International Journal of Research in Engineering and Science (IJRES)
ISSN (Online): 2320-9364, ISSN (Print): 2320-9356
www.ijres.org Volume 4 Issue 8 ǁ August. 2016 ǁ PP. 25-30

Vehicle Recognition Based On Convolutional Neural Network

XinchenWang [1], XunchengWu[1], WeiweiZhang[1]
[1](College Of Automotive Engineering, Shanghai University Of Engineering Science, China)

Abstract: *Vehicle recognition technology can play a key role in Intelligent Transport Systems.With the development of image processing, pattern recognition and deep learning, vehicle identification technology based on deep learning is becoming more and more concerned.According to the vehicle identification problem of urban traffic intersection, the theoretical basis of the convolutional neural network (CNN) is quoted,and the corresponding vehicle feature extraction algorithm is designed, and combined with the SVM classifier to optimize the convolution network, so as to build the model of vehicle identification. Experimental results show that the application of this model built on CNN identify issuescan raise the accuracy rate by5.2% higher than HOG, and 8.5% higher than the PCA + SIFT, the car is easier to distinguish with the other two models.The main reason forthe classification error to truck and bus is the appearance of the bus is very similar to a shed truck.*
Keyword : *Convolutional neural network, deep learning,vehicle identification, pattern recognition,intelligent transportation*

I. INTRODUCTION

Transportation problem is one of the crucial issues faced by many big cities.With the development of computer technology and artificial intelligence technology, intelligent transportation will become a hot research topic. Vehicle recognition is one of the key technologies of intelligent transportation.Intelligent vehicle identification in vehicle management, illegal escape vehicle, a vehicle inspection surveillance and many other issues have played a key role. At present domestic and international mainstream vehicle recognition methods are as follows: sense coil detection, infrared detection method, the dynamic piezoelectric detection method, etc. Although the recognition accuracy is still relatively high, the installation process will affect the traffic order normal operation, equipment is also easily to damage, and difficult and expensive maintenance.Vehicle recognition method based on image processing acquires the vehicle image through high precision industrial camera and image acquisition card. Then computer is used to simulate the function of human visual effect, analyze to extract the required information, such as vehicle license, color, shape and other features. Finally, we use pattern recognition method to distinguish between different models. In this paper, the method of vehicle recognition is based on image processing method, using convolutional neural network (CNN) algorithm, which is trained by a large number of sample models, thus the accuracyis improved.This method is essentially different from the traditionalalgorithm in the aspect of recognition rate and with a high quality of the recognition ratein a variety of scenarios.

Convolutional neural network (CNN) is a machine learning model for a deep supervised learning. It has made remarkable achievements in the field of speech recognition and image recognition.For example, Lecun, a professor at New York University in the United States, has usedthe CNNs for handwritten numeral recognition as a tool for Bank of America of identification documents. In 2012, deep learning leading role Prof. Hinton from University of Toronto used deep convolutional neural network image recognition andwon the first prize at a large scale datasets ImageNet.Facebook also used deep convolutional neural network to achieve recognition rate of97.25% on face recognition. Convolutional neural network can be used as the input of the original image, which avoids the additional data preprocessing process in the traditional recognition algorithm. It is similar to biological neural network weight sharing network structurecan produce biological vision of local receptive field effect, reduce network model complexity, and reduce weight number of translation,at the same time it has the high resistant to scale, inclination of visual distortion.

In this paper, CNN is adopted to build the vehicle type feature extraction algorithm. Compared with the traditional vehicle identification method, it has two advantages as follows:

(1) Traditional method is dependent on artificial feature extraction, such as SIFT and HOG algorithm. The problem of these algorithms for the models in the identification is the large amount of computation, it needspreprocessing when the application of the input image. While CNN can directly input the original image, obtain a characteristicin training by autonomous learning, thus improve the efficiency of operation.

(2) The changing shape of the vehicle, the camera shooting distance and angle to the appearance will affect the quality of the final image of the vehicle in the picture, increase the difficulty of image identification. And CNN has a high degree of resistance to the translation, scaling, tilt and other visual deformation, which can effectively overcome these difficulties.

The purpose of this paper is to explore the method of convolution neural network for vehicle image feature extraction and vehicle identification can be done with the combination of these features with SVM classifier. In the end, the results and analysisof the experiment is obtained.

II. VEHICLE RECOGNITION SYSTEM

2.1 Vehicle recognition system structure

The overall process of vehicle recognition system is shown in Figure 1. To begin with, the training samples are applied to train the CNN model for feature extraction. The characteristics of training samples can be obtained.Then the obtained features of the training samples to train the SVM classifier, at the same time the trained SVM classifier model is gotten;nextthe test vehicle images are input to the trained CNN feature extraction model, characteristics of the test vehicleare acquired.At lastextracted feature is input to SVM classifier to obtain the final models by classification.

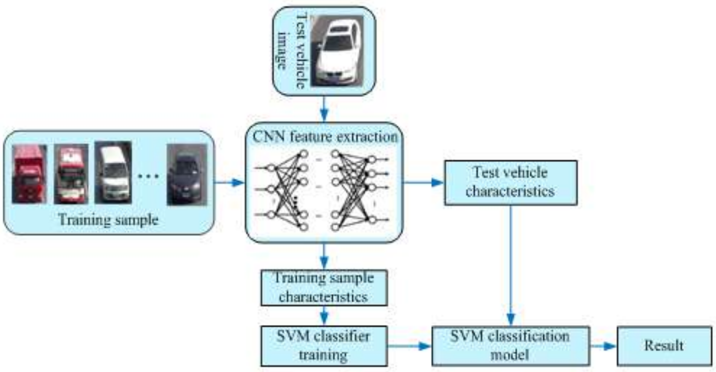


Fig.1 Vehicle identification process blocks diagram

CNNvehicle recognition model includes two parts, such as CNN feature extraction model and SVM classifier. Figure 2 is CNN feature extraction model constructed in this paper. The design of network layers is 5 layers. The number ofC_1 and C_3convolution filters are 6 and 12. The size of the filter is$5\times5=25$.

The Tier I is vehicle grayscale image.Because the color of the vehicle for the vehicle type classification has almost no effect.In order to improve the network efficiency,cvCvtColor function of OpenCVis used to transform vehicle image into gray scale.The size ofgrayscale images is100×80.$C_1$is the first convolutional layer. Itis mainly to extractlow level feature of vehicle image. With 6filters, which the size is5×5,and the input of the grayscale image of the convolution to get 6 feature maps, the size ofeach feature map is96×76. $S_2$is the first subsampling.Subsampling layer is the output of a massive regional sampling into an output value.The mean pooling method adds every bit then calculates the average value, thus replace the region. In this paper, the subsampling layer uses mean pooling method, in the neighborhood of four pixels are summed for average value, thereby to replace the region, not only reduces the dimensions, but can achieve the effect of translation invariant, so S_2layer has six feature maps, the size of each feature map is48×38. The subsampling layer also needs an activation function.The introduction of the activation function can improve the network and the nonlinear characteristics of the decision function.The sigmoid activation function ($f(x) = 1/(1+e^{\wedge}(-x))$) is chosen in this paper. C_3 is the second convolutional layer of 12filters, which the size of each is5×5 and S_2layer is obtained by convoluting 12 feature maps and the size of each feature map is44×34. S_4 is the second subsampling layer, the way is the same as S_2, so the S_4layer has 12 feature maps, the size of each feature map is 22×17.F_5isa completed connected layer. CNN high-level features are obtainedby the last completed connected layer. The connected layer will connect all neurons in frontier layers. The obtained feature maps will be arranged in a column vector to obtain the final feature vector.As a result, one dimensional column vector is acquired. In addition, the last layer also need an activation function, then select the ReLu function as the activation function ($f(x)=max(0, x)$).

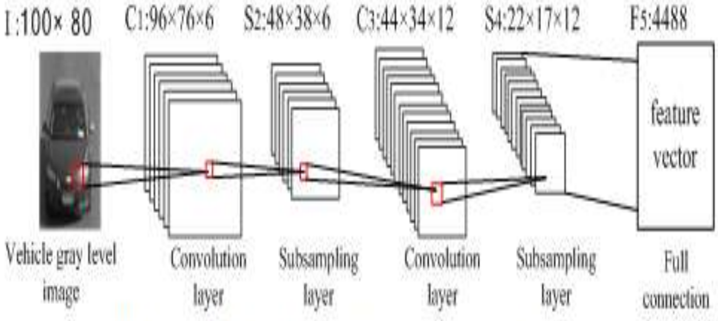


Fig.2 Model of vehicle recognition CNN features extraction

Fig.2 Model of vehicle recognition CNN features extraction

2.2 convolutional neural network training

Convolutional neural network training consists of two stages, forward and backward propagation.The forward and backward propagation is different according to the network layer.

(1) Forward propagation. The parameters of the forward propagation network are kept constant, and the output of the network is calculated according to the excitation function of each layer node, and then the difference between the output value and the known sample label is compared. The parameters are optimized according to the error value from the last layer of the network layer by layer. Assuming a multi-class problem, the number of categories is C, the number of total training samples is N, select the square error as the cost function, the system error can be calculated by formula (1):

$$E^N=\frac{1}{2}\sum_{n=1}^{c}\sum_{k=1}^{c}(t_k^c-y_k^n)^2 \quad (1)$$

In the formula, t_k^nrepresents the label corresponding to the samplen on the classK, and y_k^n indicates the predicted resultsof the system. In the traditional fully connected neural network, the weights of each connection in the network are obtained by optimizing E.The current layer is assumed to be the *l* layer, the output of the current layer is as follows:

$$x^l=f(u^l),\quad u^l=W^l x^{l-1}+b^l \quad (2)$$

As the output activation function, $f(u^l)$can choose different functionaccording to the specific situation, u^l is on and the connection before a layer of all the nodes in the output weighted summation, layer by layer iteration to the last layer of the output is the system output y_k^n. This process is called the forward propagation of the convolutional neural network.

(2) Back propagation. The neural network training process of parameters is the process to optimize the system error, so every time the error in forward propagation system will do backward propagation in the network, then the need to define a variable: "sensitivity" node, which represents the error rate varies with the change of a parameter, with.For example, according to the derivation of the chain rule,deviation unit value (bias) parameteris defined as:

$$\frac{\partial E}{\partial b}=\frac{\partial E}{\partial u}\frac{\partial u}{\partial b}=\delta \quad (3)$$

The recurrence relation between adjacent network layers is:

$$\delta^l=(W^{l+1})^T\delta^{l+1}\circ f'(u^l) \quad (4)$$

∘ representseach element to do multiplication one by one. And each layer is updated according to the corresponding delta value through the formula (5) and (6) to the weights of each connection.

$$\frac{\partial E}{\partial W^l}=x^{l-1}(\delta^l)^T \quad (5)$$

$$\Delta W^l=-\eta\frac{\partial E}{\partial W^l} \quad (6)$$

2.3 Using SVM classifier for convolutional neural network optimization

Support vector machine (Vector Machine Support, SVM) is the linear separable data classification, and the linear nonseparable samples, analgorithm to find the mapping function to map the data from low dimensional space to a high dimensional space. In the paper, using linear SVM method for vehicle recognition. In the experiment, using the LibSVM software package was used for training the classifiers, LibSVM is inventedby National Taiwan University Professor Lin Chih-Jen, such as design and development of a simple, easy to use and fast and efficient SVM pattern recognition and regression software package, it not only provides compiled in the windows series system executable file,but also provides source code, to facilitate improved, modification and application in other operating systems. Therefore, the use of LibSVM software package for the training and verification of features.

III. EXPERIMENT

Fig.3 Part of the model sample pictures

Table.1 Vehicle identification data set

models	Training sample	Test sample
car	2000	1000
bus	1000	500
truck	1000	800

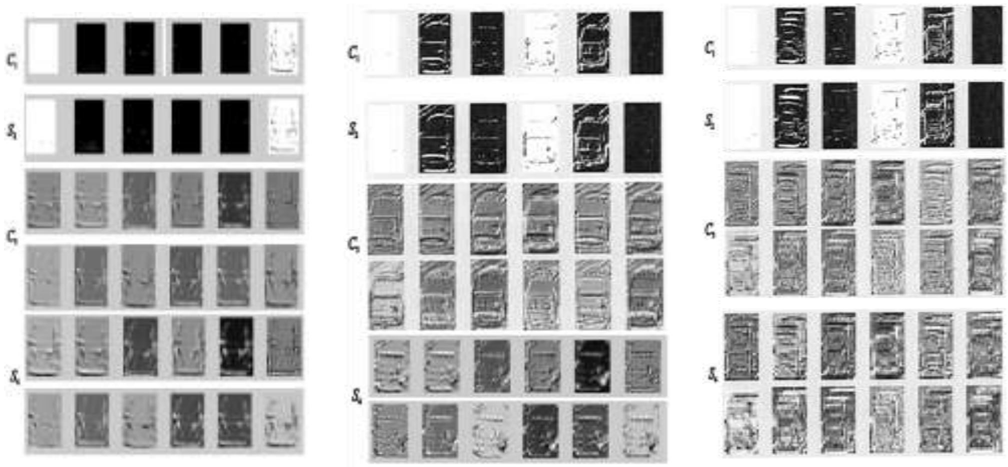
3.1 Experimental data set

This paper adopts the highway surveillance video interception of vehicle image recognition as an object.In order to ensure the accuracy and reliability of the experiment, artificial intercepted and labeling of the image of three models of cars, buses and trucks as test samples and training samples. The total collection of different scales, light and angle is6300pictures. According to the actual situation of the proportion, 3000 car pictures, 1500 bus pictures and 1800 truck pictures are selected. In order to meet the requirements of the convolutional neural network input, the images of three kinds of models are normalized to 100×80 pixels size picture, and then transform into grayscale imagein the OpenCV, the part of the pictures as shown in Figure 3. In the end, 2000 car images, 1000truck and bus picturesare selected as the training samples and the rest is asthe test samples.Table 1 shows the data set of the specific circumstances.

3.2 experimental results and analysis

The features of each model are extracted, as shown in figure 4. Comparing the training of different times of CNN learning characteristics and HOG features,the experimental results is shown in table 2. From accuracy, training of 5 times get the feature classification effect than training of 1 time classification feature increased by 1.1%; training of 10 times than training of 5 times accuracies only increased by 0.1%, but the training time of the former is about two times of the latter, therefore, comprehensive consideration of the accuracy and training time of network, training of 5 times is more appropriate. From the time consumption, the difference between feature extraction times of three kinds of network is small.The training of 5 times of the network to get the feature classifier training time and recognition time is the shortest. Above all, according to the data set of this paper, the number of selected is training of 5 times for feature extraction. Table 2 also makes a comparison of the recognition results between the commonly used HOG features and the PCA + SIFTS features in vehicle identification. In character recognition accuracy rate, we construct the CNNis 5.2%higher than HOG, and 8.5%higherthan PCA +SIFT; from the feature extraction speed of view, CNN is about 2.5 timesfaster than HOG, and about 98 times faster than PCA +SIFT; PCA + SIFT is fastest in the classifier training, it is the 215 times fasterthan CNN, and the 340 timesfasterthan the hog; in classification and recognition time, PCA+SIFT isthe 9.5 timesfasterthan CNN,and the 10 times fasterthan the HOG. Comprehensive accuracy and time consumption, the effect of CNN is the best.

Three models specific classification results are analyzedin Table 3. As can be seen from the table,the recognition rate of the caris the highestby three methods. From the results, the car is easier to distinguish between the other two vehicle types, trucks and buses easily lead to misclassification.The main reasonis that the covered trucks and buses have very similar appearance, coupled with lighting, shooting angles and other factors, which bring more recognition difficulty. From the characteristic graph extracted from Figure 4, buses and trucks extracted characteristic is similar to a certain extent, thus it has a certain impact on the final classification accuracy.



(a) Feature of each layer of car **(b)** Feature of each layer of truck **(c)** Feature of each layer of bus
Fig.4 Vehicle features map

Table.2CNN characteristic performance comparison

CNN					
CNN training times	n=1	n=5	n=10	HOG	PCA+SIFT
CNN training time /s	55.8	272.3	540.2		
Feature extraction time /ms	15	15	16	38	1475
Classifier training time /s	5.2	4.3	5.1	6.8	0.02
Classification recognition time /ms	10	9.5	9.5	10	1
accuracy rate	97.5%	98.6%	98.7%	93.4%	90.1%

Table.3 Specific classification results

Algorithm	Accuracy rate	Car		Bus		Truck	
CNN	accuracy rate	99.9%		98.4%		99.1%	
	Error identification number	Truck	Bus	Truck	Car	Bus	Car
		1	0	8	0	6	1
HOG	accuracy rate	95.3%		92.6%		95.25%	
	Error identification number	Truck	Bus	Truck	Car	Bus	Car
		15	32	25	12	30	8
PCA+SIFT	accuracy rate	96.3%		94%		96.25%	
	Error identification number	Truck	Bus	Truck	Car	Bus	Car
		12	25	20	10	24	6

IV. CONCLUSION

Based on convolutional neural network (CNN) theory, this paper is aimed at urban traffic vehicle type recognition problem. It proposes a method to design the corresponding vehicle feature extraction algorithm, and combine with the SVM classifier for convolutional network optimization, so as to construct the recognition model for vehicle classification. Cars are easier to distinguish between the other two models. The main reason for the error classification of trucks and buses is the truck with a shed is very similar to the shape of the bus.

REFERENCES

[1]. La H M, Lim R S, Du J,et al.Development of a small-scale research platform for intelligent transportation system[J].intelligent Transportation Systems， IEEE Transactions on,2012,13(4):1753-1762
[2]. Fan H, Cao Z, Jiang Y, et al. Learning Deep Face Representation [J]. EprintArxiv, 2014:11-13

[3]. Lecun Y, Boser B,Denker J S, et al. Back propagation applied to handwritten zip code recognition[J].Neural computation,1989,1(4):
[4]. Han Feng, Shan Ying, CekanderR, et al. A two-stage approach to people and vehicle detection with hog-based SVM［C］/ /Proc of Performance Metrics for Intelligent Systems Workshop. 2006: 133-140.
[5]. Kavukcuoglu K, Sermanet P, Boureau Y L, et al. Learning convolutional feature hierarchies for visual recognition［C］/ /Proc of NIPS. 2010.
[6]. Bengio Y. Learning deep architectures for AI［J］. Foundations and Trends in Machine Learning, 2009, 2(1) : 1-127.
[7]. Fan Haoqiang, Cao Zhimin, Jiang Yuning, et al. Learning deep face representation［J］. arXiv preprint arXiv: 1403. 2802, 2014.
[8]. Oliveira L, Nunes U. On integration of features and classifiers for robust vehicle detection［C］/ /Proc of the 11th International IEEE Conference on Intelligent Transportation Systems: IEEE Press, 2008: 414-419.