Learning to Detect License Plates Using Synthesized Data

Yanhui Pang, Wenzhong Wang, Aihua Zheng, and Jin Tang[✉]

Anhui University, Hefei, China
yanhuipang@foxmail.com, ahzheng214@foxmail.com,
{wenzhong,tj}@ahu.edu.cn

Abstract. Due to the lack of large-scale license plate dataset, existing license plate detection methods are usually conducted on small and unrepresentative datasets. Therefore, the training of these models maybe insufficient and only sub-optimal results can be achieved. In this paper, we propose a simple but effective method to handle this issue by automatically synthesizing license plate images. Specifically, we utilize Blender as a modeling and rendering engine to simulate various environmental factors and create scenes with diverse vehicle models. With these created models, we can obtain massive training data by synthesizing unique license plate. The benefits of our proposed method are: (1) we cannot only automatically provide pixel-level bounding box annotation of license plate, but also avoid errors caused by manual labeling. (2) the introduced algorithm is more efficient than manual labelling and thus we can generate a large-scale dataset in a rather short term. Based on these synthesized data, we propose a dilated convolutional attention augmentation module in conventional deep license plate detection algorithm to further boost the final detection performance. Extensive experiments on two benchmarks validate the effectiveness of our proposed algorithm.

Keywords: Data rendering · License plate detection · Attention mechanism · Dilated convolution

1 Introduction

With the development of modern traffic, license plate detection (LPD) and recognition technology have attracted more and more attention. It is commonly used in traffic monitoring, highway toll station, parking lot entrance, exit management and other actual monitoring systems. Although it has achieved great success in recent years, LPD is still a difficult task under the unconstrained scenarios, such as rotation, distortion, uneven illumination and vagueness. Most previous works [1,14,18,26] usually achieve good performance on extremely limited datasets. In some extremely complex real scenes, the performance may not satisfying due to the manual collected LP data is insufficient in both quantity and diversity. One intuitive idea is to collect and annotate a massive LP data

to obtain better detection results. However, this procedure is rather time and energy consuming. In addition, the human labelling may also introduce bias or errors.

Recently, some works [6,7] have applied synthetic training dataset to the field of object detection. Inspired by the work of Peter Slosar *et al.* [24] which uses synthetic data for vehicle detection. In this paper, we propose a novel synthetic data generation approach for license plate detection. Specifically, we use Matlab (version R2016b) and Blender[1] (version 2.79) with bundled Python (version 3.6) scripting for the rendering. We simulate various factors affecting license plate acquisition in natural scenes, and render different types of license plate images containing different regions. We extract the depth map of each rendered image for pixel-wise segmentation of the license plate in a given view, then compute an axis-aligned 2D bounding box, and realize automatic labelling of the license plate bounding box. The image of our synthesized dataset contains the LP images with various tilt angles, light intensities, and degrees of blur, which can cover a large diversity of the vehicles in the real scene.

The current state-of-the-art detection methods can be divided into two categories: the two-stage approach [5,8,9,23] and the one-stage approach [17,21]. However, these methods are not specifically designed for license plate detection, so the accuracy and efficiency of these methods for license plate detection maybe not optimal. In recent years, some works [6,16,25] have introduced dilated convolutional operation and attention mechanism into the field of object detection. Dilated convolution can help the network to expand the receptive field of convolutional kernels and obtain higher resolution features without increasing the parameter amount. Attention mechanisms can help the network better focus on the object area. Based on these observations, we propose to jointly integrate the dilated convolutional operations and attention model into a unified model, which can help the network to detect the small object license plate and improve the final detection performance. We first propose a dilated convolutional attention enhancement block for license plate detection. It introduces dilated convolutional operation into the Faster R-CNN [23] framework to increase the receptive field of convolution kernels and obtain higher resolution feature maps. Then the attention mechanism is introduced to weight the feature maps and help the neural network to achieve better classification performance.

In summary, this paper makes the following contributions:

– We propose a method to synthesize the license plate images, which can not only generate license plates of different provinces, cities and different types, but also realize the accurate labelling of the bounding box of the license plate area.
– We propose an novel license plate detection method based on Faster R-CNN. Specifically, we introduce the dilated convolutional operation and attention mechanism into the conventional convolutional network to generate more discriminative feature representations to achieve better performance of license plate detection.

---

[1] https://www.blender.org.

– Evaluations of the proposed LPD model on generated LP dataset demonstrate the validity of the synthetic license plate method and the proposed license plate detection model.

2   Related Work

2.1   Dateset of Synthetic LP

Most license plate datasets [3] tend to collect images from traffic monitoring systems, highway toll stations or parking lots. The collected license plate images usually have some shortcomings, such as small tilt angle, small number, uneven distribution of license plate types, manual annotation and so on. Therefore, these datasets can not evaluate LP detection algorithm very well. At present, the largest public license plate dataset is CCPD [27], but the CCPD dataset comes from the same city with limited types of license plate. On this basis, we propose to use the synthetic license plate dataset to simulate the real license plate data to train the detector. Our method can not only change the angle of license plate and various environmental factors, but also generate various license plate data from different provinces and cities with annotation.

2.2   LP Detection Algorithms

With the rapid development of region-based convolutional neural network [8], the currently popular object detection models have been widely applied in LP detection [10,14,15]. Faster R-CNN [23] utilizes a region proposal network which can generate high-quality region proposals for detection, so as to detect objects more accurately and quickly. YOLO [21] and YOLO9000 [22] frame object detection as a regression problem to spatially separated bounding boxes and associated class probabilities. SSD [17] combines the regression idea in YOLO and Anchor mechanism in Faster R-CNN, completely eliminate proposal generation and subsequent pixel or feature resampling stage, and encapsulate all calculations in a network.

2.3   Attention Mechanism

Attention mechanism in deep learning is essentially similar to human selective visual attention mechanism, whose goal is to select more critical information from a large number of information for the current task. At present, attention model has been widely used in various types of deep learning tasks, such as image recognition [6], speech recognition [2] and sequence learning [19], location and understanding of images [4,13]. Recently, the representative Squeeze-and-Excitation [11] reweights feature channels using signals aggregated from entire feature maps, while BAM [20] and CBAM [25] refine convolutional features independently in the channel and spatial dimensions.

Algorithm 1. License plate synthesizing algorithm

Input: M(total,savedir,camx,camy,sunx,suny,suni,brightness,contrast,licenseplate
map,backgroundmap)
Output: License plate image(Image) and bounding box coordinates(label)
 1: **function** RENDER(M)
 2:     **while** $i < total$ **do**
 3:         **if** M is True **then**
 4:             $\textbf{camera}_{horizontal} = \textbf{random}(camx)$
 5:             $\textbf{camera}_{vertical} = \textbf{random}(camy)$
 6:             $\textbf{sun}_{azimuth} = \textbf{random}(sunx)$
 7:             $\textbf{sun}_{height} = \textbf{random}(suny)$
 8:             $\textbf{sun}_{intensity} = \textbf{random}(suni)$
 9:             $\textbf{brightness} = \textbf{random}(brightness)$
10:             $\textbf{contrast} = \textbf{random}(contrast)$
11:             $\textbf{licenseplatemap} = \textbf{random}(licenseplatemap)$
12:             $\textbf{backgroundmap} = \textbf{random}(backgroundmap)$
13:         **end if**
14:         $[\textbf{Image}^{i}, \textbf{label}^{i}] = \textbf{render}(M)$
15:     **end while**
16: **end function**

3   Overview of the Synthetic Dataset

3.1   LP Rendering Methodology

The main idea of our license plate rendering system is illustrated in Fig. 1.
First, we generate different types of license plate maps (Fig. 1(a)) of different
provinces, and load these license plate maps into the defined vehicle object model
(Fig. 1(b)). The object is instantiated with a given set of parameters of mate-
rial properties (vehicle color, windows material properties, etc.). For this image,
depth map of LP is also rendered (Fig. 1(d)) which is used to determine the
bounding box of the license plate in the image. Algorithm 1 summarizes the
license plate synthesizing algorithm.
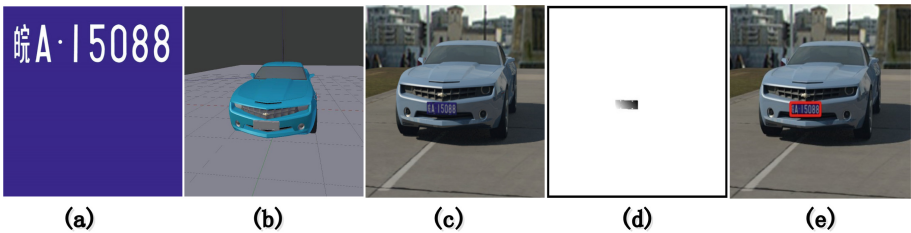
Fig. 1. The procedure of our license plate rendering framework. First we use MATLAB
to generate a license plate map (a) and load it into a defined set of vehicles (b). Then
we create a synthetic image of the vehicle and license plate from the defined collection
(c) and render a depth map of the license plate (d). The depth map is used to obtain
the precise segmentation of pixels in the license plate bounding box (e).

3.2   Viewpoint and Sunlight of LP

In order to obtain realistic and useful dataset, it is very important to control the camera angle and the illumination. As shown in Algorithm 1, we simulate the camera angle, solar direction and illumination intensity in natural scenes by specifying function parameters. We use variables **cmax** and **cmay** to control the horizontal and vertical view of the camera, **sunx** and **suny** to control the azimuth and height of the sun, **suni** to control the intensity of the sun. At the same time, we use **brightness** and **contrast** variables to adjust the brightness and contrast of the image.

4   Detection Approach

4.1   Model

The network architecture is shown in Fig. 2. We employ the Faster R-CNN [23] as the base network.

We propose a dilated convolutional attention block based on RFNet [16] and the attention model proposed by Woo *et al.* [25] called DCA Block. The feature map **FM** of the base network is fed into the dilated convolution layer and the attention module respectively. Then the merged feature map is used as the input feature map of the RPN network and the detection network. Compared with the whole image, the license plate is small. Therefore features from the convolution layer of the original model can not accurately describe the license plate, so the dilated convolution layer is added to obtain higher resolution features. Additionally, we apply channel attention and spatial attention to the feature map to enhance the features of the object area. The whole process is summarized as follows:

$$
\begin{aligned}
FM^{'} &= M_{dc}(FM) \\
FM^{''} &= [M_{sp}(M_{ch}(FM) \otimes FM)] \otimes [M_{ch}(FM) \otimes FM] \\
FM^{'''} &= FM^{'} + FM^{''}
\end{aligned}
\tag{1}
$$

where $\otimes$ denotes element-wise multiplication. $\mathbf{M_{dc}}$ is a dilated convolution attention map, $\mathbf{M_{ch}}$ is a one-dimensional channel attention map and $\mathbf{M_{sp}}$ is a two-dimensional spatial attention map. $\mathbf{FM}^{'''}$ is the final refined output.

4.2   Dilated Convolution

As shown in Fig. 2, we utilize the combination of multi-branch convolution layer and dilated pool. Specifically, we first use the bottleneck structure in each branch, consisting of a 1×1 convolution layer plus an $n \times n$ convolution layer, followed by a pooling or convolution layer with a corresponding dilation. The detailed process is as follows:
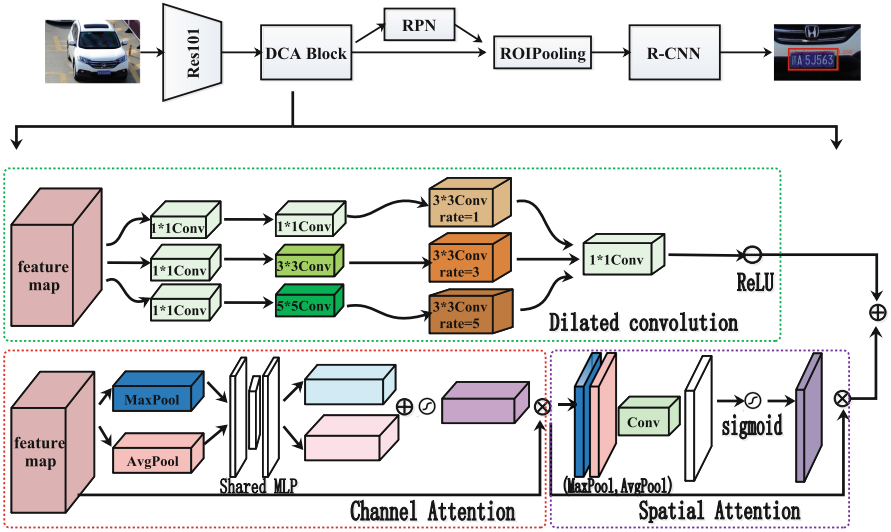
Fig. 2. The overall structure of our network. In the Faster R-CNN base network output feature map, we introduce the dilated convolution layer to obtain higher resolution features while maintaining the parameters and the same receptive field. In addition, we use the feature map channel and spatial attention to enhance the object features. Finally, we fuse the two output results and feed them to RPN and subsequent detection network.

$$
\begin{aligned}
M_{dc}(FM) =& \varphi[f^{1\times1}(concat(f^{3\times3}_{r=1}(f^{1\times1}(f^{1\times1}(FM))), \\
& f^{3\times3}_{r=3}(f^{3\times3}(f^{1\times1}(FM))), f^{3\times3}_{r=5}(f^{5\times5}(f^{1\times1}(FM))))))] \qquad (2) \\
=& \varphi[f^{1\times1}(concat(FM_{r=1}, FM_{r=3}, FM_{r=5}))]
\end{aligned}
$$

where φ denotes the Relu activation function and $f^{n\times n}$ represents a convolution operation with the filter size of n × n, $f^{3\times3}_r$ denotes an dilated convolution operation with a convolution kernel of size 3 and an dilation rate of r.

4.3 Channel Attention

As in Fig. 2, we utilize average pooling and maximum pooling operations to aggregate the spatial information of the feature map, and generate two different spatial average pool features and maximum pool features. Then both are fed to the shared network to generate our channel attention map. The shared network consists of a multilayer perceptron (MLP) and a hidden layer. After sharing the network layer, we use element summation to merge the output eigenvectors.

$$
M_{ch}(FM) = \sigma(MLP(AvgPool(FM)) + MLP(MaxPool(FM))) \qquad (3)
$$

4.4   Spatial Attention

Following the idea of CBAM [25] module, the feature map produced by channel attention module is used as the input feature map of spatial attention module. First, we make a global max pooling and global average pooling on channel, and then concatenate the two results. After a convolution operation, the dimension is reduced to one channel, then a spatial attention feature can be acquired via sigmoid function. Finally, the feature is multiplied with the input feature of the module to obtain the final feature, as shown in Fig. 2. The specific calculation is as follows:

$$M_{sp}(FM) = \sigma(f^{7 \times 7}([AvgPool(M_{ch}(FM)), MaxPool(M_{ch}(FM))])) \qquad (4)$$

where σ denotes the sigmoid function, the MLP weight shared and $f^{7 \times 7}$ represents a convolution operation with the filter size of 7×7.

Training. We choose Resnet-101 as backbone. Our model is pre-trained on ImageNet and then fine-tuned on synthetic dataset of license plate. The experiment is implemented in pytorch and trained end-to-end on a group with four Tesla P100 GPUs, with Stochastic Gradient Descent (SGD) and a weight decay of 0.0001 and momentum of 0.9. At the beginning of the training process, the learning rate is set to 0.001. After 20 epochs, the learning rate decreases by 0.1 times for every 5 epochs.

5   Experiments

In this section, we give a detailed description of our synthetic dataset, further more, we evaluate the the performance of different detectors on our synthetic dataset and CCPD dataset. We show that the detectors trained with the synthetic dataset are comparable with those trained with the real license plate dataset. Finally, we synthesize a dataset of 20,000 yellow, blue and new energy license plates, and compare the performance of the prevalent detection algorithm with our algorithm. We show that the proposed method improves the accuracy of license plate detection compared with the original method.

5.1   Data Preparation

As aforementioned in Sect. 3, we render and synthesize a large license plate dataset (SLPD100) containing only **blue plates**, which contains about 100 K images with resolution of 800 (Width) × 1160 (Height) × 3 (Channels). For each image, the bounding box label contains (x, y) coordinates of the top left and bottom right corner of the bounding box are used to locate the minimum bounding rectangle of LP. The CCPD dataset is the largest license plate dataset in public, which contains about 250 k images. We divide CCPD into two parts, the default training set containing about 100 k images, and the default evaluation set containing about 20,000 images. The training set and test set of our experiment are shown in Table 1.

Table 1. The division of training set and test set and the numbers in brackets represent the size of dataset.

Training set	Test set	License plate type
SLPD100 (100 k)	CCPD [27] (20 k)	Blue license plate
CCPD [27] (100 k)	CCPD [27] (20 k)	Blue license plate
SLPD20 (20 k)	LPD3000 (3 k)	Blue, yellow and new energy license plate

In order to verify the validity of our proposed license plate detection method, we also synthesize another dataset of about 20 k (SLPD20), including **yellow license plate, blue license plate and new energy license plate**. Our test dataset for evaluating detector performance is about 3,000 license plate images (LPD3000) taken by surveillance cameras and hand-held cameras.

5.2   Experiment Analysis

Evaluation Criterion. We follow the standard protocol (Intersection-over-Union (IoU)) [12] of object detection. The bounding box is considered to be correct if and only if its IoU with the ground-truth bounding box is more than 70% (IoU > 0.7).

Experimental Results on Different Datasets. We synthesize a dataset similar to the CCPD dataset and conduct experiments by the current prevalent YOLO9000, SSD, Faster R-CNN detection algorithm. Table 2 shows the experimental results. In the experiment, we set the same parameters for the same dataset. As shown in Table 2, we use the synthetic license plate datasets (SLPD100) and CCPD datasets to train SSD, Faster R-CNN and YOLO9000 detectors and use CCPD (20k) as the test set. The test accuracy on SSD and Faster R-CNN are about 2% to 3% lower than the real dataset CCPD, and the performance on YOLO9000 is only 0.2% higher. The main reason may be that the data distribution of the synthetic datasets is not comprehensive enough. Because we do not consider these factors such as license plate occlusion, rain, snow and fog, which may be a gap between our synthetic datasets and the license plate image collected by the natural scene. In future work, we will consider further increasing the data diversity in the synthetic datasets. Generally speaking, the detector trained with the synthetic license plate dataset is comparable to the detector trained with the real license plate dataset, which shows the effectiveness the synthetic license plate datasets.

Experimental Results on Different Detectors. We evaluate different detectors on the synthetic dataset. The experimental results are shown in Table 3. We evaluate the current prevalent detectors, SSD [17], YOLO9000 [22], R-FCN [5] and Faster R-CNN [23]. The results show that the performance of our detector based on Faster R-CNN improves about 2.2% compared with other detectors. In Sect. 5.3, we will analyse the effectiveness of our proposed method in detail.

Table 2. Comparison of license plate testing accuracy between different datasets of the same detector.

Dataset	Model	blue_licenceplate(%)
CCPD [27] (100 k)	SSD300+VGG16	99.1
	YOLO9000+Darknet19	98.8
	Faster R-CNN+Resnet-101	98.1
OurDataset (100 k)	SSD300+VGG16	**97.4**
	YOLO9000+Darknet19	**99.0**
	Faster R-CNN+Resnet-101	**96.2**

Table 3. Test accuracy of different detectors.

NAME	mAP(%)	yellowlp(%)	bluelp(%)	nelp(%)
R-FCN [5]	87.9	86.1	89.3	88.4
SSD300 [17]	86.0	86.5	88.1	83.4
YOLO9000 [22]	84.3	83.4	89.0	80.6
Faster R-CNN [23](**baseline**)	88.9	87.9	88.8	90.1
Faster R-CNN+Channel Attention	89.5	89.3	89.3	89.8
Faster R-CNN+Dual Attention	90.0	90.4	89.5	90.3
Faster R-CNN+Dilated convolution	89.8	90.0	89.3	90.2
Faster R-CNN+DCABlock(our)	**91.1**	**90.9**	**91.3**	**91.2**

5.3   Ablation Studies

For ablation study, we use the SLPD20 and LPD3000 as training and test dataset. In the experiment, we progressively introduce the channel attention on Faster R-CNN, the spatial attention, and then the dilation convolution module, and report the results on Table 3.
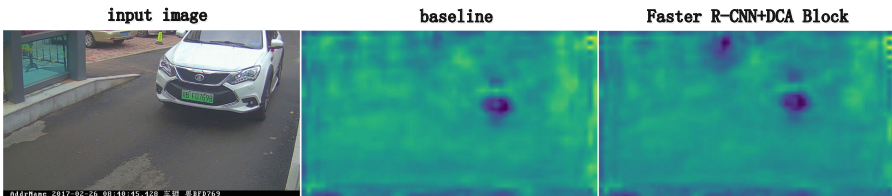
Fig. 3. Feature visualization. We compare the visualization results of our method (Faster R-CNN + DCA Block) with baseline (Faster R-CNN).

Dilated Convolution. As shown in Table 3, after the introducing the dilated convolution module, the detection accuracy on the test set increases by 0.9% (from 88.9% to 89.8%) compared to the baseline.

Channel Attention. We choose Faster R-CNN as our baseline, and we introduce the channel attention module between the base network Resnet-101 and the RPN network. As shown in Table 3, introducing channel attention, the mAP increases after by about 0.6% (from 88.9% to 89.5%), which demonstrates the effectiveness of the channel attention model.

Dual Attention. As shown in the experiment in Table 3, the introduction of the channel and spatial dual attention model improves the license plate detection accuracy by 1.1% (from 88.9% to 90.0%), which demonstrates the effectiveness of the dual attention model.

DCA Block. Based on the above analysis, we consider combining the dilated convolution and the dual attention module. The feature maps generated by the dilated convolution module and the dual attention module are fused. The experimental results show that the accuracy of our method is about 2.2% (from 88.9% to 91.1%) higher. For qualitative analysis, we compare the visualization results of our method (Faster R-CNN + DCA Block) with baseline (Faster R-CNN) in Fig. 3. We can see that our method pays more attention to the object area than baseline. Meanwhile, as in Fig. 4, our method can detect almost all the objects in the image. The result shows the effectiveness of the method.

Fig. 4. Comparison of results. The first line is the baseline method test result picture, in which the blue rectangle represents the undetected object, the second line is our method test result picture.

6   Conclusions

In this paper, we present a method to synthesize license plate datasets and a dilated convolutional attention augmentation module in conventional deep license plate detection. The proposed license plate synthesis method can not only simulate the real scene by controlling the illumination intensity and other environmental factors of the synthetic images, but also can automatically label the license plate area as ground truth. It is very useful to solve the problems of limited license plates in training dataset and high cost manual labeling under some specific conditions. The proposed dilated convolutional attention augmentation module uses the dilated convolutional operation with different dilation rates to increase the receptive field of convolution kernels and obtain the higher resolution feature maps. In addition, the attention mechanism is added to learn the weight map for better classification. Extensive evaluations on two benchmarks demonstrate that our method improves the performance of license plate detection over the baseline methods.

References

1. Al-Shemarry, M.S., Li, Y., Abdulla, S.: Ensemble of adaboost cascades of 3l-lbps classifiers for license plates detection with low quality images. Exp. Syst. Appl. **92**, 216–235 (2018)
2. Bluche, T.: Joint line segmentation and transcription for end-to-end handwritten paragraph recognition. In: Advances in Neural Information Processing Systems (2016)
3. Caltech: Caltech licese plate dataset. http://www.vision.caltech.edu/html-files/archive.html
4. Cao, C., et al.: Look and think twice: capturing top-down visual attention with feedback convolutional neural networks. In: Proceedings of the IEEE International Conference on Computer Vision (2015)
5. Dai, J., Li, Y., He, K., Sun, J.: R-fcn: Object detection via region-based fully convolutional networks. In: Advances in Neural Information Processing Systems (2016)
6. Duan, S., Hu, W., Li, R., Li, W., Sun, S.: Attention enhanced convnet-RNN for Chinese vehicle license plate recognition. In: Lai, J.H., et al. (eds.) PRCV 2018. LNCS, vol. 11257, pp. 417–428. Springer, Cham (2018). https://doi.org/10.1007/978-3-030-03335-4_36
7. Georgakis, G., Mousavian, A., Berg, A.C., Kosecka, J.: Synthesizing training data for object detection in indoor scenes (2017)
8. Girshick, R.: Fast R-CNN. In: The IEEE International Conference on Computer Vision (ICCV) (2015)
9. Girshick, R., Donahue, J., Darrell, T., Malik, J.: Rich feature hierarchies for accurate object detection and semantic segmentation. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (2014)

10. Hsu, G.S., Ambikapathi, A., Chung, S.L., Su, C.P.: Robust license plate detection in the wild. In: 2017 14th IEEE International Conference on Advanced Video and Signal Based Surveillance (AVSS) (2017)
11. Hu, J., Shen, L., Sun, G.: Squeeze-and-excitation networks. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (2018)
12. Hui, L., Peng, W., Shen, C., Hui, L., Peng, W., Shen, C.: Towards end-to-end car license plates detection and recognition with deep neural networks. IEEE Trans. Intell. Transp. Syst. (2017)
13. Jaderberg, M., Simonyan, K., Zisserman, A., et al.: Spatial transformer networks. In: Advances in Neural Information Processing Systems (2015)
14. Laroca, R., et al.: A robust real-time automatic license plate recognition based on the yolo detector. In: 2018 International Joint Conference on Neural Networks (IJCNN) (2018)
15. Li, H., Wang, P., Shen, C.: Towards end-to-end car license plates detection and recognition with deep neural networks. corr abs/1709.08828 (2017)
16. Liu, S., Huang, D., et al.: Receptive field block net for accurate and fast object detection. In: Proceedings of the European Conference on Computer Vision (ECCV) (2018)
17. Liu, W., et al.: SSD: single shot multibox detector. In: Leibe, B., Matas, J., Sebe, N., Welling, M. (eds.) ECCV 2016. LNCS, vol. 9905, pp. 21–37. Springer, Cham (2016). https://doi.org/10.1007/978-3-319-46448-0_2
18. Masood, S.Z., Shu, G., Dehghan, A., Ortiz, E.G.: License plate detection and recognition using deeply learned convolutional neural networks (2017)
19. Miech, A., Laptev, I., Sivic, J.: Learnable pooling with context gating for video classification. arXiv preprint arXiv:1706.06905 (2017)
20. Park, J., Woo, S., Lee, J.Y., Kweon, I.S.: Bam: bottleneck attention module (2018)
21. Redmon, J., Divvala, S., Girshick, R., Farhadi, A.: You only look once: unified, real-time object detection. In: Proceedings of the IEEE conference on computer vision and pattern recognition (2016)
22. Redmon, J., Farhadi, A.: Yolo9000: better, faster, stronger. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (2017)
23. Ren, S., He, K., Girshick, R., Sun, J.: Faster R-CNN: towards real-time object detection with region proposal networks. In: Advances in Neural Information Processing Systems (2015)
24. Šlosár, P., Juránek, R., Herout, A.: Cheap rendering vs. costly annotation: rendered omnidirectional dataset of vehicles. In: Proceedings of the 30th Spring Conference on Computer Graphics. ACM (2014)
25. Woo, S., Park, J., Lee, J.Y., So Kweon, I.: Cbam: convolutional block attention module. In: Proceedings of the European Conference on Computer Vision (ECCV) (2018)
26. Xie, L., Ahmad, T., Jin, L., Liu, Y., Sheng, Z.: A new CNN-based method for multi-directional car license plate detection. IEEE Trans. Intell. Transp. Syst. **19**, 507–517 (2018)
27. Xu, Z., et al.: Towards end-to-end license plate detection and recognition: a large dataset and baseline. In: Proceedings of the European Conference on Computer Vision (ECCV) (2018)