Journal of Theoretical and Applied Information Technology
31st August 2020. Vol.98. No 16

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

VEHICLE LOGO RECOGNITION BASED ON VEHICLE REGION AND MULTI-SCALE FEATURE FUSION

HOANH NGUYEN

Faculty of Electrical Engineering Technology, Industrial University of Ho Chi Minh City, Ho Chi Minh City, Vietnam

E-mail: nguyenhoanh@iuh.edu.vn

ABSTRACT

This paper presents a deep learning-based vehicle logo recognition framework based on vehicle regions and single shot framework with multi-scale feature fusion. To exactly locate front and rear regions of a vehicle, a vehicle region detection network based on region proposal network is designed. The vehicle region detection network uses ResNet-50 as the base network to generate base convolution feature maps. Atrous convolution is used after the base network to enlarge the receptive field to incorporate larger context information without increasing the number of parameters or the amount of computation. Each of detected vehicle region is fed into a vehicle logo recognition network based on single shot framework. The vehicle logo recognition network adopts Darknet-53 as the base network to generate proposals from each vehicle region. To enhance the performance of the proposed framework on small vehicle logo recognition, this paper designs a multi-scale feature fusion subnet which generates high-level semantic feature maps from base convolution feature maps. The tasks of detection predictions and object classification are performed on the multi-scale high-level semantic feature maps. Furthermore, a collected dataset based on a large public vehicle dataset is used to evaluate the proposed method. Experimental results on the collected dataset show the effectiveness of the proposed method on vehicle logo recognition. The proposed method provided a promising solution to vehicle logo recognition applications.

Keywords: *Vehicle Logo Recognition, Deep Learning, Multi-Scale Feature Fusion, Atrous Convolution, Single Shot Framework*

1. INTRODUCTION

Vehicle logo recognition is widely used in advanced driver assistance systems, intelligent parking systems, traffic surveillance. Exactly recognize logo of a vehicle will provide inherent properties associated with that vehicle such as vehicle type, dimensions, carrying capacity, generation, and so on. These properties can help to effectively improve the robustness and reliability of vehicle analysis activities used in intelligent transport systems. A vehicle logo recognition framework normally includes two stages: vehicle logo detection and logo classification. Vehicle logo detection aims to exactly locate the position of a logo in an image, and logo classification aims to assign a certain logo class label to each detected vehicle logo. Many methods have been conducted to perform vehicle logo recognition in the literature. These methods can be divided into two groups: traditional methods and deep learning-based methods. Traditional methods design different handcrafted features such as feature points, edges, and invariant moments to train a detector. Ou et al. [16] proposed to used regions of interest based on AdaBoost-based detector at first to coverage of the vehicle logo. A weighted spatial pyramid framework is then used to extract feature vectors from these regions. A linear support vector machine classifier is adopted to classify the logos based on max pooling of local descriptors. Chen et al. [17] proposed a framework based on a spatial Scale Invariant Feature Transform combined with a logistic regression classifier for vehicle logo recognition. Sun et al. [18] first adopted LBP features combined with the AdaBoost method to detect vehicle logos, and then exploited HOG features combined with a support vector machine classifier for recognition. Chen et al. [19] proposed to use a feature-match based symmetry detection algorithm for detecting symmetric axis in the image and rotating the image. A learning-based license plate localization approach and a coarse logo

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

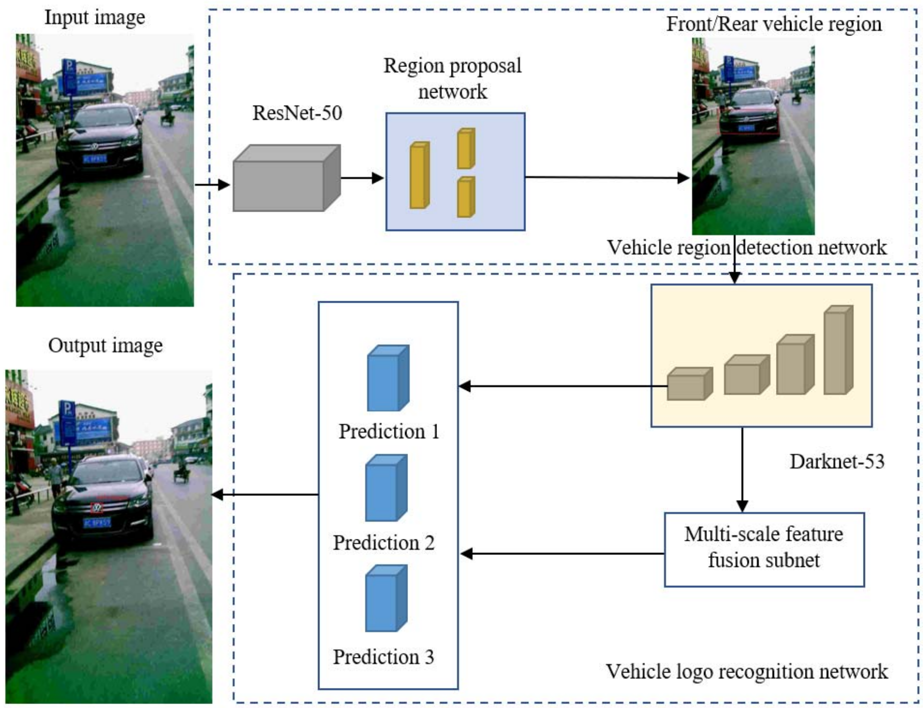

Figure 1: The Overview Structure of The Proposed Method.

location algorithm is then designed by using prior knowledge to obtain a coarse logo region. In addition, a new vehicle logo location method based on Gabor Filters is put forward in order to localize the exact region of logo. Yu et al. [20] proposed descriptors combining Oriented Edge Magnitudes and overlapping LBP features for vehicle logo recognition. Traditional methods are easy to understand and implement. However, these methods with specially designed features can only solve a single problem, which is inadequate to recognize vehicle logos under various complex environmental situations, such as night illumination, logo tilt, and bad weather.

Recently, deep learning-based object recognition methods have become dominant in computer vision tasks. Thus, many approaches based on deep learning for vehicle logo recognition have been proposed and showed better performance compared with traditional approaches. Huang et al. [21] proposed a convolutional neural network (CNN) model for vehicle logo recognition, which consists of two convolution layers, two pooling layers and two fully connected layers. In addition, a pretraining strategy is applied to accelerate the training procedure to render the vehicle logo recognition system suitable for real-world applications. Li and Hu [22] propose a Mapreduce based CNN called MRCNN to train the networks, which significantly increases the training speed and reduces the computation cost simultaneously. Furthermore, a novel genetic algorithm global optimization and Bayesian regularization approach is designed to initialize the weights of classifier, which help prevent the overfitting and avoid the local optima. Xia et al. [23] proposed to combine a CNN with the multitask learning approach and used an adaptive weight training strategy to accelerate the convergence of the model. Soon et al. [24] presented a method that aimed to automatically search and optimize a CNN architecture for vehicle logo recognition. Huan et al. [25] used the Hough transform to achieve accurate vehicle logo detection based on the locations of a vehicle's logo and license plate. Then, vehicle logo classification was performed with deep belief networks. Liu et al. [26] proposed a framework with three stages for vehicle logo recognition. First, a constrained region extraction method based on segmentation of the car head and car tail is proposed to accurately extract the candidate region of logo. Then, an enhanced matching method is proposed to improve the detection performance of small objects. Finally, a single deep neural network based on a reduced ResNeXt model and Feature Pyramid Networks is proposed to improve classification performance of the network and retain more detailed information for small-sized vehicle logo detection. Yuxin and H.

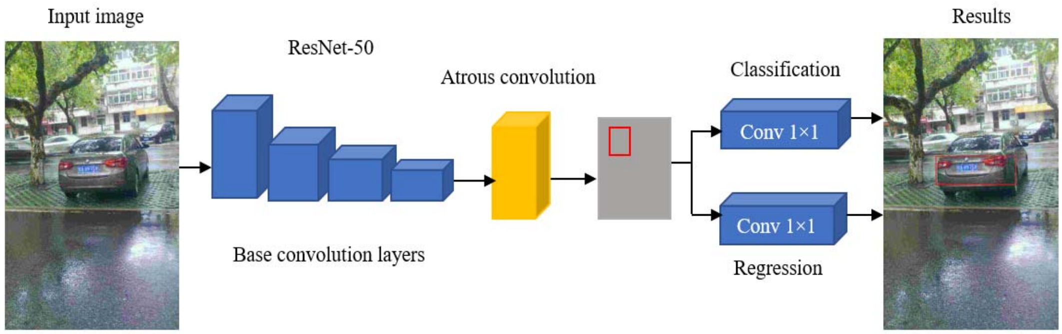

Figure 2: The Structure of The Vehicle Region Detection Network Proposed in This Paper.

Peifeng [27] proposed to combine the logo information and license plate information to improve the reliability of vehicle identification. In addition, a highway entrance vehicle logo recognition system is designed by using CNN to detect and classify the vehicle logo in real time.

2. PROPOSED APPROACH

2.1 Overview of Proposed Method

The overview structure of the proposed method is shown in Figure 1. As shown, the proposed method includes two stages. In the first stage, a vehicle region detection network is designed based on region proposal network [1] to locate front or rear region of a vehicle. The vehicle region detection network first uses ResNet-50 [2] as the base network to generate base convolution feature maps. Then, an atrous convolution layer is applied on the highest-level feature map to enlarge the receptive field. Finally, two parallel sibling layers are used to classify and regress each of detected vehicle region. In the second stage, a vehicle logo recognition network is designed based on single shot framework. The extracted vehicle regions generated by the first stage are fed into the vehicle logo recognition network to exactly locate and classify vehicle logos into the background and different categories of vehicle logos. In this stage, Darknet-53 [3] architecture is first adopted to generate base convolution feature maps. To enhance the performance of the proposed framework on small vehicle logo recognition, a multi-scale feature fusion subnet which generates high-level semantic feature maps from base convolution feature maps is designed. Finally, the tasks of detection predictions and object classification are performed on the multi-scale high-level semantic feature maps. Details of each stage will be explained in the following sections.

2.2 Vehicle Region Detection Network Based on Region Proposal Network

Figure 2 illustrates the structure of the vehicle region detection network proposed in this paper. The vehicle region detection network takes images captured in traffic scenes and containing vehicle as the inputs and outputs front or rear region of a vehicle. As shown in Figure 2, the vehicle region detection network can be roughly partitioned into three stages. The first stage contains base convolution layers to generate initial feature maps. The second stage includes an atrous convolution layer acting on the highest-level feature map of the first stage to enlarge the receptive field to incorporate larger context without increasing the number of parameters or the amount of computation. The third stage contains two parallel sibling layers: classification layer for classifying each object proposals into vehicle regions (front region or rear region) and background; regression layer for estimating a set of object proposals centered at each position of the feature maps. By fusing the information from the two sibling layers, only the object proposals labeled as vehicle regions are selected as the final region proposals.

In the first stage, ResNet-50 [2] is used as base network to generate the initial feature maps. ResNet is an efficient network which adopted residual learning to every few stacked layers such that the training of networks can be eased and substantially deeper than others. ResNet-50 is not only more accurate than VGG-16 [4] but also faster than VGG-16. Supposing the size of input images is 224×224, the structure of ResNet-50 used in this paper is shown in Table 1. As shown, average pooling layer and fully connected layer in original ResNet architecture are discarded, and the highest-level feature layer is adopted as input feature to the second stage. In the second stage, atrous convolution, which is a powerful tool in dense prediction tasks, can effectively enlarge the field of view of filters to incorporate larger context without increasing the

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

Table 1: The Structure of ResNet-50 Used in This Paper.

Layer name	Kernel size	Output size
Conv1	7 × 7 × 64, stride = 2	112 × 112
Conv2_x	$\begin{bmatrix} 1\times1\times64 \\ 3\times3\times64 \\ 1\times1\times256 \end{bmatrix}\times3$	56 × 56
Conv3_x	$\begin{bmatrix} 1\times1\times128 \\ 3\times3\times128 \\ 1\times1\times512 \end{bmatrix}\times4$	28 × 28
Conv4_x	$\begin{bmatrix} 1\times1\times256 \\ 3\times3\times256 \\ 1\times1\times1024 \end{bmatrix}\times6$	14 × 14
Conv5_x	$\begin{bmatrix} 1\times1\times512 \\ 3\times3\times512 \\ 1\times1\times2048 \end{bmatrix}\times3$	7 × 7

number of parameters or the amount of computation. Atrous convolution with a rate r introduces $(r-1)$ zeros between consecutive filter values, effectively enlarging the kernel size of $k \times k$ filter to $k' \times k'$ by using equation (1) without increasing the number of parameters or the amount of computation.

$$k' = k + (k-1)(r-1) \quad (1)$$

In this paper, atrous convolution is used with a 3×3 kernel size and rate r = 2. Therefore, the kernel size is enlarged to $k' = 5$, which means that more context information can be obtained.

To generate vehicle region proposals, this paper designs a set of anchors with different scales and aspect ratios centered at each position of the highest-level feature map. In this paper, an anchor is designed to be a rectangular shape since front and rear vehicle region are usually in rectangular shape. More specific, this paper uses four scales and one aspect ratios, yielding $n = 4$ anchors at each position. In total, for the highest-level feature map with a size of $W \times H$, there are total $W \times H \times k$ anchors distributing on the input image. The regression layer aims to adjust the raw anchors at each position to predict a tight bounding box for each anchor to encapsulate an object. The regression layer outputs $4k$ values for respectively adjusting the position, width, and height of each of the anchors. The classification layer aims at estimating k predicted proposals at each position to deduce the objectness. Thus, each position should output $2k$ scores to indicate the probabilities of objectness for each of the proposals. The classification and regression layers can be easily derived with two sibling convolutional layers with kernel sizes of 1×1 operated on the highest-level feature maps of the first stage. The numbers of the resultant feature maps are $2k$ and $4k$ for the classification and regression layers, respectively. Thus, the accurate location information of the vehicle regions can be obtained through the classification and bounding box regressor. To train the vehicle region proposal network, this paper collects 3,000 images containing front and rear regions of vehicle. The front region and rear region of each vehicle in image in this collected dataset are carefully annotated. Details of training process will be introduced in next section.

2.3 Vehicle Logo Recognition Network Based on Single Shot Framework

The extracted region proposals generated by the vehicle region detection network are fed into a vehicle logo recognition network to exactly locate and classify vehicle logos into the background and different categories of vehicle logos. To effectively recognize vehicle logos of diverse conditions, a one-stage framework is proposed for logo recognition in this paper. The structure of the proposed vehicle logo recognition network is illustrated in Figure 3. As shown, Darknet-53 [3] is first applied to extract features from vehicle regions. An additional layer (conv_6) is added after the highest-level convolution layer of the base network to make the first prediction. The additional layer with high-level context and semantic information will improve the detection performance on large vehicle logos. Since vehicle logos usually occupy a small region on traffic scene image, a multi-scale feature fusion subnet with two feature fusion modules is designed to enhance the

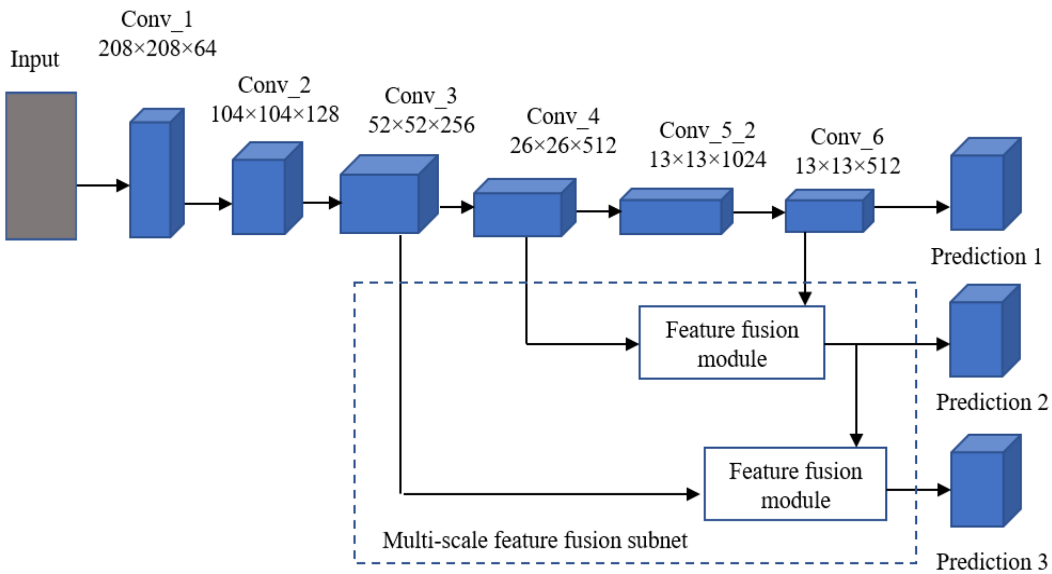

Figure 3: The Structure of The Proposed Vehicle Logo Recognition Network.

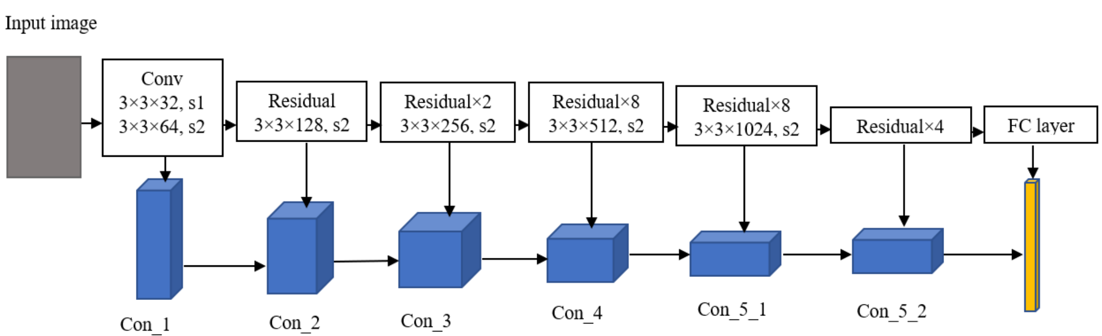

Figure 4: The Architecture of Darknet-53 Used in This Paper.

performance on small vehicle logos. Finally, the tasks of detection prediction and object classification are performed on the multi-scale high-level semantic feature maps. Details of each module will be explained in the next subsections.

2.3.1 Feature Extraction Module

In this paper, Darknet-53 architecture is adopted to generate base convolution feature maps. Figure 4 shows the structure of Darknet-53 used in this paper. There are total 53 convolutional layers without any pooling layer. Instead of using max-pooling or average-pooling for decreasing feature map size, the size of the feature map is decreased by applying stride at 2 at each block. Darknet-53 adopts residual blocks to ease the training process of deep network. Each residual block contains 1×1 and 3×3 convolutional filters, and there are total 23 residual blocks used in the network. A batch normalization (BN) operation [5] is applied after each convolutional layer in the whole framework to control overfitting and convergence during the training process. In addition, the Leaky ReLU [6] is used as the activation function in each convolutional layer.

2.3.2 Multi-Scale Feature Fusion Subnet

To enhance the performance of the proposed framework on small vehicle logo recognition, this paper designs a multi-scale feature fusion subnet which generates high-level semantic feature maps from base convolution feature maps. Figure 5 shows the structure of the proposed multi-scale feature fusion subnet. As shown, the high-level feature maps (e.g., layer of conv_6 and the first feature fusion module) are upsampled with a factor of 2 by performing upsampling operation. The upsampled feature maps will have the same size as the low-level feature maps (e.g., layer of conv_4 and conv_3). Another 1 × 1 convolutional layer is added before the upsampling operation to ensure that the channel dimensions of the upsampled feature maps are the

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

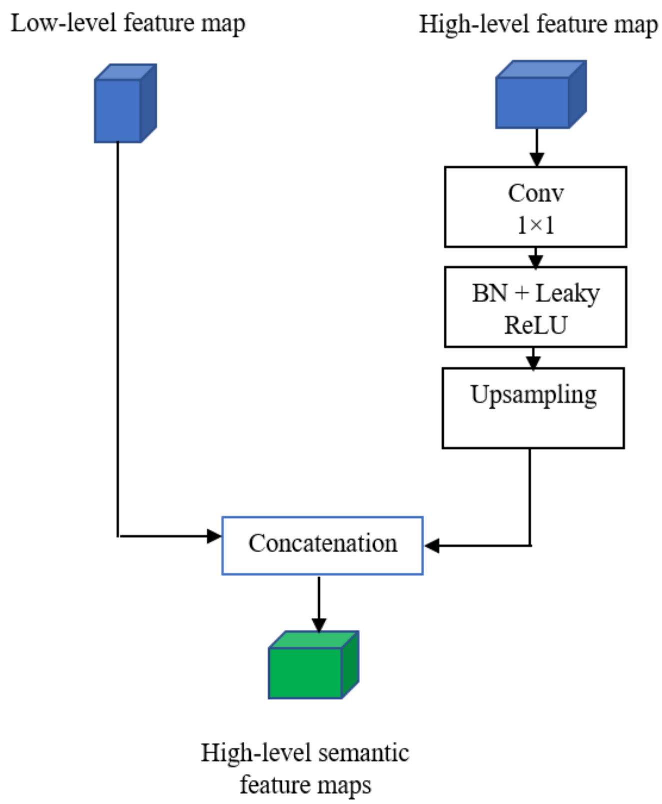

Figure 5: The Structure of The Proposed Multi-Scale Feature Fusion Subnet.

same as the corresponding low-level feature maps. Each convolutional layer is followed with a BN layer and a Leaky ReLU layer. Lastly, concatenation is performed to merge the upsampled feature maps with the corresponding low-level feature maps. Thus, high-level semantic feature maps are built to enhance the detection performance of the proposed method on small vehicle logo recognition.

2.3.3 Multi-scale Object Recognition

The tasks of detection predictions and object classification are performed on the multi-scale high-level semantic feature maps, including conv_6 layer and two feature fusion modules. There are three feature maps used for predictions, and each feature map generates a fixed set of default boxes. To solve the problem of multiple-scale object detection, different default boxes are used with particular designed scales (scales = [13, 26, 52]) and five aspect ratios (aspect ratios = [1, 2, 3, 1/2, 1/3]). During the training process, 13 labels are assigned to each ground-truth box according to such categories as vehicle logos and background. The overall objective loss function is defined as follows:

$$L = \frac{1}{N} L_{conf} + \propto L_{loc} \tag{2}$$

where N represents the number of matched default boxes, the confidence loss L_{conf} is a softmax loss over 13 classes, and the localization loss L_{loc} is the default bounding box regression for positive boxes.

3. EXPERIMENTAL RESULTS

3.1 Dataset

This paper locates and recognizes vehicle logo from images captured in natural scenes and containing vehicles. Since currently there are few publicly vehicle image datasets for vehicle logo recognition, this paper generates a self-collected dataset to evaluate the performance of the proposed vehicle logo recognition approach. Images in the collected dataset are taken from the Chinese City Parking Dataset (CCPD) [7], which is used to evaluate license plate detection methods. The CCPD provides over 250k unique over 250k unique car images under diverse environments. The resolution of each image is 720×1160. Based on the CCPD dataset, 8,000 images are chosen carefully to generate the collected dataset. There are 13 classes of vehicle logos in this dataset as shown in Figure 6. The resolution of logo instances ranges from 10×10

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

Figure 6: 13 Classes of Vehicle Logos in The Collected Dataset.

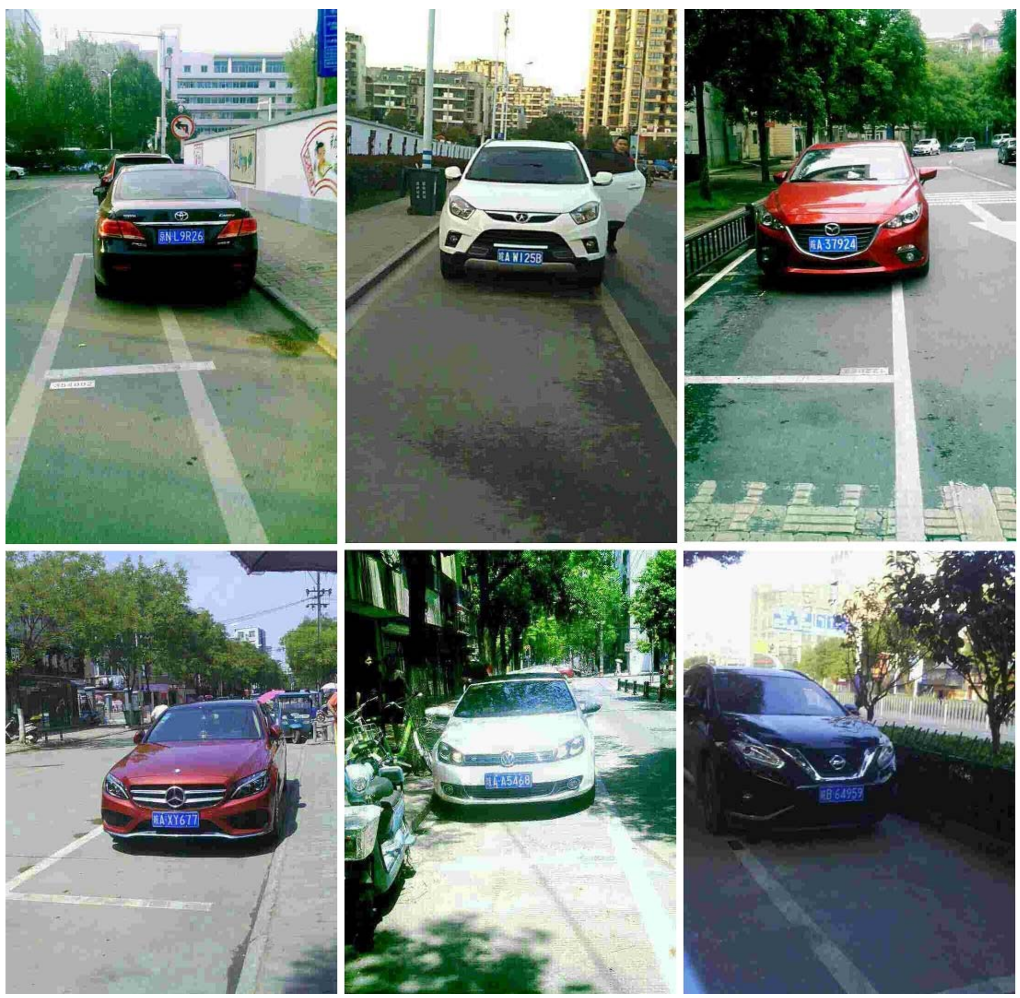

Figure 7: Example Images in The Collected Dataset.

pixels to 150×150 pixels. In this paper, 7,000 images are used for training and 1,000 images are used for testing. Of this, each logo class has at least 500 images as training samples and 70 images as testing samples. Figure 7 shows some example images in the collected dataset.

Table 2: Comparison of The Detection Results.

Methods	Base network	mAP (%)	Inference time (s)
Faster R-CNN	VGG-16	82.1	0.74
	RestNet-50	85.3	0.92
YOLOv3	Darknet-53	83.0	0.20
	RestNet-50	87.8	0.55
SSD	VGG-16	80.2	0.24
	RestNet-50	82.7	0.76
Proposed method	Darknet-53	90.5	0.85

3.2 Implementation Details

The proposed method was implemented using Python 3.8, and Pytorch library [8] is adopted for implementing deep CNN framework. Comprehensive experiments are conducted on the collected dataset to evaluate the performance of the proposed method and compare the recognition results with several state-of-the-art methods. The evaluation is performed on a Window system machine with an Intel Core i7 CPU, 16 GB of memory and NVIDIA 1080 GPU.

The training process of the whole approach includes two phases. The vehicle region detection network is trained first, and then both the vehicle region detection network and the vehicle logo recognition network are trained. For training the vehicle region detection network, the dataset aforementioned in section 3.1 is used. Front region and rear region of each vehicle in image in this dataset are carefully annotated. To increase the number of images for training, data augmentation is performed as in [9]. For the base network, the pretrained ResNet-50 trained on ImageNet dataset [10] is applied. The vehicle region detection network is trained end-to-end by backpropagation and stochastic gradient descent. Each image in training set contains a set of positive and negative anchors for classification and regression. This paper randomly samples 256 anchors in an image to calculate the multi-task loss function for a batch, where the numbers of the positive and negative anchors are identical. To fine-tune the network, this paper uses the initial learning rate of 0.001 for the first 200,000 iterations, and the learning rate of 0.0001 for the remaining 100,000 iterations. The momentum and weight decay are configured as 0.9 and 0.0005, respectively. For training the vehicle logo recognition network, the dataset aforementioned in section 4.1 is used. During the training procedure, the parameters of Darknet-53 are initialized with the pretrained weights on the large PASCAL VOC dataset [11]. The learning rate starts from 0.001 and then decays by a factor of 0.1 when the training loss does not reduce in five epochs. An Adam method [12] is adopted for efficient stochastic optimization.

3.3 Experimental Results

This paper conducts experiments on the collected dataset and compares the recognition results with the results of recent deep CNN-based object detector, including Faster R-CNN [13], YOLOv3 [14], SSD [15]. For each of the competitive network, this paper conducts experiments on both original base network and ResNet-50. RestNet-50 is adopted to replace original base network to increase the recognition performance on small vehicle logos. More specific, Faster R-CNN with RestNet-50 adopts the highest-level feature map with size of 17×17 for generating proposals and making detection predictions. SSD with RestNet-50 uses extra feature layer with size of 65×65 to retain more detail information. YOLOv3 with RestNet-50 adopts three feature maps with size of 65×65, 33×33 and 13×13 to making detection predictions. Table 2 shows the detection results of the proposed method and all competitive methods on the collected dataset. As shown in Table 2, the proposed method achieves the best detection accuracy on the collected dataset. More specific, the mAP of the proposed method is improved by 5.2%, 2.7%, 7.8% compared with Faster R-CNN, YOLOv3, and SSD respectively. Furthermore, the results in Table 2 show that using RestNet-50 instead of VGG-16 and Darknet-53 will improve the detection accuracy. This result shows the effectiveness of RestNet-50 in designing object detection framework. For the inference time, the proposed method takes 0.76 second to process an image. As shown in Table 2, YOLOv3 with Darknet-53 as the base network is the fastest method. However, YOLOv3 achieves lower accuracy than the proposed method. The superior performance of

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

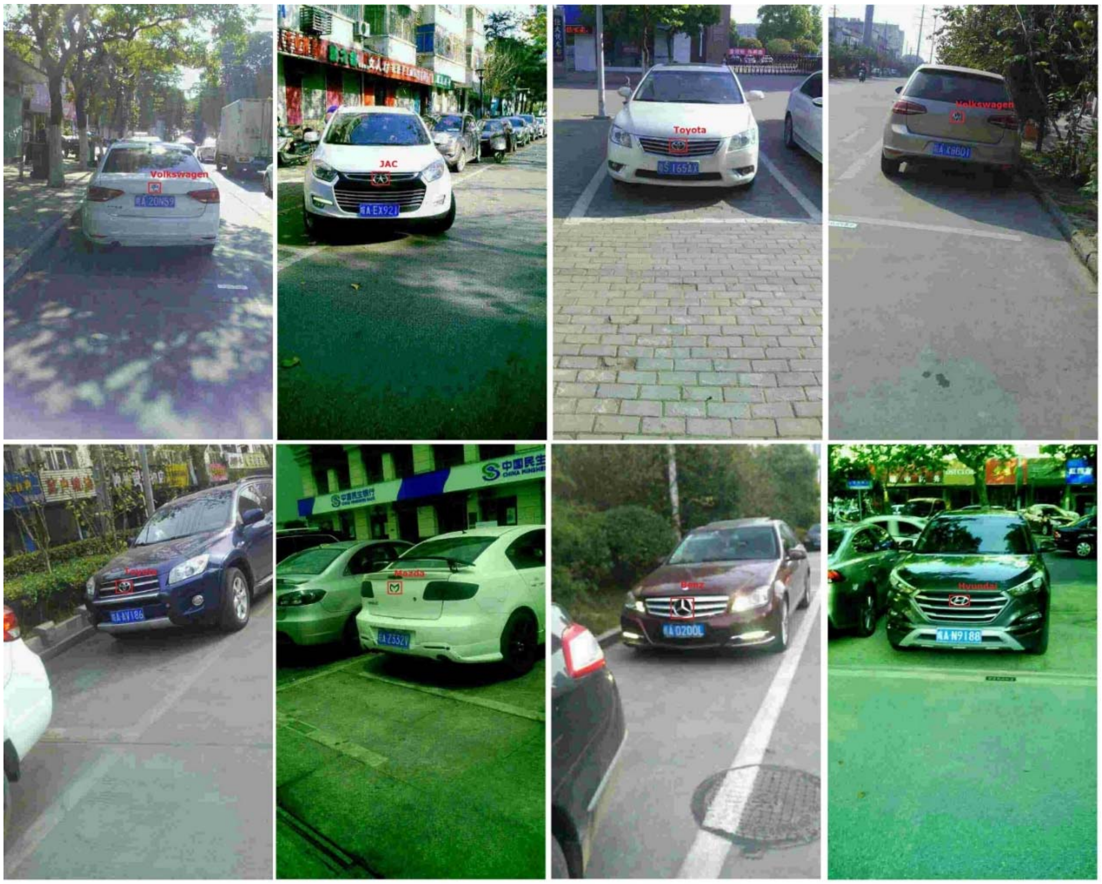

Figure 8: Example of Detection Results of The Proposed Method on The Collected Dataset.

the proposed framework is benefitted from the design of the vehicle logo recognition network with multi-scale feature fusion subnet, which can extract robust and distinctive features of vehicle logos, especially with small logos. Thus, the proposed framework provided a promising solution to vehicle logo recognition applications. Figure 8 shows some example results of the proposed method on the collected dataset. As shown, the proposed network can exactly locate small logos in difficult environment conditions.

4. CONCLUSIONS AND FUTURE WORK

This paper presents a deep learning-based framework for vehicle logo recognition. In the proposed framework, a vehicle region detection network is first designed to detect front and rear regions of a vehicle. Atrous convolution is adopted on the highest-level feature map to enlarge the receptive field, thus enhancing the performance of vehicle region detection. The extracted vehicle regions generated by the vehicle region detection network are fed into a vehicle logo recognition network to exactly locate and classify vehicle logos into the background and different categories of vehicle logos. Furthermore, a multi-scale feature fusion subnet which generates high-level semantic feature maps from base convolution feature maps is designed to enhance the performance of the proposed framework on small vehicle logo recognition. Experimental results on the collected dataset show that the proposed framework provides a promising solution to vehicle logo recognition applications. In the future, this paper will explore more improvements to enhance the detection performance and the inference speed.

REFERENCES

[1] Ren, Shaoqing, Kaiming He, Ross Girshick, and Jian Sun. "Faster r-cnn: Towards real-time object detection with region proposal networks." In *Advances in neural information processing systems*, pp. 91-99. 2015.

[2] He, Kaiming, Xiangyu Zhang, Shaoqing Ren, and Jian Sun. "Deep residual learning

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

for image recognition." In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp. 770-778. 2016.

[3] Redmon, Joseph, and Ali Farhadi. "Yolov3: An incremental improvement." *arXiv preprint arXiv:1804.02767* (2018).

[4] Simonyan, Karen, and Andrew Zisserman. "Very deep convolutional networks for large-scale image recognition." *arXiv preprint arXiv:1409.1556* (2014).

[5] Ioffe, Sergey, and Christian Szegedy. "Batch normalization: Accelerating deep network training by reducing internal covariate shift." *arXiv preprint arXiv:1502.03167* (2015).

[6] Maas, Andrew L., Awni Y. Hannun, and Andrew Y. Ng. "Rectifier nonlinearities improve neural network acoustic models." In *Proc. icml*, vol. 30, no. 1, p. 3. 2013.

[7] Xu, Zhenbo, Wei Yang, Ajin Meng, Nanxue Lu, Huan Huang, Changchun Ying, and Liusheng Huang. "Towards end-to-end license plate detection and recognition: A large dataset and baseline." In *Proceedings of the European Conference on Computer Vision (ECCV)*, pp. 255-271. 2018.

[8] Paszke, Adam, Sam Gross, Francisco Massa, Adam Lerer, James Bradbury, Gregory Chanan, Trevor Killeen et al. "PyTorch: An imperative style, high-performance deep learning library." In *Advances in Neural Information Processing Systems*, pp. 8024-8035. 2019.

[9] KHALIFA, NOUR ELDEEN M., MOHAMED LOEY, and MOHAMED HAMED N. TAHA. "Insect pests recognition based on deep transfer learning models." *Journal of Theoretical and Applied Information Technology* 98, no. 01 (2020).

[10] Deng, Jia, Wei Dong, Richard Socher, Li-Jia Li, Kai Li, and Li Fei-Fei. "Imagenet: A large-scale hierarchical image database." In *2009 IEEE conference on computer vision and pattern recognition*, pp. 248-255. Ieee, 2009.

[11] Everingham, Mark, Luc Van Gool, Christopher KI Williams, John Winn, and Andrew Zisserman. "The pascal visual object classes (voc) challenge." *International journal of computer vision* 88, no. 2 (2010): 303-338.

[12] Kingma, Diederik P., and Jimmy Ba. "Adam: A method for stochastic optimization." *arXiv preprint arXiv:1412.6980* (2014).

[13] Ren, Shaoqing, Kaiming He, Ross Girshick, and Jian Sun. "Faster r-cnn: Towards real-time object detection with region proposal networks." In *Advances in neural information processing systems*, pp. 91-99. 2015.

[14] Redmon, Joseph, and Ali Farhadi. "Yolov3: An incremental improvement." *arXiv preprint arXiv:1804.02767* (2018).

[15] Liu, Wei, Dragomir Anguelov, Dumitru Erhan, Christian Szegedy, Scott Reed, Cheng-Yang Fu, and Alexander C. Berg. "Ssd: Single shot multibox detector." In *European conference on computer vision*, pp. 21-37. Springer, Cham, 2016.

[16] Ou, Yuanchang, Huicheng Zheng, Shuyue Chen, and Jiangtao Chen. "Vehicle logo recognition based on a weighted spatial pyramid framework." In *17th International IEEE Conference on Intelligent Transportation Systems (ITSC)*, pp. 1238-1244. IEEE, 2014.

[17] Chen, Ruilong, Matthew Hawes, Lyudmila Mihaylova, Jingjing Xiao, and Wei Liu. "Vehicle logo recognition by spatial-SIFT combined with logistic regression." In *2016 19th International Conference on Information Fusion (FUSION)*, pp. 1228-1235. IEEE, 2016.

[18] Sun, Quan, Xiaobo Lu, Lin Chen, and Haihui Hu. "An improved vehicle logo recognition method for road surveillance images." In *2014 Seventh International Symposium on Computational Intelligence and Design*, vol. 1, pp. 373-376. IEEE, 2014.

[19] Chen, Cong, Xiaobo Lu, Shengqin Jiang, and Jiaji Song. "An effective vehicle logo recognition method for road surveillance images." In *2016 2nd IEEE International Conference on Computer and Communications (ICCC)*, pp. 728-732. IEEE, 2016.

[20] Yu, Ye, Jun Wang, Jingting Lu, Yang Xie, and Zhenxing Nie. "Vehicle logo recognition based on overlapping enhanced patterns of oriented edge magnitudes." *Computers & Electrical Engineering* 71 (2018): 273-283.

[21] Huang, Yue, Ruiwen Wu, Ye Sun, Wei Wang, and Xinghao Ding. "Vehicle logo recognition system based on convolutional neural networks with a pretraining

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

strategy." *IEEE Transactions on Intelligent Transportation Systems* 16, no. 4 (2015): 1951-1960.

[22] Li, Binquan, and Xiaohui Hu. "Effective vehicle logo recognition in real-world application using mapreduce based convolutional neural networks with a pre-training strategy." *Journal of Intelligent & Fuzzy Systems* 34, no. 3 (2018): 1985-1994.

[23] Xia, Yizhang, Jing Feng, and Bailing Zhang. "Vehicle Logo Recognition and attributes prediction by multi-task learning with CNN." In *2016 12th International Conference on Natural Computation, Fuzzy Systems and Knowledge Discovery (ICNC-FSKD)*, pp. 668-672. IEEE, 2016.

[24] Soon, Foo Chong, Hui Ying Khaw, Joon Huang Chuah, and Jeevan Kanesan. "Hyper-parameters optimisation of deep CNN architecture for vehicle logo recognition." *IET Intelligent Transport Systems* 12, no. 8 (2018): 939-946.

[25] Huan, Li, Qin Yujian, and Wang Li. "Vehicle logo retrieval based on hough transform and deep learning." In *Proceedings of the IEEE International Conference on Computer Vision Workshops*, pp. 967-973. 2017.

[26] Liu, Ruikang, Qing Han, Weidong Min, Linghua Zhou, and Jianqiang Xu. "Vehicle Logo Recognition Based on Enhanced Matching for Small Objects, Constrained Region and SSFPD Network." *Sensors* 19, no. 20 (2019): 4528.

[27] Yuxin, Mao, and Hao Peifeng. "A Highway Entrance Vehicle Logo Recognition System Based on Convolutional Neural Network." In *2019 2nd International Conference on Artificial Intelligence and Big Data (ICAIBD)*, pp. 282-285. IEEE, 2019.