License Plate Location and Recognition on Neural Network

Zhilong He¹, Zhongjun Xiao¹,²,* and Zhiguo Yan¹

¹School of Electrical Engineering and Automation, Qilu University of Technology (Shandong Academy of Sciences), Jinan 250353, China.
²Key Laboratory of Pulp and Paper Science and Technology of Ministry of Education, Qilu University of Technology (Shandong Academy of Sciences), Jinan, 250353, Shandong, China.

*Email: jinanxiaoxiao2012@163.com.

Abstract. This paper takes autonomous driving technology as the research background, and aims to improve the efficiency of the traffic system in real life, and proposes a neural network recognition method for license plates. The method realizes accurate location and recognition of LP (license plate) by constructing two neural recognition networks. For the LP location, a 11-layer convolution recognition network (CNN) is constructed. The basic step is to roughly locate the LP after preprocessing the image, and use the neural network to accurately locate the LP. For LP recognition, the 10-layer convolutional neural network constructed by inputting the segmented LP is used to realize character recognition, and the output result is obtained. The method is efficient and accurate, and can be applied to handheld devices in transportation systems.

1. Introduction

With the rapid development of the social economy, the number of cars is increasing year by year, and the problems of public transportation are becoming more and more prominent. In order to solve complex traffic management problems, the development of Intelligent Transport System (ITS) [1, 2] has been valued by countries all over the world, and LP positioning and recognition as an important link in the intelligent transportation system.

The methods of LP location are based on edge [3], color [4], texture [5] and deep learning [6]. Such as, the LP edge density and other information are used to locate the LP, but for images with complex or unclear edges, the recognition accuracy is poor. The color-based recognition method locates through the inherent color features of the LP, but this method is extremely susceptible to light, making the positioning inaccurate. The method of LP texture feature is to locate the LP by sliding the frame to locate the difference in texture in the image. Deep learning is one of the key research areas in recent years. This method has good generalization ability, and the feature selection is done by the computer, which greatly improves the positioning efficiency and accuracy. However, the model established by this method often has a large parameter scale, which limits the actual deployment.

The methods of LP character recognition include template matching [7], feature matching [8, 9] and neural network [10]. Based on the template matching method, the divided characters are compared with a preset character template, and the character with the highest matching degree is output. The template matching method is simple to use, but it is difficult to identify characters with severe noise and distortion. The method of character feature matching is to extract the character structure features
of the LP for character recognition, but this method is more sensitive to external interference and requires more calculation. Feature selection of feature value matching includes (directional gradient histogram, HOG) feature, (scale invariant feature transformation, SIFT) feature, and these features can be used (support vector machine, SVM) [11] for character recognition, which is a commonly used feature matching recognition method. The character recognition method of neural network is a popular method in recent years. This method does not require complex feature extraction on the image, and can directly use the original image as the input image. Recognize images by extracting image features layer by layer and extracting high-dimensional features of images. Compared with traditional methods, this method has higher recognition accuracy, but the established model is often very large and the training process is longer.

Based on the above large model and long training process, in this paper a method of LP recognition based on CNN is proposed. The specific steps are: first, the input image is preprocessed, and then the LP is located by mathematical morphology. Second, the candidate LPs are sent to the CNN model to accurately locate the LPs, and a single LP image is projected to segment the characters. Finally, the CNN network performs character recognition.

2. Our approach

2.1. LP location method

First, the image is preprocessed. The steps of image preprocessing are histogram equalization, filtering and edge detection. Histogram equalization is a commonly used image transformation method. Its essence is to calculate the probability of the gray value of the pixel value and the cumulative probability of each gray level by counting the number of pixel values of the image. The result of image equalization is obtained by the probability of gray mapping, as shown in Equations (1-2).

\[ T(r) = L_i^T p(t) dt \]  \hspace{1cm} (1)

\[ T(r) = \sum_{i=0}^{n} \frac{n_i}{n} \]  \hspace{1cm} (2)

Where \( r \) is the pixel value of the original image, \( T \) is the mapped pixel value, \( L \) is the gray level, is the probability distribution of pixels, \( n \) is the sum of pixels, and \( N_i \) is the pixel value corresponding to the number of pixels. Equations (1) and (2) are continuous and discrete, respectively.

The mean value filtering can filter out the high-frequency signal of the image and retain the low-frequency signal of the image, thereby eliminating the noise of the image and realizing the smoothness of the image, as shown in Equation (3).

\[ f(x, y) = \frac{1}{mn} \sum_{(x,y) \in S_{oo}} g(s,t) \]  \hspace{1cm} (3)

Where \( f(x, y) \) is the filtered image, \( g(s, t) \) is the original image, and \( mn \) is the size of the filter (m, n).

Edge detection can detect the edge of the image, which not only reduces the amount of image data, but also retains the key information of the image. This paper uses the Sobel operator, which has high edge detection efficiency and can highlight the potential position of LP.

Second, the color space of the original image is converted from RGB to HSV, and the thresholds of H, S, and V are set according to the color of the license plate, and finally an operation is performed with the edge-processed image to obtain an image that highlights the edge of the license plate position. Morphological processing is the final step of image preprocessing. Through multiple opening and closing operations, a binary image of the license plate is obtained, which prepares for the next positioning of the license plate. Based on the prior knowledge of the license plate (size, length to width ratio), determine whether it is a license plate area, and correct the tilt of the obtained license plate area. Color conversion Equations (4-6) as follows:
Where r, g, b represent the values of RGB components. Cmax and Cmin are the maximum and minimum values of the three components of r, g and b, and $\Delta$ is the difference between max and min.

Finally, the size of the obtained candidate license plate is normalized, adjusted to the specified size, and sent to the CNN network to screen the candidate license plate, so as to obtain the accurately positioned license plate and extract the image of the license plate position.

In this paper, ELU is the activation function of this experiment, as shown in Equations (7).

$$f(x) = \begin{cases} 
  x, & \text{if } x \geq 0 \\
  \alpha (e^x - 1), & \text{if } x < 0 
\end{cases}$$

Where $x > 0$, the output is the same as ReLU; $x < 0$, the output is not zero, and $\alpha$ is a non-zero constant.

In order to retain more texture information of LPs characters, in this article the pooling method is maximum pooling. Maximum pooling can reduce the dimension of features formed by convolution in a way to obtain the maximum value of the region. The specific operation is the size of the convolution kernel of the input image is divided into different regions, and each region to a maximum value as an output.

A CNN LP location network (LPL - CNN) was built by the analysis of LP images, LPL - CNN network structure is shown in Table 1.

### Table 1. LPL – CNN network structure

| Layer | Type   | Maps & size             | Kernel/stride |
|-------|--------|-------------------------|---------------|
| 0     | Input  | 3 m. of 36×136 s        |               |
| 1     | Conv   | 32 m. of 36×136 s       | 3×3/1         |
| 2     | Max-Pool | 32 m. of 18×68 s       | 2×2/2         |
| 3     | Conv   | 64 m. of 18×68 s        | 3×3/1         |
| 4     | Max-Pool | 64 m. of 9×34 s        | 2×2/2         |
| 5     | Conv   | 128 m. of 9×34 s        | 3×3/1         |
| 6     | Conv   | 128 m. of 9×34 s        | 3×3/1         |
| 7     | Max-Pool | 128 m. of 5×17 s       | 2×2/2         |
| 8     | FC     | 1×1024s                 |               |
| 9     | FC     | 1×1024s                 |               |
| 10    | FC     | 1×2s                    |               |
| 11    | Soft-Max |                     |               |
Three convolutions and pooling are used to build a LPL-CNN network. The size of the convolution kernel is $3 \times 3$. The size of the core of the pooling layer is $2 \times 2$. As the network deepens, the number of convolution features increases. Multiple convolutional feature extraction is performed on the $136 \times 3$ size image. Finally, a fully connected layer is used to connect feature subsets, reduce the feature size and output the recognition result. After many attempts, the learning rate was set to 0.001, and the dropout was used to randomly drop neurons, which improved the overfitting during the training process, thereby improving the training effect.

2.2. LP recognition method

1) Character segmentation

Before character recognition is performed on LP, single character segmentation needs to be performed on the located and extracted LP. The segmentation method used in this paper is the vertical projection method. The basic process is as follows:

First, the LP obtained by binarization is extracted to reduce the storage space and image noise. The image is projected vertically on the Y axis to obtain the pixel value of the image.

Second, the program scans the entire projected image to determine the size and continuity of pixel values, and crops the irrelevant LP edges to obtain a separate image of the license plate. The reasonable threshold is set by the program to determine whether the character size conforms based on the prior knowledge of the character size, and the matching character is cut to obtain the final character recognition of the character cut result.

Finally, the size of the single character after segmentation is normalized to facilitate subsequent CNN to accurately identify the character.

2) Character recognition

Due to the large number of characters to be recognized, a character recognition network (CR-CNN) was built. We increase the convolutional layer of the network to obtain higher-dimensional image feature maps, increase the number of nodes in the fully connected layer, and increase the number of feature maps to improve the recognition ability of the network. The network structure is shown in Table 2.

| Layer | Type    | Maps & size   | Kernel/stride |
|-------|---------|---------------|---------------|
| 0     | Input   | 1 m. of 20×20 s |               |
| 1     | Conv    | 32 m. of 20×20 s | 3×3/1         |
| 2     | Max-Pool| 32 m. of 10×10 s | 2×2/2         |
| 3     | Conv    | 64 m. of 10×10 s | 3×3/1         |
| 4     | Max-Pool| 64 m. of 5×5 s  | 2×2/2         |
| 5     | Conv    | 128 m. of 5×5 s | 3×3/1         |
| 6     | Max-Pool| 128 m. of 3×3 s | 3×3/1         |
| 7     | FC      | 1×256s         | 2×2/2         |
| 8     | FC      | 1×256s         |               |
| 9     | FC      | 1×256s         |               |
| 10    | Soft-Max|               |               |

3. Experiment and analysis

3.1. Datasets and configurations

The vehicle LP data set comes from the network collection and pictures taken by the handheld camera. The data is divided into a test set and a training set according to 1: 4, and the data is fed into the built network for training. Part of the data set image is as shown in Figure 1.
In Figure 1, the first five lines are the LP data set, which includes the LP itself and non-plate license images, which are used to accurately locate the LP; the last line is the LP character recognition data set, which includes three categories of different forms of letters, numbers and Chinese characters.

The configuration used in this experiment is Intel (R) Core (TM) i5-7300HQ CPU @ 2.50GHz processor, 20GB memory, 1T hard disk capacity, 4GB GTX 1050 GPU graphics card, and the operating system is ubuntu18.04. In order to speed up the training process, configure NVIDIA CUDA and CUDNN environment, use GPU acceleration training. LPL-CNN and CR-CNN are trained separately, then they can identify whether the extracted area is LP and a single character after segmentation, respectively.

![Examples of datasets.](image1.png)

**Figure 1.** Examples of datasets.

### 3.2. Experimental results and analysis

#### Table 3. Model performance

| Model     | Precision(%) | Recall(%) | Detection times(s) |
|-----------|--------------|-----------|--------------------|
| VGG16     | 96.7         | 95.7      | 3.4                |
| LPL-CNN   | 97.5         | 96.8      | 1.58               |
| CR-CNN    | 98.8         | 97.6      | 0.58               |

As shown in Table 3, the LP detection and character recognition models have good detection accuracy and recall rate, the detection time is also very short, and can be deployed in embedded devices. The experimental results show that the LP recognition and character recognition network constructed in this paper can maintain high accuracy under the consideration of speed.

![Examples of car license plate detection and recognition](image2.png)

**Figure 2.** Examples of car license plate detection and recognition
Figure 2 shows an example of using this method to detect and recognize LP images. The model carries out targeted training on the numbers and characters of the LP, which improves the recognition accuracy.

4. Conclusion
In this paper, we propose a CNN neural network LP detection and recognition method. By constructing two different CNN networks, the tasks of LP detection and recognition are realized separately. For LP detection, the LP is first coarsely located by threshold segmentation, and the LP is accurately located and extracted through the CNN network; CNN network training, recognize the input segmented characters, and get a better recognition effect. The experimental results show that the method has a higher recognition accuracy for the LP in the actual scene, which verifies the effectiveness and accuracy of the algorithm.

Acknowledgments
This research was supported by grants from the National Natural Science Foundation of China (No. 61877062).

References
[1] Sładkowski, Aleksander, and Wiesław Pamuła, eds. Intelligent transportation systems-problems and perspectives. Vol. 303. Springer International Publishing, 2016.
[2] Bertozzi, M., Broggi, A. and Fascioli, A., 2000. Vision-based intelligent vehicles: State of the art and perspectives. Robotics and Autonomous systems, 32(1), pp.1-16.
[3] Tan, J.L., Abu-Bakar, S.A. and Mokji, M.M., 2013, September. LP localization based on edge-geometrical features using morphological approach. In 2013 IEEE International Conference on Image Processing , pp. 4549-4553.
[4] Zunino, R. and Rovetta, S., 2000. Vector quantization for license-plate location and image coding. IEEE Transactions on Industrial Electronics, 47(1), pp.159-167.
[5] Yu, S., Li, B., Zhang, Q., Liu, C. and Meng, M.Q.H., 2015. A novel LP location method based on wavelet transform and EMD analysis. Pattern Recognition, 48(1), pp.114-125.
[6] Yang, Y., Li, D. and Duan, Z., 2017. Chinese vehicle LP recognition using kernel-based extreme learning machine with deep convolutional features. IET Intelligent Transport Systems, 12(3), pp.213-219.
[7] Farhat, A., Al-Zuwqari, A., Al-Qahtani, A., Hommos, O., Bensaali, F., Amira, A. and Zhai, X., 2016, March. OCR based feature extraction and template matching algorithms for Qatari number plate. In 2016 International Conference on Industrial Informatics and Computer Systems (CIICS), pp. 1-5.
[8] Dalal, N. and Triggs, B., 2005, June. Histograms of oriented gradients for human detection. In 2005 IEEE computer society conference on computer vision and pattern recognition (CVPR'05) Vol. 1, pp. 886-893.
[9] Lim, H.W. and Tay, Y.H., 2010, November. Detection of LP characters in natural scene with MSER and SIFT umigram classifier. In 2010 IEEE Conference on Sustainable Utilization and Development in Engineering and Technology, pp. 95-98.
[10] Liu, P., Li, G. and Tu, D., 2015, December. Low-quality LP character recognition based on CNN. In 2015 8th International Symposium on Computational Intelligence and Design (ISCID) Vol. 2, pp. 53-58.
[11] Kim, K.K., Kim, K.I., Kim, J.B. and Kim, H.J., 2000, December. Learning-based approach for LP recognition. In Neural Networks for Signal Processing X. Proceedings of the 2000 IEEE Signal Processing Society Workshop (Cat. No. 00TH8501) Vol. 2, pp. 614-623.