Journal of Network Intelligence
Taiwan Ubiquitous Information

Volume **5**, Number **3**, August **2020**

A Novel License Plate Location Method Based on Deep Learning

Li-Sang Liu[1,2,3,4*], Dong-Wei He[1,2], Ying Ma[1,3], Xiu-Zhen Zhang[1],
Jing Huang[1,3], Jun-Nan Li[1], Jin-xin Yao[1,3]

[1]School of Information Science and Engineering,
Fujian University of Technology, Fujian, China

[2]FuJian Key Laboratory of A.E.D, Fujian University of Technology, Fujian, China

[3]National Demonstration Center for Experimental Electronic Information and
Electrical Technology Education, Fujian University of Technology, Fujian, China

[4]School of Computer Science and Electronic Engineering,
University of Essex, Wivenhoe Park, Colchester, UK

[*]Corresponding author: liulisang@fjut.edu.cn

Received August 2019; revised June 2020

Abstract. *The effectiveness of license plate location greatly affects the accuracy of license plate recognition technology. In view of the particularity of Chinese license plates, this paper applies the SSD algorithm to license plate location, and introduces the method of deconvolution fusion to optimize the SSD algorithm. Through the deep learning network training on the test data set, our method is proved to be the most efficient and most accurate model. Then we experimented again in our own dataset via the embedded vehicle system. The results verified that our method improves the accuracy of detecting small target license plates in complex environments such as dark, tilt, and multiple license plates, and has better accuracy, versatility and adaptability of license plate location.*
Keywords: License plate location, SSD, Deep learning, Small target detection, Embedded vehicle system

---

1. **Introduction.** Vehicle license plate recognition technology is an important part of urban intelligent transportation system management, which has wide application such as public security law enforcement, highway toll collection, urban road monitoring, intelligent parking lot management, traffic flow monitoring and so on. License plate recognition technology generally consists of three parts, vehicle license plate location, character segmentation and vehicle license plate recognition. Vehicle license plate location is, among which, the base of vehicle license plate recognition, and is also the most important factor of accuracy of license plate character segmentation and recognition results [1]. An accurate and effective license plate location can simplify character segmentation and recognition algorithm especially when in complex environment, such as rainy day, night, complex background, inclined angle and multi-license plates, reducing deployment costs and improving the applicability of license plate recognition technology.

Chinese license plate has its own special characteristics. There are mainly four color types of license plate, blue background with white characters, yellow background with black characters, white background with black characters, and black background with white characters. The background color of license plate accounts for about 70% of the license plate area, and contrasts sharply with the character color. Besides, vehicle license plate contains rich features of texture and edges due to the inclusion of Chinese characters[2].

According to the color and texture characteristics of Chinese license plates, traditional license plate location methods can be generally divided into two categories. One type is based on the gray-scale image processing of the license plate. These kind of typical algorithms include edge detection based location algorithm[3], mathematical morphology based location algorithm[4], wavelet transform based location algorithm[5], etc. The other type is based on the color characteristics of the license plate, for example, location algorithm based on the combination of color space conversion and morphology[6], or that based on color edge detection[7-8]. These traditional license plate location algorithms can achieve good detection accuracy under certain circumstances, however, their versatility and applicability are unable to be obtained because of uncertain factors of natural environment such as weather, illumination, complex background, shooting angle and distance. In recent years, the versatility of license plate location technology in the natural environment has been greatly improved as the license plate location and recognition algorithm are combined with convolutional neural network[9-12]. Moreover, license plate location and recognition can be extended to other application, for example, security analysis and improvement of image encryption scheme for some security consideration[13], or analysis of seismic site emergency rescue traffic path[14].

In this paper, SSD algorithm is applied to license plate location and is optimized according to the special characteristics of Chinese license plates. The inverse convolution fusion method is introduced to improve the ability of detection of small targets. The algorithm proposed in this paper is also applied to embedded vehicle system, and the testing results show that the method has a better location ability of license plate.

2. Proposed Method.

2.1. **SSD Algorithm.** The SSD(Single Shot MultiBox Detector) algorithm, proposed by Wei Liu in 2016 [15], is an end-to-end multi-frame detection depth neural network model, which combines the regression idea of YOLO and the candidate frame mechanism of Faster R-CNN. The regression idea of YOLO helps SSD algorithm to greatly reduce the computational complexity of the neural network and improve the speed of the algorithm. Candidate frame or region proposal does not need to be generated, while the feature map of the input image is extracted as well as the bounding boxes of the position are regressed on the feature map directly, then the objects are classified. The local feature extraction method is used to obtain the features of different positions, different aspect ratios and sizes, which is more efficient than the global feature extraction method of YOLO V2 for a certain position. In addition, in order to increase the robustness of the model to detect objects of different sizes, the SSD algorithm selects multiple levels of feature maps in the network for prediction.

2.2. **SSD Structure.** The SSD structure is based on VGG-16 network as shown in Fig.1. Subsequently, a convolutional feature map with progressively decreasing resolution is added. Each layer of the feature map has different receptive fields, and a multi-scale feature box can be obtained. The feature map can predict the object classification and the offset of the target frame. Finally, the final detection result is generated by Non-Maximum Suppression (NMS) to achieve multi-scale target detection. Compared with fast RCNN, the algorithm has much faster detection speed due to no region proposals generated, while compared with YOLO V2, the algorithm has better detection accuracy because of using multi-scale feature box to judge comprehensively.

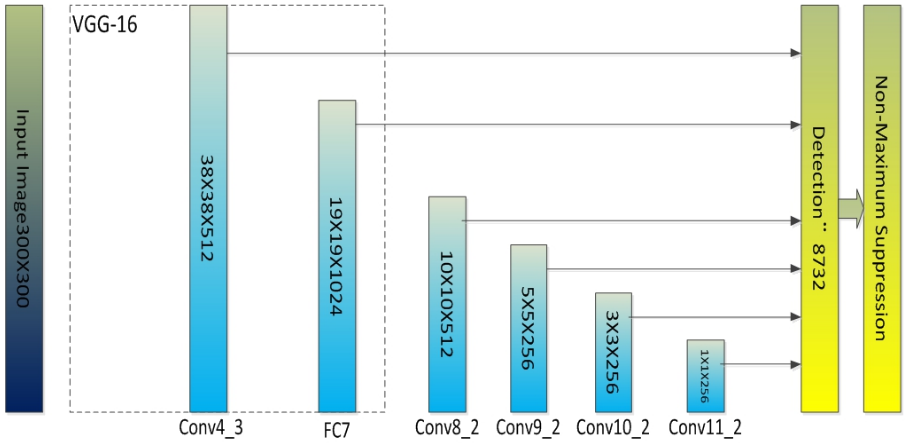


FIGURE 1. SSD structure based on VGG-16

2.3. **Shortages of SSD.** SSD has its own shortage because only when the IOU between the prior box and GroundTruth greater than 0.5 then can be put into the network for training during the training process. Larger goals are likely to have greater IOU values, as a result more prior boxes are included, and they can be adequately trained. On the contrary, smaller goals have fewer prior boxes to cover itself and few IOUs are unlikely to larger than 0.5, the training of small goals are not adequate either.

2.4. **Optimization of SSD.**

1. **Using MobilenetV2 in place of VGG-16.** Google released MobileNet V2 in April 2018 [16]. MobileNet V2 optimizes the customization for mobile embedded devices, reduces a large number of operands, maintains a certain level of detection accuracy, and greatly reduces the computing speed of embedded devices and main storage access requirements. MobileNet has built in a new set of network frameworks with two important technical components: depth-wise separable convolution [17-18] and inverse residual with bottleneck [19].

   Using depth-wise separable convolution can reduce the time complexity and space complexity of convolution layer exponentially. As can be seen from the following formula(1), since the size K of the convolution kernel is usually much smaller than the number of output channels C_{out}, the computational complexity of the standard convolution is approximately k^2 times that of the depth-wise separable convolution. MobileNetV2 sets k at 3(k=3)A so the computational cost is 8 to 9 times smaller than that of standard convolutions at only a small reduction in accuracy.

$$\text{Complexity}\frac{Depth - wiseSeparable\ CONV}{Standard\ CONV} = \frac{1}{K^2} + \frac{1}{C_{out}} \sim \frac{1}{K^2} \tag{1}$$

   The micro-structure of MobileNetV1AMobileNetV2 and ResNet are shown in Fig.2. The micro-structure of MobileNet V2 is called bottleneck inverse residual block[16]. It can be seen that the basic structure of MobileNet V2 and ResNet are similar. Both of them use shortcut to add input and output. However, ResNet first reduces dimensions (0.25 times), extracts features, and then increases dimensions. Visually, ResNet's microstructure is hourglass-shaped while MobileNetV2 is spindle-shaped, which first increases dimensions (6 times), extracts features, and then reduces dimensions. The bottlenecks encode the model's intermediate inputs and outputs while the inner layer encapsulates the model's ability to transform from lower-level concepts such as pixels to higher level descriptors such as image categories. Finally, as with traditional residual connections, shortcuts enable faster training and better accuracy. The nature of our networks allows us to utilize much smaller input and output dimensions. In Fig.2 we compare the needed sizes for each resolution between MobileNetV1, MobileNetV2 and ResNet.

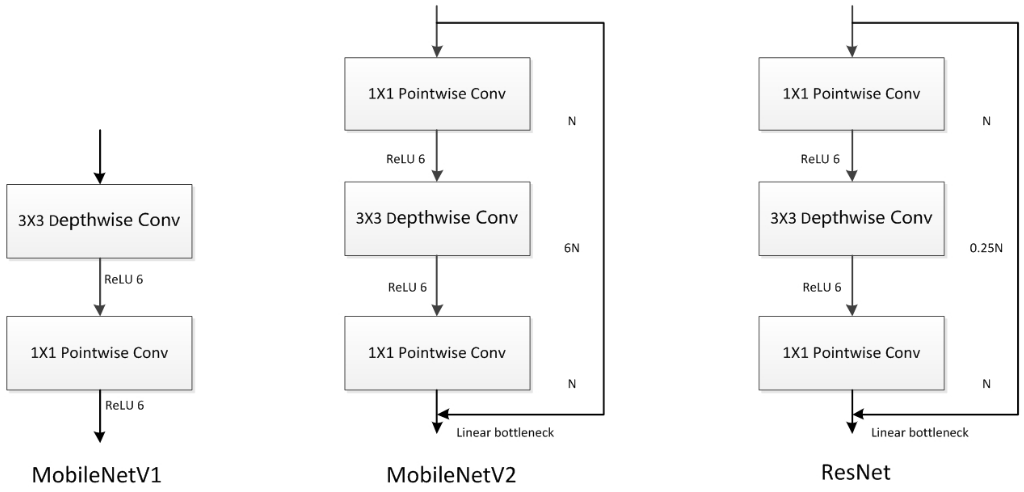


FIGURE 2. The micro-structure of MobileNetV1, MobileNetV2 and ResNet

2. Deconvolution

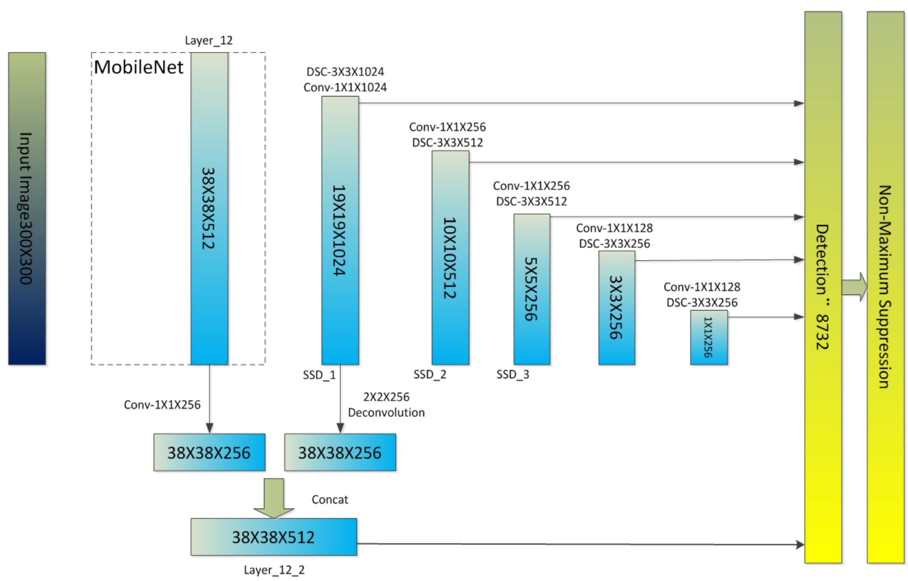


FIGURE 3. Proposed SSD structure based on mobilenetV2

The structure of SSD is similar to image pyramid. The high-level feature maps have large receptive field, low resolution, strong semantic information representation ability and weak geometric details representation ability. On the contrary, the low-level feature maps are just the opposite, having relatively small receptive fields, high resolution, strong representation ability of geometric details, but weak representation ability of semantic information. However, the detection of small targets depends on the low-level feature maps and does not have the high-level semantic information contained in the high-level feature maps. Therefore, the SSD detection of small targets is ineffective. In order to strengthen the detection of small targets, this paper introduces inverse convolution to deconvolute high-level feature maps, and fuse them with adjacent low-level feature maps to enrich the semantic information of low-level feature maps. Detection performance will be degraded, however, due to excessive deconvolution layers. Thus the proper number of deconvolution layers need to be carefully considered. Taking into account the detection accuracy and efficiency, this paper introduces the network framework of deconvolution layer as shown in Figure 3 through the experiments.

In the framework , we can see that the SSD_1 convolutional layer is deconvolved firstly to obtain the feature map with the same resolution as the Layer_12 layer, and then concat with the original Layer_12 layer feature map, so that the Layer_12 layer obtains better high-level semantics information. The information is supplemented and the fusion Layer_12_2 is formed to transfer to the full connection layer.

3. Experiment and Result.

3.1. Simulation experiments. We trained our model with the union of PASCAL VOC2007 and PASCAL VOC2012, and evaluate the results on PASCAL VOC2007 test. The algorithm training mentioned in this paper is divided into three steps: first, training MobileNet V2; secondly, adding fusion model and training; finally, fine-tuning the entire network.

In the first training stage, ImageNet is used to train MobileNet V2 as our pre-trained model. Batch size is selected as 16, the learning rate is10^{-3} in the first 80K iterations, then continues training for 30K iterations with 10^{-4} and 30K for 10^{-5} respectively. In the second stage of training, the fusion network layer is trained iteratively for 30K at a learning rate of 10^{-3}. Then the learning rate decreases to 10^{-4} for the next 25K iterations. Lastly, we fine-tune the entire network by using learning rate 10^{-3} for the first 60K iterations, and then continue training 30K with 10^{-4} and 30K with 10^{-5}, last 20K iterations with 10^{-6}. The results on PASCAL VOC2007 test detection are shown in Table 1. It shows the improving the detection performance of our model. The original SSD can get 77.5% mAP. Our model can reach 71.6% mAP., which is higher than MobileNetSSD by 4.1%.

Table 1. PASCAL VOC2007 TEST detection results

Method	mAP	Tv	Train	Sofa	Sheep	Plant	Person	Mbike	Horse	Dog	Table	Cow	Chair	Cat	Car	Bus	Bottle	Boat	Bird	Bike	aero
SSD300 [97]	77.5	76.8	87.6	79.5	77.9	52.3	79.4	84.0	87.5	86.1	77.0	81.5	60.3	88.1	85.7	87.0	50.5	69.6	76.0	83.9	79.5
Mobilenet SSD[96]	67.5	64.5	82.7	75.4	60.7	40.2	70.4	79.3	82.0	78.4	70.6	62.4	51.4	83.6	72.8	77.4	34.7	56.1	61.2	79.1	67.4
Our method	71.6	68.5	83.3	78.4	62.5	44.8	72.2	80.6	82.2	78.4	73.7	62.8	53.7	82.9	74.4	78.8	44.6	54.7	36.4	78.2	69.7

Table 2. The running time on 2016 COCO TEST

Method	Base NetWork	mAP	Madds	Params
SSD300[97]	VGG-16	21.2	34.9B	14.8M
MobileNetSSD	MobileNetV1	19.3	1.3B	6.8M
Our Method	MobileNetV2	22.4	0.9B	4.6M

The running time on 2016 COCO TEST is shown in Table 2, undertaking with batch size 1 using Titan X and cuDNN 6.0.21 with Intel Core(TM) i7-4790K@4.00GHz. Our Method is not only the most efficient model, but also the most accurate of the three.

3.2. **Testing experiments.** We built a set of 12151 marked training pictures, 5000 pure license plate pictures, 6451 pictures with license plate and complex background, including different shooting angles, distances, different weather conditions, different number of license plates and so on. In order to improve the generalization ability of the algorithm, 5000 pure license plate images were scaled to expand the data set. The ratio of the size of each sample image block to the original input image is between [0.1, 1], and the aspect ratio is between 1/2 and 2. If the center of the real frame is in the sampled image block, the overlap between the real frame and the image block is preserved. After sampling, each image block is set to a fixed size of 300 x 300 and flipped at a probability level of 0.5.

3.3. **Embedded Platform Transplantation.** This paper adopts Rockchip's RK1608 Pre-ISP chip, which integrates two CEMH-XM4 with 600MHz frequency, and provides up to 256*0.6GMIPS computing power. The built-in 256/128MB DDR and embedded VPU (Vector Process Unit) supports convolution acceleration with various precision. Target detection is more accurate and faster. The main AI application scenarios include face access control attendance, security monitoring, image and video auditing, face beauty, object recognition, all kinds of whiter appliances, and commercial display. Single core 7.5fp, dual core 11fps. The RK1608 dual-core, whose power consumption is around 1.5W, can support MobileNetV2_ssd300.

3.4. **Visual Experimental Results.** The plate location method based on color segmentation and Sobel operator [20] was compared in our experiments. The algorithm proposed in this paper can locate the license plate efficiently and quickly, overcoming the influence of background environment and shooting angle. When comparing under different tilt angles circumstance, the proposed algorithm not only can accurately locate the tilted and rotated license plates, but also can locate multiple license plates. When comparing in the dark and long-distance environment, the proposed algorithm still can locate quickly under the environment of poor illumination. These show our algorithm has good environmental robustness. While compared with the traditional SSD algorithm, the algorithm also has obvious advantages in small target license plate location.

3.4.1. *Comparison Testing with Tilt Angles.*

3.4.2. *Comparison Testing in Dark and Long-Distance Environment.*

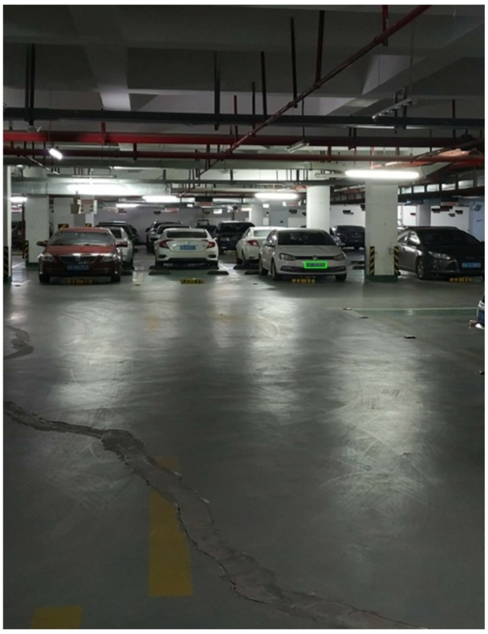
3.4.3. *Comparison Testing of Small Target Location.* Even the occluded small target, the red car's license plate obscured by a motorcycle, can still be identified, as shown in Fig.6(a).

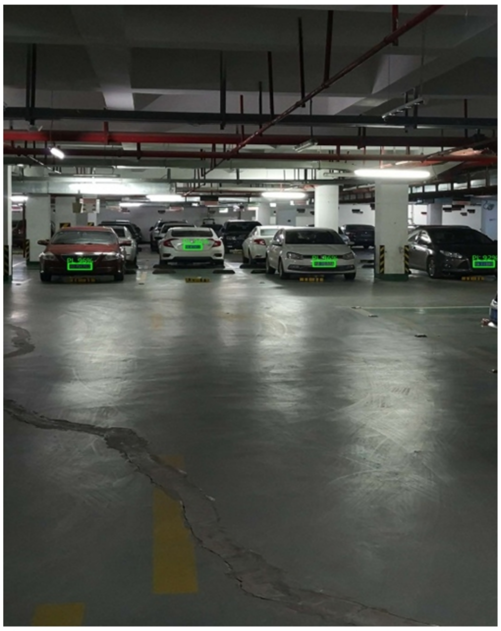
4. **Conclusion.** Based on SSD algorithm, this paper proposed a novel license plate location method on the basis of deep learning, optimizing the basic network and replacing the traditional VGG network with MobileNetV2. This algorithm can improve the efficiency of algorithm execution and is very suitable for embedded platforms. Beside, aiming to solve the problem of poor positioning ability of SSD small target license plate, we proposed an approach to improve the algorithm by providing the high-level semantic information for the low-level network. The testing shows our method can improve the ability of locating small target license plates and occluded license plates. It is applied in the actual embedded hand-held license plate detection equipment and can accurately locate multiple license plates in different natural environments with good environmental adaptability.

(a)

(b)

FIGURE 4. Tilt angles testing: (a)using our algorithm, (b)using HOU's method[18]





(a) (b)

FIGURE 5. Dark environment testing: (a)using our algorithm, (b)using HOU's method[18]

(a) (b)

Figure 6. Small targets testing: (a)using our algorithm, (b)using traditional SSD algorithm

Acknowledgment. The authors would like to thank the anonymous reviewers for their valuable comments. This work is supported by Initial Scientific Research Fund of FJUT(GY-Z12079), Fujian Provincial Education Department Youth Fund(JAT170367,JAT170369), Natural Science Foundation of Fujian Province(2019J01773, 2018J01640) and China Scholarship Council(201709360002).

REFERENCES

[1] Yin Xu. The License plate position research. *Computer Knowledge and Technology*, vol.6, no.14, pp. 3729-3730, 2010.
[2] Zhang Xiaomin, Chang Hao. Research progress on license plate location algorithm. *Natural Science Journal of Harbin Normal University*, vol.33, no.1, pp.44-46+59, 2017.
[3] Jinn-Li Tan, Syed A.R Abu-Bakar and Musa M. Mokji. License plate localization based on edge-geometrical features using morphological approach. *IEEE International Conference on Image Processing. IEEE*, pp.4549-4553, 2014.
[4] GAN Ling, SUN Bo. Multiple license plate location based on separation projective and morphology operation. *Application Research of Computers*, vol.29, no.7, 2730-2732, 2012.
[5] Yu S, Li B, Zhang Q, et al. A novel license plate location method based on wavelet transform and EMD analysis. *Pattern Recognition*, vol.48, no. 1, pp. 114-125, 2015.
[6] Chang Qiaohong Gao Mantun. Research on license plate location based on HSV color space and Mathematical Morphology. *Journal of Graphics*, vol.34, no. 4, 159-162, 2013.
[7] Ashtari A H, Nordin M J, Fathy M. An Iranian license plate recognition system based on color features. *IEEE Transactions on Intelligent Transportation Systems*, vol.15, no.4, pp.1690-1705, 2014.
[8] Zhang Xiaona, He Ren,Chen Shian. Vehicle license plate location using active learning AdaBoost algorithm and color feature. *Journal of Traffic and Transportation Engineering*, vol.13, no.1, pp.121-126, 2013.
[9] GUO Keyou, JIA Haijing, GUO Xiaoli. Application of CNN in license plate classifier. *Computer Engineering and Applications*, vol.53, no.14, pp. 209-213, 2017.

[10] Yao D, Zhu W, Chen Y, et al. Chinese license plate character recognition based on convolution neural network. *2017 Chinese Automation Congress (CAC). IEEE*, 2017.

[11] Dorbe N, Jaundalders A, Kadikis R, et al. FCN and LSTM based computer vision system for recognition of vehicle type, license plate number, and registration country. *Automatic Control and Computer Sciences*, vol.52, no.2, pp.146-154, 2018.

[12] Laroca R , Severo E , Zanlorensi L A , et al. A robust real-time automatic license plate recognition based on the YOLO detector. *2018 International Joint Conference on Neural Networks (IJCNN)*, pp.1-10, 2018.

[13] Tsu-Yang Wu, Xiaoning Fan, King-Hang Wang, Jeng-Shyang Pan, Chien-Ming Chen and Jimmy Ming-Tai Wu. Security analysis and improvement of an image encryption scheme based on chaotic tent map. *Journal of Information Hiding and Multimedia Signal Processing(JIHMSP)*, vol. 9, no. 4, pp. 1050-1057, July 2018.

[14] Ying Ma, Nannan Wu. Research on seismic site emergency rescue traffic path analysis system based on public image information. *Journal of Information Hiding and Multimedia Signal Processing(JIHMSP)*, vol. 9, no. 3, pp. 577-585, May 2018.

[15] Liu W, Anguelov D, Erhan D, et al. SSD: single shot multiBox detector. *European Conference on Computer Vision*. pp. 21-37, 2016.

[16] Mark Sandler, Andrew Howard, Menglong Zhu, Andrey Zhmoginov, Liang-Chieh Chen. MobileNetV2: inverted residuals and linear bottlenecks. .https://arxiv.org/abs/1801.04381v4, April 2018.

[17] Francois Chollet. Xception: Deep learning with depthwise separable convolutions. https://arxiv.org/abs/1610.02357v3, April 2017.

[18] Xiangyu Zhang, Xinyu Zhou, Mengxiao Lin, and Jian Sun. Shufflenet: An extremely efficient convolutional neural network for mobile devices. 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), https://arxiv.org/abs/1707.01083v2, Dec 2017.

[19] Andrew G. Howard, Menglong Zhu, Bo Chen, Dmitry Kalenichenko, Weijun Wang, TobiasWeyand, Marco Andreetto, and Hartwig Adam. Mobilenets: Efficient convolutional neural networks for mobile vision applications. https://arxiv.org/abs/1704.04861v1, April 2017.

[20] HOU Xiangning, LIU Huachun. Plate location based on color segmentation and Sobel operator. *Computer technology and Development*, vol.28, no.8, pp.156-159, 2018.