CrossMark

Research on key technologies of intelligent transportation based on image recognition and anti-fatigue driving

Jun Wang[1], Xiaoping Yu[3*], Qiang Liu[1,2] and Zhou Yang[1]

Abstract

Intelligent transportation system needs to solve the main problems in traffic safety. This paper focuses on the traffic safety caused by fatigue driving based on image recognition of key technologies for research and analysis. This paper proposes that the location of face and facial feature points and the classification of fatigue detection are the key links to determine the fatigue driving detection rate. In the analysis of face localization algorithm based on skin color modeling, a corner-based optimization method is proposed to optimize the face region. Based on the analysis of the binary algorithm of human eye localization algorithm, a bi-directional integral projection method is proposed to achieve accurate human eye localization. Then the commonly used fatigue classification algorithm (KNN algorithm) is analyzed. Finally, the proposed method is verified by the simulation test of fatigue driving. Experimental results show that the algorithm based on skin color modeling can accurately locate the driver's face region. The eye location algorithm based on the two-valued algorithm can also locate the eye location of the tester accurately. The accuracy of KNN fatigue detection model is 87.82%. It can identify driver's fatigue state with high accuracy.

Keywords: Image recognition, Fatigue driving, Intelligent traffic

1 Introduction

At the same time of large-scale urban expansion, the reform of infrastructure construction and management mode is relatively lagging behind, resulting in "urban disease" becoming more and more serious. The explosive growth of urban population and the rapid increase of the number of vehicles in the city have led to urban traffic obstacles and development bottlenecks. The main obstacles and problems are as follows [1]: serious urban traffic congestion, resulting in increased travel time and consumption of large amounts of energy, serious traffic safety problems, and frequent accidents; noise pollution and air pollution are becoming increasingly serious. Traffic safety is one of the main problems in the development of urban transportation, and it needs to be solved in time [2]. In the global human casualty accidents, traffic casualties are one of the main causes of human casualties. Statistics show that in 2016, 86.443 million traffic accidents occurred in

China, resulting in 63,093 deaths and 1.21 billion yuan of direct property losses; [3] in May 2017, traffic network data show that 787 traffic accidents occurred in Huai'an section of Beijing-Shanghai Expressway in 2016, including 414 traffic accidents caused by fatigue driving, and it accounts for about 52.6% of the total accident. Thus, fatigue driving is the main cause of major traffic accidents, so real-time monitoring of driver fatigue state has important practical significance in reducing traffic accidents and casualties.

In order to solve the problem of traffic safety, many countries in the world have given comprehensive consideration to the driving process, vehicle scheduling, and the overall control safety of vehicle operation. Intelligent transportation system (ITS) [4] emerged and developed continuously. Intelligent transportation system (ITS) makes full use of Internet of Things, cloud computing, Internet, artificial intelligence, automatic control, mobile Internet, and other technologies in the field of transportation. It collects traffic information through high tech and manages traffic, transportation, public travel, and

* Correspondence: yuxiaoping@bjtu.edu.cn
[3]School of Architecture and Design, Beijing Jiaotong University, Beijing, China
Full list of author information is available at the end of the article

other traffic areas in all aspects as well as the entire process of traffic construction and management to support the management, so that the transport system in the region, the city, and even a larger space-time range with the perception, interconnection, analysis, prediction, control, and other capabilities could fully protect traffic safety, play the effectiveness of transport infrastructure, and enhance transportation system operation efficiency and management level, for smooth public travel and sustainable economic development services. At present, the intelligent transportation system has also been widely applied. For example, [5] the intelligent traffic prediction system (ITPS) in Singapore consists of a computerized traffic signal system, an electronic scanning system, a city expressway monitoring system, a joint electronic eye, and a road pricing system to predict traffic flow over a predetermined period of time. It can help traffic controllers to predict traffic flow and prevent traffic congestion. Stockholm, Sweden, has introduced a new intelligent toll system, which reduces traffic by 22% and emissions by 12 to 40%. The goal of intelligent transportation system is to improve transportation efficiency, ease traffic congestion, improve the capacity of road network, and reduce traffic accidents through the harmony and close cooperation of people, vehicles, and roads. At present, there are many researches on intelligent transportation system. For example, Zhang [6] and others analyzed the architecture of intelligent transportation system and gave the overall framework, system functions, database structure, and the best path analysis method of Luoyang intelligent transportation system. Xie [7] and others put forward an intelligent urban traffic system based on the Internet of Things, which uses the technology of group intelligence perception to realize information collection and uses radio and television technology, mobile phone technology, and vehicle network technology to realize information sharing. Wang [8] and others have analyzed the key technologies of the existing intelligent transportation system and pointed out the problems that need to be solved urgently and the prospects of its research. However, there are few studies on intelligent traffic technology for fatigue driving-caused traffic safety. Because fatigue driving is the main cause of traffic accidents, it is necessary to study the technology of fatigue driving in intelligent transportation.

In the fatigue driving detection methods, there are mainly based on the driver's physiological signal detection, based on the driver's operation behavior and vehicle state detection, and based on the driver's facial expression detection. Most of these tests rely on image processing technology to get driver's fatigue characteristic data. Therefore, this paper analyzes and studies the key technology of anti-fatigue driving based on image recognition and improves the common key technology. Finally, the eye fatigue data of drivers during driving are collected through experiments. The improved key technologies are applied to verify the effectiveness of the technology.

2 Method

2.1 Intelligent transportation system

Intelligent transportation system (ITS) [9] integrates advanced technologies such as Internet of Things, big data, cloud computing, and wireless sensor and makes people, cars, and roads more coordinated, making public transport services more humane and intelligent travel service system. It covers railways, highways, civil aviation, and other fields. As the internal management system of each field is relatively mature, the problem to be solved by ITS is how to integrate information within multiple platforms, analyze potential data after mining the data, and provide users with better service [10]. In intelligent transportation systems, pedestrians, surrounding traffic lights, cameras, vehicle signs, and other infrastructures are connected as sensing terminals to form an urban road network information system. The terminals are identified intelligently through radio frequency identification (RFID), GPS, and infrared induction lights, and a continuous exchange of information is carried out according to certain agreements (Fig. 1).

2.2 Key technologies of intelligent transportation

The framework of intelligent transport [11] results is shown in Fig. 2.

The intelligent traffic sensor layer [12] is mainly responsible for collecting data such as two-dimensional code or bar code, which is read by intelligent identification equipment, and the network layer is mainly responsible for transmitting data information, which transfers data collected from each point to the data center by means of Internet, wireless network, or mobile network. Support layer mainly realizes parallel processing and optimization of massive information and dynamic allocation and deployment of storage resources; application layer mainly includes information storage and processing system and integrated control system. The system involves the collection of large data, the storage of large data, and the integration, processing, and mining of different types of data. Therefore, a large number of technologies are used to complete the work of each layer in the intelligent transportation system. The key technologies are:

1. Intelligent recognition and wireless sensing technology

Intelligent recognition and wireless sensing technology is the most important technology for identifying and sensing objects and is the foundation of the whole intelligent transportation construction. Intelligent identification reads the unique barcode, two-dimensional code, or RFID tag of an item through intelligent devices. By reading these electronic

Fig. 1 Schematic diagram of intelligent traffic
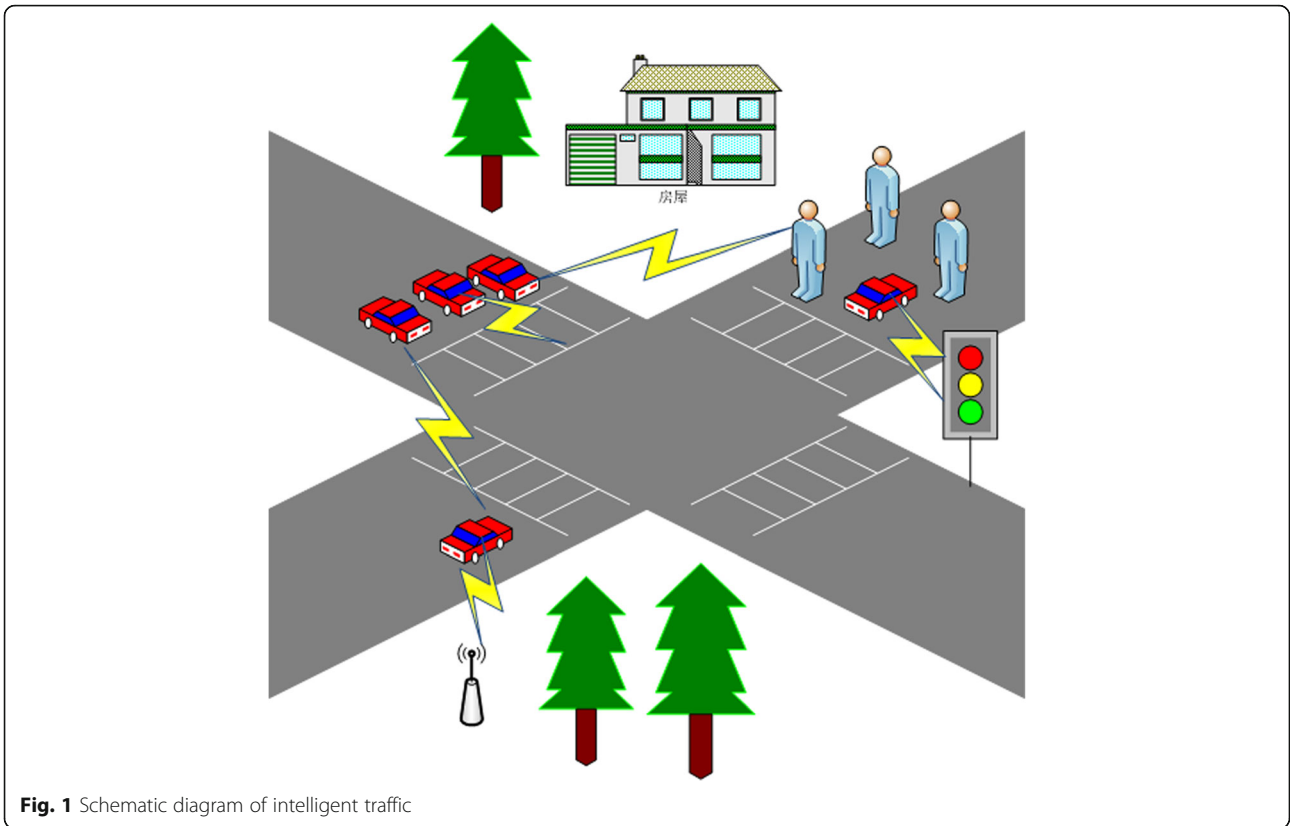
tags, it reads the unique features and location information of the item and then transmits these information to the upper system for recognition and final decision-making. In intelligent transportation network, each information collection point is equivalent to a set node in wireless sensor. They are responsible for the collection and processing of traffic environment information and then sent to other nodes or aggregation nodes; aggregation nodes will receive the information of each node after fusion processing and then transmitted to the next level. As the underlying network of the Internet of Things, wireless sensor networks provide a more secure, reliable, and sensitive solution for intelligent transportation.

Fig. 2 System diagram of intelligent transportation system
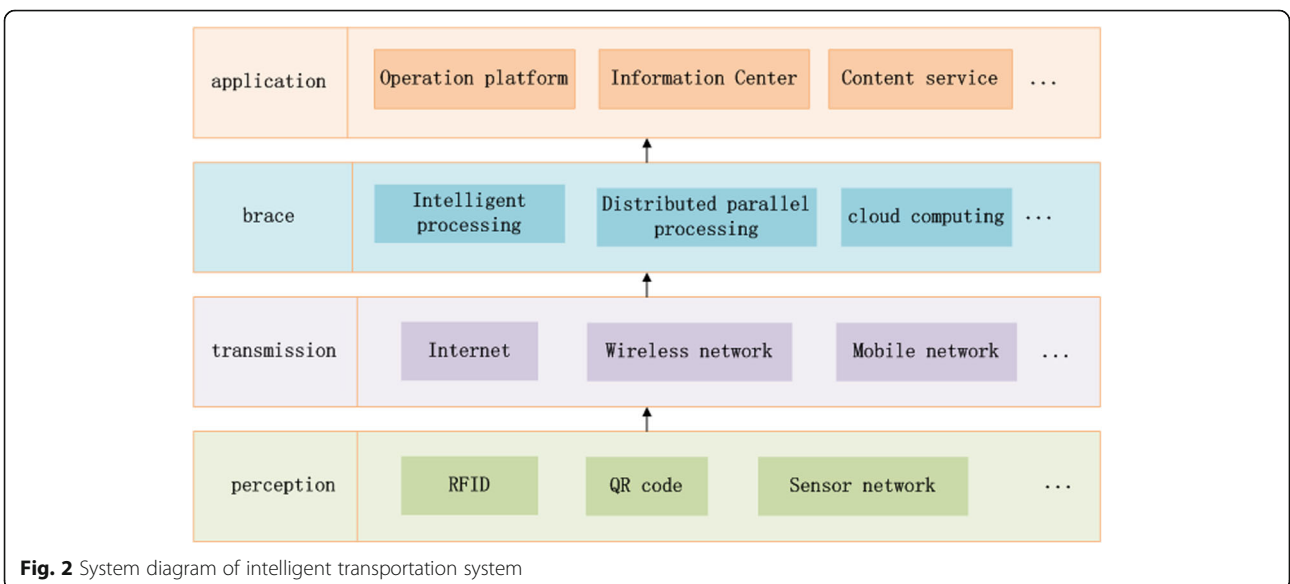
2. Distributed storage technology for big data

In the field of intelligent transportation, the whole system is in an independent state of information, and the data are difficult to transmit each other. Therefore, intelligent transportation system through cloud computing technology forms intelligent transportation for traffic data management. Intelligent transportation cloud makes full use of the advantages of cloud computing, such as mass storage, information security, and unified processing of resources, and provides a new solution for data sharing and effective management in the field of transportation.

3. Data processing technology

Because of the large amount of data in intelligent transportation and the diversity and heterogeneity of data at the same time, in addition, data processing often requires to be in real time and accurate. Therefore, intelligent transportation uses data fusion, data mining, data activation, data visualization, and other data processing technologies. Data fusion technology is a comprehensive data processing technology involving artificial intelligence, communication, decision-making, and other fields. It can detect, communicate, and analyze multi-source information from three levels: data layer, feature layer, and decision-making layer. Data activation is a new data organization and processing technology, which has the ability of storage, mapping, computing, and so on. It can evolve independently with the changes of objects and adapt to its own data reorganization with user behavior.

4. Application technology of image intelligent analysis

Because there are a lot of video images and other data in ITS, image intelligent analysis technology is used to process video image data in ITS. Intelligent image analysis and processing technology uses intelligent neural network technology to separate useful people or objects from video images by layered processing. With the help of the powerful data processing function of the computer, this technology can analyze the video image data quickly and filter out the redundant information. Automatic analysis and extraction of key information in video source will provide useful information to monitor. For example, based on the image recognition technology, the passing data can be used to recognize the license plate number, vehicle brand, and so on. In order to search for pictures, we can intercept vehicle characteristics to search for vehicles. By analyzing the driver's video, we can judge whether the driver is tired or not.

2.3 Image recognition technology

Image recognition is a basic human intelligence, widely used in people's daily life. With the rapid development of computer technology and electronic technology, computer can process image in real time, and the efficient image processing algorithm and image recognition technology occupy an important position in the intelligent transportation system. Image recognition technology is a research direction of artificial intelligence. Image recognition technology is based on the main characteristics of the image [13]. In the process of image recognition, the image must be preprocessed, the redundant information of the image should be removed, and the key information (i.e., features) should be extracted. Then the classifier can be obtained by classifying the training samples. Then classify and recognize the recognized image.

The specific image recognition process is shown in Fig. 3.

2.3.1 Image preprocessing advanced

The images obtained by the camera, scanner, and other acquisition devices are inputted to the computer. In order to improve the recognition rate, image processing is the first step. Preprocessing can effectively reduce the information without recognition value in the image and eliminate the noise and redundant information in the original image. The general preprocessing steps include grayscale, two-valued, denoising, image segmentation and so on.

- Image grayscale

Because the content of color image palette is complex, many image processing algorithms cannot be processed,
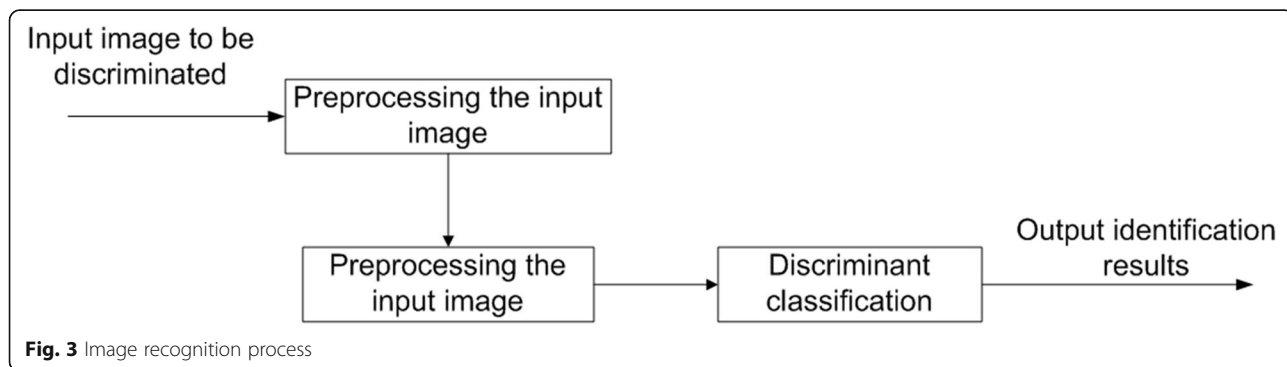

Fig. 3 Image recognition process

so it is necessary to process the color image grayscale [14]. The so-called grayscale image value of R, G, and B components of each pixel of the image is equal. In general, the grayscale of an image is to weigh the R, G, and B components of the image to get the final gray value. The common methods are the average method, the maximum method, and the weighted average method. The average method calculates the average brightness of each pixel R, G, and B and takes the average brightness as the gray value of the pixel, that is:

$$R = G = B = (R + G + B)/3 \quad (1)$$

The maximum rule is to use the maximum brightness value in the three color components of R, G, and B as the gray level of the pixel, that is:

$$R = G = B = \max(R, G, B) \quad (2)$$

The weighted average method assigns different weights to R, G, and B according to the sensitivity of human eyes to the three color components. Each pixel R, G, and B is weighted by three color values to get the average value, that is:

$$R = G = B = \mathrm{WRR} + \mathrm{WGG} + \mathrm{WBB} \quad (3)$$

Among them, WR, WG, and WB represent the weights of three colors of R, G, and B respectively.

- Two values of grayscale image

The two values of [15] are that the pixels in the image are divided into two colors according to certain criteria. Generally, there are adaptive thresholds and a given threshold method. The basic principle of the threshold method is to set the image $f(x, y)$, the gray level range is $[z1, z2]$, and select a suitable gray value t between $z1$ and $z2$, $t \in [z_1, z_2]$

$$f(x, y) = \begin{cases} 0 & g(x,y) < T \quad \text{background} \\ 1 & g(x,y) \geq T \quad \text{object} \end{cases} \quad (4)$$

where (x, y) is the coordinates of pixels and T is the threshold. $f(x, y)$ is a grayscale image processed by two values. $g(x, y)$ is the original grayscale image. When $g(x, y) < T$, the value of $f(x, y)$ is 0, and this point is set as the background image point. When $g(x, y) \geq T$, the value of $f(x, y)$ is 1, and this point is set as the target image point. Thresholding is the key to two thresholding. When T is too large, the occasional object will be considered as the background. If T is too small, it will mix with noise. In the adaptive threshold method, iterative threshold selection is used to binarize the image according to the selected threshold. The object is black and the background is white.

- Image denoising

Image denoising is filtering processing. Because the image denoising process to maintain the authenticity of the original image itself, so the selection of filtering methods is based on the noise category to determine. Generally, median filtering, adaptive median filtering, and mean filtering are used for filtering.

- Image segmentation

In the research and application of image, people often separate and extract the specific and unique region from the applied image. Image segmentation is the technology and process of dividing the image into different regions and extracting the object of interest. The features here can be grayscale, color, texture, and so on. At present, for complex images, such as remote sensing images and CT images, the traditional image segmentation algorithm segmentation effect has been unable to meet, so the application of neural network in image segmentation appears. Generally, current image segmentation algorithms can be divided into neural network algorithm based on pixel data and neural network algorithm based on feature data.

2.3.2 Image feature extraction

Image feature is the distinguishing feature of different images. Different targets can be distinguished by image features. Extracting image features is to find out the features belonging to the image itself from the substitution matching image to complete the matching with the template image. The features of [16] include color, texture, shape, and spatial features. There are different ways to extract each feature. For example, color moments, color histograms, and color correlation diagrams are commonly used to extract color features. Statistical methods, geometric methods, and model methods are commonly used for shape features.

2.3.3 Image classification

Image classification [17] is the process of dividing a set of metrics into different kinds of tags. Image classification is the core of pattern recognition. Image recognition and classification should be based on specific situations, using different classifiers. The commonly used classification method is statistical method. It also includes supervised classification method and unsupervised classification method. Supervised classification method [18] calculates the distribution of each class in the feature space according to the training samples with known class names in advance and then classifies the unknown data with it. Common supervised learning algorithms are regression analysis and statistical

classification. The most typical algorithms are KNN and SVM. Unsupervised classification method [19] is an exploratory analysis. It does not rely on pre-defined classes or training instances with class markers. It needs clustering learning algorithm to automatically determine the classification markers. Unsupervised learning methods can be divided into two categories: one is the direct method based on the estimation of probability density function, which means trying to find the distribution parameters in the feature space and then classifying them. The other is a simple clustering method called similarity measure between samples: the principle is to try to identify the core or initial kernel of different categories, and then clustering samples into different categories according to the similarity measure between samples and core. Clustering results can be used to extract hidden information from data sets and classify and predict future data.

2.4 Fatigue driving detection technology based on image recognition

2.4.1 Fatigue driving detection method and process

The ultimate goal of intelligent transportation is to use the Internet of Things, cloud computing, Internet, artificial intelligence, automatic control, mobile Internet, and other technologies to serve the smooth public travel and sustainable economic development. Fatigue driving has become one of the most important causes of traffic accidents worldwide, so real-time detection of drivers in the state of fatigue driving, and in the case of fatigue driving, giving an effective early warning is of great assistance to the establishment of intelligent transportation. From the current technology, the detection methods of fatigue driving are mainly divided into three categories [20]: detection methods based on physiological indicators, detection methods based on driver behavior characteristics analysis, and detection methods based on facial expression recognition.

2.4.1.1 Detection methods based on physiological indexes The detection method based on physiological index adopts the method of contact measurement [21]. Generally, the driver's fatigue state can be inferred by testing the driver's physiological signals. This method is mainly used in the laboratory or simulated driving environment as the control group of other fatigue detection methods.

**2.4.1.2 Detection method based on driver's behavior characteristics** Fatigue detection method based on driver behavior characteristics [22] infers driver fatigue state by analyzing driver's steering wheel, pedal operation characteristics, or vehicle trajectory characteristics. This method can achieve a certain recognition accuracy, and the measurement process will not cause interference to the driver. But the driver's operation is not only related to fatigue, but also affected by road environment, driving speed, personal habits, operating skills, and so on. Its accuracy and robustness still need to be improved.

2.4.1.3 Detection method based on facial expression Fatigue detection method based on facial expression is to use machine vision as a means [23] to collect the driver's facial image by image sensor, and judge the fatigue state by analyzing the driver's facial expression characteristics. This method is a relatively mature and widely applied technology. This method can detect the driver's fatigue by taking the driver's eye and mouth characteristics as fatigue characteristics. The related information of eye features is a widely used fatigue feature. The main steps of judging driver fatigue state based on machine vision are shown in Fig. 4.

- Face detection

The purpose of face detection is to separate the region of the face from the image, thus reducing the computational complexity of subsequent processing. Through the analysis and processing of the input image, it uses the method of knowledge or statistical learning to model the face, compares the matching degree between all possible regions to be detected and the face model, and

Fig. 4 Driving fatigue detection process based on machine vision
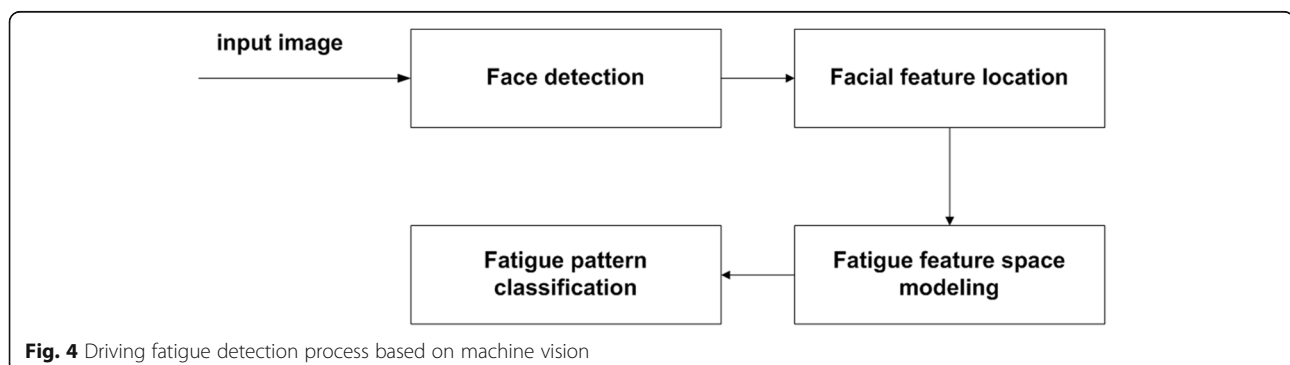
determines whether there is a face in the image. If there is, it outputs the information of the position, size, and position of the face.

- Position of facial feature points

The location of facial feature points is to search the position, key points, or contours of some or all facial features in a given area of the image. For example, eye detection is used to locate the eyes. The driver's binocular position is extracted from the face image. Eye location can be located by binarization and gray integration projection or face edge feature.

- Feature space modeling

Fatigue feature space modeling is a process of extracting parameters with significant differences under different fatigue degrees by analyzing a large number of samples. According to the results of key feature points location, each part of the image is segmented to extract the fatigue characteristics. The fusion of these features constitutes the feature space in the fatigue detection model. For example, if fatigue is detected by eye features, PERCLOS and pupil area are calculated according to eye closure.

- Fatigue pattern classification

Fatigue pattern classification is a process in which a suitable classifier is designed according to the distribution of feature points in the fatigue feature space, and the classifier is generalized for on-line identification of other drivers'fatigue states. Generally, the driver's mental state is classified (such as wakefulness, mild fatigue, fatigue, moderate fatigue, etc.). These states constitute the state space of fatigue mode. At this point, a suitable classifier is designed to map the feature space and state space to classify the feature space.

In the whole process of fatigue detection, the location of facial feature points and the classification of fatigue detection are the key links to determine the fatigue detection rate. Therefore, for the process of fatigue monitoring based on eye features, improving the face (eye) tracking algorithm and fatigue detection classification algorithm is the key technology of fatigue driving detection based on machine vision.

2.4.2 Key technologies of fatigue driving detection
Face localization and tracking algorithm

Location and tracking of driver's face area is the premise and basis for fatigue information extraction. Therefore, reliable face detection and tracking is the first solution to fatigue driving recognition. Most of the face detection algorithms are based on skin color modeling algorithm. The main steps of algorithm [24] based on skin color modeling are shown in Fig. 5.

The choice of color space

Because the color distribution of RGB color space is not continuous, RGB color space is not suitable for establishing skin color model. Terrillon et al. compared the results of face detection in nine different color spaces and found that Tint-Saturation-Luma (TSL) color space is most suitable for skin color segmentation under Gaussian model (including single Gaussian and mixed Gaussian), while YCbCr color space and TSL color space have perceptual consistency, so YCbCr color is often chosen in practical applications of color space for skin color modeling. After the skin color is modeled, the illumination compensation is preprocessed.
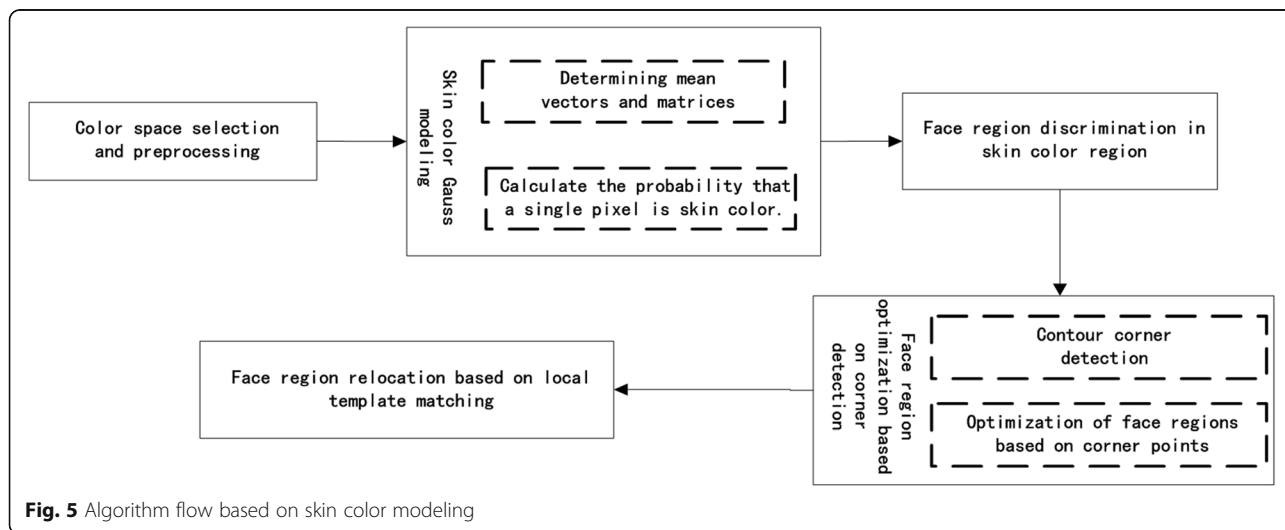

Fig. 5 Algorithm flow based on skin color modeling

Skin color Gauss modeling

Constructing skin color Gauss modeling and calculating the probability that pixels belong to skin color are as follows:

1. Determine the two parameters of the two dimensional Gauss model G (m, c) (mean vector m and covariance matrix C).

$$m = (\overline{C}_b, \overline{C}_r) \qquad (5)$$

Among them, $\overline{C}_b = \frac{1}{N}\sum\limits_{i=1}^{N} C_{bi}$, $\overline{C}_r = \frac{1}{N}\sum\limits_{i=1}^{N} C_{ri}$, $C = \text{cov}(C_b,$ $C_r)$, and N is the number of pixels used for Gauss modeling.

2. Calculate the probability that a single pixel is skin color.

The formula for calculating probability is:

$$p(x/\text{skin}) = \exp\left[-0.5(x-m)^T C^{-1}(x-m)\right] \qquad (6)$$

M is the mean value of random vectors. C is a covariance matrix. $x = (CbCr)^r$, $C = E(x - m)(x - m^r)$.

Face location and region optimization based on skin color information

The biggest difference between the face area and the normal skin area is that the face is a connected area with several holes. By analyzing the topological structure of the connected region, we can judge whether the region contains holes and then determine whether it is a human face. The Euler number is used to calculate the number of holes in the connected region. The relationship between Euler number and image hole number is defined as follows:

$$E = C - H \qquad (7)$$

where E is the Euler number, C represents the connected component, and H represents the number of holes in the region. By calculating whether each connected area contains holes, the face regions are finally distinguished.

Face region optimization based on corner detection

Sometimes clothing background affects the distinction of face regions. Therefore, the face region is optimized based on corner points. The optimization steps are divided into contour corner detection and corner-based face region optimization.

(1) Contour corner detection

Firstly, the corners on the contour are found and the region is subtracted and compensated by the relationship between adjacent corners. The adaptive corner detection method is used to detect contour points.

(2) Optimization of face regions based on corner points

When the corners of the contour are found, the original contour is replaced by a short fitting curve, and the area of the new enclosure and the ratio of length to width are judged to be reasonable. If it is unreasonable, the original contour is replaced by a short fitting curve.

Face region relocalization based on local template matching

When the exact face region is obtained, the position of eyes and mouth can be easily located by combining the hole in the face skin image and the gray information of the hole. After the correction of the image direction, the fatigue characteristics of each organ can be extracted from all regions. In this paper, we use the grayscale of the exact face as a template to search the face within three times of the human face. The formula for template similarity is as follows: [25]:

$$\text{corr} = \frac{\sum m \sum n (I_{m,n} - \overline{I})(T_{m,n} - \overline{T})}{\sqrt{\sum m \sum n ((I_{m,n} - \overline{I})^2}\sqrt{\sum m \sum n (T_{m,n} - \overline{T})^2}} \qquad (8)$$

Among them, I and T represent the area to be tested and the face template respectively.

2.4.2.1 Eye location and tracking algorithm Eye features are often used in fatigue detection. Therefore, accurate eye location is one of the key technologies in fatigue detection. In this paper, we use the two-valued eye location algorithm. The algorithm uses different areas of the face to binarize the different reflectivity of the light, obtains binary images containing the human eye, and then uses bi-directional integral projection to accurately locate the position of the human eye. The specific steps are shown in Fig. 6:

• Calculation of image reflection components

The image is represented by $f(x, y)$ function, which is represented by the total amount of light incident to the scene and the total amount of light reflected by the object in the scene. The total amount of light incident into the observation scene is called the incident component, expressed by $i(x, y)$; the total amount of light reflected by the object in the scene is called the reflection component, expressed by $r(x, y)$. The nature of the incident component depends on the source. It can be measured by a low-pass filter; the reflection component determines the intrinsic properties of the image, and by setting the

Fig. 6 Operation process of eye location algorithm based on two values
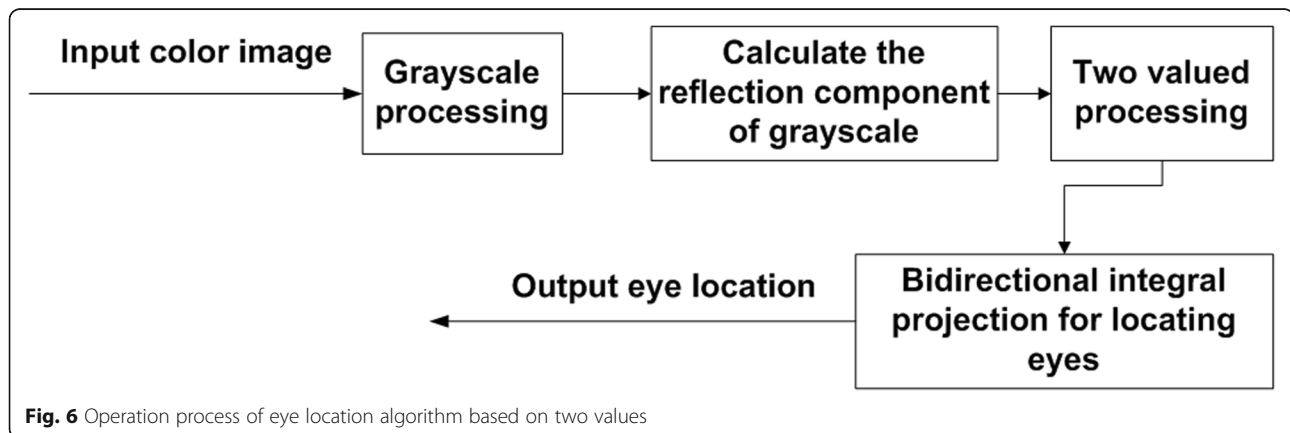
threshold value of the reflection component, the dark objects such as the eyes and eyebrows are segmented from the image where the reflection component is small. In this paper, the image is processed by low-pass filtering in the spatial domain, and each pixel and all the pixels in the surrounding area are averaged by a given size template. The result is the incident component of the pixel. The formula is:

$$i(x, y) = \frac{1}{(2N + 1)^2} \sum_{u=x-N}^{x+N} \sum_{v=y-N}^{y+N} f(u, v) \qquad (9)$$

N represents the total number of pixels in the image, and X and Y respectively represent the horizontal axis coordinates of the image. According to the relationship between $F(x, y)$, $I(x, y)$, and $R(x, y)$, formulas can be used.

$$r(x, y) = f(x, y)/i(x, y) \qquad (10)$$

Calculate the value of the entry and exit component $R(x, y)$.

- Two value processing

In this paper, we use $R(x, y)$ to separate dark objects such as human eyes and eyebrows. A threshold is given here. When the $R(x, y) >$ threshold value is in the image, the pixel will be set to 0 and the rest will be set to 1, so as to achieve two value.

- Precise positioning of human eyes

After two-value processing, the location of human eyes is roughly positioned. In order to complete the precise positioning of human eyes, we use the bidirectional integral projection method to complete the eye location. Firstly, the approximate region of the two eyes is selected, and then the two-way integral projection in the horizontal direction and the vertical direction is carried out on the region of the left and right eyes respectively. The intersection of the two directional maxima is the center of the eyeball.

2.4.2.2 Fatigue detection and classification algorithm
A typical fatigue driving detection system generally includes data acquisition, data preprocessing, feature extraction, and classification verification. The classification of fatigue detection is to map the fatigue feature space and driving state space to classify the feature space. Generally used classification algorithms are KNN algorithm, SVM algorithm, and ensemble learning algorithm. This paper introduces the principle and process of classification using KNN algorithm.

KNN is a commonly used classification algorithm of machine learning [26]. It classifies different distances between different eigenvalues. The idea is that if most of the k most similar (that is, the nearest neighbor in the feature space) samples in a feature space belong to a certain category, then the sample also belongs to this category, where K is usually no more than 20 integers. In the KNN algorithm, the selected neighbors are all correct categorization objects. This method only determines the category of the sample to be classified according to the nearest one or several samples. The classification idea is shown in Fig. 7 [27].

The three points adjacent to the X point in the graph are red and belong to the w1 category. The other point adjacent to the X point is green and belongs to the w3 category. Because most of the points adjacent to X belong to w1 class, the sample X points are also classified as w1 classes.

The idea of KNN algorithm is to input test data when the training set data and labels are known [28], compare the characteristics of test data with those of training set, and find the first K data which is the most similar to the training set, then the corresponding category of the test data is the most frequent

occurrence of *K* data. The classification of the algorithm is described as follows:

1. Calculate the distance between the test data and the training data; there are many ways to measure the distance. Commonly used is the Minkowski distance (Minkowski distance), which is defined as:

$$D(x, y) = \left(\sum_{i=1}^{m} |x_i - y_i|^p \right)^{\frac{1}{p}} \qquad (11)$$

where *X* and *Y* denote the *i* elements of two *m* dimensional variables$_i$ and *p* is a variable parameter. When *p* equals 2, it is Euclidean distance. When *p* equals 1, it is Manhattan distance.

2. Sort according to the increasing relationship of distance.
3. Select *K* points with minimum distance.

The choice of *K* will have a significant impact on the result of the algorithm. If the *K* value is small, it is equivalent to using a training example in a small neighborhood to predict. In extreme case, $k = 1$, the test case is only related to the nearest sample, and the training error is very small (0), but if the sample is just noise, the prediction will be wrong, and the test error is very large. That is to say, when the *K* value is small, the phenomenon of over fitting will occur. If the value of *K* is large, it is equivalent to using training instances in a large neighborhood to predict, the extreme case is $k = n$, and the result of test instances is the largest class in the training data set, which will produce under-fitting. In applications, *K* is generally smaller and *K* is odd. Cross-validation is usually used to select suitable *K* values.

4. Determine the frequency of the categories of the preceding *K* points.
5. The categories with the highest frequency in the *K* points before returning are used as prediction classifications of test data.

3 Experimental results and discussions

In this experiment, a 3Y-31D automobile driving simulator developed by Beijing Sino-Commercial Union Instruction Equipment Co., Ltd. was used as a driving simulator to take videos of 10 drivers (3 females and 7 males). In this paper, blink frequency and PERCLOS are used as fatigue characteristic parameters. Blinking frequency refers to the number of blinking times per unit time. Its formula in the *I* window is:

$$\mathrm{BF} = \frac{\mathrm{BT_{ei}} - \mathrm{BT_{si}}}{T_{\mathrm{bf}}}, \mathrm{BT}_i \neq 0, i = 1, 2, \dots n \qquad (12)$$

Formula: $\mathrm{BT_{ei}}$ is the total number of blinks at the *i* of the first window, $\mathrm{BT_{si}}$ is the total number of blinks at
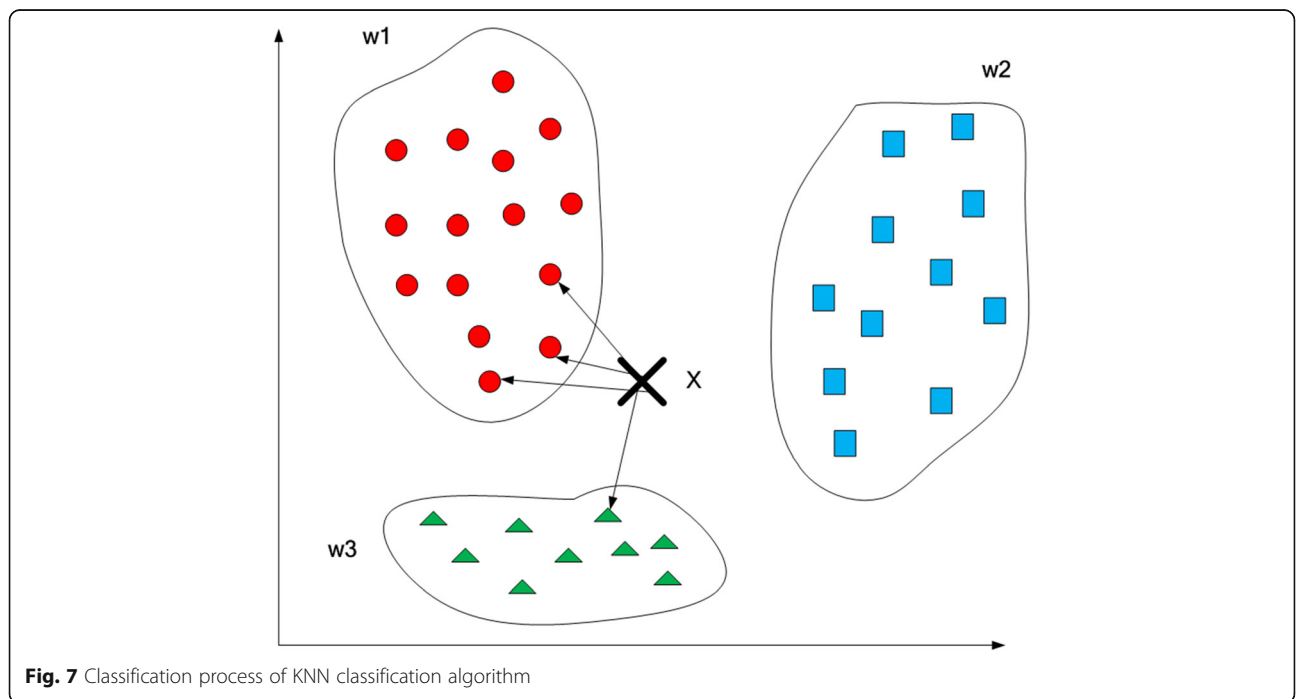

Fig. 7 Classification process of KNN classification algorithm

the beginning of the i window, and the size of T_{bf} is the calculated time window for BF. In PERCLOS use P_{80}, it is the proportion of eye opening less than 20% per unit time. Its formula is:

$$p_{80} = \frac{n_p}{T_{P_{80}} \times f_0} \qquad (13)$$

Formula: f_0 is the sample sampling frequency, T_{P80} is the calculated time window size of P_{80}, and n_p is the unit time window the degree of eye closure exceeds 80% times. Sliding time window is used to fuse the two feature data, and KNN classification method is used to build a fatigue driving detection model to judge the fatigue of the test samples.

4 Result analysis

After data acquisition, the collected video data is grayscale, two-valued, and denoised preprocessing. Then face localization based on skin color modeling algorithm, and then locate by two valued human eyes. Finally, the method of bi-directional integral projection is used to locate the human eye accurately. Some video image positioning results are shown in Fig. 8.

As can be seen from Fig. 8, the faces of the two testers were correctly positioned, including the eyes, nose, and mouth. The area of the face is precise from the chin to the forehead, from the left ear to the right ear. The location area is more precise and does not include too many non-facial areas. Therefore, the algorithm based on skin color modeling can accurately locate the driver's face area. The binary eye localization algorithm can also accurately locate the eyes of the two testers. The location of the eye area includes the entire eye area of the tester and contains less non-ocular regions. Therefore, the localization algorithm based on the two-valued algorithm is also effective. In addition, the testing accuracy rate of wearing glasses is better. The reason is that the reflection of the spectacles is less in the cab environment.

After accurately locating the position of the human eye, the input X = [blinking frequency, P80] of the KNN model is obtained by calculating the corresponding blinking frequency and p80 through the formula 10 and 11. The effective eye movement characteristic parameters group was composed of fatigue windows driving and normal driving by sliding window. A total of 2500 normal driving samples and 2000 groups of fatigue driving samples were obtained. Two thousand groups of normal driving samples and 1200 groups of fatigue driving samples were selected randomly as training set. Other remaining samples were used as test samples. The optimal K value is 7. The KNN model was tested with the remaining 1500 normal driving samples and 800 fatigue driving samples. The model test results are shown in Table 1.

According to Table 1, the number of fatigue driving samples correctly identified by the model is 700, the number of false recognition is 100, the number of
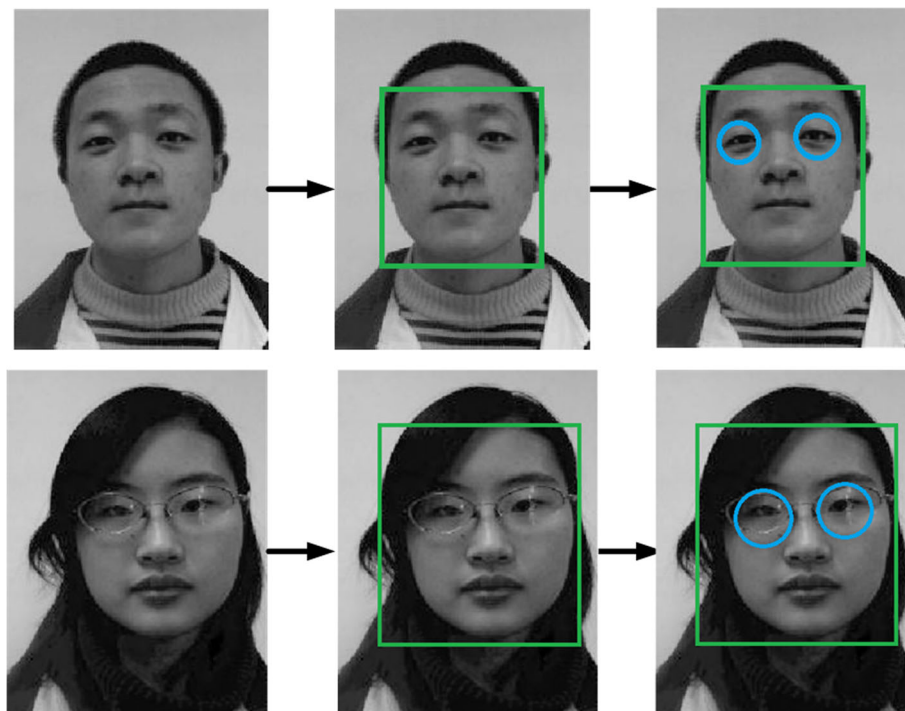

Fig. 8 Face and eye location in video images

Table 1 Output of KNN model

Model output	Driver's actual driving state	
	Fatigue driving	Normal driving
Fatigue driving	700	180
Normal driving	100	1320

normal driving samples correctly identified is 1320, and the number of false recognition is 180. Therefore, the number of correct identification is 2020, and the number of false identification is 280. The accuracy of model checking was 87.82%. In addition, the sensitivity of the model is 89.12%. The specificity is 83.28%. The analysis shows that the misjudged data can be found in two drivers. Due to individual differences among drivers, the difference between individual drivers' PERCLOS and blinking frequency and other drivers is too great, which leads to misjudgment.

5 Conclusion

With the increase of the number of automobiles in the world, intelligent transportation system has become an important means to solve modern traffic problems. Intelligent transportation system (ITS) involves image processing, intelligent recognition, machine vision, and other interdisciplinary. Fatigue driving early warning technology based on image recognition theory is the key technology widely used in ITS. In this paper, the key technologies of fatigue driving detection based on machine vision are analyzed and studied, and the results are illustrated by experiments. The main works in this paper are as follows:

1. The key technologies of fatigue driving detection based on machine vision include face localization algorithm, eye localization algorithm, and fatigue detection classification algorithm.
2. The commonly used algorithms of face location, eye location, and fatigue detection and classification are analyzed and improved. In the analysis of face localization algorithm based on skin color modeling, the corner-based optimization of face region is proposed; in the analysis of eye localization algorithm based on binary algorithm, a bi-directional integral projection method is proposed to achieve accurate eye localization.
3. Through the fatigue driving experiment, the data of normal driving and fatigue driving are collected, and the positioning and judging efficiency of the improved algorithm is verified by the collected data.

Abbreviations
ITPS: Intelligent traffic prediction system; ITS: Intelligent transportation system; RFID: Radio frequency identification; TSL: Tint-Saturation-Luma

Availability of data and materials
Please contact author for data requests.

About the author
Jun Wang was born in Changchun, Jilin,P.R. China, in 1973. He received the Master degree from Jinlin University, P.R. China. Now, he works in the specialty of vehicle inspection and maintenance, in the Jilin Engineering Normal University, as the director of the lecturer teaching and research office. His research interests include intelligent vehicle management,etc.
Xiaoping Yu was born in Harbin, Heilongjiang, P.R. China, in 1983. She received the Doctor degree from Beijing Jiaotong University, P.R. China. Now, she works in School of Architecture and Design, Beijing Jiaotong University. Her research interests include data visualization and urban transportation.
Qiang Liu was born in Jinzhou, Liaoning,P.R. China, in 1979. He received the Master degree fromChangchun University of Science and Technology,P.R. China. Now, He works in the specialtyof vehicle inspection and maintenance, in the Jilin Engineering Normal University. He research interests include engine emission and control, newenergyvehicles,etc.
Qiang Liu1,2,a, 1 State Key Laboratory of Automotive Simulation and Control, Jilin University, Changchun, China
2 Innovative Research Team of Jilin Engineering NormalUniversity(IRTJLENU), Jilin Engineering Normal University, Changchun, China.
Zhou Yang was born in Huludao, Liaoning,P.R. China, in 1989. She received the Master degree from Jinlin University, P.R. China. Now, she works in the specialty of vehicle inspection and maintenance, in the Jilin Engineering Normal University. Her research interests include engine emission and control,etc.

Authors' contributions
All authors take part in the discussion of the work described in this paper. JW wrote the first version of the paper. QL and ZY did part experiments of the paper. All authors read and approved the final manuscript.

Competing interests
The authors declare that they have no competing interests.

Publisher's Note
Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Author details
[1]School of Automotive Engineering, Jilin Engineering Normal University, No. 3050 kaixuan road, Changchun, China. [2]State Key Laboratory of Automotive Simulation and Control, Jilin University, Changchun, China. [3]School of Architecture and Design, Beijing Jiaotong University, Beijing, China.

References
1. Gao T, Liu Z G, Yue S H, et al. Moving vehicle tracking algorithm used for intelligent traffic[J]. China J. Highway Transport., 2010, 23(3):89–94
2. Tang Y, Zhang C, Gu R, et al. Vehicle detection and recognition for intelligent traffic surveillance system[J]. Multimed. Tools Appl., 2015, 76(4):1–16
3. Chen B, Xie Y, Tong W, et al. A comprehensive study of advanced information feedbacks in real-time intelligent traffic systems[J]. Physica A Statistical Mechanics and Its Applications, 2012, 391(8):2730–2739

4. Dorrian J, Roach G D, Fletcher A, et al. Simulated train driving: fatigue, self-awareness and cognitive disengagement[J]. Appl. Ergon., 2007, 38(2):155–166
5. Shao-Bin W U, Li G, Wang L A. Detecting driving fatigue based on electroencephalogram[J] Trans. Beijing Institute of Technology, 2009, 29(12): 1072–1075
6. Zhang Kaiguang, Baming Ting, Meng Hongling, et al. Research on the optimal path algorithm of Luoyang intelligent transportation system[J]. Henan Science, 2012, 30 (5): 635–639
7. Xie Shuyun, Ran Jie, Yang Cedar. Research on intelligent urban transportation system based on group intelligence perception [J]. Electron. Des. Eng., 2014, 22 (20): 49–51
8. Wang Shaohua, Lu Hao, Huang Qian, et al. Research on key technologies of intelligent transportation system [J]. Surveying and Spatial Geogr. Inf., 2013 (s1): 88–91
9. Ganasindu K S, Smithashekar B, Harish G. An approach for intelligent traffic splitting for sudden changes of traffic, dynamics[J]. Iran. J. Clin. Infect. Dis., 2011, 20(2):167–169
10. Dong C, Ma X, Wang B, et al. Effects of prediction feedback in multi-route intelligent traffic systems ☆[J]. Physica A Statistical Mechanics & Its Applications, 2012, 389(16):3274–3281
11. Patel A, Kaushik P. Improving QoS of VANET Using Adaptive CCA Range and Transmission Range both for Intelligent Transportation System[J]. Wireless Personal Communications, 2018, (3):1–36
12. Ngoduy D. Platoon-based macroscopic model for intelligent traffic flow[J]. Transportmetrica B Transport Dynamics, 2013, 1(2):153–169
13. Sun Q S, Zeng S G, Liu Y, et al. A new method of feature fusion and its application in image recognition[J]. Pattern Recogn., 2005, 38(12):2437–2448
14. Baidyk T, Kussul E, Makeyev O, et al. Flat image recognition in the process of microdevice assembly[J]. Pattern Recogn. Lett., 2004, 25(1):107–118
15. Mendoza O, Melin P, Licea G. A hybrid approach for image recognition combining type-2 fuzzy logic, modular neural networks and the Sugeno integral[J]. Inf. Sci., 2009, 179(13):2078–2101
16. López M B, Hannuksela J, Silvén O. Accelerating image recognition on mobile devices using GPGPU[J]. Proc. SPIE Int. Soc. Opt. Eng., 2011, 7872(4):389–393
17. Perronnin F, Mensink T. Improving the fisher kernel for large-scale image classification[J]. Eccv, 2010, 115(7):143–156
18. Lu D, Weng Q. A survey of image classification methods and techniques for improving classification performance[J]. Int. J. Remote Sens., 2007, 28(5):823–870
19. Sanchez, Jorge, Perronnin F, et al. Image classification with the fisher vector: theory and practice[J]. Int. J. Comput. Vis., 2013, 105(3):222–245
20. Camps-Valls G, Gomez-Chova L, Munoz-Mari J, et al. Composite kernels for hyperspectral image classification[J]. IEEE Geoscience & Remote Sensing Letters, 2006, 3(1):93–97
21. Camps-Valls G, Bruzzone L. Kernel-based methods for hyperspectral image classification[J]. IEEE Transactions on Geoscience & Remote Sensing, 2005, 43(6):1351–1362
22. Foody G M, Mathur A. The use of small training sets containing mixed pixels for accurate hard image classification: training on mixed spectral responses for classification by a SVM[J]. Remote Sens. Environ., 2006, 103(2):179–189
23. Wu C, He. Adaptive illumination detection system for fatigue driving[J]. J. Electron. Meas. Instrument, 2012, 26(1):60–66
24. Radun I, Radun J E, Summala H, et al. Fatal road accidents among Finnish military conscripts: fatigue-impaired driving.[J]. Mil. Med., 2007, 172(11):1204
25. Zhang L W, Yang Y F, Mei-Bin Q I, et al. Detection of fatigue driving based on facial features[J]. Journal of Hefei University of Technology, 2013, 36(4):448–451
26. X. Zhou, X. Liang, X. Du, J. Zhao, Structure based user identification across social networks. IEEE Trans. Knowl. Data Eng. 30(6), 1178–1119 (2018)
27. D. Lu, X. Huang, G. Zhang, X. Zheng, H. Liu, Trusted device-to-device based heterogeneous cellular networks: a new framework for connectivity optimization. IEEE Trans. Veh. Technol. 67(11)11219–11233 (2018)
28. H. Fu, Z. Liu, M. Wang, Z. Wang, Big data digging of the public's cognition about recycled water reuse based on the BP neural network. Complexity (2018) doi.org/10.1155/2018/1876861