Vehicle tracking based on Adaboost cascade classifier

Xu Long¹, Hu Ming², Tao Tao¹*

¹School of computer science and technology, Anhui University of Technology, Ma'anshan, Anhui, 243032
²School of electrical and information engineering, Ma'anshan, Anhui, 243032

*Corresponding author: taotao@ahut.edu.cn

Abstract. Driving safety and unobstructed travel have always been the common pursuit of drivers. In the actual environment, traffic elements are complex and changeable, and a large number of dynamic and uncertain factors bring great challenges to the driving behavior decision of smart vehicles. Machine vision because of its advantages of wide detection range and complete road information are widely used in smart vehicles. In this paper, we take the road environment as the main research object and carry out research on target detection and tracking in smart vehicles. First, we propose a vehicle detection method based on multi-class feature fusion, which combines the SVM classifier. The method can effectively improve the accuracy of vehicle recognition. Then a vehicle detection method based on the Adaboost cascade classifier algorithm is proposed, and the experimental verification of the accuracy and robustness of the vehicle detection algorithm in different road scenarios is completed. Based on the principles and ideas of Kalman filtering, vehicle tracking is realized on the basis of vehicle detection, and the effect of vehicle tracking algorithm based on Kalman filtering is verified under different road scenarios.

1. Background and meaning

1.1. Research Background

In recent years, China's technology has developed rapidly, and mobile Internet, big data, and cloud computing have quickly become the mainstream of the new era in the 21st century. "Internet +" Development Strategy "and" Made in China 2025 "were officially proposed by the Chinese government, which re-planned China The development direction of science and technology vigorously promotes the upgrading of China's technology industry and the optimization and adjustment of its structure. In 2016, the "China Intelligent Connected Vehicle Technology Roadmap” was announced, and the 2017 "China Intelligent Connected Vehicle Industry Development White Paper 2017” was renewed by China. Adjustment and update [1], the introduction of these two documents also pointed out the direction of the development of the intelligent connected car of the century. Generally, the intelligent connected car refers to the use of its own sensors, communication equipment to the surrounding environment of the vehicle A new generation of cars that senses and communicates with road infrastructure and surrounding vehicles, can use the controller carried by the vehicle to make reasonable and safe decisions based on the sensed information, and can use actuators to execute decisions.

Generally, a smart connected car refers to the use of its own sensors and communication equipment to...
Perceives the environment and communicates with road infrastructure and surrounding vehicles, based on the perceived information.

Use the controller carried by the vehicle to make reasonable and safe decisions, and use the actuator to perform new decisions.

Generation of cars.

Intelligent connected cars not only make the travel ways safer, more environmentally friendly, more convenient, and more energy efficient, but also have great significance for building urban intelligent transportation systems, building a green car society, and reshaping cars and related industries. UMC has great potential in solving traffic safety, road congestion, and improving the driving experience. It has attracted great attention worldwide and is an internationally recognized development direction [2]. The corresponding guiding policies are also in the United States, the introduction of the European Union and other countries has improved the development efficiency of enterprises. Autonomous vehicles are a comprehensive automobile body that integrates high-tech such as artificial intelligence, vehicle engineering, and communications. The research and development technology is very complex, and autonomous vehicles have diverse operating environments. Sexual requirements are high, so there are still many deficiencies in its development process, such as: low accuracy of environmental perception, poor intelligence of decision-making, these issues need to be improved.

1.2. Research Significance

Vehicle detection based on multi-class feature fusion, this method mainly realizes vehicle detection by extracting the salient features of the vehicle in the image. For example, when detecting a vehicle, the vehicle's geometry, symmetry, rear lights, color features, bottom Shadows can be a significant feature of vehicles driving on roads in good light conditions. However, weather conditions, lighting conditions, and even early morning or afternoon have a great impact on the shadow characteristics of the bottom of the vehicle, relying solely on this Features to achieve vehicle detection, its accuracy will be greatly reduced.

Research data show that if the driver can take 0.5s more reaction time, it can reduce accidents at intersections by about 60% and collisions by about 30% .If the alarm system on the car can send an alert to the driver 1s in advance Will reduce the occurrence of traffic accidents by more than 90%; and sensors and controllers such as smart cars can effectively sense environmental information and make reasonable decisions through advanced algorithms, which can assist. Even in the future, it will replace the driver to drive the vehicle, thereby improving travel efficiency and safety, so more and more companies are currently focusing on the research of smart car related technologies.

According to China's intelligent connected car technology roadmap, it is known that the development of smart cars will ultimately require driverless driving, but in the process of driving cars, due to the complicated road environment, there are still many issues that need to be resolved to fully realize autonomous driving. Although driverless driving cannot be fully realized at this stage in the short term, it is possible to assist the driver in specific scenarios by studying the driving assistance system based on machine vision. This will partially solve the current problems in the transportation field and also achieve the final realization. Self-driving has accumulated technology. Therefore, it is of great research significance and practical value to carry out vehicle tracking based on the Adaboost cascade classifier.

2. Vehicle detection based on multi-class feature fusion

In image-based vehicle detection, feature extraction and detection methods are the two main factors that affect the effectiveness of vehicle detection. From the perspective of feature extraction, vehicle detection methods can be divided into three major detections based on texture features, shape features, and color features. Method [2]; from the perspective of recognition methods, vehicle detection methods can be further divided into detection methods based on Ada Boost, support vector machine SVM, random forest and other classifiers. At present, both from the perspective of feature extraction and detection methods from the perspective, many scholars have done a lot of research and different attempts on object detection and its related fields.
2.1. Feature extraction

2.1.1. hog feature extraction. HOG feature is a kind of feature descriptor [3] of the local area of the image. Because it can maintain good invariance to illumination and geometric deformation, it is often used for object detection and recognition. (U, v) represents the input image. Pixels, the formula for calculating the gradient of pixels in the image is:

\[
G_u(u, v) = H(u+1, v) - H(u-1, v)
\] (1)

\[
G_v(u, v) = H(u, v+1) - H(u, v-1)
\] (2)

Where \(G_u(u, v), G_v(u, v), H(u, v)\), respectively, represent the horizontal gradient, vertical gradient, and Pixel value. The formula for calculating the gradient magnitude and gradient direction at the pixel point \((u, v)\) is:

\[
G(u, v) = \sqrt{G_u(u, v)^2 + G_v(u, v)^2}
\] (3)

\[
\alpha(u, v) = \tan^{-1}\left(\frac{G_v(u, v)}{G_u(u, v)}\right)
\] (4)

The hog feature extraction process is as follows:

1. Grayscale the original image and use \(\text{Gamma} \) Correction method for color space normalization;
2. Divide the image into multiple 3 * 3 cells and divide \([-\pi, \pi]\) The gradient direction is divided into 9 bins and the gradient (including size and direction) of each pixel in the cell is calculated;
3. Perform histogram statistics on the gradient amplitudes of all pixels in each cell in bins in all directions to obtain a 9-dimensional feature vector;
4. Every four adjacent cells form a square area, and the features of all cells in each square area are connected in series to obtain the 36-dimensional feature vector of the square area, as shown in the figure below;

**Figure 1.** block diagram
4

(5) The hog feature descriptors of all square regions in the original image are serially connected in order to obtain a feature vector with a dimension of 9396 dimensions.

2.1.2. Dimension reduction of hog features. After using the HOG feature extraction algorithm in this paper to extract the HOG features, the HOG features of an image obtained are 9396 dimensions, the HOG feature dimension is too high, the amount of data is huge, and there is a large amount of data redundancy. Testing of sample images brings a large amount of calculations. In order to eliminate redundant information between features and improve the efficiency of image processing, this article uses Principle Component Analysis (PCA) to perform dimensionality reduction operations, and then extracts HOG features.

With n original images and extracting the p-dimensional features of each image, this original feature set can be expressed as:

\[
X = \begin{bmatrix}
X_{11}, X_{12}, \ldots, X_{1p} \\
X_{21}, X_{22}, \ldots, X_{2p} \\
\vdots \\
X_{n1}, X_{n2}, \ldots, X_{np}
\end{bmatrix} = \begin{bmatrix}
X_1, X_2, \ldots, X_p
\end{bmatrix}
\]

(5)

The new comprehensive variable is linearly represented by the original variable, and we get:

\[
\begin{align*}
y_1 &= \alpha_{11}X_1 + \alpha_{12}X_2 + \cdots + \alpha_{1p}X_p \\
y_2 &= \alpha_{21}X_1 + \alpha_{22}X_2 + \cdots + \alpha_{2p}X_p \\
&\vdots \\
y_p &= \alpha_{p1}X_1 + \alpha_{p2}X_2 + \cdots + \alpha_{pp}X_p
\end{align*}
\]

(6)

Satisfaction: \(\alpha_{i1} + \alpha_{i2} + \cdots + \alpha_{ip} = 1, i = 1, 2, \ldots, p\). In order to ensure a high recognition rate, the threshold of the cumulative contribution rate in this paper is set to 99\%, and a 376-dimensional HOG feature vector after dimensionality reduction is obtained through matlab calculation.

2.2. Data normalization

Different features of the same image have significant differences in value after being extracted. In order to avoid the effect of this disparity caused by different feature values on the recognition result of the image, a range transformation method is proposed in this paper. The data is feature-normalized with the following formula:

\[
f : \tilde{x} \rightarrow \tilde{y} = \frac{\tilde{x} - \tilde{x}_{\min}}{\tilde{x}_{\max} - \tilde{x}_{\min}}
\]

(7)

Where \(x\) is any component of a certain eigenvector, \(\tilde{x}_{\min} = \min(\tilde{x})\), \(\tilde{x}_{\max} = \max(\tilde{x})\).
2.3. feature fusion
Let the feature vectors of the extracted hog feature, moment invariant feature, and gray level co-
ocurrence matrix feature of an image be $[a_1, a_2, a_3, \ldots, a_n]$, $[b_1, b_2, b_3, \ldots, b_k]$, $[c_1, c_2, c_3, \ldots, c_m]$ And fuse it to get the fused feature vector $v$:

$$V = [a_1, a_2, a_3, \ldots, a_n, b_1, b_2, b_3, \ldots, b_k, c_1, c_2, c_3, \ldots, c_m]$$

(8)

After the image in this paper is calculated by the previous algorithm, the dimensionality of the feature vector obtained by the dimensional reduction operation of the proposed hog feature is 376, the feature vector of the invariant moment feature is 7 dimensional, and the feature vector of the gray co-occurrence matrix the dimension is 5 dimensions. The above three types of features are normalized. Based on the characteristics that the linear combination of related feature vectors does not affect its invariance, a linear fusion method is used to directly connect the three feature vectors in series to obtain a new 388-dimensional fusion feature vector.

2.4. svm classifier
svm is a machine learning method developed on the basis of statistical learning theory, which will optimize

The learning problem is transformed into a convex quadratic optimization problem, thereby avoiding local minima and effectively solving the over-learning problem, because it has good generalization ability and high classification accuracy, and uses limited training samples to obtain independent decision rules. The test set can still get the characteristics of smaller errors. This paper chooses svm as the classifier for classification and recognition. The architecture of the support vector machine is shown in the following figure:

![Support Vector Machine Architecture](image)

**Figure 2. Support Vector Machine Architecture**

Since the sample in this article is divided into a vehicle sample and a non-vehicle sample, its essence is a binary classification problem.

A two-class support vector machine is used, that is, a c-svc model.

2.5. Simulation
In order to verify the effectiveness of the vehicle identification method proposed in this article, simulation experiments were selected on the matlab platform. This experiment collected a total of 872 images, which belong to vehicle and non-vehicle sample data, of which 436 are vehicle sample data.
436 non-vehicle sample data, all images are normalized to a size of 30 * 30 pixels, and some vehicle and non-vehicle sample data are shown below:

![Vehicle and non-vehicle sample data](image)

**Figure 3.** Samples of some vehicles and non-vehicles

In order to avoid the mutual influence and interference between the vehicle sample data and non-vehicle sample data, training sample data and test sample data, and affect the accuracy of the experimental results, the 872 sample data are first divided into 4 types of sample data with the number of samples being 400, 36, 400, and 36. Then the four types of samples are respectively extracted hog feature, moment invariant feature, and gray co-occurrence matrix feature, and the extracted data features The vectors are 9396, 7 and 5 dimensions. After the hog feature extraction is completed, the dimensionality reduction operation is performed using the pca algorithm. On the premise of ensuring that the compressed data can represent the original data to 99%, the final result is calculated. Get the new 376-dimensional hog feature vector.

The extracted three types of feature data have large differences before normalization. Take the vehicle image as an example. Part of the feature data is shown in Table 1. Without normalization processing, it is not easy to compare and find the distance. The parameters and accuracy will be affected. In order to improve the speed and accuracy of the calculation, normalization of the characteristic data is required.

**Table 1.** Various features of the image before normalization

| HOG Feature | Invariant moment feature | Gray level co-occurrence moment feature |
|-------------|--------------------------|-----------------------------------------|
| 0.36305286  | 0.001173486              | 0.0026000981                           |
| 0           | 2.34E-08                 | -96.29677172                           |
| Vehicle image | 0.095928                | 5.96E-12                                |
| ... ...     | 3.62E-12                 | 15347.09156                             |
| 0           | -9.75E-24                | 0.012291759                             |
| 0           | -5.41E-16                |                                         |
| 0           | 1.37E-23                 |                                         |
Table 2. Various features of the image after normalization

| HOG Feature   | Invariant moment feature | Gray level co-occurrence moment feature |
|---------------|--------------------------|----------------------------------------|
| 0.363052868   | 1                        | 0                                      |
| 0             | 1.99E-05                 | 0.0062356375265                       |
| 0.338210523   | 5.08E-09                 | 0.0063807046606                       |
| ...           | 3.09E-09                 | 1                                      |
| 0.323116909   | 4.61E-13                 | 0.0062362650870                       |
| 0.356681006   | 0                        |                                        |
| 0.311446088   | 4.61E-13                 |                                        |

The experiment uses a supervised learning method. Before the experiment, 872 data are first labeled with category labels. In this paper, the vehicle sample label is recorded as 1, and the non-vehicle sample label is recorded as 0. A 872-dimensional label vector is obtained. For the experimental test results, the actual classification results of the 36 vehicle and 36 non-vehicle test samples are represented by circles, and the predicted classification results of the corresponding test samples are represented by stars. This experiment is divided into two groups, which are single feature experiments and fusion feature experiments.

First, use the normalized training data of the three types of features to separately train the svm model to obtain the corresponding svm classifier model, and then enter the test data of the three types of features into the trained svm classifier. The test results are shown in Figure Shown:

![hog feature classification test results](image)

Figure 4. hog feature classification test results

The blue circle represents the classification of the actual test set, and the red circle represents the classification of the predicted test set. The optimal result of the experiment is that the actual test sample represented by the light blue circle and the prediction test sample represented by the red asterisk completely coincide. This will show that the svm-based classifier has completely identified positive and negative samples. 72 samples were actually tested in this experiment, of which positive samples (vehicle images) and negative samples (non-vehicle images) each accounted for 36.

The extracted hog feature is the edge feature of the image, the invariant moment feature is the shape feature of the image, and the gray level co-occurrence matrix feature is the texture feature of the image.
They are only features of one aspect of the image, and the amount of information extracted is insufficient. When a certain aspect of the image is unknown, the vehicle recognition method based on this single feature cannot guarantee the accuracy of the recognition, resulting in inaccurate vehicle recognition. The rate is low, and the above three types of single features of the extracted image are merged into new features before vehicle recognition, which overcomes the disadvantages of vehicle recognition based on a single feature and improves the accuracy of recognition.

3. Vehicle detection based on Adaboost algorithm

3.1. Haar-like feature extraction and calculation

The Haar-like feature was proposed by Papageorigiou, and after the development of Viola et al., it currently includes three categories and four forms, namely edge features, linear features, point features, and diagonal features. Considering the symmetry of the vehicle objects examined in this paper, this article only considers the two-rectangular and three-rectangular features of edge features and linear features, and calls this rectangular feature a feature template. The two types of feature templates are shown below:

![Haar-like rectangular features](image)

The eigenvalue is equal to the sum of the pixel values in the black area minus the sum of the pixel values in the white area. The formula is as follows:

\[ v = \text{Sum(black)} - \text{Sum(white)} \] (9)

This feature template can be enlarged to form a series of sub-rectangular features, but no matter what kind of rectangular feature, the ratio of the black and white area area always remains the same. Haar-like feature template and a series of rectangular features formed by its enlargement are aligned. Traversing an image will get thousands of eigenvalues, so it will take a lot of time to extract the features of an image. Therefore, the integration map is introduced to accelerate the calculation process of Haar-like eigenvalues, and the integration map is used. Do the calculation, which is:

\[ v = 2^{*}\text{integral}_{b,1} + \text{integral}_{b,3} - 2^{*}\text{integral}_{b,2} \] (10)

3.2. Adaboost Classifier

The idea of the Adaboost algorithm is to combine the classification effects of multiple weak classifiers to form a strong classifier to obtain valid output. Its main steps are: first build a neural network model as a “weak” classifier, and iterate through T iterations. The network weights are updated to obtain T neural network weak classifiers with different weights, and each operation gives greater weight to the individuals who fail to classify, making the next individual operation pay more attention to these trained individuals. The weak classifier is generated after multiple operation iterations. The sequences a1, ..., aT are evaluated to evaluate the classification effect of the T neural network weak classifiers respectively. The better the classification effect, the larger the corresponding value. Finally, the Adaboost model
combines the classification effects of multiple neural network weak classifiers to form a strong classifier. Specific steps are as follows:

1. First, select and initialize data, randomly select m sets of training data from the sample space, initialize the distribution weight of the data \( D_t(i) = 1/m \), determine the neural network according to the number of input and output samples, and initialize the BP Neural network weights and thresholds;

2. Weak classifier prediction, training. When a weak classifier is used, we train the bp neural network with the training data and predict the output to get the predicted sequence \( g(t) \). Prediction error \( e_t \) and \( e_t \). The formula for calculating the sum of errors is

\[
e_t = \sum_i D_t(i) \quad (g(i) \neq y)
\]

Where, \( g(t) \) is the predicted sequence; \( y \) is the expected classification result; \( D_t(i) \) is the distribution weight of the test data.

3. Calculate the weight of the prediction sequence, according to the prediction sequence \( g(t) \). Prediction error \( e_t \) and \( e_t \). Calculate the weight \( a_t \) of the sequence, the weight calculation formula is

\[
a_t = \frac{1}{2} \ln \left\{ \frac{1-e_t}{e_t} \right\}
\]

4. Test data weight adjustment, adjust the weight of the next round of training samples according to the predicted sequence weight, and the adjustment formula is:

\[
D_{t+1}(i) = \frac{D_t(i)}{B_t} \times \exp[-a_t y_t g_t(x_t)] \quad i = 1, 2, \ldots, m
\]

In the formula, \( B_t \) is a normalization factor, the purpose is to make the distribution weight sum to 1 while the weight ratio is constant;

5. Get t group of weak classifiers after training t rounds \( f(g_t, at) \) weak classifier by \( t f(g_t, at) \) weak classifier by \( t f(g_t, at) \) strong classifier;

6. Set the normal sample to 1 and the abnormal sample to -1 in the training sample. Use the strong classifier and the feature data of the video to be tested to obtain the classification result.

The training samples are divided into positive samples and negative samples, where positive samples are vehicle images cropped close to the vehicle edges; negative samples can be any images that do not contain vehicle features.

The positive samples in this experiment are obtained by downloading the image database on the Internet. This method can save the production time of the positive samples; then, the positive samples obtained are intercepted. The principle is to intercept the rectangles closely to the target vehicle. Finally, after the interception, the positive sample images are normalized, and all processed into images with a size of 24 * 24 pixels. This can improve the training speed of the computer and save memory.

First of all, determine the proportion of positive and negative samples. According to a large amount of literature consulted, it is known that the proportion of positive and negative samples should be maintained at 1:5. A small ratio will not only increase the time it takes to train the classifier, but also cause training overfitting. Affects the detection effect of the classifier. Then, take a negative sample and substitute it into the Adaboost classifier for calculation. The test results are shown in the following figure:
The Adaboost-based cascade classifier trained in this paper obtained a detection result with an overall accuracy of 87.64%, and the overall detection effect is good. However, from the situation reflected by the above experimental results, the detection effect of the classifier will be affected by weather conditions and roads. The impact of environmental complexity. The detection effect of the classifier is the best under good weather conditions and simple road conditions. In other cases, the detection effect of the classifier will be reduced, because it is affected by various environmental noise. Impact, reducing image quality, affecting vehicle feature extraction, and then affecting the detection performance of the classifier.

4. Vehicle tracking based on Kalman filtering

4.1. Kalman filtering principle

Based on the camera’s target vehicle has a small amount of movement between consecutive frames, the position of the vehicle in the previous frame can be used to predict its position in the next frame of the image, thereby improving the efficiency and accuracy of vehicle detection. When it is difficult to meet the real-time detection requirements based on the vehicle detection algorithm alone, the vehicle should be used.

The vehicle tracking algorithm quickly detects the vehicle. The core idea of Kalman filtering is to obtain the estimated value of the state at the next moment by combining the observed value of the state at the current moment with the predicted value of the state at the next moment. This article uses the Kalman filter algorithm to detect the vehicle is marked to achieve the purpose of tracking the target vehicle.

Assume the system dynamic equation is:

\[
X(k) = F(k)X(k-1) + B(k)u(k - 1) + w(k - 1)
\]

\[
Z(k) = H(k)X(k) + v(k)
\]

The Kalman algorithm derivation is divided into two processes: the prediction process and the update process.

Prediction process: (1) Substituting the state estimation value \(X(k-1)\) at the time \(k-1\) into the determining part of the dynamic model of the system to obtain the predicted value \(X(k|k-1)\) at the time \(k\), calculated as follows:

\[
\hat{X}(k | k-1) = F(k)\hat{X}(k-1 | k-1) + B(k)u(k - 1)
\]

(2) Calculate the prior estimation error covariance matrix \(P(k|k-1)\), which is used to measure the accuracy of the prediction. The calculation formula is as follows:
Update process: (1) Based on whether the estimated value $X(\text{k} - 1)$ of the estimated value of time $k$ at time $k - 1$ is correct or biased, we use the actual measured value $Z(k)$ and the estimated output value $Z(\text{k} - 1)$ to measure and consider using this error to compensate, the calculation formula is as follows:

$$P(k|k-1) = \text{cov}(X(k) - \hat{X}(k|k-1)) = F(k-1)P(k-1|k-1)F^T(k-1) + Q(k-1)$$

(17)

(2) Filtering is to suppress interference noise and make the estimated value as close as possible to the true value. The second norm square sum of the error between the real state at time $k$ and the estimated state at $k$ can be used to describe it, which is also equivalent. Since the trace of the covariance matrix is the smallest, the calculation formula is as follows:

$$\hat{y}(k) = Z(k) - \hat{X}(k | k - 1) = Z(k) - H(k)\hat{X}(k | k - 1)$$

(18)

4.2. Vehicle tracking

After the detection of the target vehicle, the Kalman filter algorithm is used to track the detected vehicle. First, the detected vehicle is positioned, and the center position of the vehicle positioning rectangle is used as the state variable; then feature matching, such as the vehicle in the image, is performed. The center position determines whether the two targets are the same through the center position of the vehicle between the two frames of images; finally, the equations of motion are updated.

Since the position of the target vehicle based on the camera does not change much in two consecutive frames, its movement state can be regarded as uniform motion. In this way, the position change of the target vehicle in the image can also be regarded as uniform velocity change.

The camera assigns the detected vehicle position and speed data to the detection list, and judges the status of the detection target by comparing the data stored in the detection list at adjacent moments. If the target vehicle is detected in the detection lists at the previous and subsequent moments, Then the target vehicle is confirmed to be tracked; if the target is detected at the previous time but not detected at the next time, the target is temporarily considered to be missing, and the target will continue to be detected, if re-detected When the target is reached, the target vehicle is confirmed as tracking, if it is not detected again, the target vehicle is deleted; if a new target vehicle is detected at the current moment, the target vehicle is determined to be active. Continue to detect the new target vehicle, if the target can still be detected at a later time, confirm the target as tracking, otherwise delete the new target vehicle. After determining the status of the detected target vehicle, It is possible to use multiple Kalman filters to predict and update the state of the detected target vehicle, thereby achieving vehicle tracking and partial tracking results. As shown below:

![Figure 7. Vehicle tracking results](image)
5. Summary
This paper studies two vehicle detection methods, one is a vehicle detection method based on multi-class feature fusion proposed in this paper, and the other is a vehicle detection method based on Haar-like features combined with the Adaboost algorithm. Compared with the machine learning method based on a single feature, the vehicle detection method proposed in this paper has better detection effect; in the vehicle tracking part, the principle of kalman filtering is analyzed and the vehicle tracking is realized based on kalman filtering. Finally, the effectiveness of the vehicle detection and tracking method was experimentally verified in different road scenarios, which has profound significance for future vehicle detection and vehicle tracking research and applications.

Acknowledgments
This work was supported by the Key Research and Development Program Project of Anhui Province of China (Grant No. 201904d07020020), the Natural Science Foundation Project of Anhui Province of China (Grant No. 1908085MF212), and the Key Program of the Natural Science Foundation of the Educational Commission of Anhui Province of China (Grant No. KJ2017A063).

References
[1] Li Keqiang, Dai Yifan, Li Shengbo, etc. Development status and trends of intelligent connected car (icc) technology [j]. Journal of Automotive Safety and Energy, 2017, 8 (01): 1-14.
[2] Li Hao, Zhang Yunsheng, Lian Jie, Li Zeping. Multi-angle vehicle dynamic detection method based on deep learning [j]. Traffic Information and Safety, 2017, 35 (05): 37-44.
[3] Song Huansheng, Zhang Xiangqing, Zheng Baofeng, Yan Teng. Vehicle target detection in complex scenes based on deep learning method [j]. Journal of Computer Applications, 2018, 35 (04): 1270-1273.