SLPNet: Towards End-to-End Car License Plate Detection and Recognition Using Lightweight CNN

Wei Zhang[1], Yaobin Mao[1(✉)], and Yi Han[2]

[1] Nanjing University of Science and Technology, Nanjing 210094, China
edward.w.zhang@outlook.com, maoyaobin@njust.edu.cn
[2] Zhejiang Huayun Information Technology Co. LTD, Hangzhou 310008, China
hanyiuse@163.com

Abstract. As a core component, Automatic License Plate Recognition (ALPR) plays an important role in modern Intelligent Transportation System (ITS). Due to the complexity in real world, many existing license plate detection and recognition approaches are not robust and efficient enough for practical applications, therefore ALPR still a challenging task both for engineers and researchers. In this paper, a Convolutional Neural Network (CNN) based lightweight segmentation-free ALPR framework, namely SLPNet is established, which succinctly takes license plate detection and recognition as two associated parts and is trained end-to-end. The framework not only accelerates the processing speed, but also achieves a better match between the two tasks. Other contributions includes an anchor-free LP localization network based on corners using a novel MG loss is proposed and a multi-resolution input image strategy is adopted for different tasks to balance the operation speed and accuracy. Experimental results on CCPD data set show the effectiveness and efficiency of our proposed approach. The resulting best model can achieve a recognition accuracy of 98.6% with only 3.4M parameters, while the inference speed is about 25 FPS on a NVIDIA Titan V GPU. Code is available at https://github.com/JackEasson/SLPNet_pytorch.

Keywords: License plate detection · License plate recognition · End to end training · Segmentation-free · Lightweight CNN

1 Introduction

Automatic License Plate Recognition (ALPR) plays an important role in Intelligent Transportation System (ITS) which is widely used in traffic management, intelligent surveillance and parking management in large cities [20], hence, attracts considerable research attentions in recent two decades.

Even significant progress has been made especially with the help of deep learning in recent years, ALPR in nature environment is still a challenging task,

W. Zhang—Student author.

Y. Peng et al. (Eds.): PRCV 2020, LNCS 12306, pp. 290–302, 2020.
https://doi.org/10.1007/978-3-030-60639-8_25

due to the complexity and uncertainty of the environment and uncontrollable photographing conditions. Generally, an ALPR task can be divided into three subtasks according to its operating process, namely license plate detection, character segmentation and character recognition. With the popular of deep learning, a new trend to combine some of the above steps into one is appeared. For instance, character segmentation is merged with recognition in certain methods that directly predict the whole license plate numbers in form of sequences, which is known as segmentation-free license plate recognition [12,22,28]. Furthermore, some recent work has utilized a single network to simultaneously localize license plates and recognize the characters in a single forward inference. Such network not only can be trained end-to-end, but can avoid intermediate error accumulation [20] as well. However, it suffers from a dilemma of image resolution that the network operates at, since large images can get high recognition accuracy while at the expense of slow operation speed.

In this paper, a lightweight deep network named as SLPNet is proposed for ALPR task. The network architecture is based on lightweight fully convolutional network and designed specifically to reduce the intermediate error as well as reduce the impact from the unbefitting resolution of input images at the same time. The whole ALPR framework is segmentation-free and can be trained end-to-end with the advantages of high accuracy and low computational cost. The main contributions of the paper are summarized as follows:

(1) An anchor-free method for license plate detection based on corners instead of regions is introduced. As the detected corners can provide more geometrical distortion information that will be further used for perspective correction, license plate recognition can benefit from this design.
(2) A Multiple Constraints Gaussian Distance loss (MG loss for short) is put forward to improve corner localization precision, which is demonstrated to make license plate detector's training more stable and efficient.
(3) To further improve the recognition rate, we integrate a multi-resolution strategy into the end-to-end network architecture. For different subtask networks, an image is decomposed into different resolutions that are utilized as respect inputs.

The rest of the paper is organized as follows. Section 2 provides a brief review on traditional and current methods for ALPR. The proposed approach is presented in Sect. 3 by detailing the detection and recognition subtasks, network architectures as well as the loss function design. The experimental results are reported and discussed in Sect. 4. Finally, in Sect. 5 a conclusion is drawn.

2 Related Work

There are two kinds of approaches to perform ALPR in terms of feature extraction, i.e., manual feature extraction and automatic feature extraction. Generally, manual features are extracted by image analysis while automatic feature extractions are depended upon machine learning.

2.1 Methods Based on Manual Features

Traditional ALPR systems are usually based on manual features and can be divided into three separate subtasks, namely license plate detection, character segmentation and character recognition.

Existing license plate detection methods can be roughly classified into four categories based on the features they used, namely edge based [26], color based [1], texture based [25] and character feature based [17] method. Edge and color are obvious features that are easily affected by the variation of illumination, while texture and character features can provide more fine distributed information thus are more robust.

Traditional segmentation methods based on manual features often uses pixel connectivity and some prior knowledge of characters [2] to perform segmentation. They are simple but lack of robustness. Meanwhile, some more sophisticated approaches like character-contour-based methods [4] are complex and slow.

For license plate recognition task, each segmented character is subject to classification by Optical Character Recognition (OCR) technique. Template matching [5] is a simple and straightforward way, but lack of reliability. Thus, many new methods extract more efficient features with some advanced filters like Gabor filter [10] to further improve recognition accuracy.

2.2 Methods Based on Machine Learning

Shallow Learning. Early ALPR systems use shallow machine learning to substitute manual feature selection and classification in detection, segmentation and recognition subtasks. Classic machine learning algorithms such as SVM, AdaBoost are combined with different composite features conveniently to achieve better performance. In [8], an initial set of possible character regions are obtained by AdaBoost classifiers and then passed to a support vector machine (SVM) where noncharacter regions are rejected. For license plate recognition, in addition to SVM, many classifiers can be employed to recognize characters with effective feature extraction such as ANN [9].

Deep Learning. With the remarkable development of deep learning in recent years, detection and recognition tasks can reach better precision and robustness with the help of deep neural networks, freeing people from manual feature selection. Rayson Laroca et al. [11] used a one-stage detector to efficiently localize license plate regions, while Z. Selmi et al. [19] use simple convolutional neural network (CNN) to complete the task of single character classification and the method achieves high recognition accuracy.

Moreover, some current state-of-the-art ALPR frameworks adopt segmentation free methods and the whole network can be trained end to end, which leads to an efficient learning process and achieves excellent performance. Typical work such as RPNet [22] adopts a simple CNN as backbone for license plate detection and employs fully connected layers to classify characters in each detected image.

Besides, in [12], Bidirectional Recurrent Neural Networks (BRNNs) with Connectionist Temporal Classification (CTC) [6] are adopted to label the sequential data without character separation, leading to a high recognition accuracy. To further accelerate the process speed, S. Zherzdew et al. [28] use lightweight CNNs to extract features and train the model with CTC loss [6].

3 SLPNet

Different from aforementioned methods described in Sect. 2, an ALPR framework called Skip-shuffle License Plate Network (SLPNet) based on lightweight fully convolutional networks (FCNs) is proposed here. As illustrated in Fig. 1, our approach divides the whole framework into two associated parts: detection part and recognition part. Our method completes the detection task by localizing four corners of each license plate (LP). Then, the detected and cropped LP region will be processed by perspective correction to effectively reduce the recognition difficulty. For LP recognition, we treat it as a sequence labeling problem similar to LPRNet. Although the networks in different part are designed separately, the network as a whole can be trained end to end and be optimized with a joint detection and recognition loss function. In the following subsections, we will give a detailed description about each components.

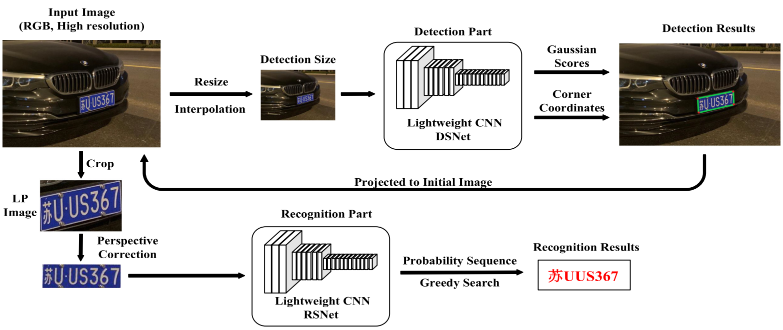

Fig. 1. The overall structure of our ALPR framework.

3.1 Detection Subnetwork

Corner Localization. To perform anchor-free object detection, an effective box is proposed in FSAF [29] where area inside is regarded as effective region (or positive region). Only cells from effective region are subject to object coordinates regression. Our approach also utilizes effective region to localize LP corners as shown in Fig. 2. We represent a LP region with an ordinary quadrangle according to the positions of its four corners, rather than the shape of a straight rectangle.

Similar to FSAF, two shrunk factors, δ_1 and δ_2 ($\delta_1 < \delta_2$) are selected to obtain effective boxes and ignore boxes as illustrated in Fig. 2(b). In our method,

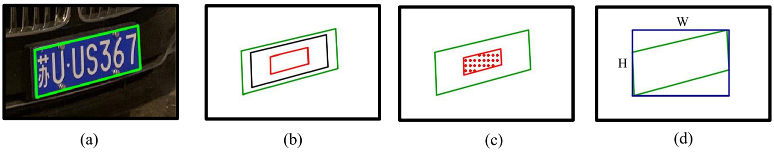

Fig. 2. An example of generating effective regions from ground truth. (a) A license plate with ground-truth bounding box; (b) Different types of boxes: bounding box (green), ignore box (black) and effective box (red); (c) Cells in positive area; (d) Rectangular bounding (blue) determined by a bounding box (green). (Color figure online)

three pairs of shrunk factors, namely $(0.8, 1.2)$, $(1.0, 1.5)$ and $(0.6, 0.9)$ are used respectively for different size of LP images. Large shrunk factor pairs are used for small size LPs while smaller ones are for large size LPs. That means more attention is paid to LPs with small size in training process. We divide the LPs into three classes namely small, middle and large, according to its size using k-means clustering based on the size of the rectangular bounding box determined by ground-truth corners as shown in Fig. 2(d).

Nonlinear Transformation. Each ground-truth bounding box consists of 4 corners and is represented by a vector $\boldsymbol{g} = (x_1, y_1, x_2, y_2, x_3, y_3, x_4, y_4)$. Our goal is to learn a nonlinear transformation that maps the network's output, $\boldsymbol{o} = (t_{x_1}, t_{y_1}, t_{x_2}, t_{y_2}, t_{x_3}, t_{y_3}, t_{x_4}, t_{y_4})$ to the ground-truth $\boldsymbol{g}$. Each (t_{x_i}, t_{y_i}) $i \in \{1, 2, 3, 4\}$ is the offset of the i-th corner to the cell center (x, y).

$$t_{x_i} = \sqrt[3]{\frac{x_i - 2^l(x + 0.5)}{z}}, t_{y_i} = \sqrt[3]{\frac{y_i - 2^l(y + 0.5)}{z}}. \tag{1}$$

where l is the pyramid level and z is an integer factor that shrinks the output ranges. Equation (1) first maps the coordinate (x, y) to the input image, then compute the offsets between the projected coordinates and $\boldsymbol{g}$ and regularizes the results with a cube root function.

Confidence of Detection Output. The normalized 2D Gaussian function is used to work out a score between the ground-truth corners and those predicted from the detection network. The score, also known as confidence, is denoted as Gaussian Scores that is described in formula (2).

$$\begin{aligned} G(x,y) &= Ae^{-(\frac{(x-x_0)^2}{2\sigma_x^2} + \frac{(y-y_0)^2}{2\sigma_y^2})}, \\ \sigma_x &= \alpha W, \sigma_y = \alpha H, A = 1. \end{aligned} \tag{2}$$

where (x, y) is a pair of the predicted corner coordinates and (x_0, y_0) is the ground-truth. As shown in Fig. 2(d), W and H are the width and the height of a rectangular bounding box determined by the 4 corners. Moreover, a scale factor $\alpha \in (0, 1)$ is used to control the Gaussian variance.

Detection Network. A lightweight fully convolutional network called DSNet is proposed for LP detection, where ShuffleNetv2 [16] units are adopted as basic blocks and skip connection [7] between different basic blocks is added as illustrated in Fig. 3. To further improve the performance, several Global context (GC) blocks [3] are employed to enhance the ability of feature representation. Specially, a stem block [21] for spatial downsampling is utilized in DSNet and complex feature maps are generated by three pyramid feature maps in different stages. In our implementation, the input image size is set to 512×512, thus, the output map size is 32×32. Therefore, the pyramid level l is equivalent to 4 and z can be set to 128 in formula (1).

The Loss Function for Detection Network. The targets of the detection network are to work out the nonlinear transformation of LP corners by regression and get high Gaussian Scores in cells from all regions. To achieve the goal, we proposed a Multiple Constraints Gaussian Distance loss (MG loss) inspired by CIoU loss [27]. The MG loss for each cell in positive regions is defined as

$$\begin{aligned} L_{MG} &= (1 - Conf) + \frac{\rho^2(b, b^{gt})}{c^2} + \alpha v + \beta d, \\ \alpha &= \frac{v}{(1 - Conf) + v}, \\ v &= \frac{2}{\pi^2}[(\arctan \frac{w_1}{h_1} - \arctan \frac{w_1^{gt}}{h_1^{gt}}) + (\arctan \frac{w_2}{h_2} - \arctan \frac{w_2^{gt}}{h_2^{gt}})], \\ \beta &= \frac{d}{(1 - Conf) + d}, d = \sqrt[2]{\frac{1}{4}\sum_{i=1}^{4}(Gs_i - Gs_i^{gt})^2}. \end{aligned} \tag{3}$$

where $Conf$ represents LP confidence that is averaged from four corners' real Gaussian Scores in a cell. $\rho(b, b^{gt})$ and c are distance and scale factor that are similar to CIoU loss. We enumerate four corners on each LP clockwise, therefore, in above formula (w_1, h_1) are worked out from Corner 1 and Corner 3. $(w_1^{gt}, h_1^{gt}, w_2^{gt}, h_2^{gt})$ are widths and heights generated by corners from ground-truth. Gs_i represents the predicted Gaussian Score of the i-th Corner, while Gs_i^{gt} represents a real Gaussian Score.

From formula (3), one can see MG loss consists of four terms: localization loss, distance loss, bounding shape loss and corners dispersion loss. The first item is the main loss and the others are constraints to make the learning process more stable and efficient. Considering the imbalance between positive and negative samples, we calculate Gaussian Score loss in each cell with Focal loss [14].

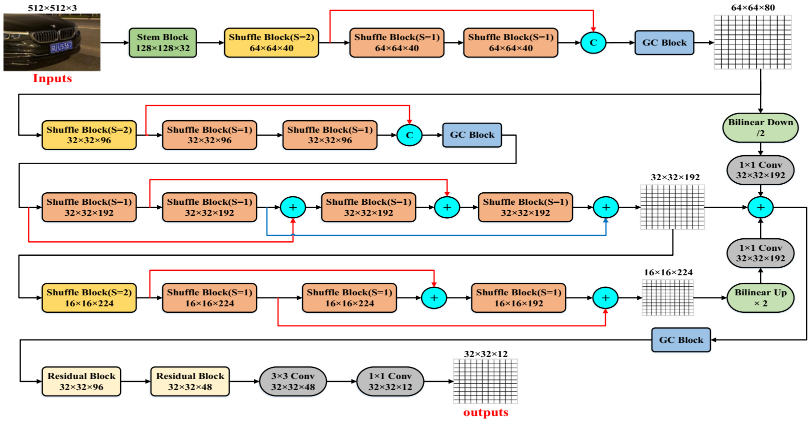

Fig. 3. The network structure of DSNet. 'C' represents concatenation operation while '+' represents addition. 'S' is the stride of convolutions. Bilinear interpolation is used to resize the pyramid features.

3.2 Recognition Subnetwork

Sequence Prediction. Since the number of characters on Chinese LP is unfixed depending on the type of vehicle, we treat the LP recognition as a sequence labeling problem by use of CTC employed in LPRNet [28], a kind of segmentation-free license plate recognition method.

Recognition Network. The recognition network, named as RSNet is also based on ShuffleNetv2 units, as illustrated in Fig. 4, where four kinds of blocks, $PDB(a) \sim PDB(d)$ for different spatial downsampling are designed. These blocks benefit from the parallel structure and are good at extracting features through downsampling. Like DSNet mentioned above, the intermediate feature maps are augmented with the global context embedding blocks (GCE blocks) [15]. To improve recognition performance, we also aggregate feature maps from different pyramid levels to get more complex maps.

The Loss Function for Recognition Network. Since the RSNet is based on CTC, the CTC loss is adopted in training recognition network.

3.3 Network Cascade

For end-to-end training, we need to link up the two separated networks and process the detection and recognition tasks sequentially. The predicted corners from DSNet are used to crop LP regions from raw images, which will be fed into RSNet then. Thus RSNet can be compatible with DSNet better, leading to less intermediate error propagation and achieving higher recognition accuracy. Since the DSNet uses small size images to perform detection, the detected and cropped LP regions are not suitable for recognition. To solve the problem, we re-map the

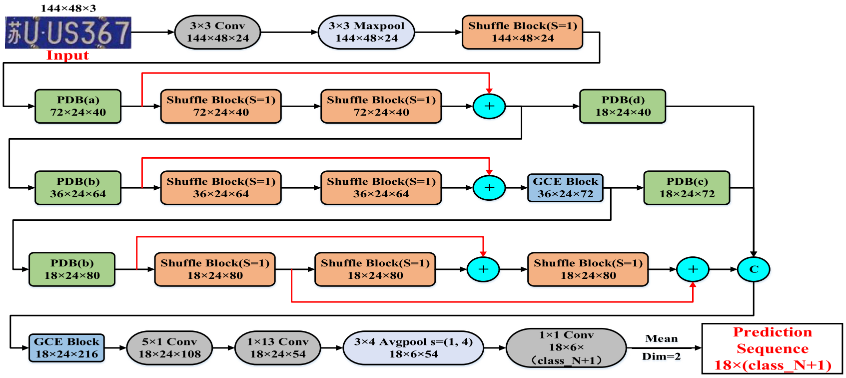

Fig. 4. The network structure of RSNet. '$PDB(a) \sim PDB(d)$' are four kinds of parallel downsampling blocks with different pooling layers or different strides. '$class\_N$' is the total number of character classes. '$Mean$' is an operator that averages the output maps in the 2nd dimension.

detected region to raw image and perform perspective correction according to the LP's coordinates of the DSNet's output. This multi-resolutions strategy can improve recognition accuracy effectively.

We utilize NMS [18] to get final detection results according to the predicted confidence. When decoding the recognition outputs, greedy search is applied on the output sequence of the RSNet to get the class with max probability.

4 Experiments

4.1 Chinese Licence Plate Dataset

In this section, the experimental results are reported. The performance of our method is compared with other state-of-the-art models in terms of a Chinese City Parking Dataset (CCPD)[22]. The CCPD dataset is the largest publicly available labeled license plate dataset in China by far. We randomly select 60,000 images from CCPD for experiment. As usual, all the images are subsequently split into 3 subsets in the proportion of 8:1:1 respectively for training, validation and testing.

4.2 Training Details

All training experiments are performed by pytorch on a NVIDIA Titan V GPU with 12 GB memory. We set Gaussian scale factor to 0.2 in training and use Adam optimizer with batch size of 22. The initial learning rate is set to 0.005. We drop the learning rate every epoch with exponential decay and the decay factor is 0.98. To make the training more stable, the whole process is consists of two stages. In the first phase, the detection network is trained until the average

Gaussian Scores of predicted corners in validation set is greater than 0.6, then the two subnetworks are trained jointly. In this stage, a loss function compound with both detection and recognition losses as described in formula (4) are used, where λ is a weighted factor, and $\lambda \in (0, 1)$:

$$L_{lp} = L_{det} + \lambda L_{reg} \tag{4}$$

Here, λ is set to 0.5. The model is trained jointly for 70 epochs in total. All hyper-parameters are adjusted through experiments and can be optimized further.

4.3 Experimental Results

During the training process, no data augmentation is performed and only the 48,000 raw images in the training set are utilized. We evaluate the performance of the proposed SLPNet with other publicly reported models on Chinese LP recognition. The rule for the calculation of the recognition accuracy is described as follows: only when a LP is detected successfully ($IoU \geq 0.5$ or $GaussianScore \geq 0.6$) and all the characters of the LP on an image are correctly recognized, the result is considered to be correct. The recognition accuracy for different methods is illustrated in Table 1.

Table 1. Experiment results of different methods on CCPD.

Model	End-to-end	Accuracy	Frame	Parameters
HyperLPR [24]	No	78.8%	–	11M
MTCNN+LPRNet [23,28]	No	91.8%	12FPS	3.6M
RPNet [22]	Yes	93.4%	58FPS	210M
SLPNet(ours)	Yes	98.6%	25FPS	3.4M

Other three publicly available models that we compared with are HyperLPR [24], MTCNN+LPRNet [23] and RPNet [22]. HyperLPR [24] is an open source Chinese license plate detection and recognition framework with high speed. The framework use a mixture of deep neural networks and classic image processing algorithms to perform detection, segmentation and recognition. MTCNN+LPRNet [23] is another open source lightweight ALPR framework based on LPRNet. It uses MTCNN to detect license plate and uses LPRNet, a segmentation-free method to perform recognition. RPNet [22] is an excellent end-to-end LP recognition model that first issued the CCPD dataset. As one can see, compared with other models, the SLPNet achieves the highest recognition accuracy at 98.6% with the least parameters and the inference speed arrives at 25 FPS. Some detection and recognition results are shown in Fig. 5, from which one finds that even under clutter scenes and uneven illumination condition our approach still can get stable results.

Fig. 5. Some typical detection and recognition results on CCPD with our SLPNet model.

To demonstrate the feasibility of the proposed MG loss, two other loss functions, namely Gaussian loss which is the first term of MG loss, and traditional smooth L1 loss are also adopted for comparison in term of LP detection. The detection accuracy with different loss functions is illustrated in Table 2.

Table 2. Detection accuracy with different loss functions.

Loss function	MG loss	Gaussian loss	Smooth L1 loss [11]
Detection accuracy	98.2%	97.7%	96.3%

LP detection can be viewed as a special nature scene text detection problem and DB [13] is a popular real-time scene text detector recently. As the last experiment, DB is utilized to detect LPs, compared with our SLPNet. The detection results are shown in Table 3. It's obvious that our SLPNet is much more efficient in LP detection task. And for better LP detection and recognition performance, an appropriative LP detector is considered to be necessary.

Table 3. Detection results on CCPD with different detectors.

Model	Precision	Recall	F1-Score
DB [13]	44.55%	81.00%	57.49%
SLPNet(ours)	99.87%	99.25%	99.56%

4.4 Performance Analysis

It should be noted that a segmentation-free ALPR framework with end-to-end training not only leads to easier learning process but also achieves a better balance between different subnetworks. The integration of the recognition subnetwork with the detection subnetwork can make the framework more consistent and matched for each other in end-to-end training. The networks in our method are built based on lightweight convolutional blocks and enhancement modules with low computational cost so that we can improve the detection and recognition performance with less parameters. Compared with RPNet which is also trained on CCPD, our method needs only a subset of the dataset while achieves higher recognition accuracy, owing to the multi-resolution training strategy and corner based detector with MG loss.

5 Concluding Remarks

In this paper, we introduce SLPNet, a segmentation-free end-to-end framework for efficient license palate detection and recognition, which can achieve up to 98.6% recognition accuracy. The model is based on lightweight convolutional networks therefore it can run fast and the total parameters are only 3.4M. To raise the detection rate, the proposed detection subnetwork uses corners instead of regions to locate license plates and a new MG loss function is introduced. The perspective transformation is utilized to correct LP images so that character recognition rate is improved. To gain better performance, a multi-resolutions strategy is adopted without adding any computational cost nearly. Compared with existing ALPR methods, our approach exhibits a noteworthy performance and great potentiality for LP detection and recognition.

References

1. Ashtari, A.H., Nordin, M.J., Fathy, M.: An iranian license plate recognition system based on color features. IEEE Trans. Intell. Transp. Syst. **15**(4), 1690–1705 (2014)
2. Busch, C., Domer, R., Freytag, C., Ziegler, H.: Feature based recognition of traffic video streams for online route tracing. IEEE Veh. Technol. Conf. **3**, 1790–1794 (1998)
3. Cao, Y., Xu, J., Lin, S., Fangyun, W., Hu, H.: GCNET: non-local networks meet squeeze-excitation networks and beyond (2019). https://arxiv.org/pdf/1904.11492.pdf
4. Capar, A., Gokmen, M.: Concurrent segmentation and recognition with shape-driven fast marching methods. In: 18th International Conference on Pattern Recognition, vol. 1, pp. 155–158 (2006)
5. Comelli, P., Ferragina, P., Granieri, M.N., Stabile, F.: Optical recognition of motor vehicle license plates. IEEE Trans. Veh. Tech. **44**(4), 790–799 (2002)
6. Graves, A., Fernández, S., Gomez, F.: Connectionist temporal classification: labelling unsegmented sequence data with recurrent neural networks. In: International Conference on Machine Learning, pp. 369–376 (2006)
7. He, K., Zhang, X., Ren, S., Sun, J.: Deep residual learning for image recognition. In: IEEE Conference on Computer Vision & Pattern Recognition, pp. 770–778 (2016)
8. Ho, W.T., Lim, H.W., Tay, Y.H.: Two-stage license plate detection using gentle adaboost and sift-svm. In: Asian Conference on Intelligent Information & Database Systems, pp. 109–114 (2009)
9. Jiao, J., Ye, Q., Huang, Q.: A configurable method for multi-style license plate recognition. Pattern Recogn. **42**(3), 358–369 (2009)

10. Kimand, M.K., Kwon, Y.B.: Recognition of gray character using gabor filters. Proc. Int. Conf. Inform. Fusion. **1**, 419–424 (2002)
11. Laroca, R., Severo, E., Zanlorensi, L.A.: A robust real-time automatic license plate recognition based on the yolo detector. In: 2018 International Joint Conference on Neural Networks, pp. 6005–6014 (2018)
12. 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 (2019)
13. Liao, M., Wan, Z., Yao, C., Chen, K., Bai, X.: Real-time scene text detection with differentiable binarization. In: AAAI Conference on Artificial Intelligence, pp. 11474–11481 (2020)
14. Lin, T.Y., Goyal, P., Girshick, R., He, K.: Focal loss for dense object detection (2017). https://arxiv.org/pdf/1708.02002.pdf
15. Liu, W., Rabinovich, A., Berg, A.C.: Parsenet: looking wider to see better (2015). https://arxiv.org/pdf/1506.04579.pdf
16. Ma, N., Zhang, X., Zheng, H.T., Sun, J.: Shufflenetv2: practical guidelines for efficient CNN architecture design. In: 2018 European Conference on Computer Vision, pp. 116–131 (2018)
17. Matas, J., Zimmermann, K.: Unconstrained licence plate and text localization and recognition. In: IEEE Intelligent Transportation Systems, pp. 572–577 (2005)
18. Neubeck, A., Gool, L.J.V.: Efficient non-maximum suppression. In: International Conference on Pattern Recognition, pp. 20–24 (2006)
19. Selmi, Z., Halima, M.B., Alimi, A.M.: Deep learning system for automatic license plate detection and recognition. In: 14th IAPR International Conference on Pattern Recognition, pp. 1132–1138 (2017)
20. Tian, Y., Lu, X., Li, W.: License plate detection and localization in complex scenes based on deep learning. In: The 30th Chinese Control And Decision Conference, pp. 6569–6574 (2018)
21. Wang, R.J., Li, X., Ling, C.X.: PELEE: a real-time object detection system on mobile devices. In: 32nd Conference on Neural Information Processing Systems, pp. 7466–7476 (2018)
22. Xu, Z., Yang, W., Meng, A., Lu, N., Huang, L.: Towards end-to-end license plate detection and recognition: a large dataset and baseline. In: 2018 European Conference on Computer Vision, pp. 261–277 (2018)
23. Xue, X.: A two stage lightweight and high performance license plate recognition in MTCNN and LPRNet (2019). https://github.com/xuexingyu24/License_Plate_Detection_Pytorch
24. Zeussees: High performance chinese license plate recognition framework (2020). https://github.com/zeusees/HyperLPR
25. Zhang, H., Jia, W., He, X., Qiang, W.: Learning-based license plate detection using global and local features. In: International Conference on Pattern Recognition. vol. 2, pp. 1102–1105 (2006)
26. Zheng, D., Zhao, Y., Wang, J.: An efficient method of license plate location. Pattern Recogn. Lett. **26**(15), 2431–2438 (2005)
27. Zheng, Z., Wang, P., Liu, W., Li, J., Ren, D.: Distance-IOU loss: faster and better learning for bounding box regression. In: AAAI Conference on Artificial Intelligence, vol. 34, pp. 12993–13000 (2020)

28. Zherzdev, S., Gruzdev, A.: LPRNet license plate recognition via deep neural networks (2018). https://arxiv.org/pdf/1806.10447.pdf
29. Zhu, C., He, Y., Savvides, M.: Feature selective anchor-free module for single-shot object detection. In: IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp. 840–849 (2019)