Car Plate Detection Based on Yolov3

Siyuan Shen1, Lidan Wang1,3,4,*, Shukai Duan2,3,4,5, Xin He1

1College of Electronic and Information Engineering, Southwest University, Chongqing, 400715, China
2College of Artificial Intelligence, Southwest University, Chongqing, 400715, China
3Brain-inspired Computing & Intelligent Control of Chongqing Key Lab, Chongqing, 400715, China
4National & Local Joint Engineering Laboratory of Intelligent Transmission and Control Technology, Chongqing, 400715, China
5Chongqing Brain Science Collaborative Innovation Center, Chongqing, 400715, China
*Corresponding author’s e-mail: ldwang@swu.edu.cn

Abstract. An intelligent car plate detection method can make the travel more convenient and efficient. However, traditional methods are reasonably effective under the specific circumstances or strong assumptions only, and there are few databases for car plate detection. Therefore, a novel real-time car plate detection method based on improved Yolov3 has been proposed. In order to select the more precise number of candidate anchor boxed and aspect ratio dimensions, the K-Means algorithm is utilized. To solve the short of the available car plate database, a car plate database which has 6668 pictures has been established. As shown in the experimental results, the method which is proposed by this paper is better than original Yolov3. Thanks to the car plate database, the proposed method obtained better results even in the situation of inclination, too bright or too dark, different weather and so on.

1. Introduction

With the number of vehicles increased dramatically, the traffic pressure is becoming more and more serious. Not only traffic jams, frequent traffic accidents, and even the management of vehicles, we need to spend a lot of human and material resources, which makes our life less convenient. As one of the core technologies in intelligent transportation, car plate detection has been widely concerned in recent years[1-4]. If there is deviation in car plate detection, the subsequent operation will be meaningless.

Traditional car plate detection has reached a high precision in a specific environment. But in the complex environment, such as illumination, shooting distance and bad angle, the traditional car plate detection algorithm has some limitations. Traditional detection algorithms include AdaBoost's license plate detection algorithm[5] which is fast but could be failed if the distance of the car plate too large, detection algorithm based on texture features[6][7] which has the higher precision under the conditions of high contrast, clear texture, the mathematical morphology detection algorithm[8][9] is fast but needs to be combined with other detection algorithms, detection algorithm based on cellular nonlinear[10], the detection algorithm based on colored image segmentation[11][12] has better precision but it is slow. Traditional car detection methods have great limitations, and the different scenes in the real world make license plate detection very challenging[13]. Therefore, in the complex environment, it is difficult to
propose a robust algorithm. Although people adopt a variety of independent features and merge many models[14], it is difficult to detect those challenging car plates.

With the improvement of computing power, deep network has returned to people's vision which has been applied to all areas of life[15-18]. The convolutional neural network method is applied to the car plate detection, which can automatically learn features from the database. As for the car plate detection, the ability of achieving real-time is very important. The current relatively mainstream detection algorithms can be approximately divided into two categories. One is to focus on classification as the theme, such as R-CNN[19], Fast R-CNN[20], Faster-RCNN[21]. Generally, the method based on classification has higher accuracy but does not have an advantage in detection time. The other is based on regression, such as Yolo[22], SSD[23], Yolov2[24], Yolov3[25]. Generally, the method based on regression is faster and can achieve real-time results, but it has lower accuracy.

Data means a lot in the time of deep learning era, the scholar need data to train their model. Although car plate detection has developed for a long time, there is still a lack of standard database. Most of the work is based on the scholars' own database[26][27], which is not open-source. So a lot of work can't be compared. The database provided to the public is also in a limited range. To solve the shortage of the available car plate database, a car plate database has been established.

2. Database

Data means a lot in the deep learning era. A database with sufficient diversity could prompt the system performance. Nowadays, car plate detection methods are tested in a small or unrepresentative database. Because there are too few license plate databases for scholars to use. Many databases of car plate detection come from traffic monitoring system, high-speed toll station, parking lot, etc[28]. Those photos are usually obtained with sufficient light source or supplementary light source. The Caltech[29] and Zemris[30] databases used high-resolution cameras to collect close to 700 photos on the road, but they were not diverse enough to cover different weather conditions. Azam[31] and Aolpe[32] point out that there are few researches on license plate detection under extreme conditions, and data are collected under various conditions, such as blur, weak light, bad weather, etc.

In this paper, a car plate database has been established. There are 6668 images in the proposed database. Each of the image is 720*1160. In the image, only one car plate is presented. Besides, the proposed database has enough diversity. Images in the database has the situation include the insufficient light intensity, Shooting distance too close or too far, different weather, and inclination (see Figure 1). In the proposed database, each image is labeled by hand.

Figure 1. Samples from the Database
Table 1. Comparison of Car Plate Detection Database

| Year | Zemris | Azam | AOLPE | OURs |
|------|--------|------|-------|------|
| 2002 | 2015   | 2017 | 2019  |      |
| Number | 510 | 850 | 4200 | 6668 |
| Distance | ✗  | ✗  | ✗  | ✔   |
| Dark | ✗  | ✔  | ✔  | ✔   |
| Weather | ✔  | ✔  | ✔  | ✔   |
| Incline | ✗  | ✔  | ✔  | ✔   |
| Annotation | ✗  | ✗  | ✔  | ✔   |

As a result, our database is more diverse. On the other hand, the diverse database make the model training more effective and more accurate (see Table 1).

3. Proposed model

3.1. K-means Algorithm
The better prior boxes we pick, the better results we get. That means the quality of prior boxes is of vital important to our model. The anchor boxes defined in the original YOLO is not suitable for our model. In this paper, we choose K-means as the clustering algorithm. K-means algorithm helps to find the more precise priors than hand pick. In traditional K-means clustering, indirect clustering is carried out by measuring the similarity between samples, usually using the Euclidean distance or Manhattan distance as the measurement formula. In yolov, as a result, the bigger border box will get more errors than the smaller one. Therefore, Yolo utilizes IOU to reflect the error between the candidate box and the truth, and its distance formula is:

\[
d(box, centroid) = 1 - IOU(box, centroid)
\]

\[d(box, centroid) = 1 - \text{IOU}(box, centroid)\] (1)

\(\text{IOU}(box, centroid)\) calculates the Intersection over Union (IoU) between a box and k clusters. We run k-means algorithm with different K values. The Figure 2 shows the results between the K and average IOU.

![Figure 2. K-means clustering analysis result](image)

In this paper, we choose K=10 as the K clusters. As seen in Figure 2, when K<10, the average IOU is not so good, when K>10, the average IOU is higher but make the model more complex. In consideration of the complexity and the precision, we choose K=10.

3.2. The Improved Model
Yolov3 is a model with both detection speed and detection accuracy, which has a wide range of applications. In this paper, our model is based on Yolov3. The more convolutional layer used, the better result we get. Inspired by the above thought, the model proposed by this paper has a deeper structure which not only more suitable for our database, but also get the ability of detect target in the more fine-grained level. In the original YOLOv3, the Darknet-53 is used as features extraction. In this paper, we improved the Darknet-53 which the input has been changed from 416*416 to 832*832, because the picture in the database is 720*1160. The reason why we expand the dimension of the input is we want...
the model as far as possible do not loss information. That will promote the model performance. Logically, the last three convolutional layer will get the 104*104, 52*52, 26*26 feature map. We still want to get the more fine-grained level features, because the high-level features has more semantic information. And the low-level features has more spatial details. To get the smaller feature map, the improved model add an additional convolutional layer behind the 26*26 feature map which stride is 2. Then the 13*13 feature will concatenate with the 26*26 feature map which do the maxpooling operation. Then the results will detect target in the high level. After the concatenate operation, the results will conduct up sampling and concatenate with the 26*26 feature map and the 52*52 feature map which do the maxpooling operation. The results will detect target in the middle level. The middle level’s feature map will do upsampling operation and concatenate with the 52*52 feature map. That will do the low level detection.

In this paper, the proposed model adopts the multi feature map fusion makes model more effective. The improved model will be described in the Figure 3.

As seen in the Figure 3, our model will detect target in three times but in different dimension. We will take the best results as the final result. The fusion of the different dimension feature maps makes the model get multi scale information. This paper will test on the proposed car plate database, and the experimental results will be shown in the next part.

3.3. The Experimental Results

In this paper, all the experiments will be conducted in the proposed car plate database. We set 5500 pictures for training, and 1168 samples for testing. Our experimental environment is listed as table 2.

| Table 2. Experimental Environment |
|-----------------------------------|
| OS      | CPU        | GPU       | RAM     |
| Ubuntu16.04 | Intel Core i7-8700k | GTX1080Ti | 16GB   |

The initial model parameters are set as follows: learning rate: in the first 50 epochs, a learning rate of 0.0001 is used to obtain a good model, and the next 150 epochs are trained with a learning rate of 0.00001, batch size: 1, number of the epochs: 200, the optimizer we utilize the Adam.

In this paper, we will utilize the proposed network structure for verification. We will compare and analyse the results from the predicted and the ground truth. In this paper, IOU is utilized to judge whether the prediction is precise when comparing the predicted result of the model and the ground truth. When IOU > 0.5, it is judged as precise prediction, otherwise it is judged as false prediction.
In the experiment, FN is used to represent the false negative, which represents the number of car plates that are labeled but not predicted. FP is used to represent false positive, which stands for the number of car plates predicted by the model but not labeled. RT indicates that the prediction is correct, which means that the number of car plates predicted correct by the model. Precision, Recall rate are used to evaluate the model.

When the model is used for car plate detection, precision represents the prediction accuracy of the model on the car plate. Recall represents the undetected car plate of the model. The car plate database has marked 6668 pictures. In this paper, 5500 pictures will be trained and 1168 pictures will be used to test. Some results are presented in the Figure 4.
From Figure 4, the proposed model completes the detection task excellent. In the Figure 4, we can see different situation in the image. Such as different weather, insufficient of the light, different shooting distance. The results show the model work excellent. The model not only detect the position of the car plate, but also get the high confidence value. That means our model is very confident about where the car plate is.

The Precision and Recall is listed as Table 3

| Model          | FN  | FP  | RT   | Precision | Recall |
|----------------|-----|-----|------|-----------|--------|
| Yolov3         | 140 | 96  | 932  | 90.75%    | 86.94% |
| Improved       | 49  | 25  | 1094 | 97.77%    | 95.71% |

As seen in Table 3, our model detect most of the car plates. It fails in some challenging images which are under the conditions of blur and too bright. In conclusion, the proposed model can detect car plate excellent.

As comparison, we still run the database on the original Yolov3. As listed in the Table 4, both method could detect the car plate, the proposed model is better than the original Yolov3.

Table 4. Comparison between Yolov3 and Proposed Model

| Model          | FN  | FP  | RT   | Precision | Recall |
|----------------|-----|-----|------|-----------|--------|
| Yolov3         | 140 | 96  | 932  | 90.75%    | 86.94% |
| Improved       | 49  | 25  | 1094 | 97.77%    | 95.71% |

As seen in Table 4, the original YOLOv3 has lower precision and recall. Our model is better in visual. As a result, the model we proposed works excellent. The model can handle the database more effective. The less information loss, the better results we get. At the mean time, our model fusions more feature maps to detect target than the original YOLOv3.

As we all known, Yolo is fast. It can detect object in real-time. In this paper, the method we proposed or the original Yolov3 can achieve the real-time level. When the model we proposed to test the car plate, the FPS can achieve 33 in one GTX1080TI GPU. That means the proposed method don’t consume much hardware.

4. Conclusion

In the deep learning era, a database with sufficient diversity could prompt the system performance. In this paper, we introduce a car plate database established by ourselves first. In consideration of above thought, the car plate database includes different situations. Such as inclination, too bright or too dark, different weather. It helps to train the model. In order to better adapt to the proposed database, we choose K-means as the clustering algorithm which helps to find the more precise priors than hand pick. The focus of this paper is to propose an improved model based on yolov3. In order not to lose too much image information, we expand the input size. We add an additional convolutional layer to get the smaller feature map. Through the concatenate and the upsampling operations, the model can fuse different dimensions of feature maps. The fusion of the different dimension feature maps makes the model get multi scale information. In the experimental results, the improved model could detect the car plate better than original Yolov3. No matter in the visual effect or the evaluation index, the proposed model is better.

Our models not only accurately detect the license plate, but also have confidence in the location of the license plate. Through the experimental results, our model doesn’t work well when the image is blur. In the future researches, We will focus on the detection of motion blurred images.

Acknowledgments

The work was supported by National Key R&D Program of China (Grant Nos. 2018YFB1306600), National Natural Science Foundation of China (Grant Nos. 61571372, 61672436, 61601376), Fundamental Science and Advanced Technology Research Foundation of Chongqing (cstc2017jcyjBX0050, cstc2016jcyjA0547), Fundamental Research Funds for the Central Universities (Grant Nos. XDJK2016A001, XDJK2017A005).
References

[1] Karwal H, Girdhar. A. (2015) Vehicle number plate detection system for Indian vehicles. In IEEE International Conference on Computational Intelligence and Communication Technology. pp. 8-12.

[2] Patel C, Patel A, Shah D. (2015) A novel approach for detecting number plate based on overlapping window and region clustering for Indian conditions. International Journal of Image, Graphics and Signal Processing, 7(5): 58.

[3] Samma H, Lim C P, Saleh J M, et al. (2016) A memetic-based fuzzy support vector machine model and its application to license plate recognition. Memetic computing, 2016, 8(3): 235-251.

[4] Wang R, Sang N, Wang R, et al. (2014) Detection and tracking strategy for license plate detection in video. Optik-International Journal for Light and Electron Optics, 125(10): 2283-2288.

[5] Zhang X, Shen P, Bai J, et al. (2010) License plate location based on adaboost. In the 2010 IEEE International Conference on Information and Automation. Pp. 1705-1710.

[6] Yee M S, Hanzo L.(1999) Radial basis function decision feedback equaliser assisted burst-by-burst adaptive modulation. In Seamless Interconnection for Universal Services. Global Telecommunications Conference. pp. 2183-2187.

[7] Wu Y, Liu S, Wang X.(2013) License plate location method based on texture and color. In 2013 IEEE 4th International Conference on Software Engineering and Service Science. pp. 361-364.

[8] Kim S, Kim D, Ryu Y, et al. (2002) A robust license-plate extraction method under complex image conditions. In Object recognition supported by user interaction for service robots. pp. 216-219.

[9] Jin B, Jia Y. (2011) Algorithmic research of speeding license plate location based on mathematics morphology. In 2011 2nd International Conference on Artificial Intelligence, Management Science and Electronic Commerce. pp. 948-950.

[10] Zhang F Q, Duan S K, Wang L D. (2013) Application of memristor cell neural network in license plate location. Computer Science, 2013 (S1): 58-60.

[11] Yu S, Li B, Zhang Q, et al. (2015) A novel license plate location method based on wavelet transform and EMD analysis. Pattern Recognition, 48(1): 114-125.

[12] Abolghasemi V, Ahmadyfard A. (2009) An edge-based color-aided method for license plate detection. Image and Vision Computing, 27(8): 1134-1142.

[13] Wazalwar D, Oruklu E, Saniee J. (2011) Design flow for robust license plate localization In 2011 IEEE International Conference on Electro/Information Technology. pp: 1-5.

[14] Wang Q, Yuan Y, Yan P, et al.(2013) Saliency detection by multiple-instance learning. IEEE transactions on cybernetics, 43(2): 660-672.

[15] Dai J, Li Y, He K, et al.(2016) R-fcn: Object detection via region-based fully convolutional networks. In Advances in neural information processing systems. pp. 379-387.

[16] He K, Zhang X, Ren S, et al. (2016) Deep residual learning for image recognition. In Proceedings of the IEEE conference on computer vision and pattern recognition. pp. 770-778.

[17] Szegedy C, Liu W, Jia Y, et al. (2015) Going deeper with convolutions. In Proceedings of the IEEE conference on computer vision and pattern recognition. pp. 1-9.

[18] Simonyan K, Zisserman A. (2014) Very deep convolutional networks for large-scale image recognition. arXiv preprint arXiv:1409.1556.

[19] Girshick R, Donahue J, Darrell T, et al. (2014) Rich feature hierarchies for accurate object detection and semantic segmentation. In Proceedings of the IEEE conference on computer vision and pattern recognition. pp. 580-587.

[20] Girshick R. (2015) Fast r-cnn. In Proceedings of the IEEE international conference on computer vision. pp. 1440-1448.

[21] Ren S, He K, Girshick R, et al. (2015) Faster r-cnn: Towards real-time object detection with region proposal networks In Advances in neural information processing systems. pp. 91-99.
[22] Redmon J, Divvala S, Girshick R, et al. (2016) You only look once: Unified, real-time object detection. In Proceedings of the IEEE conference on computer vision and pattern recognition. pp. 779-788.

[23] Liu W, Anguelov D, Erhan D, et al. (2016) Ssd: Single shot multibox detector. In European conference on computer vision. pp. 21-37.

[24] Redmon J, Farhadi (2017) A. YOLO9000: better, faster, stronger In Proceedings of the IEEE conference on computer vision and pattern recognition. pp. 7263-7271.

[25] Redmon J, Farhadi A. (2018) YOLOv3: An Incremental Improvement. arXiv preprint arXiv:1804.02767.

[26] Jiao J, Ye Q, Huang Q. (2009) A configurable method for multi-style license plate recognition. Pattern Recognition, 42(3): 358-369.

[27] Wen Y, Lu Y, Yan J, et al. (2011) An algorithm for license plate recognition applied to intelligent transportation system. IEEE Transactions on Intelligent Transportation Systems, 12(3): 830-845.

[28] Špaňhel, J., Sochor, J., Juránek, R., Herout, A., Maršík, L., and Zemčík, P. (2017). Holistic recognition of low quality license plates by CNN using track annotated data. In 2017 14th IEEE International Conference on Advanced Video and Signal Based Surveillance. pp. 1-6.

[29] Caltech: CaltechLicensePlate Dataset (2001)
   http://www.vision.caltech.edu/html-files/archive.html

[30] Zemris:ZemrisLicensePlateDataset(2003)
   http://www.zemris.fer.hr/projects/LicensePlates/hrvatski/rezultati.shtml

[31] Azam S, Islam M M. (2016) Automatic license plate detection in hazardous condition. Journal of Visual Communication and Image Representation, 36: 172-186.

[32] Hsu G S, Ambikapathi A M, Chung S L, et al. (2017) Robust license plate detection in the wild. In 2017 14th IEEE International Conference on Advanced Video and Signal Based Surveillance pp. 1-6.