International Journal of Advanced Development in Science and Technology
Volume : 3 Issue : 03 April 2021 www.ijadst.com e-ISSN: 2582-1059

URBAN SURVEILLANCE SYSTEMS IN SMART CITIES USING R - CNN

[1] Kumaran. V, [2] Vijay Sai. V,
[1, 2] Prince Shri Venkateshwara Padmavathy Engineering College, Ponmar,
[1] hari07vijay@gmail.com.

ABSTRACT

Vehicle and license plate detection plays an important role, we tackle the problem of car license plate detection and recognition in natural scene images. Here we use a unified deep neural network, which can localize license plates and recognize the letters simultaneously in a single forward pass. The whole network can be trained end-to-end. In contrast to existing approaches which take license plate detection and recognition as two separate tasks and settle them step by step, our method jointly solves these two tasks by a single network. It not only avoids intermediate error accumulation but also accelerates the processing speed. For performance evaluation, four data sets including images captured from various scenes under different conditions are tested. Extensive experiments show the effectiveness and the efficiency of our proposed approach.

Keywords : license plate detection , R-CNN

1. INTRODUCTION

The aim of this project is to implement intelligent urban surveillance system for automated Number plate Recognition. The automated object detection algorithm is really important component in the smart cities application. In urban surveillance application the image sensor / camera plays an important role in digitizing the scene or environment. Before the deep learning era, most object detection methods need to specifically design hand-engineered features for dif- ferent objects. For the vehicle, most detection approaches [1] usually utilized information about symmetry, color, shadow, geometrical features (e.g., corners, horizontal/vertical edges), texture features and vehicle lights.

As for the license plate, the available detection methods [2] can be roughly classified into five categories: edge-based, connectivity-based, texture- based, color-based and character-based methods. Recently, the deep Convolutional Neural Networks (R - CNNs) can learn features automatically from a large amount of training data. In [3], [4], R - CNN-based methods are utilized to detect the vehicle only, and [5]–[7] are proposed to detect the license plate directly. Moreover, some methods use a cascaded strategy to detect the vehicle and the license plate, where vehicles are firstly detected, and the license plate is correspondingly localized in each vehicle region [8]–[10]. However, the above- mentioned methods either regard vehicle detection and license plate detection as two independent tasks, or detect the license plate in cascaded ways, which are less efficient. Moreover, in a cascaded way, the detection of the license plate depends on the quality of the vehicle proposals, and it is certain to be failed if the corresponding vehicle is not detected. One better way is to detect the vehicle and the license plate simultaneously as a multi-task learning system.

There have been several powerful object detection methods, e.g., SSD [11], YOLO [12] and Faster R - CNN [13]. However, we find it difficult to detect the vehicle and license plate simultaneously using these prestigious frameworks. As seen in Figure 1, some distinct license plates are unexpectedly failed to be detected, and the confidence of detected ones is also at a low level.

In deep neural networks, the vehicle and the license plate generally share the same head networks and anchor boxes. Thus, license plate detection is easily affected by the vehicle because of their inclusion relation. In this paper, we propose an end-to-end multi-branch attention neural network for simultaneously detecting the vehicle and the license plate in a given image, where two separate branches with different convolutional layers are stemmed from the backbone network to detect the vehicle and the license plate respectively. In general, the low-level features of R - CNNs have high resolution with weak semantics.

To process the digitized images searching for a particular object, smart vehicle license plate is a huge task as it will need a high CPU and memory power. To achieve this kind of functionality with distributing the processing is best way to solve. The Image processing technology to search for a number plate in a given image frame is an important task.

2. RELATED WORK

License plate detection has drawn considerable research attention. Previously, most methods often extracted hand-engineered features, such as texture features and edge features, for object detection. Recently, people usually utilize Deep Neural Networks (DNNs) for feature representation.

R - CNNs for Object Detection Over the past few decades, methods with economic features and inference schemes have been popular for efficiency, such as DPM [19]. In recent years, the DNNs have been driving the advance of object detection due to the powerful ability of feature representation, and the R - CNN-based approaches have achieved state-of-the- art performances. R-R - CNN [20] is a milestone for object detection, which utilizes Selective Search [21] to generate excessive region proposals and then apply R - CNNs to classify each region. The follow-up Faster R - CNN [13] proposes the region proposal network (RPN), and combines it with the detection block [22] into an end-to-end detection framework with two stages. Moreover, YOLO [12] and SSD [11] can directly predict/regress object bounding boxes using an end- to-end network in a single shot, where SSD [11] can detect the object of various scales by combining multi-scale features. YOLOv2 [16] proposes a dimension clustering strategy to automatically find better priors for better detections. FPN [23] attempts to create feature pyramids that have strong semantics at all scales by combining low-level features and high-level features. FAN [18] utilizes an attention mechanism to improve the detection of the occluded faces [24].

Direct License Plate Detection [26] proposes a novel method to detect the license plate by principal visual word

3. METHODOLOGY

We propose an end-to-end multi-branch attention neural network to detect license plate simultaneously, where two separate branches with different convolutional layers are designed for vehicle detection and license plate detection respectively. The license plates are detected with low-level features and the vehicles are localized with multi- level features. Moreover, a task-specific anchor design strategy is applied for better object predictions. Besides, we employ the attention mechanisms and feature-fusion strategies to improve the recall of small-scale cases. The overall network archi- tecture is demonstrated in Figure 2.

Note: The depth of the attention masks, classification, and regression head layers is 1, and Figure 2 is only for demonstration discovery and local feature matching, which can adaptively cope with different changes of the license plate, such as rotation, scaling, illumination. Reference [27] presents a robust and efficient approach for license plate detection, which firstly accelerates the license plate localization using an effective image down-scaling method, and then utilizes dense filters to extract candidate regions, and finally identifies the true license plates using a cascaded classifier. Reference [5] utilizes customized YOLO [12] and YOLOv2 [16] to handle license plate detection in the wild, which deals with the license plates captured under conditions like bad weathers, lighting, traffics. Reference [6] presents a method for license plate detection aiming at images captured with low-resolution cameras from a long distance. Reference [7] proposes a R - CNN-based MD-YOLO framework for multi-directional license plate detection.

Cascaded License Plate Detection [8] proposes a method for license plate detection using vehicle region extraction, which utilizes

R-R - CNN [20] to generate vehicle proposals and then localize the license plate in each vehicle region. Reference [9] proposes a cascaded convolutional neural network for license plate detection, which firstly applies the RPN module to generate candidate vehicle proposals and then detects the license plate based on each proposal. Reference [10] introduces a novel R - CNN framework capable of detecting and rectifying multi directional license plates in a cascaded way.

Automatic car license plate detection and recognition plays an important role in intelligent transportation systems. It has a variety of potential applications ranging from security to traffic control, and attracts considerable research attentions during recent years. However, most of the existing algorithms only work well either under controlled conditions or with sophisticated image capture systems. It is still a challenging task to read license plates accurately in an uncontrolled environment.

The difficulty lies in the highly complicated backgrounds, like the general text in shop boards, windows, guardrail or bricks, and random photographing conditions, such as illumination, distortion, occlusion or blurring. Previous work on license plate detection and recognition usually considers plate detection and recognition as two separate tasks, and solves them respectively by different methods. However, the tasks of plate detection and recognition are highly correlated. Accurate bounding boxes obtained via detection methods can improve the recognition accuracy, while the recognition result can be used to eliminate false positives vice versa. Thus in this paper, we propose a unified framework to jointly tackle these two tasks at the same level.

A deep neural network is designed, which takes an image as input and outputs the locations of license plates as well as plate labels simultaneously, with both high efficiency and accuracy. We prove that the low level features can be used for both detection and recognition. The whole network can be trained end to- end, without using any heuristic rule.

An overview of the network architecture is shown in Figure 1. To our knowledge, this is the first work that integrates both license plate detection and recognition into a single end-to-end trainable network and solves them at the same time. The main contributions of this work are as follows:

In this paper, a novel approach, which is based on DL, is proposed for detection of license plates in

given images. The proposed method does not use any pre-processing for improving the quality of the input

images. More specifically, the proposed approach uses Faster RCNN for detection of the license plates [15].

International Journal of Advanced Development in Science and Technology
IJADST Volume : 3 Issue : 03 April 2021 www.ijadst.com e-ISSN: 2582-1059

Three Faster RCNN modules are used where each of them uses a pre-trained CNN model such as AlexNet, VGG16 and VGG19, respectively. Each faster RCNN produces a rectangle that is defined by four parameters such as X and Y coordinates of the upper corner of the rectangle and width (W) and height (H) of the rectangle. Thus, three rectangles are produced by three faster RCNN models. A fusion layer is used after faster RCNN models. The fusion is handled by using average operator on column wise for X and Y coordinates from all faster RCNNs and maximum operators is used on column wise for W and H values that obtained from all faster RCNNs. A dataset, that contains 502 images, is used in our experiments.

The images were collected from various environments and various conditions for making the dataset more realistic. These all images are used for training of the faster RCNN models and the flipped and rotated versions of the images are used in testing of the proposed method. The obtained results show that the proposed method is quite successful in detection of the license plates.

The original contribution of this work is as following.

The pre-trained CNN models based faster RCNN models are generally used in various object detection applications. However, a fusion operation over multiple faster RCNN models has not been applied on license plate detection until so far. The fusion operation is justified both mathematically and experimentally. While single faster RCNN models produce low detection rates, the fused model highly improves the experimental results.

The rest of the paper is organized as follow. Next section briefly introduces the theory of the faster RCNN. Section 3 describes the proposed method in detail. Experimental works and results are given in Section 4.The paper is concluded in Section 5.

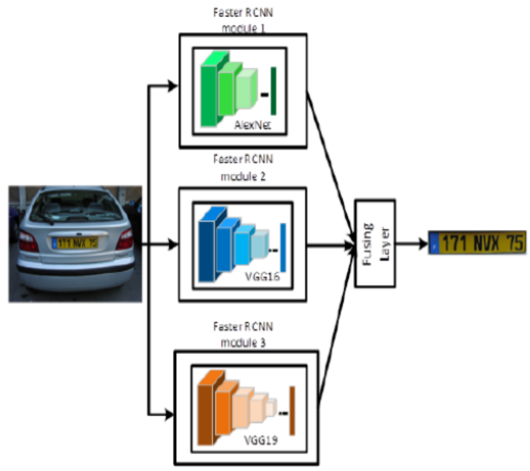

Fig 1 :

Formula

$$Bounding\_box_{AlexNet} = [X_{AlexNet}\ Y_{AlexNet}\ W_{AlexNet}\ H_{AlexNet}]$$

$$Bounding\_box_{VGG16} = [X_{VGG16}\ Y_{VGG16}\ W_{VGG16}\ H_{VGG16}]$$

$$Bounding\_box_{VGG19} = [X_{VGG19}\ Y_{VGG19}\ W_{VGG19}\ H_{VGG19}]$$

$$Fused\_Bounding\_box = \left[average \begin{bmatrix} X_{AlexNet} & Y_{AlexNet} \\ X_{VGG16} & Y_{VGG16} \\ X_{VGG19} & Y_{VGG19} \end{bmatrix} max \begin{bmatrix} W_{AlexNet} & H_{AlexNet} \\ W_{VGG16} & H_{VGG16} \\ W_{VGG19} & H_{VGG19} \end{bmatrix} \right]$$

Faster RCNNs for efficient detection of the License plates.

It is worth of mentioning that the proposed method has training and testing phases. Initially, a training image set is used to train the all faster RCNN modules. After training of the all faster RCNN modules, the testing phase can be applied on the testing images

- A single unified deep neural network is proposed, which can detect license plates from an image and recognize the labels all at once. The whole framework involves no heuristic processes, such as the use of plate colors or character space, and avoids intermediate procedures like character grouping or separation. It can be trained end-to-end, with only the image, plate positions and labels needed for training. The resulting system achieves high accuracy on both plate detection and letter recognition.

- Secondly, the convolutional features are shared by both detection and recognition, which leads to fewer parameters compared to using separated models. Moreover, with the joint optimization of both detection and recognition losses, the extracted features would have richer information. Experiments show that both detection and recognition performance can be boosted via using the jointly trained model.

- By integrating plate recognition directly into the detection pipeline, instead of addressing them by separate models, the resulting system is more efficient. With our framework, we do not need to crop the detected license plates from the input image and then recognize them by a separate network. The whole framework takes about 0.31 second for an input image of 600×600 pixels on a Titan X GPU. It should be note that although a number of methods has been proposed for both text detection and recognition in natural scene, e.g., [1]–[8], our method is quite different from them.

- The most remarkable difference is that our network can be trained end-to-end, while other methods combine results from separately trained models to obtain the final detection and recognition results. With this innovation, some pre-processing, like character detection or character grouping, are eliminated, and the intermediate errors can be avoided. The learned features can be more discriminative and lead to a better performance. We will investigate some related work in the following section. The rest of the paper is organized as follows. Section 2.

Region proposal network (RPN) aims to generate potential regions and it employs a network to determine if the potential regions contain any objects [24, 25]. The region proposals are generated by the selective search algorithm. The produced regions are ranked by the RPN and the ones most likely containing objects are selected.

It can be trained end-to-end, with only the image, plate positions and labels needed for training. The resulting system achieves high accuracy on both plate detection and letter recognition Moreover, with the joint optimization of both detection and recognition losses, the extracted features would have richer information. Experiments show that both detection and recognition performance can be boosted via using the jointly trained model.

Brief discussion on related work. Section 3 presents the integrated model, and introduces each part in detail. Experimental verifications are followed in Section 4, and conclusions are drawn in Section 5.

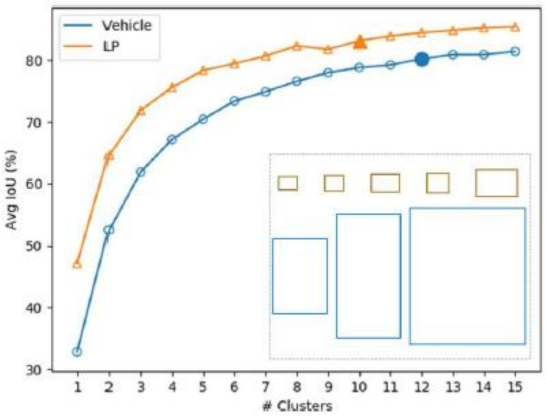

Fig 2:

4. EXPERIMENTS

The experiments were conducted on a computer having an Intel CPU and 64 GB memory. A publicly available dataset was used in experiments that contains 502 images [29]. The vehicle images were collected under various environmental conditions such as cloudy day, rainy day and night lighting. The images cover different types of vehicle such as cars, trucks, buses and mini buses. All 502 images were used in training of the proposed approach and for testing procedure, randomly selected 100 images were flipped and rotated in 5 and 10 degrees. While Figure 4 shows some training sample images, Figure 5 shows some test sample images.

Figure 4. Some sample images that were used in training of the proposed method

The training parameters of the faster RCNN modules set as following. The stochastic gradient descent with momentum optimizer was utilized in the training process. Maximum epoch, mini-batch size, and initial learning rate were set to 10, 1, and 0.001, respectively. In addition, the positive and negative overlap ranges were scaled to the [0 - 0.3] and [0.6 - 1] ranges, respectively. The number of region proposals to randomly sample from each training image was selected as [256 128]. Box pyramid scale, which was named as anchor box pyramid scale factor, is also 1.2.

Table 1. The training procedure of the faster RCNN with VGG16 model

Epoch	Iteration	Time Elapsed	Mini-batch Loss	Mini-batch Accuracy	Mini-batch RMSE	Base Learning Rate
10	4500	00:22:53	0.0031	100.00%	0.37	0.0010
10	4550	00:23:08	0.0219	100.00%	0.54	0.0010
10	4600	00:23:22	0.0468	99.22%	0.76	0.0010
10	4650	00:23:36	0.0138	100.00%	0.36	0.0010
10	4700	00:23:50	0.0033	100.00%	0.39	0.0010
10	4750	00:24:04	0.0200	100.00%	0.36	0.0010
10	4800	00:24:18	0.0172	100.00%	0.39	0.0010
10	4850	00:24:32	0.0305	99.61%	0.47	0.0010
10	4900	00:24:46	0.0270	100.00%	0.44	0.0010
10	4950	00:25:00	0.0153	100.00%	0.44	0.0010

Table 1 shows the training iterations the Faster RCNN structure. The columns of the

International Journal of Advanced Development in Science and Technology
IJADST Volume : 3 Issue : 03 April 2021 www.ijadst.com e-ISSN: 2582-1059

Table 1 show epoch, iteration, time elapsed, Mini-batch loss, Mini-batch accuracy, Mini-batch RMSE and base learning rate, respectively. As seen in Table 1, the training procedure reached the maximum accuracy (100%) at 10 epoch and 4950 iterations. The mini-batch RMSE value was 0.44. The learning rate was not changed during the iterations. The obtained test results are given in Figures. 6, 7, 8 and 9, respectively. While Figure 6 shows the AlexNet based faster-RCNN's results,Figures 7, 8 and 9 show the VGG16 based faster RCNN's results, VGG19 based faster RCNN's results, and the fused results. The detected license plate regions were indicated with colored rectangles for each method. Figure 6 shows the test results for pre-trained AlexNet model based faster-RCNN. As seen in given sample images, the "yellow" rectangles were used to locate the license plate regions. While all car's licenseplates were detected correctly, the license plates for tracks were not detected.

Table 2. Performance comparison of individual Faster-RCNN modules and fusing layer

Pre-trained CNN models	Accuracy (%)
AlexNet	74
VGG16	93
VGG19	87
Fusing Layer	97

In Table 2, the AlexNet model produces 74% accuracy score which is the worst among all results. The second worst result 87% is obtained by the VGG19 model. VGG16 model produces 93% accuracy score which is better result than AlexNet and VGG19 models. The Fusing Layer produces the best result which is 97%.

5.CONCLUSION

In this paper, We are targeting to solve the problem that the vehicle affects license plate detection when detecting the vehicle and license plate simultaneously. By adding a task-specific anchor design strategy, the network can obtain better predictions. Finally, we validate our method of high accuracy, generalization capability and efficiency using images collected from real scenes and public datasets. For future work, we hope to evaluate whether our proposed multi-branch strategy can be applied to other prestigious frameworks, like YOLO the open challenges and suggested future research directions for ALPR solutions.

6. REFERENCES

[1] L. Neumann and J. Matas, "A method for text localization and recognition in real-world images," in Proc. Asian Conf. Comput. Vis., 2011, pp. 770–783.

[2] L. Neumann and J. Matas, "Real-time scene text localization and recognition,"

in Proc. IEEE Conf. Comput. Vis. Pattern Recognit., Jun. 2012, pp. 3538–3545.

[3] K. Wang, B. Babenko, and S. Belongie, "End-to-end scene text recognition,"

in Proc. IEEE Int. Conf. Comput. Vis., Nov. 2011, pp. 1457–1464.

[4] T. Wang, D. Wu, A. Coates, and A. Y. Ng, "End-to-end text recognition with convolutional neural networks," in Proc. IEEE Int. Conf. Pattern Recognit., Nov. 2012, pp. 3304–3308.

[5] A. Bissacco, M. Cummins, Y. Netzer, and H. Neven, "PhotoOCR: Reading text in uncontrolled conditions," in Proc. IEEE Int. Conf. Comput. Vis., Dec. 2013, pp. 785–792.

[6] M. Jaderberg, A. Vedaldi, and A. Zisserman, "Deep features for text spotting," in Proc. Eur. Conf. Comput. Vis., 2014, pp. 512–528.

[7] M. Jaderberg, K. Simonyan, A. Vedaldi, and A. Zisserman, "Readingtext in the wild with convolutional neural networks," Int. J. Comput. Vis., vol. 116, no. 1, pp. 1–20, 2016. 55

[8]M Liao,B Shi,X Wang,and W.liu,"TextBoxes: A fast text detector with single deep neural network".

[9] S. Du, M. Ibrahim, M. Shehata, and W. Badawy, "Automatic license plate recognition (ALPR): A state-of-the-art review," IEEE Trans. Circuits Syst. Video Technol., vol. 23, no. 2, pp. 311–325, Feb. 2013.

[10] W. Zhou, H. Li, Y. Lu, and Q. Tian, "Principal visual word discovery for automatic license plate detection," IEEE Trans. Image Process., vol. 21, no. 9, pp. 4269–4279, Sep. 2012.

[11] C. N. E. Anagnostopoulos, I. E. Anagnostopoulos, V. Loumos, and E. Kayafas, "A license plate-recognition algorithm for intelligent transportation system applications," IEEE Trans. Intell. Transp. Syst., vol. 7,no. 3, pp. 377–392, Sep. 2006.

[12] G.-S. Hsu, J.-C. Chen, and Y.-Z. Chung, "Application-oriented license plate recognition," IEEE Trans. Veh. Technol., vol. 62, no. 2, pp. 552–561, Feb. 2013.