown in Equation(4). These parameters are known. The main purpose of Equations (5) is to solve the rotation vector $R$ and translation vector $T$ of the head relative to the camera through these known parameters. This is actually the problem of solving the head pose by the PNP method. As early as 1989, Horaud [15] has described the mathematical method of solving $R$ and $T$ by these formulas, and this will not be repeated in this paper.

Then, the rotation vector is converted into a rotation matrix according to the Rodriguez formula to calculate the Euler angle of the head. Denote the yaw angle, pitch angle, and roll angle of the head as $Y$, $P$, and $R$, respectively. After the head angle is obtained, this information is integrated into the eye opening and closing threshold. Thus, Equation (2) is rewritten as Equation (6).

$$\begin{align*}
RL_{y,p,r} &= \frac{W_1}{L_1}, \quad (Y, P, R) \\
RR_{y,p,r} &= \frac{W_2}{L_2}, \quad (Y, P, R)
\end{align*}$$

(6)
where $RL_{y,p,r}$ and $RR_{y,p,r}$ are respectively the opening and closing thresholds of left and right eyes when the head Euler angle is $Y$, $P$, $R$, i.e., a set of eye thresholds will be saved under a combination of angles. Obviously, this combination method is complicated and is not conducive to practical applications, so the number of parameters needs to be reduced. The experiment under the camera position shown in Figure 2 indicates that it is meaningful to detect the eye threshold only when the driver’s left-right shaking angle $Y$ is within $\pm 60^\circ$ and the angle $P$ of the lowering head is within positive $30^\circ$ to negative $45^\circ$. This is because when the head posture exceeds this angle range, one eye or both eyes have been obscured. In addition, it was found that the head sticking to the shoulder with ears in the direction of head rotation does not affect the eye opening and closing threshold, so the angle $R$ is negligible. From this analysis, the effective range of the driver’s head angle is obtained, which is shown in Equation 7.

$$\begin{cases} 
-60^\circ \leq Y \leq 60^\circ \\
-45^\circ \leq P \leq 30^\circ
\end{cases}$$ (7)

Therefore, the number of eye opening and closing threshold parameters to be saved, i.e., $Len\ (list)$, can be calculated by Equation 8, where $Len\ (Y)$ is the value of the yaw angle and $Len\ (P)$ is the value of the pitch angle. The angle combination can be calculated by multiplying the two values:

$$Len\ (list) = Len\ (Y) \times Len\ (P)$$ (8)

According to Equations 7 and 8, when the yaw angle and pitch angle are divided by 1°, $Len\ (Y)$ is equal to 120, $Len\ (P)$ is equal to 75, and $Len\ (list) = 120\times 75 = 9000$. This quantity is still too large. Considering that the eye threshold under each head pose should not be omitted and the amount of parameters should be reduced, we adopted the idea of fuzzing and divided yaw and pitch by 15° conservatively. This reduces the amount of parameters without missing too much information. For example, a threshold ranges from 0° to 15° and from 15° to 30°. That is, $Len\ (Y) = 120/15$, $Len\ (P) = 75/15$, and $Len\ (list) = 40$. In this way, the data to be saved will be greatly reduced, which is conducive to practical applications. The value under each head angle interval is denoted as shown in Equation 9.

$$List\ [i], \ i = [0 \sim Len\ (list))$$ (9)

By combining Equation 6 and Equation 9, Equation 10 can be obtained:

$$\begin{cases} 
List\ [i]\ [0] = \frac{\sum_{i=0}^{st} RL_{y,p,r}}{st}, \ i = [0 \sim Len\ (list)) \\
List\ [i]\ [1] = \frac{\sum_{i=0}^{st} RR_{y,p,r}}{st}, \ i = [0 \sim Len\ (list))
\end{cases}$$ (10)

where, $List\ [i]\ [0]$ and $List\ [i]\ [1]$ respectively represent the opening and closing thresholds of the left eye and the right eye under the angle interval $i$, and $st$ represents the effective times to be counted under each angle interval. When the times reach $st$, the eye threshold under the current angle is equal to the average of these values. If the number of times does not reach $st$, then the eye opening and closing threshold under the current angle interval will not take effect, and the driver’s eye information will continue to be collected without perception. Thus, $st$ is a sensitive parameter of this system. If the value of $st$ is too small, the eye threshold calibration stage will end quickly and enter the fatigue analysis stage; if the value of $st$ is too large, the calibration phase of eye threshold will be long, but the statistical data will be highly stable. In this experiment, the sensitive parameter $st$ is set to 35, that is, when the effective collection times in an angle interval reach 35, the fatigue analysis function in the current angle interval will be activated, and the angle interval before 35 times will not be activated.

Here, the effectiveness of $List\ [i]$ is considered. Not every collected $List\ [i]$ will “add 1” to $st$, which is the key to achieving “perception-free calibration”. There are two important assumptions in this paper: (1) the driver has just gone into the car and started the engine. After the speed exceeds 30 km/h for the first time, he/she must be awake for a period of time; (2) the driver has far more eyes open than closed. The significance of the first assumption is that when the vehicle speed reaches 30 km/h for the first time after the engine start, the eye opening and closing threshold can be collected. Since many drivers will show “pseudo-fatigue” when they just get in the car, such as yawning and listlessness, but when the speed rises to 30 km/h for the first time, the driver will gradually wake up. In addition, the reason why 30 km/h is set here is that above this speed, the level of traffic accidents will rise, so this speed is the initial condition for awakening fatigue warning. The significance of the second assumption is that the individual thresholds are much smaller than most thresholds in a certain angle interval, and these individual thresholds are considered to be the driver’s blinking eyes, that is, this data can be saved as a closed eye state. Here, the principle of Gaussian normal distribution is used to distinguish the data with open eyes from those with closed eyes. Specifically, the data within $-2\sigma$ to $\infty$ are saved as the eye-opening threshold, and the data less than $-2\sigma$ and greater than 0 are saved as the eye-closing threshold. By using the same method, the mouth opening and closing threshold can be determined. However, since the mouth opening and closing boundary is very obvious and easy to judge, the mouth opening and closing threshold is set as a fixed value in this study.

IV. FATIGUE JUDGMENT LOGIC BASED ON MULTIPLE FACIAL FEATURES

The most prominent characteristics of driver fatigue are eye closure and yawning, which reflects the degree and time of opening and closing the eyes and mouth. Therefore, if the opening and closing threshold of eyes and mouth can be set accurately, and the threshold of opening and
closing time can be defined by formulated strategies, whether the driver is in fatigue can be determined. In Section 2, how to set the opening and closing threshold of eyes has been introduced in detail. In this section, the post-processing logic of driver fatigue determination is mainly elaborated.

First, the judgment logic of driver fatigue is shown in Figure 4, where two methods are used to determine whether the driver is tired.

A. DETERMINE EYE FATIGUE

The most classic method to judge fatigue by eye state is PERCLOS [16], i.e., Percentage of Eyelid Closure over the Pupil. The core idea is to obtain the proportion of frames closed by dividing the number of frames closed by the total number of frames detected. In the specific experiment, there are three measurement methods, P70, P80, and EM. Among them, P80 is considered the optimal method to characterize driver fatigue. This method thinks that the area of the eyelid covering the pupil is at least greater than the area of the eye, and the percentage of closing time is taken as the criterion. This method can be understood in two aspects. The first aspect is the need to judge in what state the eyes closed, and the other aspect is the expression of the time the eyes closed as fatigue. Our research still adopts this classic idea, but considering the facial differences of diverse users, the following optimizations are made.

The experimental results show that when the yaw angle Y and pitch angle P are within the range of ±15°, the eye opening and closing threshold is between 0.12 and 0.45, with obvious differentiation, and the threshold for most people concentrates on about 0.26. If the threshold of eye closing is set as a fixed value such as 0.2, it will inevitably cause fatigue misidentification of people with small eyes and misjudge eye opening as eye closing. Meanwhile, if the head posture is not considered, when the head-turning angle is large, the actual eye opening and closing ratio captured by the camera at the fixed position will be reduced, resulting in misjudgment. Therefore, it is necessary to calibrate the eye opening and closing threshold with head postures.

The next step is to count the time spent on closing and opening eyes to determine fatigue. In continuous frame detection, the proportion of frames with eyes closed in frames with full detection is defined as the fatigue index, denoted as S. In Equation (11), \( FPS_{\text{close}} \) denotes the number of frames with eyes closed, and \( FPS_{\text{all}} \) denotes the number of frames with full detection.

\[
S = \frac{FPS_{\text{close}}}{FPS_{\text{all}}} \tag{11}
\]

When \( S \leq 0.25 \), the driver is in a normal state; when \( 0.25 \leq S \leq 0.4 \), the driver is in grade one fatigue; when \( 0.4 \leq S \), the driver is in grade two fatigue. In actual engineering applications, corresponding vigilance measures should be formulated according to the fatigue grade. In this experiment, 180 frames are set as a statistical period. That is, if 180 * 0.25 = 45 fps occurs in this detection period, it is considered that the driver enters the fatigue stage, and the first-level fatigue warning is required.

B. DETERMINE THE FREQUENCY OF YAWNING

Yawning is mainly judged according to the opening and closing degree and time of the mouth, and six key points of the mouth are selected as the judgment basis, as shown in Figure 2. These six points are collected for \( P = P_i^{x,y} \) (i=1~6), where \( P_i^{x,y} \) represents point i in the x and y coordinates. The ratio of mouth opening and closing can be calculated by the following equations:
value of $P$, as shown in Equation 12, where distance represents the Euclidean distance and $R_{\text{mouth}}$ represents the actual ratio of mouth opening and closing.

$$R_{\text{mouth}} = \frac{\text{distance}(P_3^{x,y} + P_5^{x,y}) + \text{distance}(P_4^{x,y} + P_6^{x,y})}{2 \times \text{distance}(P_1^{x,y} + P_2^{x,y})}$$ (12)

Since it is easy to distinguish the opening and closing of the mouth, in this study, the threshold of mouth opening and closing ratio is set to a fixed value of 0.2. When $R_{\text{mouth}}$ is less than this value, it indicates mouth closing. To distinguish between yawning and speech opening, a yawn is detected by mouth opening in consecutive frames, and the number of consecutive frames must exceed a set value $K$. In the experiment, to enhance user experience, after repeated debugging, $K$ is set to 35, that is, 35 consecutive frames of mouth opening is considered as one yawn. When two yawns appear within 40 seconds, it is considered grade-1 fatigue; when there are four yawns in 40 seconds, it is considered grade-2 fatigue.

V. SIMULATION EXPERIMENT

The controller used in this study is TITAN4C, which is based on Nvidia Xavier and NXP MPC57XX series MCU. It adopts GPU, CPU, and MCU to achieve a heterogeneous mode of high computing power, high performance, and high reliability. It integrates two Nvidia Xavier core computing units with a computing capability of 64 TOPS and runs Ubuntu 18.04 operating system. The camera is an SG-V4 model with a horizontal and vertical angle of $\pm 60^\circ$ and $\pm 40^\circ$, and the frame rate is 30 FPS.

First, Bazel-5.2.0 is installed on TITAN4C ARM64, and the OpenCV dependency library is configured. Meanwhile, the ROS intermediate layer environment is built, the RViz visualization software is installed, and the original Mediapipe file is downloaded through Github. After setting up the environment and installing the relevant software, the executable program is compiled by using Bazel, and the program is run. Then, the RViz software is opened to view the visualization result in the terminal.

In the simulation experiment, the CANalyzer tool was used to simulate the speed signal. When the simulated speed signal reached 30 km/h, the fatigue monitoring system was awakened, the eye opening and closing threshold under various head postures was calibrated, and the fatigue state was monitored in real-time. Twenty-three staff members with significant differences in appearance, gender, and age were invited to participate in the experiment, and the YawDD dataset [17] was used for testing.

Firstly, the participants were invited, and the eye threshold values were collected under various head postures by a non-perceptual eye threshold calibration method. The larger threshold value of the left and right eye is abbreviated as TS in Table 1. Then the self-made data and public data were filtered and clarified to construct a validation set. A total of 169 video pieces of video were collected as validation data, and each video was about 28 seconds, of which 92 were self-made data and 77 were obtained from YawDD public data. Then, these videos were divided into a total of 7581 images by video frame splitting method, and the images were labeled respectively. Label information includes eye opening and closing status, head pose Angle, and eye threshold. We took the eye threshold as the primary key and counted the photos of different eye states. The distribution of data samples is shown in Table 1.

The visualization effect is presented in Figure 7, where the eye and mouth thresholds and driver status are displayed on the interface for direct scoring, frames per second (fps) is also
TABLE 1. Validation data distribution.

| Number | Threshold value | Open eye | Close eye |
|--------|-----------------|----------|-----------|
| 1      | 0.12<TS≤0.14    | 307      | 105       |
| 2      | 0.14<TS≤0.16    | 361      | 78        |
| 3      | 0.16<TS≤0.18    | 234      | 198       |
| 4      | 0.18<TS≤0.20    | 552      | 290       |
| 5      | 0.20<TS≤0.22    | 752      | 168       |
| 6      | 0.22<TS≤0.24    | 983      | 373       |
| 7      | 0.24<TS≤0.26    | 1281     | 202       |
| 8      | 0.26<TS         | 1409     | 288       |

TABLE 2. Precision and recall for each threshold.

| Threshold value | fixed thresholds method | no-perception eye calibration method |
|-----------------|-------------------------|--------------------------------------|
|                 | Precision  | Recall  | Precision | Recall  |
| 0.12<TS≤0.14    | 0.371      | 0.156   | 0.735     | 0.645   |
| 0.14<TS≤0.16    | 0.681      | 0.612   | 0.886     | 0.861   |
| 0.16<TS≤0.18    | 0.900      | 0.842   | 0.951     | 0.932   |
| 0.18<TS≤0.20    | 0.964      | 0.958   | 0.975     | 0.974   |
| 0.20<TS≤0.22    | 0.987      | 0.995   | 0.983     | 0.992   |
| 0.22<TS≤0.24    | 0.990      | 0.992   | 0.981     | 0.979   |
| 0.24<TS≤0.26    | 0.982      | 0.986   | 0.972     | 0.980   |
| 0.26<TS         | 0.984      | 0.983   | 0.977     | 0.983   |

It needs to be explained here, because the participants are unwilling to disclose their photos in the article, and this is their right. So in Figure 7, there are only pictures of the public data. Meanwhile, some participants have actually been shown in Figure 1. Many feature points have been marked on these figures, which makes it difficult to carry out face recognition. Therefore, participants agree with the presentation method in Figure 1. The confusion matrix for each threshold is shown in Figure 8, A means eyes open and B means eyes closed. The precision and recall for each threshold are listed in Table 2. The front and back comparison chart after adding the no-perception eye calibration method is presented in Figure 9.

The analysis of the experimental data led to the following observations: (1) in the face of the camera, the detection precision of the fixed-threshold algorithm started to drop off like a cliff for people whose average eye-opening ratio is less than 0.16, and the fixed-thresholds fatigue detection system is no longer practical. When eyes are small, the fixed threshold method can easily misidentify the open state as the closed eye; (2) the detection precision was significantly improved after the addition of the no-perception eye calibration method. Compared with the non-opening of the calibration procedure, the eye state recognition precision was up to 36.4% when eyes threshold value are smallest, and the subjective evaluation experience of the fatigue monitoring system was better; (3) After the addition of the no-perception eye calibration method, frame rate can still reach above 30fps, the average reasoning time of a single frame image only increases by 0.039 seconds. The computer only needs to calculate the head pose and contrast the eye threshold according to the face feature points, and the detection of face feature points is shown.
also needed in the method of fixed threshold. Therefore, this method does not add additional machine learning models that lead to a long reasoning time, and this is also verified by the experimental frame rate.

VI. CONCLUSION
Driver fatigue monitoring is an important application scenario for visual inspection schemes in the field of automobile safety. This scheme has the advantages of low cost and remarkable effect, but there is still a certain gap between scientific research results and application requirements. With the rapid popularization of this system, the problem of false recognition caused by individual differences in the facial features of users is common. Also, this system is not accurate enough for people with small eyes, which has become a serious problem in the industry. The main purpose of our research is to increase the practicability of the driver fatigue monitoring system and solve the practical problems encountered at present. This study proposes a perception-free calibration scheme, which effectively increases the precision of eye state recognition and improves the practical value of this system. In our follow-up research, we will explore more effective scientific methods from practical applications.

ACKNOWLEDGMENT
Here, I sincerely thank all the reviewers, as well as MJEditor (www.mjeditor.com) for providing English editing services during the preparation of this manuscript.

REFERENCES
[1] B. C. Tefft, “Acute sleep deprivation and culpable motor vehicle crash involvement,” Sleep, vol. 41, no. 10, Oct. 2018, doi: 10.1093/sleep/csy144.
[2] L. Geng, X. Liang, Z. Xiao, and Y. Li, “Real-time driver fatigue detection based on morphology infrared features and deep learning,” Infr. Laser Eng., vol. 47, no. 2, 2018, Art.no. 203009, doi: 10.3788/irla201847.0203009.
[3] W. Zhang and J. Su, “Driver yawning detection based on long short term memory networks,” in Proc. IEEE Symp. Ser. Comput. Intell. (SSCI), Nov. 2017, pp. 1–5, doi: 10.1109/SSCI.2017.8285343.
[4] D. Tran, L. Bourdev, R. Fergus, L. Torresani, and M. Paluri, “Learning spatiotemporal features with 3D convolutional networks,” in Proc. IEEE Int. Conf. Comput. Vis. (ICCV), Dec. 2015, pp. 4489–4497, doi: 10.1109/iccv.2015.510.
[5] Z. Zhuang and Y. Qi, “Fatigue driving detection based on pseudo-3D convolutional neural network and attention mechanism,” Chin. J. Image Graph., vol. 26, no. 1, pp. 143–153, 2018.
[6] R. P. Balandong, R. F. Ahmad, M. N. M. Saad, and A. S. Malik, “A review on EEG-based automatic sleepiness detection systems for driver,” IEEE Access, vol. 6, pp. 22908–22919, 2018, doi: 10.1109/ACCESS.2018.2811723.
[7] Z. Li, S. E. Li, R. Li, B. Cheng, and J. Shi, “Online detection of driver fatigue using steering wheel angles for real driving conditions,” Sensors, vol. 17, no. 3, p. 495, 2017.
[8] L. Xie, “Fatigue detection based on eye texture features and identity recognition,” Shanghai Jiao Tong Univ., Shanghai, China, 2019.
[9] V. Bazarevsky, Y. Kartyunik, A. Vakanov, K. Raveendran, and M. Grundmann, “BlazeFace: Sub-millisecond neural face detection on mobile GPUs,” 2019, arXiv:1907.05047.
[10] T. Zhu, C. Zhang, T. Wu, Z. Ouyang, H. Li, X. Na, J. Liang, and W. Li, “Research on a real-time driver fatigue detection algorithm based on facial video sequences,” Appl. Sci., vol. 12, no. 4, p. 2224, Feb. 2022, doi: 10.3390/app12042224.
[11] J. Pan, Z. Liu, and Q. Wang, “Fatigue driving detection based on eye self-quotient graph and gradient graph co-occurrence matrix,” China Image Graph., vol. 26, no. 1, p. 11, 2021.
[12] H. Ding, S. K. Zhou, and R. Chellappa, “FaceNet2ExpNet: Regularizing a deep face recognition net for expression recognition,” in Proc. 12th IEEE Int. Conf. Autom. Face Gesture Recognit. (FG), May 2017, pp. 118–126, doi: 10.1109/FG.2017.23.
[13] H. Dai, S. Tan, and W. Zhang, “A survey of head pose Estimation methods,” Modern Comput., no. 7, pp. 130–144, 2021.
[14] Z. Zhang, “A flexible new technique for camera calibration,” IEEE Trans. Pattern Anal. Mach. Intell., vol. 22, no. 11, pp. 1330–1334, 2000.
[15] R. Horaud, B. Conio, O. Lebouleux, and B. Lacolle, “An analytic solution for the perspective 4-point problem,” Comput. Vis., Graph., Image Process., vol. 47, no. 1, pp. 33–44, Jul. 1989.
[16] D. F. Dinges and R. Grace, “Perclos: A valid psychophysiological measure of alertness as assessed by psychomotor vigilance,” US Dept. Transp., Federal Highway Admin., Washington, DC, USA, Tech. Rep. MCRT-98-006, 1998.
[17] S. Abtahi, M. Omideyeghan, S. Shirmohammadi, and B. Hariiri, “YawDD: A yawning detection dataset,” in Proc. 5th ACM Multimedia Syst. Conf., 2014, pp. 24–28.

CHENG MING was born in Wuhan, Hubei, China. He received the master’s degree in electronic information integrated circuits from Wuhan University. He is currently pursuing the Ph.D. degree in autonomous driving with the School of Automobile and Transportation, Wuhan University of Science and Technology. Since 2019, he has been working with the Dongfeng Automobile Technology Center, Intelligent Software Department, and engaged in the development of an intelligent vehicle driving assistance system. He leads and takes charge of the research, development, and application of L2-L3 level autonomous vehicle driver fatigue monitoring system, and is committed to promoting the engineering implementation of deep learning technology in the field of visual perception and decision planning.

Mr. Ming won the Silver Award of the National Innovation and Entrepreneurship Competition with the project on key technology in the intelligent cockpit, in 2020.

YAN YUNBING was born in Hubei, China, in 1968. He received the B.S. degree in vehicle engineering from Chongqing University, in 1990, and the M.S. and Ph.D. degrees in vehicle engineering from the Wuhan University of Technology, in 1995 and 2008, respectively.

From 1998 to 2008, he was an Assistant Professor at the School of Machinery and Automation, Wuhan University of Science and Technology. Since 2008, he has been a Professor with the School of Vehicle and Traffic Engineering. He is the author of three books, more than 100 articles, and more than 30 inventions. His main research interests include vehicle dynamics and its control, electric vehicle energy management, and intelligent vehicle technology.

* * *