



License plate detection and recognition based on YOLOv3 and ILPRNET

Yongjie Zou¹ · Yongjun Zhang¹ · Jun Yan² · Xiaoxu Jiang² · Tengjie Huang² · Haisheng Fan² · Zhongwei Cui³

Received: 16 October 2020 / Revised: 19 May 2021 / Accepted: 29 June 2021 / Published online: 6 August 2021
© The Author(s), under exclusive licence to Springer-Verlag London Ltd., part of Springer Nature 2021

Abstract

This paper is concerned with the detection and recognition of Chinese license plates in complex backgrounds. Most applications are currently focused on good conditions. In complex natural scenes such as CCPD-DB, CCPD-FN, CCPD-Rotate, CCPD-Tile, CCPD-Weather, and CCPD-Challenge from the Chinese City Parking Dataset (CCPD), inaccurate localization and poor character recognition accuracy issues appear towards existing license plates. Therefore, this paper proposes a two-stage license plate recognition algorithm based on YOLOv3 and Improved License Plate Recognition Net (ILPRNET). In the first stage, YOLOv3 is adopted to detect the position of the license plate and then extract the license plate. In the second stage, the ILPRNET license plate recognition network is used to perform localization of license plate characters and the 2D attentional-based license plate recognizer with an CNN encoder is capable of recognizing license plates accurately. The test results indicate that our proposed algorithm performs well in a variety of complex scenarios. Especially in sub-datasets like CCPD-Base, CCPD-DB, CCPD-FN, CCPD-Weather, and CCPD-Challenge, the recognition accuracy achieved 99.2%, 98.1%, 98.5%, 97.8%, and 86.2%, respectively.

Keywords License plate detection · Character localization · License plate recognition

1 Introduction

License plate detection and recognition play a very important role in intelligent transportation systems. It has a very wide range of applications in traffic control, self-driving, parking toll stations, etc. At present, license plate recognition is still a challenging task under unrestricted conditions. The difficulty lies in accurately identifying license plate characters under extreme conditions, such as occlusion, uneven illumination, rotation (large angle), vagueness, etc. This work is worthy of further study, and there are some practical signifi-

cances. There are two methods of licence plate recognition, segmented [5,19], and unsegmented [8,24,30,31]. The first one is to segment each character of the licence plate and then recognize it using the Optical Character Recognition(OCR) [2,19] model, but there are some problems, if the segmented characters are incomplete, it will affect the accuracy of OCR recognition. This method may also be affected by conditions such as light intensity which also lowers the accuracy. In recent years, the second method has been favoured by researchers, which extracts the features of the characters to avoid the accuracy degradation caused by incomplete segmented characters. The method proposed in [8] converts character recognition into a series annotation problem and uses a Long Short Term Memory (LSTM) Network to recognize the entire sequence of plate features which does not require segmentation of characters and can tap into the contextual information of the plate with high accuracy, while consuming much computational power. In order to recognize complex scenarios, Xu et al [24] proposed a end-to-end model. An open-source dataset containing a variety of complex scenes is proposed in their work. The purposed method achieved 98.5% accuracy on this dataset but could only identify images containing one license plate at a time. Inaccurate localization problems are also found by testing on

✉ Yongjun Zhang
zyj6667@126.com

✉ Jun Yan
yan@myorbita.net

¹ Key Laboratory of Intelligent Medical Image Analysis and Precise Diagnosis of Guizhou Province, College of Computer Science and Technology, Guizhou University, Guiyang 550025, China

² Zhuhai Orbita Aerospace Science and Technology Co, Ltd, Orbita Tech Park, 1 BaiSha Road, TangJia DongAn, ZhuHai 519000, China

³ Big Data Science and Intelligent Engineering Research Institute, Guizhou Education University, Guiyang 550018, China

other open-source datasets. Li et al. [9] introduced a unified network that can locate license plates while recognizing letters simultaneously, but is less friendly to multi-oriented license plates. [19] proposed a new convolutional neural network (CNN) capable of detecting and correcting multiple distorted license plates in a single image. [29] used a combination of an LSTM sequence decoder and a GAN network for license plate recognition, which is currently the state-of-the-art method on CCPD dataset, with the drawback that it cannot be trained in parallel over time steps. In [31], we use the LSTM method to have a good performance in license plate recognition. Unlike the above methods, this paper proposed a method that combines YOLOv3 with ILPRNET, which takes advantage of YOLOv3 to detect license plates very accurately with decent generalization ability and ILPRNET to recognize license plates, which not need an extra module to handle the irregularity of license plates or segment each character for recognition and can accurately extract local feature information of each character and avoid character recognition errors caused by inaccurate character feature segmentation.

Our contributions are as follows:

1. It puts forward a new type of license plate recognition network (ILPRNET).
2. A spatial domain mechanism is designed, which compresses channel information and makes the network pay attention to the spatial information of license plate characters.
3. An encoding–decoding algorithm is developed to locate license plate characters, extract the feature vector for each character by multiplying the 2D attentional weights of the different channels with the license plate feature matrix, and classify them by a character classifier.

Through evaluating ILPRNET algorithm on public datasets CCPD [24], CLPD [29] and ALOP [6], compared with other methods, our proposed model gives a better result.

2 Related work

2.1 Licence plate detection

Licence plate(LP) detection algorithms mainly consist of traditional and deep learning methods. Traditional methods manually extract features such as color information, edge contour information, etc. In order to deal with complex background, a robust method based on wavelet transform is presented, Yu et al. [27]. By performing wavelet transform on images of vehicles and projecting the obtained image details, the license plate peak is generated, with a final accuracy of 97.91%. A framework based on gradient information and cascade detection was put forward by Wang [20] to detect license plates. In [7], a LP detection algorithm based on a local structure pattern was introduced, using post-processing of location and color information of license plates to reduce the false pos-

itive rate. Yao [25] exploited a license plate (LP) detection technique based on multi-level information fusion to reduce the high false positive rate in traditional Adaboost detectors. In the field of object detection, region-based convolutional neural networks [4] are frequently used. Faster-RCNN [17] adopts the Regional Proposal Network (RPN) to share convolutional features of the full image and generate high-quality regional proposal candidate frames that can accurately and quickly detect objects yet with low efficiency. SSD [11] eliminates the generation of regionally proposed candidate boxes and uses multi-scale feature maps to detect targets, which is extremely sensitive to small objects. Compared with other state-of-the-art methods, we found that the method of RPNet [24] does not do well in generalization and cannot accurately predict the location of the license plate for other datasets. The algorithm proposed by Zhang et al. [29] uses YOLOv2 to detect licence plates with very high licence plate recognition accuracy and consumes more computational power. However, with the upgrade of YOLO series algorithm [1,14–16], either YOLOv3 [16] or YOLOv4 [1] reaches a higher accuracy than YOLOv2. YOLOv3 runs three times faster than SSD with comparable performance, which can achieve the purpose of real-time detection. Therefore, we use YOLOv3 and YOLOv4 as the license plate detection algorithm preliminary. After experimenting as shown in Table 1, we found that the accuracy of YOLOv4 is not as high as YOLOv3 while consumes more computational power, so we finally use YOLOv3 as the license plate detection algorithm.

2.2 Text recognition

Text recognition and license plate recognition are very similar. For text recognition, current methods are mainly based on an attentional encoder–decoder to make a mapping between the input image and the output sequence. Thus, we can use text recognition to recognize license plates. Shi et al. [18] proposed a flexible thin-plate spline transformation method capable of handling irregular text in various scenes. Cheng et al. [3] improved the accuracy of text recognition by pulling back the drifting attention using a focused attention mechanism. A multi-objective corrective attention network that enables recognizing of regular or irregular scene texts was put forward by Luo et al. [12] to reduce the recognition difficulty and make it easier to read irregular text based on the attention sequence recognition network. Li et al. [10] used an LSTM-based encoder–decoder framework and a two-dimensional attention module to achieve state-of-the-art performance in regular and irregular scene text recognition. Similar to our approach, but we were able to accurately extract character features and raise the accuracy of license plate recognition using the proposed convolutional neural network and 2D attention module.

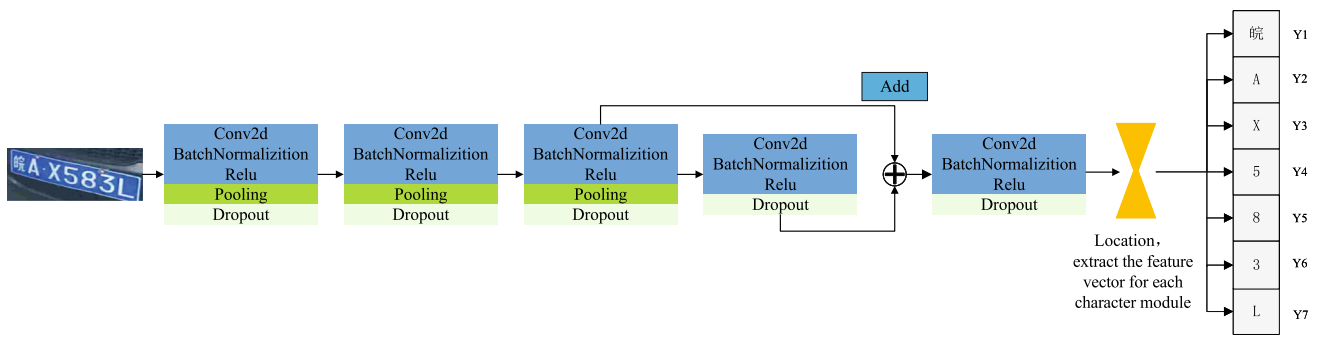


Fig. 1 License plate recognition modules. The orange regions are a character localization, front is a license plate character extraction, followed by character classifier (color figure online)

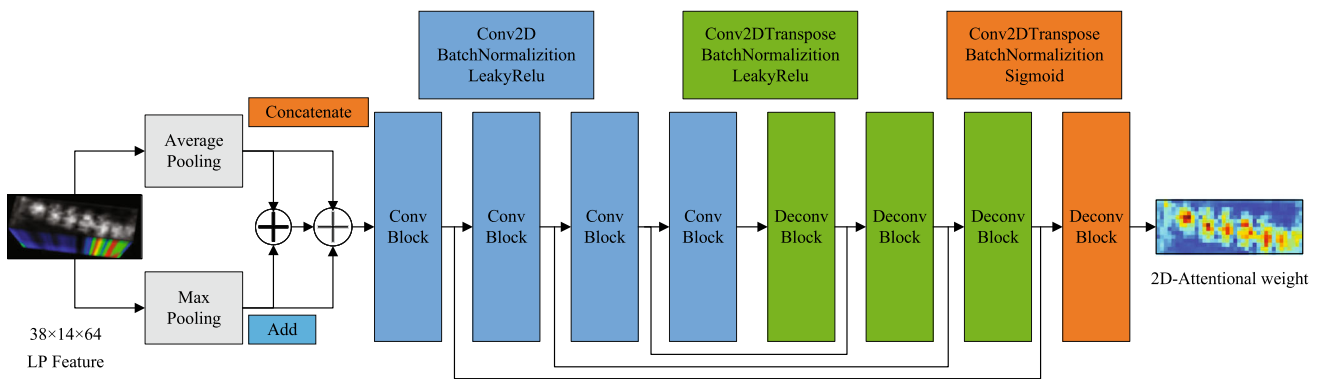


Fig. 2 License plate character location module, “LP feature” represents visualized license plate feature, and “2D-Attention” indicates the feature weights for each character visualized

3 Proposed method

3.1 License plate recognition

Figure 1 represents the process of license plate recognition, which consists of extracting LP features, locating LP characters, and extracting a feature vector for each character, character classifier.

(1) Extracting LP features

We use Convolutional Neural Networks to extract license plate features, add Batch Normalization to avoid gradient disappearance, and adopt ReLU (Rectified Linear Unit) to enhance the nonlinearity of the network. Finally, the use of Dropout after the pooling layer to avoid overfitting the network. The third- and fourth-layer features are fused to enhance the feature information of license plate characters, enabling the network to fully extract license plate features.

(2) Locating the license plate characters

The network structure of license plate character localization is shown in Fig. 2. We designed a method of U-shaped network structure incorporating spatial attention mechanism, using maximum pooling and average pooling to compress the number of channels of the extracted license plate feature information, and fuse the two features to make the network

focus on the spatial information of the license plate characters. The U-shaped network structure is then used to locate the position of the licence plate characters by combining the context of different layers. The low-level features contain high-resolution information passed directly from the encoder to the corresponding decoder after the concatenate operation, which can provide finer features for locating the characters, and the high-level features contain low-resolution information after multiple downsampling, which can provide finer context information for locating characters. Finally, the decoding layer uses a sigmoid function with values ranging from 0 to 1 aims at obtaining a 2D-Attentional weight for each character, the higher the level of activation, the more interested the position is. In Fig. 2, we visualize the attentional weights of each character and it can be found that our algorithm is able to locate each character accurately.

(3) Extracting the feature vector for each character

The process of extracting the feature vector of the licence plate character is shown in Fig. 3, where the character ‘5’ at position Y4 is extracted as an example. The 2D-Attentional weight of the character ‘5’ is obtained by the localization module and then multiplied with the licence plate feature vector to obtain the feature vector of the character ‘5’. The feature information not related to the character ‘5’ will be suppressed

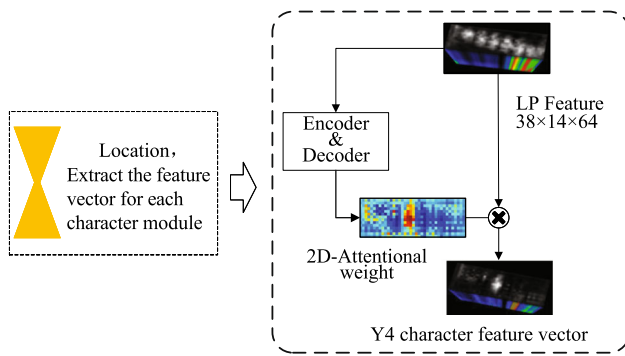


Fig. 3 Module for extracting feature vectors for each licence plate character. The red regions represent the Y4 character feature weight (color figure online)

while the relevant feature information is enhanced to ensure that the feature information of the specified character will not be interfered with and enabling the accurate extraction of the features. The heat maps of the location of license plate character features on CCPD are shown in Fig. 4. The brighter the color, the more attention is paid to the features of the location. From the heat map, we can see that the model is able to locate and extract the characteristic information of each character on the license plate. Therefore, the network is able to recognize characters accurately.

(4) License plate character classifier

Once the features of each character extracted, we feed the features into the character classifier in turn to calculate the score of each character by using the softmax activation function. We adopt the index of the character with the highest score and use the character decoder to obtain the correct character.

3.2 Training parameters

The size of the input image of the detection network is 736×736 . We use the Adam optimizer to train for 128 epochs. During pre-training, the batch size of 16 and the learning rate

of $1e-3$ are used for training. During training, the batch size of 4 and the learning rate of $1e-4$ are used. When the learning rate is not decreasing for every 3 epochs, the learning rate is multiplied by 0.1.

The size of the input image of the recognition network is 152×56 , and we use Category Cross-Entropy Loss and Adam optimizer to train for 200 epochs. During training, the batch size of 32 and the learning rate of $1e-4$ are used for training. When the learning rate is not decreasing for every 3 epochs, the learning rate is multiplied by 0.5. Category Cross-Entropy Loss is shown in equation 1, where C indicates the number of categories, y_j represents the ground truth and $y_predict_j$ represents the predicted value.

$$\text{loss} = - \sum_j^C y_j \log(y_predict_j) \quad (1)$$

4 Experimental setting

4.1 Datasets

CCPD [24] provides over 290k unique LP images, each of the images contains only one LP and each LP number is composed of a Chinese character, a English character, and five English characters or numbers. The CCPD dataset contains nine data subsets: CCPD-Base (200k), CCPD-DB (20k), CCPD-FN (20k), CCPD-Rotate (10k), CCPD-Tilt (10k), CCPD-Weather (10k), CCPD-Challenge (10k), etc. The resolution of each image is 720 (width) \times 1160 (height) \times 3 (channels). For a fair comparison, we trained our model using the same dataset standard as in [24], using half of CCPD's subset CCPD-Base as the training set and the other half as the validation set, while the remaining subsets CCPD-DB, CCPD-FN, CCPD-Rotate, CCPD-Tilt, CCPD-Weather and CCPD-Challenge and CCPD-Base (the validation set that was not used for training) were all used for testing. Compared

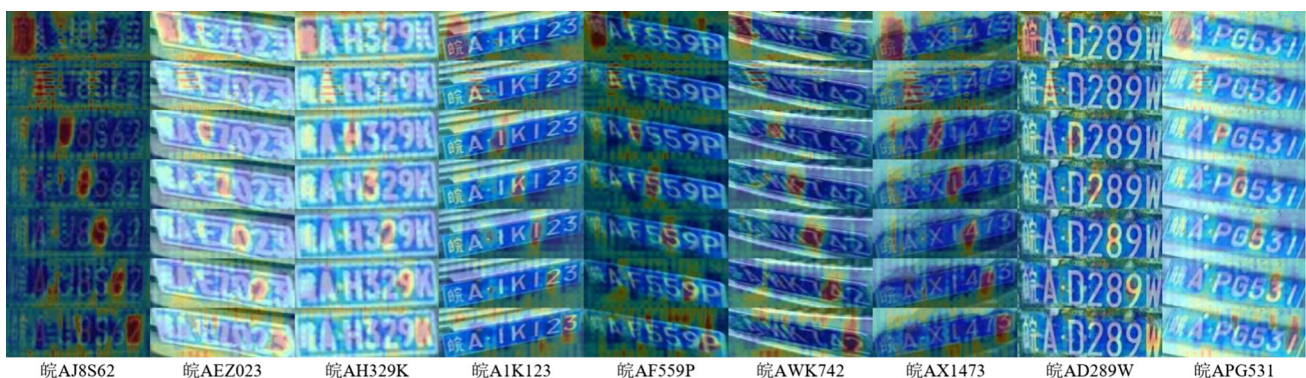


Fig. 4 Heat maps of character locations in CCPD license plates. The results show that the 2D-attention model can handle challenging cases

Table 1 mAP comparison of CCPD license plates detection, AP represents average accuracy, Base(100K) represents the validation set that CCPD is not using for training, IoU = 0.7(%)

Subset methods	AP	Base(100k)	DB(20k)	FN(20k)	Rotate(10k)	Tilt(10k)	Weather(10k)	Challenge(10k)
CascadeClassifier [21]	47.2	55.4	49.2	52.7	0.4	0.6	51.5	27.5
SSD300 [11]	94.4	99.1	89.2	84.7	95.6	94.9	83.4	93.1
YOLO9000 [15]	93.1	98.8	89.6	77.3	93.3	91.8	84.2	88.6
Faster-RCNN [17]	92.9	98.1	92.1	83.7	91.8	89.4	81.8	83.9
TE2E [9]	94.2	98.5	91.7	83.8	95.1	94.5	83.6	93.1
RPnet [24]	94.5	99.3	89.5	85.3	94.7	93.2	84.1	92.8
YOLOv4	95.1	96.8	93.7	93.1	93.5	94.7	96.6	85.5
YOLOv3	96.0	97.1	97.2	93.3	91.6	94.6	97.9	90.5

Bold indicates the largest value in this column of the Table

Table 2 Accuracy comparison of CCPD license plates recognition, Base(100K) represents the validation set that CCPD is not using for training, IoU = 0.6(%)

Subset methods	AP	Base(100k)	DB(20k)	FN(20k)	Rotate(10k)	Tilt(10k)	Weather(10k)	Challenge(10k)
Ren et al. [17]	92.8	97.2	94.4	90.9	82.9	87.3	85.5	76.3
Liu et al. [11]	95.2	98.3	96.6	95.9	88.4	91.5	87.3	83.8
Joseph et al. [15]	93.7	98.1	96.0	88.2	84.5	88.5	87.0	80.5
Li et al. [9]	94.4	97.8	94.8	94.5	87.9	92.1	86.8	81.2
Zherzdev et al. [30]	93.0	97.8	92.2	91.9	79.4	85.8	92.0	69.8
Xu et al. [24]	95.5	98.5	96.9	94.3	90.8	92.5	87.9	85.1
Zhang et al. [28,30]	93.0	99.1	96.3	97.3	95.1	96.4	97.1	83.2
Wang et al. [22]	96.6	98.9	96.1	96.4	91.9	93.7	95.4	83.1
Ours	97.5	99.2	98.1	98.5	90.3	95.2	97.8	86.2

Bold indicates the largest value in this column of the Table

with CCPD, most published datasets have limited number of images and little variation in the shooting distance of images.

CLPD [29] collected 1200 LP images of various vehicle types in all 31 provinces of mainland China, these images are from various real-life scenarios, such as photos shot with mobile phones and driving recorders, and images from the Internet. CLPD includes various information such as camera angle, time, resolution, and background. However, they are only used to evaluate our proposed model of the license plate recognition.

AOLP [6] is a license plate dataset from Taiwan. It has three sub-datasets for different scenarios: Access Control (AC), Traffic Law Enforcement (LE), and Road Patrol (RP). There are a total of 2019 license plate images, which AC has 681 images, LE has 757 images, and the RP subset has 611 images.

4.2 Evaluation metrics

We use the same IoU (Intersection over Union) metric as in [24] to evaluate the performance of the algorithm, where IoU represents the overlap between the detected license plate by YOLOv3 and the ground truth, and a license plate recogni-

tion result is considered correct only if IoU is greater than 0.6 and all characters in the license plate are recognized correctly, otherwise it is considered wrong. As shown in formula 2, pb represents the license plate box detected by YOLO algorithm and gb represents the ground truth of license plate. All experiments were implemented on an NVIDIA TITAN Xp 12GB.

$$\text{IoU} = \frac{\text{area}(\text{pb} \cap \text{gb})}{\text{area}(\text{pb} \cup \text{gb})} \quad (2)$$

5 Results

5.1 Experiments on CCPD dataset

It can be seen from Table 1 that the YOLOv3 algorithm outperforms other algorithms in terms of the seven subsets CCPD-Base, CCPD-DB, CCPD-FN, CCPD-Rotate, CCPD-Tilt, CCPD-Weather and CCPD-Challenge. Compared to the state-of-the-art methods, the average accuracy of our method improved by 1.5% percentage points on the overall dataset,

Table 3 Accuracy comparison on CCPD license plates recognition of different bounding boxes

Bounding boxes	AP	Base(100k)	DB(20k)	FN(20k)	Rotate(10k)	Tilt(10k)	Weather(10k)	Challenge(10k)
Yolo3 detection	97.5	99.2	98.1	98.5	90.3	95.2	97.8	86.2
Ground truth	98.2	99.3	98.6	98.9	97.1	98.0	98.3	87.8

Bold indicates the largest value in this column of the Table



Fig. 5 Recognition results of some license plates image, the model can recognize license plates in various scenarios

Fig. 6 Compare samples of licence plates recognized by different methods. The ground truth is shown in the parentheses

Xu [24]	皖 AY188E	皖 AC2550	皖 AD1K85	皖 AVL287	皖 AF4A44	皖 ALZ316	皖 AL733J
	(皖 AY188E)	(皖 AC256P)	(皖 AD1K85)	(琼 AVL287)	(皖 AF4A44)	(皖 AL3316)	(皖 AL733J)
Zhang [29]	皖 AY188F	皖 AC2566P	皖 A01K85	渝 AVL287	皖 AF4A44	皖 AL3816	皖 AL733U
	(皖 AY188E)	(皖 AC256P)	(皖 AD1K85)	(琼 AVL287)	(皖 AF4A44)	(皖 AL3316)	(皖 AL733J)
Ours	皖 AY188E	皖 AC258P	皖 AD1889	皖 AVL287	皖 AF4A44	皖 AL3316	皖 AL733J
	(皖 AY188E)	(皖 AC256P)	(皖 AD1K85)	(琼 AVL287)	(皖 AF4A44)	(皖 AL3316)	(皖 AL733J)

and by 5.1%, 8% and 13.7% percentage points on the DB, FN and Weather subsets, respectively. YOLOv3 is adopted for detection, which can locate the license plates accurately, and it is a real-time detection framework that can meet the requirements of road monitoring. Although RPnet [24] is an end-to-end framework that detect and recognize simultaneously. The problem of inaccurate localization leads to inaccurate recognition.

we compare our model with other state-of-the-art methods for license plate recognition. The results in Table 2 show that our approach performs better than other methods on five subsets. In particular, our method leads to the accuracy increments of 0.9% on all subsets, 0.1% on Base, 1.2% on DB, 1.2% on FN, 0.7% on weather and 1.1% on challenge, compared to the second best results. The only exception is that some algorithms are better than ours on the rotate and tilt subsets, in the future, we could correct the rotating image and improve the recognition accuracy. We compared the accuracy in different license plate bounding boxes. As shown in Tables 3 and 6, the bounding boxes detected by YOLOv3 have a very

close to the ground truth in recognition accuracy. Therefore, we can see that YOLOv3 is reliable as a license plate detector compared to a real license plate bounding boxes. Partial test results are shown in Fig. 5. Comparisons with the current the state-of-the-art methods [29] are shown in Fig. 6, where red indicates that the character was incorrectly recognized and black indicates that the character was correctly recognized, and the performance of our algorithm performs better on these samples. Some characters are misrecognized by other methods while correctly recognized by our method.

5.2 Experiments on CLPD dataset

The model trained on the CCPD-Base dataset is adopted to test the CLPD dataset. The recognition results are shown in Table 4, indicating the advantage of our model. It leads to the highest accuracy (ACC) no matter region code (a Chinese characters) is considered or not. Compared to the second best results, without the region code, ACC is increased by 6.9%, and including the region code ACC is improved by 1.9%

Table 4 Accuracy comparison of CLPD license plates recognition

Bounding boxes	Without REGION CODE ACC	ACC
Masood et al. [13]	85.2	–
Xu et al. [24]	78.9	66.5
Zhang et al. (real data only) [29]	86.1	70.8
Zhang et al. (real + synthetic data) [29]	87.6	76.8
Ours (real data only)	94.5	78.7

Bold indicates the largest value in this column of the Table

Table 5 Accuracy comparison of ALOP license plates recognition(%)

Methods	AC	LE	RP
Li et al. [9]	95.29	96.57	83.63
Li et al. [8]	94.85	94.19	88.38
Hsu et al. [6]	88.5	86.6	85.7
DenseNet [23]	96.61	97.8	91
M et al. [26]	97.63	97.65	94.57
Ours	96.32	97.90	94.96

Bold indicates the largest value in this column of the Table

Table 6 Accuracy comparison on ALOP license plates recognition of different bounding boxes(%)

Bounding boxes	AC	LE	RP
Yolo3 detection	96.32	97.90	94.96
Ground truth	99.27	98.71	95.08

Bold indicates the largest value in this column of the Table

with the region code. The result indicates that the generalization ability of our model is reliable, and the license plate recognition model owns robustness.

5.3 Experiments on AOLP dataset

Table 5 proves that the accuracy of our proposed license plate recognition algorithm has been improved in LE and RP subsets. Compared to the second best results, the overall license plate recognition accuracy is improved by 0.25% in LE subset. For the RP subset, the recognition accuracy of the overall license plate is enhanced by 0.39%.

6 Conclusion

In this paper, we proposed a two-stage license plate recognition algorithm based on YOLOv3 and ILPRNET. We use the ILPRNET network to recognize license plates detected by YOLOv3. The proposed algorithm contains feature extraction, character localization, character feature extraction, and character classification. Compared with other algorithms, the advantages of our algorithm mainly lie in the use of a U-shaped network structure which avoids character recognition errors caused by inaccurate character feature segmentation. Our method incorporates a spatial attention mechanism to

accurately localize the licence plate characters, obtain the 2D-Attention weights of each character, and then multiply them with the licence plate feature vector to calculate the feature vector of each character, which is finally classified by a classifier. Through extensive experiments, our results show that the proposed model performs well in complex or conventional scenarios. Some problems are spotted in our method compared to the state-of-the-art methods: low recognition rate for rotating number plates, fixed number of plate characters. Therefore, in future work, we will improve the algorithm by correcting for rotating license plates, improve the algorithm to recognize license plate characters of different lengths and we may even get better results by feeding the loss of character localization back to the network. As for dataset, we can expand the dataset with a GAN network to make a balanced distribution of characters in order to improve the accuracy of license plate recognition.

Acknowledgements This work was supported by 2017 Zhuhai introduces innovation and entrepreneurship team (ZH01110405170027P WC, No.ZH0405-1900-01PWC) and Cloud Service Platform and Applications based on “ZHUHAI No.1” Constellation & Remote Sensing Big Data.(ZDXK[2018]007), Key Supported Disciplines of Guizhou Province-Computer Application Technology (No. QianXueWeiHeZi ZDXK [2016]20), and the work was also supported by National Natural Science Foundation of China (61462013, 61661010)

References

1. Bochkovskiy, A., Wang, C.Y., Liao, H.Y.M.: Yolov4: Optimal speed and accuracy of object detection. Preprint [arXiv:2004.10934](https://arxiv.org/abs/2004.10934) (2020)
2. Bulan, O., Kozitsky, V., Ramesh, P., Shreve, M.: Segmentation-and annotation-free license plate recognition with deep localization and failure identification. *IEEE Trans. Intell. Transp. Syst.* **18**(9), 2351–2363 (2017)
3. Cheng, Z., Bai, F., Xu, Y., Zheng, G., Pu, S., Zhou, S.: Focusing attention: towards accurate text recognition in natural images. In: *IEEE International Conference on Computer Vision*, pp. 5076–5084 (2017)
4. Girshick, R.: Fast r-cnn. In: *IEEE International Conference on Computer Vision*, pp. 1440–1448 (2015)
5. Gonçalves, G.R., da Silva, S.P.G., Menotti, D., Schwartz, W.R.: Benchmark for license plate character segmentation. *J. Electron. Imaging* **25**(5), 053034 (2016)
6. Hsu, G.S., Chen, J.C., Chung, Y.Z.: Application-oriented license plate recognition. *IEEE Trans. Veh. Technol.* **62**(2), 552–561 (2012)
7. Lee, Y., Song, T., Ku, B., Jeon, S., Han, D.K., Ko, H.: License plate detection using local structure patterns. In: *IEEE International Conference on Advanced Video and Signal Based Surveillance*, pp. 574–579 (2010)
8. Li, H., Shen, C.: Reading car license plates using deep convolutional neural networks and lstms. Preprint [arXiv:1601.05610](https://arxiv.org/abs/1601.05610) (2016)
9. Li, H., Wang, P., 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 (2018)
10. Li, H., Wang, P., Shen, C., Zhang, G.: Show, attend and read: a simple and strong baseline for irregular text recognition. *AAAI Conf. Artif. Intell.* **33**, 8610–8617 (2019)
11. Liu, W., Anguelov, D., Erhan, D., Szegedy, C., Reed, S., Fu, C.Y., Berg, A.C.: Ssd: Single shot multibox detector. In: *European Conference on Computer Vision*, pp. 21–37 (2016)
12. Luo, C., Jin, L., Sun, Z.: A multi-object rectified attention network for scene text recognition. Preprint [arXiv:1901.03003](https://arxiv.org/abs/1901.03003) (2019)
13. Masood, S.Z., Shu, G., Dehghan, A., Ortiz, E.G.: License plate detection and recognition using deeply learned convolutional neural networks. Preprint [arXiv:1703.07330](https://arxiv.org/abs/1703.07330) (2017)
14. Redmon, J., Divvala, S., Girshick, R., Farhadi, A.: You only look once: unified, real-time object detection. In: *IEEE Conference on Computer Vision and Pattern Recognition*, pp. 779–788 (2016)
15. Redmon, J., Farhadi, A.: Yolo9000: better, faster, stronger. In: *IEEE Conference on Computer Vision and Pattern Recognition*, pp. 7263–7271 (2017)
16. Redmon, J., Farhadi, A.: Yolov3: an incremental improvement. Preprint [arXiv:1804.02767](https://arxiv.org/abs/1804.02767) (2018)
17. Ren, S., He, K., Girshick, R., Sun, J.: Faster r-cnn: towards real-time object detection with region proposal networks. Preprint [arXiv:1506.01497](https://arxiv.org/abs/1506.01497) (2015)
18. Shi, B., Yang, M., Wang, X., Lyu, P., Yao, C., Bai, X.: Aster: an attentional scene text recognizer with flexible rectification. *IEEE Trans. Pattern Anal. Mach. Intell.* **41**(9), 2035–2048 (2018)
19. Silva, S.M., Jung, C.R.: License plate detection and recognition in unconstrained scenarios. In: *European Conference on Computer Vision (ECCV)*, pp. 580–596 (2018)
20. Wang, R., Sang, N., Huang, R., Wang, Y.: License plate detection using gradient information and cascade detectors. *Optik* **125**(1), 186–190 (2014)
21. Wang, S.Z., Lee, H.J.: A cascade framework for a real-time statistical license plate recognition system. *IEEE Trans. Inf. For. Secur.* **2**(2), 267–282 (2007)
22. Wang, T., Zhu, Y., Jin, L., Luo, C., Chen, X., Wu, Y., Wang, Q., Cai, M.: Decoupled attention network for text recognition. *AAAI Conf. Artif. Intell.* **34**, 12216–12224 (2020)
23. Wu, C., Xu, S., Song, G., Zhang, S.: How many labeled license plates are needed? In: *Chinese Conference on Pattern Recognition and Computer Vision (PRCV)*, pp. 334–346 (2018)
24. Xu, Z., Yang, W., Meng, A., Lu, N., Huang, H., Ying, C., Huang, L.: Towards end-to-end license plate detection and recognition: a large dataset and baseline. In: *Proceedings of the European Conference on Computer Vision (ECCV)*, pp. 255–271 (2018)
25. Yao, Z., Yi, W.: License plate detection based on multistage information fusion. *Inf. Fusion* **18**, 78–85 (2014)
26. Yousef, M., Hussain, K.F., Mohammed, U.S.: Accurate, data-efficient, unconstrained text recognition with convolutional neural networks. *Pattern Recogn.* **108**, 107482 (2020)
27. Yu, S., Li, B., Zhang, Q., Liu, C., Meng, M.Q.H.: A novel license plate location method based on wavelet transform and emd analysis. *Pattern Recogn.* **48**(1), 114–125 (2015)
28. Zhang, K., Zhang, Z., Li, Z., Qiao, Y.: Joint face detection and alignment using multitask cascaded convolutional networks. *IEEE Signal Process. Lett.* **23**(10), 1499–1503 (2016)
29. Zhang, L., Wang, P., Li, H., Li, Z., Shen, C., Zhang, Y.: A robust attentional framework for license plate recognition in the wild. *IEEE Trans. Intell. Transp. Syst.* (2020)
30. Zherzdev, S., Gruzdev, A.: Lprnet: license plate recognition via deep neural networks. Preprint [arXiv:1806.10447](https://arxiv.org/abs/1806.10447) (2018)
31. Zou, Y., Zhang, Y., Yan, J., Jiang, X., Huang, T., Fan, H., Cui, Z.: A robust license plate recognition model based on bi-lstm. *IEEE Access* **8**, 211630–211641 (2020)

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.