Chinese License Plate Recognition Using a Convolutional Neural Network

Conference Paper · January 2009

DOI: 10.1109/PACIIA.2008.196 · Source: IEEE Xplore

Zhao Zhihong
Shijiazhuang Tiedao University

Shaopu Yang
Shijiazhuang Tiedao University

2008 IEEE Pacific-Asia Workshop on Computational Intelligence and Industrial Application

Chinese License Plate Recognition Using a Convolutional Neural Network

Zhihong Zhao, Shaopu Yang, Xinna Ma
Shijiazhuang Railway Institute
Shijiazhuang, Hebei Province,050043, china
hb_zhaozhihong@126.com; yangsp@sjzri.edu.cn

Abstract

In this paper, a new method was introduced in the Chinese license plate recognition. We propose a convolutional neural network architecture designed to recognize license plate directly from pixel images with no preprocessing. We present the image transformation applied on the original license plate to increase the training database. We also provide experimental results to demonstrate the robustness of our approach and the recognition rate on the license plate and non-license plate testing set.

1. Introduction

License plate recognition is a very important research topic, due to its wide range of applications in intelligent traffic applications, such as the payment of parking fee, highway toll fee, traffic data collection, etc. Because of ambient lighting conditions, image perspective distortion, it is difficult to efficiently recognize license plate in complex conditions.

Numerous approaches for license plate recognition have been proposed in the last decade. Most methods are based on features of the license plates. Features commonly employed have been derived from the license plate format and the alphanumeric characters constituting license plate numbers. The features regarding license plate format include edge[1], symmetry[2] , color[3,4], texture of grayness[5] .

Learning-based methods have widely used in license plate recognition recent years. These methods need positive samples and negative samples in the training procedure. Usually the positive samples are obtained through labeling the license plate regions from the vehicle images. The negative samples are randomly extracted from different images which do not contain license plate. K.K.Kim [6] used neural network to recognize license plate in traffic image. Huaifeng Zhang[7] used AdaBoost learning algorithm to build up the weak classifiers based on local Haar-like features.

Most of these techniques rely on image preprocessing before the training and classification stages. In this paper, we propose a new method based on convolutional neural network that have been introduced by Le Cun et al. and successfully applied to handwritten character recognition[8]. Convolutional neural networks avoiding local preprocessing to ensure that no unrealistic assumption is made about the data and that the richness of the original signal is preserved.

Convolutional Neural Networks have been used successfully in a number of vision applications such as handwritten character recognition[8,9], generic object recognition[10] , face recognition[11], pedestrian detection[12], etc. The use of receptive fields, shared weights and spatial subsampling in convolutional network provides much higher degrees of invariance to translation, rotation and scale, while strongly reducing the number of adjustable weights to learn.

The remainder of this paper is organized as follows. The convolutional Neural Networks are introduced in section 2. The details architecture of the convolutional Neural Network used in this paper is described in section 3. The training set and testing set is described in section 4. We present experimental results in section 5. Finally we draw conclusions in section 6.

2. Convolutional Neural Networks

Convolutional neural networks[9] are “end-to-end” trainable system that can operate on raw pixel images and learn low-level features and high-level representation in an integrated fashion. Convolutional neural networks are advantageous because they can easily learn the types of shift-invariant local features that are relevant to image recognition; and more importantly, they can be replicated over large images (swept over every location) at a fraction of the cost of replicating more traditional classifiers[9]. This is a considerable advantage for building real-time systems.

Convolutional neural networks are an attempt to solve the dilemma between small networks that cannot learn the training set, and large networks that seem over-parameterized. Three architectural ideas used in Convolutional Neural Networks to ensure some degree of shift, scale, and distortion invariance are local receptive fields, shared weights, and spatial or temporal sub-sampling. Local receptive fields refer to each unit in a layer receives inputs from a set of units located in a small

978-0-7695-3490-9/08 $25.00 © 2008 IEEE
DOI 10.1109/PACIIA.2008.196

neighborhood in the previous layer. With local receptive fields, neurons can extract elementary visual features such as oriented edges, corners, etc. Shared weights allow to significantly reduce the number of free parameters, which in turn improves the generalization ability. Sub-sampling reduce the resolution of the image and reducing the sensitivity of the output to shifts and distortions.

Typically convolutional neural network include two different kinds of layers: convolutional layer, sub-sampling layer. At a convolution layer, the previous layer's feature maps are convolved with learnable kernels and put through the activation function to form the output feature map. Each output map may combine convolutions with multiple input maps. A sub-sampling layer produces down-sampled versions of the input maps. If there are M input maps, then there will be exactly M output maps, although the output maps will be smaller.

3. The Proposed Convolutional Neural Network

We designed a Convolutional Neural Network for Chinese License Plate recognition. The architecture as shown in figure 1 comprises 7 layers, not counting the input, all of which contain trainable weights. The actual license plate size is 20×60 (because the aspect ratio of Chinese license plate is 1:3), and padding it to 24×64 to extract the feature in the border of the license plate.

In the following, convolutional layers are labeled Cx, sub-sampling layers are labeled Sx, and fully-connected layers are labeled Fx, where x is the layer index.

Layer C1 is a convolutinal layer with 4 feature maps. Each unit in each feature map is connected to a 5×5 neighborhood in the input. The size of the feature maps is 20×60.

Layer S2 is a sub-sampling layer with 4 feature maps of size 10×30. Each unit in each feature map is connected to a 2×2 neighborhood in the corresponding feature map in C1. The two inputs to a unit in S2 are added, then multiplied by a trainable coefficient, and added to a trainable bias. The result is passed through a sigmoid function.

Layer C3 is a convolutional layer with 11 feature maps. Each unit in each feature map is connected to several 5×5 neighborhoods at identical locations in a subset of S2's feature maps. Table 1 shows the set of feature maps combined by each C3 feature map. Each Column indicates which feature map in S2 are combined by the units in a particular feature map of C3. The size of the feature maps is 6×26.

Table 1 Connections between layer S2 and layer C3

	0	1	2	3	4	5	6	7	8	9	10
0	X			X	X		X		X	X	X
1	X	X				X	X	X		X	X
2		X	X		X		X	X	X		X
3			X	X		X		X	X	X	X

Layer S4 is a sub-sampling layer with 11 feature maps of size 3×13. Each unit in each feature map is connected to a 2×2 neighborhood in the corresponding feature map in layer C3.

Layer C5 is a convolutional layer with 120 feature maps. Each unit is connected to a 3×13 neighborhood on all 11 of S4's feature maps. We use the size of receptive field 3×13 to make the size of C5's feature map 1×1. Thus the C5 layer and the following F6 layer and output layer can easily make the classifier.

Layer F6, contains 84 units and is fully connected to C5. This layer is like the hidden layer of the BP neural network.

Units in layers from C1 to F6 compute a dot product between their input vector and their weight vector, to which a bias is added. This weighted sum is then passed through a scaled hyperbolic tangent sigmoid squashing function.

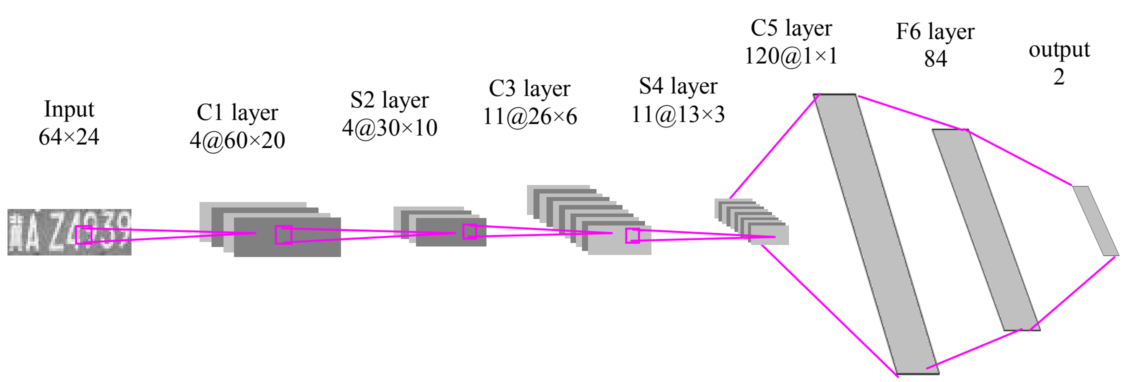

Fig. 1. Architecture of the Convolutional Neural Network

Finally, the output layer is composed of Euclidean Radial Basis Function units (RBF), one for license plate, the other for non-license plate, with 84 inputs each.

A sample of the feature maps of each convolutional neural network layer are shown in figure 2. The left is the input image, from the left to right are feature maps of C1 layer, feature maps of S1 layer, feature maps of C2 layer and feature maps of S2 layer.

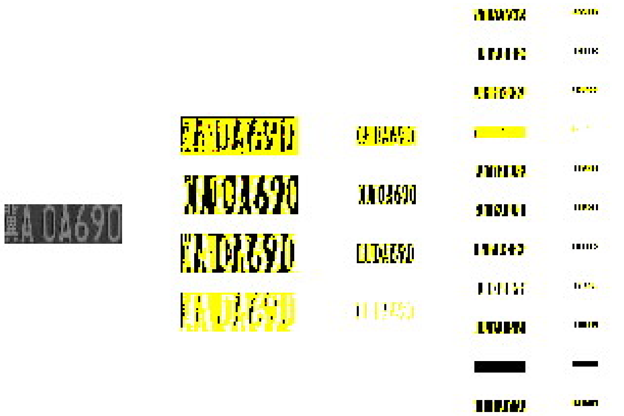

Fig. 2. A Sample of Feature Maps of the Proposed Convolutional Neural Network

4. Dataset

Positive samples and negative samples are needed in the training procedure. We built positive training set by manually cropping 400 variable license plate areas in a collection of images obtained from real traffic scenes. Most of the learn-based approaches for Chinese license plate recognition in the literature[7] use an input window of dimension around 16×48, reported as being the smallest window one can use without loosing critical information. We have chosen the window for the central part of the license plate but the size 20×60, considering the process of plate character recognition. No intensity normalization is applied on the cropped license plates, such as histogram equalization and overall brightness correction. As mentioned earlier, convolutional neural network is robust in scale and position, and we aim at enforcing this robustness by providing samples that are not normalized.

In order to create more examples and to enhance the capabilities of invariance to intensity, the reverse color transformation and some contrast reduction transformation are applied to all the original training examples. The final positive training set including 2400 license plates. Some samples are shown in Fig. 3. The left are the original images, from the left to right are the reverse color image, and the different contrast reduction of the original image.

Fig.3. Some positive samples of the training set

Negative samples are produced by random cropped from the traffic images that are not license plate. Some negative samples are shown in Fig. 4.

Fig.4. Some negative samples of the training set

5. Experimental Results

In this section, we aim at presenting the performance of the proposed convolutional neural network. All experiments were trained on 2400 license plates and 4000 non-license plates. There was no overlap between the training and testing sets.

5.1. Performance of the Network

To evaluate the proposed license plate classifier, 4400 test images with 400 license plates and 4000 non-license plates were used in the testing procedure. The experimental results are listed in Table 2. After training 500 epochs, the license plate recognition rate achieves 98.25% and the non-license plat recognition rate achieved 100%. This result shows that the non-license plate recognition rate is better than that of the license plate. This perhaps because that the training samples of the non-license plates are more than that of the license plates.

Table 2 Recognition rate after training 500 epochs

	Training set	License plate	Non-license plate
recognition rate	100%	98.25%	100%

5.2 Recognition Rate Curve

Over-training is a serious problem when used neural network. When over-training occurs, the training error keeps decreasing over time, but the test error goes through a minimum and starts increasing after a certain number of iterations. Many authors have reported observing the common phenomenon of over-training when training neural networks on various tasks. In [9], LeCun declared over-training was not observed in the convolutional neural network architecture LeNet-5. But we observed a phenomenon of that the recognition rate was not steady. The learning curve is shown in figure 5. The license plate recognition rate is vibrating during the training procedure. But the vibrating amplitude was decreased during the training.

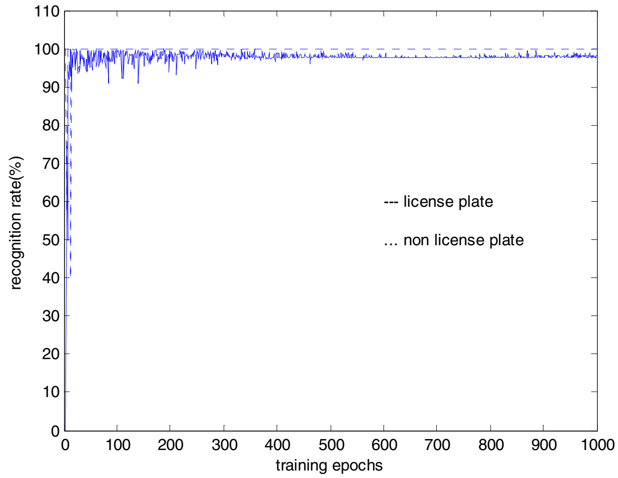

Fig.5. the recognition rate of license plate and non-license plate

5.3 Analysis of the Errors

The 7 misclassified license plate images out of 400 test samples are shown in figure 6. From the 7 recognition errors, three are due to the hard recognition of the plate image, one are due to the skew of the plate image. There are also three samples can easily recognize by humans. The reason maybe the training data is not large enough. This shows that further improvements are to be expected with more training data.

Fig.6. misclassified by the convolutional neural network

6. Conclusion

Our experiments have shown that using convolutional neural networks for Chinese license plate recognition is a very promising approach. The training set was increased by contrast reduction and reverse color transformation. The proposed approach need not the intensity normalization and is robustness to lighting conditions. Tested on the database, the recognition rate of the 400 license plates was 98.25%, and the recognition rate of the 4000 non-license plates was 100%. We plan to use this method in character recognition and in vehicle recognition.

Acknowledgement

The authors are grateful to the support by the National Science Fund for Distinguished Young Scholars of China (50625518).

References

[1] Vladimir Shapiro, Dimo Dimov, Stefan Bonchev (2004). “Adaptive license plate image extraction”. Proceedings of the 5th international conference on Computer systems and technologies. pp. 1 - 7

[2] D. S. Kim and S. I. Chien (2001). “Automatic car license plate extraction using modified generalized symmetry transform and image warping”. Proc. IEEE Int. Symp. Industrial Electronics, vol. 3, pp. 2022–2027.

[3] S. K. Kim, D. W. Kim, and H. J. Kim (1996). “A recognition of vehicle license plate using a genetic algorithm based segmentation”, Proc. Int. Conf. Image Processing, vol. 2, pp. 661–664.

[4] Zhu, W. G., Hou, G. J., & Jia, X (2002). “A study of locating vehicle license plate based on color feature and mathematical morphology”. 6th international conference on signal process.Vol. 1, pp. 748–751.

[5] M. H. T. Brugge, J. H. Stevens, J. A. G. Nijhuis, and L. Spaanenburg (1998). “License plate recognition using DTCNNs”. in Proc. 5th IEEE Int. Workshop on Cellular Neural Networks and Their Applications, pp. 212–217.

[6] Kim, K.K., Kim, K.I., Kim, J.B., and Kim, H.J (2000). “Learning-based approach for license plate recognition”. IEEE Int. Workshop on Neural Netw. Signal Process. vol. 2, pp. 614–623

[7] Huaifeng Zhang, Wenjing Jia, Xiangjian He, and Qiang Wu (2006). “Learning-Based License Plate Detection Using Global and Local Features”. Proceedings of the 18th International Conference on Pattern Recognition.

[8] Y. Le Cun, B. Boser, J.S. Denker, D. Henderson, R. Howard, W. Hubbard, and L. Jackel (1990). “Handwritten digit recognition with a backpropagation neural network”. Advances in Neural Information Processing Systems 2, pp.396–404.

[9] LeCun Y, Bottou L., Bengio Y. and Haffner P (1998). “Gradient-based learning applied to document recognition”. Proc. IEEE. 2278-2324

[10] Y. LeCun, F.J. Huang, and L. Bottou (2004). “Learning methods for generic object recognition with invariance to pose and lighting”. Proceedings of CVPR.

[11] Steve Lawrence, C. Lee Giles, Ah Chung Tsoi, Andrew D. Back (1997). “Face Recognition: A Convolutional Neural Network Approach”. IEEE Trans. on Neural Networks, vol. 8(1), pp.98–113

[12] Szarvas Mate, Yoshizawa Akira, Yamamoto Munetaka, Ogata Juncz (2005). “Pedestrian detection with convolutional neural networks”. IEEE Intelligent Vehicles Symposium Proceedings, pp.224-229