A DISTRACTED DRIVING DISCRIMINATION METHOD BASED ON THE FACIAL FEATURE TRIANGLE AND BAYESIAN NETWORK

TIANLIU FENG[1], LINGXIANG WEI[2,3], WENJUAN E[1*], PENGFEI ZHAO[4], ZHE LI[1], YUCHUAN JI[3]

[1] *School of Rail Transportation, Soochow University, Suzhou, China*

[2] *School of Defense Engineering, Army Engineering University of PLA, Nanjing, China*

[3] *School of Materials Science and Engineering, Yancheng Institute of Technology, Yancheng, China*

[4] *School of Civil and Transportation Engineering, Beijing University of Civil Engineering and Architecture, Beijing, China*

Abstract. Distracted driving is one of the main causes of road crashes. Therefore, effective distinguishing of distracted driving behaviour and its category is the key to reducing the incidence of road crashes. To identify distracted driving behaviour accurately and effectively, this paper uses the head posture as a relevant variable and realizes the classification of distracted driving behaviour based on the relevant literature and investigation. A

---

* Corresponding author. E-mail: wje@suda.edu.cn

Tianliu FENG (ORCID ID 0000-0002-9863-3839)
Lingxiang WEI (ORCID ID 0000-0002-7117-0301)
Wenjuan E (ORCID ID 0000-0001-8502-6266)
Pengfei ZHAO (ORCID ID 0000-0001-7443-5857)
Zhe LI (ORCID ID 0000-0003-3211-366X)
Yuchuan JI (ORCID ID 0000-0002-4640-6685)

Tianliu Feng,
Lingxiang Wei,
Wenjuan E,
Pengfei Zhao,
Zhe Li,
Yuchuan Ji

A Distracted Driving Discrimination Method Based on the Facial Feature Triangle and Bayesian Network

distracted driving discrimination algorithm based on the facial feature triangle is proposed. In the proposed algorithm, the Bayesian network is employed to judge driving behaviour categories. The proposed algorithm is verified by experiments using data from 20 volunteers. The experimental results show that the discrimination accuracy of the proposed algorithm is as high as 90%, which indicates that the head posture parameters used in this study are closely related to the distracted driving state. The results show that the proposed algorithm achieves high accuracy in the discrimination and classification of distracted driving behaviour and can effectively reduce the accident rate caused by distracted driving. Moreover, it can provide a basis for the research of distracted driving behaviour and is conducive to the formulation of the corresponding laws and regulations.

Keywords: Bayesian network, distracted driving behaviour, facial feature triangle, feature point recognition, head posture, traffic safety.

Introduction

The road crash causing factors include human, vehicle, road, and environmental factors. Among human factors, distracted driving behaviour has been considered the leading factor causing road crashes (Née et al., 2019; Kidd & Chaudhary, 2019). Distracted driving generally refers to the driver behaviour phenomenon where driver attention is directed to activities that are not related to driving, resulting in the decline in the driving operation ability (Pope et al., 2017; Craig et al., 2021), and the distracted driving behaviour has a direct impact on the probability of collision (Shaaban et al., 2020). Due to the insufficient accumulation of distraction category data, it is challenging to prevent distracted driving behaviour effectively, and thus it is difficult to formulate relevant regulations to reduce distracted driving behaviour (Nevin et al., 2017). Therefore, the way of distinguishing and classifying drivers' distracted driving behaviour accurately, preventing distracted driving behaviour effectively, and avoiding road crashes caused by distracted driving successfully has been an urgent problem to be solved in the field of distracted driving (Tung & Khattak, 2015). Since distracted driving is difficult to be measured using quantitative physiological indicators and there have been no unified laws and regulations to supervise and restrict this type of driving behaviour, developing an accurate discrimination method of distracted driving behaviour could be an effective way to prevent distracted driving (Wei et al., 2021).

Parameter acquisition is the primary problem to solve when studying distracted driving behaviour, and selecting appropriate indicators is a precondition of parameter acquisition. Parameter acquisition methods can be roughly divided into invasive and non-invasive (Sun et al., 2017).

The invasive detection methods use electroencephalogram (EEG) and electrocardiogram (ECG). Although the invasive index is highly related to distracted driving, the contact between the equipment and a driver during the detection process greatly affects the driving state, which is not suitable for on-board equipment (Wang et al., 2015). The non-invasive detection methods consider the vehicle driving characteristics and external environmental parameters. These indicators can be detected without affecting the driver, but the detection accuracy has been difficult to improve (Siddiqui et al., 2021). With the development of the face recognition technology, the face feature index detection based on face recognition has not only ensured the non-invasive approach to drivers but has also provided the theoretical feasibility to improve the detection progress because of the close correlation between its detection parameters and distracted driving features (Lei et al., 2017). Therefore, this paper selects the eyes and mouth parameters as indicators because they can effectively represent the distracted driving characteristics as a detection index of face recognition and constructs a distracted driving discrimination method based on the facial feature triangle.

Related research has shown that the most common distracted driving behaviours include talking to passengers, smoking, and using mobile phones (Parr et al., 2016). Drivers' distracted driving behaviour categories present different risk levels (Fu et al., 2022), and during these behaviours, a driver's head posture changes significantly. In this study, the yaw, pitch, and roll angles are used to quantify the head posture, and the head posture is used to judge distracted driving behaviour. The Bayesian network model is employed to fuse prior knowledge and actual data and learn causality (Chen et al., 2018). The Bayesian network has been selected due to its advantages of stable classification efficiency and insensitivity to missing data, and the judgment of distracted driving behaviour is a typical classification and discrimination problem. Therefore, this paper proposes a distracting action recognition method based on the Bayesian network.

The contributions of this study are mainly reflected in two aspects. First, this paper establishes a head pose estimation method using facial feature triangles to estimate a driver's head pose accurately and effectively. Second, this paper uses the Bayesian network to establish a distracted driving behaviour discrimination model, which can effectively and accurately distinguish and classify distracted driving behaviour, so as to accumulate distracted category data and prevent distracted driving.

The rest of this paper is organised as follows. In Section 2, the current research and related studies are reviewed. Section 3 presents the index selection method, facial feature point recognition algorithm, and

Tianliu Feng,
Lingxiang Wei,
Wenjuan E,
Pengfei Zhao,
Zhe Li,
Yuchuan Ji

A Distracted Driving
Discrimination
Method Based on
the Facial Feature
Triangle and
Bayesian Network

distracted driving discrimination algorithm. The head pose estimation method is introduced in Section 4. The proposed distracted driving category judgment method is explained in Section 5. The analysis of experimental and prediction results is given in Section 6. The conclusions and future research areas are presented in Section 7.

1. Literature review

In this study, a facial feature triangle is constructed by using the information on the eyes and mouth of a driver to reflect the driver's head posture, and then the distracted driving state is distinguished by the Bayesian network.

1.1. Head pose estimation method

In recent years, many studies have been conducted on the driver's head posture detection. Fice et al. (2018) studied the driver's head deflection angle and duration under normal conditions and found that compared with the stationary state, during the movement of a vehicle, the driver's head attitude deflection amplitude was smaller and the duration of driver's head attitude deflection was shorter. Zhao et al. (2020) used the head posture as the evaluation parameter of distracted driving. The experimental results show that the distracted driving state can be effectively judged according to head posture. Yan et al. (2022) found that different combinations of head, hand and object positions constituted a complex category of driver posture. The research shows that it has continuity, diversity, superposition, similarity, transition and interaction. Teyfouri et al. (2021) designed a fatigue warning system based on the driver's neck position and blinking frequency based on the fact that the driver's head drooping drives the change in the neck position during sleep. He et al. (2015) established the evaluation model of driver fatigue using the EEG data and a driver's head nodding angle as detection indicators. The results showed that this method could effectively prevent driver fatigue during driving. Therefore, the distracted driving state can be distinguished based on the head state, so this paper selects the head posture as an evaluation index of the driving state.

In the non-invasive detection indicators, the eyes and mouth on a relatively stable face can be used as feature points to construct a facial feature triangle, which can reflect the driver's head posture (Ghimire et al., 2017). On this basis, the paper uses three parameters, namely, the yaw, pitch, and roll angles, to quantify the head deflection angle to improve the detection accuracy of Head Pose Estimation.

1.2. Facial feature point recognition algorithm

Based on the increasingly mature image classification and detection technology, the recognition methods of a driver's distracted driving face can be mainly divided into traditional computer vision (CV) algorithms and deep learning-based algorithms. Driver distraction detection based on the traditional CV algorithms extracts image features using the scale invariant feature transform (SIFT), histogram of oriented gradient (HOG), and other feature operators (Zhang, Tang, & He, 2019), then combines them with the support vector machine (SVM) and establishes a classification model. However, traditional CV algorithms have the disadvantages of high requirements for the environment, a narrow application range, numerous parameters, and a large amount of calculation. Convolutional neural networks (CNNs) have been proven to be the most effective technology to achieve high precision in face recognition (Zhang et al., 2019). With the rapid development of deep learning, CNNs have been applied to many computer vision tasks, such as image recognition and target detection (Ding & Tao, 2018). The test results of face recognition show that the performance of image recognition and detection could be significantly improved using the CNNs (Hu et al., 2019). The comparison of different recognition algorithms shows that the aggregating handcrafted and deep CNN features can make up for the deficiency of deep learning with higher accuracy (Alkinani et al., 2022). The recognition methods based on deep learning have attracted great research attention in recent years. By using an instrument panel camera in a car to record a driver's driving process and a pretrained neural network model to detect and recognise the captured image, extracting the local neighbourhood texture information of a grey image, it is found that the mouth and eyes contribute the most to facial expression, the deep learning-based algorithm can obtain a high recognition rate (Zhang & Hua, 2015). Therefore, this study uses a deep learning-based method to recognise a driver's face and mark the facial feature points.

In this study, the eyes and mouth are selected as facial feature points. At present, three types of methods have been commonly used for human eye positioning:

1. Feature-based methods, such as the projection method, which have fast processing speed but are greatly affected by the face pose transformation;
2. Shape-based methods, such as the template matching method, which can achieve accurate positioning, but usually include a large amount of calculation and have insufficient real-time performance;

Tianliu Feng,
Lingxiang Wei,
Wenjuan E,
Pengfei Zhao,
Zhe Li,
Yuchuan Ji

A Distracted Driving
Discrimination
Method Based on
the Facial Feature
Triangle and
Bayesian Network

3. Performance-based methods, such as the AdaBoost algorithm, which have strong robustness but often require a large number of training samples.

This paper combines the projection method and Hough transform to detect the circle, locate the eyeball accurately, and overcome their shortcomings. The combined algorithm improves the accuracy and speed of human eye positioning. In addition, on the basis of the eye positions, the mouth is located according to the distribution characteristics of facial organs.

1.3. Distracted driving behaviour discrimination method

The in-depth study of distracted driving has shown that the parameter curves of a distracting action often have similar morphological characteristics. Therefore, similar parameter curves can be classified, and then the description characteristics can be obtained according to the characteristics of the parameter curves and distracting action. By using multiple description features to express and describe the knowledge about distracting action, the morphological features of the parameter curve can be recognised, and then the distracting action can be determined. There are four main types of distracted driving: cognitive, visual, audio (Babić et al., 2021; van der Zwaag et al., 2012; Warren & Micha, 2011; Catalina et al., 2020) and manual distraction (Zhang & Hua, 2015). The Bayesian networks can learn causal relationships. Therefore, it is an ideal model to fuse a priori knowledge and data and realise reasoning under the condition of incomplete and uncertain information. It can learn from practice and optimize the network structure and parameters (Yang et al., 2010; Fasanmade et al., 2020). The Bayesian network has many advantages in classification. For instance, it uses a graphical method to describe the relationship between data, which is simple and efficient; it is beneficial to deal with incomplete datasets; it can deal with the causal relationship between variables. In addition, combining the Bayesian statistics can make full use of information on the domain knowledge and sample data (Ruitao, 2021). Therefore, in this study, the Bayesian network is used, and a distracting action recognition method based on the Bayesian network is proposed. First, parameters related to a distracting action in distracted driving data are divided into multiple morphological feature classes using the time series hierarchical clustering method based on the DTW (Distance to Waypoint) distance (Wan et al., 2017). Next, the descriptive characteristics of each parameter curve are determined by the method based on the statistical dependency analysis to distinguish various parameter sequences. Then, a Bayesian network for distracting action

recognition is constructed by fusing multiple description features (Liang & Lee, 2014). Finally, the recognition of distracting actions is realised by the Bayesian network reasoning.

2. Feature point recognition of human eyes and mouth

This section may be divided by subheadings. It should provide a concise and precise description of the experimental results, their interpretation, as well as the experimental conclusions that can be drawn.

2.1. CNN-based face recognition

The CNN model used in this paper consists of two convolution layers and two maximum pooling layers. To improve the accuracy of face recognition, in the proposed CNN model, the 5×1 and 1×5 convolution kernels are used instead of the 5×5 convolution kernel to reduce the parameter number and calculation amount and improve the accuracy. After being processed by the connection layer, the input is sent to the output layer, and the softmax function is used as a classification function in the classification process. The CNN model structure and calculation steps are shown in Figure 1.
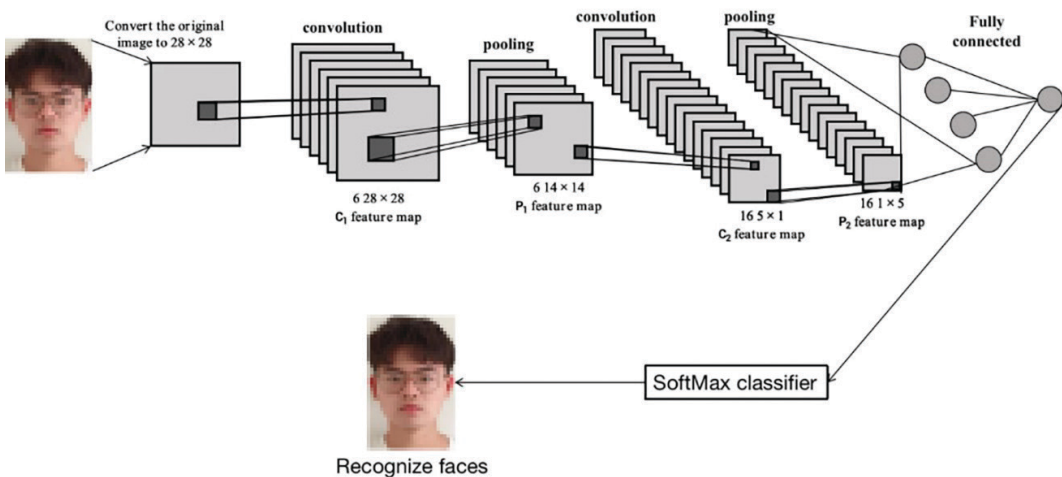

Figure 1. The CNN model structure

2.2. Eye feature point positioning based on Hough transform

In a face image, first, the face area is cropped. Then, the human eyes are roughly located through two integral projections, and the edge of the roughly located eye grey image is extracted after enhancement processing. The boundary of the eyeball region is obtained by processing the previous section. The idea of using the Hough transform is as follows. First, two points m and n on a circle, having the coordinates of x_m, y_m and x_n, y_n respectively, are recorded. The centre of the circle must be on the vertical line connecting the two points, and this vertical line is defined by:

$$(x-x_m)^2+(y-y_m)^2=(x-x_n)^2+(y-y_n)^2. \tag{1}$$

Then, a value of one is added to the accumulated value of all points on this vertical line in the parameter space. Finally, the point with the largest accumulated value in the parameter space is extracted. This point represents the centre of the detection circle. The detected centre is the position of the pupil, which denotes the position of the eye feature point. The positions of eye feature points are marked as B and C, as shown in Figure 2.
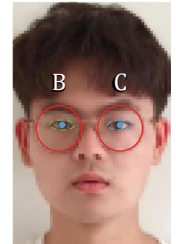

Figure 2. Recognition results of eye feature points

2.3. Mouth positioning

Once the eyes have been located, the mouth location can be obtained based on the eye locations. According to the distribution characteristics of facial organs, the mouth area can be roughly divided, as shown in Figure 3(a). Also, according to the distribution law of facial organs, it is assumed that the distance between two eyes is a; the size of the mouth area is $0.4a \times 1.3a$, and this area is located below the eyes at a distance of a. The experimental analysis has shown that in the HIS space, selecting H component can accurately distinguish the colour of skin and lip.

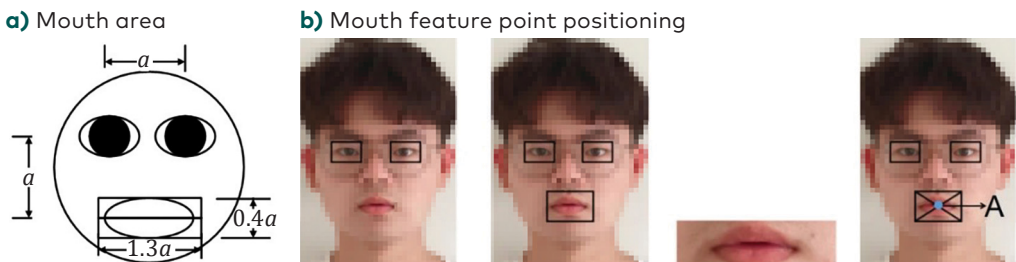**a)** Mouth area   **b)** Mouth feature point positioning

Figure 3. Mouth positioning.

Therefore, in the *H* component, after morphological image processing, the mouth area is selected by a rectangular frame, and the centre of the rectangle is taken as the centre of gravity of the mouth. The centre of the mouth represents the feature point of the mouth, which is marked as *A*. The mouth detection process is shown in Figure 3(b).

3. Facial feature triangle-based driving attitude estimation

In the head movement analysis, the eyes and mouth with a stable facial movement are selected as feature points to construct the facial feature triangle. Using the geometric changes in the facial feature triangle, the yaw, pitch, and roll angles of the head can be determined and used as parameters to infer the head posture. The changes in the yaw, pitch, and roll angles can be used to judge whether the driver is distracted.

Attitude parameter estimation mainly refers to calculating the deflection angle of the head relative to the three coordinate axes, namely, yaw, pitch, and roll, as shown in Figure 4. The vertical direction of a driver's head image is set as the *z*-axis, the horizontal direction is set as the *x*-axis, and the direction perpendicular to the driver's image is set as the *y*-axis.

In a video image, the relative position of eyes and mouth change with the change in the head posture, and they show certain geometric features. Therefore, as long as the positions of eyes and mouth are located, the head posture can be preliminarily estimated.

Assume the coordinates of feature points *A*, *B*, and *C* in the pixel coordinate system are denoted by (x_1, y_1), (x_2, y_2) and (x_3, y_3),
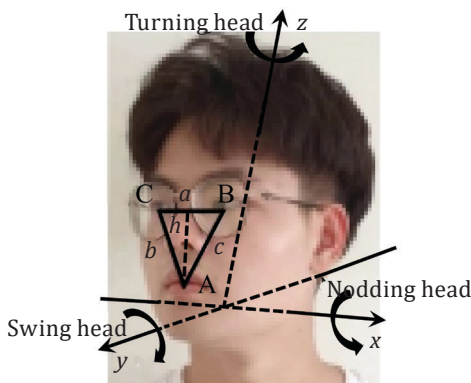

Figure 4. The head pose model

respectively. According to the geometric relationship characteristics of the human face, when the head is facing a camera, $\triangle ABC$ is an isosceles triangle, and when the head posture changes, the $\triangle ABC$ geometry in an image also changes. Therefore, the current head posture can be judged by analysing the $\triangle ABC$ geometric features.

According to the three vertex coordinates of the face feature triangle, the three side lengths of $\triangle ABC$ can be obtained as follows:

$$a = \sqrt{(x_3 - x_2)^2 + (y_3 - y_2)^2}, \tag{2}$$

$$b = \sqrt{(x_3 - x_1)^2 + (y_3 - y_1)^2}, \tag{3}$$

$$c = \sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2}. \tag{4}$$

The three sides a, b, and c of $\triangle ABC$ correspond to angles $\angle A$, $\angle B$ and $\angle C$, respectively, which are obtained as follows:

$$\angle A = \pi - \angle B - \angle C, \tag{5}$$

$$\angle B = \arccos\left(\frac{1}{2c}\left(\frac{1}{a}\left(c^2 - b^2\right) + a\right)\right), \tag{6}$$

$$\angle C = \arccos\left(\frac{1}{2b}\left(a - \frac{1}{a}\left(c^2 - b^2\right)\right)\right). \tag{7}$$

The height corresponding to side a of $\triangle ABC$ is given by:

$$h = \frac{2\sqrt{p(p-a)(p-b)(p-c)}}{a}, \tag{8}$$

where $p = (a + b + c)/2$.

3.1. Yaw attitude analysis

Yaw indicates that a driver's head rotates around the z-axis, as shown in Figure 5(a). As shown in Figure 5(b), in the video image (i.e., the *x-z* plane), $\triangle A_p B_p C_p$ is no longer an isosceles triangle, and its properties change. Further, as also presented in Figure 5(b), in which direction the head rotates, in that direction, the triangle top angle will become larger. For instance, if $\angle C > \angle B$, the head rotates to the right, but if $\angle B > \angle C$, turn the head rotates to the left. For the convenience of calculation, two triangles are projected onto the *x-y* plane, as shown in Figure 5(c).

Set the rotation angle as α; then, it holds that:

$$\alpha = \arccos\frac{a}{a'}, \tag{9}$$

where α denotes the distance between the eyes after turning the head; a' is the distance between the eyes in the frontal image of the head.

When turning the head, the side length of the feature triangle changes, but the corresponding height h stays unchanged. According to the positional relationship between the mouth and eyes presented in Figure 5(a), it holds that
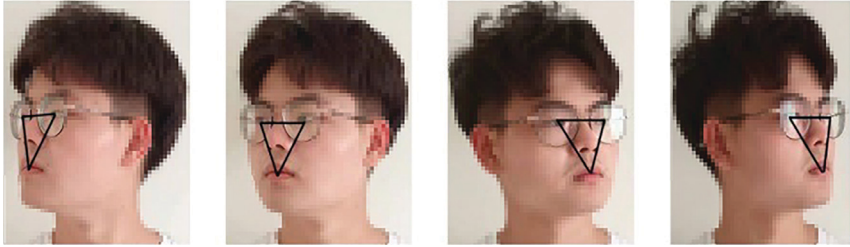
$$a' = \frac{h}{1.2}. \tag{10}$$

Therefore, it can be written as:
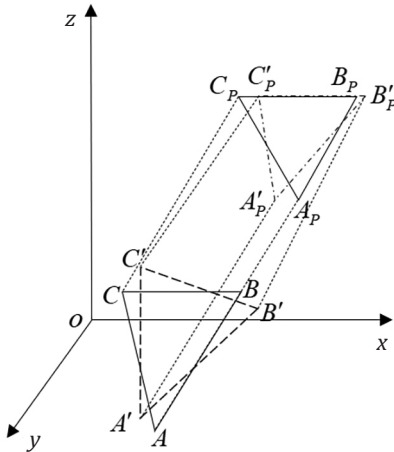
$$\alpha = \arccos\frac{1.2a}{h}. \tag{11}$$

3.2. Pitch attitude analysis

Pitching refers to the rotation of the head around the x-axis, as shown in Figure 6(a). In a video image (i.e., the *x-z* plane), as shown in Figure 6(b), *ABC* and 123 denote isosceles triangles, and the height h of

a) The head rotation model about the *z*-axis

b) The *x-z* plane corresponding to the video image

c) The projection of the two triangles onto the *x-y* plane
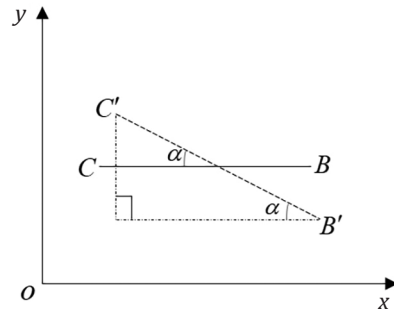

Figure 5. The head turn model.

the characteristic triangle changes. Set the pitch angle as BB. For the convenience of calculation, project two triangles onto the y-z plane, as shown in Figure 6(c).

It can be written as:

$$\beta = \arccos \frac{h}{h'}, \tag{12}$$

where h denotes the height of the triangle in the image after nodding, and h' is the height of the triangle in the frontal image of the head.

When nodding, although the coordinate positions of the eyes in an image are changing, the relative distance between the eyes remains unchanged, i.e., a remains unchanged. According to the previous analysis, for the front face image, the centre of the mouth should be located at $1.2a$ below the eyes, so it can be written as:
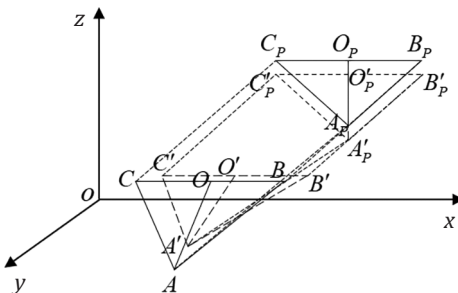
$$h' = 1.2a. \tag{13}$$

Therefore, it holds that

$$\beta = \arccos \frac{2\sqrt{p(p-a)(p-b)(p-c)}}{1.2a^2}. \tag{14}$$

a) The head rotates around the x-axis
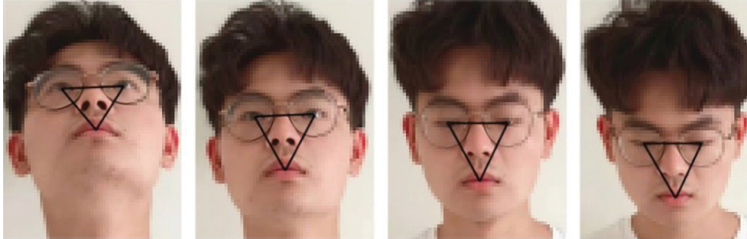
b) The x-z plane corresponding to the video image

c) The projection of the two triangles onto the y-z plane
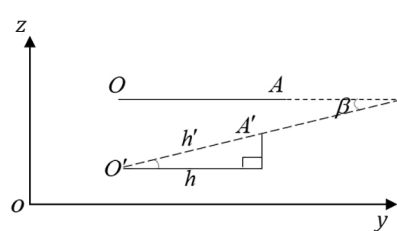

Figure 6. The nodding model

3.3. Roll attitude analysis

Tumbling means that the head rotates around the y-axis, as shown in Figure 7(a). In a video image (i.e., the x-z plane), as shown in Figure 7(b), $\Delta A_\mathrm{p}B_\mathrm{p}C_\mathrm{p}$ and $\Delta A'_\mathrm{p}B'_\mathrm{p}C'_\mathrm{p}$ denote isosceles triangles, and the distance between the two eyes remains unchanged. The only change in the image after the head rotation around the y-axis is that the height h of $\Delta A_\mathrm{p}B_\mathrm{p}C_\mathrm{p}$ is no longer parallel to the z-axis, but forms a certain angle with the z-axis, which is denoted by γ. For the convenience of calculation, the two triangles are projected onto the x-z plane, as shown in Figure 7(c).
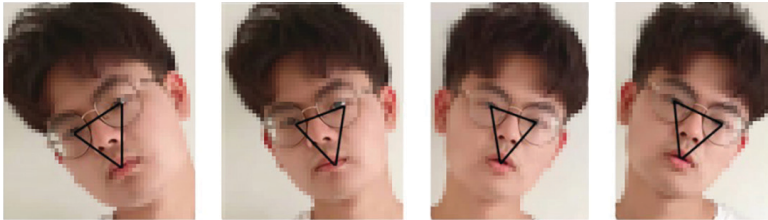
Considering the actual situation, the motion range of the head swing is small. Define the motion range in the image pixel coordinate system as $x_3 < x_1 < x_2$. Then, if $y_2 > y_3$, the head deviates to the left, the deflection angle γ is given by:

$$\gamma = \arccos\left(\frac{y_2 - y_1}{c}\right) - \frac{1}{2}\angle C, \tag{15}$$
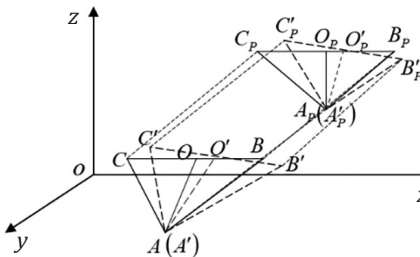
If $y_2 < y_3$, the head deviates to the right, and the deflection angle γ is given by:

$$\gamma = \arccos\left(\frac{y_3 - y_1}{c}\right) - \frac{1}{2}\angle C. \tag{16}$$

a) The head rotates around the y-axis

b) The x-z plane corresponds to the video image

c) The projection of the two triangles onto the y-z plane
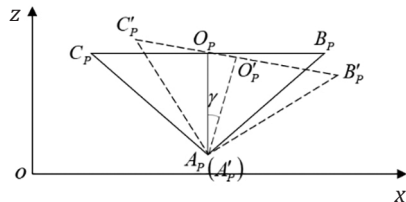

Figure 7. The head swing model

4. Bayesian network-based distracted driving category judgment method

4.1. Morphological and description feature node construction

Descriptive and morphological feature nodes refer to the first and second-layer nodes in the Bayesian network model, respectively. These nodes are the key to the Bayesian network construction. They include the classification of parameter sequences, as well as the discretization and selection of descriptive features.

4.1.1. Parameter sequence classification based on hierarchical clustering

This paper adopts condensed hierarchical clustering; it starts from a single distracted driving behaviour sample, then merges smaller distracted driving behaviour samples and, finally, forms a distracted driving behaviour category containing all samples. In this study, relevant parameter sequences of simulated distracted driving behaviours are clustered based on the DTW distance. According to the clustering situation, the distance threshold is set for classification, and multi-class parameter sequences are obtained as morphological feature nodes. Fifteen typical distracted driving behaviours, including visual, cognitive, operational, and auditory distractions, are considered, as shown in Table 1.

Table 1. Fifteen typical distracted driving behaviours

Distraction type	Behaviour type	Behaviour description
Visual distraction	Behaviour A	Checking a mobile phone
	Behaviour B	Looking at the dashboard
	Behaviour C	Looking out a window of a vehicle
	Behaviour D	Looking at a pedestrian on the roadside
Cognitive distraction	Behaviour E	Thinking about things unrelated to driving
Operational distraction	Behaviour F	Tuning instruments (radio, air condition, or navigation)
	Behaviour G	Lighting a cigarette
	Behaviour H	Doing makeup or looking in the mirror
	Behaviour I	Eating food or drinking the water
	Behaviour J	Phone
	Behaviour K	Sending a message
	Behaviour L	Combing the hair
	Behaviour M	Reaching out to the back
Auditory distraction	Behaviour N	Talking to passengers
	Behaviour O	Listening to the radio and music

4.1.2. Description features and their discretization

According to the specific analysis of the distracted driving behaviour and related parameters, the description characteristics of a parameter sequence are determined. The description characteristics and their descriptions are shown in Table 2. The discretization results of describing the features are shown in Table 3. The continuous quantity uses the heuristic Gaussian cloud algorithm to mine the qualitative concept of each description feature of distracted driving behaviour and divides the interval with the intersection of the expected curve.

Table 2. Description features and their descriptions

Nature	Descriptive feature	Description
Continuous quantity	Absolute cumulative variation	Sum of absolute variation values
	Cumulative variation	Sum of changes of parameter
	Maximum value	Maximum absolute value of a parameter
	Maximum difference	Difference between the maximum and minimum values of a parameter
	Duration	Time needed to complete an action
	Changing process	The overall changing trend of a parameter

Table 3. The characteristics and their discretization

Parameter	Descriptive feature	Discrete interval / Changing process
Yaw angle	Absolute cumulative variation	(0, 20), [20, 50), [50, 90), [90, 135), [135, 180), [180, 270), [270, +∞)
	Cumulative variation	(−∞, −5), [−5, 5), [5, +∞)
	Maximum	(0, 60), [60, 90), [90, 135), [135, 180)
	Maximum difference	(0, 30), [30, 45), [45, 60), [60, 90), [90, 135), [135, 360)
Pitch angle	Absolute cumulative variation	(0, 70), [70, 157), [157, 203), [203, 282), [282, +∞)
	Cumulative variation	(−∞, −45), [−45, 45), [45, +∞)
	Maximum	(0, 15), [15, 30), [30, 45), [45, 60), [60, 90)
	Maximum difference	(0, 45), [45, 90), [90, 180)
Roll angle	Absolute cumulative variation	(0, 160), [160, 200), [200, 340), [340, +∞)
	Cumulative variation	(0, 27), [27, 140), [140, +∞)
	Maximum	(0, 30), [30, 45), [45, 75)
	Maximum difference	(0, 90), [90, 150)
Public	Duration	(0, 2), [2, 4.5), [4.5, 7), [7, +∞)

Tianliu Feng,
Lingxiang Wei,
Wenjuan E,
Pengfei Zhao,
Zhe Li,
Yuchuan Ji

A Distracted Driving
Discrimination
Method Based on
the Facial Feature
Triangle and
Bayesian Network

The obtained interval is used for node discretization of the description features of distracted driving behaviour. The change in the continuous quantity corresponds to the overall changing trend of a parameter without considering small fluctuations in data.

4.1.3. Node selection of description feature class based on statistical dependency statistical analysis

The dependency relationship is usually determined by the mutual information or conditional mutual information of two nodes, which indicates the correlation degree of the two nodes. The dependency relationship between the feature nodes and morphological feature nodes can be analysed and described by statistical or information theory methods.

Let us assume that X and Y represent two nodes, having the values of x and y, respectively. Then, the mutual information between nodes X and Y is given by:

$$I(X,Y) = \sum_{x,y} p(x,y) \log \frac{p(x,y)}{p(x)p(y)}. \qquad (17)$$

The greater the value of $I(X,Y)$ is, the stronger the dependency between nodes X and Y is, i.e., the greater the correlation between them is. If $I(X,Y)$ is less than the set threshold ξ, nodes X and Y are considered independent. Therefore, Equation (5) has been usually used for the conditional independence test of nodes.

The statistical dependency analysis is used to calculate the mutual information I between a description feature class node and a morphological feature class node of the corresponding parameter. By setting the threshold ξ, the description feature class node with a strong dependency is selected to construct the Bayesian network.

The calculation results of mutual information between the description feature nodes of the Bayesian network and their corresponding morphological feature nodes are shown in Table 4. In this study, the threshold is set as $\xi = 0.5$ according to the actual demand. The description feature nodes of mutual information $I > \xi$ are selected as morphological feature nodes, and the selection results are shown in Table 5.

Table 4. Calculation results of mutual information between the nodes

Description class feature node	Morphological feature class node		
	Yaw angle	Pitch angle	Roll angle
Absolute cumulative variation	1.314	0.919	1.095
Cumulative variation	0.788	0.297	0.211
Maximum value	1.189	0.961	0.989
Maximum difference	0.710	0.787	0.845
Duration	0.755	0.578	0.514

Table 5. The results of the feature class nodes

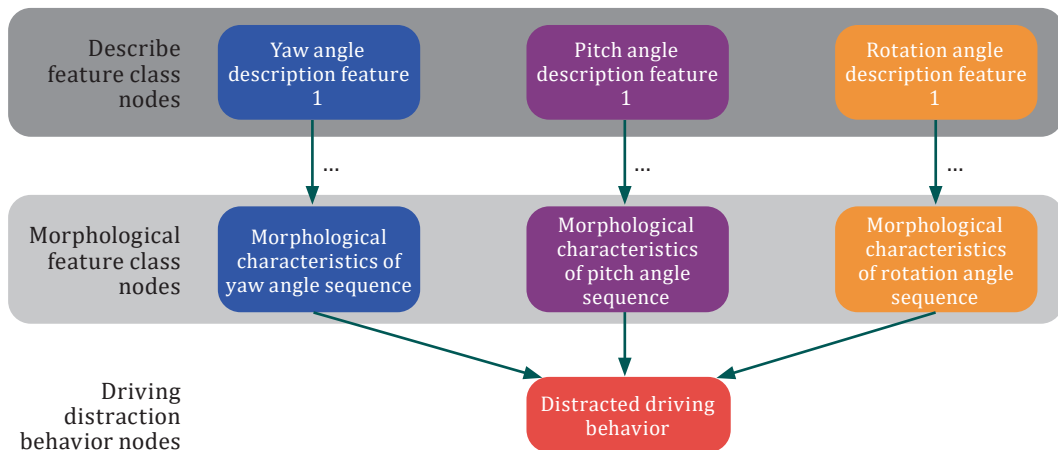| Morphological feature class node	Feature class node description
Yaw angle	Absolute cumulative variation, maximum value, maximum difference, duration
Pitch angle	Absolute cumulative variation, maximum value, maximum difference, duration variation process, duration
Roll angle	Absolute cumulative variation, maximum value, maximum difference, duration

Figure 8. The structure of the Bayesian network model used for driving behaviour recognition

Tianliu Feng,
Lingxiang Wei,
Wenjuan E,
Pengfei Zhao,
Zhe Li,
Yuchuan Ji

A Distracted Driving
Discrimination
Method Based on
the Facial Feature
Triangle and
Bayesian Network

4.2. Bayesian network-based distracted driving behaviour recognition method

4.2.1. Bayesian network construction

The Bayesian network model used in the proposed driving behaviour recognition method includes three layers, namely, description feature layer, morphological feature layer, and driving behaviour layer, as shown in Figure 8. The description feature layer is the first layer, and it includes the description features of each head pose parameter. The morphological feature layer is the second layer, and it indicates the morphological characteristics of each parameter sequence, including analogue quantity information and switching quantity information. The driving behaviour layer is the third layer, and it represents various distracted driving behaviours.

4.2.2. Identification process

The driving behaviour recognition process based on the Bayesian network includes two main steps, the Bayesian network learning and Bayesian network reasoning, as shown in Figure 9.
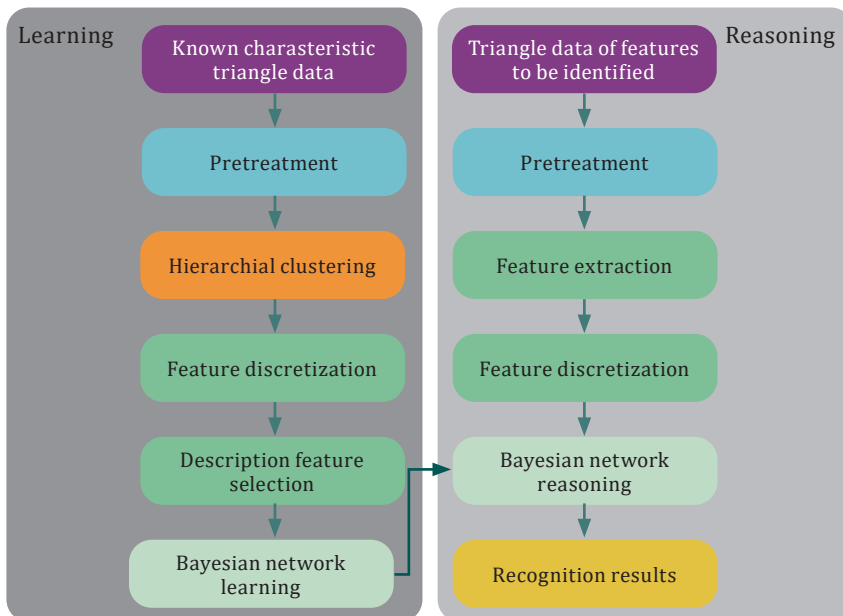

Figure 9. The block diagram of the driving behaviour recognition system

The Bayesian network reasoning is based on the conditional probability calculation when the network model is known. Through the analysis and application of known conditions, the target node probability is calculated. The change characteristics of each parameter are extracted from the driving data to be identified and then discretized and input into the Bayesian network. The probability obtained in the Bayesian learning process is used for a priori probability calculation to infer the probability that the input data belongs to the corresponding action.

5. Case study and discussion

The characteristic physiological data of the driver were obtained using the CLT-353 simulation driving platform. The driving device had the same operation mode as the real car and was equipped with a force feedback system, which could simulate the real driving environment. In the experiment, the triangle index data of the driver's facial features were obtained by an infrared anti-shake camera.

To test the actual detection accuracy of the proposed method, 20 volunteers of different genders were selected to participate in the test. The volunteers had different driving experiences. There was no significant difference in the relationship between executive functions and distracted driving behaviours (Zhang et al., 2020); thus, these volunteers were classified into the age groups of 18–30 (5 male and 3 female), 30–45 (3male and 3 female), and 45–70 (4 male and 2 female). The volunteers were assigned numbers from the range of 1–20, and the simulation test was performed on the above-mentioned driving simulator.

5.1. Facial feature triangle-based head pose detection

To test the accuracy of the facial feature triangle used for head pose estimation, the following tests were conducted. The camera was placed at a fixed position G; the seat was fixed in front of the camera; the intersection between the front of the camera and the top of the seat was denoted as point O; points A, B, C, A', B', C' were marked on the left and right sides of the camera head; the angles of OA, OB, OC, OA', OB', OC', and OG were 20°, 30°, 40°, 20°, 30°, and 40°, respectively. Similarly, points D, E, F, D', E', F' were denoted at angles of 20°, 30°, and 40° above and below of the camera; strip marks were denoted at 20°, 30°, and 40° on the left and right of the longitudinal axis. During the test, the volunteers only turned their heads, while pupil, sight, and body remained unchanged, and looked at each mark in turn. For each volunteer, data of 100 rounds

Tianliu Feng,
Lingxiang Wei,
Wenjuan E,
Pengfei Zhao,
Zhe Li,
Yuchuan Ji

A Distracted Driving
Discrimination
Method Based on
the Facial Feature
Triangle and
Bayesian Network

were collected, summarised, and sorted, and the average error was calculated by:

$$\bar{X} = \frac{\sum\limits_{i=1}^{10}(|x_i - \varepsilon|)}{10}, \tag{18}$$

where $\bar{X}$ represents the average error of a data reading, x_i represents the ith reading at the same position, and ε represents the angle marked by the fixation mark point.

The head pose estimation was performed using the proposed facial feature triangle algorithm. The target angle of gaze was compared to evaluate the accuracy of the facial feature triangle-based head pose estimation.
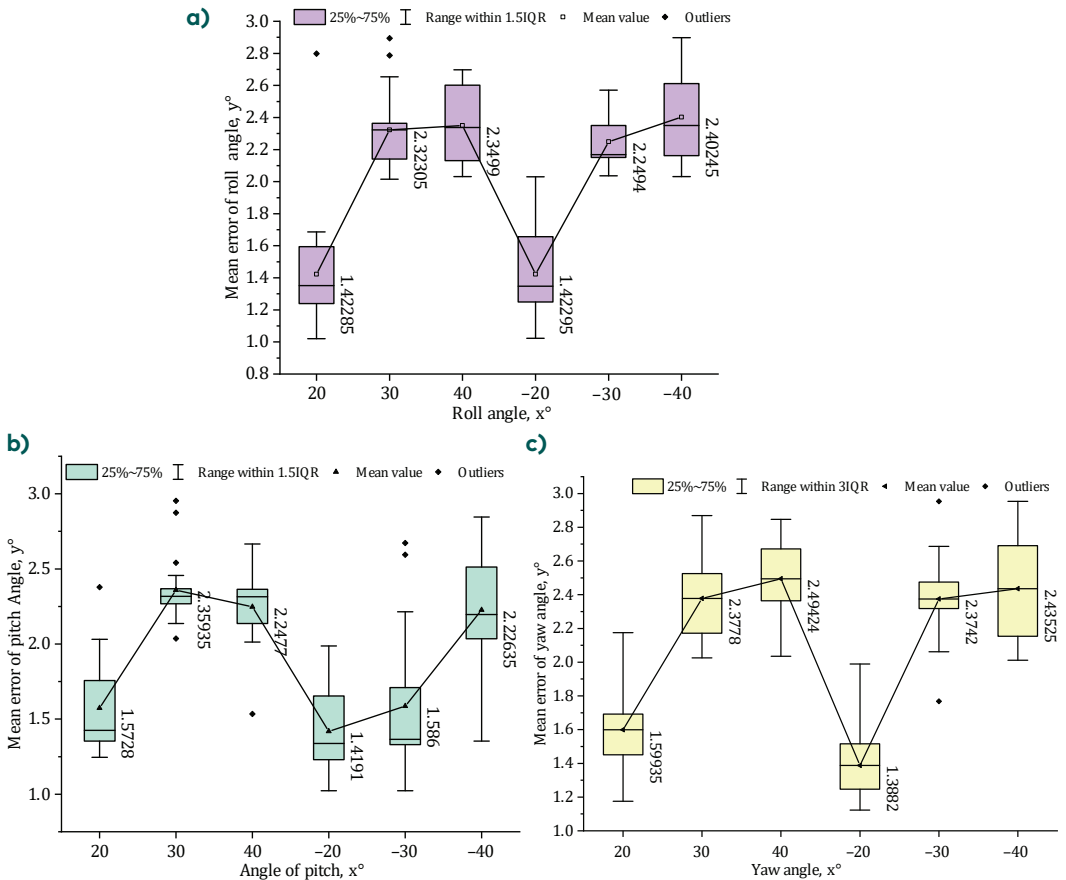
Figure 10. The mean errors obtained based on the facial feature triangle: (a) roll angle; (b) pitch angle; (c) yaw angle

Integrating all experimental results, the absolute value of the angle error was calculated to be within 3°. The test results for different angles are shown in Figure 10. As shown in Figure 10, the head pose estimation result was in line with the actual situation, and high accuracy with a small error range was achieved. Unlike the traditional attitude estimation, the proposed algorithm does not require information on the initial attitude, is easy to implement, and has a faster calculation speed.

5.2. Facial feature triangle-based distracted driving detection

To test the distracted driving detection accuracy of the proposed algorithm, the following experiments were conducted. Volunteers were asked to drive normally and make some distracted driving behaviours randomly during the driving process, such as looking at a mobile phone, lighting a cigarette, eating food, and taking certain things. The video images of the drivers' driving state were collected and edited. Each video lasted 3–5 s. Humans judged the driver's driving state to determine whether the driver is in a distracted driving state and in what kind of distracted driving state. For each driver, 100 segments were selected, and the proposed algorithm was used to distinguish each of the segments. The correct output of the driving state discrimination was denoted by "1" and the wrong output was denoted by "0". The sum value M of the judgment results was obtained by:

$$P_n = \frac{M_n}{100}, \tag{19}$$

where P_n is the discrimination accuracy of a driver n, and M_n is the sum of the judgment results of the driver n.

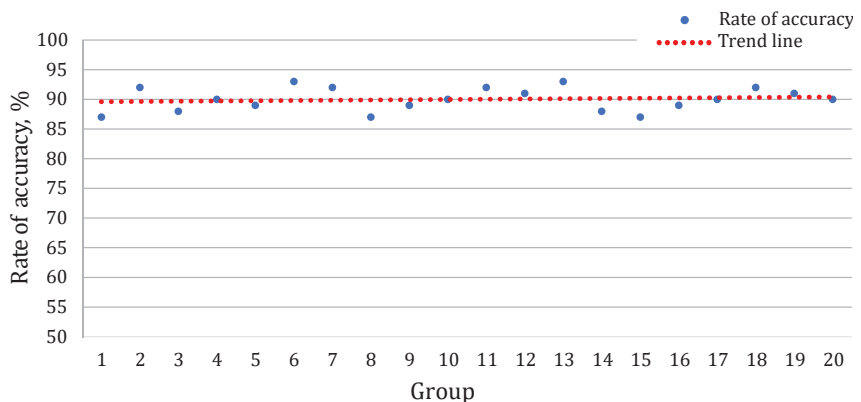The detection accuracy of 20 drivers is shown in Figure 11.

Figure 11. The overall detection accuracy of 20 drivers

As shown in Figure 11, the distraction discrimination algorithm based on the facial feature triangle had high recognition accuracy, which could provide a guarantee for the smooth and accurate operation of the system and ensure the feasibility of subsequent research.

5.3. Bayesian network-based distracted driving behaviour judgment

To verify the accuracy and effectiveness of the proposed judgment method of the distracted driving category based on the Bayesian network, the simulation tests were conducted using the driving simulator. In the early Bayesian network parameter learning, 10 volunteers were randomly selected from 20 volunteers to simulate 15 typical distracted driving behaviours. Each behaviour included 10 groups of video data, so a total of 1500 groups of sample videos were collected in the experiment. The Bayesian network model was constructed using the Netica software developed by Norsys company in Canada. For each action, 100 groups of data were preprocessed, and feature sequences were extracted. The feature class nodes were discretized, and the parameters were optimized combined with the EM algorithm to construct the CPT to realise the parameter learning of the Bayesian network. Taking the yaw angle model as an example, after the CPT was applied, the Bayesian network yaw angle model for distracting action recognition was obtained, as shown in Figure 12.

The driving data of another 10 drivers were selected as test samples. Each driver performed 10 distracted driving actions randomly in the
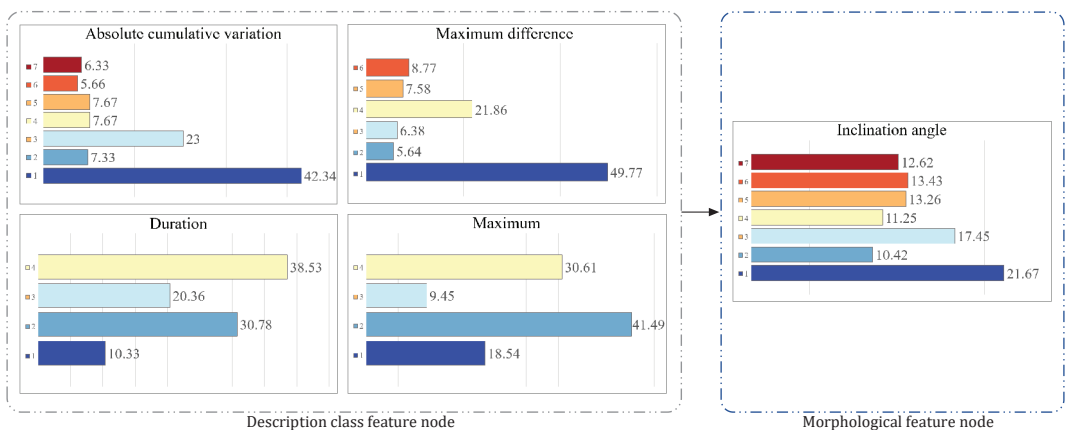

Figure 12. The Bayesian network model used for describing the characteristic nodes and morphological characteristic nodes of the yaw angle

test; 100 groups of test samples were collected and used as the input of the Bayesian network model. According to the edge probability of class nodes and the conditional probability of characteristic nodes determined by the learning process of the Bayesian network, the Bayesian reasoning was performed. The probability of class nodes of distracted driving behaviour was used as a basis for judging the distracted driving action type; namely, the distracted driving action corresponding to the place with the largest probability was considered the recognition result. The results of the first and fourth groups of test samples are shown in Figure 13. As displayed in Figure 13, the probabilities of behaviours J and F were the largest, so it was judged that the action recognition results of distracted driving behaviour of the first and fourth groups of test samples were action J (phone) and F (adjusting instruments such as radio, air condition, or navigation), respectively. In the experiment, for 91 groups of 100 test samples, the distracted driving behaviour was correctly judged, and the results of only 11 groups were wrong; thus, the discrimination accuracy of 91% was achieved. Therefore, the proposed solution could achieve high discrimination accuracy of distracted driving behaviour.

5.4. Discussion

This experiment verified the feasibility of distinguishing distracted driving based on facial feature triangles. In addition, compared with other classification models, because of the advantages of Bayesian network classification efficiency stability and insensitivity to missing data, this experiment has good performance in judging distracted
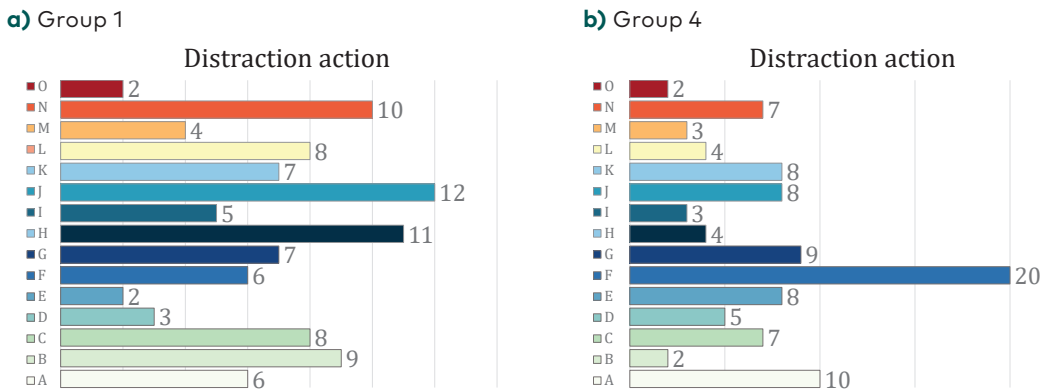
a) Group 1

b) Group 4

Figure 13. Probabilistic reasoning results

Tianliu Feng,
Lingxiang Wei,
Wenjuan E,
Pengfei Zhao,
Zhe Li,
Yuchuan Ji

A Distracted Driving
Discrimination
Method Based on
the Facial Feature
Triangle and
Bayesian Network

driving behaviour. However, this experiment also has some limitations. For example, the size of samples collected is relatively small, and the age of participants is not divided in detail, which is also the scope of our next research.

Conclusions

Distracted driving has always been one of the crucial factors causing road crashes. This driving behaviour is related to drivers' driving habits, road environment, and other related factors. When analysing the distracted driving of a driver, this paper takes the head posture as a detection index. Based on the real driving data of 20 drivers, a driver distraction discrimination and classification model is developed. The image sensor is used for data acquisition, and the Bayesian network is used to identify the distracted driving behaviour category. The proposed method adopts a non-contact detection index, which can distinguish and classify distracted driving behaviour without affecting the driver's normal driving. The experimental results show that the proposed method has good recognition performance and can identify and classify a driver's distracted state accurately and rapidly. The results confirm that the head posture parameters are closely related to the distracted driving state. Compared with other methods, Bayesian network has the characteristics of stable efficiency and insensitivity to missing data, so it can more accurately distinguish the distracted driving behaviour category. The proposed method can provide a reliable data basis for standardizing the driver driving behaviour, thus preventing the driver distracted driving, and help define relevant laws and regulations.

In the future, the proposed method framework could be improved. Namely, for small or difficult targets in images or videos of the human head, the weight in the proposed method could be adjusted appropriately to improve the overall recognition effect of the method. In addition, depth images could be used to estimate the head posture. The datasets used in this study include the RGB images. Although such training data are easy to obtain and suitable for extensive research, in the process of feature training, the RGB-D images could provide sufficient feature information and more dimensional information. Namely, the RGB-D images have more advantages than ordinary images under extreme conditions and can meet the requirements under strict accuracy requirements.

Funding

REFERENCES

Aljasim, M., & Kashef, R. (2022). E2DR: A deep learning ensemble-based driver distraction detection with recommendations model. *Sensors*, *22*, Article 1858. https://doi.org/10.3390/s22051858

Alkinani, M. H., Khan, W. Z., Arshad, Q., & Raza, M. (2022). HSDDD: A hybrid scheme for the detection of distracted driving through fusion of deep learning and handcrafted features. *Sensors*, *22*(5), Article 1864. https://doi.org/10.3390/s22051864

Babić, D., Babić, D., Sucha, M., Stanić, V., & Toman, M. (2021). The influence of music genres on the driving behaviour of young drivers and their visual scanning of the environment. *Transportation Research Part F: Psychology and Behaviour*, *81*, 396–407. https://doi.org/10.1016/j.trf.2021.07.001

Catalina, C.A., García, H.S., Cabrerizo, E., Herrera, S., García, P.S., Mohamadi, F., & Mariscal, M.A. (2020). Music distraction among young drivers: Analysis by gender and experience. *Journal of Advanced Transportation*, *2020*, Article 6039762. https://doi.org/10.1155/2020/6039762

Chen, C., Chen, Y. Y., Ma, J. M., Zhang, G. H., & Walton, C. M. (2018). Driver behaviour formulation in intersection dilemma zones with phone use distraction via a logit-Bayesian network hybrid approach. *Journal of Intelligent Transportation Systems*, *22*(4), 311–324. https://doi.org/10.1080/15472450.2017.1350921

Craig, L., Steve, B., Ward, V., & Robyn, R. (2021). Prevalence and trends of distracted driving in Canada. *Journal of Safety Research*, *76*, 118–126. https://doi.org/10.1016/j.jsr.2020.12.005

Ding, C. X., & Tao, D. C. (2018). Trunk-branch ensemble convolutional neural networks for video-based face recognition. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *40*(4), 1002–1014. https://doi.org/10.1109/TPAMI.2017.2700390

Fasanmade, A., He, Y., Al-Bayatti, A. H., Morden, J. N., Aliyu, S. O., Alfakeeh, A. S., & Alsayed, A. O. (2020). A fuzzy-logic approach to dynamic Bayesian severity level classification of driver distraction using image recognition. *IEEE Access*, *8*, 95197–95207. https://doi.org/10.1109/ACCESS.2020.2994811

Fice, J. B., Blouin, J. S., & Siegmund, G. P. (2018). Head postures during naturalistic driving. *Traffic Injury Prevention*, *19*(6), 637–643. https://doi.org/10.1080/15389588.2018.1493582

Tianliu Feng,
Lingxiang Wei,
Wenjuan E,
Pengfei Zhao,
Zhe Li,
Yuchuan Ji

A Distracted Driving
Discrimination
Method Based on
the Facial Feature
Triangle and
Bayesian Network

Fu, X., Meng, H., Wang, X., Yang, H., & Wang, J. (2022). A hybrid neural network for driving behaviour risk prediction based on distracted driving behaviour data. *PLoS ONE*, *17*(1), Article e0263030. https://doi.org/10.1371/journal.pone.0263030

Ghimire, D., Lee, J. W., Li, Z. N., & Jeong, S. H. (2017). Recognition of facial expressions based on salient geometric features and support vector machines. *Multimedia Tools and Applications*, *76*(6), 7921–7946. https://doi.org/10.1007/s11042-016-3428-9

He, Q. C., Li, W., Fan, X. M., & Fei, Z. M. (2015). Driver fatigue evaluation model with integration of multi-indicators based on dynamic Bayesian network. *IET Intelligent Transport Systems*, *9*(5), 547–554. https://doi.org/10.1049/iet-its.2014.0103

Hu, C.H., Zhang, Y., Wu, F., Lu, X.B., Liu, P., & Jing, X.Y. (2019). Toward driver face recognition in the intelligent traffic monitoring systems. *IEEE Transactions on Intelligent Transportation Systems*, *21*(12), 4958–4971. https://doi.org/10.1109/TITS.2019.2945923

Kidd, D. G., & Chaudhary, N. K. (2019). Changes in the sources of distracted driving among Northern Virginia drivers in 2014 and 2018: A comparison of results from two roadside observation surveys. Journal of Safety Research, 68, 131–138. https://doi.org/10.1016/j.jsr.2018.12.004

Lei, J. M., Han, Q.W., Chen, L., Lai, Z. D., Zeng, L. Q., & Liu, X. (2017). A novel side face contour extraction algorithm for driving fatigue statue recognition. *IEEE Access*, *5*, 5723–5730. https://doi.org/10.1109/ACCESS.2017.2686424

Liang, Y. L., & Lee, J. D. (2014). A hybrid Bayesian network approach to detect driver cognitive distraction. *Transportation Research Part C*, *38*, 146–155. https://doi.org/10.1016/j.trc.2013.10.004

Née, M., Contrand, B., Orriols, L., Gil-Jardiné, C., Galéra, C., & Lagarde, E. (2019). Road safety and distraction, results from a responsibility case-control study among a sample of road users interviewed at the emergency room. *Accident Analysis and Prevention*, *122*, 19–24. https://doi.org/10.1016/j.aap.2018.09.032

Nevin, P. E., Blannar, L., Kirk, A. P., Freedheim, A., Kaufman, R., Hitchcock, L., Maeser, J. D., & Ebel, B. E. (2017). "I wasn't texting; I was just reading an email…: a qualitative study of distracted driving enforcement in Washington State. *Injury Prevention*, *23*(3), 165–170. https://doi.org/10.1136/injuryprev-2016-042021

Niloufar, T., Hossein, S., & Alireza, S. (2021). Designing a glass mounted warning system to prevent drivers to fall in sleep based on neck posture and blinking duration. *Journal of Medical Signals and Sensors*, *11*(3), 217–221. https://doi.org/10.4103/jmss.JMSS_31_20

Parr, M. N., Ross, L. A., McManus, B., Bishop H. J., Wittig S. M. O, & Stavrinos, D. (2016). Differential impact of personality traits on distracted driving behaviours in teens and older adults. *Accident Analysis and Prevention*, *92*, 107–112. https://doi.org/10.1016/j.aap.2016.03.011

Pope, C. N., Bell, T. R., & Stavrinos, D. (2017). Mechanisms behind distracted driving behaviour: The role of age and executive function in the engagement of distracted driving. *Accident Analysis and Prevention*, *98*, 123–129. https://doi.org/10.1016/j.aap.2016.09.030

Ruitao, S. (2021). Driver intention prediction using model-added Bayesian network. *Proceedings of the Institution of Mechanical Engineers*, *235*(5), 1236–1244. https://doi.org/10.1177/0954407020968967

Shaaban, K., Gaweesh, S., & Ahmed, M. M. (2020). Investigating in-vehicle distracting activities and crash risks for young drivers using structural equation modeling. *PLOS ONE*, *15*(7), Article e0235325. https://doi.org/10.1371/journal.pone.0235325

Siddiqui, H.U.R., Ali, S.A., Robert, B., Bahattin, B., Ernesto, L., Furqan, R., & Sandra, D. (2021). Non-invasive driver drowsiness detection system. *Sensors*, *21*(14), Article 4833. https://doi.org/10.3390/s21144833

Sun, W., Zhang, X. R., Peeta, S., He, X. Z., & Li, Y. F. (2017). A real-time fatigue driving recognition method incorporating contextual features and two fusion levels. *IEEE Transactions on Intelligent Transportation Systems*, *18*(12), 3408–3420. https://doi.org/10.1109/TITS.2017.2690914

Teyfouri, N., Shirvani, H., Shamsoddini, A. (2021). Designing a glass mounted warning system to prevent drivers to fall in sleep based on neck posture and blinking duration. *Journal of Medical Signals and Sensors*, *11*(3), 217–221. https://doi.org/10.4103/JMSS.JMSS_31_20

Tung, L.W., & Khattak, A. (2015). Distracted motor vehicle driving at highway-rail grade crossings. *Transportation Research Record*, *2476*(1), 77–84. https://doi.org/10.3141/2476-11

van der Zwaag, M. D., Dijksterhuis, C., de Waard, D., Mulder, B. L. J. M., Westerink, J. H. D. M., & Brookhuis, K. A. (2012). The influence of music on mood and performance while driving. *Ergonomics*, *55*(1), 12–22. https://doi.org/10.1080/00140139.2011.638403

Wan, Y, Chen, X. L., & Shi, Y. (2017). Adaptive cost dynamic time warping distance in time series analysis for classification. *Journal of Computational and Applied Mathematics*, *319*, 514–520. https://doi.org/10.1016/j.cam.2017.01.004

Wang, Y. K., Jung, T. P., & Lin, C. T. (2015). EEG-based attention tracking during distracted driving. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, *23*(6), 1085–1094. https://doi.org/10.1109/TNSRE.2015.2415520

Warren, B., & Micha, K. (2011). Exploring an alternative in-car music background designed for driver safety. *Transportation Research Part F: Psychology and Behaviour*, *15*(2),162–173. https://doi.org/10.1016/j.trf.2011.12.001

Wei, L. X., Feng, T. L., Zhao, P. F., & Liao, M. J. (2021). Driver sleepiness detection algorithm based on relevance vector machine. *The Baltic Journal of Road and Bridge Engineering*, *16*(1),118–139. https://doi.org/10.7250/bjrbe.2021-16.518

Yan, X. T., He, J., Wu, G., Zhang, C., & Wang, C. (2022). A proactive recognition system for detecting commercial vehicle driver's distracted behaviour. *Sensors*, *22*, Article 2373. https://doi.org/10.3390/s22062373

Tianliu Feng,
Lingxiang Wei,
Wenjuan E,
Pengfei Zhao,
Zhe Li,
Yuchuan Ji

A Distracted Driving
Discrimination
Method Based on
the Facial Feature
Triangle and
Bayesian Network

Yang, G. S., Lin, Y. Z., & Bhattacharya, P. (2010). A driver fatigue recognition model based on information fusion and dynamic Bayesian network. *Information Sciences*, *180*(10), 1942–1954. https://doi.org/10.1016/j.ins.2010.01.011

Zhang, H. L., Tang, L. M., & He, C. J. (2019). A variational level set model for multiscale image segmentation. *Information Sciences*, *493*, 152–175. https://doi.org/10.1016/j.ins.2019.04.048

Zhang, Y., & Hua, C. J. (2015). Driver fatigue recognition based on facial expression analysis using local binary patterns. *Optik*, *126*(23), 4501–4505. https://doi.org/10.1016/j.ijleo.2015.08.185

Zhang, Y., Li, J. J., Guo, Y. H., Xu, C. N., Bao, J., & Song, Y. P. (2019). Vehicle driving behaviour recognition based on multi-view convolutional neural network with joint data augmentation. *IEEE Transactions on Vehicular Technology*, *68*(5), 4223–4234. https://doi.org/10.1109/TVT.2019.2903110

Zhang, Z., Guo, Y., Fu, R., Yuan, W., & Wang, C. (2020). Linking executive functions to distracted driving, does it differ between young and mature drivers? *PLOS ONE*, *15*(9), Article e0239596. https://doi.org/10.1371/journal.pone.0239596

Zhao, Z. P., Xia, S. L., Xu, X. Z., Zhang, L., Yan, H. L., Xu, Y., & Zhang, Z. X. (2020). Driver distraction detection method based on continuous head pose estimation. *Computational Intelligence and Neuroscience*, *2020*(4), Article 9606908. https://doi.org/10.1155/2020/9606908