

## Article

# Driver Emotion and Fatigue State Detection Based on Time Series Fusion

Yucheng Shang <sup>1</sup>, Mutian Yang <sup>2</sup>, Jianwei Cui <sup>1,\*</sup>, Linwei Cui <sup>1</sup>, Zizheng Huang <sup>1</sup> and Xiang Li <sup>1</sup>

<sup>1</sup> Institute of Instrument Science and Engineering, Southeast University, Nanjing 210096, China

<sup>2</sup> School of Information Science and Engineering, China University of Petroleum, Beijing 266580, China

\* Correspondence: [cjw@seu.edu.cn](mailto:cjw@seu.edu.cn)

**Abstract:** Studies have shown that driver fatigue or unpleasant emotions significantly increase driving risks. Detecting driver emotions and fatigue states and providing timely warnings can effectively minimize the incidence of traffic accidents. However, existing models rarely combine driver emotion and fatigue detection, and there is space to improve the accuracy of recognition. In this paper, we propose a non-invasive and efficient detection method for driver fatigue and emotional state, which is the first time to combine them in the detection of driver state. Firstly, the captured video image sequences are preprocessed, and Dlib (image open source processing library) is used to locate face regions and mark key points; secondly, facial features are extracted, and fatigue indicators, such as driver eye closure time (PERCLOS) and yawn frequency are calculated using the dual-threshold method and fused by mathematical methods; thirdly, an improved lightweight RM-Xception convolutional neural network is introduced to identify the driver's emotional state; finally, the two indicators are fused based on time series to obtain a comprehensive score for evaluating the driver's state. The results show that the fatigue detection algorithm proposed in this paper has high accuracy, and the accuracy of the emotion recognition network reaches an accuracy rate of 73.32% on the Fer2013 dataset. The composite score calculated based on time series fusion can comprehensively and accurately reflect the driver state in different environments and make a contribution to future research in the field of assisted safe driving.



**Citation:** Shang, Y.; Yang, M.; Cui, J.; Cui, L.; Huang, Z.; Li, X. Driver Emotion and Fatigue State Detection Based on Time Series Fusion. *Electronics* **2023**, *12*, 26. <https://doi.org/10.3390/electronics12010026>

Academic Editors: Silvia Liberata Ullo and Donghyeon Cho

Received: 7 November 2022

Revised: 28 November 2022

Accepted: 16 December 2022

Published: 21 December 2022



**Copyright:** © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

## 1. Introduction

According to statistics, the number of deaths caused by traffic accidents has reached 1.35 million worldwide [1], and among them, traffic accidents caused by fatigued driving account for approximately 20%–30% of all traffic accidents [2]. Studies have shown that uncontrollable emotions are one of the primary elements that raise driving risks [3], such as anger that may cause road rage [4], and sadness and stress that can reduce driver concentration [5]. Approximately 90% of traffic accidents can be avoided if drivers are warned before they occur [6]. Therefore, in order to reduce and avoid traffic accidents, it is significant to identify driver fatigue and emotional state and warn them with the help of assisted driving systems.

The existing driver fatigue detection methods can be divided into three main categories: based on vehicle behavior [7], based on physiological signals [8–11], and based on visual features [12–14]. Based on vehicle behavior, the driver's fatigue can be indirectly judged, such as whether the car is pressing the line, whether the distance between the car and another car is too close, etc. However, due to the complex road conditions of the actual scene and the great differences in drivers' driving habits, it is difficult to develop a unified standard to determine fatigue, and its main drawback is the low accuracy rate; common physiological signals used to detect fatigue include electrocardiogram (ECG),

electroencephalogram (EEG), and electrooculogram (EOG), which have the advantages of fast detection speed and high accuracy rate, but complex wearable devices are usually costly, their operation complicated and inconvenient to use, and to a certain extent they will interfere with drivers' driving. Extracting visual information of faces using cameras and performing fatigue detection based on these features is an effective method [15].

The detection and identification of driver emotions has become an emerging topic in human-machine systems for intelligent vehicles [16]. While driver emotion recognition is mainly achieved through visual information, deep learning-based feature extraction methods are superior to traditional manual feature extraction. However, the complex network model poses new challenges to the computational power of computers.

The above research can achieve driver emotion and fatigue detection, but driver fatigue and emotional state both increases driving risk, and both of them affect each other, so the combined research can effectively improve the accuracy of driver state detection; meanwhile, both emotional and fatigue states cannot be accurately determined in real-time based on the facial features of a single frame [17,18]. Therefore, this paper proposes a non-invasive and efficient detection method based on time series fusion to identify driver fatigue and emotional states. This method can simultaneously detect the driver's emotional and fatigue states and provide early warning of potential driver-induced driving risks based on the fusion index scores, contributing to future research in the field of assisted safe driving.

The innovative work in this paper includes three main aspects:

- (1) Firstly, the established multi-feature dual-threshold fatigue detection model incorporates fatigue metrics, such as head posture, fatigue eye closure frequency, eye closure duration, yawn frequency, etc., and shows superior performance compared with several classical fatigue detection algorithms;
- (2) Secondly, the improved lightweight RM-Xception convolutional neural network model for emotion recognition, which performs well in expression feature extraction capability, achieving an accuracy of 73.32% on the Fer2013 expression dataset;
- (3) Thirdly, the method proposed in this paper combines driver fatigue and emotional state for the first time, based on time series fusion metrics, which more accurately and comprehensively reflects the driver state.

The remainder of the paper is organized as follows. In Section 2 we present the algorithm design for driver emotion and fatigue detection and how to integrate the two. Section 3 conducts experimental tests of the proposed algorithm. Section 4 discusses the main contributions made in this paper and proposes future research directions.

## 2. Related Work

### 2.1. Fatigue Detection Methods

For driver fatigue detection based on vehicle information, Li et al. detected driver fatigue by detecting the driver's grip on the steering wheel, extracted fatigue features using wavelet transform, and compared the performance of algorithms, such as SVM and K-neighborhood in distinguishing driver status [19]. Zhang et al. proposed a driver fatigue detection method based on steering wheel angle features, built a detection model using support vector machines, and optimized the model parameters through cross-validation [20].

For driver fatigue detection based on physiological signals, Sobhan et al. proposed a fatigue detection model based on a deep convolutional neural network-long and short term memory network to extract fatigue features from six active regions and raw EEG data [10]. Chai et al. proposed an EEG-based dichotomous fatigue detection method using autoregressive modeling for feature extraction and Bayesian neural networks for feature classification [11]. Lin et al. proposed a novel brain-computer interface system to detect human physiological states by acquiring EEG signals in real time. The system is implemented to detect drowsiness in real time based on EEG and provide warning information to the user when there is a need [21].

For driver fatigue detection based on visual information, Zhu et al. designed a driver fatigue detection algorithm based on facial key points by constructing a deep convolutional network to detect face regions, calculating eye aspect ratio (EAR), mouth aspect ratio (MAR) and eye closure time the driver fatigue assessment model was established by calculating eye aspect ratio (EAR), mouth aspect ratio (MAR) and eye closure time percentage (PERCLOS) based on facial key points [22]. He et al. proposed a fatigue detection method based on two CNN cascades, and the two CNNs were used for face feature detection and eye and mouth state classification, respectively [23]. Bin et al. proposed a fatigue detection algorithm based on facial multi-feature fusion with blink and yawn frequencies [24]. Li et al. designed a fatigue driving detection algorithm based on facial multi-feature fusion, which introduces an improved YOLOv3 algorithm to capture facial regions and calculate the driver's eye closure time, blink frequency, and yawn frequency to assess the fatigue state through eye and mouth feature vectors [25]. Chen et al. proposed a fatigue detection model based on BP neural network and time-accumulation effect to reflect the accumulation process of fatigue with time [15]. Yu et al. proposed a fatigue detection method based on 3D deep convolutional neural networks, with input multiple frames to generate a spatio-temporal representation, and combined with scene conditions to generate fused features for discriminating sleepiness detection [17]. Facial multi-feature fusion can greatly improve the accuracy and robustness of fatigue detection.

## 2.2. Emotion Recognition Methods

For emotion recognition based on physiological signals, Robert et al. proposed an EEG-based feature extraction method for emotion recognition, using a machine learning technique to filter and compare features on a dataset they created [26]. Joao et al. used facial electromyography for the detection of emotional states and proposed a framework that combines electromyographic detection of all over expressions and oculogram-based sweep detection to classify emotions into four categories: neutral, sad, angry and happy [27].

To recognize people's mental states based on speech signals, Renato et al. propose emotion-related audio functions to advance music emotion recognition techniques. The paper proposes algorithms related to musical textures and expression techniques, and creates a public dataset of 900 audios to evaluate the algorithms [28]. Han et al. used deep neural networks to generate probability distributions of emotional states for speech segments, then constructed discourse-level features from the probability distributions and then fed these features into an extreme learning machine to achieve speech emotion recognition [29].

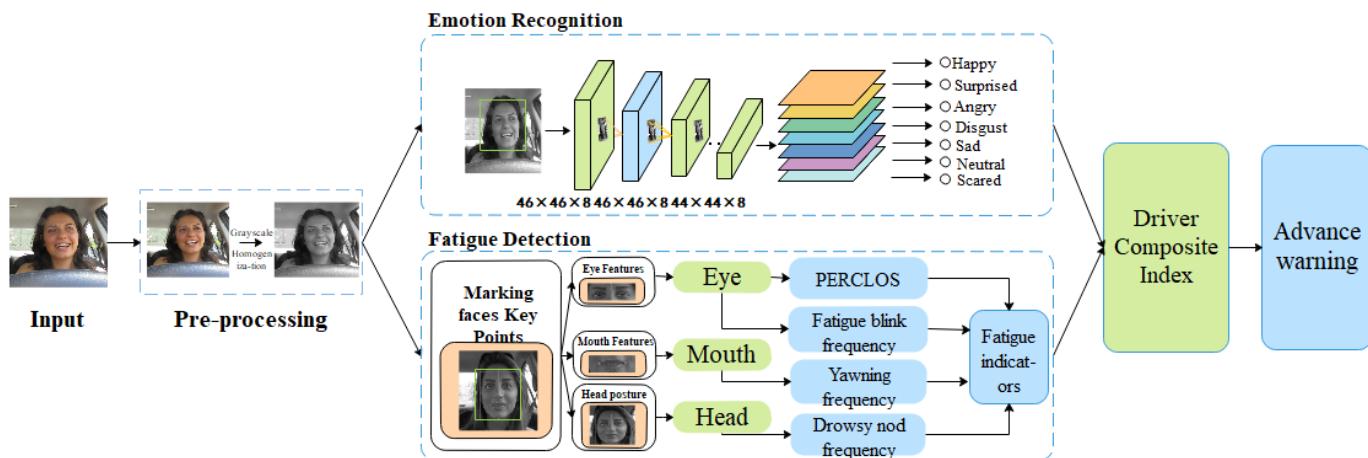
For emotion recognition based on visual features, Mohan et al. designed a facial expression recognition model based on deep convolutional networks using a hierarchical fusion method of local feature classification and overall feature classification [12]. Minaee et al. proposed a facial expression recognition method using attentional convolutional networks [13]. Xiao et al. designed a facial expression-based driver emotion recognition network called FERDERnet, which mainly consists of three parts: a face detection module, a data enhancement and resampling module, and an emotion recognition module [14]. Li et al. proposed an emotion recognition method based on video sequences, using the visual information of facial expression sequences and the speech information of audio to fuse the judgment, and using convolutional neural networks to improve the recognition performance of facial expressions [18]. Kansizoglu et al. implemented continuous emotion recognition via recurrent neural networks for long-term behavior modeling [30]. The above emotion recognition methods, which build more complex models and require processing a large amount of image data, present new challenges to the computing power of computers.

## 3. Materials and Methods

This system collects and processes the driver's facial information, analyzes the driver's emotion and fatigue level in real-time, effectively monitors and actively warns of accidents

due to the driver's causes, and realizes a kind of assisted driving early warning system for real-time status monitoring of the driver.

The hardware part of this system uses a Raspberry Pi as the central controller and a CSI camera for video image data acquisition. The Raspbian operating system is installed on the Raspberry Pi, and the OpenCV and Dlib libraries are installed in the operating system for image video processing and facial key point identification. The CSI camera was selected mainly for its low cost, the limited computing resources of the Raspberry Pi, and its CSI camera with a better resolution. The system block diagram is presented in Figure 1.



**Figure 1.** System architecture block diagram.

### 3.1. Image Pre-Processing and Face Detection

#### 3.1.1. Image Pre-Processing

In this paper, we employ Raspberry and CSI cameras to capture video data. The acquired color image stream will be grayed out frame by frame, and the grayed-out image can reduce the impact of external factors, such as illumination, while preserving the image information intact, and the data volume is small and easy to conduct matrix operations.

When the influence of environmental factors, such as illumination on the image is high, the grayscale operation can play a limited role, and the extracted information does not adequately reflect the emotional and fatigue information. In this paper, histogram equalization is performed on the grayed-out image, and the local information in the processed image changes significantly, and the darker parts become brighter while the brighter parts do not appear exposed, which can solve the impact of uneven illumination on feature extraction.

#### 3.1.2. Face and Key Point Detection

Dlib is an open-source toolkit built on C++ that includes a Python development interface. Among them, shape\_predictor\_68\_face\_landmarks.dat is used for facial 68 key points detection. This model library includes two important methods, face detector and face key points predictor, for face feature extraction, returning the coordinates of face feature points, face angle, and other parameters. The Dlib library detects and extracts the facial regions of interest in images quickly.

After pre-processing, the image is given to the face detector. After the face is successfully detected, a bounding box is applied to the image to extract the region of interest (ROI) for further analysis. If the face is not detected, the next frame will be processed. The extracted ROI will be fed into the face key point predictor to mark 68 key points of the face, as shown in Figure 2.



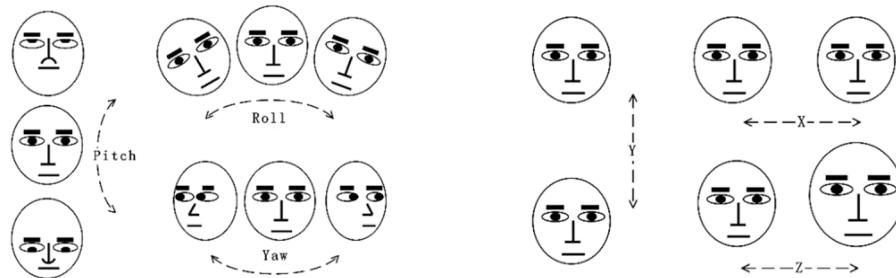
**Figure 2.** Image pre-processing and face detection.

### 3.2. Multi-Feature Double-Threshold Fatigue Recognition Algorithm

#### 3.2.1. Key Features Selection

##### Head Posture Estimation

Head posture is one of the indispensable indicators for fatigue detection of the driver. When employing the camera for head posture estimation of the driver, the conversion of the coordinate system is indispensable. The coordinate systems involved include four types of coordinate systems: world coordinate system, camera coordinate system, image center coordinate system, and pixel coordinate system. A 3D rigid body has two types of motion relative to the camera, translation and rotation. The translation motion includes X, Y, and Z degrees of freedom movement, while the rotation motion is described by Euler angles, including three types of horizontal roll, vertical pitch, and yaw. The essence of the driver's head posture estimation is to find these six parameters, as shown in Figure 3.



**Figure 3.** Estimated parameters of head posture.

Suppose that a point P ( $U, V, W$ ) in the world coordinate system is known. Assuming that the rotation matrix and translation vector are known to be  $R$  and  $t$ , respectively, we can then calculate the position of P in the camera coordinate system ( $X, Y, Z$ ) as follows.

$$\begin{bmatrix} X \\ Y \\ Z \end{bmatrix} = R \begin{bmatrix} U \\ V \\ W \end{bmatrix} + t \Rightarrow \begin{bmatrix} X \\ Y \\ Z \end{bmatrix} = [R|t] \begin{bmatrix} U \\ V \\ W \\ 1 \end{bmatrix} \quad (1)$$

From the camera, the coordinate system to the pixel coordinate system can be calculated from Equation (2).

$$\begin{bmatrix} x \\ y \\ 1 \end{bmatrix} = s \begin{bmatrix} f_x & 0 & c_x \\ 0 & f_y & c_y \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \end{bmatrix} \quad (2)$$

where  $f_x$  and  $f_y$  are the lengths of the focal lengths in the x- and y-axis directions, respectively,  $(c_x, c_y)$  is the optical center, and for practical applications, the radial aberration parameter is omitted, and  $S$  is the scale factor.

Therefore, the relationship between the pixel coordinate system and the world coordinate system is shown as follows:

$$s \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} = \begin{bmatrix} fx & 0 & cx \\ 0 & fy & cy \\ 0 & 0 & 1 \end{bmatrix} [R|t] \begin{bmatrix} U \\ V \\ W \\ 1 \end{bmatrix} \quad (3)$$

The equation can be solved by DLT (Direct Linear Transform) and least squares, which can calculate the rotation and translation matrices from which the Euler angles can be found.

$$R = \begin{bmatrix} r_{00} & r_{01} & r_{02} \\ r_{10} & r_{11} & r_{12} \\ r_{20} & r_{21} & r_{22} \end{bmatrix} = \begin{bmatrix} \cos \varphi \cos \gamma & -\cos \varphi \sin \gamma & \sin \varphi \\ \cos \varphi \sin \gamma + \sin \varphi \sin \gamma \cos \gamma & \cos \varphi \cos \gamma + \sin \varphi \sin \gamma \sin \gamma & -\sin \varphi \sin \varphi \\ \sin \varphi \sin \gamma - \cos \varphi \cos \gamma \cos \gamma & \sin \varphi \cos \gamma + \cos \varphi \sin \gamma \sin \gamma & \cos \varphi \sin \varphi \end{bmatrix} \quad (4)$$

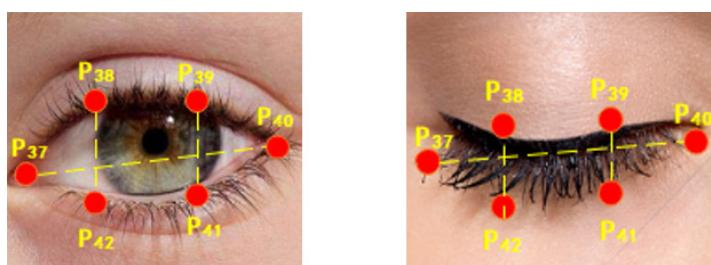
$$\left\{ \begin{array}{l} \varphi = \text{atan}(-r_{12}, r_{22}) \\ \varphi = \text{atan}(r_{02}, \sqrt{r_{12}^2 + r_{22}^2}) \\ \gamma = \text{atan}(-r_{01}, r_{00}) \end{array} \right. \quad (5)$$

OpenCV includes APIs for head pose estimation: solvePnP and solvePnPRansac. In this paper, solvePnP is used to solve the matrix equation.

#### Eye and Mouth Aspect Ratio Definition

The blink information of the driver's eye is one of the most important indicators to reflect the fatigue status. To determine whether the driver blinks or not, the eye aspect ratio (Eye Aspect Ratio,  $E_{AR}$ ) can be calculated by using the Euclidean distance ratio of the longitudinal coordinates of points 38, 42, 39 and 41 of the eye and the transverse coordinates of points 37 and 40 of the eye to calculate the degree of eye-opening, as shown in Figure 4. Taking the left eye as an example,  $E_{AR}$  is calculated as follows:

$$E_{AR} = \frac{||P_{38} - P_{42}|| + ||P_{39} - P_{41}||}{2||P_{37} - P_{40}||} \quad (6)$$

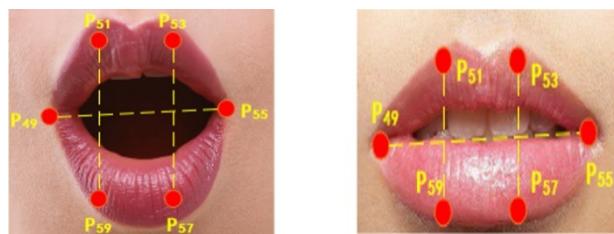


**Figure 4.** Eye feature points.

When the driver's eyes are open,  $E_{AR}$  fluctuates around a particular value to maintain dynamic equilibrium, whereas when the eyes are closed,  $E_{AR}$  decreases rapidly. When  $E_{AR}$  drops below a certain threshold, the human eye is in a closed state. The complex blink discrimination problem is transformed into calculating the Euclidean distance ratio of the eye feature points.

The mouth feature can also be used as an important basis for fatigue discrimination. Similar to the definition of  $E_{AR}$ , the Euclidean distance ratio  $M_{AR}$  is calculated using the vertical coordinates of 51, 59, 53, and 57 at the mouth and the horizontal coordinates of 49 and 55 to determine the degree of mouth opening, as shown in Figure 5. The calculation formula is as follows:

$$M_{AR} = \frac{||P_{51} - P_{59}|| + ||P_{53} - P_{57}||}{2||P_{49} - P_{55}||} \quad (7)$$



**Figure 5.** Mouth feature points.

### 3.2.2. Double-Threshold Fatigue Index Calculation

Head posture, eye closure and mouth opening are all important indicators to identify driver fatigue [31–33]. In this paper, by identifying the driver's head, eyes and mouth states, we use the double-threshold method to calculate four fatigue indicators, such as driver drowsy head nod frequency, fatigue blink frequency, yawn frequency and eye closure rate and then determine the driver's fatigue level.

#### Header Indicator

The head rotation angle is also a significant indicator of fatigue discrimination. When the driver is tired, the head will do a similar nodding or tilting posture, which mainly refers to the change of Pitch and Roll angles in the rotation vector but not much change in the Yaw direction. We can determine whether the driver has performed a head nodding or head tilting movement by comparing the angle of the Pitch or Roll change with the set threshold value.

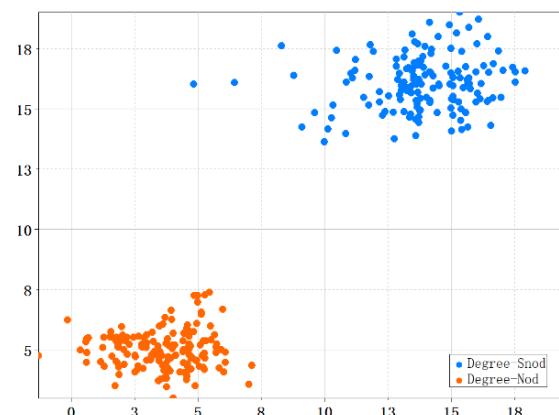
$$\text{Pitch} \geq |T_{h1}| \quad (8)$$

$$\text{Roll} \geq |T_{h2}| \quad (9)$$

In this paper, we choose pitch change to reflect the change of the driver's head posture, the amplitude of the driver's head movement when talking to people is different from the amplitude of drowsy nodding when fatigued, set the threshold to  $T_h$  to distinguish the general head movement and drowsy nodding, and drowsy nodding will last for a period of time, set the threshold  $F_{\text{Hset}}$  by double threshold comparison method can determine whether the driver's head movement is drowsy nodding, as shown in Equation (10).

$$F_{\text{HC}} \geq F_{\text{Eset}} \quad (10)$$

The driver's general head movement Pitch shifts between  $2^\circ$  and  $8^\circ$ , while the Pitch amplitude is higher during drowsy head nodding, between  $12^\circ$  and  $20^\circ$ , as shown in Figure 6.



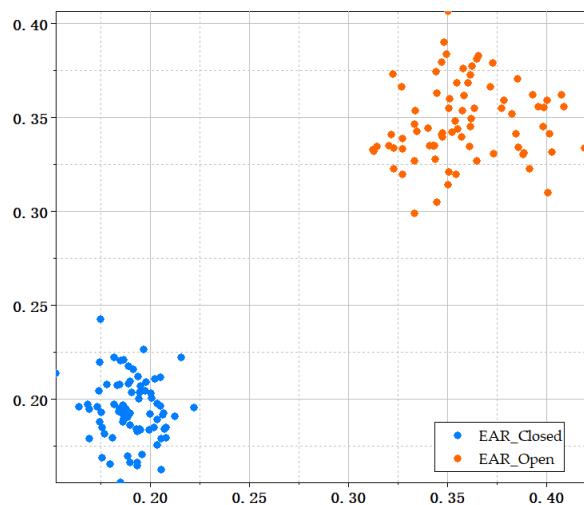
**Figure 6.** Magnitude of pitch variation in different cases.

### Fatigue Eye Closure Indicator

In this paper, a double-threshold comparison method is used to determine whether a driver's blink is a fatigue eye closure. Firstly, the average value of  $E_{AR}$  of the left and right eyes of the current frame of the video is computed and compared with the set threshold  $T_E$ , which in turn determines whether the eyes are open or closed. Secondly, when the eyes are closed in fatigue, the eyes are closed for a longer period of time, and the number of frames  $F_{EC}$  with eyes closed is compared with the set threshold  $F_{Eset}$ , which in turn determines whether the driver is closing his eyes in fatigue. Yawning discrimination and drowsy nodding is the same, the  $E_{AR}$  values of eyes open and closed are shown in Figure 7, and the  $E_{AR}$  values of left and right eyes are calculated and averaged as the driver's real-time  $E_{AR}$  value.

$$E_{AR} \geq T_E \quad (11)$$

$$F_{EC} \geq F_{Eset} \quad (12)$$



**Figure 7.** Difference between EAR when eyes are open and closed.

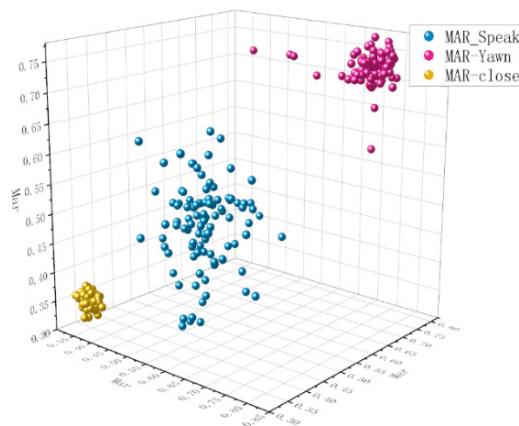
### Yawning Indicator

Since there is a more obvious change in the mouth in the behaviors of yawning, talking, and eating, the threshold  $T_M$  is set to distinguish the change in mouth length-width ratio between yawning and other situations, similar to judging fatigue eye closure, and the double threshold comparison method is used to determine whether the driver is yawning, as shown in Equations (13) and (14), to prevent misjudgment.

$$MAR \geq T_M \quad (13)$$

$$F_{MO} \geq F_{Mset} \quad (14)$$

After testing, the  $M_{AR}$  threshold fluctuated approximately 0.35 when the driver shut up, 0.5 when talking, and 0.73 when yawning, as shown in Figure 8.



**Figure 8.**  $M_{AR}$  values in different cases.

#### Eye Closure Rate (PERCLOS)

The eye closure rate is the percentage of total time spent in eye closure per unit of time. PERCLOS (percentage of eyelid closure over the pupil over time) is one of the most important indicators of driver fatigue. PERCLOS usually has three measurement standards P70, P80, and EM. According to studies, P80 has the strongest correlation with the degree of fatigue, the proportion of time that the eyelid covers more than 80% of the region. Under the condition that the frame rate of the image captured by the camera is constant per unit of time, PERCLOS can be considered as the ratio of the number of frames of eye closure per unit of time  $F_{EC}$  to the total number of frames  $F_{Total}$ . The calculation formula is as follows.

$$\text{PERCLOS} = \frac{F_{EC}}{F_{Total}} \times 100\% \quad (15)$$

#### 3.2.3. Fatigue Recognition Algorithm with Multi-Feature Fusion

Through the previous subsection, narrative can be seen, and the use of Raspberry Pi can be real-time detection of the input video, driver fatigue closed eyes, yawning, drowsy nodding, PERCLOS, and other indicators. With the help of the double threshold comparison method can be counted in the driver unit time  $T$  fatigue closed eyes, yawning, and drowsy nodding the number of times  $N$ .

The formula for calculating the frequency of fatigue eye closure per unit time  $T$  is as follows:

$$F_{blink} = \frac{N_{blink}}{T} \quad (16)$$

The formula for calculating the frequency of yawning per unit of time  $T$  is as follows:

$$F_{yawn} = \frac{N_{yawn}}{T} \quad (17)$$

The formula for calculating the frequency of drowsy nodding per unit of time  $T$  is as follows:

$$F_{nod} = \frac{N_{nod}}{T} \quad (18)$$

Fatigue detection requires the fusion of four indicators, and given the different magnitudes of the four fatigue indicators, a comprehensive evaluation of the fatigue degree requires the normalization of the data and the normalized conversion of the data using the inverse tangent function, as shown in Equation (19). The normalized data are provided in Table 1.

$$x' = \arctan(x) * \frac{2}{\pi} \quad (19)$$

**Table 1.** Fatigue indicators.

Fatigue Indicators	Value	Normalization
Frequency of eye closure for fatigue	$F_{\text{blink}} / (\text{Times} \cdot S^{-1})$	$F'_{\text{blink}}$
Yawning frequency	$F_{\text{Yawn}} / (\text{Times} \cdot S^{-1})$	$F'_{\text{Yawn}}$
Sleepy nod frequency	$F_{\text{Nod}} / (\text{Times} \cdot S^{-1})$	$F'_{\text{Nod}}$
PERCLOS	P / %	$P'$

According to the influence of each index on the driver fatigue degree set different weights, the comprehensive evaluation index  $F$  of the fatigue degree can be calculated as follows:

$$F = \frac{1}{2} \left( \frac{1}{2} F'_{\text{blink}} + \frac{3}{10} F'_{\text{Yawn}} + \frac{1}{5} F'_{\text{Nod}} + P' \right) \quad (20)$$

### 3.3. Improved RM-Xception Emotion Recognition Algorithm

#### 3.3.1. Convolutional Neural Network

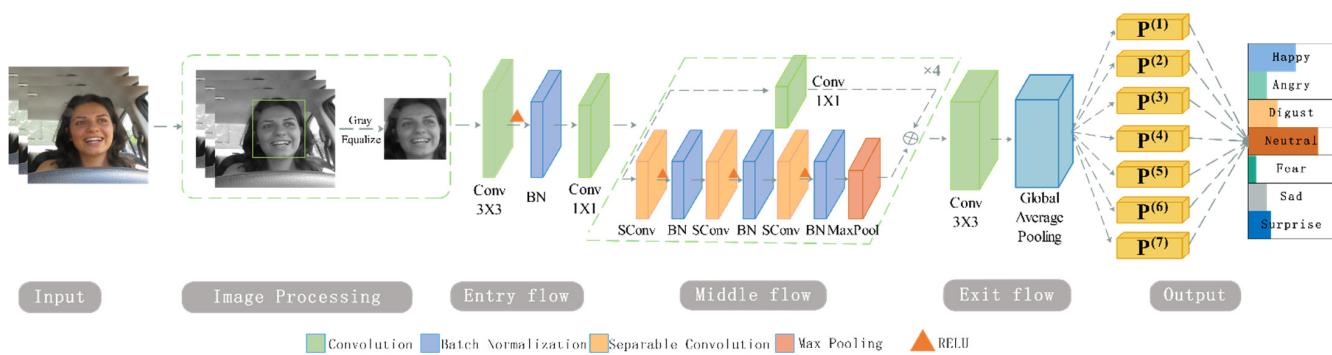
The convolutional neural network evolved from the traditional multilayer neural network development, adding a convolutional layer, a pooling layer, and a feature extraction part to reduce the training parameters while effectively extracting feature information and reducing the complexity of the network. The fully connected layer is used to calculate the loss and obtain the classification results. In this paper, the driver emotion recognition model is trained in an improved RM-Xception convolutional neural network.

#### 3.3.2. Improved RM-Xception Emotion Recognition Algorithm

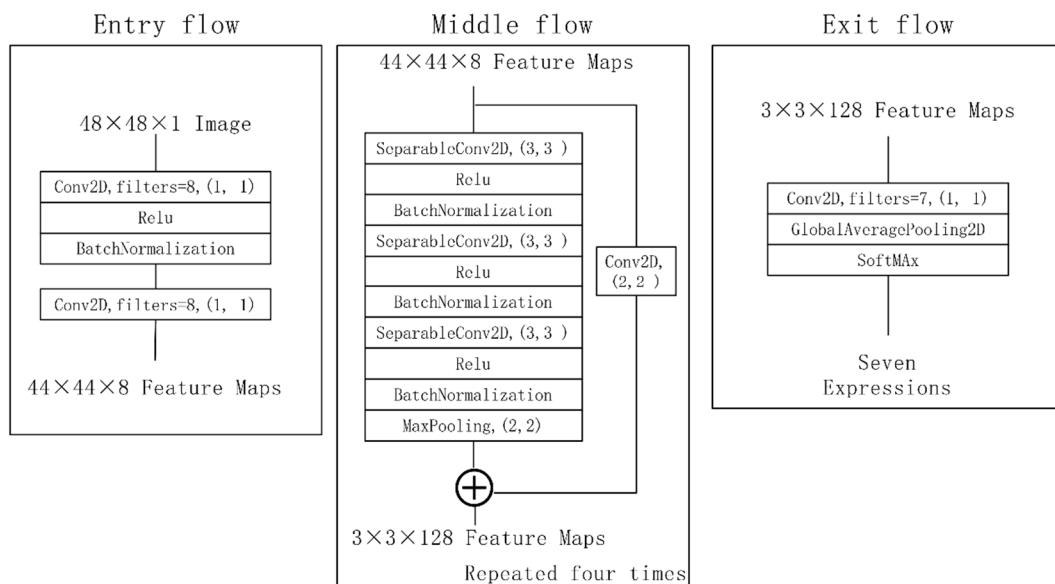
The improved RM-Xception emotion recognition algorithm, firstly, in the selection of activation function, chooses the commonly used activation function RELU in neural networks, also known as modified linear unit, as shown in Equation (21). It does not require exponential operations and has a small computation; it does not have gradient disappearance and can effectively solve the gradient saturation problem.

$$f(x) = \max(0, x) \quad (21)$$

Next, the RM-Xception network is lightly processed, and the overall parameters of the network body are 75,143 among which 73,687 parameters are used for training. The overall structure of the network is shown in Figures 9 and 10, which is divided into three parts: Entry flow, Middle flow, and Exit flow. Firstly, the Entry flow part performs  $3 \times 3$  convolution on the face images of the input network and batch normalization after the activation of the relu function. The relu function and batch normalization operation can reduce the data divergence and further enhance the nonlinear expression capability of the model; secondly, the Middle flow part sends the convolved data into four depth-separable convolution modules with direct residual connection, and each convolution module performs 3 times of depth-separable convolution, activation, batch normalization and then  $1 \times 1$  convolution with direct residual connection; finally, the Exit flow part sends the output of the last module after  $1 \times 1$  convolution and global mean pooling operation. Then sent to the SoftMax classifier to classify and get seven emotions: angry, disgusted, scared, happy, sad, surprised, and neutral.



**Figure 9.** RM-Xception network structure diagram.



**Figure 10.** RM-Xception network structure parameters diagram.

### 3.4. Time Series-Based Emotional Fatigue Feature Fusion Algorithm

When we judge the driver's emotion and fatigue state based on a single frame, there is generally a high degree of uncertainty. This can affect the precision of the system. Emotional and fatigue states are usually expressed as a process, and the driver's state cannot be determined by a single frame of facial information alone. Therefore, this paper constructs a driver state recognition method based on the fusion of emotional and fatigue features in time series, capturing the contextual information of the input video sequence to achieve accurate recognition of the driver condition.

In the field of emotion recognition, psychologists led by Ekman classified people's basic emotions into six categories, namely, happiness, anger, sadness, disgust, fear, and surprise, to which emotions, such as neutrality were later added [34]. In this paper, the calculation of emotions is based on the above seven categories of emotions. When the driver is tired, the emotions flowing from the face are mostly neutral, while tension and fear appear to distract the driver's attention, and anger may cause road rage and increase the driver's safety risk. Therefore, the two modalities of the driver's facial emotion and fatigue status are identified and fused, and the driver's status level is classified according to the score of the fusion index and then the driver is warned in advance or actively intervened.

When the driver is in a happy mood, the auxiliary safety driving system does not need to intervene; when the driver expresses emotions, such as anger, fear, and dread, the system needs to identify and intervene in time. Therefore, according to the actual situation

in which the auxiliary safety system needs to interfere and the impact of different emotions on the driver, the seven emotions are scored, and the scoring table is shown in Table 2.

**Table 2.** Real-time sentiment score table.

Real-Time Emotions	Score	Real-Time Emotions	Score
Happy	-0.001	Anger	0.002
Neutral	0.000	Sadness	0.002
Disgust	0.001	Fear	0.003
Surprise	0.001		

According to the real-time sentiment scoring table, the sentiment score of each frame is obtained in real-time, and the cumulative score NScore in T time is calculated based on the captured video sequence context information and then the sentiment score  $S_T$  per unit time is derived.

$$S_T = \frac{NScore}{T} \quad (22)$$

Fusing the mood score with the composite index score of fatigue, the equation was obtained as follows:

$$S = \frac{1}{2}S_T + \frac{1}{2}F = \frac{1}{2}\frac{NScore}{T} + \frac{1}{2}\left(\frac{1}{2}F'_{\text{blink}} + \frac{3}{10}F'_{\text{yawn}} + \frac{1}{5}F'_{\text{Nod}} + P'\right) \quad (23)$$

The status of the driver is divided into four levels according to the score: suitable for driving, lower risk, higher risk, and unsuitable for driving, as indicated in Table 3.

**Table 3.** Driver status classification.

S Takes the Value of	Status Level	Fatigue Behavioral Manifestations	Advance Warning Measures
<0.01	Suitable for driving	Driver mood and fatigue indicators are normal	None
0.01~0.02	Lower risk	Individual indicators began to increase	Intermittent alerts
0.02~0.03	Higher risks	Indicators with higher scores emerged	Increased alarm frequency
>0.03	Unsuitable for driving	Fatigue or mood scores near maximum, or both at moderate to high levels	Continuous alerts

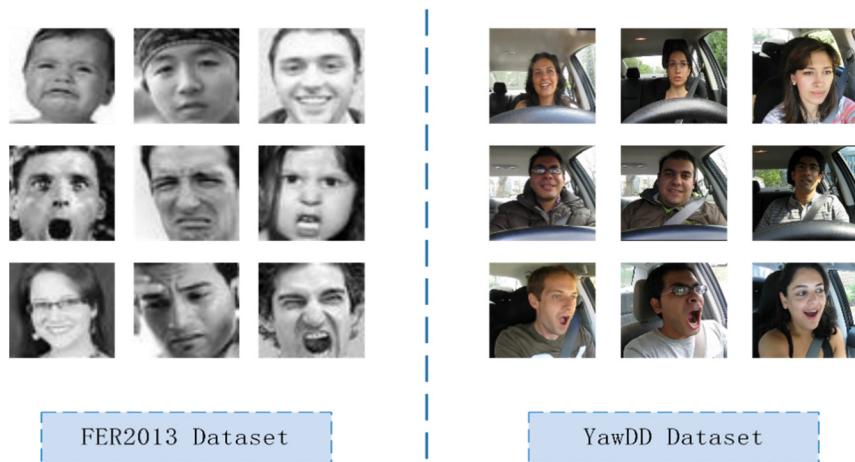
## 4. Results

### 4.1. Experimental Platform and Dataset

In this experimental environment, the hardware configuration for neural network training is an Intel(R) Core(TM) i5-8300H CPU, an NVIDIA GeForce GTX 1050Ti GPU, and Windows 10. The hardware platform on which the system runs is a Raspberry Pi.

The dataset used for emotion recognition is FER2013, which was proposed in the Kaggle facial expression analysis competition and contains 28,709 training samples of seven emotions: angry, disgusted, scared, happy, sad, surprised, and neutral, and 3859 validation and test sets. The human eye recognition rate of this database is roughly between 60% and 70%.

Since there is no publicly available dataset for fatigue detection, the YAWDD dataset is chosen as the sample for experimental validation [35]. The YAWDD dataset is a video dataset recorded by the in-car camera, which records the real-time states of the driver chatting, silence, yawning, etc. In the car under different lighting conditions, as shown in Figure 11.



**Figure 11.** Fer2013 and YawDD datasets.

#### 4.2. Fatigue Detection Experiment

The video stream in YawDD is randomly selected to detect the fatigue characteristic indexes of the driver's eyes, mouth, and head, and the results are recorded in Tables 4 and 5, and the fatigue state of the driver is comprehensively evaluated according to the actual data of the experiment, and the higher the fatigue comprehensive index of fatigue degree represents the more tired is the driver, as shown in Table 6.

**Table 4.** Accuracy of eye fatigue index detection.

Test Number	Number of Actual Blinks (Times/min)	Detects the Number of Blink Counts (Times/min)	Accuracy (%)	The Actual Number of Eye Closures (Times/min)	Detects the Number of Eye Closures (Times/min)	Accuracy (%)
1	9	9	100%	1	1	100%
2	20	21	95.2%	4	4	100%
3	15	15	100%	2	2	100%
4	19	19	100%	1	1	100%
5	23	23	100%	5	5	100%

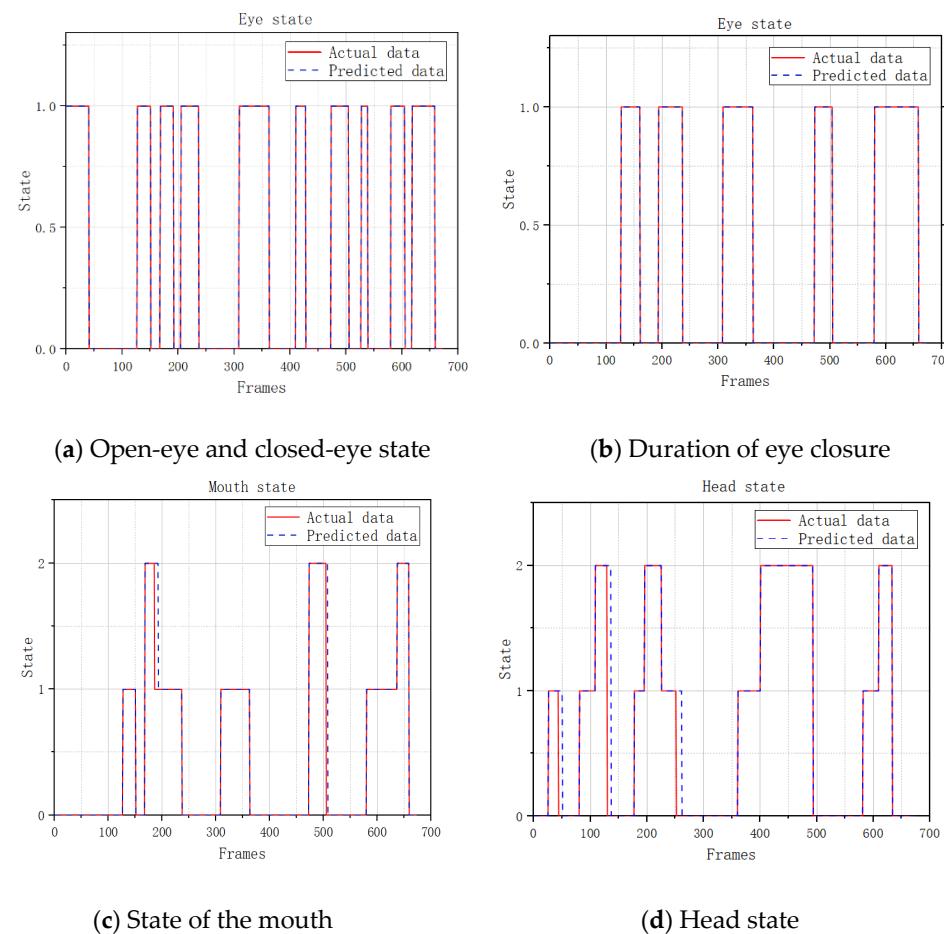
**Table 5.** Accuracy of mouth fatigue index detection.

Test Number	Number of Actual Yawning (Times/min)	Detect the Number of Yawning (Times/min)	Accuracy (%)
1	2	2	100%
2	4	4	100%
3	1	1	100%
4	3	3	100%
5	2	2	100%

**Table 6.** Comprehensive fatigue index experiments.

Test Number	Number of Eye Closures for Fatigue (Times/min)	Number of Yawning (Times/min)	Number of Drowsy Nods (Times/min)	PERC LOS/%	Fatigue Composite Index
1	1	4	0	0.0119	0.0128
2	5	2	4	0.0241	0.0283
3	4	12	0	0.0250	0.0424
4	3	9	1	0.0198	0.0331
5	4	15	4	0.0308	0.0501

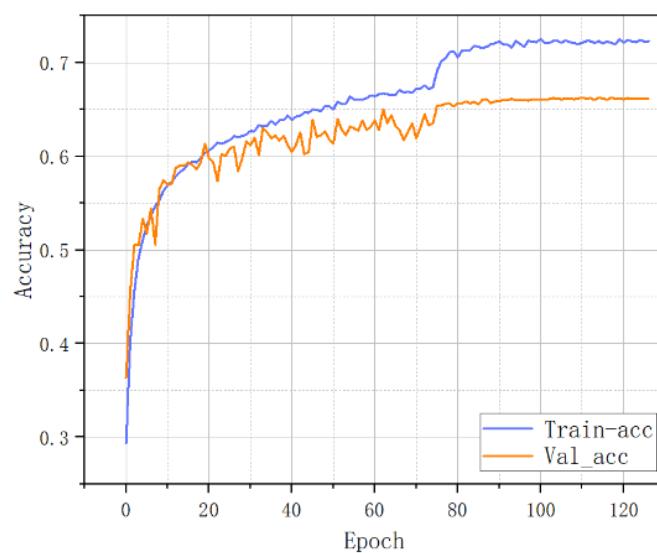
To verify the accuracy of the proposed fatigue detection method, the states of the driver's eyes, mouth, and head were recorded on the Raspberry Pi for a specified number of frames in this paper, as shown in Figure 12. Figure 12a records the frames of the driver's eye-opening and fatigue closing, where 1 represents eye-opening and 0 represents eye-closing; Figure 12b records the duration of eye-opening and eye-closing, where 1 represents eye-opening and 0 represents eye-closing, from the figure, the longest frame of driver's eye closing is 160 frames, which is approximately 10 s, and the driver is in the state of severe fatigue almost sleeping; Figure 12c records the state of driver's mouth, where 0 represents the mouth closed, 1 represents the driver talking, 2 represents the driver yawning; Figure 12d records the driver's head posture, where 0 represents the driver's head in a normal posture, 1 represents a small movement of the head, 2 represents the driver nodding in sleep. After testing, the actual data curve and the recorded data curve almost overlap, which indicates that the fatigue detection method is more accurate.

**Figure 12.** Fatigue indicator accuracy test.

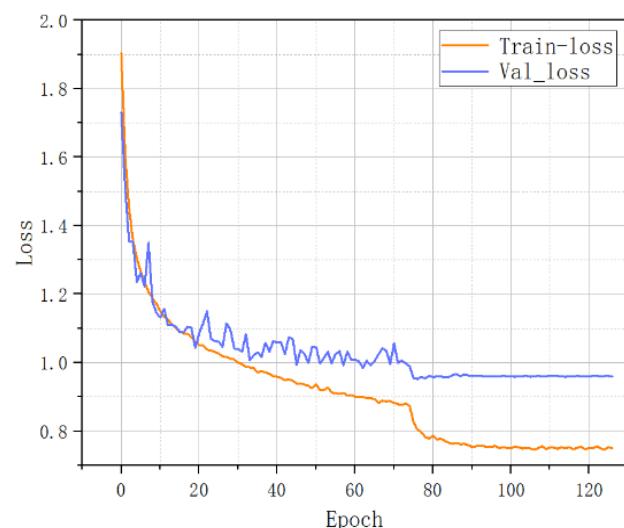
#### 4.3. Emotion Recognition Experiment

In this paper, we use the Fer2013 dataset to train the neural network model, whose data volume is minimal. To strengthen the robustness of the training network, this paper performs data augmentation on the Fer2013 dataset. Data enhancement means artificially flipping, cutting, and rotating the images. Common data enhancement methods include rotation and cropping, color dithering, and so on. In this paper, we set the range of random image rotation to 10 degrees and the range of random scaling to 0.1 without decentering and normalization.

In the process of training, the batch size of sample training is set to 64, the total number of training rounds (epchos) is set to 200, the number of classifications is set to 7, and the Adam optimization algorithm is selected as the optimizer to lower the loss. Convergence speed is fast and the learning effect is good. As the number of training iterations increases, the accuracy of the improved RM-Xception network recognition also increases, and after 127 Epochs, the accuracy reaches 73.32%, as shown in Figure 13. The iteration loss is displayed in Figure 14.



**Figure 13.** Accuracy curve.



**Figure 14.** Loss function curve.

In the classification problem, Precision indicates the probability of actual positive samples among the positive samples predicted by the classifier; Recall indicates the probability of being correctly predicted as positive samples among all positive samples. Precision and recall are calculated as follows:

$$\text{Precision} = \frac{\text{TP}}{\text{TP} + \text{FP}} \quad (24)$$

$$\text{Recall} = \frac{\text{TP}}{\text{TP} + \text{FN}} \quad (25)$$

where TP indicates that the sample is positive and the prediction result is also positive, FP indicates that the sample is negative but the prediction result is positive, and FN indicates that the sample is positive but the prediction result obtained is negative. The accuracy obtained by the method in this paper is 80.82% and the recall rate is 63.01% after 127 rounds of iterative calculation.

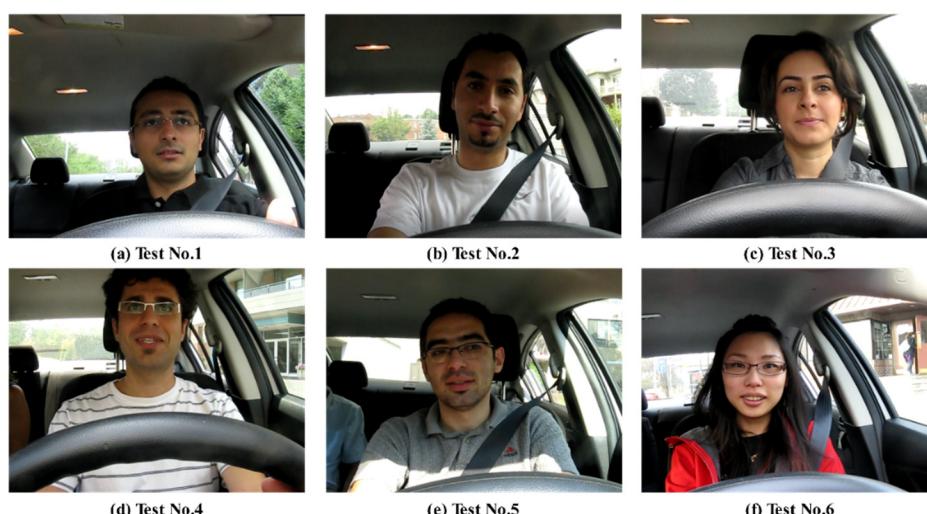
The performance of the method in this paper is compared with other methods in the Table 7.

**Table 7.** Recognition accuracy of different methods on fer2013.

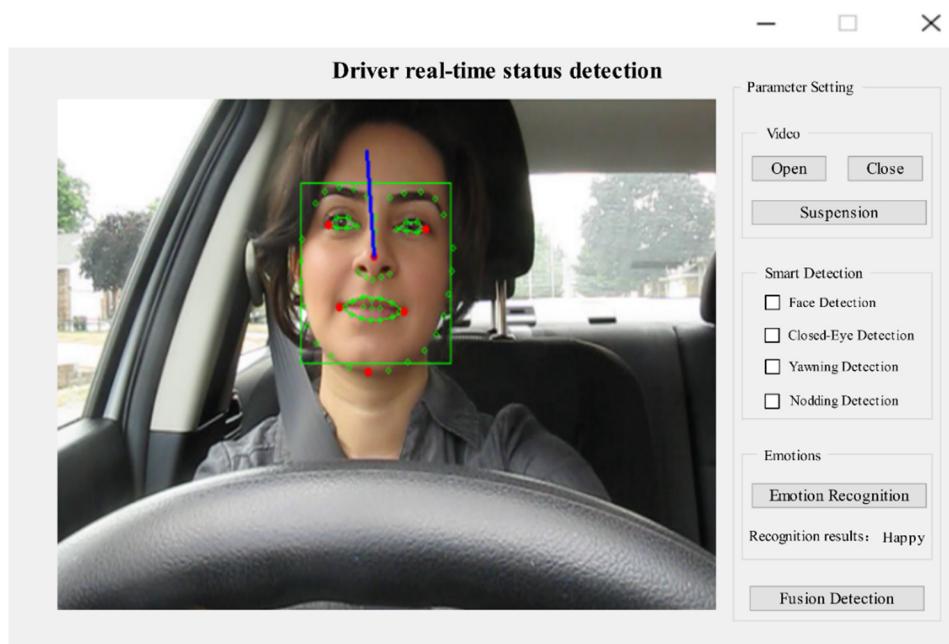
Algorithm	Accuracy/%
Xception	66.80
CNN	65.00
Inception V4	67.01
The algorithm in this paper	73.32

#### 4.4. Driver Status Detection Experiment

To verify the merits of the time-series-based emotional fatigue feature fusion model, six video streams in the YawDD dataset were randomly selected for driver real-time state detection in this paper, as shown in Figure 15. The detection system calculates the driver fatigue and emotional scores in a single frame, based on the time-series accumulation, and calculates the driver state scores per unit of time T according to the method mentioned above. The system operates at a frame rate of approximately 4fps on the Raspberry Pi 4B. Six video sequences were tested in the experiment for one minute of driver status data, the interface of the test system is presented in Figure 16, and the test data are presented in Table 8.



**Figure 15.** Test video number.



**Figure 16.** Driver real-time status test interface.

**Table 8.** Driver status classification experimental test.

Test Number	Fatigue Eyes Closed Times	Yawning Times	Number of Drowsy Nods	PERC LOS/%	Fatigue Comprehensive Indicators	Emotions Score	Comprehensive Status Indicators	Predicted Driving States	Actual Driving Condition
1	2	3	0	1.36	0.015	0.000	0.008	Suitable for driving	Suitable for driving
2	4	1	4	4.53	0.031	0.003	0.017	Lower risk	Lower risk
3	2	2	0	2.35	0.016	-0.009	0.004	Suitable for driving	Suitable for driving
4	1	13	0	1.92	0.029	0.000	0.015	Lower risk	Lower risk
5	4	7	0	2.98	0.031	0.005	0.018	Lower risk	Lower risk
6	8	6	0	3.21	0.041	0.021	0.031	Unsuitable for driving	Unsuitable for driving

From the experimental data, it can be seen that among the tested video sequences, two have a comprehensive driver state score of less than 0.01, which is suitable for driving; three have a comprehensive driver score between 0.01 and 0.02, which has a low driving risk; and one has a driver state score higher than 0.03, which is judged unsuitable for driving by the system. The predicted driver status of the system matches the actual status, proving that the model can truly reflect the driver's status. The video sequence with test number 3 has a negative emotional score, indicating that the driver is in a happy state during that test time, while the driver's eyes close to a certain extent during the process of laughing. If only the fatigue state is identified for the driver, the system is likely to misjudge it as a mild risk, but when combined with the emotion recognition, the system predicts it as suitable for driving, which better expresses the driver's real driving state. The facial state of driver fatigue and emotional performance are inextricably linked. When some drivers are happy, the eye closure increases, which will increase the risk of system misjudgment; in a sad or neutral state, the corresponding fatigue indicators will also increase. Combining the analysis of emotion and fatigue state will more accurately express the driver's current state and increase the robustness of driver state identification. The integrated state indicators will not affect the system's determination due to the level of individual indicators and are more adaptable to complex driving environments.

This study realizes real-time state monitoring of driver emotion and fatigue, which can effectively prevent safety accidents caused by driver fatigue and emotion fluctua-

tions. The experimental data can more realistically and accurately reflect the driver's state, contributing to future research in the field of assisted safe driving.

## 5. Conclusions

This paper implements a time series-based driver fatigue and emotion recognition algorithm for accurate detection of the driver's real-time status with some robustness for complex driving environments. The main findings are as follows:

First, a dual-threshold fatigue recognition algorithm with multi-feature fusion is proposed, which can greatly reduce the external environment's interference and improve the system's accuracy in determining driver fatigue by graying and histogram equalization of the face images captured by the micro camera. Following that, the driver's head posture, and eye and mouth states are recognized, fatigue indicators are calculated, and the fatigue composite score obtained by fusion according to mathematical methods can rapidly and accurately reflect the driver's fatigue level.

Second, the improved RM-Xception algorithm, which introduces a depth-separable convolution module and a residual module, lightens the Xception processing, significantly reduces the computational resources required to train the network, and trains a model with high emotion extraction capability. Meanwhile, by training the model using the data-enhanced images, the obtained network has better robustness, and the model finally achieves 73.32% accuracy on the Fer2013 dataset. Based on the time series to calculate the driver's emotional indicators over a period of time, it can truly reflect the driver's emotional state at a certain time and make a certain contribution to the field of assisted safe driving in the future.

In future work, this paper will test and improve the driving state recognition algorithm in more realistic scenarios and complex environments, consider combining driving data collected by multiple sensors, consider the applicability of the algorithm under different light intensities, and further investigate the relationship between a driver's facial state and his or her emotions when fatigued.

**Author Contributions:** Y.S.: design of the driver emotion recognition algorithm, preparation and writing of the draft. J.C.: funding acquisition, thesis revision work and project design. M.Y. and L.C.: design of driver fatigue detection algorithm. Z.H. and X.L.: preparation of the dataset and training of the neural network. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research was funded by China's National Key R&D Program (2020YFC2007401): Multimodal intelligent sensing, human-machine interaction and active safety technology.

**Institutional Review Board Statement:** The study does not require ethical approval.

**Informed Consent Statement:** Informed consent was obtained from all subjects involved in the study.

**Data Availability Statement:** The data presented in this study are available on request from the corresponding author.

**Conflicts of Interest:** The authors declare no conflict of interest.

## References

1. World Health Organization. *Global Status Report on Road Safety 2018: Summary*; Technical Report; World Health Organization: Geneva, Switzerland, 2018.
2. Alvaro, P.K.; Burnett, N.M.; Kennedy, G.A.; Min, W.Y.X.; McMahon, M.; Barnes, M.; Jackson, M.; Howard, M.E. Driver education: Enhancing knowledge of sleep, fatigue and risky behaviour to improve decision making in young drivers. *Accid. Anal. Prev.* **2018**, *112*, 77–83. [[CrossRef](#)] [[PubMed](#)]
3. Li, G.; Lai, W.; Sui, X.; Li, X.; Qu, X.; Zhang, T.; Li, Y. Influence of traffic congestion on driver behavior in post-congestion driving. *Accid. Anal. Prev.* **2020**, *141*, 105508. [[CrossRef](#)] [[PubMed](#)]
4. Eon, M. Towards affect-integrated driving behavior research. *Theor. Issues Ergon. Sci.* **2015**, *16*, 553–585.
5. Lee, Y.C. Measuring drivers' frustration in a driving simulator. In *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*; Sage Publications: Los Angeles, CA, USA, 2010; Volume 54.

6. Koh, S.; Cho, B.R.; Lee, J.; Kwon, S.-O.; Lee, S.; Lim, J.B.; Lee, S.B.; Kweon, H.-D. Driver drowsiness detection via PPG biosignals by using multimodal head support. In Proceedings of the 2017 4th International Conference on Control, Decision and Information Technologies (CoDiT), Barcelona, Spain, 5–7 April 2017; pp. 383–388.
7. Kulathumani, A.; Soua, R.; Karray, F.; Kamel, M.S. Recent trends in driver safety monitoring systems: State of the art and challenges. *IEEE Trans. Veh. Technol.* **2017**, *66*, 4550–4563.
8. Balandong, R.P.; Ahmad, R.F.; Saad, M.N.M.; Malik, A.S. A review on EEG-based automatic sleepiness detection systems for driver. *IEEE Access* **2018**, *6*, 2290822919. [\[CrossRef\]](#)
9. Rohit, F.; Kulathumani, V.; Kavi, R.; Elwarfalli, I.; Kecojevic, V.; Nimbarate, A. Real-time drowsiness detection using wearable, lightweight brain sensing headbands. *IET Intell. Transp. Syst.* **2017**, *11*, 255–263. [\[CrossRef\]](#)
10. Sheykhhivand, S.; Rezaii, T.Y.; Mousavi, Z.; Meshgini, S.; Makouei, S.; Farzamnia, A.; Danishvar, S.; Teo Tze Kin, K. Automatic Detection of Driver Fatigue Based on EEG Signals Using a Developed Deep Neural Network. *Electronics* **2022**, *11*, 2169. [\[CrossRef\]](#)
11. Chai, R.; Naik, G.R.; Nguyen, T.N.; Ling, S.H.; Tran, Y.; Craig, A.; Nguyen, H.T. Driver fatigue classification with independent component by entropy rate bound minimization analysis in an EEG-based system. *IEEE J. Biomed. Health Inform.* **2017**, *21*, 715–724. [\[CrossRef\]](#)
12. Mohan, K.; Seal, A.; Krejcar, O.; Yazidi, A. Facial Expression Recognition Using Local Gravitational Force Descriptor-Based Deep Convolution Neural Networks. *IEEE Trans. Instrum. Meas.* **2020**, *70*, 5003512. [\[CrossRef\]](#)
13. Minaee, S.; Minaei, M.; Abdolrashidi, A. Deep-emotion: Facial expression recognition using the attentional convolutional network. *Sensors* **2021**, *21*, 3046. [\[CrossRef\]](#)
14. Xiao, H.; Li, W.; Zeng, G.; Wu, Y.; Xue, J.; Zhang, J.; Li, C.; Guo, G. On-Road Driver Emotion Recognition Using Facial Expression. *Appl. Sci.* **2022**, *12*, 807. [\[CrossRef\]](#)
15. Chen, J.; Yan, M.; Zhu, F.; Xu, J.; Li, H.; Sun, X. Fatigue Driving Detection Method Based on Combination of BP Neural Network and Time Cumulative Effect. *Sensors* **2022**, *22*, 4717. [\[CrossRef\]](#) [\[PubMed\]](#)
16. Braun, M.; Chadowitz, R.; Alt, F. User Experience of Driver State Visualizations: A Look at Demographics and Personalities. In Proceedings of the IFIP Conference on Human-Computer Interaction, Paphos, Cyprus, 2–6 September 2019; Springer: Cham, Switzerland, 2019; pp. 158–176.
17. Yu, J.; Park, S.; Lee, S.; Jeon, M. Representation Learning, Scene Understanding, and Feature Fusion for Drowsiness Detection. In *Computer Vision—Accv 2016 Workshops, Pt Iii.*; Chen, C.S., Lu, J., Ma, K.K., Eds.; Springer International Publishing Ag: Cham, Switzerland, 2017; Volume 10118, pp. 165–177.
18. Li, S.; Zheng, W.; Zong, Y.; Lu, C.; Tang, C.; Jiang, X.; Liu, J.; Xia, W. Bi-modality Fusion for Emotion Recognition in the Wild. In Proceedings of the 2019 International Conference on Multimodal Interaction Icmi’19, Suzhou, China, 14–18 October 2019; Assoc Computing Machinery: New York, NY, USA, 2019; pp. 589–594.
19. Li, F.; Wang, X.W.; Lu, B.L. Detection of Driving Fatigue Based on Grip Force on Steering Wheel with Wavelet Transformation and Support Vector Machine. In *ICONIP 2013: Neural Information Processing; Lecture Notes in Computer Science*; Springer: Berlin/Heidelberg, Germany, 2013; Volume 8228.
20. Zhang, L.; Yang, D.; Ni, H.; Yu, T. Driver Fatigue Detection Based on SVM and Steering Wheel Angle Characteristics. In Proceedings of the 19th Asia Pacific Automotive Engineering Conference & SAE-China Congress 2017: Selected Papers, Shanghai, China, 24–26 October 2017; Lecture Notes in Electrical Engineering. Springer: Singapore, 2017; Volume 486, pp. 729–738.
21. Lin, C.T.; Chen, Y.C.; Huang, T.Y.; Chiu, T.T.; Ko, L.W.; Liang, S.F.; Hsieh, H.Y.; Hsu, S.H.; Duann, J.R. Development of Wireless Brain Computer Interface with Embedded Multitask Scheduling and its Application on Real-time Driver’s Drowsiness Detection and Warning. *IEEE Trans. Bio-Med. Eng.* **2008**, *55*, 1582–1591. [\[CrossRef\]](#)
22. Zhu, T.; Zhang, C.; Wu, T.; Ouyang, Z.; Li, H.; Na, X.; Liang, J.; Li, W. Research on a Real-Time Driver Fatigue Detection Algorithm Based on Facial Video Sequences. *Appl. Sci.* **2022**, *12*, 2224. [\[CrossRef\]](#)
23. He, H.; Zhang, X.; Jiang, F.; Wang, C.; Yang, Y.; Liu, W.; Peng, J. A Real-time Driver Fatigue Detection Method Based on Two-Stage Convolutional Neural Network. *IFAC-PapersOnLine* **2020**, *53*, 15374–15379. [\[CrossRef\]](#)
24. Fang, B.; Xu, S.; Feng, X. A Fatigue Driving Detection Method Based on Multi Facial Features Fusion. In Proceedings of the 2019 11th International Conference on Measuring Technology and Mechatronics Automation (ICMTMA), Qiqihar, China, 28–29 April 2019; pp. 225–229.
25. Li, K.; Gong, Y.; Ren, Z. A Fatigue Driving Detection Algorithm Based on Facial Multi-Feature Fusion. *IEEE Access* **2020**, *8*, 101244–101259. [\[CrossRef\]](#)
26. Jenke, R.; Peer, A.; Buss, M. Feature Extraction and Selection for Emotion Recognition from EEG. *IEEE Trans. Affect. Comput.* **2014**, *5*, 327–339. [\[CrossRef\]](#)
27. Perdiz, J.; Pires, G.; Nunes, U.J. Emotional State Detection Based on EMG and EOG Biosignals: A Short Survey. In Proceedings of the 2017 IEEE 5th Portuguese Meeting on Bioengineering (Enbeng), Coimbra, Portugal, 16–18 February 2017.
28. Panda, R.; Malheiro, R.; Paiva, R.P. Novel Audio Features for Music Emotion Recognition. *IEEE Trans. Affect. Comput.* **2020**, *11*, 614–626. [\[CrossRef\]](#)
29. Kun, H.; Yu, D.; Tashev, I. Speech emotion recognition using deep neural network and extreme learning machine. In Proceedings of the Fifteenth Annual Conference of the International Speech Communication Association, Singapore, 14–18 September 2014.
30. Kansizoglou, I.; Misirlis, E.; Tsintotas, K.; Gasteratos, A. Continuous Emotion Recognition for Long-Term Behavior Modeling through Recurrent Neural Networks. *Technologies* **2022**, *10*, 59. [\[CrossRef\]](#)

31. Xu, L.; Ren, X.; Chen, R. Fatigue driving detection based on eye state recognition. *Sci. Technol. Eng.* **2020**, *20*, 8292–8299.
32. Shang, L.; Shi, Q.; Fang, J. Eye detection and fatigue judgment based on OpenCV. *Electron. World* **2018**, *23*, 19–20.
33. Sun, W.; Zhang, X.; Wang, J.; He, J.; Peeta, S. Blink number forecasting based on improved bayesian fusion algorithm for fatigue driving detection. *Math. Probl. Eng.* **2015**, *1*, 832621. [[CrossRef](#)]
34. Ekman, P.; Friesen, W.V. Facial Action Coding System(FACS): A technique for the measurement of facial actions. *Riv. Di Psichiatr.* **1978**, *47*, 126–138.
35. Abtahi, S.; Omidyeganeh, M.; Shirmohammadi, S.; Hariri, B. YawDD: A yawning detection dataset. In Proceedings of the 5th ACM Multimedia Systems Conference, Singapore, 19–21 March 2014; ACM: New York, NY, USA; pp. 24–28.

**Disclaimer/Publisher’s Note:** The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.