IOP Conf. Series: Earth and Environmental Science **153** (2018) 032052 doi:10.1088/1755-1315/153/3/032052

A CNN Vehicle Recognition Algorithm based on Reinforcement Learning Error and Error-prone Samples

Zhang Xuze, Niu Shengsuo, Huang Teng

Department of Electronic Engineering, North China Electric Power University, Baoding, China, 071003

809892878@qq.com, 11689214@qq.com, 799057162@qq.com

Abstract: Focused on the issue that in the latter part of training the recognition accuracy of convolutional neural network increased slowly and the training process was time-consuming, a reinforcement learning error training samples and error prone training samples CNN recognition algorithm was proposed. The algorithm uses the advantage that the network error rate decreases rapidly in the initial epochs of training. In the initial stage of training, reduce the training epochs and after each training, error training samples and error prone training samples are updated to the training set. Apply this algorithm to vehicle recognition. The accuracy of the experiment was 98.33%, and the training time was shortened to 30.97% of the original algorithm. Experimental results show that in the process of training learning error samples and error prone samples repeatedly improves the recognition accuracy and reduces the training time of CNN.

1. Introduction

In recent years, depth learning theory has developed rapidly, which is different from the traditional feature extraction algorithm which relies on prior knowledge. Depth neural network can construct feature description adaptively under the driving of training data, so it has higher flexibility and universality. Convolutional Neural Network (CNN), as a kind of deep learning model, it is an efficient recognition method developed in recent years. Convolutional neural networks reduce the computational complexity of the network through local area sensing, weight sharing and subsampling [1] to reduce the parameters that the network needs to learn. The training process of CNN requires a large number of labeled samples, the convergence rate is slow and the time complexity is high. In addition, the cost of sample tagging is also very high. Therefore, the literature [2] used GPU to accelerate the training and testing of CNN. Sun Yanfeng et al. [3] proposed a deep learning algorithm based on Fisher criterion to solve the problem of degraded performance when the number of iterations is reduced. Literature [4] proposed a periodic CNN network to integrate the input information to improve the recognition accuracy.

As an important part of intelligent transportation, vehicle recognition has attracted people's attention in recent years. Munroe [5] extracted the global features representing the edges in the region of interest of the image and classified the models using the K nearest neighbor algorithm. Literature [6] combined local information and global information to classify vehicle images. Pearce and Nick Pears [7] proposed a vehicle identification method based on Harris corner. Bailing Zhang [8] used Gabor transform and hierarchical gradient histogram features to train integrated cascade classifiers with rejecting options for vehicle type recognition. The above method used artificial features, lack of expressiveness, low recognition accuracy. The reference [8] achieved high recognition accuracy,

IOP Conf. Series: Earth and Environmental Science **153** (2018) 032052 doi:10.1088/1755-1315/153/3/032052

but it was calculated based on the removal of some uncertain samples.

In this paper, a method of CNN recognition for reinforcement learning error samples and error-prone samples is proposed. Taking advantage of the characteristics of rapid increase in the recognition rate at the beginning of training, the number of training samples is reduced properly at the early stage of training, and the identification samples and error-prone samples are constantly updated to the training set. The CNN recognition algorithm is applied to the fine recognition of vehicle brands and models. The experimental results show that the proposed method can effectively shorten the training time and improve the recognition accuracy.

2. Convolution neural network

The typical network structure of a convolutional neural network is shown in Figure 1, generally including input layer, several convolution layers and pool layer, full connection layer, classifier. The convolution layer preserves different local features, and the extracted features are rotation-invariant and translational invariant. The pool layer samples the features obtained from the convolution layer, which makes the extracted features more robust to the small deformation and simplifies the output of the convolution layer. The average value pool, maximum value pool, random pool and so on are generally used. The output layer classifies by using the feature vectors we have learned before, and the classifier selects the classifier according to different problems. Commonly used are Softmax, Support Vector Machine (SVM), etc.

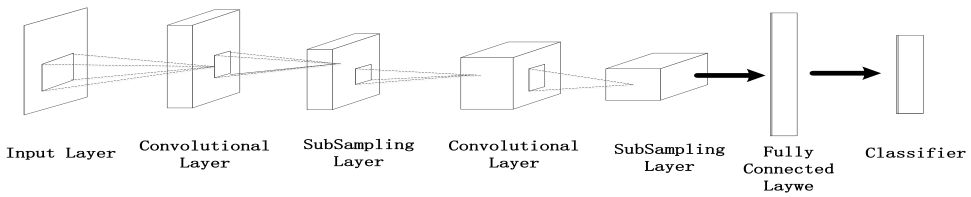

Figure 1. The structure of CNN

Traditional CNN only uses top-level information for classification, which cannot well represent the input image. The fine identification of vehicle brands and models needs to extract more expressive features. The thermal convolutional neural network [9] uses the local information of the lower layer and the global information of the upper layer to represent the input image, making the features more expressive. The structure is shown in Figure 2. The pool layer S2 connection in Fig. 2 is different from Fig. 1. A part of S2 is transformed into a feature vector, a part of it is input to the next level of convolution layer, and then the convolution calculation is carried out, and then the eigenvector is obtained after passing through the pool layer S3. The feature vectors of two parts together form the feature representation of the image.

IOP Conf. Series: Earth and Environmental Science **153** (2018) 032052 doi:10.1088/1755-1315/153/3/032052

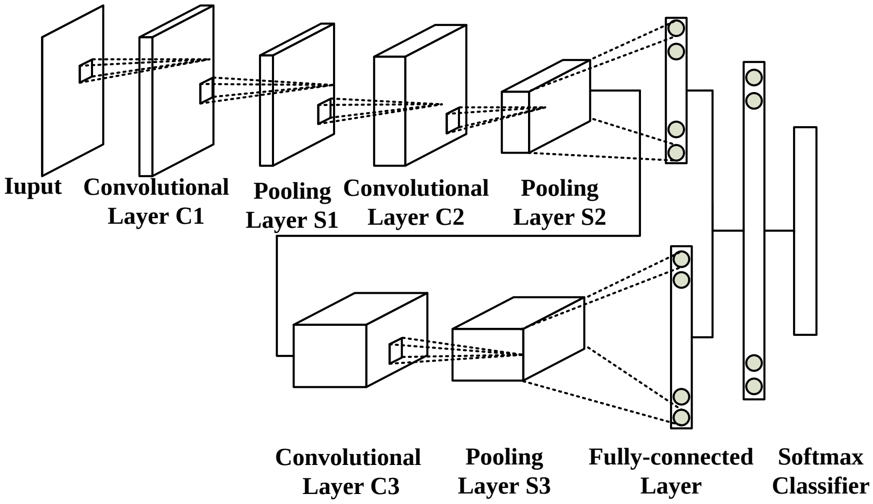

Figure 2. The structure of layer-skipping CNN

The training process of convolution neural network is a supervised learning process, which includes the forward propagation stage and the back propagation stage. In the forward propagation stage, the training samples are input into the network to calculate the actual output of the network. The error between the actual output and the ideal output is calculated in the backpropagation stage, and the parameters are adjusted by the back-propagation algorithm and the stochastic gradient descent method to optimize the network and make the network adjust to the global optimal direction.

3. CNN recognition algorithm based on reinforcement Learning error sample and Error-prone sample

In the initial training of convolutional neural network, the error rate of network recognition decreases rapidly, but with the increase of training times, the error rate decreases slowly. When the training reaches a certain number of times, the accuracy rate remains basically unchanged, even if the training times are increased, it is difficult to increase the accuracy rate.

As shown in Fig. 3, Fig. 3 shows the relationship between the number of training times and the recognition accuracy when a leapfrog CNN is trained on a vehicle identification data set using a traditional training algorithm. From Figure 3, we can see that the number of iterations is less than 20:00, and the accuracy rate increases gradually with the number of iterations rising in a straight line after 20 times. When the number of times of training reaches 60 times, with the number of times of training increasing greatly, there's only a slight change in accuracy.

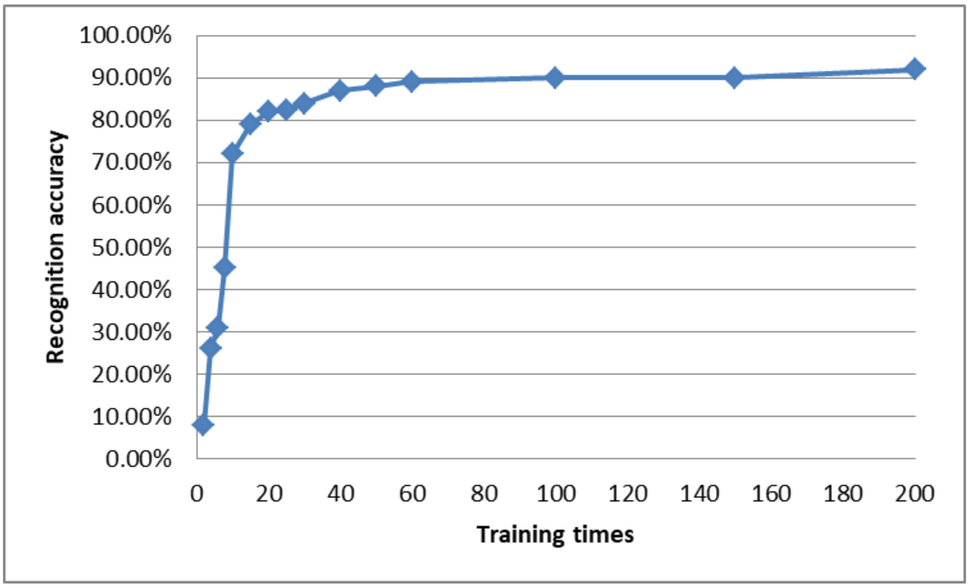

Figure 3. The relationship between the epochs of layer-skipping CNN and recognition accuracy

When people learn new knowledge, it takes different time to understand and master different knowledge because of the different degree of difficulty of knowledge. Easy knowledge can be mastered quickly, hard to understand knowledge needs to be consolidated many times, draw from one another. Inspired by the human learning process, this paper presents a CNN recognition algorithm for reinforcement learning error samples and error-prone samples, taking advantage of the rapid reduction of the initial error rate in the training of convolutional neural networks. At the beginning of the training, the training frequency of samples is reduced appropriately, and the samples that identify errors and the samples that are prone to errors are updated to the training set to strengthen the training at the same time. Further improve the accuracy of recognition, reduce the training time.

3.1Definition of error and error-prone samples

Error samples and error-prone samples are just like the knowledge difficult to master in human learning. Through learning from one another, multiple learning can play a better learning effect. First, we define the error sample and the error-prone sample. The error sample is the sample which is different from the label after the recognition of convolution neural network. The error-prone sample is the same as the label, but the network output is less than a threshold.

For example, in vehicle recognition, if a car image is a Volkswagen speed, after recognition, the network determines that it is Audi, then this sample is the wrong sample; If the vehicle is identified as mass speed, but the output value of the network is less than the set threshold, such as the threshold of 0.8, the network output is 0.7, indicating that the network recognition result of the sample cannot reach a high degree of certainty, this kind of sample is an error-prone sample.

3.2 CNN recognition algorithm for reinforcement learning error samples and error-prone samples

The CNN recognition algorithms based on reinforcement learning error samples and error-prone samples are as follows:

- The data set is divided into three parts, the test set T_1, the data set to be identified T_2 and the training set V. The training set V is used to train the convolutional neural network, the test set is T_2 used to test the effect of the initial training network, and the training set V is adjusted. The data set T_1 to be identified is a sample to be identified.
- At the beginning of the training, the image of the training set V is taken as the input of the CNN, and a CNN model is obtained for each training process.

- Test the CNN model with T_2. Find out the number of general training sessions that end at the beginning of the training and stop training.
- The wrong samples and error-prone samples are selected from the test results. The training samples are scaled and rotated and updated to the training set to get the new training set V'.
- Using a new training set, repeat (2)~(4) continue to update to V' enable the network to learn more information.
- When the recognition accuracy reaches a certain requirement, stop training. The image of the data set T_1 to be identified is input into the network trained by the previous steps, and the recognition result is obtained.

This method takes advantage of the rapid reduction of the error rate in the initial training of convolution neural network, and only carries out a small number of times of training to the network, saves the training time, and adjusts the training set. The accuracy of network recognition is improved by learning error samples and error-prone samples many times.

4. Experiment

The experiment is carried out in i52.67GHz processor and 6GB memory environment, and is completed with Matlab software.

4.1 Data sets

The vehicle brands and models used in the experiment contain 3180 images, of which 300 images are to be identified and 290 images are included in the test data set. The training set contains 2, 590 pictures, covering 15 categories of vehicles from 10 brands. Audi, BMW and Benz(2 Types), Peugeot(2 Types)Buick(2 Types), The public(2 Types)Toyota, Ford, Geely, Hyundai(2 Types)。

The face image shows all parts of the edge of the tire up to the edge of the lamp. The image is normalized to pixel, and the image is processed as a gray image.

4.2 Network structure and parameter setting

Grayscale car face picture as the input of the network, the size of the picture is. Convolution layer C1 layer and C2 layer convolution kernel size are, the number of characteristic graphs is respectively 12 and 24. The size of the C3 layer convolution kernel is 36.The pool zone of the pool layer is the pool layer. The learning rate was set to 0.1.

4.3 Experimental results and analysis

The following experiments were performed on the above data sets: (1) The training algorithm of this paper is applied to the fine identification of vehicle brands and models; (2) Histogram of Oriented Gradient (HOG) + SVM is applied to vehicle model recognition and is compared with experiment (1) In the recognition of the results were compared.

- The CNN recognition algorithm for reinforcement learning error samples and error-prone samples proposed in this paper is applied to vehicle recognition. First, set samples with output values less than 0.5 as error-prone samples. In the experiment, first with the training set of network training 20 times, set test, the accuracy rate was 90.33%, find the error and error prone sample, update the training set, and then training 40 times to set test, the accuracy rate was 99.66%, achieved high accuracy, stop training, training a total of 60 times. Finally, the trained network is used to classify test sets, and 5 images are identified for 300 images, and the accuracy of vehicle recognition is 98.33%.
- Experiments were carried out by direct training of the cline CNN network and the HOG+SVM method respectively.(1)The results of the recognition are compared, as shown in Table 1.Table 1 shows that the recognition accuracy of this algorithm is much higher than that of direct training method and artificial feature extraction algorithm.

Table 1. Experimental results of different recognition algorithm

Experimental method	Algorithm in this paper	Direct training method	HOG+SVM
Type of car	Error recognition number		
Audi	0	0	16
BMW	0	1	0
Mercedes Benz C system	0	2	10
Mercedes Benz S system	0	3	0
Peugeot 301	0	2	12
Peugeot 308S	0	0	0
Buick Weilang	0	0	3
Buick Hideo	0	1	13
Volkswagen Passat	0	1	0
Volkswagen Sagitar	0	2	15
Toyota	1	1	4
Ford	1	1	0
Auspicious	1	5	18
Modern movement	2	2	0
Hyundai	0	1	5
Total	5	22	96
accuracy rate	98.33%	92.67%	68%

Table 2 shows the comparison between the improved algorithm and the direct training method in training time and recognition accuracy. As shown in Table 2, the improved training method was trained 60 times, 300 test picture recognition errors, and the recognition accuracy of 98.33%.The training time was 3.10h for 60 times. The accuracy rate of the direct training network is 92.67%, and the time of sharing is 10.01h, and the time is shortened to 30.97%.

Table 2. Comparison between the training algorithm in this paper and the direct training algorithm

	This paper method	Direct training method
Training times / times	60	200
Training time /h	3.10	10.01
Error recognition number	5	22
Recognition accuracy	98.33%	92.67%

5 Conclusion

In this paper, the CNN recognition algorithm for reinforcement learning errors and error-prone samples is proposed and applied to the fine vehicle recognition of vehicle brands and models. The recognition accuracy of the algorithm is obviously better than that of the traditional convolution neural network and the artificial feature extraction method, ant the training time is shortened. However, due to the limitation of the type of vehicle and the number of samples used in the experiment, it has a certain impact on the experimental results. Therefore, the next step is to expand the sample types and the number of samples.

References

[1] LeCun Y, Kavukcuoglu K, Farabet C. Convolutional networks and applications in vision [C]//Circuits and Systems (ISCAS),Proceedings of 2010 IEEE International Symposium on. IEEE, 2010: 253-256.
[2] Krizhevsky A, Sutskever I, Hinton G E. Imagenet classification with deep convolutional neural networks [C]//Advances in neural information processing systems. 2012: 1097-1105.
[3] Sun Yanfeng, Qi Guanglei, Hu Yongli, and so on. A deep convolution neural network recognition algorithm based on improved Fisher criterion [J]. Journal of Beijing University of Technology, 2015, 41(6): 835-841.
[4] Liang M, Hu X. Recurrent convolutional neural network for object recognition [C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. 2015: 3367-3375.
[5] Munroe D T, Madden M G. Multi-class and single-class classification approaches to vehicle model recognition from images[J]. Proc. AICS, 2005.
[6] Pearce G, Pears N. Automatic make and model recognition from frontal images of cars [C]//Advanced Video and Signal-Based Surveillance (AVSS), 2011 8th IEEE International Conference on. IEEE, 2011: 373-378.
[7] Saravi S, Edirisinghe E A. Vehicle make and model recognition in CCTV footage [C]//Digital Signal Processing (DSP), 2013 18th International Conference on. IEEE, 2013: 1-6.
[8] Zhang B. Reliable classification of vehicle types based on cascade classifier ensembles [J]. Intelligent Transportation Systems, IEEE Transactions on, 2013, 14(1): 322-332.
[9] Dong Z, Pei M, He Y, et al. Vehicle type classification using unsupervised convolutional neural network [C]//Pattern Recognition (ICPR), 2014 22nd International Conference on. IEEE, 2014: 172-177.