

Received 26 August 2023, accepted 4 September 2023, date of publication 12 September 2023, date of current version 21 September 2023.

Digital Object Identifier 10.1109/ACCESS.2023.3314665

RESEARCH ARTICLE

Face Fatigue Feature Detection Based on Improved D-S Model in Complex Scenes

JIANFENG CAI¹, XU LIAO¹, JUNBO BAI², ZHEN LUO¹, LI LI¹, AND JUNJIE BAI¹

¹College of Electrical Engineering, Chongqing University of Science and Technology, Chongqing 401331, China

²Shanghai Ear Technology Company Ltd., Jiading, Shanghai 201800, China

Corresponding author: Junjie Bai (baijj@cqust.edu.cn)

This work was supported in part by the Chongqing Natural Science Foundation Project under Grant cstc2020jcyj-msxmX0452, and in part by the Industry-University-Research Innovation Fund of the Ministry of Education of China under Grant 2020HYA06001.

This work involved human subjects or animals in its research. Approval of all ethical and experimental procedures and protocols was granted by Chongqing University of Science and Technology Committee.

ABSTRACT Fatigued driving is one of the main causes of road traffic accidents. In the process of fatigued driving detection, the evaluation based on a single sign is biased. To improve the adaptability and accuracy of fatigued driving detection, this paper proposes an improved D-S evidence theory-based algorithm for detecting facial fatigue signs. This algorithm uses the multi-thread-optimized Dlib to track and locate the image of the face, captures the 68 key points of the driver's face with reference to the Dlib open-source library, and narrows the target areas to the eyes and mouth regions. Based on the video stream, it calculates the horizontal and vertical ratios of the eyes and mouth to determine the fatigue sign subsets based on the EAR and MAR within a unit cycle, and calculates the Pitch angle after converting the head pose from 2D images to 3D models, which is used for determining the status of the head pose. The algorithm then fuses multiple feature subsets and uses the improved D-S evidence theory to optimize the weights of the three subsets to mitigate the influence of the "general conflict" and "one veto" problems, and to increase the temporal correlation coefficient of a single fatigue sign. Experimental results show that the improved D-S face detection algorithm can effectively solve the problems of lighting and partial occlusion in complex environments, with an accuracy rate of 93.8% for detecting facial fatigue signs.

INDEX TERMS Fatigued driving, Dlib, improved D-S evidence theory, machine vision.

I. INTRODUCTION

With the significant increase in the number of vehicles worldwide, road traffic safety problems become more and more serious. Studies show that fatigued driving behavior is one of the main causes of road safety problems [1]. The rapid development of machine vision enables, the quick and effective detection of drivers' facial fatigue features. The facial fatigue signs are: eye closing, yawning and head posture deviation. Non-contact fatigued driving detection through multi-feature information fusion has become a research hotspot [2].

Xiaofeng [3] detected the driver's face according to the LBP features, obtained the facial sign points through the

multi-cascade residual regression tree, and optimized as well as accelerated the facial detection process by using background subtraction to narrow the detection area and reduce the size of the video frame image. Liu et al. [4] used Dlib library to capture facial fatigue signs, calculated EAR values corresponding to the eyes and MAR values corresponding to the mouth respectively, and then fused the signs of two parts with SVM with a accuracy rate of 84.29%. Wang et al. [5] used multi-threaded Dlib to detect, and fused the data of, blink frequency, eye closing frequency, percentage of eye closing time and yawn frequency, with an accuracy of 91.5%. But this method underperformed in complex environments. Bo et al. [6] used D-S to carry out decision-level fusion of several fatigue discrimination indexes such as PERCLOS, longest eye closing time and percentage of stasis period in

The associate editor coordinating the review of this manuscript and approving it for publication was Zahid Akhtar¹.

steering wheel, and the detection accuracy reached 91%, higher than those using a single index. Zhao et al. [7] fused six fatigue sign subsets, such as blink, yawn, head posture and corresponding time and frequency, and used weighted feature subsets to determine the fatigue level.

At present, the main problems in face fatigue sign detection are as follows: the determination depends on a single fatigue feature; face feature extraction is affected by lighting conditions and partial occlusion; and the weight coefficient of facial fatigue sign fusion process is not balanced [8]. Therefore, an improved D-S evidence theory fusion algorithm is proposed in this paper for the fusion and determination of multi-sign subsets of faces [9]. This algorithm integrates Opencv and Dlib for real-time detection of yawning, blinking and head migration, and adopts improved D-S for fusion decision, and determines fatigue level, which effectively improves its adaptive ability and robustness [10].

II. SYSTEM OVERALL RESEARCH PROGRAM

Three regions for fatigue sign detection, namely, the driver’s eyes, mouth and head, are selected, and the corresponding threshold and feature values of these regions defined. The location of the key points of the face and the division of the corresponding facial regions are realized through the modeling of the key points of the face [11]. On this basis, the fatigue characteristic values of eyes, mouth and head regions were calculated respectively. The working flow chart of the facial fatigue signs extraction model is shown in Figure 1. This process mainly bases on the key point model of the face and various fatigue sign extraction schemes [12] to realize key point positioning and fatigue sign extraction in the driver’s face image. Specific steps: 1. Upon input of the face image, the algorithm extracts the face region and converts it into gray-scale image; 2. detects face key points; 3. confirms whether the face detection is successful; 4. extracts fatigue signs in the eye region and compare the reference values for determination; 5. extracts fatigue signs in the mouth region to determine whether the characteristic value of the mouth region exceeds the threshold; 6. extracts head pose signs to determine whether the value exceeds the threshold; 7. three face feature sequences that meet the discriminant conditions are stored and their feature values are calculated respectively; 8. fuses three subsets using the improved D-S; and 9. outputs sign data and the detection loops [13].

III. FACIAL FATIGUE FEATURE DETECTION

A. DLIB LIBRARY TRACKS KEY POINTS OF FACES

The face detection process involves collecting video footage using the front camera, which is then divided into segments based on time series [14], [15], [16], [17], [18], [19]. The pre-processed image is used to extract facial feature points in each frame using Dlib and OpenCV, allowing for external features of the face contour to be retrieved and analyzed. Based on the main features of the Dlib face recognition library, blink, mouth, and head posture are calculated, and

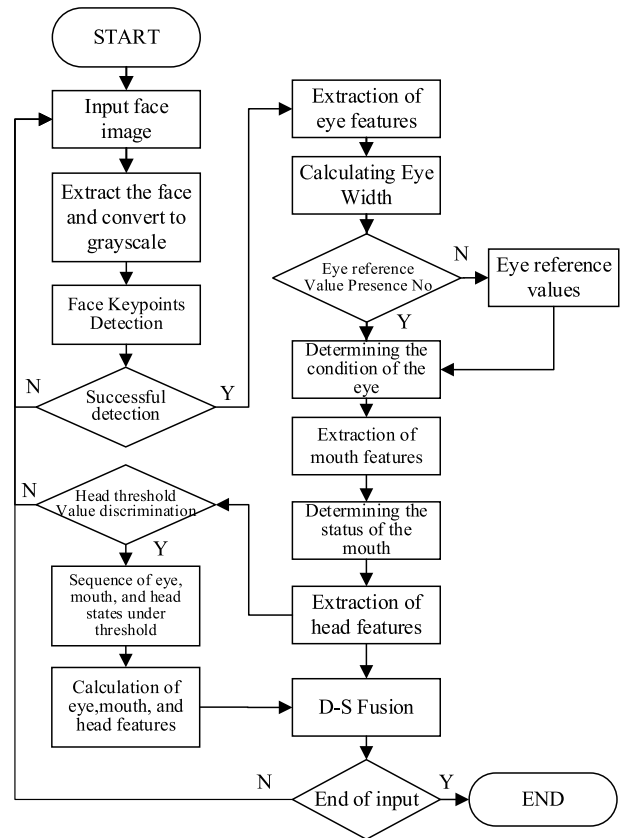


FIGURE 1. Multi-parameter fusion scheme.

the head posture angle is used to correct the mouth feature structure and associated data.

The physiological and physical experiments of the characteristic values collected from the eyes, mouth and head are comprehensively judged, and the reference values corresponding to the three fatigue characteristics are compared and analyzed, that is, the threshold values corresponding to the fatigue characteristics. Finally, the recognition states of the three subsets are adjusted and the thresholds are set according to the driver’s behaviour characteristics and habits, so as to further determine the level of fatigue driving and provide a reliable basis for the subsequent multi-parameter information fusion. In this paper, Dlib and Opencv were used to trace and mark 68 key points of the face, and 68 points of external features of the face contour were calibrated according to the model “shape_predictor_68_face_landmark.dat”. The specific procedure is shown in Figure 2.

The dot image of the tracing of the facial features and its effect are shown in Figure 3.

B. BLINK FEATURE DETECTION

To further effectively detect eye closure, the scores of the left and right eyes were calculated and averaged according to the concept of eye length/width ratio (EAR) as a final score. Less than the threshold then adds one; If it was less than the threshold value for three consecutive times, a blink activity

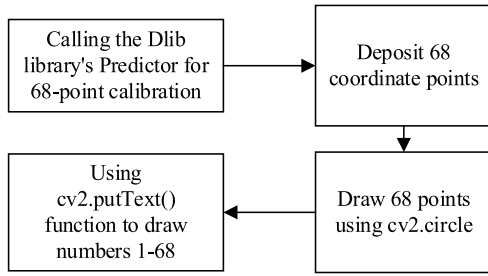


FIGURE 2. Face feature point retrieval process.



FIGURE 3. Face recognition tracing point effect.

was performed. By establishing the time correlation dimension, each blink is associated with the time dimension and the blink frequency is statistically measured when calculating the degree of opening and closing [20].

$$EAR = \frac{\|p_2 - p_6\| + \|p_3 - p_5\|}{2\|p_1 - p_4\|} \quad (1)$$

In the formula, p_2 minus p_6 is the horizontal coordinate difference, p_3 minus p_5 is similar, p_1 minus p_4 is the longitudinal coordinate difference, and the horizontal and vertical ratio is used to measure the driver's eye opening and closing degree. In this paper, a non-contact machine vision detection method is adopted to collect and identify facial fatigue features. Blink detection, as one of the most important subsets of fatigue driving feature detection, is a prerequisite for accurately identifying the driver's eye state to ensure the driver's safe driving [21]. The characteristics of drivers' eye movements during driving can be divided into closed eyes, normal frequency blinking and fatigue blinking. Short and long periods of eye closure are red lines for dangerous driving behavior. The duration of fatigue blinking is longer than that of normal frequency blinking. As shown in Figure 4.

The black box represents the blinking state (including habitual blinking and fatigue blinking) and the red box represents the closed eye state. Therefore, the alert EAR value of eye opening and closing is set to 0.15, which can detect the driver's eye opening and closing state according to different individuals. After completing the characteristics of the opening and closing states of the eye movement, the blinking state is further distinguished. Blink state is mainly divided into habitual blinking and fatigue blinking [22]. Without the driver having different driving habits and eye movement habits,

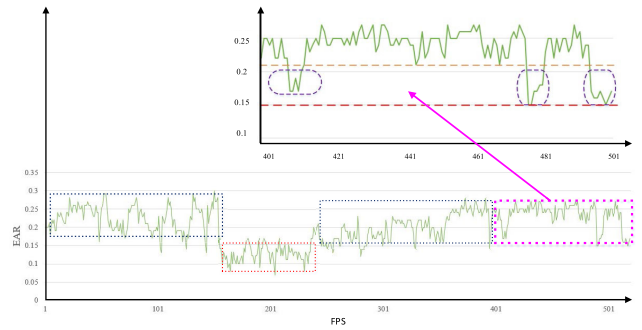


FIGURE 4. Comparative analysis of closed eyes and normal blinking.

facial feature information is collected every 3 frames during the experiment. The driver's habitual eye movement features are shown as 'burrs' on the acquisition map. If the EAR value is lower than 0.22 in a short time, it is habitual blinking with a duration of 0.2-0.4 seconds, while the duration of fatigue blinking is much longer than 0.4 seconds, which can effectively distinguish habitual blinking from fatigue blinking [23].

C. MOUTH DYNAMIC FEATURE DETECTION

Yawn detection is more complicated than blink detection. Generally, the duration of a yawn is between 3.5 seconds and 4.2 seconds. The duration of the yawning process can be effectively separated from the speech process, and the degree of mouth opening and closing during speech can also be repeated, but there are obvious differences in frequency and the actual effective process [24]. In this paper, speech and yawning are distinguished by time frame rate. First, coordinate transformation is performed on the characteristic points of the mouth to obtain the current coordinate feature points, and then the opening and closing ratio and aspect ratio are calculated. Thresholds are set through several trials and the number and frequency of yawns are counted.

The mouth opening and closing degree recognition process mainly uses the Dlib library to calculate the mouth feature points. Compared to the blink detection process, the mouth feature points are more densely distributed. The key points are selected to detect and calculate yawns. The opening and closing degree of the mouth is calculated according to the formula (2).

$$MAR = \frac{(y_{51} + y_{59}) - (y_{53} + y_{57})}{2(x_{49} - x_{55})} \quad (2)$$

In the formula, y_{51} plus y_{59} , y_{53} plus y_{57} , the difference between the two is the ordinate difference, x_{49} minus x_{55} is the horizontal difference, and the ratio of the two is used to measure the opening and closing degree of the driver's mouth. To further distinguish the difference between yawn detection and speech detection in the process of mouth detection, a time threshold was set for the degree of mouth opening and closing. The characteristic of speech is high frequency of mouth opening and closing, short time of single opening

and closing. The threshold time of speech characteristics and yawning characteristics was set as $T1$, and the speech process was defined as the single duration less than $T1$. Deep and shallow yawn detection was set. A single mouth opening and closing time greater than $T1$ and less than $T2$ was considered a shallow yawn, while a single mouth opening and closing time greater than $T2$ was considered a deep yawn. Mouth speech recognition and yawning recognition were effectively discriminated according to different duration and frequency characteristics of the opening and closing degree. The analysis of the change process of yawn, speech and MAR value is shown in Figure 5.

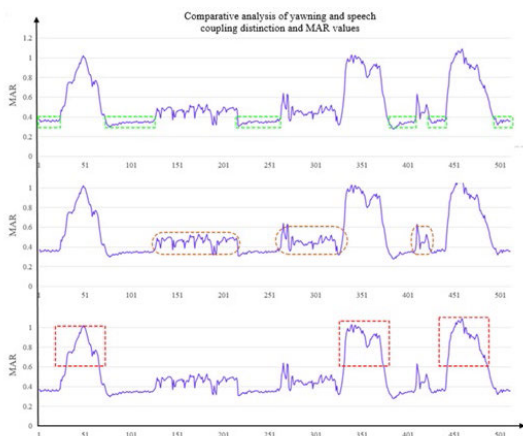


FIGURE 5. Analysis of the changes of yawn, speech and MAR.

In Figure 5, the frame rate marked in the green box means that the driver’s mouth is normally closed. The MAR value fluctuates slightly around 0.37, maintaining a relatively stable state, and the up and down curve of the mouth point cloud map changes slightly. The brown boxes represent the state of the driver’s mouth when speaking. The MAR value fluctuates a lot compared to the closed state. The MAR value fluctuates between 0.37 and 0.6, and the mutation coefficient of the spectral peak is large, which corresponds to the instantaneous state of open and closed during speaking. The graph marked in red shows the driver’s yawning state, which lasts for a long time and has a large degree of opening. The opening and closing of the mouth means relatively slow, the whole process takes about 6 seconds and the MAR value is above 0.6. Compared to the closing state and the speaking state, the MAR value of yawning shows a peak state with a change in the number of frames.

D. HEAD POSTURE DETECTION

Starting from three dimensions and four coordinate systems, this paper connects the head attitude with four coordinate systems to improve the detection accuracy of the head attitude. Blink detection and yawn detection are mainly based on extracting the feature points of the image, describing the face outline, and calculating the local information value of each frame image feature position for threshold determination

and detection. However, the head attitude detection needs to further understand the three-dimensional condition of the head attitude and carry out the transformation of various coordinate systems. The world coordinate system is centered on the coordinates of point $P(X, Y, Z)$, and any axis can be specified. $X_c, Y_c,$ and Z_c constitute the camera coordinate system, and Z -axis and radial axis are overlapped. The image coordinate system is composed of x, y and z . The xy plane is coincident with the projector plane, and the origin is coincident with the imaging hole. uv constitutes the pixel coordinate system, whose plane coincides with the projector plane, and the origin is located in the upper left corner [25].

The basic principle of camera in the imaging process is small-hole imaging. Camera parameters need to be integrated into the conversion process of various coordinate systems, so that real image parameter values can be obtained in the image processing process. The transformation relationship of the four coordinate systems is shown in Figure 6.

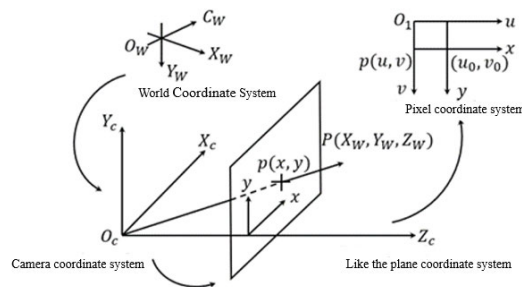


FIGURE 6. Coordinate system conversion relationship diagram.

This design process uses the transformation of pixel coordinate system and world coordinate system. Where M is the parameter matrix in the camera, R is the rotation matrix, t is the translation matrix, is the value of the target point in the direction of world coordinate Z , and s is the scaling factor.

$$s \begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = M \left(R \begin{bmatrix} X \\ Y \\ Z_{const} \end{bmatrix} + t \right) \tag{3}$$

The first matrix on the right side of the equation is the camera’s internal parameters, and the second matrix is the camera’s external parameters.

$$s \begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = \begin{bmatrix} f_x & 0 & c_x \\ 0 & f_x & c_y \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} r_{11} & r_{12} & r_{13} & t_1 \\ r_{21} & r_{22} & r_{23} & t_2 \\ r_{31} & r_{32} & r_{33} & t_3 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \\ 1 \end{bmatrix} \tag{4}$$

After transformation, we can obtain:

$$R^{-1}M^{-1}s \begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = \begin{bmatrix} X \\ Y \\ Z_{const} \end{bmatrix} + R^{-1}t \tag{5}$$

Through 3D head pose detection, calculate the real-time pose value of three dimensions, as shown in Figure 7 contains Pitch-X, Yaw-Y, Roll-Z, combined with the driver’s real

car environment, the experimental process is mainly to nod Pitch-X value as the first element, when the deflection angle is too large when the face detection is lost. x value in the range of 0-8 belongs to the normal jitter, more than 8 and above belongs to nod movement.

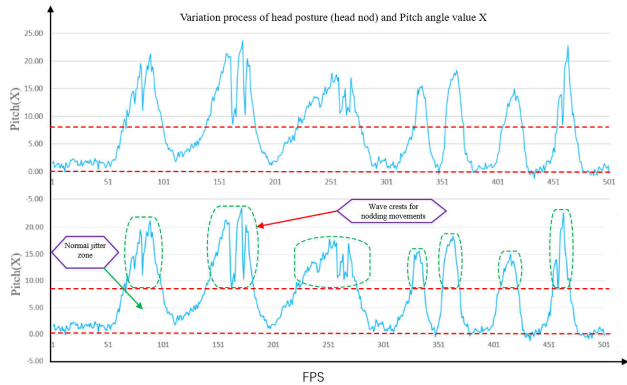


FIGURE 7. Analysis diagram of head posture characteristics.

IV. D-S EVIDENCE THEORY FUSION MULTI-FEATURE A. TYPES OF GRAPHICS

Experimental environment: Windows10 system, GTX1650Ti graphics card, development environment is Python3.7.0, Pycharm2020, deep learning framework is Tensorflow2.1, Oni S500 infrared binocular camera. Combined with the features of the eyes, mouth and head posture, the fatigue driving is divided into four levels (normal driving, mild fatigue, moderate fatigue, severe fatigue). Combined with the facial feature information recognition results of the three regions, the comprehensive decision is made to realize D-S information fusion and hierarchical warning, and the weight coefficient correlation calculation is carried out for the three parts. Improved general conflict and one-vote issue [25]. The information fusion flow chart of partial face recognition features is shown in Figure 8.

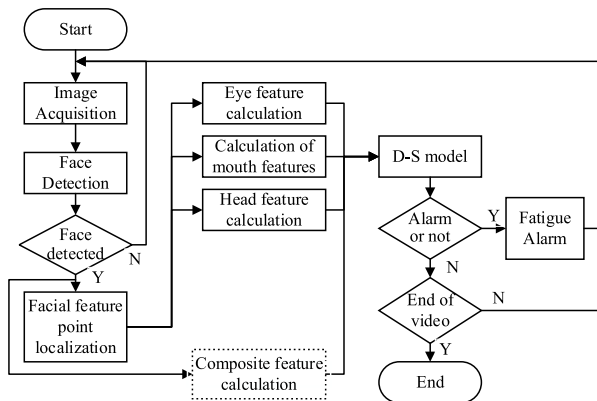


FIGURE 8. Flow chart of information fusion of partial face recognition features.

A. IMPROVED D-S ALGORITHM

Traditional D-S algorithm has problems such as unbalanced weight distribution and large conflicts influenced by

a single subset in the decision process of feature subsets [26]. An improved D-S fusion model is proposed in this paper. The feature subset is composed of facial fatigue feature information. For random variable X, the information entropy can be expressed as:

$$H(X) = - \sum_{i=1}^n p(x_i) \log p(x_i) \tag{6}$$

Then, the dependence degree of random variable X relative to another variable Y can be expressed by conditional entropy as:

$$H(X|Y) = - \sum_{j=1}^m p(y_j) \sum_{i=1}^n p(x_i|y_j) \log_2 p(x_i|y_j) \tag{7}$$

The $p(x_i)$ as the prior probability of a random variable X, $p(x_i|y_j)$ of the random variable Y under the condition of the a posteriori probability of X.

Thus, information gain $I(X; Y)$ can (7) calculated by the formula, can be used to represent variables X and Y correlation, namely $I(X; Y)$, the greater the correlation of two variables.

$$I(X; Y) = H(X) - H(X|Y) \tag{8}$$

The addition of new features with low information growth rate to the selected feature subset will lead to the moderation of the correlation degree within the subset [27]. Therefore, although the information gain can better characterize the degree of nonlinear correlation between two variables, further consideration should be given to the information growth rate between variables in order to carry out sustained feature selection [28]. After normalization of information gain, the symmetric uncertainty SU between variables can be measured according to formula (9).

$$SU(X, Y) = 2 \left[\frac{I(X; Y)}{H(X) + H(Y)} \right] \tag{9}$$

According to Equation (9), the larger the value of SU between a feature and a class label, the higher the relevance of the feature to the class label, while the larger the value of SU between two features, the higher the relevance between the features, i.e., containing more information redundancy [29]. Based on the principles of maximum correlation and minimum redundancy of feature selection, a subset of features with strong correlation with class labels needs to be selected, so the feature selection of feature f_i can be performed according to the evaluation function of Equation (10).

$$J(f_i) = SU(S \cup f_i, C) - \frac{1}{|S|} \sum_{f_s \in S} SU(f_s, f_i) \tag{10}$$

$SU(S \cup f_i, C)$ is the characteristic feature subset S join for f_i after the formation of the new collection with nonlinear correlation, the label C $SU(f_s, f_i)$ for the proposed feature f_i and feature subset S of correlation between the variables in the |S| for the number of features or first-order norm,

Fatigue A formulation process and weighting coefficients constitute the early warning algorithm A

Inputs: training data, number of iterations n , number of frames acquired, coordinates of human eye feature points a , coordinates of mouth features b , coordinates of head pose c

Output: MAR , EAR , $Nod(X, Y, Z)$ and fatigue warning level

```

1 BEGIN
2 Repeat
3 For  $i=1,2,3,\dots,n$  do//nis the total amount of feature data collected (number of iterations)
4 Video  $img = cv2.imread((file), cv2.IMREAD_COLOR)$  perform single-frame image feature conversion
5 Get new_img for pre-processing and capturing facial features (tracing points)
6 End for
7 Calculation of  $EAR$ ,  $MAR$  and  $Nod(X, Y, Z)$ 
8 Calculation of fatigue indicators based on set thresholds.
9 Close eyes( $EAR < 0.15$ ), Routine blinking( $0.16 < EAR < 0.22$  &  $0.5s < T < 1s$ ), Fatigue blinking( $0.16 < EAR < 0.22$  &  $T > 1s$ );
Speaking( $MAR \approx 0.37$ ), Regular Speaking( $1 < MAR < 0.6$  &  $0 < T < 3s$ ), Yawn( $1 < MAR$  &  $T > 4.5s$ ); XYZ Offset factor  $\beta > 0.5$  &  $T > 1s$  &  $n > 5$ )
10 If ( $EAR < 0.15$  &  $T > 1s$  &  $n > 5$ )
    Print{ $EAR, MAR, Nod(X, Y, Z)$  & Fatigue A}
    else if ( $0.16 < EAR < 0.22$  &  $T > 1s$  &  $0.6 < MAR$  &  $T > 4.5s$  &  $n$ )
    Print{ $EAR, MAR, Nod(X, Y, Z)$  & Fatigue A}
    else if ( $0.16 < EAR < 0.22$  &  $T > 1s$ , XYZ Offset factor  $\beta > 0.5$  &  $T > 0.5s$  & Number of times  $> 5$ )
    Print{ $EAR, MAR, Nod(X, Y, Z)$  & Fatigue A}
    else if ( $0.6 < MAR$  &  $T > 4.5s$ , XYZ Offset factor  $\beta > 0.5$  &  $T > 0.5s$  & Number of times  $> 5$ )
    Print{ $EAR, MAR, Nod(X, Y, Z)$  & Fatigue A}
    else if ( $0.16 < EAR < 0.22$  &  $T > 1s$ ,  $0.6 < MAR$  &  $T > 4.5s$ , XYZ Offset factor  $\beta > 0.5$  &  $T > 0.5s$  & Number of times  $> 5$ )
    Print{ $EAR, MAR, Nod(X, Y, Z)$  & Fatigue A}
    else (Does not meet Fatigue A)
    Print{Other grades( $B, C, D$ )}
11 End
    
```

So according to the evaluation function $J(f_i)$ values decide whether to feature f_i elected to the optimal feature subset.

In the improved D-S fusion model, the corresponding weight coefficient $M_i(A_j)$ is added to a single index, and A_j is a subset of facial fatigue feature. There are 5 dimensional fatigue feature data in this paper. Three weight coefficients $M_1(A_i) M_2(A_j) M_3(A_l)$ are set for fatigue blink, fatigue yawn and fatigue bow. The product of weight coefficients of feature subsets is derived, and conflict function is attached to the sum of derivative coefficients [30], to obtain the weighted result $m(A)$ after the fusion of feature subsets, as shown in Equation 11.

$$m(A) = \begin{cases} \frac{\sum_{A_i, A_j, A_l, \dots = A} ds_i M_1(A_i) M_2(A_j) M_3(A_l) \dots}{1 - k} \\ + (1 - ds_i) f(A), A \neq \Phi \\ 0, A = \Phi \end{cases} \quad (11)$$

The weight calculation of the coupling coefficients of the remaining three fatigue feature subsets is shown in Equation 12.

$$\sum_{A_i, A_j, A_l, \dots = \Phi} ds_i M_1(A_i) M_2(A_j) M_3(A_l) = \frac{DS(M_i, M_j)}{DS(m_i, m_j) + DS(M_i, M_j)} \quad i, j = 1, 2, \dots, n \quad (12)$$

The improved D-S model is capable of weighting all input single features and multiple features to output fatigue levels for early warning.

B. DETERMINATION OF FATIGUE CHARACTERISTICS

In the process of driver’s face feature information collection, the face is affected by daytime and night light. The infrared binocular camera is used to collect the face feature information, and then the three parts of the face fatigue feature representation interest region is extracted for real-time detection [31]. In order to further improve the effectiveness of multi-feature fusion, blink, yawn and head posture were judged by multi-feature fusion. MAR , EAR and Nod constituted the decision fusion model of dangerous driving warning system [32]. Severe fatigue, moderate fatigue, mild fatigue and Normal driving correspond to *Fatigue A*, *Fatigue B*, *Fatigue C* and *Normal* driving states respectively. Take *Fatigue A* as an example:

V. ANALYSIS AND COMPARISON OF EXPERIMENTAL RESULTS

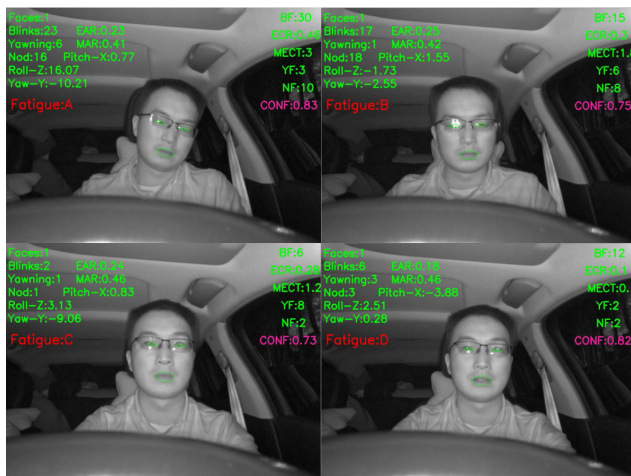
A. RESULTS OF FACE DETECTION EXPERIMENTS

Facial fatigue sign detection consists of 5 fatigue sign subsets, including blink frequency, yawn frequency, nod frequency, blink percentage and longest eye closing time. Each subset is assigned a corresponding weight. The subsets are input into the improved D-S model to obtain the evaluation results of this multi-parameter model after weighted calculation. Indices of these signs were obtained according to the fatigue sign coefficient corresponding to the subsets, and based on the output of the multi-parameter fusion model, the comprehensive decision was made to realize the feedforward differential fatigue sign index calculation.

Based on the improved fusion model, the detection of serious, moderate, and mild fatigue states as well as normal



(a) Daytime test results



(b) Night test results

FIGURE 9. Face fatigue detection effect of improved D-S algorithm.

driving state is realized in practical environments, and the adaptive experimental results classified into Class A, B, C and D are printed. The results of facial fatigue sign detection with the improved D-S algorithm are shown in Figure 9.

B. COMPARISON OF D-S IMPROVED ALGORITHM RESULTS

According to the previous research, fatigued driving can be divided into four classes: Class A, Class B fatigue, Class C and Class D. Class A is the most serious fatigue state, Class B fatigue is moderate fatigue, Class C fatigue is mild fatigue, and Class D fatigue is normal driving. The results of real vehicle tests conducted using the improved D-S algorithm are compared with the actual states, as shown in Table 1.

The improved D-S fusion algorithm is applied in the real vehicle detection, and a total of 1,000 fatigue characteristic detection experiments are carried out. According to Table 4.1, when the actual state of the driver is Class A, the accurate determinations number 148, among which 5 are inaccurate.

TABLE 1. The improved D-S algorithm for real vehicle testing.

Indicators	Status	Improved D-S algorithm detection/sub			
		A	B	C	N
The actual state of the driver/session	A	148	2	4	0
	B	3	228	0	2
	C	2	3	296	4
	N	0	4	2	302

TABLE 2. Comparison of results of different evaluation models in the practical environment.

Different face features detection	Day time/%			Night /%		
	Data efficiency	T/s	Precision	Data efficiency	T/s	Precision
Yawn	93.2	0.059	90.2	93.2	0.56	82.6
Head	94.1	0.076	88.6	94.1	0.59	79.3
D-S	95.3	0.052	91.0	95.3	0.51	84.1
Improved D-S	94.5	0.046	93.8	94.5	0.49	85.6

The accurate determinations number respectively 228 in the Class B state, 296 in the Class C state, and 302 in the Class D state.

After 1,000 tests were completed in the experimental environment with sufficient lighting, 500 times tests were conducted to determine the characteristics of fatigue at night. Data efficiency is the effective value of face fatigue feature detection in the collection process. In complex and multi-scene, driver face fatigue features are collected by infrared camera, and the collected driver face features are compared with real-time video frames to further ensure the effectiveness and rationality of the detection experiment, and the comparative analysis results of experimental data is shown in Table 2.

In order to solve the class imbalance problem, the official fatigue data set and self-made data set are used for comprehensive training in the process of model training, and the face fatigue feature data in different environments are used.

Through physiological and physical experiments, the driver has facial features when he is tired, such as yawning, blinking, drowsiness and other head features. Based on the above features, the driver's facial fatigue feature judgment threshold is formulated. Finally, the fusion judgment is made according to the three fatigue feature values, and the conflict factor of the fusion process is changed to achieve comprehensive judgment.

D-S fusion model is a classical data fusion model. In this paper, blink features, yawn features and nodding frequency are systematically fused, and weight factors are added to the three data features to optimize the fusion model.

Blink, yawn and head respectively represent a single feature value in the process of face fatigue driving detection. Multiple feature values are not fused, and fatigue and fatigue detection accuracy are judged only according to a single feature of the face. The D-S corresponds to the fatigue determination of the results obtained after the fusion of the three face feature detection values. The improved D-S model is also the same evaluation method.

The accuracy of the fusion model is higher than that of the model with single sign subsets, and the accuracy of the improved D-S algorithm reaches 93.8% in daytime, and 85.6% at night. The improved D-S model is more generalizable. Distributing weight after inputting a single feature subset can effectively reduce the conflict among fatigue sign subsets and comprehensively improve the decision-making accuracy of the fusion model. Figure 10 compares the accuracy between D-S before and after improvement as well as single sign subsets.

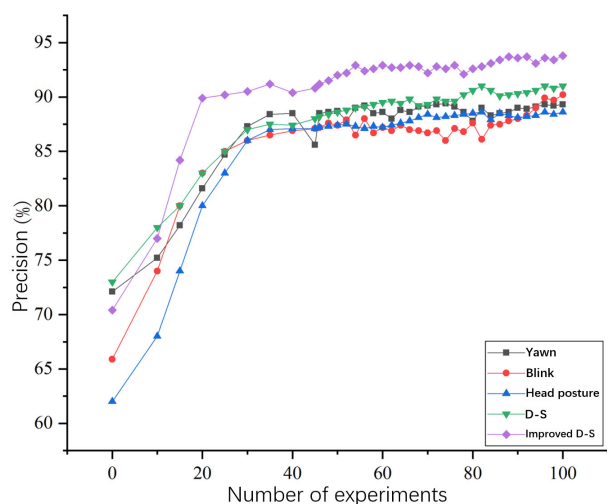


FIGURE 10. Precision comparison between D-S before and after improvement and single feature subset.

VI. CONCLUSION

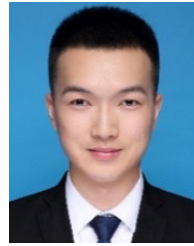
Starting with video sequence and non-contact fatigue sign detection, this paper proposes a facial fatigue sign detection algorithm based on improved D-S evidence theory, which uses multi-threaded optimized Dlib to track and locate fatigue signs, and determines fatigue sign subsets according to the coupling between EAR and MAR and time flow. By calculating the Pitch Angle through the conversion of head pose from 2D image to 3D data, the head pose model was constructed based on the fatigue sign subsets, and the three fatigue sign subsets were correlated and coupled to further improve the

effectiveness of the fatigue level corresponding to the subsets. The fusion of multiple subsets is realized by the improved D-S evidence theory, and the weight of three feature subsets is optimized to mitigate the influence of the “general conflict” and “one vote rejection” problems. The experimental results show that the improved D-S face detection algorithm can effectively solve the problem of face illumination and partial occlusion in complex environment, and the face fatigue sign detection accuracy reaches 93.8%. It can meet the practical demand of fatigued driving detection and provide support for driving safety warning, which presents its practicability. In the experimental process of this paper, there are still some standards that are not uniform enough, which brings some unavoidable interference and influence factors to the experiment. Judging fatigue by facial fatigue features is a hot spot in current and future research. At the same time, it is also faced with environmental impacts, such as night detection, light impact, and differences in evaluation criteria for different facial features. However, in the future intelligent driving development process, the above problems will be the focus to solve. At the same time, the decision method based on multi-source information feature fusion is the advantage of intelligent integration.

REFERENCES

- [1] X. Rui and D. Yuan, “Analysis of factors influencing the severity of traffic accidents involving fatigue driving,” *China Prod. Saf. Sci. Technol.*, vol. 18, pp. 20–26, Jun. 2019.
- [2] L. Yan, Y. Shi, M. Wei, and Y. Wu, “Multi-feature fusing local directional ternary pattern for facial expressions signal recognition based on video communication system,” *Alexandria Eng. J.*, vol. 63, pp. 307–320, Jan. 2023.
- [3] F. Xiaofeng, “Fatigue driving detection method based on fused eye-mouth head features,” *J. Saf. Environ.*, vol. 22, pp. 263–270, Jan. 2022.
- [4] L. Mengjia and Z. Jianguo, “SVM-based algorithm for fatigue driving face image feature detection,” *Semicond. optoelectron.*, vol. 41, pp. 278–282, Feb. 2020.
- [5] H. Wang and H. Bai, “Driver fatigue state detection and early warning technology based on computer vision,” *Sci. Technol. Eng.*, vol. 22, pp. 4887–4894, Dec. 2022.
- [6] C. Bo, F. Ruijia, Z. Wei, L. Jiawen, and Z. Xibo, “Research on the monitoring and early warning method of driver fatigue based on multi-source information fusion,” *Road Traffic Sci. Technol.*, vol. 26, pp. 13–18, Feb. 2009.
- [7] Q. Zhao, J. Jiang, Z. Lei, and J. Yi, “Detection method of eyes opening and closing ratio for driver’s fatigue monitoring,” *IET Intell. Transp. Syst.*, vol. 15, no. 1, pp. 31–42, Jan. 2021.
- [8] Z. Yan, X. Zhongzhi, Y. Wen, L. Shusheng, and Z. Xun, “Fatigue driving detection technology based on facial feature points in low light environment,” *J. Automot. Saf. Energy Conservation*, vol. 13, pp. 282–289, Feb. 2020.
- [9] G. Zhao, A. Chen, G. Lu, and W. Liu, “Data fusion algorithm based on fuzzy sets and D-S theory of evidence,” *Tsinghua Sci. Technol.*, vol. 25, no. 1, pp. 12–19, Feb. 2020.
- [10] Y. Pan, L. Zhang, Z. Li, and L. Ding, “Improved fuzzy Bayesian network-based risk analysis with interval-valued fuzzy sets and D-S evidence theory,” *IEEE Trans. Fuzzy Syst.*, vol. 28, no. 9, pp. 2063–2077, Sep. 2020.
- [11] Z. Zhang, H. Zhang, H. Liu, S. Xin, N. Xiao, and L. Zhang, “Frontal face generation based multi-angle face identification system,” in *Proc. Int. Conf. Comput., Control Robot. (ICCCR)*, Jan. 2021, pp. 329–334.
- [12] S. Murugan, J. Selvaraj, and A. Sahayadhas, “Analysis of different measures to detect driver states: A review,” in *Proc. IEEE Int. Conf. Syst., Comput., Autom. Netw. (ICSCAN)*, Mar. 2019, pp. 1–6.

- [13] S. Mohanty, S. V. Hegde, S. Prasad, and J. Manikandan, "Design of real-time drowsiness detection system using Dlib," in *Proc. IEEE Int. WIE Conf. Electr. Comput. Eng. (WIECON-ECE)*, Nov. 2019, pp. 1–4.
- [14] E. K. M. Sanchez, N. B. Linsangan, and R. E. Angelia, "Three triangle method for face recognition using Dlib and OpenCV," in *Proc. IEEE 12th Int. Conf. Humanoid, Nanotechnol., Inf. Technol., Commun. Control, Environ., Manag. (HNICEM)*, Dec. 2020, pp. 1–5.
- [15] S. Li and B. Chao, "A multi-modal information fusion-based method for detecting fatigue driving of train drivers," *J. Railways*, vol. 44, pp. 59–65, Jun. 2022.
- [16] N. Boyko, O. Basystiuk, and N. Shakhovska, "Performance evaluation and comparison of software for face recognition, based on Dlib and OpenCV library," in *Proc. IEEE 2nd Int. Conf. Data Stream Mining Process.*, Aug. 2018, pp. 478–482.
- [17] A. Kuwahara, R. Hirakawa, H. Kawano, K. Nakashi, and Y. Nakatoh, "Eye fatigue prediction system using blink detection based on eye image," in *Proc. IEEE Int. Conf. Consum. Electron. (ICCE)*, Jan. 2021, pp. 1–3.
- [18] H. Li, Q. Zhao, W. Zhao, and Y. Wu, "Driver-car natural interaction method based on head-eye behaviors," in *Proc. 11th Int. Conf. Intell. Human-Machine Syst. Cybern. (IHMSC)*, vol. 1, Aug. 2019, pp. 161–164.
- [19] H. Garg, "Drowsiness detection of a driver using conventional computer vision application," in *Proc. Int. Conf. Power Electron. IoT Appl. Renew. Energy Control (PARC)*, Feb. 2020, pp. 50–53.
- [20] A. M. Al-madani, A. T. Gaikwad, V. Mahale, Z. A. T. Ahmed, and A. A. A. Shareef, "Real-time driver drowsiness detection based on eye movement and yawning using facial landmark," in *Proc. Int. Conf. Comput. Commun. Informat. (ICCCI)*, Jan. 2021, pp. 1–4.
- [21] G. Huili, W. Zim, and G. Hao, "Research on fatigue driving warning system based on facial multi-features," *J. Commun.*, vol. 39, pp. 22–29, Sep. 2018.
- [22] K. Li, Y. Gong, and Z. Ren, "A fatigue driving detection algorithm based on facial multi-feature fusion," *IEEE Access*, vol. 8, pp. 101244–101259, 2020.
- [23] K. Rezaee, S. R. Alavi, M. Madanian, M. R. Ghezlbash, H. Khavari, and J. Haddadnia, "Real-time intelligent alarm system of driver fatigue based on video sequences," in *Proc. 1st RSII/ISM Int. Conf. Robot. Mechatronics (ICRoM)*, Feb. 2013, pp. 378–383.
- [24] I. Lashkov, A. Kashevnik, N. Shilov, V. Parfenov, and A. Shabaev, "Driver dangerous state detection based on OpenCV & Dlib libraries using mobile video processing," in *Proc. IEEE Int. Conf. Comput. Sci. Eng. (CSE) IEEE Int. Conf. Embedded Ubiquitous Comput. (EUC)*, Aug. 2019, pp. 74–79.
- [25] A. Nemcová, V. Svozilová, K. Bucsuházy, R. Smíšek, M. Mézl, B. Hesko, M. Belák, M. Bilík, P. Maxera, M. Seitl, T. Dominik, M. Semela, M. Sucha, and R. Kolár, "Multimodal features for detection of driver stress and fatigue: Review," *IEEE Trans. Intell. Transp. Syst.*, vol. 22, no. 6, pp. 3214–3233, Jun. 2021.
- [26] X. Zhang, H. Gan, and L. Chen, "A fatigue driving detection algorithm based on weighted D-S evidence theory," in *Proc. Chin. Autom. Congr. (CAC)*, Nov. 2020, pp. 2705–2710.
- [27] C. Zhu, B. Qin, F. Xiao, Z. Cao, and H. M. Pandey, "A fuzzy preference-based Dempster-Shafer evidence theory for decision fusion," *Inf. Sci.*, vol. 570, pp. 306–322, Sep. 2021.
- [28] Y. Shang, X. Xu, Y. Jiao, Z. Wang, Z. Hua, and H. Song, "Using lightweight deep learning algorithm for real-time detection of apple flowers in natural environments," *Comput. Electron. Agricult.*, vol. 207, Apr. 2023, Art. no. 107765.
- [29] P. Ping, C. Huang, W. Ding, Y. Liu, M. Chiyomi, and T. Kazuya, "Distraction driving detection based on the fusion of deep learning and causal reasoning," *Inf. Fusion*, vol. 89, pp. 121–142, Jan. 2023.
- [30] Q. Zhang, Z. Zhu, Y. Bai, G. Liao, and T. Liu, "Distraction driving detection based on the improved CenterNet with attention mechanism," *Multimedia Tools Appl.*, vol. 81, no. 6, pp. 7993–8005, Mar. 2022.
- [31] S. Bag, R. Golder, S. Sarkar, and S. Maity, "SENE: A novel manifold learning approach for distracted driving analysis with spatio-temporal and driver praxeological features," *Eng. Appl. Artif. Intell.*, vol. 123, Aug. 2023, Art. no. 106332.
- [32] M. S. Ansar, Y. Ma, S. Chen, K. Tang, and Z. Zhang, "Investigating the trip configured causal effect of distracted driving on aggressive driving behavior for e-hailing taxi drivers," *J. Traffic Transp. Eng.*, vol. 8, no. 5, pp. 725–734, Oct. 2021.



JIANFENG CAI received the B.S. degree in automation from Southwest Petroleum University, in 2018, and the M.S. degree in resources and environment from the Chongqing Institute of Science and Technology, China, in 2023. He was with the Institute of Computer Applications, Chengdu Branch, Chinese Academy of Sciences. During his postgraduate study, his main research directions were digital image processing and multi-source information fusion, and also research on dangerous driving detection and computer vision intelligent driving assistance system.



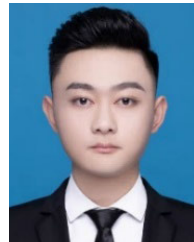
XU LIAO was born in 1997. He is currently pursuing the master's degree resources and environment with the Chongqing University of Science and Technology, Chongqing, China. His research interests include digital image processing, emotion recognition, robot control technology, and assisted driving.



JUNBO BAI was born in 1977. He received the B.S. and M.S. degrees in marine and offshore engineering from the Wuhan University of Technology, in 2000 and 2003, respectively. He was the CEO of several companies engaged in research on intelligent vehicles and artificial intelligence-related technologies, and developed a number of "driverless vehicles" used in logistics and distribution, large ports. He currently serves as the Vice President of Shanghai Ecar Technology Company Ltd.



ZHEN LUO was born in Chongqing, China, in 1997. She is currently pursuing the master's degree in resources and environment with the Chongqing University of Science and Technology, Chongqing. Her research interests include digital image processing, robot control technology, and assisted driving.



LILI was born in Chongqing, China, in 2000. He is currently pursuing the master's degree with the Chongqing University of Science and Technology. His research interest includes robot intelligent control technology.



JUNJIE BAI received the master's degree in control theory and engineering from the Wuhan University of Science and Technology, Wuhan, China, and the Ph.D. degree in instrument science and technology from Southeast University, Nanjing, China. He is currently a Professor with the Chongqing University of Science and Technology, China. In 2017, he was a Visiting Scholar in artificial intelligence technology with the Department of Electrical Engineering, University of Calgary, Canada. He is also working on computer vision technology, intelligent driving technology, robotics and intelligent systems, tactile perception, and human-computer interaction.

• • •