The Network Design of License Plate Recognition Based on the Convolutional Neural Network

Xingzhen Tao[1](✉), Fahui Gu[2], and Shumin Xie[1]

[1] College of Information Engineering, Jiangxi College of Applied Technology, Ganzhou, China
348627805@qq.com
[2] Office of Academic Studies, Guangdong Vocational and Technical College, Guangzhou 341000, Jiangxi, China

Abstract. Due to the generation of data sets and the rapid improvement of GPU computing performance, in-depth learning has undergone qualitative changes and development in the past decade. Various excellent convolutional neural network models have been implemented and verified, which accelerates the application of the convolutional neural network in various fields. A license plate recognition based on the convolutional neural network is pro-posed for the application of large underground parking lots. An end-to-end identification network framework without segmentation characters is designed. At the same time, sequence information is added to the convolutional neural network for improving the license plate recognition rate. Compared with the existing step-by-step license plate detection and recognition method, the joint solution of a single network can avoid the error accumulation in the intermediate process. At the same time, it can improve the accuracy rate, save the recognition time, accelerate the vehicle entering and leaving time and avoid traffic congestion.

Keywords: Convolutional neural network · License plate recognition · End-to-end identification network

1 Introduction

The phenomenon of "difficult parking and disorderly parking" on the ground is getting worse and worse, which has seriously affected the urban environment and traffic order. It is undoubtedly the most effective way to solve this problem by vigorously promoting the construction of urban underground parking lots through digging tapping the potential, making a renovation and starting a new construction. In the traditional underground parking lot, there are vehicle railings at the entrance and exit. When entering the parking lot, vehicles need to be intercepted manually for registration or card taking registration. When leaving the parking lot, the vehicles can be removed only after completing the manual payment, which is inefficient. In recent years, the intelligent recognition system is widely used in underground parking, but the parking lot gate is usually installed at the underground exit, which has the problems of dim

K. Li et al. (Eds.): ISICA 2019, CCIS 1205, pp. 749–758, 2020.
https://doi.org/10.1007/978-981-15-5577-0_60

light, large slope, narrow access road and so on. Therefore, the correct recognition rate and time of vehicle license plates are very important.

2 Related Work

The intelligent license plate recognition system is an indispensable part of intelligent underground parking lots [1], which is of great significance to realize automatic management of parking lots. The accurate recognition rate and recognition time of license plate are important indicators to consider. The main problems of license plate recognition technology include license plate location and character recognition [2]. The steps of a common license plate recognition method are shown in Fig. 1.

Fig. 1. The steps of license plate recognition.

2.1 License Plate Location

For license plate location, the currently used techniques are divided into two categories [3]. The first category is based on feature extraction methods and can be divided into gray feature based methods and color feature based methods. The second category is based on machine learning methods. It can be divided into a method based on pattern recognition and a method based on deep learning. The document [4] uses the unique color of the license plate to locate the license plate. However, when the light is weak, the resolution is low and the license plate picture is unclear, so the recognition rate of the license plate will be affected. The gray scale feature of license plate area contains enough information for license plate location, but the license plate location method based on gray scale feature requires a lot of calculation time. The method of license plate location based on edge and color assistance proposed by Abolghasemi et al. increases the comparison of similar regions of license plates, thus avoiding erroneous location of license plate regions, but this method is especially affected by illumination.

In recent years, due to the development of the big data and computer capability, and the rapid development of deep learning [5–7], license plate detection based on depth learning has been applied in intelligent license plate recognition system and achieved good results. Mainstream target recognition frameworks include Faster R-CNN [8, 9]. The document designed a license plate detector based on YOLO [10] and achieved good recognition accuracy. The document [11] reduces the difficulty of image recognition by preprocessing the image, and then locates the license plate through CNN network. These methods regard the license plate as an approximate rectangle. Although it has high accuracy for license plate recognition in complex environment, there is a large amount of redundancy of background information for the recognition results of tilted license plates, which will cause great interference for subsequent character cutting and recognition.

2.2 Character Segmentation

The domestic standard license plate contains 7 characters, the first character is Chinese characters, the second is uppercase English characters, and the last 5 characters are a mixture of numbers and letters. It is a key problem to How to accurately and effectively identify the 7 characters segmented by vertical projection is a key problem. Currently, the commonly used license plate character segmentation technologies include projection-based license plate character segmentation method and stroke-based license plate character segmentation method [12]. The traditional template matching method has a simple process and is greatly disturbed by additional factors, such as uneven illumination or shape change of the obtained image, so its recognition rate and robustness are not high. The projection-based license plate character segmentation method is simple and effective. Once the license plate area is determined, the characters can be segmented according to the length and height of the characters; however, such methods have poor adaptability, low resolution, and very poor detection effect for license plate images with poor light. The license plate character segmentation method based on stroke construction has a good adaptability to noise, dust, angle and other conditions [13]. However, this kind of algorithm is very complex, and its running efficiency is low, so it cannot reach the level of real-time application. The license plate character detection is a key part of license plate recognition system. The main task is to separate the characters in the license plate one by one for subsequent license plate character recognition. The effect of license plate character detection directly determines the effect of license plate character recognition. However, at the entrance and exit of the underground parking lot, it is a difficult problem to accurately segment each character in the license plate picture because the license plate is affected by factors such as illumination, angle change of the license plate, dust cover on the license plate, etc.

2.3 The License Plate Recognition

After the license plate characters are segmented, the license plate characters need to be identified. The method based on template matching is widely used in the license plate character recognition technology. However, this method is susceptible to factors such as the segmentation results of license plate characters, the noise of license plate region, and whether the license plate region is clear or not. The support vector machine (SVM) classifier is also used for the license plate character recognition [14]. The SVM-based license plate character recognition method needs to specify the extracted features, and the selection of different features will directly affect the effect of the license plate character recognition. In addition, there are also methods based on BP neural network [15]. The license plate character recognition method based on BP neural network has high learning ability, but it has the defects of complex structure and slow convergence. In recent years, methods based on depth learning have also been applied to the license plate character recognition, but they are all based on license plate character segmentation, which requires relatively long recognition time. The accuracy of the license plate character recognition depends on the effect of license plate character segmentation. For all the license plate recognition methods which are based on license plate character segmentation, the vast majority of experiments are carried out in an ideal environment.

When applied to underground parking lots, the accuracy of the license plate recognition will be greatly reduced due to weak illumination, slope inclination and other conditions.

2.4 The End-to-End Identification

With the development of the in-depth neural network and the improvement of data sets, the direct whole recognition method of license plate without character segmentation has been continuously proposed. Li et al. considered the license plate recognition as a sequence labeling problem and the convolutional neural network (CNNs) is used to extract a series of feature vectors from the license plate boundary frame in the way of sliding window [16]. Then RNN with CTC loss is used to label sequence data, and finally the recognition result without dividing characters is obtained. The recognition accuracy of the model is high, but it is time consuming.

3 System Network Model

Previous research works on the license plate detection and recognition usually divides the detection and recognition of license plates into two sub-tasks, which are solved by different methods. Before the completion of identification, the license plate should be detected firstly so as to find the position of the license plate on the map, and latter to divide the license plate character [17]. actually, the two sub-tasks of the license plate detection and recognition are highly correlated in nature [18]. The accurate license plate detection helps to improve the recognition accuracy, and accurate recognition results can be used to eliminate false alarms during the detection process. Therefore, this paper proposes a unified neural network framework, which unifies the license plate detection and recognition tasks into one framework. Through a deep neural network, the two sub-tasks of license plate location and recognition can be completed simultaneously, which will improve the license plate recognition efficiency. The specific network structure is shown in the Fig. 2.

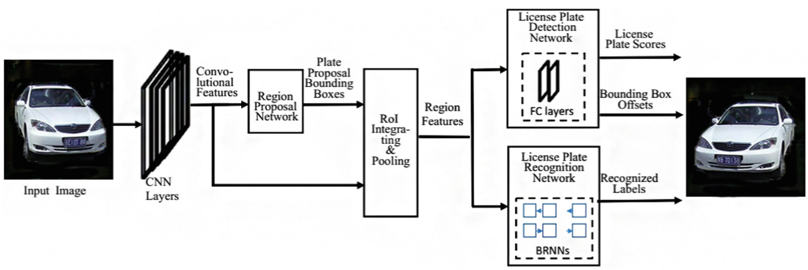

Fig. 2. The specific network structure.

3.1 Model Network Structures

As shown in Fig. 2 that the overall structure of our model consists of several convolutional layers, a region proposal network for license plate proposals generation, proposal integrating and pooling layer, multi-layer perceptrons for plate detection and bounding box regression, and RNNs for plate recognition. Given an input RGB image with a single forward evalution, the net-work outputs scores of predicted bounding boxes being license plates, bounding box offsets with a scale-invariant translation and log-space height/width shift relative to a proposal, as well as the recognized license plate labels at the same time. The extracted region features are used by both detection and recognition, which not only shares computation, but also reduces model size.

The license plate detection and recognition algorithm based on depth neural networks proposed in this paper can complete the task of license plate detection and recognition at the same time. In the network structure, the convolution features of the shallow layer and the convolution features of the deep layer are simultaneously utilized, thus improving the expressiveness of the convolutional neural network and further improving the recognition accuracy of the model. The model in this paper can be trained end-to-end and the license plate detection and recognition share a set of shallow convolution feature extraction networks. Therefore, the model can greatly reduce the amount of computation in the previous convolution feature extraction process, improve the recognition efficiency and speed up the vehicle entering and leaving time.

1) CNN Layers: shallow feature extraction networks. The residual network in ResNet is used to reduce the calculation and parameter amount, and the residual learning structure can be realized through forward neural network +shortcut connection, as shown in the structure diagram. Moreover, the shortcut connection is equivalent to simply performing equivalent mapping, which will not generate additional parameters or increase computational complexity. Moreover, the entire network can still undergo end-to-end back propagation trainings.

2) Candidate box generation networks: The RPN network layer is proposed in the Faster R-CNN network model designed by Ren et al. This is a full volume network, which inputs an image of any size and outputs a set of candidate boxes [19]. The generated candidate frames are sent to the training network of Fast R-CNN for target detection and accurate positioning. After inputting a feature map, The RNP uses a sliding window of n * n (where n is 3, i.e. a sliding window of 3 * 3). Each sliding window is mapped to a low-dimensional feature (ZF network is used in this paper, and the feature is 256-d). This feature is input into two full connection layers, one is used for border regression and the other is used for classification. The position of each sliding window is simultaneously predicted for multiple areas. The maximum number recommended for each region is k (where k is 9), so the border regression layer has 4k outputs and the classification layer has 2k outputs. The value output by the classification layer cls indicates that the candidate frame is the foreground or the background, while the border regression layer reg outputs the coordinates and offsets of the candidate frame. The RPN network structure is shown in the Fig. 3.

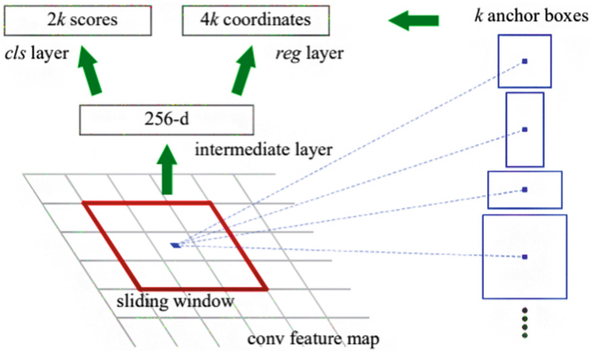

Fig. 3. The RPN network structure.

According to the size and aspect ratio of the China's license plate, three different scales (128, 192, 256) and two different aspect ratios (3, 2) are designed, corresponding to each position on the convolution map. The k (k = 6) anchor of different scales and aspect ratios generated in advance can be obtained through the RPN network. The license plate classification layer outputs 2K scores, corresponding to the probability of each anchor frame at each position being a license plate. Reg layer outputs 4K values, which respectively correspond to the offset coefficient of the nearest calibration frame at each anchor frame. Given an anchor frame with a center of (xa, ya) and a length and width of wa and ha respectively, the four values output by the regression layer are (tx, ty, tw, th), where tx, ty is the scaling factor and tw, th is the logarithmic value of the offset at the anchor frame center point. The new position calculation formula of anchor frame after passing through regression layer is as follows:

$$\mathrm{x} = x_a + t_x w_a \quad \mathrm{y} = y_a + t_y h_a$$

$$\mathrm{w} = w_a \exp(t_w) \quad \mathrm{h} = h_a \exp(t_h)$$

Where x and y are the coordinates of the center point of the anchor frame after the regression, and w and h are its width and height. Both the cls layer and reg layer will receive their own loss function and give the value of the loss function. At the same time, they will give back-propagation data according to the result of derivation. The loss function is calculated as follows:

$$\mathrm{L}(\{p_i\}, \{\mathrm{t}_i\}) = \frac{1}{N_{cls}} \sum\nolimits_i L_{cls}(p_i, p_i^*) + \lambda \frac{1}{N_{reg}} \sum\nolimits_i p_i^* L_{reg}(t_i, t_i^*)$$

$$t_x = (x - x_a)/w_a \quad t_y = (y - y_a)/h_a$$

$$t_w = \log(w/w_a) \quad t_h = \log(h/h_a)$$

$$t_x^* = (x^* - x_a)/w_a \quad t_y^* = (y^* - y_a)/h_a$$

$$t_w^* = \log(w^*/w_a) \quad t_h^* = \log(h^*/h_a)$$

3) The region normalization: The RoI Pooling operation is performed for each region candidate frame on the feature layer to obtain a fixed size feature representation. The RoI pooling layer can significantly accelerate training and testing, improve detection accuracy, and allow end-to-end trainings of the target detection system. The specific operation of ROI pooling is as follow:

a) According to the input image, the ROI is mapped to the corresponding position of the feature map;
b) Divide the mapped area into sections of the same size (the number of sections is the same as the dimension of the output);
c) Perform a max pooling operation on each section.

4) The License Plate Detection Network: The goal of the license plate detection network is to determine whether the candidate Region of Interests is a license plate and refine the coordinates of the candidate to complete the location of license plates.

5) The License Plate Recognition Network: The target of license plate recognition network is to recognize the characters of the Region of Interests based on the extracted features. In order to avoid the challenges brought by the license plate segmentation, the license plate recognition problem is regarded as a sequence labeling problem. A bidirectional RNN network based on BLSTM + CTC is proposed in the system model, as shown in the figure. The convolution layer proposes a feature sequence from the input picture and inputs it to the bidirectional regression neural network layer. in the feature sequence x = x1, x2, … xT, the regression layer predicts the distribution yi of each frame label xi. Finally, the preliminary prediction is converted into a tag sequence through a transcription layer. In order to solve the problem that the length of the character sequence directly recognized by CNN is much smaller than the input feature frame sequence, the CTC loss function is added at the end of the network. A blank is added to the label symbol set, and then the label is marked with RNN. Finally blank symbols and predicted repeated symbols are eliminated (Fig. 4).

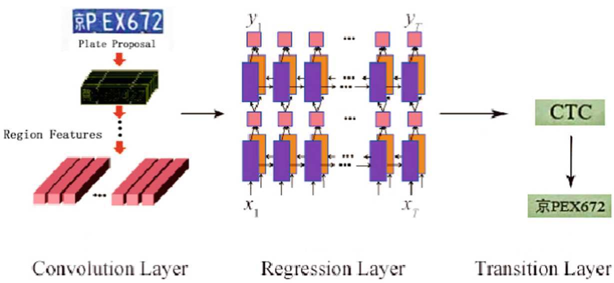

Fig. 4. The BLSTM + CTC system model.

4 The Experiments and Conclusions

4.1 The Database

The experimental data in this paper are provided by OpenITS, it is provided by Sun Yat-Sen University. The open dataset VRID for vehicle re-identification contains 10,000 images, which are captured by 326 surveillance cameras within 14 days. The resolutions of images are distributed from 400 $\times$ 424 to 990 $\times$ 1134. VRID collects 1000 vehicle IDs (vehicle identities) of top 10 common vehicle models (Table 1) to reconstruct the interference with the same vehicle model in the real world. The vehicle IDs belong to the same model have very similar appearance and their differences appears in the area of the logo and accessories. Besides, each vehicle IDs contains 10 images which are in various illuminations, poses and weather condition.

Table 1. The 10 vehicle models in the dataset

Vehicle model	Vehicle IDs	Total images
Audi_A4	100	1000
Honda_Accord	100	1000
Buick_Lacrosse	100	1000
Volkswagen_Magotan	100	1000
Toyota_Corolla_I	100	1000
Toyota_Corolla_II	100	1000
Toyota_Camry	100	1000
Ford_Focus	100	1000
Nissan_Tiida	100	1000
Nissan_Sylphy	100	1000

4.2 The Experimental Result

The experimental running environment is Win10 64 bit operation system, Python language is used for program editing, and the end-to-end license plate recognition neural network is constructed. Firstly, input data to train the neural network, get the classification model of convolution neural network according to the output results, then input the test data set to test the recognition results, and finally output the correct rate and recognition time of network recognition. The recognition accuracy rate of end-to-end license plate recognition neural network proposed in this paper is 95.4%, and the time is 200 ms. Compared with the document [14], the recognition accuracy is improved and the recognition time is greatly reduced (Table 2), which is conducive to speeding up the speed of vehicles entering and leaving the underground parking lot.

Table 2. The results of two algorithms

Method	Performance (%)	Speed (per image) (ms)
Ours (End-to-End)	95.4	300
Document [14] (SVM)	94.6	643

4.3 Conclusions

The system model proposed in this paper can be used for end-to-end training and complete the license plate detection and recognition tasks simultaneously in one time. The entire framework does not require artificially extracted features, all of which are obtained by deep neural networks through learning, and can be trained end-to-end, greatly shortening the training time of the model. The license plate detection and license plate recognition tasks are combined together, and the problem of license plate character segmentation is avoided at the same time. Compared with the two separate sub-tasks, the method is more efficient, and the character recognition is no longer affected by the character segmentation results.

Acknowledgement. This work was supported by the Key Research and Development Project of Ganzhou,the name is "Research and Application of Key Technologies of License Plate Recognition and Parking Space Guidance in Intelligent Parking Lot".

References

1. Mohandes, M., et al.: Preference-based smart parking system in a university campus. IET Intell. Transp. Syst. **13**(2), 417–423 (2019)
2. Saha, S.: A Review on Automatic License Plate Recognition System (2019)
3. Bei, C., Cao, W., He, Y.: An automated vision system for vehicle license plate localization, segmentation in real life scene. Int. J. Pattern Recogn. Artif. Intell. **32**(7), 1855010 (2017)
4. Wang, C.Q., et al.: The research on locating license plate of vehicle based on color conversion and projection. In: Advanced Materials Research, vol. 655–657, pp. 769–772 (2013)
5. Namysl, M., Konya, I.: Efficient, lexicon-free OCR using deep learning (2019)
6. Yao, Z., Yi, W.: Bionic vision system and its application in license plate recognition. Nat. Comput. **3**, 1–11 (2019)
7. Li, H., et al.: Reading car license plates using deep neural networks. Image Vis. Comput. **72**, 14–23 (2018). S0262885618300155
8. Ren, S., et al.: Faster R-CNN: towards real-time object detection with region proposal networks. IEEE Trans. Pattern Anal. Mach. Intell. **39**(6), 1137–1149 (2017)
9. Jacobs, I.S., Bean, C.P.: Fine particles, thin films and exchange anisotropy. In: Rado, G.T., Suhl, H. (eds.) Magnetism, vol. III, pp. 271–350. Academic, New York (1963)
10. Redmon, J., et al.: You only look once: unified, real-time object detection (2015)
11. Laroca, R., et al.: A robust real-time automatic license plate recognition based on the YOLO detector, pp. 1–10 (2018)
12. Lu, Y., Shridhar, M.: Character segmentation in handwritten words—an overview. Pattern Recogn. **1**(1), 77–96 (1996)

13. Khamdamov, R.Kh., Rakhmanov, H.E.: A character segmentation algorithm for vehicle license plates. Cybern. Syst. Anal. **55**(4), 649–654 (2019). https://doi.org/10.1007/s10559-019-00173-0
14. Khan, M.A., et al.: License number plate recognition system using entropy-based features selection approach with SVM. IET Image Process. **12**(2), 200–209 (2018)
15. Liu, W.B., Wang,T.: The character recognition of vehicle's license plate Based on BP neural networks. In: Applied Mechanics and Materials, vol. 513–517, 3805–3808 (2014)
16. Hui, L., Peng, W., Shen., C.: Toward end-to-end car license plate detection and recognition with deep neural networks. IEEE Trans. Intell. Transp. Syst. **20**(3), 1126–1136 (2019)
17. Yang, X., Wang, X.: Recognizing License Plates in Real-Time (2019)
18. Chen, L.W., Ho, Y.F.: Centimeter-grade metropolitan positioning for lane-level intelligent transportation systems based on the internet of vehicles. IEEE Trans. Ind. Inf. **15**(3), 1 (2019)
19. Nabati, R., Qi, H.: RRPN: Radar Region Proposal Network for Object Detection in Autonomous Vehicles (2019)