

## Research Article

# Vehicle Re-Identification System Based on Appearance Features

Dawei Xu <sup>1,2</sup>, Yunfan Yang <sup>2</sup>, Liehuang Zhu <sup>1</sup>, Cheng Dai,<sup>2</sup> Tianxin Chen,<sup>2</sup> and Jian Zhao <sup>2</sup>

<sup>1</sup>School of Cyberspace Science and Technology, Beijing Institute of Technology, Beijing 100081, China

<sup>2</sup>College of Cybersecurity, Changchun University, Changchun 130022, China

Correspondence should be addressed to Liehuang Zhu; [liehuangz@bit.edu.cn](mailto:liehuangz@bit.edu.cn)

Received 28 October 2021; Accepted 22 March 2022; Published 5 May 2022

Academic Editor: Gu Zhaoquan

Copyright © 2022 Dawei Xu et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Aiming at the low recognition accuracy caused by the problems of angle, illumination, and occlusion in vehicle re-identification based on deep learning, a vehicle re-identification method based on multibranch network feature extraction and two-stage retrieval feature is proposed. The multibranch feature extraction module uses ResNet-50 as the backbone network to extract the vehicle's attribute features and apparent features, respectively, and uses the attribute features for rough retrieval. On this basis, the attribute features and apparent features are fused for fine retrieval. Through experiments, the accuracy of vehicle re-identification on Veri-776 data set and VehicleID datasets is significantly improved. In addition, based on the improved algorithm, this paper designs and develops a vehicle re-identification system, which realizes the functions of inputting file directory, selecting target image, and querying result image, and provides a visual technical scheme for vehicle re-identification and retrieval in the real scene.

## 1. Introduction

With the development of the monitoring system, more and more traffic data can be collected, which makes the management and supervision of traffic safety by the traffic control department more convenient. However, how to deal with thousands of monitoring data also poses problems for people. In order to improve the inefficient way of manual screening, more and more researchers have begun to use computers to screen and monitor systems semi-autonomously or autonomously.

In the early days, sensors could be used for specific monitoring tasks such as vehicle counting, license plate recognition, and vehicle detection. However, most of these methods need to deploy control coils on the paved road, which makes the road easily damaged and increases the cost of road construction, and changes in external temperature will affect the current and reduce the accuracy of vehicle identification. With the development of technology, video vehicle detection technology gradually emerges and assumes the huge role of planning and monitoring in smart transportation. The computer analyzes the images collected by the surveillance cameras and further processes the images to

obtain vehicle information such as license plate, vehicle color, car model, and so on in order to more intuitively grasp the traffic information.

Vehicle re-identification can extract various information of the vehicle from the complex picture based on the characteristics of the vehicle (color and car model) and so on in a given picture containing the vehicle to be identified and then identify the system of the vehicle to be detected. In the actual traffic safety management system, vehicle re-identification can be used to monitor, track, and find vehicles. It is important in the fields of the red-light recognition system, speeding alarm system, blacklist inspection system, highway toll system, vehicle dispatching system, vehicle Internet security [1–3], and so on. The application of the vehicle has brought great convenience to the traffic management system, and the field of vehicle re-identification has also been vigorously developed.

In the early days, because the license plate format was the same and the recognition was simple, in theory, the vehicle recognition rate was better, and it was usually used as a general method for vehicle re-identification. In the license plate recognition system, the most important part is image processing, including smoothing, binarization, edge processing,

image segmentation [4], and so on. With the development of computer technology, a variety of processing methods have been derived and the detection accuracy has been greatly improved. However, the image of the vehicle usually captured cannot be clear and complete. The occlusion of the license plate and the deck of the vehicle are the main reasons for reducing the accuracy of vehicle re-identification. Therefore, the current vehicle re-identification training usually considers the situation that the license plate has been occluded. For example, the vehicle data picture in the VehicleID [5] data set has already been occluded the license plate. Therefore, the recognition of vehicles needs to be judged by appearance or attributes, so extracting vehicle features with higher discrimination is the key to people's research. There are several key issues when recognizing vehicles: different cameras can capture images of the same vehicle at different angles, which may be very different; there are many vehicles with similar appearance on the market; due to the material problem of the appearance of the vehicle, different lighting under the same conditions, the same color can sometimes make a big difference. Reducing the impact of these constraints is very important to improve the accuracy of vehicle re-identification, and it is of vital value in the field of transportation and social public safety.

## 2. Related Works

Vehicle re-identification technology refers to a vision technology that retrieves vehicle targets in multiple cameras under a specific monitoring perspective, extracts their features, and matches after similarity measures to determine whether a given vehicle image is the same target.

*2.1. Sensor-Based Vehicle Re-Identification.* Early vehicle re-identification has vehicle re-identification based on sensors and artificial design features. Early vehicle re-identification was mostly based on methods such as ultrasonic, microwave radar, infrared lidar, non-imaging passive infrared, video image processing with visible spectrum, and infrared spectrum images [6–8], on highways and ground street sites. The magnetometer detector technology was evaluated by comparing the actual traffic ground data obtained by counting the number of vehicles in the recorded video image with the count output by the detector [9].

Sensor-based vehicle re-identification is actually to install sensors on the road that the vehicle must pass. The controller will receive the signal generated by the sensor after detecting the passage of the vehicle and calculate the axle load, vehicle speed, and distance from the generated signal. Among them, wireless magnetic sensors [10] and multidimensional sensors [11] were popular at the time. Emerging sensor technologies include radio frequency and so on, such as combining radio frequency identification technology with GPS and global mobile communication system to design and implement accurate vehicle positioning systems [12, 13]. Although sensor-based vehicle re-recognition does not require training and learning, it has high requirements on the external environment. When temperature changes or the sensor is damaged by the outside world, it will increase the

complexity of vehicle re-recognition, the recognition rate is relatively low, and a large amount of hardware is required.

### *2.2. Vehicle Re-Identification Based on Traditional Features.*

With the rise of machine learning, people began to try machine learning methods to solve practical problems, such as using machine learning algorithms to build anomaly, intrusion, and network attack traffic identification models in Internet of things security analysis [14]. In terms of learning methods, it is divided into supervised learning and unsupervised learning. Supervised learning is to solve some classification and regression problems by labeling data tables. Unsupervised learning is not to label the data, which is often used in feature extraction. Feature extraction [15–17] is a key part of machine learning. For example, a new feature selection measurement method CorrAUC is proposed and used to accurately identify intelligent Internet of things anomalies and intrusion traffic [18].

There are many methods for feature extraction, among which SIFT and HOG methods are more common. That is, some of the most representative and stable features are used to represent the object, and it has good stability when the object is blocked or the object changes due to external factors such as light as shown in Figure 1.

HOG has two different calculation methods, static and dynamic. For pictures, we use static calculations to calculate gradients based on histograms, divide the image into cells, calculate the gradients in the blocks, and collect features. HOG divides the image into grids, which is more suitable for some stationary objects. The study by Shafiq et al. [19] is based on the BOW-SIFT method to extract local features. Since SIFT is a local feature of an image, it is relatively stable for some aspects. In the early research on computer vision, Liu et al. [20] proposed a method to re-identify vehicles based on 3D vehicle models and the extraction of the color of the vehicle's top surface, which corrected the shadows and light reflections on wet streets and achieved high recognition accuracy. Woesler [21] used linear regression, color histogram, and directional gradient histogram to solve the re-identification problem. In the study of re-identification, there are methods based on global features [22, 23] and local features [24, 25] in apparent features. Taking face recognition as an example, select the face range in the image and use all the features in this area to represent the object [26]. Such features have a lot of redundant information.

## 3. Preliminaries

### *3.1. Vehicle Re-Identification Based on Deep Learning.*

Deep learning has been widely used since it was proposed, such as distributed deep learning Web attack detection system [27]; neighborhood-based travel time estimation (TTE) depth learning method query path [28]; a microblog emotion classification model based on convolutional neural network (CNN) [29]; Wang et al. [30] proposed an intent prediction-based approach named LocJury. LocJury provides location privacy by learning and estimating the intent of location access and will penalize those malicious location accesses.

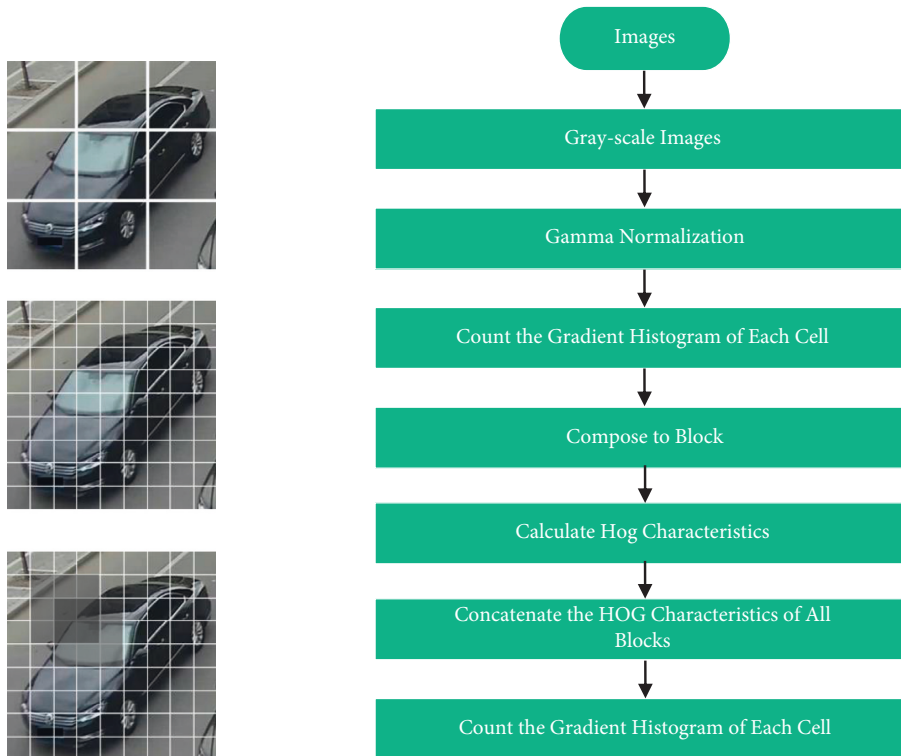


FIGURE 1: HOG model.

This algorithm has gradually become the main research method for vehicle re-identification in recent years. Deep learning has not been popular in the past mainly because the computing power of the equipment is low, but with the generation of a large amount of available data, the birth of more powerful computing equipment and tools has made deep learning a great success in image retrieval, classification [31], and target detection [32]. The typical ones are convolutional neural networks [33, 34] and recurrent neural networks [35]. DenseNet [36] slows down the disappearance of gradient, enhances feature propagation, promotes feature reuse, and reduces network parameters. VGG [37] pushes the depth to 16–19 weight layers, realizing a significant improvement on the configuration of the prior art. GoogLeNet [38] has a deeper level and better performance. Through the comparison of the above network models, the superior performance of the model in this paper is verified.

In order to solve the impact of different camera perspectives on the appearance of vehicles and the problems of similar vehicles, Tang et al. [39] eliminated the need to manually label attribute information and created a randomly synthesized data set with automatically marked vehicle attribute characteristics. Explicitly infer vehicle pose and shape via keypoints, heatmaps, and segments from pose estimation. Semantic vehicle attributes (colors and types) are classified through multitask learning with the embedded pose representations. Sun et al. [40] realized multiscale feature extraction, extracting high-resolution feature maps and low-resolution feature maps in parallel. In addition to extracting global features, Liu et al. [41] also extracted features from a series of local regions and used vehicle ID, model, and color to train RAM.

The current vehicle re-identification always uses a classification search method from coarse to fine, such as screening by vehicle color and vehicle model, and then adding some unique characteristics of the vehicle to search after narrowing the scope. Since deep learning is mainly based on neural networks, it has good learning ability and adaptability, can extract features more efficiently, can be used in problems in different fields, and is easy to transplant.

First, let us introduce the CNN network. It is divided into input layer, convolutional layer, pooling layer, and fully connected layer. Each layer of the network has multiple neurons, which are mapped through activation functions [42]. In the convolutional layer, we use the convolution operation to make the picture smaller. In the pooling layer, we need to down-sample. Common pooling methods are maximum pooling and average pooling. The former is to take the maximum value of the sliding window, and the latter is to take the average value. Parameters are reduced by extracting features from convolution kernels shared by neurons in the same layer. This step is repeated to extract as many features as possible and finally perform classification in a fully connected layer.

When the data are passed through the network, we need to classify the results obtained from the input values, update the weights, and backpropagate them. To strengthen the expression ability of the model, we need the activation function. Common activation functions are Sigmoid, ReLU, Leak ReLU functions, and so on. Sigmoid can be used in the last layer of the neural network, which can change the continuous input value into the output between 0 and 1 [43]; the formula is as follows:

$$f(x) = \frac{1}{1 + e^{-x}}. \quad (1)$$

Among them, when the input value of the ReLU function is less than 0, the output value is 0; when the input value is greater than or equal to 0, the output value is equal to the input value [44]; the formula is as follows:

$$f(x) = \max(0, x). \quad (2)$$

**3.2. Difficulties about Identifying Vehicles.** When dealing with the problem of vehicle re-identification, there are situations such as unlicensed vehicles, false license plate information, occlusion, and defacement. The traditional joint classifier based on artificially extracted features has a low correct rate of vehicle re-identification. First, the change in illumination makes the resolution of the image lower, and the color of the vehicle in the image may be quite different from the actual color. In addition, in the case of similar appearance, occluding the logo will affect the recognition accuracy. Therefore, it is not enough to be based on the appearance characteristics. It is necessary to introduce the attribute characteristics of the vehicle to enhance the discrimination of the vehicle re-identification.

**3.3. Main Work of the Paper.** This paper studies an appearance-based vehicle re-recognition algorithm. It extracts the attributes and appearance features of vehicles through a multibranch network, uses the attribute features to perform a rough search on the vehicle, and then fuses the attributes and appearance features for a fine search. This work will focus on how to optimize the network structure to improve the discrimination of vehicles under occlusion or overexposure. The comparative experimental analysis using the public large-scale Veri-776 data set and VehicleID [5] data set shows that the algorithm can really improve the generalization ability of the vehicle re-recognition model. This has a great impact and positive significance on the promotion of intelligent transportation construction and other fields.

## 4. Research Motivation and Overall Design

**4.1. Overall Design.** The entire model framework of this article is based on ResNet-50. First, a residual network ResNet-50 built on the ImageNet database is used to extract the basic feature vector of the vehicle. The entire model is shown in Figure 2.

The whole model has two branch networks to calculate the attributes and appearance characteristics of the vehicle. Since the network can perform input and output at the same time, it constitutes a multibranch network that simultaneously extracts attribute features and apparent features. Among them, the multibranch feature extraction module uses ResNet-50 as the backbone network to extract attribute features and apparent features and uses Triplet loss joint training to make the two influence each other and improve feature discrimination; two-stage retrieval makes full use of

the two extracted features; the re-identification of vehicles is carried out from coarse to fine; first the attribute characteristics are used to eliminate the vehicle color; and vehicles of different models, after narrowing the scope, merge the apparent characteristics and attribute characteristics to screen the details, so that the classification accuracy of the vehicles is improved. The similarity measure obtains the 10 pictures that are closest to the target to be detected.

### 4.2. Feature Learning Network Based on Joint Multiattribute Branches

**4.2.1. Deep Learning Network.** Common high-precision networks are AlexNet, VGGNet, GoogLeNet, ResNet, DenseNet, and so on based on the ImageNet model. These networks can have stronger expressive power after multiple linear transformations and can also integrate multi-perspective and multiscale information and carry out learning from different perspectives. They have more network paths and have both multimodel fusion and layering in the network. Direct supervision can have better characterization ability for images with multiple features.

**4.2.2. Loss Function.** This article uses the triplet loss function to achieve sample similarity calculation by optimizing the distance between the sample and the positive sample to be less than the distance between the sample and the negative sample [45]; the formula is as follows:

$$L = \max(d(a, b) - d(a, c) + \text{margin}, 0), \quad (3)$$

where  $a$  is a sample,  $b$  is a sample of the same type as  $a$ ,  $c$  is a sample of a different type from  $a$ , and  $d$  is the distance between the two. The purpose is to shorten the distance between  $a$  and  $b$ .

**4.3. Two-Stage Search from Coarse to Fine.** Since vehicles have a very similar appearance when the manufacturer, model, and color are the same, at this time, only the attributes and characteristics of the vehicle, such as the color and model of the vehicle, cannot be accurately identified. When the apparent characteristics of vehicles such as license plates cannot be accurately identified due to problems such as light and occlusion, the discrimination of similar vehicles will be reduced. Therefore, this paper proposes the contribution of two-stage retrieval. When performing vehicle re-identification, the common classification mainly relies on vehicle ID, color, car model, and so on, and the backbone network ResNet-50 is used to extract the attribute characteristics and appearance characteristics of the vehicle volume. The low-level network does not share the weight, and the high-level network shares weight. The two branch networks, respectively, calculate the attributes and characteristics of the vehicle such as the color and vehicle type and the apparent characteristics of the vehicle, as shown in Figure 3.

First, a rough search is performed using attributes such as the color and model of the vehicle; then, different features are assigned corresponding weights and then serially

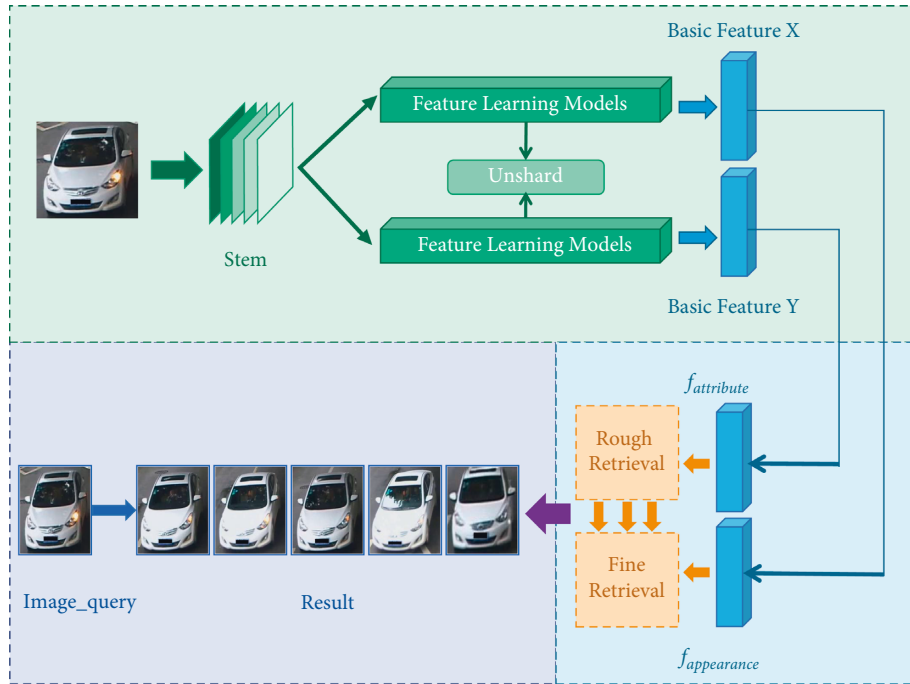


FIGURE 2: System design model.

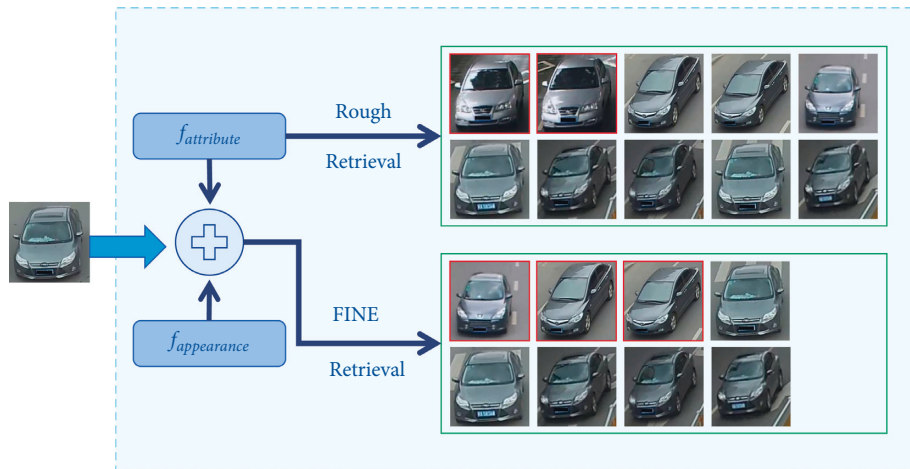


FIGURE 3: A two-stage retrieval scheme from coarse to fine.

stitched. The attributes and appearance features of the vehicle are fused together to form the attribute identity feature of the vehicle, which is used for the vehicle and finely searched.

The feature vector of each ID has an obvious clustering effect, and it can be concluded that different classification basis affects the feature distribution. Whether it is the color of the vehicle or the appearance of the vehicle, the feature vector can be restrained. In the use of attribute features for rough retrieval, when the threshold is less than 0.5, it can be regarded as the same class of samples; on this basis, the vehicle appearance classification constraint is added. After measuring the similarity between the appearance and the vehicle attributes, the attribute vector and the appearance vector are merged, and the same color or the eigenvectors of the samples of the same car model are closer. When the

threshold is less than 0.3, it can be regarded as the same sample. This recognition method is also closer to the general process of re-recognition of vehicles. After a rough search, you can get a picture of a vehicle with the same color, and the wrong picture is selected in the red box. After a fine search, you can get the correct picture. For the results obtained, this paper uses the Euclidean distance to measure the similarity and can get the sorting table of similarity from high to low.

## 5. Experimental Analysis and System Implementation

To verify the performance of the algorithm proposed in this experiment, experiments with different models will be verified on the Veri-776 data set and VehicleID data set. The first part is an experiment based on the basic model, the

second part of the experiment is a model with attribute branching, and the third part of the experiment is a model based on attribute branching and two-stage retrieval. The input picture size is all scaled to 256×256 pixels, the Euclidean distance is used to measure the similarity, and the loss function is calculated using Triplet.

### 5.1. Data set Introduction

**5.1.1. Veri-776 Dataset.** The data of the VeRi-776 data set are derived from images captured by real traffic cameras, and the VeRi-776 data set is actually an extension of VeRi, which is a large-scale data set based on urban traffic. The real traffic scenes covered by this data set include surveillance cameras from 20 different shooting angles. The pictures of each vehicle are captured from 2 to 18 angles. There are also different environmental conditions such as lighting and resolution as shown in Table 1.

Data images usually include two-lane, four-lane, and intersections, containing 776 vehicles and more than 50,000 images. The pictures are marked with different attributes, such as license plate border, type, color, and brand, as well as sufficient license plate and spatiotemporal information, the time stamp of the shot, and the distance between adjacent cameras. Different information such as vehicle ID and color have interrelated tags.

**5.1.2. VehicleID Dataset.** Although the data available to us during the large-scale construction of monitoring and supervision systems in cities have exploded, compared with the more mature and leading recognition research on humans, vehicle re-identification does not have a large number of data sets for learning in the early days.

The National Laboratory of Video Technology of Peking University has created a super-large database VehicleID based on real-scene surveillance cameras, which includes multiple images of the same car captured by different real-world cameras scattered in a city during the day, as shown in Table 2.

Each car has more than one photo at both front and rear angles. On average, there are about eight photos of each vehicle. The dataset consists of 221,763 photos and contains views from two cameras. There are a total of 26,267 cars in the dataset. All images have identification numbers that indicate their real identities. In addition, 10,319 vehicles were manually labeled with vehicle model information for a total of 90,196 images. The car model information marked in these 90,000 photos is very detailed and covers more than two hundred common car models on the market. The training set has 13,134 cars and 110178 pictures, and the test set has 13,133 cars with 111585 pictures, which shows that the data set is very large.

### 5.2. Comparative Analysis on Multiple Data

**5.2.1. Evaluating Indicator.** Rank- $k$  is the probability that the top  $n$  graphs in the search results have correct results. Rank-1 is the first hit. Rank- $k$  is to hit within the  $k$ -th time.

TABLE 1: Veri-776 dataset.

File name	File content
Query data set	Contains 1678 query images
Gallery data set	Contains 11579 test images
Training image	Contains 37778 training images

TABLE 2: VehicleID dataset.

Name	Capacity
Vehicle colors	7
Number of images	221763
Number of vehicles	26267
Test image	111585
Training image	110178

For each picture in the query set, we will get a table sorted according to the similarity, and formula is as follows:

$$S_M = \sqrt{\sum_{k=1}^N (q_k - a_{Mk})^2}, \quad (4)$$

where  $q_k$  is the feature vector of the query image and  $a_{Mk}$  is the feature vector of the  $m$ -th image in the gallery image library. The smaller the value of  $S_M$  is, the more similar the images are. On the contrary, the larger the value of  $S_M$  is, the lower the similarity is between images. Ranking  $S_M$  from low to high, the accuracy of Rank- $k$  can be calculated, respectively.

mAP is the average accuracy of average, which is mainly used for classification tasks and as a common evaluation index. mAP adds the average accuracy of each classification result and then averages it. Formula is as follows:

$$ap = ap + (\text{recall} - \text{oldrecall}) * \left( \frac{\text{oldprecision} + \text{precision}}{2} \right). \quad (5)$$

Precision stands for accuracy, which is the ratio of the number of correct information extracted to the number of information extracted. Recall stands for recall rate, which is the ratio of the number of correct information extracted to the amount of information in the sample.

**5.2.2. Experimental Analysis.** The experiment mainly analyzes the influence of different characteristics on the result of vehicle re-identification.

Baseline + Attri is an improved vehicle re-recognition model that includes attribute branch networks, and Baseline + Attri + Par is a two-stage retrieval model that includes attribute characteristics and appearance characteristics. This article compares other different vehicle re-identification methods, as shown in Table 3.

DenseNet121 [46], PROVID [47], and VGG + CTS [48] are 45.1%,53.4%,and 58.3% on mAP, respectively. The improved vehicle re-identification model has different degrees of improvement on mAP, Rank-1, Rank-5, and Rank-10. It shows that the improved method in this paper has a good effect on the accuracy of vehicle re-identification in terms of

TABLE 3: Comparative analysis results based on Veri-776 dataset.

Model	Rank-1 (%)	Rank-5 (%)	mAP (%)
DenseNet121 [46]	80.3	91.1	45.1
PROVID [47]	81.6	95.1	53.4
VGG + CTS [48]	83.5	90.0	58.3
<b>Baseline + Attri + Part</b>	<b>94.4</b>	<b>97.8</b>	<b>70.92</b>

TABLE 4: Comparative analysis results based on VehicleID dataset.

Model	Rank-1 (%)	Rank-5 (%)
BOW-CN [49]	13.14	22.69
GoogLeNet [50]	47.90	67.43
FACT [20]	49.53	67.96
<b>Baseline + Attri + Part</b>	<b>73.9</b>	<b>85.1</b>

appearance features and attribute features, as shown in Table 4.

BOW-CN [49], GoogLeNet [50], and FACT [20] can see that the improved vehicle re-identification model has different degrees of improvement in Rank-1 and Rank-5. It shows that the improved method in this paper has a better effect on the accuracy of vehicle re-recognition in terms of appearance characteristics and attribute characteristics. Compared with other mainstream methods, it can be seen that the method in this paper has more advantages.

*5.3. Ablation Analysis of Multiattribute Branch and Two-Stage Retrieval.* When vehicle attribute parameters are added, mAP increases by 1.53% and 1.8%, respectively, compared to the base model. Rank-1 increases by 1.05% and 2.1% on both datasets, and Rank-5 increases by 0.58% and 0.4% on both datasets. Rank-10 increases by 0.76% and 0.6% on the two datasets, as shown in Tables 5 and 6.

After adding the attribute branch, the accuracy of vehicle re-identification is significantly improved. It also illustrates the importance of adding appearance feature branches. In the case of optimizing vehicle attributes and appearance features, mAP is improved by 2.43% and 2.7% on the two datasets, Rank-1 is improved by 1.85% and 2.9%, and Rank-5 is improved by 0.84% and 0.8%. Rank-10 improves by 0.87% and 1.0%, as shown in Tables 5 and 6.

When the two-stage retrieval of vehicle attributes and appearance characteristics is added, the accuracy of vehicle re-recognition has been further improved, which also illustrates the importance of adding attribute branches.

*5.4. Visual Analysis of Vehicle Re-Identification Based on Appearance Features.* To make the result of vehicle re-identification more obvious, this article visually displays the result of vehicle re-identification. As shown in Figure 4, the left column is the selected target picture and the pictures on the right are the search results based on the target picture and from left to right are from Rank-1 to Rank-6.

*5.5. System Function Realization.* The purpose of vehicle re-identification system writing is as follows: the vehicle re-identification system is user-oriented, and the purpose is to visualize the vehicle re-identification experiment and to facilitate the presentation of the results of the vehicle re-identification experiment. The system includes a selection target picture module, a result picture display module, and a target picture display module.

The background of the vehicle re-identification system is as follows: after extracting the feature vector, the feature vector is measured for similarity, and based on the result of the measurement, the ten pictures that are most like the picture to be detected are retrieved.

The functional division of the vehicle re-identification system is as follows: (1) select the target picture file and (2) select the target picture.

The function description of the vehicle re-identification system is as follows: (1) enter the name of the folder where the target picture is located and select the target picture in the corresponding data set. (2) Select the target picture serial number; this function can show the superiority of the improved code. (3) The most similar pictures are shown. The system flow chart of the vehicle re-identification system is shown in Figure 5.

The user opens the interface and enters the system. After entering the file path, it will be connected to the database of the target image. After entering the serial number of the target image, the pictures sorted from high to low similarity will be transferred from the searched image library, and these pictures will be displayed on the system desktop and exit the system after finishing.

As shown in Figure 6, when switching the target image, you can click the forward and backward buttons. Of course, you can also enter the serial number of the target image after the serial number button below and click the search button.

For example, enter the serial number 80, as shown in Figure 7; the target image display area shows the image ranked according to the similarity with the current vehicle to be retrieved. The retrieved vehicle shown in the first three rows is consistent with the target to be detected. The last vehicle image in the second row has the same color as the current vehicle to be detected, but the vehicle direction is inconsistent.



TABLE 5: Comparative analysis results based on Veri-776 dataset.

Model	Rank-1 (%)	Rank-5 (%)	Rank-10 (%)	mAP (%)
Baseline	92.55	96.96	98.09	68.49
Baseline + Attri	93.6	97.54	98.76	70.02
<b>Baseline + Attri + Part</b>	<b>94.4</b>	<b>97.8</b>	<b>98.96</b>	<b>70.92</b>

TABLE 6: Comparative analysis results based on VehicleID dataset.

Model	Rank-1 (%)	Rank-5 (%)	Rank-10 (%)	mAP (%)
Baseline	71.0	84.3	90.3	37.8
Baseline + Attri	73.1	84.7	90.9	39.6
<b>Baseline + Attri + Part</b>	<b>73.9</b>	<b>85.1</b>	<b>91.3</b>	<b>40.5</b>

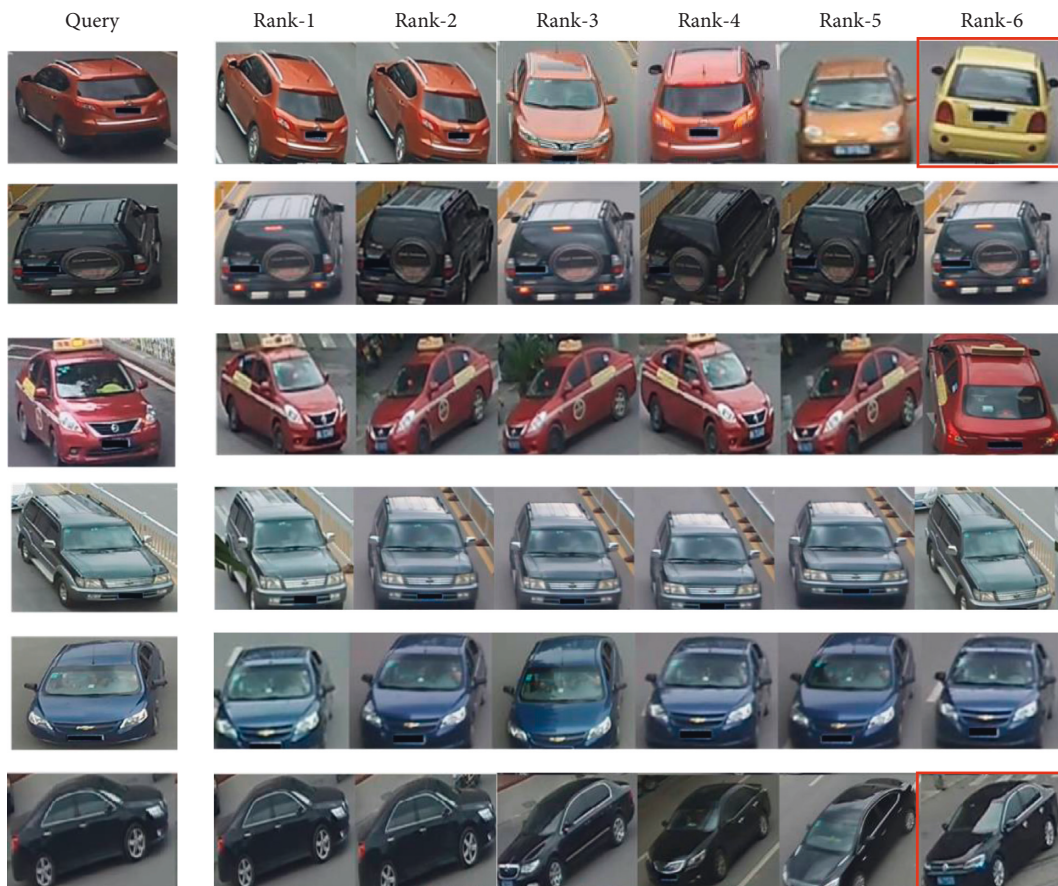


FIGURE 4: Top six similarity search results.

For Figure 8, we use the unmodified baseline to conduct the experimental test. The images in the third row and fourth column in the figure are wrong. However, for Figure 9, the

test using the method in this paper shows that the first 12 images are error free, so the improved results in this paper are still good.



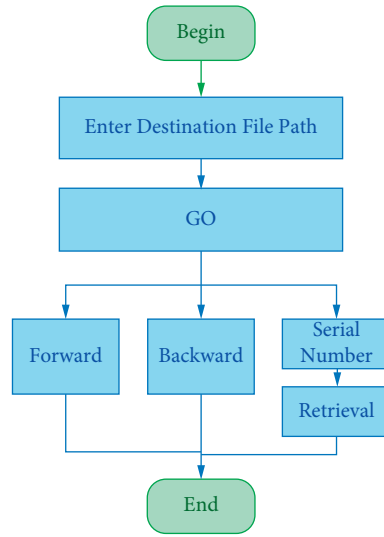


FIGURE 5: System flow chart.



FIGURE 6: Vehicle search results 1.

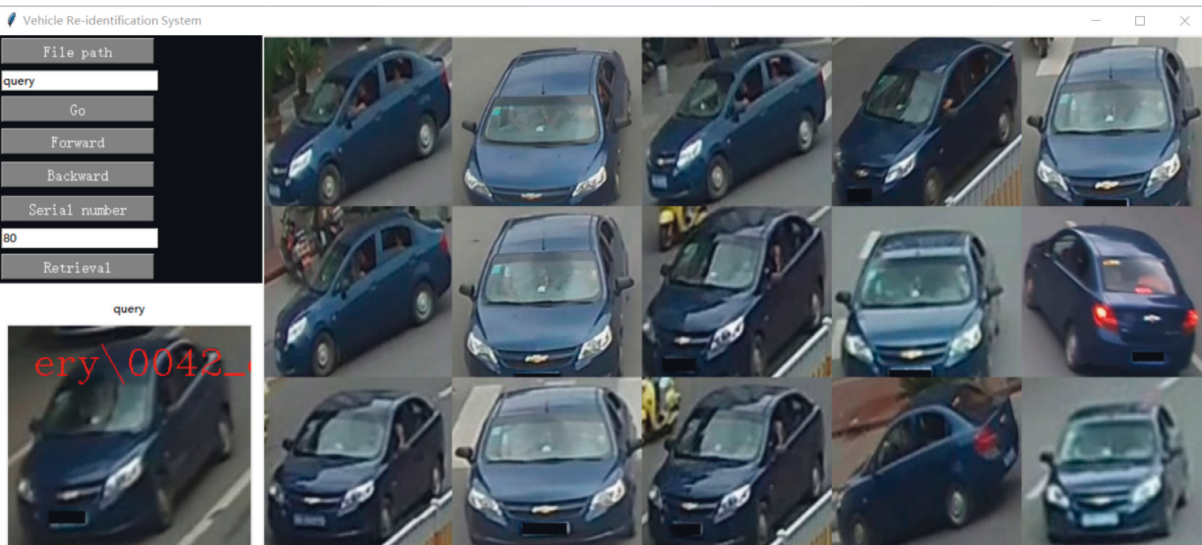


FIGURE 7: Vehicle search results 2.

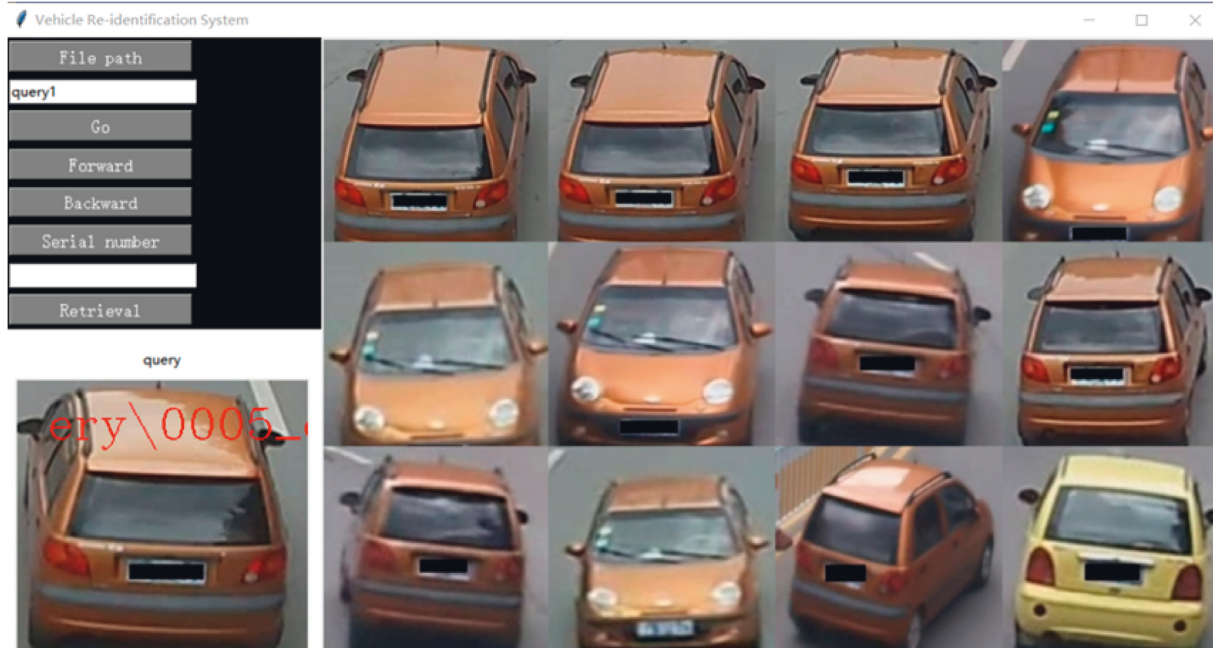


FIGURE 8: Vehicle search results 3.

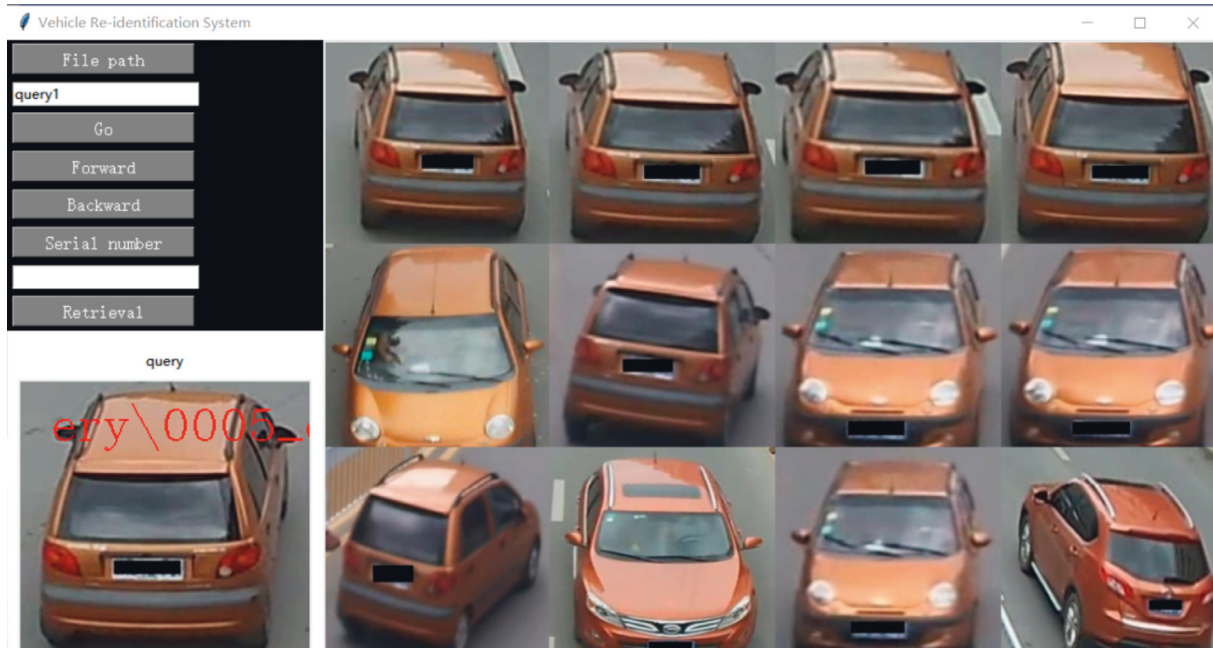


FIGURE 9: Vehicle search results 4.

## 6. Conclusion

To improve the accuracy of vehicle re-recognition, this paper proposes a vehicle re-recognition method with multibranch network feature extraction and two-stage retrieval feature. The main research contribution is the addition of attribute features. Among them, the multi-branch feature extraction module uses ResNet-50 as the backbone network to extract the attribute features and

apparent features of the vehicle, respectively, and use the attribute features to search for rough and, on this basis, use the apparent feature fine search. The attributes and characteristics of the vehicle include vehicle color and model, which can maintain good stability in different environments. It has been well verified on the Veri-776 data set and VehicleID data set. The effectiveness of the method is further proved by the verification of the comparative experiment.

## Data Availability

All data for this study and method have been included in the paper. Therefore, no additional data are required.

## Conflicts of Interest

The authors declare that there are no conflicts of interest regarding the publication of this paper.

## Acknowledgments

This research was supported by the Scientific Research Project of the Education Department of Jilin Province (no. JJKH20220602KJ).

## References

- [1] Z. Tian, X. Gao, S. Su, and J. Qiu, "Vcash: a novel reputation framework for identifying denial of traffic service in internet of connected vehicles," *IEEE Internet of Things Journal*, vol. 7, no. 5, pp. 3901–3909, 2020.
- [2] Z. Tian, X. Gao, S. Su, J. Qiu, X. Du, and M. Guizani, "Evaluating reputation management schemes of internet of vehicles based on evolutionary game theory," *IEEE Transactions on Vehicular Technology*, vol. 68, no. 6, pp. 5971–5980, 2019.
- [3] S. Su, Z. Tian, S. Liang, S. Li, S. Du, and N. Guizani, "A reputation management scheme for efficient malicious vehicle identification over 5G networks," *IEEE Wireless Communications*, vol. 27, no. 3, pp. 46–52, 2020.
- [4] C. Nie, W. Heng, J. Shi, and M. Zhang, "Optimizing actuated traffic signal control using license plate recognition data: methods for modeling and algorithm development," *Transportation Research Interdisciplinary Perspectives*, vol. 9, 2021.
- [5] H. Liu, Y. Tian, Y. Wang, P. Lu, and H. Tiejun, "Deep relative distance learning: tell the difference between similar vehicles," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 167–2175, Las Vegas, NV, USA, June 2016.
- [6] N. Sun, H. Qin, and L. Zhang, "GE Ruhai Vehicle target recognition method based on multi-sensor information fusion," *[J] Automotive Engineering*, vol. 39, no. 11, pp. 1310–1315, 2017.
- [7] A. Liu, *Research on Vehicle Detection and Recognition Algorithm Based on TMR Sensor*, Harbin Institute of technology, Harbin, China, 2016.
- [8] X. Chaonan, *Research on Vehicle Driving State Recognition Technology Based on Motion Sensor*, Beijing University of Posts and telecommunications, Beijing, China, 2019.
- [9] R. O. Sanchez, C. Flores, R. Horowitz, R. Ram, and V. Pravin, "Vehicle re-identification using wireless magnetic sensors: algorithm revision, modifications and performance analysis," in *Proceedings of the 2011 IEEE International Conference on Vehicular Electronics and Safety*, pp. 226–231, IEEE Press, Beijing, China, June 2011.
- [10] S. Charbonnier, A. C. Pitton, and A. Vassilev, "Vehicle re-identification with a single magnetic sensor," in *Proceedings of the 2012 IEEE International Instrumentation and Measurement Technology Conference*, pp. 380–385, IEEE Press, Graz, Austria, May 2012.
- [11] J. Prinsloo and R. Malekian, "Accurate vehicle location system using RFID, an internet of things approach," *Sensors*, vol. 16, no. 6, pp. 825–848, 2016.
- [12] Z. Qiang, M. Qiu, Z. Chen, D. Liu, L. Fu, and Yu Xiaoqing, "Research on long-distance high-altitude transport vehicle tracking system based on Beidou satellite positioning and RFID," *Urban housing*, vol. 28, no. 09, pp. 224–226, 2021.
- [13] Z. Yinlong, *Design and Development of Vehicle Monitoring System Based on GPS/GPRS/RFID*, Nanjing University of Aeronautics and Astronautics, Nanjing, China, 2014.
- [14] M. Shafiq, Z. Tian, Y. Sun, X. Du, and M. Guizani, "Selection of effective machine learning algorithm and Bot-IoT attacks traffic identification for internet of things in smart city," *Future Generation Computer Systems*, vol. 107, pp. 433–442, 2020.
- [15] Y. Li, Y. Zou, and J. Ma, "Document block image classification algorithm based on feature extraction and machine learning," *Signal processing*, vol. 35, no. 05, pp. 747–757, 2019.
- [16] Xu Wei, *Research on EEG Brain Age Based on Feature Extraction and Machine Learning Algorithm*, Wenzhou University, Wenzhou, China, 2018.
- [17] Yi Zheng, *Research on Capsule LSTM Feature Extraction Algorithm for Time Series Data*, Central China Normal University, Wuhan, China, 2018.
- [18] M. Shafiq, Z. Tian, A. K. Bashir, X. Du, and M. Guizani, "CorrAUC: a malicious bot-IoT traffic detection method in IoT network using machine-learning techniques," *IEEE Internet of Things Journal*, vol. 8, no. 5, pp. 3242–3254, 2021.
- [19] M. Shafiq, Z. Tian, A. Bashir, X. Du, and M. Guizani, "IoT malicious traffic identification using wrapper-based feature selection mechanisms," *Computers & Security*, vol. 94, Article ID 101863, 2020.
- [20] X. Liu, W. Liu, and H. Ma, "Large-scale vehicle re-identification in urban surveillance videos," in *Proceedings of the 2016 IEEE International Conference on Multimedia and Expo (ICME)*, pp. 1–6, IEEE Press, Seattle, WA, USA, June 2016.
- [21] R. Woesler, "Fast extraction of traffic parameters and re-identification of vehicles from video data," in *Proceedings of the 2003 IEEE International Conference on Intelligent Transportation Systems*, pp. 774–778, IEEE Press, Shanghai, China, October 2003.
- [22] Y. Wan, Y. Wang, X. Zhang, and D. Li, "Chen Xiaolin Colorization of antagonistic gray image combined with global semantic optimization," *Liquid crystal and display*, vol. 36, no. 09, pp. 1305–1313, 2021.
- [23] K. Zhu and J. Zhang, "Tan Shuqiu Pedestrian re recognition based on global features and multiple local features," *Microelectronics & Computer*, vol. 39, no. 02, pp. 43–50, 2022.
- [24] C. T. Wang, N. Zhang, and L. Yujia, "Vehicle light recognition based on convolutional neural network with local features," *Journal of Changchun University of Technology(Natural Science Edition)*, vol. 45, no. 01, pp. 16–23, 2022.
- [25] Y. Tu, He Song, and Y. Shaohua, "Pedestrian re recognition method based on local feature fusion," *Intelligent computer and application*, vol. 11, no. 12, pp. 122–125, +132, 2021.
- [26] L. Fu, *Research on Face Recognition Method Based on Multi Feature Representation*, Harbin University of technology, Harbin, China, 2020.
- [27] Z. Tian, C. Luo, J. Qiu, X. Du, and M. Guizani, "A distributed deep learning system for Web attack detection on edge devices," *IEEE Transactions on Industrial Informatics*, vol. 16, no. 3, pp. 1963–1971, 2020.
- [28] J. Qiu, L. Du, D. Zhang, S. Su, and Z. Tian, "Nei-TTE: intelligent traffic time estimation based on fine-grained time

- derivation of road segments for smart city,” *IEEE Transactions on Industrial Informatics*, vol. 16, no. 4, pp. 2659–2666, 2020.
- [29] D. Xu, Z. Tian, R. Lai, X. Kong, Z. Tan, and W. Shi, “Deep learning based emotion analysis of microblog texts,” *Information Fusion*, vol. 64, pp. 1–11, 2020.
- [30] Y. Wang, Z. Tian, Y. Sun, X. Du, and N. Guizani, “LocJury: an IBN-based location privacy preserving scheme for IoCV,” *IEEE Transactions on Intelligent Transportation Systems*, 2020.
- [31] Q. Wang, H. Jin, W. Li, and Q. Li, “Small sample image classification algorithm based on multi-scale channel attention mechanism,” *Journal of Hubei University of technology*, vol. 37, no. 01, pp. 34–39, +70, 2022.
- [32] X. Fu and D. Zhu, “Overview of deep learning target detection methods,” *Computer system application*, vol. 31, no. 02, pp. 1–12, 2022.
- [33] Y. Jianlong and H. Xinhai, “Research on image recognition algorithm based on convolutional neural network,” *Journal of Anyang Normal University*, no. 05, pp. 14–18, 2021.
- [34] K. Zhang, F. Xiaohan, Y. Guo et al., “Overview of deep convolution neural network models for image classification,” *Chinese Journal of image and graphics*, vol. 26, no. 10, pp. 2305–2325, 2021.
- [35] Y. Wang, B. Liao, P. Chen, and J. Li, “Yin Yumin Review of recurrent neural networks,” *Journal of Jishou University (Natural Sciences Edition)*, vol. 42, no. 01, pp. 41–48, 2021.
- [36] Y. Zhang, “Zhang Zan Application of densenet in voiceprint recognition,” *Computer engineering and science*, vol. 44, no. 01, pp. 132–137, 2022.
- [37] J. Cai, J. Du, Q. Wang, and H. R. Zhou, “Face emotion recognition based on VGG16 network,” *Electronic Technology Application*, vol. 48, no. 1, pp. 67–70+75, 2022.
- [38] F. Y. Zhou, L. P. Jin, and J. Dong, “Review of convolutional neural network,” *Chinese Journal of Computers*, 2017.
- [39] Z. Tang, M. Naphade, S. Birchfield et al., “PAMTRI:pose-aware multi-task learning for vehicle re-identification using highly randomized synthetic data,” in *Proceedings of the IEEE International Conference on Computer Vision*, pp. 211–220, IEEE Press, Seoul, Korea, October 2019.
- [40] K. Sun, B. Xiao, D. Liu, and J. Wang, “Deep high-resolution representation learning for human pose estimation,” in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 5693–5703, IEEE Press, Long Beach, CA, USA, June 2019.
- [41] X. Liu, S. Zhang, Q. Huang, and G. Wan, “Ram: a region-aware deep model for vehicle re-identification,” in *Proceedings of the 2018 IEEE International Conference on Multimedia and Expo (ICME)*, pp. 1–6, IEEE Press, San Diego, CA, USA, July 2018.
- [42] H. Zhang, Q. Zhang, and J. Yu, “Review on the development of activation function and its property analysis,” *Journal of Xihua University (NATURAL SCIENCE EDITION)*, vol. 40, no. 04, pp. 1–10, 2021.
- [43] Y. Lin, W. Yunlong, Q. Chen, W. Zhang, and Z. Qiu, “A sigmoid function optimization method for embedded computing platform,” *[J] Small microcomputer system*, vol. 42, no. 10, pp. 2053–2058, 2021.
- [44] angbo Jiang and W. Wang, “Study on Optimization of relu activation function,” *Sensors and Microsystems*, vol. 37, no. 02, pp. 50–52, 2018.
- [45] J. Zhao, Y. Zhang, X. Shi, and Y. Li, “Research on multi view gait recognition method based on joint loss function,” *Intelligent computer and application*, vol. 12, no. 02, pp. 13–17, 2022.
- [46] G. Huang, Z. Liu, L. V. D. Maaten, and Q. W. Kilin, “Densely connected convolutional networks,” in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Honolulu*, pp. 2261–2269, IEEE, Honolulu, HI, USA, July 2017.
- [47] X. Liu, W. Liu, T. Mei, and H. Ma, “A deep learning-based approach to progressive vehicle Re-identification for urban surveillance,” in *Proceedings of the Computer Vision - ECCV 2016 European Conference on Computer Vision*, pp. 869–884, Springer, Amsterdam, The Netherlands, October 2016.
- [48] Y. Zhang, D. Liu, and Z. J. Zha, “Improving triplet-wise training of convolutional neural network for vehicle re-identification,” in *Proceedings of the IEEE International Conference on Multimedia and Expo*, pp. 1386–1391, IEEE, Hong Kong, China, July 2017.
- [49] L. Zheng, L. Shen, T. Lu, S. Wang, J. Wang, and T. Qi, “Scalable person re-identification: a benchmark,” in *Proceedings of the IEEE International Conference on Computer Vision*, pp. 1116–1124, Santiago, Chile, December 2015.
- [50] C. Szegedy, W. Liu, Y. Jia et al., “Going deeper with convolutions,” in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 1–9, Boston, MA, USA, June 2015.