2022 21st International Symposium on Distributed Computing and Applications for Business Engineering and Science (DCABES)

Fatigue detection for public transport drivers under the normalization of epidemic prevention

Zhe Yu, Lei Li, Lijun Xu*, Kansong Chen*
School of Computer Science and Information Engineering, Hubei University
Wuhan 430062, P. R. China
kschen1999@aliyun.com (K. Chen); xulijun@hubu.edu.cn (L. Xu)

***Abstract*—Several studies have shown that fatigue driving is one of the important causes of public transport safety accidents. With the outbreak of the COVID-19 , the wearing of masks by public transport drivers presents new challenges for computer-based visual fatigue detection. In order to achieve the goal of accurately capturing the landmark information of the face even when the face is occluded by a large area, we adopt the DNN-based face detection method which has the highest accuracy and the best occlusion resistance. When the driver's face is blocked, the landmark information of the blocked face can be accurately detected by using our optimized face landmark detector. The accuracy rate of landmark recognition can reach 97.80%. On this basis, we calculate the driver's eye information, mouth information and the driver's head deflection angle information in real time as the judgment indicators of the degree of fatigue to comprehensively evaluate the driver's fatigue state. And use mathematical methods to fuse indicators in real time, classify the driver's fatigue state according to the value of the fusion indicators, and adopt different early warning methods for different levels of fatigue. In addition, in order to further improve the accuracy of the detection results and exclude the influence of other facial behaviors on our fatigue judgment indicators, we propose a kinetic energy calculation formula for facial organs based on the improved optical flow method. According to the different kinetic energy of facial organs in different states , which can accurately distinguish the different behaviors of the same facial organs such as blinking and closing eyes, yawning and speaking, which significantly increases the robustness and generalization ability of the detection program. The final experimental results show that the correct rate of the method for determining the degree of fatigue of the driver and passengers can reach 98.40% and 92.30% respectively when the driver does not wear a mask or wears a mask.**

Keywords- DNN face detection; fatigue detection; optical flow method; epidemic prevention

I. INTRODUCTION

At present, the shortcomings of computer vision-based fatigue driving detection methods in the epidemic situation are as follows:(1)The mainstream DLIB face feature point detection method does not have occlusion resistance, so the landmarks of the face after wearing a mask cannot be obtained.(2) The fatigue index for reference is too single.(3)The indicators of fatigue are easily disturbed by other facial behaviors.(4)Failure to effectively solve the problem of driver judgment and fatigue classification warning in critical fatigue state.

In order to solve the above problems, this paper proposes a fatigue driving detection method which is suitable for public transport drivers under the normalization of epidemic prevention. First of all, for the drivers who wear masks, the DNN face detection method[1] with high accuracy and strong anti-occlusion ability is used to detect the face first, and then we use our optimized and trained face landmark detector to extract the driver's eye information as a basis for judging fatigue. For drivers who do not wear masks, the traditional dlib face landmark detection algorithm is used[2]. Finally, the classification determines the fatigue level of the driver. At the same time, the kinetic energy of the driver's eyes and mouth is calculated in real time during the monitoring process, and the interference signals caused by blinking or talking are filtered out, which improves the accuracy of judgment and filters out invalid interference signals. Efficiency and real-time performance have been further improved. The flow chart of the final designed system algorithm is shown in Figure 1：

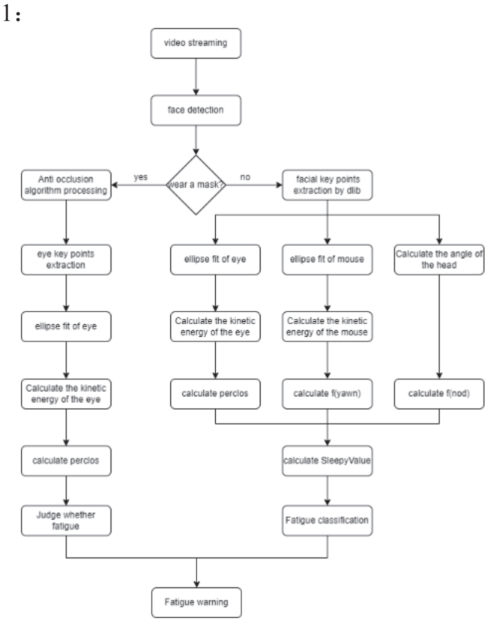

Fig 1 System algorithm flow chart

II. FATIGUE DETECTION OF DRIVERS WEARING MASKS

A. Mask wearing detection

The fatigue detection method designed in this paper adopts different facial information acquisition methods for drivers wearing masks and drivers who do not wear masks[3]. Therefore, the first step in processing video stream information is to determine whether the driver wears a mask. In order to quickly judge the situation of the driver's

2473-3636/22/$31.00 ©2022 IEEE
DOI 10.1109/DCABES57229.2022.00032

mask wearing, we adopted a relatively simple, rapid and high-accuracy HSV color threshold segmentation method. The method is mainly divided into three steps: (1) Use the DNN network to detect and locate the face. (2) In the detected face area, the skin color is extracted through the HSV threshold. (3) Determine whether the driver wears a mask by the ratio of the area of the skin contour to the ROI area of the face. In practical applications, we set this ratio to 0.65, that is, when the covered area of the face reaches 65%, it is considered that the person is wearing a mask. The specific detection effect is shown in Figure 2. The face without a mask will be framed by a red line.

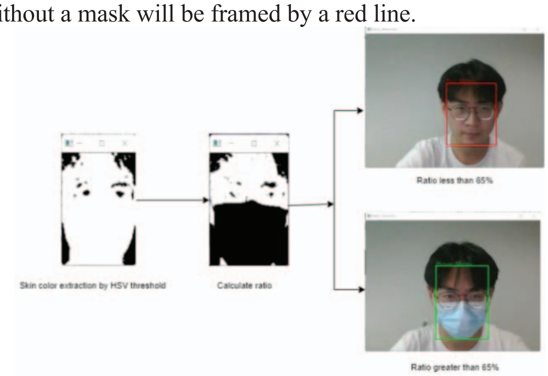

Fig 2 Schematic diagram of mask wearing detection

B. Obtaining facial information of drivers wearing masks

In order to obtain the face information that is occluded by a large area accurately and efficiently, we propose a new face landmark extraction method based on DNN face detection. The method is mainly divided into three steps: firstly, we use the DNN face detection method which has the strongest anti-occlusion ability and the highest accuracy to track and locate the face of the driver occluded by the mask. Secondly, this paper trained a new face landmark detector on iBUG 300-W data set. The face landmark detector can only calibrate the driver's eyes and several landmarks in the face contour, and the landmarks information of the nose and mouth covered by masks is discarded by our new landmark detector. Finally the face landmarks are aligned using similarity transformation.

In this paper, the method of cascade regression[4] is used for the training of face landmark detector. The detection target of facial landmarks is predictive vector (Facial Shape), which can be expressed as follows:

$$S=（x_1,y_1,x_2,y_2,\cdots\cdots,x_k,y_k）$$

Among them, k represents the number of landmarks. Because each landmark has two coordinates, horizontal and vertical, the length of the estimated vector S is 2k. For each input I, an initial shape S^0 is given, which is usually an average shape computed over the training set. And each level will output an offset estimate ΔS which is obtained from the input image, and then each level of operation will get a more accurate predicted position of the facial key point Landmark.

$$S^{t+1}=S^t+r_t(\text{Ø}(\text{I},S^t))$$

Among them, S^{t+1} and S^t represent the shape of the face predicted at level t and t+1 respectively, and represent the regression function. Finally we train our face landmark detector on the iBUG 300-W dataset. Some pictures of the iBUG 300-W dataset are shown in Figure 3. The landmarks on each person's face have been manually marked on the images of this dataset, but for practical application, what we want to obtain is only the landmark information of the driver's eyes after being blocked by the mask, as the follow-up driving. The basis for judging the degree of fatigue of personnel. So we only get the eyes landmark coordinates in the dataset as our training data. Part of the parameter settings during training and the accuracy results after training are shown in Table 1. Finally, the facial landmark detector trained in this paper can obtain 97.80% accuracy on the test set. Figure 3 shows the practical application effect of the trained face landmark detector.

parameter names	value
cascade_depth	10
tree_depth	5
num_trees_per_cascade_level	500
nu	0.15
oversampling_amount	20
feature_pool_size	1000
lambda	0.1
num_test_splits	20
padding_mode	landmark_relative
Feature_pool_regin_padding	0
Precision	0.978

Table 1 The settings of some parameters during training and the results of training

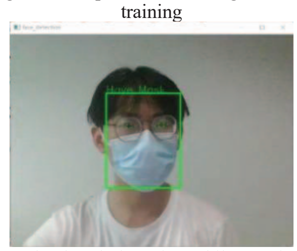

Fig 3 Eyes landmark detection while wearing a mask

C. Circumscribed oval fitting of the eye

Most of the existing fatigue detection algorithms based on computer vision evaluate the eye state by calculating the Euclidean distance ratio(E_{AR}) of the vertical coordinate and the horizontal coordinate. The calculation formula of E_{AR} is：

$$E_{AR}=\frac{|P_2-P_6|+|P_3-P_5|}{2|P_4-P_1|}$$

However, there is a problem in using this method in practical applications, that is, we cannot find an E_{AR} value that applies to all different eyes to judge a person's eye opening and closing state. For some people with small eyes, the eyes may even be judged as closed eyes. In order to solve this problem, we processed the minimum circumscribed ellipse for the landmarks of the eye, and used the area of the ellipse to fit the area of the driver's eyes, as shown in Figure 4. The algorithm will save the area of the ellipse surrounding the driver's eyes when the eyes are opened normally, and will further accurately judge the state of the eyes of different drivers according to the degree of change in the area of the external graphics, so it has wider adaptability. The final ellipse fitting effect diagram and the area change diagram of the circumscribed ellipse are shown in Figure 4.

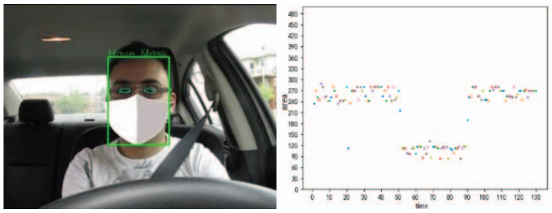

Fig 4 ellipse fitting and the area change diagram

D. Kinetic energy calculation of facial organs based on improved optical flow method

In practical application, all the pixels in the eyes and mouth are too many and the calculation is too complicated to meet the real-time requirements. Therefore, according to the definition of the optical flow method[6], we propose a simplified scheme that only calculates the speed and direction of movement of the landmarks identified in the eyes and mouth. The advantage of this method is to reduce the computational complexity and simplify the complexity of optical flow method. In addition, we use this method without considering the relocation of point A in the next frame, because the face landmark detector will help us track the position of each landmark after the change in the next frame.

Finally, we need to design the kinetic energy calculation formulas for the eyes and mouth. We named the 12 landmarks on the left and right eyes of the eyes as E_1-E_{12}, and named the 20 landmarks of the mouth as M_1-M_{20}, and set the mass m to 1, and the speed V_E and $V_M$are obtained by dividing the displacement distance of each point marked in the above frame relative to the previous frame by the displacement time. Then the final eye kinetic energy calculation formula and mouth kinetic energy calculation formula are as follows：

$$E_{e}=\sum_{E=1}^{12}\frac{1}{2}V_E^2 \qquad\qquad E_{m}=\sum_{M=1}^{20}\frac{1}{2}V_M^2$$

The following Figure 5 respectively record the kinetic energy changes of the driver's eyes and mouth. The part framed by the red rectangle indicates that the kinetic energy fluctuates greatly, and the driver is blinking and talking.

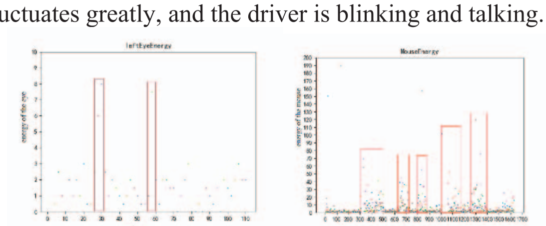

Fig 5 Kinetic energy change diagram of eyes and mouth

E. Fatigue level determination of drivers wearing masks

After a large number of experiments, it has been shown that the degree of fatigue is closely related to the length of time when the eyes are closed. The longer the eyes are closed, the more serious the degree of fatigue is[7]. PERCLOS has the best correlation with fatigue driving[8], and has high reliability and robustness in the field of fatigue detection[9]. There are usually three measurement methods for the measurement of the PERCLOS value. Here we choose the one with the highest correlation, that is, to calculate the proportion of time when the area of the eyelid covering the eyeball exceeds 80%. In order to calculate this value specifically, we have obtained the area of the minimum circumscribed ellipse of the driver's eye in the previous steps, and monitored the area change of the current minimum circumscribed ellipse of the eye in real time. When the ratio of the current graphic area to the initial graphic area is less than 20%, we record the situation. Then, the specific PERCLOS value is obtained according to the time converted by the number of frames. We record the total number of frames in the image with the ratio of the current eye area to the initial eye area less than 20% per unit time as N_e, and record the total number of frames of the video per unit time as N_P. Then there is the following expression：

$$\text{PERCLOS} = \frac{N_e}{N_P} \times 100\%$$

Finally, we use the PERCLOS value to judge whether the driver wearing a mask is in a state of fatigue.

III. FATIGUE DETECTION OF DRIVERS WITHOUT MASKS

The fatigue detection work for drivers without masks is based on the content in Chapter 1 and further expanded. Because the face is not covered by the mask, we can obtain the driver's mouth and head angle information. So we introduce two more variables, f(yawn) and f(nod), they represent the frequency of yawning and the frequency of nodding because of exhaustion. The number of yawns per unit time is represented by N_m; the number of nods in unit time is represented by N_h. Then there are the following two expressions：

$$f(yawn) = \frac{N_m}{N_P} \times 100\% \qquad\qquad f(nod) = \frac{N_h}{N_p} \times 100\%$$

At this point we have obtained the three variables PERCLOS ,f(yawn) and f(nod),which are required to calculate the driver's fatigue level. The next step is to set different weights for each indicator according to the reliability of the driver's fatigue performance, and finally calculate the SleepyValue.

$$SleepyValue = W_1 \times \text{PERCLOS} + W_2 \times f(yawn) + W_3 \times f(nod)$$

In order to determine the specific value of the weight W_1，W_2，W_3 , We searched the Internet for an open source fatigue driving dataset, which contains a total of 4125 images of two drivers in fatigue driving. Part of the data set is shown in Figure 6.

Fig 6 Fatigue driving dataset

The results of classified statistics on the images in this data set according to the behaviors of drivers such as eye closing, yawning and lowering head are shown in Table 2. It should be noted that there are a total of 4125 pictures, but not every picture has the action of closing eyes, yawning or bowing. We only include those pictures with fatigue characteristics into the statistical scope.

driver number	Closed eye	yawn	nod	sum
1	1189	249	661	2099
2	588	144	338	1070
percentage	56.5%	11.6%	31.9%	100%

Table 2 Driver fatigue behavior statistics

Therefore, according to the percentage of each fatigue feature, we set the weights ,W_1, W_2, W_3 to 0.565, 0.116, and 0.319 respectively. Therefore, the calculation formula of the final sleepiness value is：

$$SleepyValue = 0.565 \times \text{PERCLOS} + 0.116 \times f(yawn) + 0.319 \times f(nod)$$

The value range of the final calculated value of $SleepyValue$ is within the $[0，1]$. Then, according to the driver's behavior, the fatigue level is divided into four levels: sober, mild fatigue, middle fatigue and severe fatigue. The higher the degree of fatigue, the larger the $SleepyValue$. The specific fatigue level classification is shown in Table 3：

$SleepyValue$	fatigue level	fatigued behavior
<0.3	sober	Three indicators fluctuate slightly
0.3-0.4	mild fatigue	Increased frequency of eye closure and occasional yawning
0.4-0.5	middle fatigue	PERCLOS value increased significantly, yawning and nodding
>0.5	sever fatigue	The three indicators increased significantly at the same time

Table 3 Fatigue rating

IV. EXPERIMENTAL RESULTS AND ANALYSIS

The laboratory environment of this article is a general laboratory environment, using PyCharm and Anaconda as development tools, and the programming language is Python. Finally, the YAWDD data set is selected as the experimental verification sample when the driver does not wear a mask. Because there is no data set of fatigue driving of drivers wearing masks, we adopted two approaches to solve this problem for the purpose of the experiment. One is to let the students in the lab wear masks to simulate the scene of fatigue driving. The second is to process some of the existing data set videos in YAWDD through editing software, and "put on masks" for drivers who do not wear masks through editing. As a result, a total of 30 pieces of video data were obtained, which were used for experimental verification samples when drivers wore masks. Part of the processed YAWDD dataset is shown in Figure 7：

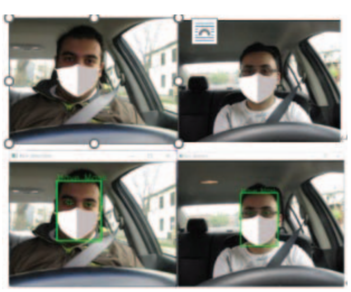

Fig 7 YAWDD dataset and processing demonstration after wearing a mask

In order to verify the accuracy of the algorithm, we selected five representative videos for experimental verification on the YAWDD data set and the self-made mask-wearing driving data set. The experimental results are shown in the table below：

Video number	closed eye frames	false detection	accuracy (%)
1	238	16	93.33
2	352	27	92.32
3	312	23	92.63
4	375	31	91.73
5	412	30	92.72

Table 4 Algorithm accuracy test when wearing a mask

Video number	1	2	3	4	5
PERCLOS	0.12	0.24	0.42	0.54	0.72
$f(yawn)$	0	0	0	0.64	0.81
$f(nod)$	0	0.17	0.38	0.36	0.63
$SleepyValue$	0.068	0.189	0.358	0.492	0.701
detection	sober	sober	mild	middle	severe
reality	sober	sober	mild	middle	severe

Table 5 Algorithm Accuracy Test When Not Wearing a Mask

ACKNOWLEDGMENT

This work was financially supported by the High Technology Key Program of Hubei Province of China (Grant No.202011901203001), and the Project of Youth Talent of Hubei Provincial Department of Education (Grant No.202011901301002),

.

REFERENCE

[1] Liu, W., et al., *SSD: Single Shot MultiBox Detector.* 2015.

[2] Mingyuan., Z. and L. Dinggen, *Research on fatigue detection method based on face feature points.* Application Research of Computers, 2020. **37**(S2): p. 305-307.(in Chinese).

[3] Singh, S.P., et al., *Face Mask Detection using Multi-Stage CNN Architecture.* SSRN Electronic Journal, 2021.

[4] Dollár, P., P. Welinder, and P. Perona, *Cascaded pose regression.* IEEE, 2010.

[5] Soukupova´, T. and J.C. Ech, *Real-Time Eye Blink Detection using Facial Landmarks.*

[6] Lucas, B.D. and T. Kanade. *An Iterative Image Registration Technique with an Application toStereo Vision*. in *Proceedings of the 7^th International Joint Conference on ArtificialIntelligence*. 1997.

[7] XQ., G., et al., *The implementation of human eye opening PERCLOS based on finite state automaton.* Application Research of Computers, 2014,31(1). **31**(1): p. 307-310.(in Chinese).

[8] Rahman, A., M. Sirshar, and A. Khan. *Real time drowsiness detection using eye blink monitoring*. in *Software Engineering Conference*. 2016.

[9] Zhou, E., et al. *Extensive Facial Landmark Localization with Coarse-to-Fine Convolutional Network Cascade*. in *IEEE International Conference on Computer Vision Workshops*. 2013.