

Research Article

A Fatigue Driving Detection Algorithm Based on Facial Motion Information Entropy

Feng You,^{1,2} Yunbo Gong,¹ Haiqing Tu,¹ Jianzhong Liang,¹ and Haiwei Wang^{1,3} 

¹School of Civil Engineering and Transportation, South China University of Technology, Guangzhou 510640, China

²State Key Lab of Subtropical Building Science, South China University of Technology, Guangzhou, China

³School of Transportation and Economic Management, Guangdong Communication Polytechnic, Guangzhou 510650, China

Correspondence should be addressed to Haiwei Wang; whw2046@126.com

Received 8 April 2020; Revised 14 May 2020; Accepted 25 May 2020; Published 15 June 2020

Academic Editor: Dongfang Ma

Copyright © 2020 Feng You et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Research studies on machine vision-based driver fatigue detection algorithm have improved traffic safety significantly. Generally, many algorithms assess the driving state according to limited video frames, thus resulting in some inaccuracy. We propose a real-time detection algorithm involved in information entropy. Particularly, this algorithm relies on the analysis of sufficient consecutive video frames. First, we introduce an improved YOLOv3-tiny convolutional neural network to capture the facial regions under complex driving conditions, eliminating the inaccuracy and affections caused by artificial feature extraction. Second, we construct a geometric area called Face Feature Triangle (FFT) based on the application of the Dlib toolkit as well as the landmarks and the coordinates of the facial regions; then we create a Face Feature Vector (FFV), which contains all the information of the area and centroid of each FFT. We use FFV as an indicator to determine whether the driver is in fatigue state. Finally, we design a sliding window to get the facial information entropy. Comparative experiments show that our algorithm performs better than the current ones on both accuracy and real-time performance. In simulated driving applications, the proposed algorithm detects the fatigue state at a speed of over 20 fps with an accuracy of 94.32%.

1. Introduction

Every year, road traffic accidents cause severe damage to human health. According to the statistics from the WHO, fatigue driving is one of the main reasons behind road traffic accidents [1]. National Sleep Foundation points out that about 32% of drivers have at least one fatigue driving experience per month [2]. Fatigue driving is a harmful threat to the driver and other traffic participants. Countries all over the world have made laws to tackle this problem. For example, the Chinese Road Traffic Safety Law stipulates that “Drivers are not allowed to drive continuously for more than 4 hours, and the rest period between every two long-duration driving should be no less than 20 minutes” [3]. In Europe, the law requires that “Drivers should stop and rest for every 4.5 hours of continuous driving, and the rest period should be no less than 20 minutes” [3]. In the United States, the law provision is that “The cumulative maximum daily driving time must not exceed 11 hours, and the continuous

daily rest time must not be less than 10 hours” [4]. As mentioned above, fatigue driving is solely associated with driving duration. It is subjective to determine whether the driver is in fatigue state or not without sufficient quantified indexes and reliable data analysis.

According to relevant data, heavy road traffic accidents caused by fatigue driving account for about 50% of all road traffic accidents [5]. Therefore, research on fatigue driving detection is inevitable. The detection algorithms are of the following types.

1.1. Detection Methods Based on Physiology and Behavior. Detection methods based on physiology and behavior are those that judge the driver’s status by installing an intrusive sensor and collecting data that characterizes the driver’s physiology, psychology, and driving operations. These detection methods include EEG signal detection [6], ECG signal detection [7], pulse beat detection [8], and EMG signal detection [9].

1.2. Detection Methods Based on Machine Vision. With distinctive characteristics of the vehicle motion and the behaviors of the driver obtained, this method assesses the driver's fatigue status. Machine vision-based detection has become the widely used method in fatigue driving detection due to its noninvasion and higher accuracy. This method applies core technologies including face detection, eye positioning, and fatigue assessment. Yan et al. [10] used the mask to locate the eye position by obtaining the driver's facial image and used PERCLOS to evaluate the driver's fatigue state. This method has better performance on individuals with conspicuous features, but the fabrication of the mask has a significant influence on the generalization performance of the model. Niu and Wang [11] divided the face image in the sequence image into nonoverlapping blocks of the same size. Then, they managed to use Gabor wavelet transform to extract multiscale features. In order to select the most recognizable ones, they applied AdaBoost algorithm. This method can effectively recognize different genders and postures under various illumination conditions. Using "bright eye effect," Bergasa and Nuevo [12] located eye position with active near-infrared light source equipment. They used finite-state machine to confirm whether the eye is closed. They also applied fuzzy system to evaluate the fatigue state. However, Bergasa's algorithm depends highly on hardware level; on the other hand, the effectiveness of the "bright eye effect" strictly relies on surrounding light conditions. You et al. [13] applied the CAMShift tracking algorithm to make the targeted areas detectable even they were under occlusion. Then the eye feature points were obtained according to the specific proportion relationship of the facial organs. Finally, they used PERCLOS to determine driver fatigue state.

1.3. Detection Methods Based on Information Fusion. Any fatigue detection method has its advantages and disadvantages. So comprehensive monitoring of driver fatigue status by various methods is promising. "AWAKE" [14] launched by the European Union is a driving behavior comprehensive monitoring system. It used many sensors such as images and pressures to synthesize the driver's eye movement, the direction of eyesight, steering wheel grip, and other driving conditions. Then, it made comprehensive detection and evaluation. Seeing Machines [15] conducts multifeature information fusion by detecting facial features such as driver's head posture, eyelid movement, gaze direction, and pupil diameter. It completed real-time monitoring of driver fatigue status.

Although the technology of fatigue detection has made great progress, it can be better:

- (i) Physiology-based driver fatigue detections require a variety of additional monitoring devices or equipment. It would not only reduce comfort during driving but also make the collected data costly and vulnerable, which has set back the popularization of these methods.
- (ii) If the light condition changes or the driver's face is partially occluded, for example, wearing glasses or

sunglasses, AdaBoost fails to accurately locate the face position and give the alarm to the driver promptly.

- (iii) At present, the commonly used algorithms are based on PERCLOS, which judge fatigue by opening and closing state of the driver's eyes. However, when the driver's eyes are too small, the algorithms are easy to misjudge. Moreover, other fatigue indicators are less commonly used due to lower reliability and less robustness.

As above literature studies discussed, results of the driving fatigue detection have defects of high intrusion, low robustness, and low reliability. Therefore, we propose a fatigue driving detection algorithm based on facial motion information entropy. The innovations are as follows:

- (i) We design a driver's face detection architecture based on the improved YOLOv3-tiny convolutional neural network and train the network with the open-source data set WIDER FACE [16]. Compared with other deep learning algorithms, such as YOLOv3 [17] and MTCNN [18], the algorithm based on the improved YOLOv3-tiny network is more accurate and simplified. It has fewer calculations and thus is easy to transplant to other mobiles.
- (ii) We used the Dlib toolkit to extract facial feature points recognized by improved YOLOv3-tiny convolutional neural network. Then we created the FFT after analyzing the characteristics of the eye and mouth position. Next, we constructed FFV which contains the overall information of the area and centroid of each FFT. We calculate the FFV of each frame and write it to the database. Thereby a state analysis data set is established. In many research studies the basis for assessing the state of the driver is the recognition result of a single frame or a few frames, which reduce the accuracy of fatigue driving detection. Based on the analysis results of a large number of consecutive frames, we design sliding windows of driving fatigue analysis to obtain the statistical characteristics of the facial motion state. Therefore, the process of driver fatigue can be observed.
- (iii) To get rid of the interference that originated from the size differences between every FFT, we introduce the face projection datum plane and apply the projection principle to extract the motion feature points of the face. Then, based on the motion feature points, we propose the facial motion information entropy, which quantitatively characterizes the chaotic degree of the motion feature points of the face. Accordingly, the driver's fatigue state can be judged. At present, the commonly used algorithms are based on PERCLOS [19], which judge fatigue by opening and closing state of the driver's eyes. However, when the driver's eyes are too small, the algorithms are easy to misjudge. Therefore, we

reveal the difference in the motion characteristics between fatigue driving and nonfatigue driving by proposing facial motion information entropy.

This paper is divided into the following seven parts. The first chapter is the introduction. In this part, we introduce the background and research significance of our fatigue driving detection system and the research status from home and abroad. We propose a fatigue driving detection algorithm based on facial motion information entropy with technical innovations. In the second chapter, we explain the algorithm in detail. The structure of this algorithm is a combination of improved YOLOv3-tiny network and Dlib toolkit. The former captures ROI, while the latter obtains facial landmarks and creates a fatigue state data set. We make a description of the definition and calculation method of facial motion information entropy, which is the main index to represent the fatigue state. The third chapter is the experimental analysis. Firstly, the experimental environment and data set are introduced. Then we use qualitative description and quantitative evaluation to measure face detection and feature point location. Finally, we evaluate our fatigue driving detection algorithm in two directions: accuracy and real time. The fourth chapter is the conclusion, which mainly summarizes the main work content of this paper and analyzes the shortcomings of the system and the aspects that need to be improved. Then, we propose the future optimization direction and prospect of the algorithm. Other sections are Data Availability, Conflict of Interests, Acknowledgments, and References.

2. Methodology

The overall pipeline of our approach is shown in Figure 1. The algorithm consists of the following 4 modules.

Face Positioning. The original data source is the real-time camera video. Based on deep learning theory, we apply the improved YOLOv3-tiny network to extract suspected face regions from complex backgrounds.

Feature Vector Extraction. FFT is a geometry area in every frame that contains facial features. Based on the coordinates of the suspected face region, we obtain facial landmarks with the application of the Dlib toolkit and construct FFV by calculating the area and centroid of the driver's FFT.

Data Set Building. According to the FFV extracted in a certain period, the driver state analysis data set is established in chronological order.

Fatigue Judgment. We design a sliding window as a sampler; every time it analyzes several sequential FFVs which match with the related sequential frames by projecting the FFV on the facial projection datum. Afterwards, it loops through all FFVs and outputs a facial motion information entropy corresponding to the current facial feature point set. We then compare the facial motion information entropy with its threshold to evaluate the fatigue state of the driver.

2.1. Face Detection Based on the Improved YOLOv3-Tiny Network. Face detection location is the foundation of driver fatigue detection, and the accuracy of the results has a great impact on the algorithm's performance. So, accurate and rapid face detection is the fundamental task of the driving fatigue detection algorithm. In the traditional face detection algorithm, the face features are mostly based on prespecified features such as Haar and HOG [20, 21]. In terms of Haar features, Viola and Jones [22] propose a joint Haar feature for face detection algorithms. However, image features may lose because of inappropriate face postures, dim light conditions, noise interference, or a partially occluded face, which decreases the robustness and reliability of prespecified feature method. Recently, deep learning theory provides new ways for detection and segmentation [23]. It can be divided into 2 categories: one transfers the target detection model to face detection and segmentation process; the other is the cascade methods, such as MTCNN [24, 25] and Cascade CNN [26]. Compared with the traditional methods [27], the face detection based on convolutional neural network extracts features autonomously instead of man-made operation. With the support of data sets, face detection performance has been greatly improved.

The YOLO [28] (You Only Look Once) model is a fast target detection model based on deep learning [29]. It is a separate end-to-end network that turns target detection into a regression problem. Specifically, we can replace the sliding window in the traditional target detection to the regression method and convolutional neural network (CNN) [30]. This method of feature extraction is less affected by the external environment and has the advantage of extracting target features quickly.

Inspired by the idea of YOLO model, we transform the multiobjective regression into the single target regression, hence reducing the calculation amount. Then, we improve YOLOv3-tiny network to locate suspected face regions.

The YOLOv3-tiny network is a simplified version of YOLOv3, so it has better real time than YOLOv3. It simplifies the YOLOv3 feature detection network darknet-53 to 7 conventional convolution layers and 6 Max Pooling layers and 1 Up Sample layer. The improved network structure is shown in Figure 2. In the figure, "Darknetconv2d BN Leaky" (DBL) is the basic component of the network, "Conv" is the convolution layer, and "Leaky ReLU" is the activation function. Batch normalization (Batch Norm) is a regularization method that guarantees the algorithm convergence and avoids overfitting. Concat sandwiches a sample layer in the middle of two DBL. Nonmaximum suppression (NMS) is to eliminate the extra facial box and locate the best driver's face suspected area.

We consider that the images used for analysis for fatigue driving contain only one face. If the network shows high accuracy in multiface detection, one face detection will be more accurate. So, in the YOLOv3-tiny network training phase, we use the WIDER FACE (Face Detection Data Set and Benchmark) (<http://wider-challenge.org/2019.html>) [16] data set as the driving data. The WIDER FACE data set includes 32,203 images and 393,703 marked faces, which is one of the most common face databases. The data set

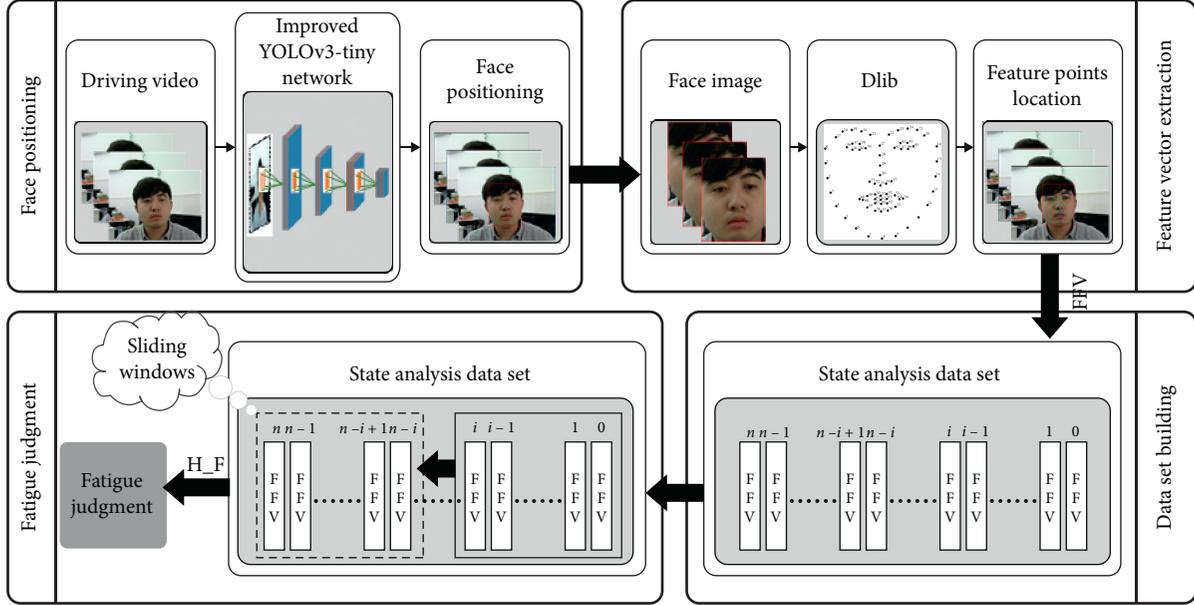


FIGURE 1: Algorithm structure diagram, where H_F is facial motion information entropy.

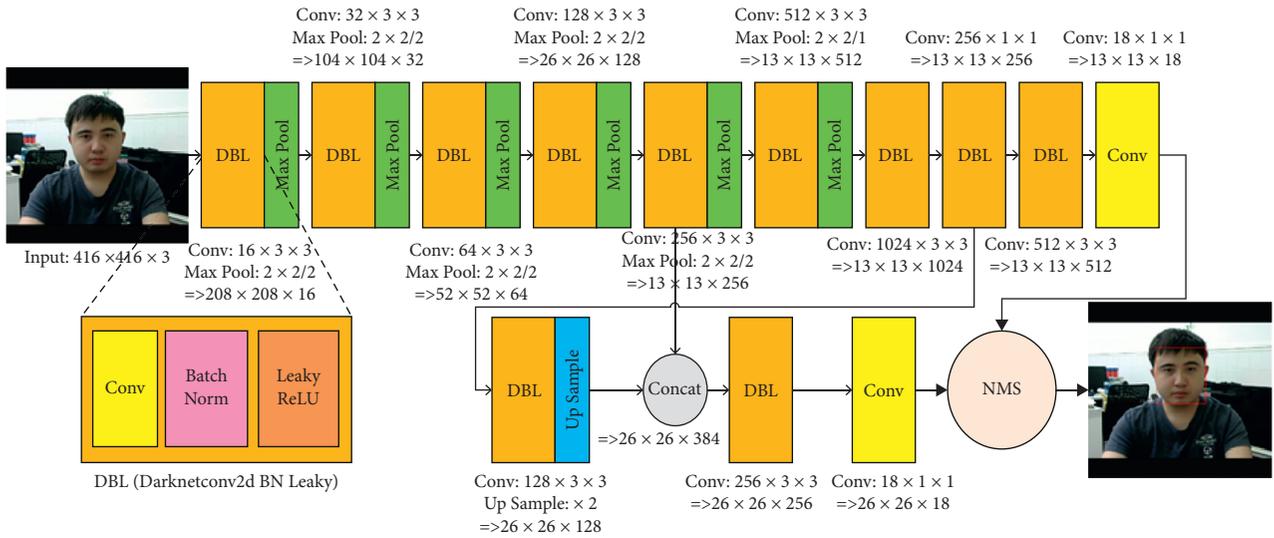


FIGURE 2: Improved YOLOv3-tiny network structure diagram.

includes different scales, poses, occlusions, expressions, makeup, and lighting, as shown in Figure 3.

The WIDER FACE data set has the following features:

- (i) The data set is divided into three types: training set, test set, and verification set, which, respectively, account for 40%, 50%, and 10% of the data set
- (ii) There are a large number of faces in each image, which contains an average of 12.2 faces
- (iii) The data set pictures are high-resolution color images

Firstly, based on the YOLOv3-tiny network, the picture of the WIDER FACE data set is adjusted to 10 different sizes, and every picture is divided into 13×13 grid cells or 26×26 grid cells. Then, we find the location of the driver's face on

the nonoverlapping grid cell and classify it. For each grid cell, the network outputs B bounding boxes as well as the corresponding confidence and the conditional probability of the driver's face. Finally, nonmaximal values are used to suppress redundant bounding boxes. The confidence formula is given as

$$\text{score} = P_r(\text{Object}) * \text{IOU}_{\text{pred}}^{\text{truth}}, \quad (1)$$

where $P_r(\text{Object})$ is the probability of the driver's face. If the face is included, $P_r(\text{Object}) = 1$; otherwise $P_r(\text{Object}) = 0$. $\text{IOU}_{\text{pred}}^{\text{truth}}$ is the intersection over union (IOU) of the bounding box to the real box.

There are four basic elements in the YOLOv3-tiny network loss function: the central error term of the bounding box, the width and high error term of the



FIGURE 3: WIDER FACE data set diagram.

bounding box, the error term of the prediction confidence, and the error term of the prediction category. We managed to use the offline trained YOLOv3-tiny network to extract the accurate face region for further research.

2.2. Driver's Facial Motion Feature Extraction

2.2.1. Face Feature Location Based on the Dlib Toolkit.

On the driver's face area located by the improved YOLOv3-tiny network, we used the face key point detection model based on the Dlib-ml [31] library to extract the fine-grained features of a driver's face (as is shown in Figure 4(a)). The Dlib library contains 68 face key points. The testing principle is applying cascading shape regression to check all the key points of the face component.

The face detection process is as follows. Firstly, the feature of the input image is extracted, including the features of the face contour, eyebrows, eyes, nose, and mouth contours. Secondly, the extracted features are mapped to the face feature points through a trained regressor; at this point, an initial shape of the key points of the human face component is generated from its original image. Thirdly, gradient boosting [32] is used to iteratively adjust the initial shape until it matches with the real shape; then the cascaded regressor of each stage is calculated with the least-square method.

The face key point detection method of the Dlib library is based on the ensemble of regression trees (ERT) algorithm [29]. It uses the regression tree set to estimate the face feature points, and the speed of calculation is fast. The detection of 68 key points in each face takes about 1 ms. Similar to [33] and [34], this cascade regressor method is available even though feature points are partially missing in the training sample set. The iterative algorithm process uses the following formula:

$$\widehat{S}^{(t+1)} = \widehat{S}^{(t)} + R_t\left(h\left(I, \widehat{S}^{(t)}\right)\right), \quad t = 1, \dots, T, \quad (2)$$

where T is the number of rounds of the regression and $\widehat{S}^{(t)}$ is the current shape estimation; each regression $R_t(\cdot)$ predicts an increment based on the input images I and $\widehat{S}^{(t)}$, that is, $R_t(h(I, \widehat{S}^{(t)}))$. The initial shape used is the average shape of the training data, and the update strategy is the Gradient Boosting Decision Tree (GBDT) algorithm [32]. Every time, for each separate subregion, we train a weak classifier whose predictive value approximates the true value of that subregion. Ultimately, the predicted value of the whole region is the weighted sum of every predictive value.

When the driver's face is detected, the feature points of the face are obtained in real time by the above algorithm, as shown in Figure 4(b).

2.2.2. Motion State Parameter Extraction.

As discussed above, drivers get exhausted naturally during driving due to physiological and psychological state changes. At that time, they are in fatigue state. Fatigue driving endangers the driver and other traffic participants, as it declines the driving cognitive and driving skills, therefore resulting in misperception, misjudge, and misoperation. To ensure driving security and traffic safety, the driver must have a clear understanding of the driving condition and surrounding road environments all the time [35]. It requires the driver to continually adjust the head orientation and the fixation point of the eye. Compared to nonfatigue driving, the driver's visual field adjustment behaviors change significantly whether in the early, middle, or late stages of fatigue [36]. The facial motion state, such as movement amplitude and frequency, is abnormal.

Hence, we propose a Face Feature Triangle to characterize the driver's facial motion state. Based on face feature location, we defined a Face Feature Triangle (FFT). As shown in Figure 5, the midpoint of left eye is A, the midpoint of right eye is B, and the midpoint of mouth is C. The three points consist of the FFT. According to the FFT, we define the Face Feature Vector (FFV), as

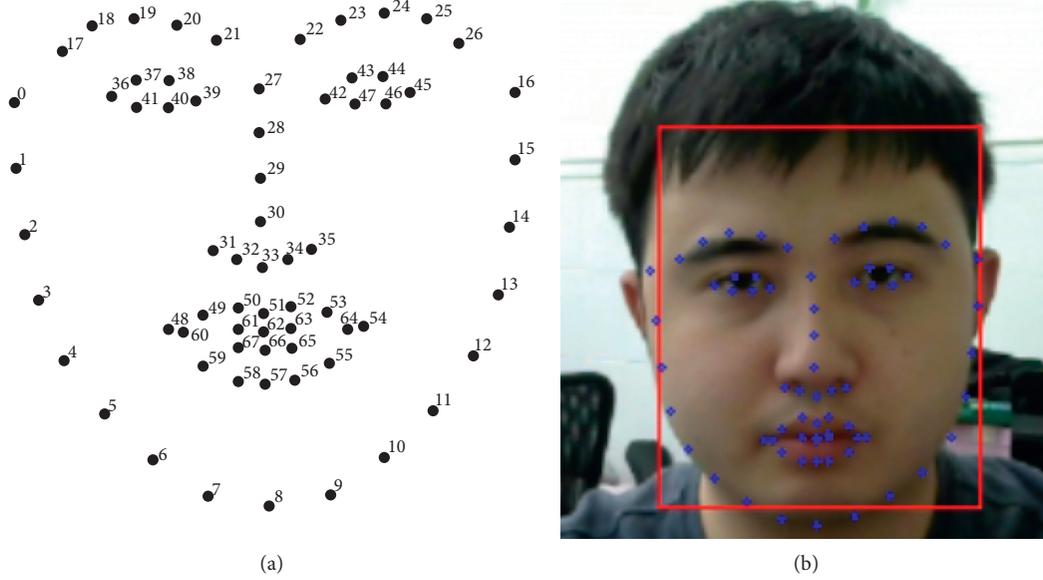


FIGURE 4: Driver's face feature point acquisition based on Dlib. (a) Dlib face feature point positioning. (b) Face feature point positioning effect.

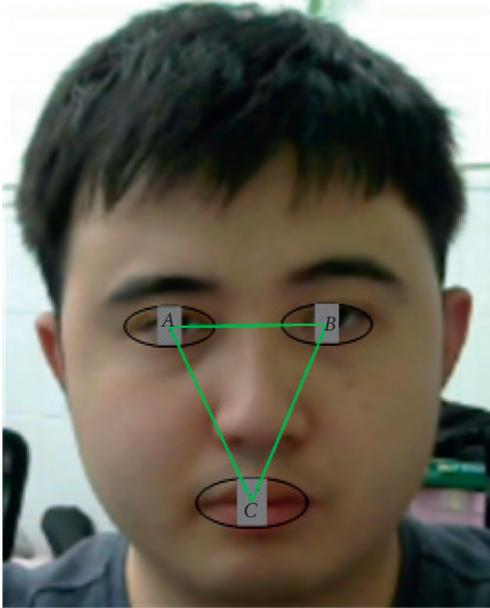


FIGURE 5: Face Feature Triangle (FFT).

$$FFV = (F_x, F_y, \sqrt{S}), \quad (3)$$

where (F_x, F_y) is the midpoint of the FFT and S is the area of the FFT. According to the plane triangle's center of gravity and area formula, F_x, F_y, S are as shown in the following equation:

$$F_x = \frac{A_x + B_x + C_x}{3},$$

$$F_y = \frac{A_y + B_y + C_y}{3},$$

$$S = \frac{|A_x * B_y - B_x * A_y + B_x * C_y - C_x * B_y + C_x * A_y - A_x * C_y|}{2}. \quad (4)$$

Among them, according to Figure 4(a), Dlib face feature point positioning, and midpoint two-dimensional coordinate formula, the coordinates (A_x, A_y) , (B_x, B_y) , and (C_x, C_y) are defined as

$$(A_x, A_y) = \left(\frac{p36_x + p39_x}{2}, \frac{p36_y + p39_y}{2} \right),$$

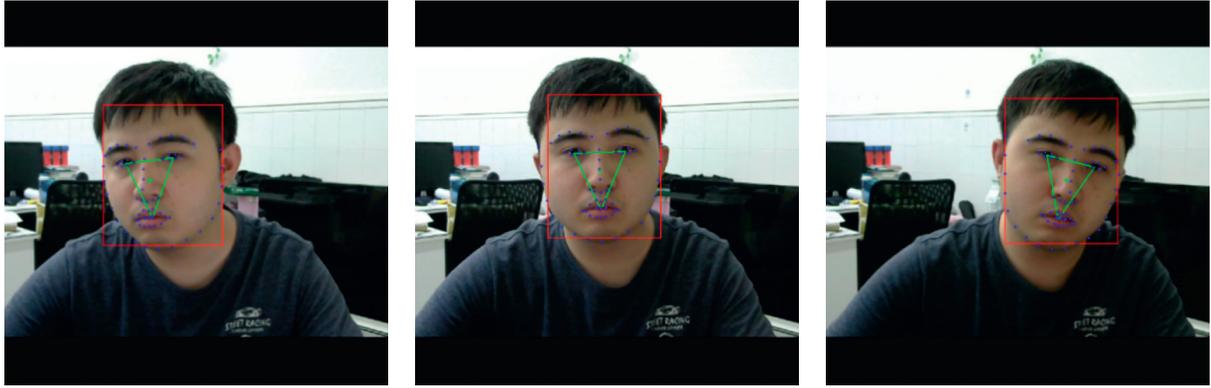
$$(B_x, B_y) = \left(\frac{p42_x + p45_x}{2}, \frac{p42_y + p45_y}{2} \right), \quad (5)$$

$$(C_x, C_y) = \left(\frac{p60_x + p64_x}{2}, \frac{p60_y + p64_y}{2} \right),$$

where $p36$ is the coordinate of point 36 in Figure 4(a).

As is shown in Figure 6, FFT varies significantly with the driver's face position; therefore the FFV is suitable for characterizing the state of facial motion in the fatigue detection algorithm.

2.3. Driver's Facial Feature Points Collection. Generally, head posture-based fatigue detection algorithms [37] depend on the characteristics of instantaneous head motions such as nodding to determine whether the driver is in fatigue state. It is challenging to judge fatigue based on a single frame or a small number of frames and there may even be misjudgment. Therefore, it is necessary to study the statistical characteristics of the driver's facial movement state during fatigue. As described in Section 2.2, to extract the statistical characteristics of facial motion and find the relationship between statistical characteristics and driving fatigue state, we define FFT. Since the area of the FFT varies with the distance between driver's head and the camera, in order to get regularized data, we apply a face projection datum plane method. As shown in Figure 7, it projects all FFTs to a preset



(a)

(b)

(c)

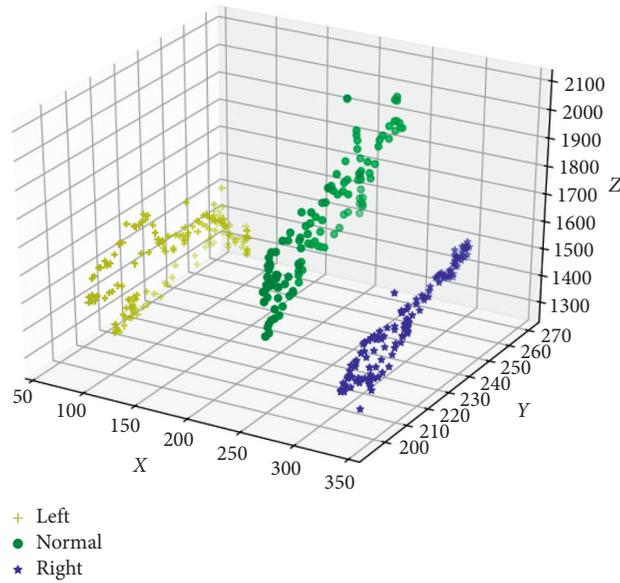


FIGURE 6: Different facial movement states and FFV differences, where X is F_x , Y is F_y , and Z is \sqrt{S} . “Left” stands for the left swing of the face, “Normal” stands for normal face posture, and “Right” stands for the right swing of the face.

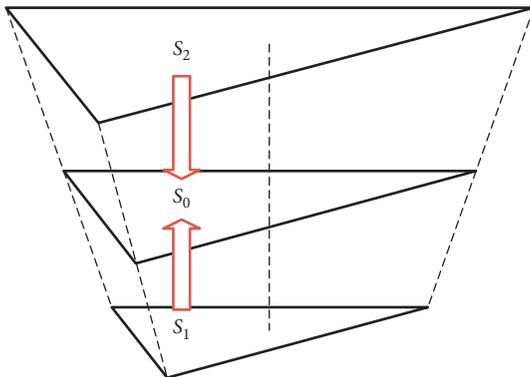


FIGURE 7: Projection schematic.

$$\begin{aligned}
 x &= \left(F_x - \frac{\text{col}}{2} \right) * \sqrt{\frac{S}{S_0}} + \frac{\text{col}}{2}, \\
 y &= \left(F_y - \frac{\text{row}}{2} \right) * \sqrt{\frac{S}{S_0}} + \frac{\text{row}}{2},
 \end{aligned}
 \tag{6}$$

where “row” and “col” are the numbers of rows and columns of the input images. A point (x, y) projected onto the datum projection plane is defined as a feature point of the driver’s facial motion. We establish the feature point set of the driver’s facial motion by counting the feature points in frames and then construct the statistical model of the driver’s facial motion state. The experimental results are shown in Figure 8.

projection datum plane and eliminates the interference that originated from the distance difference. The area of the projection datum plane is S_0 , and projection formula is shown in the following equation:

2.4. Driver Fatigue State Assessment Model Based on Facial Motion Information Entropy

2.4.1. Facial Motion Information Entropy. As mentioned above, in nonfatigue state, a driver is active to quickly switch

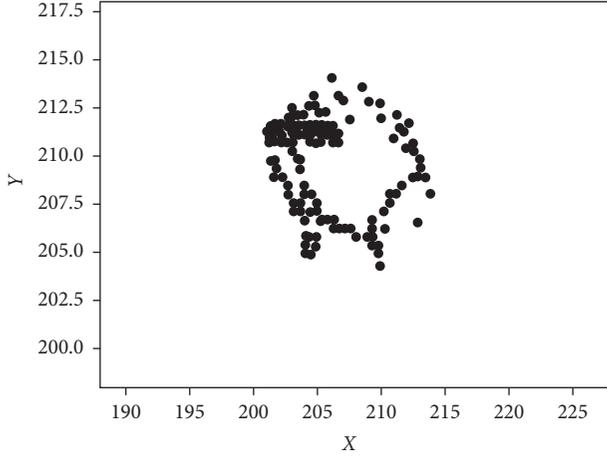


FIGURE 8: Facial motion feature point set.

the fixation point and head orientation, whereas in the opposite situation the drivers change their head position much more slowly.

To compare the difference between frequency and amplitude of the gaze point and the head orientation in the two driving states, based on the facial motion feature points, we count the set of facial motion feature points under a large number of consecutive frames. Figures 9(a) and 9(b) show the set of facial motion feature points under fatigue and nonfatigue conditions, respectively.

Accordingly, compared with the fatigued driving state, the nonfatigue facial motion feature points are more divergent and chaotic. “A Mathematical Theory of Communication” [38] pointed out that any information is redundant, and the redundancy is related to the probability or uncertainty of each symbol (number, letter, or word) in the message. That is information entropy, a concept from thermodynamics. It refers to the average amount of information after removing the redundant parts. The following equation shows the mathematical expression of information entropy:

$$H(X) = - \sum_{x \in \chi} p(X) \log p(X). \quad (7)$$

Based on the location of facial feature points in Section 2.2.1, we extract the FFV and establish the state analysis data set. Then, the facial motion information entropy is defined according to the concept of information entropy. Thus, the indicator to assess the degree of chaos of the facial feature point set is established. The calculation method is as follows:

- (1) Calculate the center point (\bar{F}_x, \bar{F}_y) of the facial motion feature point set, and N is the number of feature points, as is shown in

$$\begin{aligned} \bar{F}_x &= \frac{\sum F_x}{N}, \\ \bar{F}_y &= \frac{\sum F_y}{N}. \end{aligned} \quad (8)$$

- (2) Calculate the Euclidean distance denoted as l_i from each feature point to the center point, where $i = 1, 2, \dots, N$, as shown in

$$l_i = \sqrt{(F_x - \bar{F}_x)^2 + (F_y - \bar{F}_y)^2}. \quad (9)$$

- (3) Calculate the mean value and standard deviation of distance, as is shown in the following equation

$$\begin{aligned} \mu_l &= \frac{\sum_{i=1}^N l_i}{N}, \\ \sigma_l &= \sqrt{\frac{\sum_{i=1}^N (l_i - \mu_l)^2}{N}}. \end{aligned} \quad (10)$$

- (4) The interval I_i is defined as equation (11), where $i = 1, 2, \dots, i_{\max}$. i_{\max} is defined as equation (12):

$$I_i = \left[(i-1) * \frac{\mu_l}{\sigma_l}, i * \frac{\mu_l}{\sigma_l} \right], \quad (11)$$

$$i_{\max} = \frac{\max(l_1, l_2, \dots, l_N)}{\mu_l / \sigma_l} + 1. \quad (12)$$

- (5) According to the distance from each feature point to the center point, the number of distances falling in the interval I_i is counted as n_i .
- (6) Calculate facial motion information entropy $H_F(X)$, as is shown in

$$H_F(X) = - \sum_{i=1}^{i_{\max}} p(x_i) \log p(x_i), p(x_i) = \frac{n_i}{N}. \quad (13)$$

2.4.2. Design of Driver’s Facial Motion Information Entropy Classifier Based on SVM. As mentioned above, when drivers focus well on driving, they usually switch the fixation point and head orientation in order to get a better view of the driving environments, and the facial motion information entropy is higher. On the contrary, information entropy is much lower under fatigue driving situations. We use the training set in the open-source dataset YawDD (<http://www.site.uottawa.ca/~shervin/yawning/>) [39]. It contains fatigue driving data sets of all ages and people of all races, including different genders and facial features. It provides videos that record several common driving conditions such as driving with glasses, speaking, and singing while driving, even pretending to be simulating fatigue.

SVM [40] is a machine learning model that adopts the structural risk minimization criterion under the framework of statistical learning theory. It is a linear classifier model with the largest interval defined in the feature space. Given a training data set $S = \{(x_i, y_i), i = 1, 2, \dots, N\}$ on a feature space, $x_i \in R^d$ is the i th input sample and $y_i \in \{+1, -1\}$ is the label corresponding to x_i . When $y_i = +1$, x_i is called a positive sample, and when $y_i = -1$, x_i is a negative sample.

Generally, a linear discriminant function $f(x) = w^T x_i + b$ in a d -dimensional space can distinguish two types of data, and a classification hyperplane can be described as

$$w^{*T} \cdot x + b^* = 0. \quad (14)$$

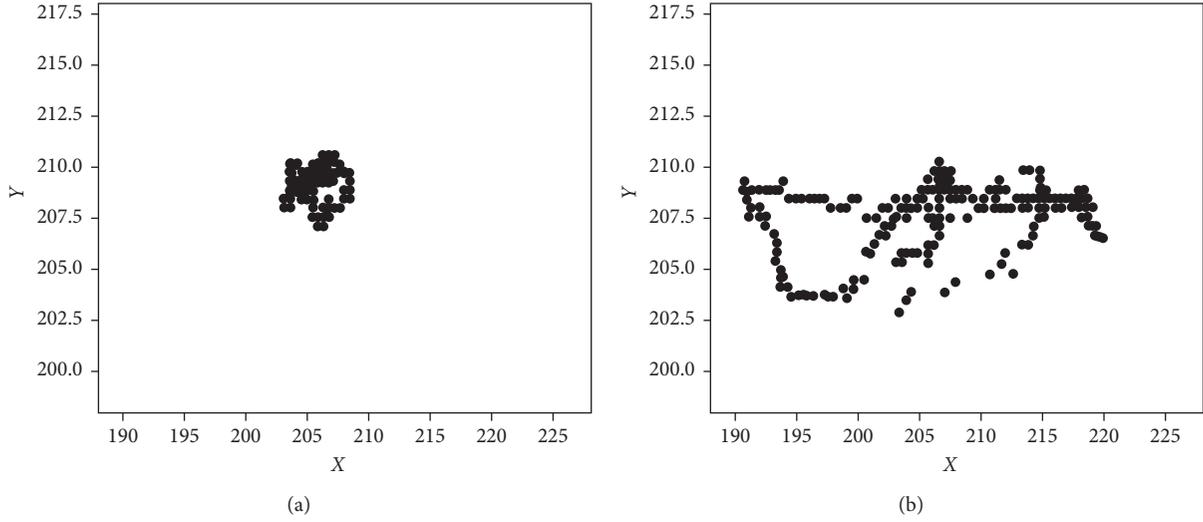


FIGURE 9: Different drive state facial motion feature point set. Facial motion feature point set in (a) fatigue and. (b) nonfatigue.

The normal vector w^T and the intercept b determine the superclass surface function. According to the basic idea of SVM, the constrained optimization problem of linear separable support vector machine can be obtained:

$$\begin{cases} \min_{w,b} & J(w) = \frac{1}{2} \|w\|_2^2 \\ \text{s.t.} & y_i(w^T \cdot x_i + b) \geq 1, i = 1, 2, \dots, N. \end{cases} \quad (15)$$

In the training phase of the driver's face mark box, the improved YOLOv3-tiny is used as the training network, and the training set is applied to detect the driver's face. As described in Section 2.4.1, the driver's facial motion information entropy is calculated based on the positioning information of the Dlib face feature points. Among them, when $y_i = +1$, x_i is a positive sample, indicating that the driver is in nonfatigue driving state, and when $y_i = -1$, x_i is a negative sample, indicating that the driver is in fatigue driving state. Combined with the constraints of equation (15), the hyperplane parameters w^T and b can be calculated to obtain the driver's facial motion information entropy classifier.

Experiments show that the projection datum area S_0 has different values, which will affect the parameters w^T and b of the driver's facial motion information entropy classifier. In the experiment, S_0 is set to 10000.

2.4.3. Fatigue Judgment Based on Facial Motion Information Entropy. As mentioned above, the original image of the driver was acquired with an in-vehicle camera, and the improved YOLOv3-tiny network was used to detect the driver's face. The face area will be extracted as an input subimage, and then the Dlib toolkit is used to obtain the facial feature points of the subimage if the face is detected in a frame image. If not, the system will determine that the driver's head posture is abnormal. If it is determined that the driver's head posture is abnormal for more than 10

consecutive frames, the system will issue an alarm. Based on the face landmarks, the FFV is calculated according to the coordinates of the eye feature points and the mouth feature points. Within a certain number of frames (the number of frames set in this paper is more than 1000 frames), we count the FFV per frame. Considering that fatigue often generates during driving, if directly calculating the facial motion information entropy of all FFVs, the result may be inaccurate. In order to improve accuracy, as is shown in Figure 10, the paper sets a sliding window to calculate the facial motion information entropy in segments on all FFVs. The window size is set to 1000, and the sliding step size is set to 100. Each time the sliding window slides, the 1000 FFVs in the current sliding window are obtained first. Then, we can obtain the set of facial motion feature points in the current window. Finally, the facial motion information entropy $H_F(X)$ in the current window is calculated. Set $\text{Th}_{H_F(X)}$ as the judgment threshold by training the SVM classifier on the YawDD training set. If $H_F(X) < \text{Th}_{H_F(X)}$, the judgment is that the driver is in fatigue state. Otherwise, the sliding window moves to the next position to continue analyzing.

The flow chart of fatigue judgment based on facial motion information entropy is shown in Figure 11.

3. Results and Discussion

In order to verify the validity of the algorithm, we evaluated the performance of the improved YOLOv3-tiny network with the public data sets WIDER FACE and YawDD. On this basis, the design comparison experiment is carried out to verify whether the fatigue driving detection algorithm based on facial motion information entropy is correct.

3.1. Experimental Environment and Data Set. The experimental platform is the Intel Core i5-8400 with x86 architecture, and the CPU clock speed is 2.80 GHz. Graphics card is GTX1060 with Pascal architecture (CUDA: 9.2;

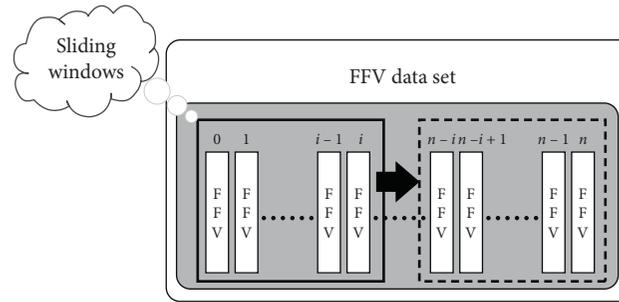


FIGURE 10: Sliding windows.

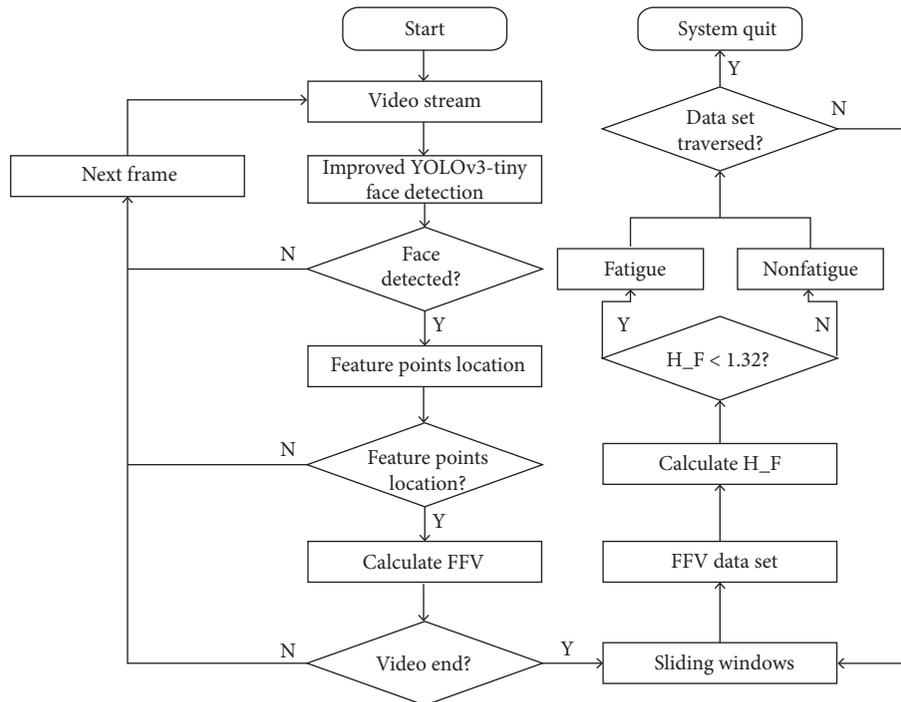


FIGURE 11: Driver fatigue state assessment model flow chart.

CUDNN: 7.2). The RAM is 8G DDR4, and the opencv3.4.6 image library is used. The deep learning computing framework is PaddlePaddle1.5. The environment of the program is python 3.6. Hardware configuration is shown in Table 1.

The data set used in the experiment included the public data sets WIDER FACE and YawDD, where the public data set WIDER FACE includes 32203 pictures and 393703 marked faces, which is used to train Yolov3-tiny's face network. However, the WIDER FACE data set only contains marker face images and does not provide any information about the driver's fatigue status. Therefore, the WIDER FACE data set cannot be used to analyze driver fatigue status. YawDD is a data set of fatigue driving detection including male and female volunteers in the naked eye, wearing glasses, normal state, speaking/singing, and simulated fatigue. So we choose YawDD data set as test set of fatigue driving detection. The detection result of the YawDD data set is shown in Figure 12.

TABLE 1: Hardware configuration table.

Type	Specific parameters
Processor	Intel(R)Core(TM)i5-8400 CPU@2.80 GHz 2.81 GHz
GPU	NVIDIA GeForce GTX1060 6 GB
Computer version	Windows 10
RAM	8.00 GB
Python version	3.6
Opencv version	3.4.6
Paddle version	1.5
CUDA version	9.2
CUDNN version	7.2

3.2. Face Detection and Feature Point Location

3.2.1. Qualitative Description. In order to verify the effectiveness of face detection based on the improved YOLOv3-tiny network and the accuracy based on the Dlib facial feature point location, the experiments were performed in the laboratory and in the vehicles.



FIGURE 12: The detect result of YawDD data set.

In the laboratory, the light is uniform and does not drastically change. The face recognition algorithm based on improved YOLOv3-tiny network can accurately detect faces from test videos. The face area can be correctly marked, as is shown in Figures 13(a) and 13(b) (1-1) and (1-2). Besides, the algorithm can detect the driver's face area and mark feature points, even in the cases of wearing glasses (as shown in Figure 13 (2-1)), head tilting (as shown in Figure 13 (1-3)), and expression changing (as shown in Figure 13 (2-2)).

In the vehicle experiment, the change of illumination may cause high interference to the driver's face detection and feature point location. So, it is crucial to verify the effectiveness of the algorithm in the real vehicle scenario. In the real driving scene, the algorithm can complete face detection and feature point location in case of uneven illumination, as is shown in Figure 13 (4-1). It can be seen that the algorithm has excellent recognition performance and robust performance in both the laboratory and real vehicle, and this will provide the basis for the driver's fatigue feature extraction and fatigue state assessment.

3.2.2. Quantitative Evaluation. The improved YOLOv3-tiny network provides face landmarks for fatigue driving detection. Its performance represents the effectiveness of the fatigue driving detection algorithm. Therefore, we quantitatively evaluate the performance of the improved YOLOv3-tiny network on the WIDER FACE data set.

In this paper, we adopt the ROC curve [41] theory for evaluation. Accuracy is the ratio of the number of correctly predicted samples to the total number of samples, and it is an intuitive evaluation index of model performance. However, the accuracy rate is difficult to express the pros and cons of the model in case of uneven distribution of positive and negative sample data. The sensitivity indicates the proportion of all positive samples correctly detected. Specificity indicates the proportion of all negative samples correctly detected. The ROC curve is a comprehensive indicator formed by the combination of sensitivity and specificity and reflects the sensitivity and specificity of continuous variables.

(1) *Accuracy (ACR)*. In the task of the driver's face detection, the ACR is the ratio of the number of correctly detected images to the total number of images:

$$ACR = \frac{N_{\text{detected}}}{N_{\text{total}}}, \quad (16)$$

where N_{detected} is the number of correctly detected images and N_{total} is the total number of images.

In the process of improving the YOLOv3-tiny network training and verification, the intersection ratio parameter (IOU) [42] is introduced to measure the similarity between the face detection area and the marked real area. IOU is a standard for measuring the accuracy of a corresponding object in a specific data set. In Figure 14, $face_d$ is the face area detected by the model, $face$ is the real area marked, and the calculation formula is given in the following equation (17): where $Area(face_d \cap face)$ is the area of $face_d \cap face$ and $Area(face_d \cup face)$ is the area of $face_d \cup face$.

$$IoU = \frac{Area(face_d \cap face)}{Area(face_d \cup face)}, \quad (17)$$

The intersection ratio indicates the degree of overlap between the model prediction area and the real area. As can be seen from Figure 14, the higher the value is, the higher the detection accuracy is. In the case where $IoU = 1$, the prediction box overlaps with the real box. Generally speaking, the object is correctly detected when the IOU is more than 0.5. In the face detection process, we adopt a higher threshold. In this paper, when the IOU is more than 0.75, the face is considered to be correctly detected. Figure 15 shows the accuracy curve of the driver's face detection during the training of the improved YOLOv3-tiny network. It can be seen that, with the increase of training rounds, the accuracy of face detection gradually increases. The improved YOLOv3-tiny network has an accuracy rate of 98.5%.

(2) *ROC Curve*. Sensitivity and specificity are important evaluation indicators of the pattern recognition model. If

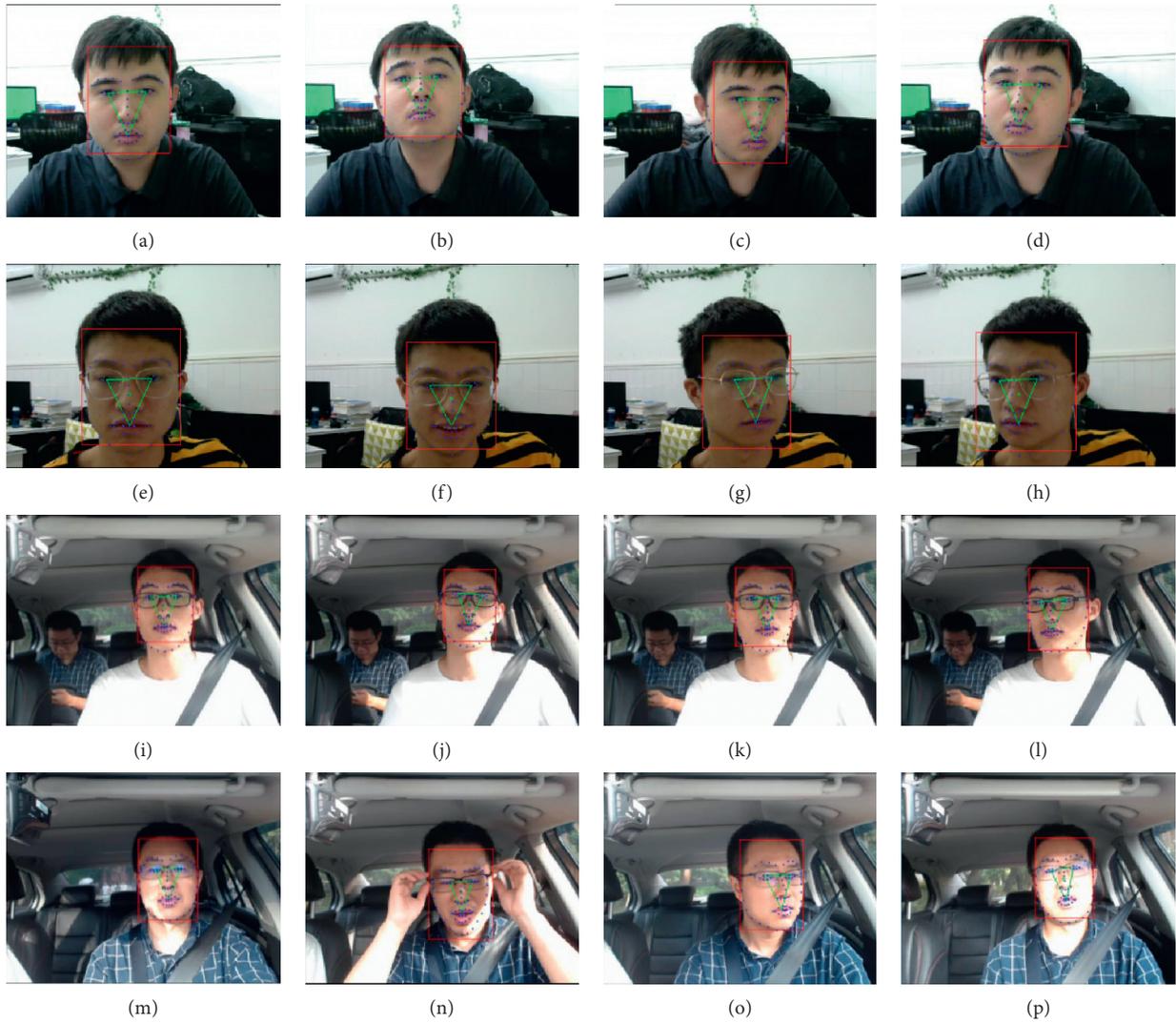


FIGURE 13: The results of face detection and feature point location. (a) (1-1). (b) (1-2). (c) (1-3). (d) (1-4). (e) (2-1). (f) (2-2). (g) (2-3). (h) (2-4). (i) (3-1). (j) (3-2). (k) (3-3). (l) (3-4). (m) (4-1). (n) (4-2). (o) (4-3). (p) (4-4).

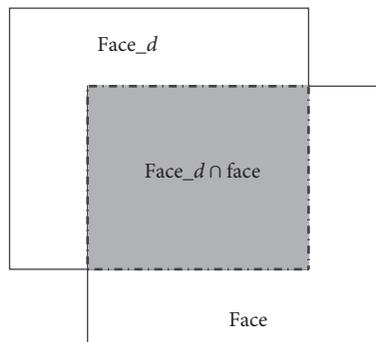


FIGURE 14: Intersection over union.

you use TP, TN, FP, and FN to indicate the number of true-positive, true-negative, false-positive, and false-negative samples, respectively, in a test, then the definitions of sensitivity S_n and specificity S_p are

$$S_n = \frac{TP}{TP + FN}, \quad S_p = \frac{TN}{TN + FP}. \quad (18)$$

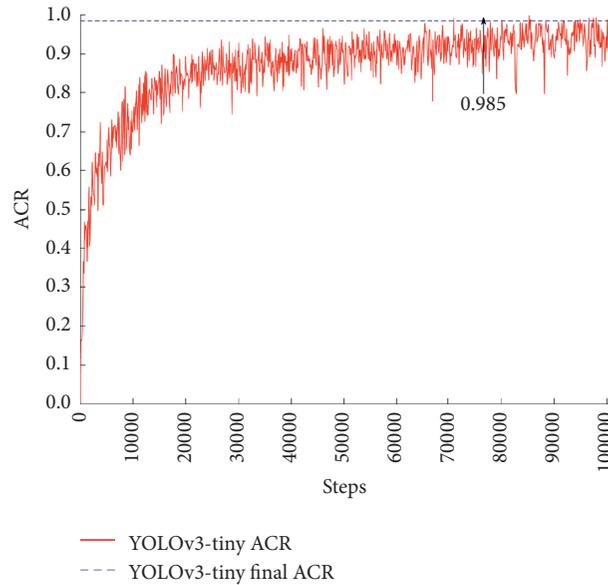


FIGURE 15: Driver face detection accuracy.

A ROC curve is a graph of the relationship between the true-positive rate (sensitivity) and the false-positive rate ($1 - \text{specificity}$). The ROC curve is one of the comprehensive indicators for characterizing the accuracy of pattern recognition tasks, and the closer the ROC curve is to the upper left corner, the better the model performance is.

Figure 16 shows the ROC curve of the driver's face detection model. As can be seen from the figure, the ROC curve corresponding to the improved YOLOv3-tiny network is close to the upper left corner of the graph, indicating high accuracy in face detection.

In summary, by evaluating the performance of the improved YOLOv3-tiny network on the WIDER FACE data set, it is shown that the improved YOLOv3-tiny network in this paper has high accuracy. Besides, the ROC curve indicates that the algorithm can effectively avoid two types of errors in the driver's face recognition, that is, to ensure that the driver's face can be correctly detected while avoiding the misjudgment on the face.

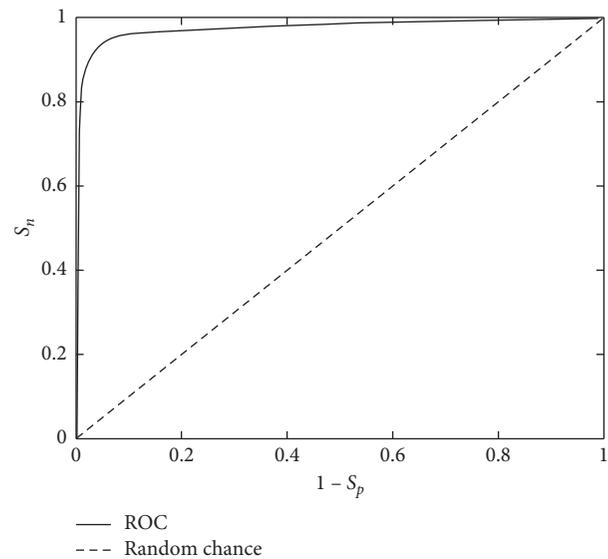


FIGURE 16: ROC curve.

3.3. Fatigue State Evaluation

3.3.1. Accuracy. We use the YawDD data set to test the performance of fatigue detection. Face detection and facial feature point location are the basis of fatigue driving detection. The FFV of each frame in the on-board video is calculated and stored based on the facial feature points. Calculate the FFVs of all video frames in a certain period, and establish a state analysis data set. The sliding window (discussed in Section 2.4.3) is applied to the state analysis data set to calculate the facial motion information entropy for each sliding. If the entropy does not exceed the threshold, we can conclude that the driver is in fatigue state. Videos are randomly selected from the data set for fatigue driving detection. The process of fatigue driving detection is shown in Figure 11.

In this paper, we randomly select ten videos from the YawDD test set, including nonfatigue driving status and fatigue driving status. The facial information entropy threshold for judging fatigue state is 1.32, and the results are shown in Table 2. It can be seen that the accuracy of the fatigue driving detection in the randomly selected ten videos is 90%, and the correct rate of the system in the entire test set of YawDD is 94.32%.

3.3.2. Speed. Based on hardware configuration as shown in Table 1, a comparison test is performed on the image source to verify the real-time performance of the system. The results are shown in Table 3.

Table 3 illustrates that YawDD Video excels at face detection time. One possible reason is the difference between

TABLE 2: Sample fatigue test table.

Sample number	Facial motion information entropy	Actual driving status	Predictive driving status
1	[1.23, 0.96, 0.56, 1.20, 1.40, 0.49, 0.65, 0.45, 0.75]	Fatigue	Fatigue
2	[1.10, 1.42, 0.86, 0.52, 0.97, 0.95, 1.50, 0.88]	Fatigue	Fatigue
3	[2.50, 2.42, 2.65, 1.93, 2.01, 2.89, 3.32, 3.21]	Nonfatigue	Nonfatigue
4	[0.57, 0.87, 0.34, 0.67, 0.95, 1.12, 1.21, 1.29, 1.01]	Fatigue	Fatigue
5	[1.98, 1.87, 1.93, 2.03, 3.23, 3.42, 3.34, 2.72]	Nonfatigue	Nonfatigue
6	[0.62, 0.57, 0.88, 1.02, 1.42, 1.45, 0.92]	Fatigue	Fatigue
7	[2.22, 1.52, 2.33, 2, 78, 3.11, 2.07, 2.98, 3.04]	Nonfatigue	Nonfatigue
8	[1.35, 1.02, 1.22, 0.78, 0.56, 0.22, 0.24, 0.31, 0.55]	Fatigue	Fatigue
9	[2.44, 2.57, 2.72, 1.98, 1.42, 1.30, 2.23, 2.89, 2.66]	Nonfatigue	Fatigue
10	[1.50, 0.89, 0.76, 0.71, 0.65, 0.88, 0.31, 0.42, 0.51]	Fatigue	Fatigue

TABLE 3: The time spent in fatigue status judgment.

Image source	Face detection time (ms)	Facial feature point positioning time (ms)	Calculate FFV time (ms)	Total time (ms)
Camera	34.52	13.91	1	49.43
YawDD Video	32.13	13.91	1	47.04

TABLE 4: Comparison of fatigue detection algorithms.

Algorithms	Accuracy (%)	Speed ($\text{ms}\cdot\text{f}^{-1}$)
AdaBoost + CNN	92.10	58.61
CNN + DF_LSTM	91.48	65.64
Algorithm in this paper	94.32	49.43

the data reading methods, and the YawDD Video method gets the data from the video stream directly.

Our algorithm shows that the system has good accuracy and high-speed performance under various conditions and can accurately judge the fatigue state of the driver. Compared with AdaBoost + CNN and CNN + DF_LSTM algorithms [43, 44], our method improves the accuracy of the fatigue driving detection algorithm. It also has better real-time performance, which meets the requirements of the fatigue driving detection system. The comparative result is shown in Table 4.

4. Conclusions and Future Directions

With the rapid increase of global car ownership, road traffic accidents have become one of the leading causes of human death in the world. Fatigue driving is one of the main causes of road traffic accidents. Fatigue driving can seriously affect driving skills and seriously threaten drivers and other traffic participants. At present, fatigue driving detection and early warning have achieved better research results, but they still need some improvements, such as high intrusiveness, poor detection performance in complex environments, and simple evaluation indicator. Therefore, we propose a new detection algorithm for fatigue driving based on facial motion information entropy. The main contributions are as follows.

- (i) We design a driver's face detection architecture based on the improved YOLOv3-tiny convolutional neural network and train the network with the open-source data set WIDER FACE. Compared

with other deep learning algorithms, such as YOLOv3 [17] and MTCNN [18], the algorithm based on the improved YOLOv3-tiny network improves the face recognition accuracy, simplifies the network structure, and reduces the amount of calculation. Then, it is more convenient to transplant to the mobile. The accuracy rate of face recognition based on the improved YOLOv3-tiny network is up to 98.5%, and single test just takes 34.52 ms.

- (ii) The Dlib toolkit is used to extract facial feature points on the face area that is located by the improved YOLOv3-tiny convolutional neural network. Then, the driver's FFT is established by analyzing the positioning characteristics of the eye and mouth. Finally, the driver's FFV is constructed by the area and centroid of FFT. We calculate the FFV of each frame and write it to the database. Thereby a state analysis data set is established. In many research studies, the basis for assessing the state of the driver is the recognition result of a single frame or few frames, which reduce the accuracy of fatigue driving detection. In this paper, based on the analysis results of a large number of consecutive frames, we design sliding windows of driving fatigue analysis to obtain the statistical characteristics of the facial motion state. Therefore, the process of driver fatigue can be observed.
- (iii) To eliminate the interference of change of the FFT's area to fatigue driving judgment, we introduce the face projection datum plane and apply the projection principle to extract the motion feature points of the face. Then, based on the motion feature points, we propose the facial motion information entropy, which quantitatively characterizes the chaotic degree of the motion feature points of the face. Then, we train the SVM classifier using the open-source data set YawDD [37]. Experiments show that the

projection datum area S_0 has different values, which will affect the parameters w^T and b of the driver's facial motion information entropy classifier. We design fatigue judgment algorithm based on facial motion information entropy, and the comparison experiments show that our algorithm has an accuracy rate of 94.32% and an algorithm speed of 49.43 ms/f, which further improve the accuracy and speed of the driver's fatigue detection algorithm.

In the future, we will focus on the following research:

- (1) Upload the results of the fatigue detection to the cloud platform and combine the big data analysis techniques to analyze the driver's fatigue period [45]
- (2) Integrate the fatigue driving detection algorithm into ADAS (Advanced Driving Assistant System) [46, 47]
- (3) Expand the applicable environment of the algorithm and explore the driver fatigue detection algorithm based on facial motion information entropy in night environment [48, 49]

Data Availability

The data used to support the findings of this study are available from the first author and the corresponding author upon request.

Conflicts of Interest

The authors declare that there are no conflicts of interest regarding the publication of this article.

Acknowledgments

This work was supported in part by the National Natural Science Foundation of China (Grant no. 51808151), Guangdong Provincial Public Welfare Research and Capacity Building Special Project (Grant no. 2016A020223002), South China University of Technology Central University Fund Project (Grant no. 2017ZD034), Guangdong Provincial Science and Technology Plan Project (Grant no. 2017A040405021), the Fundamental Research Funds for Guangdong Communication Polytechnic (Grant no. 20181014), Guangdong Provincial Natural Science Foundation (Grant no. 2020A151501842), Guangzhou 2020 R&D Plan for Key Areas (Grant no. 202007050004), and by State Key Lab of Subtropical Building Science, South China University of Technology (Grant no. 2020ZB20).

References

- [1] A. Amodio, M. Ermidoro, D. Maggi, S. Formentin, and S. M. Savaresi, "Automatic detection of driver impairment based on pupillary light reflex," *IEEE Transactions on Intelligent Transportation Systems*, vol. 20, no. 8, pp. 3038–3048, 2019.
- [2] X. Li, X. Lian, and F. Liu, "Rear-end road crash characteristics analysis based on Chinese in-depth crash study data," in *Proceedings of the 16th COTA International Conference of Transportation Professionals: Green and Multimodal Transportation and Logistics*, pp. 1536–1545, Shanghai, China, July 2016.
- [3] F. Chen and S. Chen, "Injury severities of truck drivers in single- and multi-vehicle accidents on rural highways," *Accident Analysis & Prevention*, vol. 43, no. 5, pp. 1677–1688, 2011.
- [4] X. Zhu, Z. Dai, F. Chen, X. Pan, and M. Xu, "Using the visual intervention influence of pavement markings for rutting mitigation-part I: preliminary experiments and field tests," *International Journal of Pavement Engineering*, vol. 20, no. 5, pp. 734–746, 2019.
- [5] R. Zhang, F. You, X. N. Chu, L. Guo, Z.-C. He, and R.-B. Wang, "Lane change merging control method for unmanned vehicle under V2V cooperative environment," *China Journal of Highway and Transport*, vol. 31, pp. 180–191, 2018.
- [6] Y. Wang, X. Liu, Y. Zhang, Z. Zhu, D. Liu, and J. Sun, "Driving fatigue detection based on EEG signal," in *Proceedings of the 5th International Conference on Instrumentation and Measurement, Computer, Communication, and Control*, pp. 715–718, Qinhuangdao, China, September 2015.
- [7] R. Bhardwaj, P. Natrajan, and V. Balasubramanian, "Study to determine the effectiveness of deep learning classifiers for ECG based driver fatigue classification," in *Proceedings of the 13th International Conference on Industrial and Information Systems*, pp. 98–102, Punjab, India, December 2018.
- [8] M. K. Sharma and M. M. Bunde, "Design & analysis of k-means algorithm for cognitive fatigue detection in vehicular driver using oximetry pulse signal," in *Proceedings of the IEEE International Conference on Computer, Communication and Control (IC4)*, Indore, India, September 2015.
- [9] L. Boon-Leng, L. Dae-Seok, and L. Boon-Giin, "Mobile-based wearable-type of driver fatigue detection by GSR and EMG," in *Proceedings of the TENCON 2015-2015 IEEE Region 10 Conference*, Macau, China, November 2015.
- [10] J. Yan, H. Kuo, Y. Lin, and T. Liao, "Real-time driver drowsiness detection system based on PERCLOS and grayscale image processing," in *Proceedings of the 2016 International Symposium on Computer, Consumer and Control (IS3C)*, pp. 243–246, Xi'an, China, July 2016.
- [11] G. Niu and C. Wang, "Driver fatigue features extraction," *Mathematical Problems in Engineering*, vol. 2014, Article ID 860517, 10 pages, 2014.
- [12] L. M. Bergasa and J. Nuevo, "Real-time system for monitoring driver vigilance," in *Proceedings of the IEEE International Symposium on Industrial Electronics, 2005. ISIE 2005*, pp. 1303–1308, Dubrovnik, Croatia, June 2005.
- [13] F. You, Y.-h. Li, L. Huang, K. Chen, R.-h. Zhang, and J.-m. Xu, "Monitoring drivers' sleepy status at night based on machine vision," *Multimedia Tools and Applications*, vol. 76, no. 13, pp. 14869–14886, 2017.
- [14] R.-H. Zhang, Z.-C. He, H.-W. Wang, F. You, and K.-N. Li, "Study on self-tuning tyre friction control for developing main-servo loop integrated chassis control system," *IEEE Access*, vol. 5, pp. 6649–6660, 2017.
- [15] F. Chen, M. Song, and X. Ma, "Investigation on the injury severity of drivers in rear-end collisions between cars using a random parameters bivariate ordered probit model," *International Journal of Environmental Research and Public Health*, vol. 16, no. 14, p. 2632, 2019.
- [16] S. Yang, P. Luo, C. C. Loy, and X. Tang, "Wider face: a face detection benchmark," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 5525–5533, IEEE Computer Society, Las Vegas, NV, USA, June 2016.

- [17] S. Luo, C. Xu, and H. Li, "An application of object detection based on YOLOv3 in traffic," in *Proceedings of the 2019 International Conference on Image, Video and Signal Processing - IVSP 2019*, pp. 68–72, Association for Computing Machinery, Shanghai, China, 2019.
- [18] X. Chen, X. Luo, X. Liu, and J. Fang, "Eyes localization algorithm based on prior MTCNN face detection," in *Proceedings of the 2019 IEEE 8th Joint International Information Technology and Artificial Intelligence Conference (ITAIC)*, pp. 1763–1767, Chongqing, China, May 2019.
- [19] D. Sommer and M. Golz, "Evaluation of PERCLOS based current fatigue monitoring technologies," in *Proceedings of the 2010 Annual International Conference of the IEEE Engineering in Medicine and Biology*, pp. 4456–4459, Buenos Aires, Argentina, August 2010.
- [20] X. Sun, H. Zhang, W. Meng, R. Zhang, K. Li, and T. Peng, "Primary resonance analysis and vibration suppression for the harmonically excited nonlinear suspension system using a pair of symmetric viscoelastic buffers," *Nonlinear Dynamics*, vol. 94, no. 2, pp. 1243–1265, 2018.
- [21] G. Wu, F. Chen, X. Pan, M. Xu, and X. Zhu, "Using the visual intervention influence of pavement markings for rutting mitigation—part I: preliminary experiments and field tests," *International Journal of Pavement Engineering*, vol. 20, no. 6, pp. 734–746, 2019.
- [22] P. Viola and M. Jones, "Robust real-time face detection," in *Proceedings Eighth IEEE International Conference on Computer Vision. ICCV 2001*, vol. 2, p. 747, 2001.
- [23] K. Luu, C. Zhu, C. Bhagavatula, T. H. N. Le, and M. Savvides, "A Deep learning approach to joint face detection and segmentation," in *Advances in Face Detection and Facial Image Analysis*, pp. 1–12, Springer International Publishing, Cham, Switzerland, 2016.
- [24] J. Xiang and G. Zhu, "Joint face detection and facial expression recognition with MTCNN," in *Proceedings of the 4th International Conference on Information Science and Control Engineering*, pp. 424–427, Institute of Electrical and Electronics Engineers Inc., Hunan, China, July 2017.
- [25] W. Shi, J. Li, and Y. Yang, "Face fatigue detection method based on MTCNN and machine vision," *Advances in Intelligent Systems and Computing*, Springer Verlag, vol. 1017, pp. 233–240, Huainan, China, 2020.
- [26] S. Zhao, H. Song, W. Cong, Q. Qi, and H. Tian, "End-to-end cascade cnn for simultaneously face detection and alignment," in *Proceedings of the 2017 International Conference on Virtual Reality and Visualization (ICVRV)*, pp. 35–40, Institute of Electrical and Electronics Engineers Inc., 2017, Zhengzhou, China.
- [27] M. El-Arabawy, S. Zaki, and F. Harby, "Improved AdaBoost algorithm for face detection," in *Proceedings of the 2010 International Conference on Image Processing, Computer Vision, and Pattern Recognition*, vol. 1, CSREA Press, Las Vegas, NV, USA, pp. 353–358, 2010.
- [28] J. Redmon, S. Divvala, R. Girshick, and A. Farhadi, "You only look once: unified, real-time object detection," in *Proceedings of the 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 779–788, IEEE Computer Society, Las Vegas, NV, USA, 2016.
- [29] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "ImageNet classification with deep convolutional neural networks," *Communications of the ACM*, vol. 60, no. 6, pp. 84–90, 2017.
- [30] Y. Lecun, L. Bottou, Y. Bengio, and P. Haffner, "Gradient-based learning applied to document recognition," *Proceedings of the IEEE*, vol. 86, no. 11, pp. 2278–2324, 1998.
- [31] K. D. E. Dlib-ml, "A machine learning toolkit," *Journal of Machine Learning Research*, vol. 10, pp. 1755–1758, 2009.
- [32] J. H. Friedman, "Greedy function approximation: a gradient boosting machine," *The Annals of Statistics*, vol. 29, no. 5, pp. 1189–1232, 2001.
- [33] X. Cao, Y. Wei, F. Wen, and J. Sun, "Face alignment by explicit shape regression," *International Journal of Computer Vision*, vol. 107, no. 2, pp. 177–190, 2014.
- [34] P. Dollar, P. Welinder, and P. Perona, "Cascaded pose regression," in *Proceedings of the 2010 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, CVPR 2010*, pp. 1078–1085, IEEE Computer Society, San Francisco, CA, USA, June 2010.
- [35] H. Wang, F. You, X. Chu, X. Li, and X. Sun, "Research on customer marketing acceptance for future automatic driving—a case study in China city," *IEEE Access*, vol. 7, pp. 20938–20949, 2019.
- [36] L. Jiang, H. Wang, S. Gao, and S. Jiang, "Research of the automotive driver fatigue driving early warning system," *Communications in Computer and Information Science*, Springer-Verlag, Berlin, Germany, pp. 383–391, 2011.
- [37] G. Sun, Y. Jin, Z. Li, F. Zhang, and L. Jia, "A vision-based head status judging algorithm for driving fatigue detection system," *Advances in Transportation Studies*, vol. 2015, pp. 51–64, 2015.
- [38] C. E. Shannon, "A mathematical theory of communication," *Bell System Technical Journal*, vol. 27, 1948.
- [39] S. Abtahi, M. Omidyeganeh, S. Shirmohammadi, and B. Hariri, "YawDD," in *Proceedings of the 5th ACM Multimedia Systems Conference, MMSys 2014*, pp. 24–28, Association for Computing Machinery, Singapore, March 2014.
- [40] Z. You, Y. Gao, J. Zhang, H. Zhang, M. Zhou, and C. Wu, "A study on driver fatigue recognition based on SVM method," in *Proceedings of the 4th International Conference on Transportation Information and Safety, ICTIS 2017*, pp. 693–697, Institute of Electrical and Electronics Engineers Inc., Banff, Canada, August 2017.
- [41] J. Hernández-Orallo, "ROC curves for regression," *Pattern Recognition*, vol. 46, no. 12, pp. 3395–3411, 2013.
- [42] L. Tychsen-Smith and L. Petersson, "Improving object localization with fitness NMS and bounded IoU loss," in *Proceedings of the 31st Meeting of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, CVPR 2018*, pp. 6877–6885, IEEE Computer Society, Salt Lake City, UT, USA, June 2018.
- [43] G. Lei, X. Liang, Z. Xiao, and Y. Li, "Real-time driver fatigue detection based on morphology infrared features and deep learning," *Infrared & Laser Engineering*, vol. 47, no. 2, Article ID 203009, 2018.
- [44] J. M. Guo and M. Herleeyandi, "Driver drowsiness detection using hybrid convolutional neural network and long short-term memory," *Multimedia Tools & Applications*, vol. 78, no. 20, pp. 29059–29087, 2019.
- [45] C. Xu, Y. Yang, S. Jin, Z. Qu, and L. Hou, "Potential risk and its influencing factors for separated bicycle paths," *Accident Analysis & Prevention*, vol. 87, pp. 59–67, 2016.
- [46] F. Chen, H. Peng, X. Ma, J. Liang, W. Hao, and X. Pan, "Examining the safety of trucks under crosswind at bridge-tunnel section: a driving simulator study," *Tunnelling and Underground Space Technology*, vol. 92, Article ID 103034, 2019.
- [47] H. Xiong, X. Zhu, and R. Zhang, "Energy recovery strategy numerical simulation for dual axle drive pure electric vehicle based on motor loss model and big data calculation," *Complexity*, vol. 2018, Article ID 4071743, 14 pages, 2018.

- [48] X. Qu, M. Zhou, Y. Yu, C. T. Lin, and X. Wang, "Jointly dampening traffic oscillations and improving energy consumption with electric, connected and automated vehicles: a reinforcement learning based approach," *Applied Energy*, vol. 257, Article ID 114030, 2019.
- [49] M. Zhou, Y. Yu, and X. Qu, "Development of an efficient driving strategy for connected and automated vehicles at signalized intersections: a reinforcement learning approach," *IEEE Transactions on Intelligent Transportation Systems*, vol. 21, no. 1, pp. 433–443, 2019.