Computer Fraud and Security
ISSN (online): 1873-7056

Driver's Personal Emotion Recognition for Intelligent Cockpit of New Energy Vehicles

Yuhong Chen, Peihang Shi*

School of art and design, Anyang Institute of Technology, Anyang, Henan, China

**Corresponding Author*

Abstract:

New energy vehicles pay attention to the development of intelligent driving technology, which requires high human-computer interaction technology. Therefore, this paper studies the emotion recognition of intelligent cockpit of new energy vehicles, so as to promote the emotion regulation of intelligent driving on drivers and assist drivers to complete driving behavior more efficiently and safely. In order to ensure that the angry driving data can be organically combined with the overall architecture of intelligent vehicles, the motion preview model based on integrated direction and speed control can make real-time decisions according to the vehicle motion state and current traffic information, and feedback the decision information to the vehicle control module. Based on this model, the driver's anger characteristics are considered. In addition, this paper proposes a multimodal driver emotion recognition model MDERNet based on facial expression and driving behavior. It filters and highlights the data of driving behavior modes through the temporal attention obtained from facial expression modes, so as to realize information fusion among multiple modes at the input information level. Finally, through the results of experimental research and analysis, we can see that the driver's personal emotion recognition proposed in this paper has a good effect.

Key words: new energy vehicles; intelligent cockpit; driver; emotion recognition

1 INTRODUCTION

Emotion is a person's subjective feeling and experience, and it is an adaptive means and psychological ability to drive individuals to make certain behaviors. At present, studies have shown that both negative and positive emotions will have an important impact on driving behavior. Moreover, many scholars have found that bad emotions such as anger and disgust will make drivers have more risky driving behaviors, and drivers will have inattention, easy to make mistakes in judgment and reduce emergency response ability, which is prone to traffic accidents. Emotional stress refers to the emotions that cause certain pressure encountered by drivers during driving, and most of them refer to bad emotions, such as anger, disgust, tension and irritability, etc. These negative emotions will cause psychological pressure of drivers, change drivers' cognition of traffic conditions, reduce information processing ability and produce dangerous driving behaviors.

Emotions are mostly generated under the subconscious control of individuals, and within a certain range, they are transferred by human will. Once they exceed the control range, they will lead to excessive behavior, which is subjective and mutable. Drivers may experience various emotions when encountering different driving behaviors, such as anger when encountering forced lane occupation and slow driving in overtaking lanes, irritability when encountering traffic jams during morning rush hours, sudden pedestrians while driving, and fear when encountering reverse driving[1].

Emotions can have a certain impact on people's cognition, judgment, and behavior. People react more quickly in a positive emotional state than in a negative one, and their judgments of things are more accurate. Risk taking driving behavior is prone to occur in negative emotions. There are three forms of emotional expression, namely physiological, linguistic, and behavioral. Most scholars measure emotions by observing the physiological signals of drivers through experimental methods. The physiological and psychological conditions of drivers can interact with each other, and changes in the driver's psychological signals can drive changes in the physiological signals. The driver's physiological signals can recognize corresponding emotions, including heartbeat, heart rate variability, skin and body temperature, etc. When the driver experiences angry psychological emotions, it will physiologically manifest as facial redness, increased heartbeat and breathing, sweating, increased breathing volume, and trembling[2]. At the same time, drivers may experience errors in judgment, lack of concentration, narrower vision, slower reaction speed, and decreased motor agility. In terms of language and driving behavior,

scholars generally use survey questionnaires to quantitatively study emotions. Emotions in language generally manifest as insulting others, such as using uncivilized language while driving to express one's dissatisfactionn[3]. Emotions in driving behavior are generally manifested in motor vehicle operation, such as frantically tapping on the car's interior equipment, forcefully stepping on the brake accelerator, etc. Driving behaviors that vent outward include flashing high beams towards the opposite vehicle, seizing parking space, frequently honking the horn, intentionally shortening the distance from the preceding vehicle, and braking hard. Research in reference [4] shows that drivers increase their steering angle values under various emotions. In angry and happy emotions, lane departure becomes more pronounced, making them more prone to dangerous driving behaviors. In sad and fearful emotions, the range of lane departure is smaller, tending towards safer driving behaviors.

Each emotion is a response to a specific event, such as feeling sad, angry, or happy about something. Emotions depend on the evaluation of external events and their meanings, and emotional states include four aspects: cognition, sensation, physiological changes, and behavior. The emotional pressure of drivers is dynamic and will be adjusted according to their own physiological, psychological state, and external environment[5]. Emotional stress can affect the driver's information processing and judgment abilities, ultimately affecting their decision-making choices. It is crucial to study the impact of driver emotional stress on driving behavior. Early scholars divided the emotions of drivers into two categories: one is the psychological factors of drivers, and the other is the environmental factors of driving on the road. The psychological factors of drivers include age, personality, lifestyle habits, etc. The road environment includes factors such as vehicle environment, natural environment, and cultural environment. Based on the research of previous scholars, the emotional stress of drivers can be divided into two categories: physiological factors and driving environment factors. Changes in age, environment, and lifestyle habits can all alter the driver's psychological state, leading to emotional stress and behavioral changes as emotions gradually accumulate[6]. If in a sick state, mild general illnesses have a relatively small impact on the driver's emotions, while serious illnesses directly affect driving behavior, and the driver's emotional control ability will correspondingly decrease. For example, fatigue, the degradation of physiological functions caused by prolonged behavior, can reduce a person's reaction ability and affect coordination ability. Fatigue has always been widely studied [7]. Mild fatigue can cause frequent blinking, heavy eyelids, and slow gear shifting. In severe cases, it can lead to short-term blurring of consciousness, narrowing of sight range, and unconsciously accelerating heart rate. Fatigue can lower the driver's mood and affect driving safety. According to scholars, 20% of traffic accidents are caused by driver fatigue during driving. The external environment can be divided into three categories: natural environment, cultural environment, and traffic environment. The natural environment includes weather and road conditions. Adverse weather can cause fluctuations in the driver's emotions, resulting in emotional stress. Adverse weather conditions such as rain and snow can cause drivers to experience fear or tension, as rain and snow can cause road slippage and affect driver safety[8]. The road conditions, such as the inclination of the slope and the smoothness of the road surface, can affect the driver's emotions. The traffic environment includes traffic flow, roadblocks, traffic congestion, etc. For example, when the driver encounters road congestion, they may experience a restless emotional state. When there is no police officer at an intersection or they call for a stop, they may feel nervous. The humanistic environment is the main factor that causes drivers to develop emotions. The humanistic environment refers to factors such as cognition and quality, mainly divided into the interior environment and the driving behavior of other drivers. The bad driving behavior of other drivers is most likely to cause emotional fluctuations in drivers. Incorrect use of high and low beam headlights, malicious road occupation, and malicious overtaking can all cause negative emotions among drivers[9].

The study in reference [10] shows that the higher the level of positive emotions, the more likely it is to generate aggressive driving behaviors, and female drivers are more likely to develop negative emotions than male drivers. Reference [11] found that positive emotions lead to an increase in lane control speed, while negative emotions lead to a decrease in lane control speed. Reference [12] collected emotional EEG information and driving behavior information through simulated driving experiments, constructed a model of the relationship between driver emotions and driving behavior, and pointed out that the change rate of brake pedals, the mean standard deviation of vehicle speed, throttle opening, and vehicle acceleration under different emotional states all formed specific trends, All have the highest positive emotions, followed by normal emotions, and the lowest negative

emotions.

Reference [13] used images as stimuli to induce four types of emotions in subjects, and identified emotions through EEG data features with an accuracy rate of 84.5%. Reference [14] collected emotional EEG brainwaves and surrounding physiological signals from volunteers through audiovisual stimulation using OpenBCI, and achieved good results in predicting two-dimensional emotional states through standardized self-assessment, genre labeling, and unsupervised machine learning methods. Reference [15] used music to induce emotions measured on arousal/arousal dominant scales, as well as wavelet decomposition of EEG signals. A classifier was used to identify the feature parameters fused with three features, achieving good classification results. Reference [16] extracted multi-dimensional physiological signal feature data through genetic algorithm and ant colony algorithm particle swarm algorithm, and achieved certain results in identifying six emotions such as happiness, surprise, disgust, sadness, anger, and fear. Reference [17] used threshold method for peak detection to extract characteristic parameters of electrocardiogram and pulse wave, and then adopted deep learning algorithm to recognize positive and negative emotions, achieving good recognition rate. Reference [18] identifies positive, neutral, and negative emotions of drivers by collecting physiological signals such as driver heart rate, skin surface temperature, and driver skin potential

Reference [19] combines text and EEG information, extracts traditional statistical features, demographic features, and deep abstract features from multimodal information, integrates them into multimodal emotional features, and uses long short-term memory networks to recognize emotions. Reference [20] studied the impact of different traffic environments on driver emotions and further investigated their impact on driving safety. By designing simulated driving experiments, collecting driver's EEG signal data for filtering, removing wakes, and other methods to extract features, calculating emotional index through alpha wave power, and studying the relationship between emotions and vehicle data in driving behavior. Build the relationship between emotional index and driving behavior using k-nearest neighbor algorithm. Although there are multiple modalities for expressing emotions, visual modalities dominate research in the field of vehicle engineering. Reference [21] uses facial expression recognition to identify whether drivers exhibit road rage. By collecting a sequence of driver facial images over a period of time, and using facial monitoring algorithms to crop and align the driver's facial images, inputting them into a deep neural network for angry expression recognition, the neural network performs binary recognition of whether or not to be angry. When the number of angry images recognized exceeds a certain threshold, it is determined that the driver is in an angry state and may experience road rage

New energy vehicles pay attention to the development of intelligent driving technology, and have high requirements for human-computer interaction technology. Therefore, this paper studies the emotional recognition of intelligent cockpit of new energy vehicles, so as to promote the emotional regulation of intelligent driving on drivers and assist drivers to complete driving behavior more efficiently and safely.

2 DRIVER'S PERSONAL EMOTION RECOGNITION FOR INTELLIGENT COCKPIT

2.1 Overall architecture of smart vehicles

In order to ensure that the angry driving data can be organically combined with the overall architecture of smart cars, this section needs to analyze the overall architecture of smart cars proposed by our research group, and find a reasonable embedding point. The overall architecture of smart vehicles is shown in Figure 1.

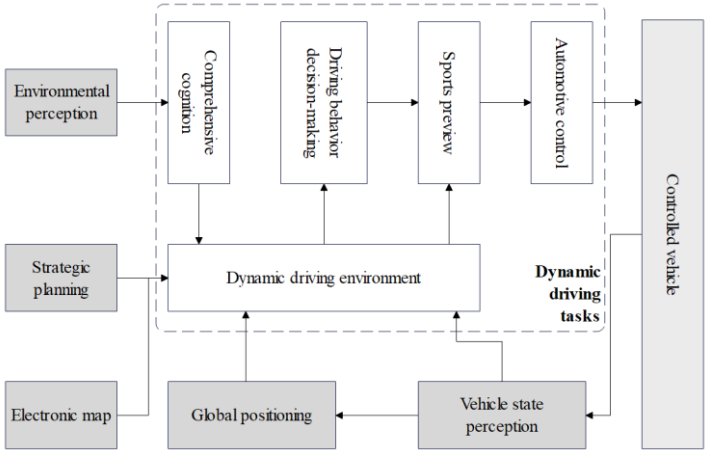


Figure 1 Overall architecture of smart vehicles

The motion preview model based on the integrated control of direction and speed can make real-time decisions according to the vehicle motion state and current traffic information, and feedback the decision information to the vehicle control module. Based on this model, the driver's anger characteristics are considered. Figure 2 shows the motion preview model.

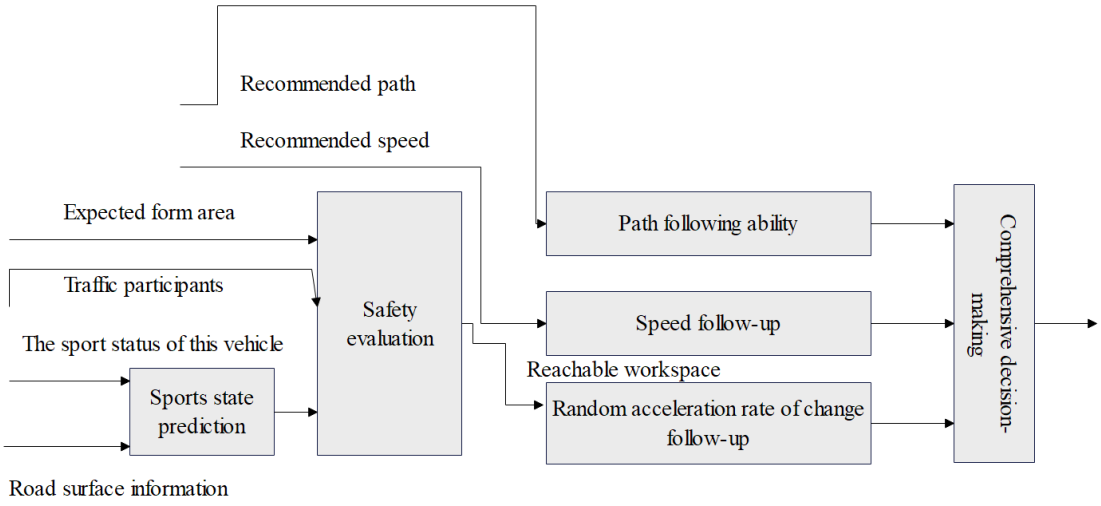


Figure 2 Motion preview model

The function of safety evaluation module is to detect collision with traffic participants according to the predicted motion state and accessible space. If it is determined that it is unsafe, the preview point is deleted, which is collision safety. At the same time, it is necessary to ensure the safety of the area, judge whether the car deviates from the expected driving area, and delete the preview point if it deviates. The input of preview acceleration synthesis decision is the recommended path, the speed random number generated according to the probability density function of driver's longitudinal speed and the acceleration change rate random number generated according to the probability density function of driver's longitudinal speed change rate.

In order to evaluate the decision-making optimality of preview acceleration, this section mainly adopts the following three evaluation indexes to evaluate: (1) path following; (2) speed following; (3) follow-up of random acceleration rate. Then, each single evaluation index is weighted and summed to obtain the value of horizontal and longitudinal preview acceleration that minimizes it, so as to realize the optimal sum evaluation of preview acceleration decision, as shown in the following equation (1):

$$grg\ min\left(w_{f,d}J_{f,d}+w_{f,\theta}J_{f,\theta}+w_{f,v}J_{f,v}+w_{h,lon}J_{h,lon}\right) \quad (1)$$

For normal drivers, it is necessary to ensure that the vehicle follows the recommended path as much as possible

during driving, and the path following index mainly describes the rational part of the driver. (1) Lateral distance following index: This index mainly measures the degree of vehicle lateral deviation from the recommended path. In this paper, the lateral distance following evaluation index is calculated by the lateral distance from the preview position of the vehicle to the recommended path at the time $t+T_p$. The lateral distance following index is calculated by the following equation (2):

$$J_{f,d} = \left(\frac{\left|d_{ep}\right| - min\left(\left|d_{ep}\right|\right)}{max\left(\left|d_{ep}\right|\right) - min\left(\left|d_{ep}\right|\right)} \right) \tag{2}$$

(2) Direction following index: This paper calculates the direction following index by the angle between the direction of the preview point and the recommended path at the time $t+T_p$. The direction following index can be obtained by the following equation (3)[22]:

$$J_{f,\theta} = \left(\frac{\theta - min(\theta)}{max(\theta) - min(\theta)} \right)^2 \tag{3}$$

In equation, $J_{f,\theta}$ is the direction following index and θ is the angle between the direction of the preview point and the recommended path.

When a normal driver drives a car, the speed change will be based on the recommended speed. By establishing the following evaluation index of recommended speed, the motion preview model can follow the recommended speed. The evaluation index of random speed following is calculated by the difference between preview longitudinal speed and random speed. It can be calculated by the following equation (4):

$$J_{f,v} = \begin{cases} \left(\frac{u_p - u_{ep}}{u_{p,max} - u_{ep}} \right) & u_p \geq u_{ep} \\ \left(\frac{u_p - u_{ep}}{u_{p,min} - u_{ep}} \right) & u_p < u_{ep} \end{cases} \tag{4}$$

When a normal driver drives, the driver's operation is in line with his usual habits and the driving rules of ordinary drivers. This paper holds that the normal driver will accelerate and decelerate when driving the car, but the rate of change of acceleration is uncertain, that is, there is some randomness. Therefore, the acceleration change rate is generated randomly according to the probability density function of the acceleration change rate, and then the preview acceleration change rate is approached to the randomly generated acceleration change rate according to the evaluation index, so that the driver's choice of acceleration change rate is random.

The preview longitudinal acceleration rate of change is shown in equation (5):

$$Jerk_{x,p} = \frac{a_{x,p} - a_x}{T} \tag{5}$$

$a_{x,p}$ is the preview longitudinal acceleration, a_x is the current vehicle longitudinal acceleration, and T is the decision time.

$$G_x = \left| Jerk_{x,p} - Jerk_{x,r} \right| \tag{6}$$

The following evaluation index of longitudinal random acceleration is:

$$J_{h,lon} = \frac{G_x - min(G_x)}{max(G_x) - min(G_x)} \tag{7}$$

2.2 Inducing methods of anger emotion

The premise of research on anger inducing factors is how to quantify the subjective phenomenon of anger through

psychological evaluation theory, so as to achieve a relatively accurate and objective evaluation of the current emotional state of the evaluation object. With the rapid development of human-computer interaction in recent years, human emotion computing has gradually become an important research field and direction. Emotion quantification tools such as Hamilton Depression Scale (SAS) and Self-rating Depression Scale (SDS) are widely used. Therefore, this topic chooses the PAD emotion scale corresponding to the PAD emotion model to evaluate the anger of the subjects.

PAD emotion scale is derived from PAD three-dimensional emotion model, in which P, A and D represent pleasure, arousal and dominance respectively. The three dimensions of P, A and D are positive and negative, and the three-dimensional space is divided into eight modules, as shown in Figure 3, which are: + P + A + D happy; -P-A-D boring; + P + A-D dependent; -P-A + D contempt; + P-A + D relaxed; -P + A-D anxiety; + P-A-D docile; -P + A + D hostile. The above eight modules are the general classification of eight kinds of emotions, and the more accurate classification of each kind of emotions can be measured by specific spatial coordinates.

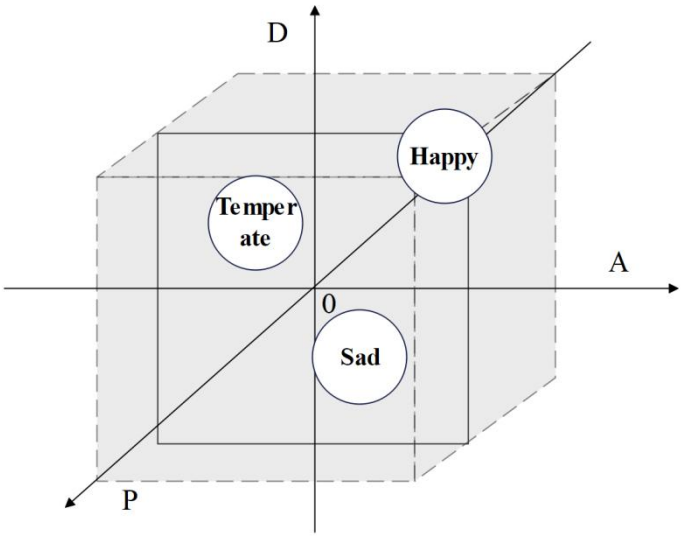


Figure 3 PAD three-dimensional emotion model

PAD emotional scale is a 9-point semantic difference scale, that is, according to the subjective feelings of the tested object, taking 0 as the benchmark, there are 4 evaluation grades indicating the corresponding degree of the current item in the positive and negative directions. Among them, 0 point is neutral degree, with a total of 9 difference evaluation points. The simplified PAD scale is shown in Table 1. Twelve evaluation items finally correspond to PAD three-dimensional values through the following formula:

$$P = \left(V_1 - V_4 + V_7 - V_{10}\right) / 4 \tag{8}$$

$$A = \left(-V_2 + V_5 - V_8 + V_{11}\right) / 4 \tag{9}$$

$$D = \left(V_3 - V_6 + V_9 - V_{12}\right) / 4 \tag{10}$$

Table 1 Simplified PAD emotional scale

-4 -3 -2 -1 0 1 2 3 4	
Angry	Interested
Awake	Sleepy
Controlled	Master controlled
Friendly	Contemptuous
Calm	Excited

Dominant	Submissive
Painful	Happy
Concerned	Relaxed
Guided	Autonomous
Excited	Irritated
Reserved	Surprised
Influential	Affected

2.3 Driver emotion recognition model

A multimodal driver emotion recognition model MDERNet based on facial expression and driving behavior is proposed. MDERNet uses two branches to extract the features of facial expression and driving behavior respectively, and the two branches correspond to different modes respectively. The facial expression modes generate temporal attention, and screen and highlight the input data of driving behavior modes, so as to refine the input features and better realize the fusion of features between the two modes.

MDERNet model is a dual-stream network that uses sparse representation and attention mechanism to recognize discrete emotion and dimensional emotion. The first stage of the network structure is preprocessing operation, which mainly includes preprocessing of facial expression data and driving behavior data. Then, the preprocessed data are processed by two branches of the model respectively, and finally the extracted multimodal features are fused to realize discrete emotion and dimensional emotion recognition of drivers. The framework of driver emotion recognition model based on facial expression and driving behavior is shown in Figure 4.

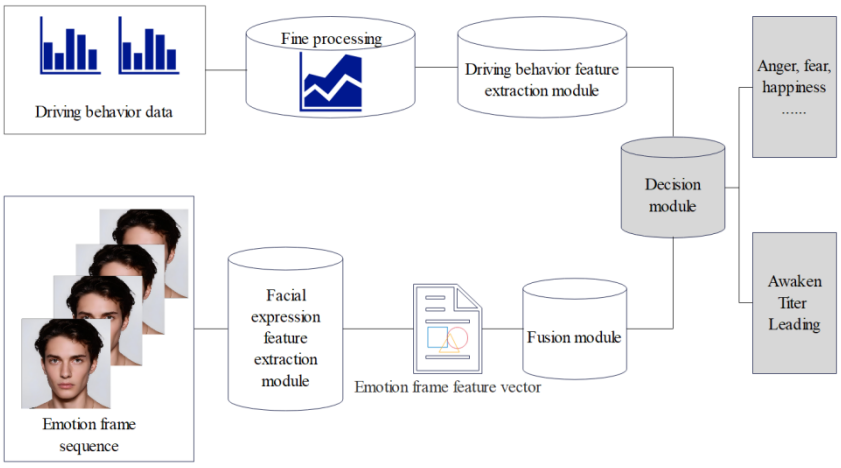


Figure 4 Architecture of driver emotion recognition model based on facial expression and driving behavior

The data preprocessing of driver's face image mainly includes face key point detection, face alignment, face image geometry normalization, brightness normalization and so on. The preprocessing process is shown in Figure 5.

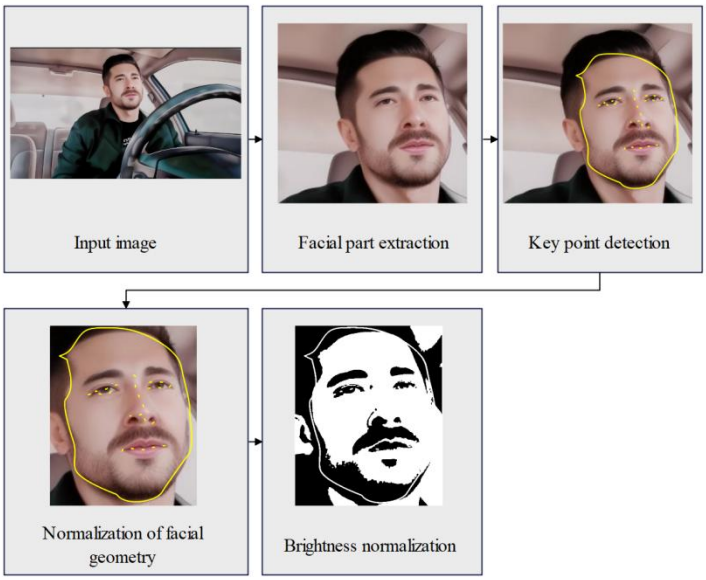


Figure 5 Operation flow of facial expression preprocessing

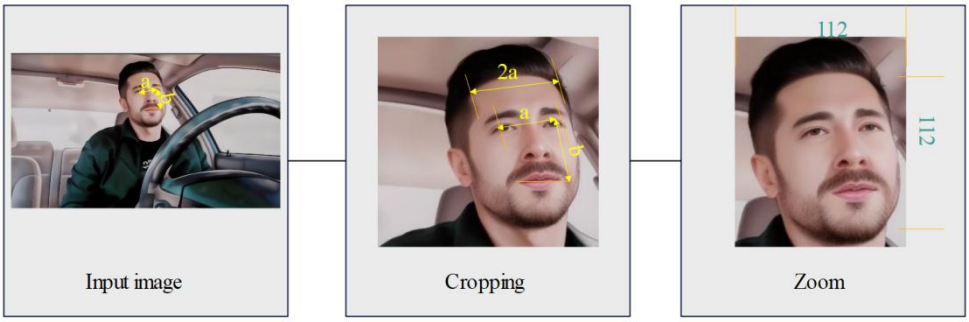


Figure 6 Face geometry normalization

Firstly, MTCNN (Multi-task Cascade Convolution Network) is used to identify face images, and 68 key points of faces are obtained. Then, the angle between the connection line between left eye center and right eye center and horizontal line is calculated, and the image is rotated to make the connection line between left eye center and right eye center horizontal to realize face alignment. After that, the distance between the center of the left eye and the center of the right eye is set as a, and the vertical difference between the center point of the center of the left eye and the center of the right eye and the center point of the mouth is set as b. The driver's face image is cut to 2a wide and 2b high, and then scaled to an image of 112 × 112 pixels, as shown in Figure 6. Scale normalization enables the same facial feature points to coexist roughly in the same position in different video frames. At the same time, because these areas do not represent specific information of facial expressions, this process discards background details and facial areas such as ears and forehead unrelated to facial expressions. In order to reduce the change of image signal caused by the change of illumination, the brightness of the cropped face image is normalized. For all experimental data sets, 30 frames of images are sampled from each video as training samples of the model.

The selected driving behavior data includes steering wheel angle, accelerator pedal opening and closing degree, brake pedal force, speed, acceleration, lateral speed and lateral acceleration. The selected input data is normalized as shown in equation (11).

$$x' = \frac{x - \mu}{\sigma} \tag{11}$$

The facial expression branch of MDERNet model mainly includes a facial expression feature extraction module FEFEM, two fusion modules FM and a frame attention module FAM.

The input of FEFEM module is a face gray image I_{fefem} with 112 × 112 resolution, and the output is a

512-dimensional 1 × 1 feature image M_{fefem}. The process is as follows:

$$M_{fefem} = f_{fefem}\left(I_{fefem}\right) \tag{12}$$

Among them, f_{fefem} is a function of FEFEM module, and FEFEM outputs corresponding feature vectors for each frame of face image, and all frame feature vectors belonging to the same video are input into FM together as a part of FAM input.

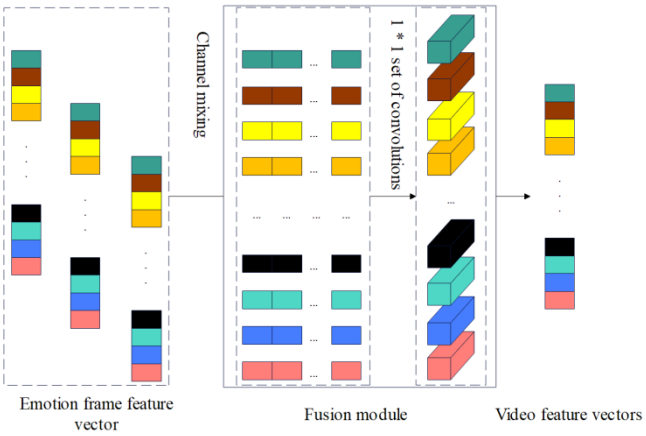


Figure 7 Fusion module

The FM module is shown in Figure 7. The input of FM is a set of feature vectors output by all frame images of the same video through the FEFEM module. The set of face feature vectors is arranged channel by channel according to time series, so that all frame feature maps of each channel can be conveniently convolved in 1 × 1 groups, and video feature maps are formed according to channels. The video feature maps of all channels together constitute the overall video features. The feature vectors extracted by FM are input into the FAM together with the frame feature vectors output by FEFEM as the whole video feature vectors. The process is as follows:

$$M_{fm} = f_{fm}\left(M_{fefem0},\ldots,M_{fefemk\text{-}1}\right) \tag{13}$$

Among them, f_{fm} is a function of FM, including channel mixing and group convolution operations, and k is the number of sampled frames of video.

As shown in Figure 8, the FAM module is similar to the SE module, including (1) two full connection layers and (2) the Sigmoid function. The implementation transformation of FAM is as follows:

$$\begin{aligned} W_{fam_i} &= f_{fam}\left(Concat\left(M_{fefem_i}, M_{fm}\right)\right) \\ M_{fam_i} &= W_{fam_i} \times M_{fefem_i}, i \in [0, K-1] \end{aligned} \tag{14}$$

Wherein, f_{fam} is the module of FAM, M_{fefem_i} is the feature map obtained by FEFEM for the i-th frame image, M_{fm} is the preliminary video feature map of the video to which the frame image belongs, and the generated W_{fam_i} is expressed as the weight value the i-th frame image, and M_{fam_i} is the feature map of the weighted i-th frame image.

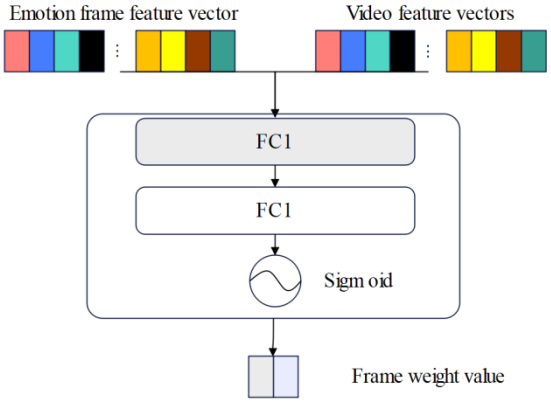


Figure 8 Frame attention Module

Driving behavior data refinement is to screen and highlight the data of driving behavior modes through the temporal attention obtained from facial expression modes, so as to realize information fusion among multiple modes at the input information level. Using W_{fam_i} to screen and highlight driving behavior data mainly includes two steps. First, W_{fam_i} and M_{fam_i} are up-sampled so that its numerical length is consistent with the length of driving behavior data Idb. Then, W_{fam_i} is binarized by setting a threshold value and the binarized W_{fam_i} is multiplied one by one by the driving behavior data. The implementation process is as follows.

$$\begin{aligned} W_{fam}^{binary} &= Binary\left(Upsample\left(W_{fam}\right)\right) \\ I_{db}^{refined} &= W_{fam}^{binary} \times I_{db} \end{aligned} \tag{15}$$

Among them, $Binary$ and $Upsample$ represents the binarization operation and the up-sampling operation, W_{fam}^{binary} is the attention weight sequence of the time sequence frame after up-sampling and binarization, and $I_{db}^{refined}$ is the refined driving behavior data.

The decision module DM is a fully connected layer, and Softmax function is added to discrete emotion classification. As shown in Figure 9, facial expression features are the output of weighted frame *image features M fami* after passing through another FM module, and the process is as follows:

$$\hat{y} = f_{dm}\left(Concat\left(M_{defem}, f_{fm}\left(M_{fam_0},...,M_{fam_{k-1}}\right)\right)\right) \tag{16}$$

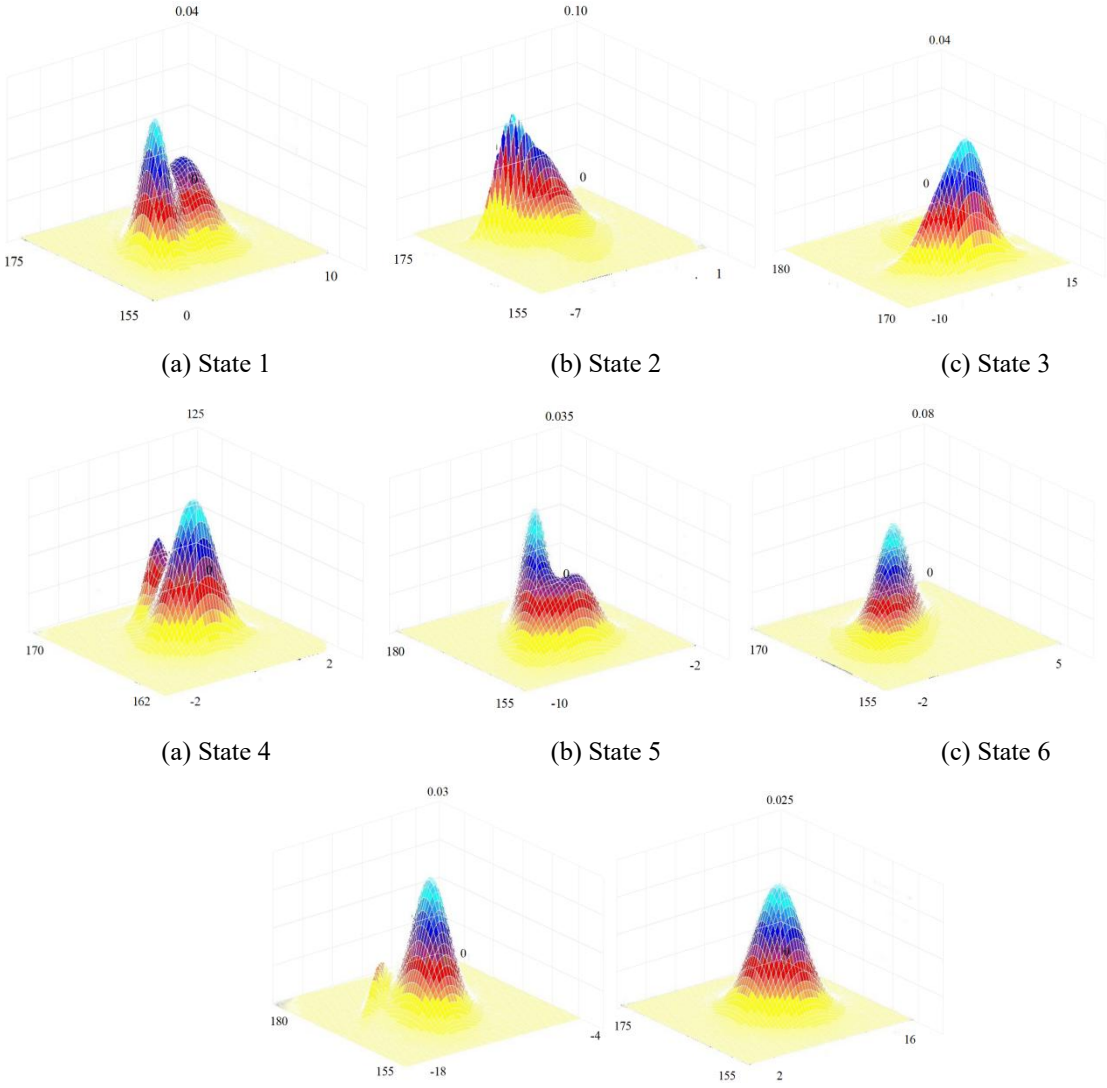
Among them, f_{dm} is the function of DM and $\hat{y}$ is the prediction result of the model.

3 EXPERIMENTAL STUDY

The experimental environment is Intel (R) Core (TM) i7-9700KCPU @ 3.60 G processor, and the running memory is 32G, and the GPU is NVIDIA GeForce RTX 2080 × 2 GPU. The implementation of the experiment is based on Tensorflow deep learning platform, which is an open source software library based on data flow diagram for numerical calculation and widely used in various machine learning algorithms.

Because there are many kinds of drivers' emotions, this paper first chooses anger as the main factor to verify the effect of this model, and anger is also the most common emotional state during driving.

According to the existing data accumulated in the laboratory, when the number of states is selected at 8, the relatively satisfactory model training effect can be obtained under the condition of less calculation workload. Therefore, the corresponding parameter is x= (y, z), the parameter is an 8 × 1 matrix, the parameter A is an 8 × 8 state transition probability matrix, and the parameter B represents the probability density distribution of the output observation under different states. Therefore, the parameter B has a corresponding probability density distribution function for each state. After training and learning, the driving model data of anger style is shown in Figure 9.



(a) State 1 (b) State 2 (c) State 3

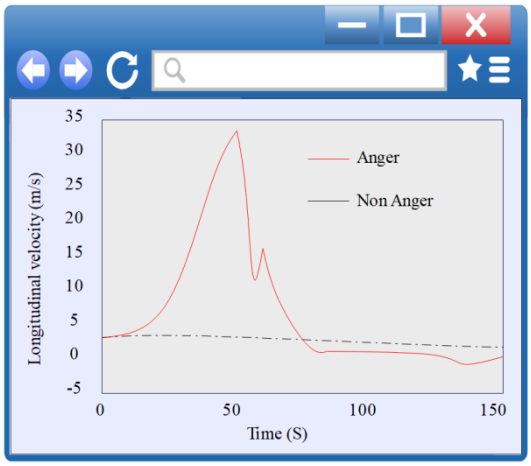
(a) State 4 (b) State 5 (c) State 6

(a) State 7 (b) State 8

Figure 9 Driving model data diagram of anger style

As shown in Figure 10 (a), through the longitudinal speed fluctuation of non-angry drivers and angry drivers under the same working conditions, it can be seen that the driving process of non-angry drivers is generally stable and there is no abnormal acceleration and deceleration action. During the driving process of angry drivers, violent driving actions such as rapid acceleration and rapid deceleration are frequent, which can express the irrational speed characteristics of angry drivers.

As shown in Figure 10 (b), through the change curves of longitudinal acceleration of non-angry and angry drivers under the same working conditions, it can be seen that during the driving process of non-angry drivers, the longitudinal acceleration generally remains stable, and the fluctuation range is extremely small. During the driving process of angry drivers, the fluctuation of longitudinal acceleration is irregular, and the change rate of longitudinal acceleration also changes greatly, which shows that the model can reflect the randomness of the acceleration and the change rate of acceleration of angry drivers, and thus express the irrational characteristics of angry drivers in the driving process.



(a) Comparison of longitudinal speed between non-angry and angry drivers

(b) Comparison of longitudinal acceleration between non-angry and angry drivers

Figure 10 Acceleration comparison in angry state

On this basis, the driver's personal emotion recognition effect in this paper is verified. On the driver's facial expression data set, the expression recognition rate before and after data enhancement is shown in Figure 11:

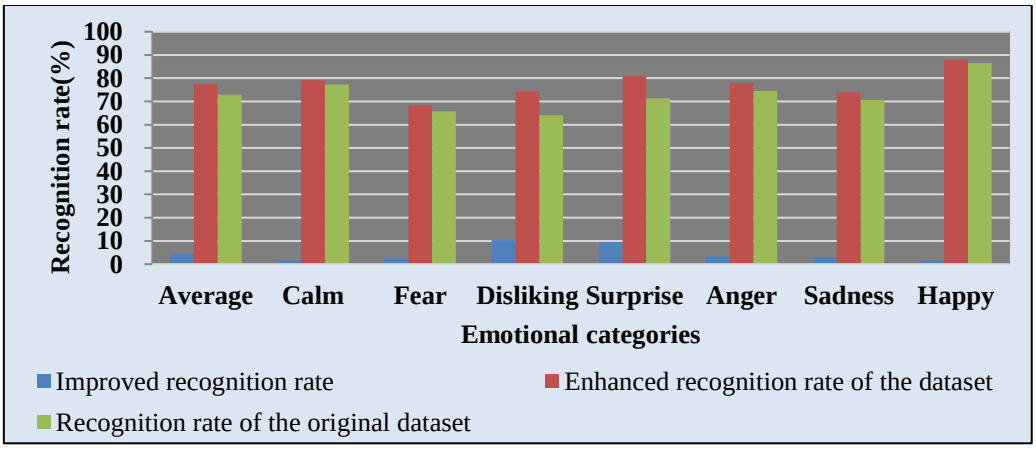


Figure 11 Expression recognition rate on driver facial expression data set

After processing the original driver facial expression data set, not only the facial image recognition rate of disgust and surprise expression categories is improved by nearly 10%, but also the facial expression recognition rate of other categories of images is improved to a certain extent. The reason is that with the increase of the number of images trained in the network, the differences between different types of expressions also increase, and the more

features obtained in the training, so the average recognition rate is higher. Therefore, it is verified that the model proposed in this paper has a good effect on driver's personal emotion recognition.

4 CONCLUSION

Emotion is a reflection of people's comprehensive state, and it also affects drivers' driving behavior. With the widespread use of cars in life and the rapid development of smart cars, people are increasingly dependent on cars when traveling, and the time people spend on cars will be longer. At home and abroad, the research on the relationship between driving behavior and emotional state is also deepening. New energy vehicles pay more attention to the development of intelligent driving technology and have higher requirements for human-computer interaction technology. Therefore, this paper studies emotion recognition in intelligent cockpit of new energy vehicles and proposes a multi-modal driver emotion recognition model MDERNet based on facial expression and driving behavior. MDERNet uses two branches to extract the features of facial expression and driving behavior respectively, and carries out fine processing of driving behavior data. Through the temporal attention obtained from facial expression modes, the data of driving behavior modes are screened and highlighted, so as to realize information fusion among multiple modes at the input information level. Finally, through the results of experimental research and analysis, we can see that the driver's personal emotion recognition proposed in this paper has a good effect.

REFERENCES

[1] Zepf, S., Hernandez, J., Schmitt, A., Minker, W., & Picard, R. W. (2020). Driver emotion recognition for intelligent vehicles: A survey. ACM Computing Surveys (CSUR), 53(3), 1-30.

[2] Du, G., Wang, Z., Gao, B., Mumtaz, S., Abualnaja, K. M., & Du, C. (2020). A convolution bidirectional long short-term memory neural network for driver emotion recognition. IEEE Transactions on Intelligent Transportation Systems, 22(7), 4570-4578.

[3] Li, W., Zeng, G., Zhang, J., Xu, Y., **ng, Y., Zhou, R., ... & Wang, F. Y. (2021). Cogemonet: A cognitive-feature-augmented driver emotion recognition model for smart cockpit. IEEE Transactions on Computational Social Systems, 9(3), 667-678.

[4] Naqvi, R. A., Arsalan, M., Rehman, A., Rehman, A. U., Loh, W. K., & Paul, A. (2020). Deep learning-based drivers emotion classification system in time series data for remote applications. Remote Sensing, 12(3), 587-601.

[5] Cordero, J., Aguilar, J., Aguilar, K., Chávez, D., & Puerto, E. (2020). Recognition of the driving style in vehicle drivers. Sensors, 20(9), 2597-2607.

[6] Pal, S., Mukhopadhyay, S., & Suryadevara, N. (2021). Development and progress in sensors and technologies for human emotion recognition. Sensors, 21(16), 5554-5570.

[7] Koduru, A., Valiveti, H. B., & Budati, A. K. (2020). Feature extraction algorithms to improve the speech emotion recognition rate. International Journal of Speech Technology, 23(1), 45-55.

[8] Park, C. Y., Cha, N., Kang, S., Kim, A., Khandoker, A. H., Hadjileontiadis, L., ... & Lee, U. (2020). K-EmoCon, a multimodal sensor dataset for continuous emotion recognition in naturalistic conversations. Scientific Data, 7(1), 293-303.

[9] Zhang, Z. (2021). Speech feature selection and emotion recognition based on weighted binary cuckoo search. Alexandria Engineering Journal, 60(1), 1499-1507.

[10] Zhao, S., Jia, G., Yang, J., Ding, G., & Keutzer, K. (2021). Emotion recognition from multiple modalities: Fundamentals and methodologies. IEEE Signal Processing Magazine, 38(6), 59-73.

[11] Li, W., Zhang, B., Wang, P., Sun, C., Zeng, G., Tang, Q., ... & Cao, D. (2021). Visual-attribute-based emotion regulation of angry driving behaviors. IEEE Intelligent Transportation Systems Magazine, 14(3), 10-28.

[12] Kuruvayil, S., & Palaniswamy, S. (2022). Emotion recognition from facial images with simultaneous occlusion, pose and illumination variations using meta-learning. Journal of King Saud University-Computer and Information Sciences, 34(9), 7271-7282.

[13] Tan, L., Yu, K., Lin, L., Cheng, X., Srivastava, G., Lin, J. C. W., & Wei, W. (2021). Speech emotion recognition enhanced traffic efficiency solution for autonomous vehicles in a 5G-enabled space–air–ground integrated intelligent transportation system. IEEE Transactions on Intelligent Transportation Systems, 23(3), 2830-2842.

[14] Zheng, X., Yu, X., Yin, Y., Li, T., & Yan, X. (2021). Three-dimensional feature maps and convolutional neural network-based emotion recognition. International Journal of Intelligent Systems, 36(11), 6312-6336.

[15] Pandey, P., & Seeja, K. R. (2022). Subject independent emotion recognition from EEG using VMD and deep learning. Journal of King Saud University-Computer and Information Sciences, 34(5), 1730-1738.

[16] Houssein, E. H., Hammad, A., & Ali, A. A. (2022). Human emotion recognition from EEG-based brain–computer interface using machine learning: a comprehensive review. Neural Computing and Applications, 34(15), 12527-12557.

[17] Yadav, S. P., Zaidi, S., Mishra, A., & Yadav, V. (2022). Survey on machine learning in speech emotion recognition and vision systems using a recurrent neural network (RNN). Archives of Computational Methods in Engineering, 29(3), 1753-1770.

[18] Wu, D., Han, X., Yang, Z., & Wang, R. (2020). Exploiting transfer learning for emotion recognition under cloud-edge-client collaborations. IEEE Journal on Selected Areas in Communications, 39(2), 479-490.

[19] Liu, X., Cheng, X., & Lee, K. (2020). GA-SVM-based facial emotion recognition using facial geometric features. IEEE Sensors Journal, 21(10), 11532-11542.

[20] Gao, Z., Wang, X., Yang, Y., Li, Y., Ma, K., & Chen, G. (2020). A channel-fused dense convolutional network for EEG-based emotion recognition. IEEE Transactions on Cognitive and Developmental Systems, 13(4), 945-954.

[21] Muhammad, G., & Hossain, M. S. (2021). Emotion recognition for cognitive edge computing using deep learning. IEEE Internet of Things Journal, 8(23), 16894-16901.