

APPLIED RESEARCH

Research on Car License Plate Recognition Based on Improved YOLOv5m and LPRNet

SHAN LUO¹ AND JIHONG LIU²¹School of Electrical and Information Engineering, Panzhihua University, Panzhihua 617000, China²School of Health and Wellness, Panzhihua University, Panzhihua 617000, China

Corresponding author: Shan Luo (407237861@qq.com)

This work was supported in part by the Sichuan Scientific and Technological Achievements Transfer and Transformation Demonstration Project of China under Grant 2022ZHCG0055.

ABSTRACT The application of license plate recognition technology is becoming more and more extensive. In view of the current practical requirements for the recognition accuracy and real-time performance of license plate recognition system in complex scenes, the existing target detection methods and license plate recognition methods are studied, and a car license plate recognition method based on improved YOLOv5m and LPRNet model is proposed. On the basis of studying the YOLOv5m algorithm and the image features of the car license plate, the YOLOv5m algorithm is improved from three aspects: the K-means++ algorithm is used to improve the matching degree between the anchor frame and the detection target, the DIOU loss function is used to improve the NMS method, and the feature map with 20×20 is removed to reduce the number of detection layers. A lightweight LPRNet network is used to realize license plate character recognition without character segmentation. Combining the improved YOLOv5m algorithm with LPRNet network, a license plate recognition system based on IYOLOv5m-LPRNet model is designed. The experimental results show that the average recognition accuracy of license plates in front, tilt, night and strong light interference scenes is more than 98%; Compared with the models of YOLOv3-LPRNet, YOLOv4-LPRNet, YOLOv5s-LPRNet and YOLOv5m-LPRNet, the recognition accuracy and recall rate of this method are improved, reaching 99.49% and 98.79% respectively; The mAP of this method is also the highest, reaching 98.56%; In terms of recognition speed, this method is also faster than the other four methods, and the number of pictures processed per second is increased by 5 compared with the YOLOv5m-LPRNet model. Therefore, the improved license plate recognition method in this paper performs well in robustness and speed.

INDEX TERMS License plate recognition, license plate detection, deep learning, YOLOv5m, LPRNet.

I. INTRODUCTION

With the continuous advancement of smart city construction, intelligent transportation system has developed rapidly, and license plate recognition system is a necessary part of intelligent transportation system, which plays an important practical role in improving traffic efficiency, traffic law enforcement and traffic safety. How to improve the accuracy and real-time performance of the license plate recognition system in a variety of complex scenes is a key problem to be solved by relevant scholars. Many scholars have carried out in-depth research on license plate recognition.

The associate editor coordinating the review of this manuscript and approving it for publication was Wei Liu.

At present, there are two main types of license plate recognition technologies. One is to use traditional image processing technology, and the other is to use deep learning methods. Deep learning methods are more robust than traditional methods and have been widely studied [1], [2]. Shi and Zhang [3] proposed to use BGRU to optimize the license plate recognition network model, and combined with the improved YOLOv3 network to locate the license plate. This method has good robustness. In order to solve the problem of fuzzy license plate character recognition, Zhang *et al.* [4] proposed an license plate character recognition algorithm without character segmentation based on improved CRNN+CTC(Convolutional Recurrent Neural Network+Connectionist Temporal Classification).

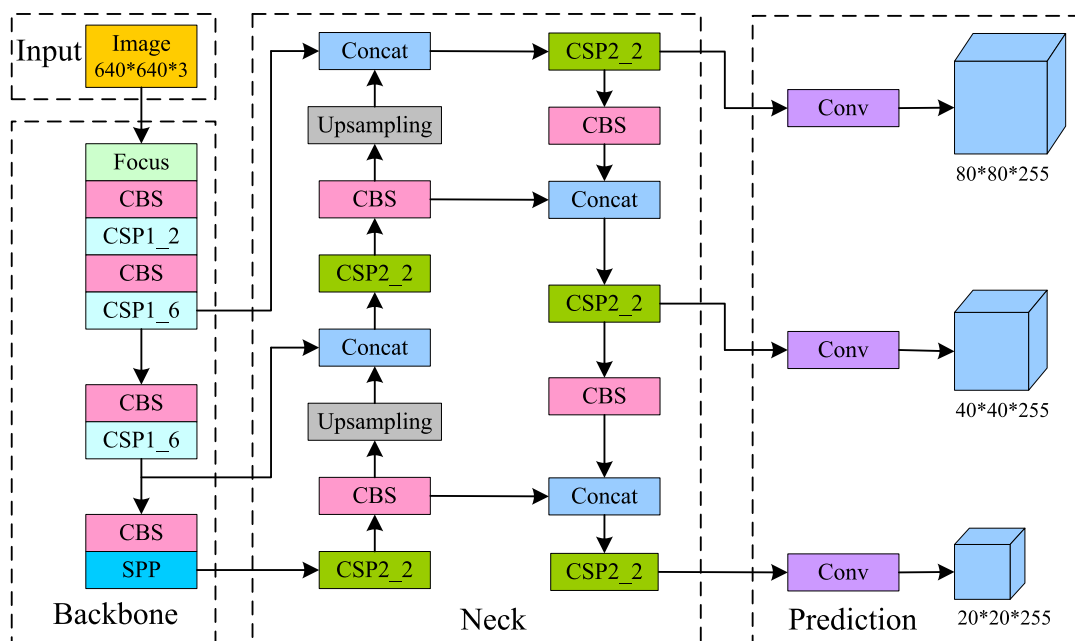


FIGURE 1. Structure of YOLOv5m network.

The recognition accuracy and speed of the algorithm have been improved to a certain extent. Li *et al.* [5] used the Faster R-CNN algorithm to locate the license plate, used the K-means++ algorithm to select the best license plate area size, and used the improved AlexNet model to achieve high accuracy license plate recognition, avoiding the impact on license plate character recognition due to the inability to accurately segment the license plate characters in the license plate recognition method based on license plate character segmentation. Li *et al.* [6] used deep neural network to realize end-to-end license plate detection and recognition, and trained the whole network end-to-end, which not only avoided the accumulation of intermediate errors, but also accelerated the processing speed. Silva and Jung [7] proposed an end-to-end adaptive license plate recognition method based on hierarchical convolutional neural network. Its core idea is to use the two passes of the same convolutional neural network to identify vehicles and license plate areas, and then use the second convolutional neural network for character recognition. Kim *et al.* [8] proposed a license plate recognition method based on deep convolution neural network on the basis of multi task learning. The network has a good generalization error rate. The research shows that this method is more accurate than using the traditional deep convolution neural network to classify numbers and characters. Khare *et al.* [9] introduced a novel method called partial character reconstruction to segment license plate characters, developed an automatic license plate recognition system that can cope with many factors, and enhanced the performance of the license plate recognition system. In order to solve the problem of license plate location in the low resolution multi vehicle environment, Zhu *et al.* [10] proposed a deep learning license plate recognition method based on improved YOLOv5s.

The algorithm has good robustness and fast operation speed. Fu and Qiu [11] used the improved YOLOv3 network structure to identify the characters of the license plate. The network is composed of seven full connection layers to predict the seven characters of the license plate, and each full connection layer accurately predicts the position and category of one character.

The existing license plate recognition methods have some shortcomings, such as low recognition accuracy and poor real-time performance, and the application of YOLOv5 algorithm in license plate recognition is rare. Therefore, this paper studies the application of YOLOv5m algorithm in license plate recognition, improves YOLOv5m algorithm according to the characteristics of license plate objects, and proposes a license plate recognition method based on improved YOLOv5m and LPRNet combined with LPRNet license plate recognition network. Taking the blue license plate of the car as the object, the recognition test is carried out, and the recognition performance is quantitatively evaluated.

II. IMPROVEMENT OF YOLOv5m ALGORITHM

A. INTRODUCTION TO YOLOv5m ALGORITHM

YOLOv5m is one of the models of YOLOv5 [12], which is characterized in that the depth of each CSP module is CSP1_2, CSP1_6, CSP1_6, CSP2_2, CSP2_2, CSP2_2, CSP2_2, CSP2_2. In Focus and CBS_1, CBS_2, CBS_3, CBS_4 the model widths (i.e. the number of convolution kernels) of each stage are 48, 96, 192, 384 and 768 respectively. YOLOv5m has more depth and width than YOLOv5s. The structure is slightly complex, but the detection accuracy is higher. YOLOv5m network is composed of four parts: Input, Backbone, Neck and Prediction. Its structural framework is shown in Figure 1.

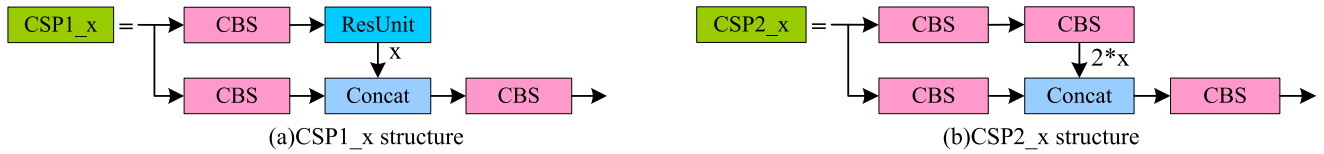


FIGURE 2. Structure of CSP.

Mosaic method is used in Input to realize data enhancement, adaptive anchor box calculation and adaptive image scaling. The idea of Mosaic method is to randomly select several pictures and splice them on one picture for data training, and then flip the picture and change the color gamut. Its advantage is that more small targets are added through random scaling, which improves the robustness.

Backbone implements image feature extraction, including Focus, CBS, CSP1_x and Spp four modules. The key function of Focus is to slice the input image, which makes the feature image smaller, reduces the number of layers and parameters, and improves the convolution operation speed. CBS is composed of convolution layer(Convolution), batch normalization layer(BN) and activation function(SiLU). Its function is to perform convolution, normalization and activation operations on slice images. Cross Stage Partial(CSP) network has CSP1_x and CSP2_x two structures, where x represents the number of residual components. CSP1_x with residual structure is used in feature extraction to optimize the gradient in the network, so that the gradient value is enhanced when the layer and layer back-propagation, so that the gradient disappearance caused by the deepening of the network can be effectively avoided. The composition of CSP is shown in Figure 2. Spatial Pyramid Pooling(SPP) [13] is composed of CBS and MaxPool, as shown in Figure 3. The most important function of the SPP module is to expand the receptive field. The SPP uses convolution kernel of different sizes to input characteristic maps of different sizes for maximum pooling, and then splice different results with the data without pool operation, so that the dimension of the output characteristic vector is the same.

Neck part realizes multi-scale feature information fusion. It adopts the structure of Feature Pyramid Network(FPN) + Pyramid Attention Network(PAN) [14], which is composed of several bottom-up paths and several top-down paths. Its structure is shown in Figure 4. The feature map is up sampled by FPN from top to bottom, and the extracted features are fused with those extracted from the backbone network; PAN is used to down sample the feature map from bottom to top, and the extracted features are fused with the features extracted from FPN layer. Through FPN+PAN network structure, the features extracted from the backbone network and detection network can be aggregated, and the feature fusion ability of the network can be improved. The extracted features are fused through CSP2_2 module. Concat is a channel splicing module to realize the combination of image features.

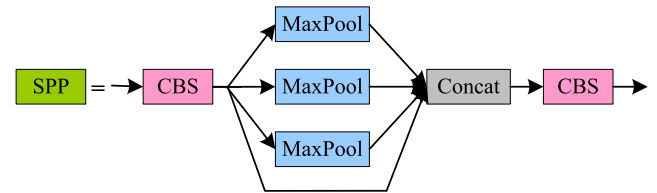


FIGURE 3. Structure of SPP.

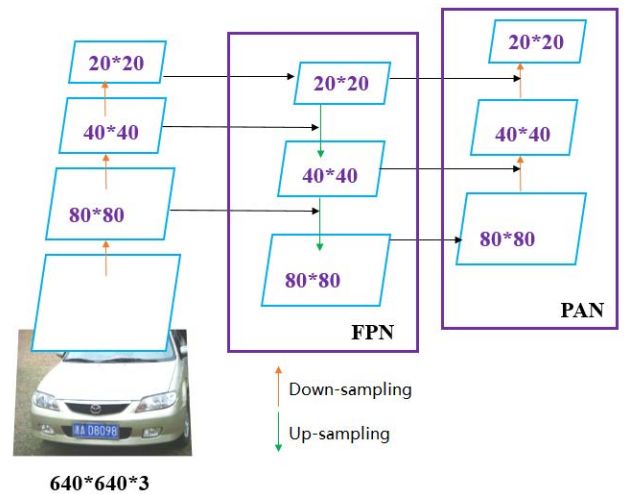


FIGURE 4. Schematic diagram of Neck structure.

Prediction section predicts the input image features, outputs the boundary box of the target category and target position with the highest confidence score, and provides three different detection scales(80 × 80, 40 × 40, 20 × 20). There are three commonly used bounding box loss functions, namely, classification loss function, confidence loss function and location loss function. The real target box is screened out by using NonMaximum Suppression(NMS) to eliminate redundant prediction target boxes, which improves the detection accuracy of the network.

B. IMPROVEMENT OF YOLOv5m ALGORITHM

1) IMPROVE THE MATCHING DEGREE BETWEEN ANCHOR FRAME AND DETECTION TARGET

K-means algorithm is a random allocation of initial clustering centers, which is not suitable for clustering license plate data sets. Therefore, in order to improve the accuracy of license plate detection, K-means++ algorithm [15] is used for multidimensional clustering of label data sets, which can effectively reduce the time for the model to find the anchor

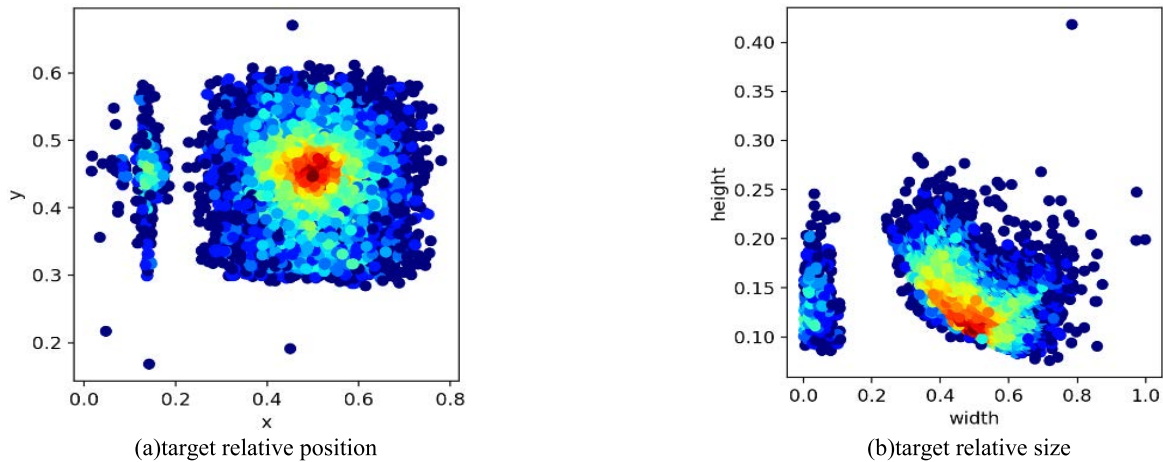


FIGURE 5. Relative position and size of target.

box and the amount of calculation in the target box matching process. In order to make the anchor frame and detection frame have a large intersection, so as to select the best a priori frame, the expression of the algorithm is:

$$d = 1 - IOU \quad (1)$$

where, IOU represents the intersection over union of the prediction frame and the real frame.

YOLOv5m network will adaptively calculate the corresponding a priori anchor box values according to different data sets, and its initial a priori anchor boxes are (10,13), (16,30), (33,23), (30,61), (62,45), (59,119), (116,90), (156,198) and (373,326). The prior anchor boxes obtained by K-means++ algorithm clustering are (12,16), (17,39), (30,52), (54,60), (33,26), (126,183), (227,283), (373,326) and (407,486) respectively. The principle of allocation is to allocate large prior anchor boxes to small targets and small prior anchor boxes to large targets. The map of the relative position and size of the target is obtained by analyzing the data on the target position and size label, as shown in Figure 5.

Figure 5(a) shows the relative position of the target. According to it, the relative position of the anchor frame in the image can be found, that is, it is concentrated between 0.4-0.6 along the X axis and between 0.4-0.5 along the Y axis. Figure 5(b) shows the relative size of the target. It can be seen that the target width accounts for 40%-60% of the image width, and the target height accounts for 10% -15% of the image height.

2) IMPROVED NMS METHOD

In view of the problem that NMS is prone to miss detection when screening the real target frame, the idea of DIOU loss function is used for reference to improve the post-processing method of NMS, and the DIOU-NMS method is obtained. DIOU loss function is:

$$R_{DIOU} = \frac{\rho^2(b, b^{st})}{l^2} \quad (2)$$

where, b is the predicted target box, b^{st} is the real target box, $\rho^2(b, b^{st})$ is the distance between the center point of the predicted target box and the real target box, and l is the distance between the diagonals of the minimum circumscribed rectangle of the two boxes.

Assuming that the network model detects a candidate box set as H_i , for the prediction target box M with the highest category confidence, the p_i of DIOU-NMS update formula is defined as:

$$p_i = \begin{cases} p_i, & IOU - R_{DIOU}(M, H_i) < \varepsilon \\ 0, & IOU - R_{DIOU}(M, H_i) \geq \varepsilon \end{cases} \quad (3)$$

where, i is the number of anchor boxes corresponding to each grid, p_i is the classification score of different category targets, IOU is the intersection and union ratio, and $R_{DIOU}(M, H_i)$ is the value of R_{DIOU} about M and H_i , ε is manually set threshold for NMS operation.

DIOU-NMS method takes into account the distance, overlapping area and aspect ratio between the predicted target frame and the real target frame. The farther the distance between the center points of the two rectangular frames, it is determined that they may be located on different detection objects. Combining the IOU of the two rectangular boxes with the distance between the center point, on the one hand, optimizes the IOU loss, on the other hand, guides the learning of the center point, and can more accurately return to the prediction target box [16].

3) REDUCE THE NUMBER OF DETECTION LAYERS

There are three detection layer characteristic maps with different scales in YOLOv5m network, which are 80×80 , 40×40 , 20×20 . The receptive field of 80×80 detection scale is 8×8 pixels region, suitable for detecting small targets; And the receptive field of 40×40 detection scale is 16×16 pixel region, mainly used to detect medium-sized targets; For 20×20 detection scale, its receptive field is 32×32 pixel region, used to detect large targets.

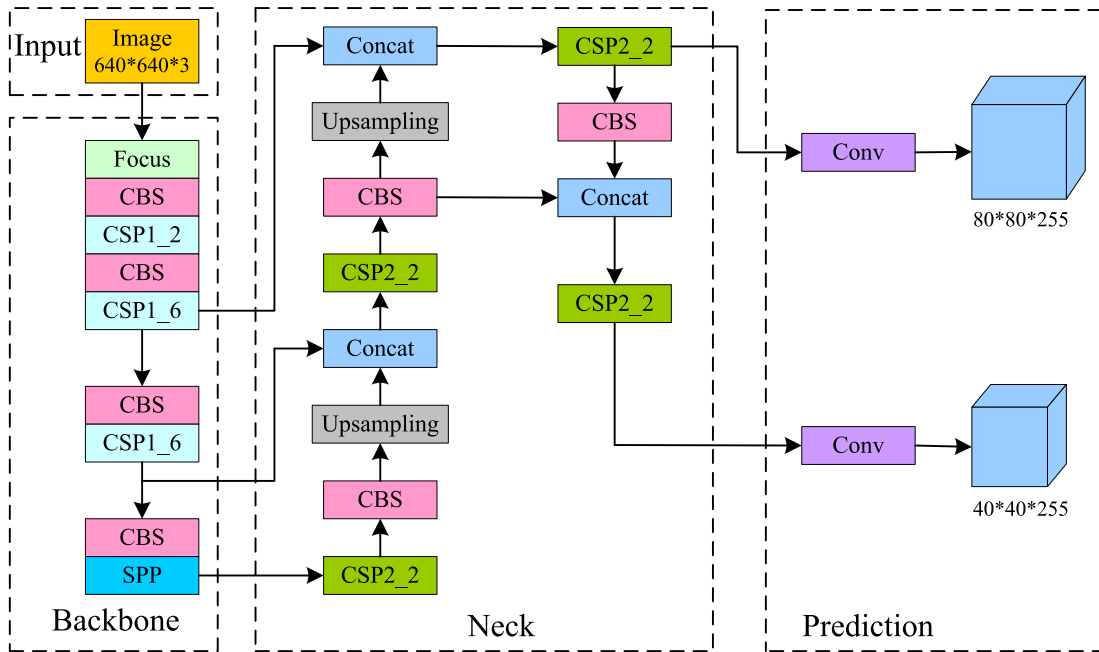


FIGURE 6. Structural diagram of IYOLOv5m.

In the traffic scene, the license plate belongs to a small target, so it is necessary to reduce the number of detection layers, that is, removing feature map of the scale with 20×20 , reserving feature map of the scale with 80×80 and 40×40 . The improved network not only meets the needs of license plate detection, but also reduces the complexity of the network and improves the running speed of the network and the speed of target detection.

The structure of the improved YOLOv5m(IYOLOv5m) model is shown in Figure 6.

III. LPRNet LICENSE PLATE RECOGNITION NETWORK

LPRNet [17] adopts a lightweight convolutional neural network structure, which can be used in license plate recognition without the need for license plate character segmentation like the traditional algorithm, and can realize end-to-end license plate character recognition with good robustness. The backbone network of LPRNet has three convolution layers, three maximum pooling layers, three basic modules, and two Dropout layers set to prevent over fitting. The network input is 94×24 image, the output layer is a convolution layer, and its structure is shown in Table 1. Each basic module contains four convolution layers, one input layer and one feature output layer. Its structure is shown in Table 2.

LPRNet uses the backbone network to extract image features, and then uses convolution kernel to convolute to obtain the license plate character sequence. However, since the LPRNet decoder output is different from the length of the character sequence, the CTC(Connectist Temporary Classification) loss function with variable length coding is used to avoid errors. LPRNet uses the beam search filter to obtain

TABLE 1. LPRNet network backbone.

Layer Type	Parameter
Input	94*24 pixels RGB image
Convolution	#64, 3*3, stride:1
Max Pooling	#64, 3*3, stride:1
Small basic block	#128, 3*3, stride:1
Max Pooling	#64, 3*3, stride:1
Small basic block	#256, 3*3, stride:1
Small basic block	#256, 3*3, stride:1
Max Pooling	#64, 3*3, stride:2*1
Dropout	0.5
Convolution	#256, 4*1, stride:1
Dropout	0.5
Convolution	#class_number, 1*13, stride:1

TABLE 2. Small basic block structure.

Layer Type	Parameter
Input	$C_{in} * H * W$ feature map
Convolution	# $C_{out}/4$, $1 * 1$, stride:1
Convolution	# $C_{out}/4$, $3 * 1$, stride:h:1, pad:h:1
Convolution	# $C_{out}/4$, $1 * 3$, stride:w:1, pad:w:1
Convolution	# C_{out} , $1 * 1$, stride:1
Output	$C_{out} * H * W$ feature map

the first n sequences with the highest probability as output, and returns the first pre-defined template set successfully

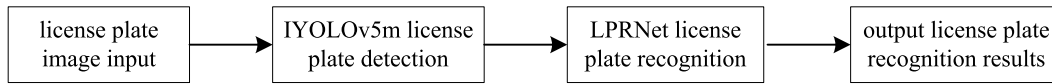


FIGURE 7. License plate recognition system process.

matched according to the national standard for motor vehicle license plates.

IV. YOLOv5m-LPRNet LICENSE PLATE RECOGNITION MODEL

Based on the above research, in view of the strong target detection ability and fast detection speed of YOLOv5m, combined with the ability of lightweight LPRNet network to realize end-to-end license plate character recognition, a car license plate recognition system based on YOLOv5m-LPRNet model is designed, as shown in Figure 7. Compared with the traditional method, this model does not need to segment the license plate characters, avoids the impact of character segmentation effect on character recognition, and improves the adaptive ability of the algorithm.

V. EXPERIMENTAL RESULTS AND ANALYSIS

A. LICENSE PLATE DATA SET

CCPD(Chinese City Parking Dataset) [18], an open-source data set of the Chinese Academy of Sciences, is composed of Chinese blue license plates. There are about 250000 independent license plate pictures, taking into account various complex traffic environments, such as different lighting environments, diversity of shooting angles, different time difference, etc. Taking into account the accuracy and operation speed, 21445 license plate pictures under different light intensity, different camera angles and different time scenes are selected for the experiment. The K-fold cross validation method is used to train the model. The license plate data set is divided into five groups, with 4289 pictures in each group. The validation set is made for each group of data respectively, and the remaining four groups of data are used as the training set, that is, the training set contains 17156 pictures.

The picture file name of the CCPD data set is the annotation information, for example, the file name of the license plate picture is “0093-2_4-326&446_470&500-470&493_332&500_326&453_464&446-10_11_26_23_24_26_25-124-72.jpg”, represents the annotation information of the picture. The file name is divided into 7 segments, including the region, horizontal and vertical tilt angles, annotation box vertex coordinates, four vertex coordinates of license plate, license plate number and other information. For example, “326&446_470&500” in paragraph 3 indicates that the vertex coordinates of the upper left corner and the lower right corner of the callout box are (326,446) and (470,500) respectively, and “10” in paragraph 5 “10_11_26_23_24_26_25” indicates “<Jiangsu>”, “11” indicates “M”, and “26_23_24_26_25” respectively indicates: 2, Z, 0, 2 and 1, that is, the license plate number is “<Jiangsu>M2Z021”.

B. EXPERIMENT CONFIGURATION AND PARAMETER SETTING

Experimental configuration: Intel CPU®Core™I7-108750 @2.60GHz×8, The GPU is GeForce GTX 3060 12GB, the RAM size is 16GB, the operating system is Windows 10, the development environment is PyCharm 2021, the framework is Pytorch 1.7, the development language is Python 3.8, CUDA 11.6, CuDNN 7.6.

Superparameter Settings: the optimizer is Adam, the initial learning rate is 0.01, the learning rate attenuation coefficient is 0.1, batch_size is 32 and the weight attenuation coefficient is 0.0005. Because the larger the epoch value is, the more stable the training model is, the higher the accuracy is and the faster the convergence is. Therefore, the epoch is determined to be 300 through experiments.

C. PERFORMANCE EVALUATION INDEX

In order to effectively evaluate the robustness of the model and the accuracy of license plate recognition, the performance of the model is evaluated by five indicators: *Precision*, *Recall*, F_{score} , mean of average precision(*mAP*) and *FPS*. F_{score} comprehensively reflects the performance of the model. *FPS* represents the number of pictures processed per second, the larger the value, the faster the operation speed. The closer *mAP* is to 1, the better the overall performance of the model. The formula for defining the evaluation index by using the confusion matrix is [19]:

$$Precision = \frac{TP}{TP + FP} \times 100\% \quad (4)$$

$$Recall = \frac{TP}{TP + FN} \times 100\% \quad (5)$$

$$F_{score} = \frac{2 \times Precision \times Recall}{Precision + Recall} \quad (6)$$

$$mAP = \frac{\sum_{i=1}^N \int_0^1 P_i(R) dR}{N} \quad (7)$$

$$FPS = \frac{Figure\ Number}{Total\ Time} \quad (8)$$

where, TP is the number of samples that are both actual and predicted to be correct, FP is the number of samples that are actually negative and predicted to be positive, FN is the number of samples that are actually positive and predicted to be negative, N is the number of sample categories, $AP_i = \int_0^1 P_i(R) dR$ is the average accuracy of category i , *FigureNumber* is the total number of processed pictures, and *TotalTime* is the total processing time.



(a)front license plate recognition results



(b)tilting license plate recognition results



(c)night license plate recognition results



(d)dark light license plate recognition results

FIGURE 8. License plate recognition test results.

D. IDENTIFICATION RESULTS AND COMPARATIVE ANALYSIS

The license plate recognition model in Figure 7 is used to test the license plate recognition of 6000 pictures with different shooting angles and different light intensities. These 6000 pictures are divided into four categories, including

each 1500 pictures collected in front, tilt, night and dark scenes. This can test the performance of the license plate recognition model in different scenes. Part of the recognition results are shown in Figure 8. The purplish red rectangular box is the detection box, and the white characters above the detection box are the recognition results, that is,

TABLE 3. Index values of license plate recognition.

Scene Category	Precision(%)	Recall(%)	F _{score}	mAP(%)	FPS(figure/s)
front	99.84	99.86	0.9985	99.34	34
tilt	98.45	98.23	0.9834	97.86	32
night	98.12	97.16	0.9764	97.48	31
dark	97.56	97.12	0.9734	98.35	31
average value	98.49	98.09	0.9829	98.26	32

TABLE 4. Performance comparison of different license plate recognition models.

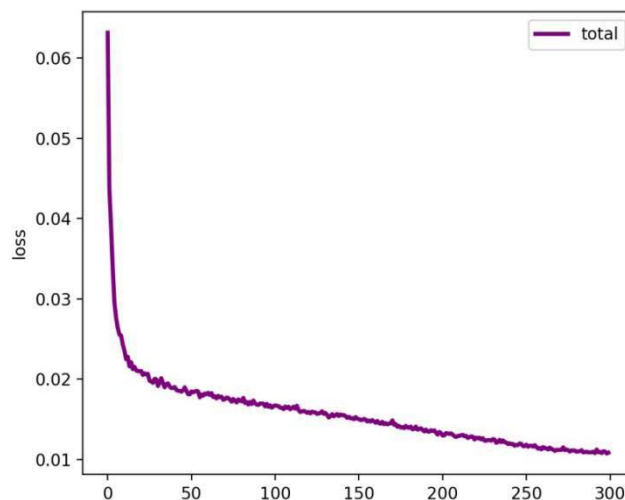
Model	Precision(%)	Recall(%)	F _{score}	mAP(%)	FPS(figure/s)
YOLOv3-LPRNet	87.46	82.56	0.8494	88.31	30
YOLOv4-LPRNet	89.23	84.31	0.8670	89.21	34
YOLOv5s-LPRNet	93.36	94.86	0.9410	93.51	40
YOLOv5m-LPRNet	94.49	95.37	0.9493	94.59	37
ours	99.49	98.79	0.9914	98.56	42

the output license plate number. The recognition results in Figure 8(a) are “<Anhui>A37095”, “<Jiangsu>E5NF78”, “<Anhui>AX8J13” and “<Anhui>AHF777” respectively, the recognition results in Figure 8(b) are “<Anhui>ANN923”, “<Anhui>AF5Z25”, “<Anhui>AN4526” and “<Anhui>ACL042” respectively, the recognition results in Figure 8(c) are “<Anhui>A71U52”, “<Anhui>HW896T”, “<Anhui>AHE328” and “<Anhui>APB097” respectively, and the recognition results in Figure 8(d) are “<Anhui>AP077L”, “<Anhui>ATJ063”, “<Anhui>A002R6” and “<Jiangsu>M2S966” respectively.

It can be seen from Figure 8 that the recognition results are basically correct for license plates in front, tilt, night and strong light interference. Figure 9 shows the change curve of the loss value of the improved model during the training process. It can be seen that the loss value drops steadily. When epoch=300, the global loss rate drops to about 1.5%, indicating that the convergence of the improved model is more stable.

In order to verify the overall performance of the license plate recognition model in this paper, use equations (4) - (8) to calculate the index values, as shown in Table 3.

According to the analysis in Table 3, the Precision of front license plate, inclined license plate and night license plate is higher than 98%, the Precision of front license plate is as high as 99.84%, only the Precision under dark light conditions is close to 98%, and the average value of Precision is more than 98%. For Recall, the front license plate and inclined license plate are higher than 98%, the front license plate is as high as 99.86%, and the Recall of night and dark light license plates is slightly higher than 97%, with an average value of more than 98%. The average value of F_{score} is 0.9829, which is close to 1. The mAP of the front license plate is as high as 99.34%, and the average value is 98.26%, indicating that the model has high accuracy. The average value of FPS is 32, that is, it processes 32 pictures per second, which is faster.

**FIGURE 9.** Loss function curve.

The model in this paper is compared with YOLOv3-LPRNet, YOLOv4-LPRNet, YOLOv5s-LPRNet and YOLOv5m-LPRNet models to conduct license plate recognition test under the same experimental conditions to verify the effectiveness of the improved model in this paper. The results are shown in Table 4.

It can be seen from Table 4 that the Precision of the method in this paper reaches 99.49%, which is 13.75%, 11.5%, 6.57% and 5.29% higher than the Precision of the YOLOv3-LPRNet, YOLOv4-LPRNet, YOLOv5s-LPRNet and YOLOv5m-LPRNet models respectively; Recall reaches 98.79%, which is 19.66%, 17.17%, 4.14% and 3.59% higher than the Recall of YOLOv3-LPRNet, YOLOv4-LPRNet, YOLOv5s-LPRNet and YOLOv5m-LPRNet models respectively. F_{score} takes into account both Precision and Recall, which comprehensively reflects the performance of the recognition model. The value of the model in this paper

is as high as 0.9914, indicating that the recognition effect of the model in this paper is the best, showing excellent recognition performance. The mAP of this method is the highest, which is 11.61%, 10.48%, 5.4% and 4.2% higher than the mAP of the YOLOv3-LPRNet, YOLOv4-LPRNet, YOLOv5s-LPRNet and YOLOv5m-LPRNet models respectively. In terms of recognition speed, the improved method in this paper is also faster than the other four methods, and the number of pictures processed per second is increased by 5 compared with the YOLOv5m-LPRNet model. Overall, the Precision, Recall, F_{score} , mAP and FPS indexes of this method are higher than those of YOLOv5m-LPRNet. Therefore, the improved license plate recognition model in this paper shows excellent adaptability and robustness, and the recognition speed is also fast.

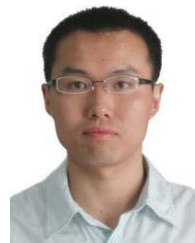
VI. CONCLUSION

The YOLOv5m algorithm is studied, and the YOLOv5m algorithm is improved from three aspects in combination with the characteristics of the license plate detection object. The improved license plate detection algorithm has faster operation speed and accurate positioning. The lightweight LPRNet network is used to realize license plate character recognition. The network has strong adaptability to complex environment and good robustness. Combining the improved YOLOv5m algorithm with LPRNet network, a car license plate recognition system based on YOLOv5m-LPRNet model is designed, and the system simulation model is established. The model is used to train and verify a large number of car license plate samples, which verifies the effectiveness of the model. The experiment on the car license plate recognition model shows that the recognition accuracy, recall and mAP of the model are improved, reaching 99.49%, 98.79% and 98.56% respectively; The recognition speed reaches 42 pictures per second. The improved method has high recognition accuracy, fast speed, excellent robustness and real-time performance, and can basically meet the needs of car license plate recognition in complex scenes.

REFERENCES

- [1] J. Y. Tu, "Research on license plate recognition algorithm based on deep learning in complex scenes," M.S. thesis, School Comput., Zhejiang Univ. Technol., Hangzhou, China, 2020.
- [2] N. Duan, J. J. Cui, L. Z. Liu, and L. R. Zheng, "An end to end recognition for license plates using convolutional neural networks," *IEEE Intell. Transp. Syst. Mag.*, vol. 13, no. 2, pp. 177–188, Summer 2021.
- [3] J. W. Shi and Y. Zhang, "License plate recognition system based on improved YOLOv3 and BGRU," *Comput. Eng. Des.*, vol. 41, no. 8, pp. 2345–2351, 2020.
- [4] C. Z. Zhang, Y. Li, B. L. Kang, and Y. Chang, "Blurred license plate character recognition algorithm based on deep learning," *Laser Optoelectron. Prog.*, vol. 58, no. 16, pp. 259–266, 2021.
- [5] X. P. Li, W. D. Min, Q. Han, and R. K. Liu, "License plate location and recognition based on deep learning," *J. Comput.-Aided Des. Comput. Graph.*, vol. 31, no. 6, pp. 979–987, 2019.
- [6] H. Li, P. Wang, and C. H. Shen, "Toward end-to-end car license plate detection and recognition with deep neural networks," *IEEE Trans. Intell. Transp. Syst.*, vol. 20, no. 3, pp. 1126–1136, Mar. 2019.
- [7] S. M. Silva and C. R. Jung, "Real-time license plate detection and recognition using deep convolutional neural networks," *J. Vis. Commun. Image Represent.*, vol. 71, Aug. 2020, Art. no. 102773.

- [8] H.-H. Kim, J.-K. Park, J.-H. Oh, and D.-J. Kang, "Multi-task convolutional neural network system for license plate recognition," *Int. J. Control Autom. Syst.*, vol. 15, no. 6, pp. 2942–2949, Dec. 2017.
- [9] V. Khare, P. Shivakumara, C. S. Chan, T. Lu, L. K. Meng, H. H. Woon, and M. Blumenstein, "A novel character segmentation-reconstruction approach for license plate recognition," *Expert Syst. Appl.*, vol. 131, pp. 219–239, Oct. 2019.
- [10] Q. Q. Zhu, Y. N. Liu, Z. H. Zhao, and X. Ma, "Research on license plate location algorithm based on YOLOv5," *J. Phys., Conf. Ser.*, vol. 2278, pp. 1458–1464, 2022, Art. no. 012040.
- [11] C. X. Fu and R. H. Qiu, "A license plate recognition system based on YOLOv3 algorithm," *Sci. Technol. Innov.*, vol. 3, no. 3, pp. 42–44, 2020.
- [12] X. Zhu, S. Lyu, X. Wang, and Q. Zhao, "TPH-YOLOv5: Improved YOLOv5 based on transformer prediction head for object detection on drone-captured scenarios," in *Proc. IEEE/CVF Int. Conf. Comput. Vis. Workshops (ICCVW)*, Oct. 2021, pp. 2778–2788.
- [13] K. He, X. Zhang, S. Ren, and J. Sun, "Spatial pyramid pooling in deep convolutional networks for visual recognition," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 37, no. 9, pp. 1904–1916, Jul. 2015.
- [14] S. Liu, L. Qi, H. F. Qin, J. P. Shi, and J. Y. Jia, "Path aggregation network for instance segmentation," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2018, pp. 8759–8768.
- [15] D. Arthur and S. Vassilvitskii, "K-means++: The advantages of careful seeding," Stanford Univ., Palo Alto, CA, USA, Tech. Rep. 4589, 2006.
- [16] J. H. Qiu, S. Y. Bei, M. F. Yin, and H. J. Qing, "Gear surface defect detection based on improved YOLOv5s," *Modern Manuf. Eng.*, vol. 3, no. 3, pp. 104–113, 2022.
- [17] S. Zherzdev and A. Gruzdev, "LPRNet: License plate recognition via deep neural networks," *Comput. Vis. Pattern Recognit.*, no. 6, pp. 248–253, 2018. [Online]. Available: <https://arxiv.org/abs/1806.10447v1>
- [18] Z. B. Xu, W. Yang, A. J. Meng, N. X. Lu, and L. S. Huang, "Towards end-to-end license plate detection and recognition: A large dataset and baseline," in *Proc. Eur. Conf. Comput. Vis.*, Sep. 2018, pp. 255–271.
- [19] Y. H. Wang, H. W. Ding, B. Li, Z. J. Yang, and J. D. Yang, "Mask wearing detection algorithm based on improved YOLOv3 in complex scenes," *Comput. Eng.*, vol. 46, no. 11, pp. 12–22, 2020.



SHAN LUO was born in Lezhi, Sichuan, China, in 1979. He received the B.S. degree in electronic information engineering from the Jiangxi University of Science and Technology, in 2004, and the M.S. degree in communication and information system from the Kunming University of Science and Technology, in 2009.

Since 2009, he has been a Teacher with the School of Electrical and Information Engineering, Panzhihua University, Panzhihua, China. He was the author of one book, more than 20 articles, and more than four inventions. His research interests include information processing, machine vision, and image processing.



JIHONG LIU was born in Lezhi, Sichuan, China, in 1982. He received the B.S. degree in nursing science from Southwest Medical University, in 2005. Since 2005, he has been a Teacher with the School of Health and Wellness, Panzhihua University, Panzhihua, China. His research interests include geriatric nursing and information processing.

• • •