2017 13th International Conference on Computational Intelligence and Security

Forward Vehicle Detection Based on Incremental Learning and Fast R-CNN

Kaijing Shi
Beijing Key Laboratory of Information Service Engineering, Beijing Union University
Beijing, China
e-mail: 151081210205@buu.edu.cn

Hong Bao*
Beijing Key Laboratory of Information Service Engineering, Beijing Union University
Beijing, China
e-mail: xxtbaohong@buu.edu.cn

Nan Ma
College of Robotics, Beijing Union University
Beijing, China
e-mail: xxtmanan@buu.edu.cn

***Abstract*—Recently the research of vehicle detection is mainly through machine learning, but it still has low detection accuracy problem. With the study of researchers, using deep learning methods of vehicle detection becomes hot. In this paper, a selective search method and a target detection model based on Fast R-CNN are used to detect vehicle. The strategy optimizes the model by preprocessing the sample image and the new network structure. Firstly, the experiment uses the public KITTI data set and self-collected BUU-T2Y data set, respectively, for training validation and test. Secondly, based on the original data set, the experiments go on through incremental learning, combining the KITTI dataset with the BUU-T2Y dataset. The experimental results show that the proposed method is superior to the result of multi - feature and classifier detection in terms of accuracy. To a large extent, the proposed method solved the problem of missing vehicle for detection and improved the accuracy of vehicle testing and robustness.**

Keywords-deep learning; image ;vehicle detection; Fast R-cnn; accurate rate;

I. INTRODUCTION

With the development of science and technology, intelligent transportation system arises out of the historic moment. Intelligent Vehicle as a part of intelligent transportation system became a hot issue of researchers to research. And vehicle detection is the important application of the field of intelligent vehicles. It plays a key role in increasing safety of autonomous driving.

Research on the forward vehicle detection methods, researchers have proposed many algorithms at home and abroad. It is mainly divided into the traditional machine learning method and deep learning. In [1], the researchers used the gradient direction histogram (HOG) combined with support vector machine (SVM) classifier to train. And the method realized image detection in engineering vehicles. But it was not good for the location about border around the vehicles. In [2], the paper improved multidimensional Haar - like features and Adaboost algorithm to realize the cascade classifier training, so as to realize reliable vehicle detection. The separation between lanes and other areas uses the lane-edge detection algorithm. The experimental results show that, compared with the traditional algorithms method, it can effectively improve the accuracy and efficiency of vehicle detection. In[3],the researchers proposed a Haar-Like characteristics based on LBP feature robust detection algorithm. Finally it uses Adaboost classifier to classify vehicles category. In addition to the machine learning method, convolution neural network has been widely applied to the field of target detection. Among them, the convolutional neural network based on region (region-based CNN, R - CNN) [4] [5] and Fast R - CNN develop very rapidly. In[4],it found the candidate boxes , then used CNN to extract the feature vector, and finally used SVM to classify feature vectors. In [5], the researchers found that RCNN is a multi-stage pipeline and time-consuming when it was training. The problems were improved on FRCNN. It mainly puts the deep web and the SVM classification together, then uses a new network to do classification and regression directly. And the Fast R - CNN in training VGG16 network is 9 times higher than RCNN at speed. In [13] ,a classic Spatial Pyramid Pooling structure was introduced into the CNN, so that the CNN can handle any size and scale of the images. This method not only improves the classification accuracy, but also is very suitable for detection. The result shows that it is more quickly and precisely than classic RNN.In[15], The researchers introduce an algorithm for incremental learning. This algorithm was inspired by AdaBoost algorithm. It utilizes ensemble of modified convolutional neural networks as classifiers by generating multiple hypotheses. The classification error achieved by this approach was highly comparable with non-incremental learning. In[16], Incremental learning inspired by Learn++ algorithm is based on ensemble of Convolutional Neural Network classifiers. Algorithm's update rule is optimized for incremental learning of new data. It does not require access to previously seen data during subsequent

* Corresponding Author

0-7695-6341-4/17/31.00 ©2017 IEEE
DOI 10.1109/CIS.2017.00024

training process and it is able to retain previously acquired knowledge.

Nowadays, there are three famous networks, namely, CaffeNet [6], GoogLeNet [7], and VGGNet [8]. Those three models opened the boom in deep learning. And it made proud achievements on ImageNet. In order to further reduces the running time detection network, Microsoft Shaoqing Ren, etal in[9], the latest target detection method is proposed: Faster-RCNN[14].They designed a region proposal network(RPN) to generate a recommendation of region proposals. The emergence of RPN replaced Selective Search [10] and Edge Boxes [11] before. It shares the full convolution features of the detection network makes proposals spending little time. But it is insensitive to small target detection problem. In[12] the convolutional neural network is improved and the unified multi-scale convolution neural network is proposed. The MS-CNN consists of a proposal sub-network and a detection sub-network. The results show that the accuracy is improved obviously, and has made great savings in memory and calculation. But it demands for high equipment performance.

In order to solve the current problems existing in the vehicle detection research, combines the actual conditions of our current research , this paper proposes a new method of forward vehicle detection, through the image preprocessing and incremental learning, to optimize the whole process of training. The network structure is based on Fast R- CNN. It adjusts the training parameters Constantly in the process of training to achieve the optimal state. Finally the experiment tests the test data set several times, then chooses the best test results. Experimental results show that the performance is improved compared with the published articles of vehicle detection method. It provides more generalization and simple solution for the forward vehicle detection problem.

The remainder of this paper is organized as follows. Section II describes the forward vehicle detection method in detail. Section III presents the experimental results and analysis. Finally, in Section IV, conclusion is drawn and it has carried on a brief summary and outlook.

II. METHODOLOGY

A.The whole process of forward vehicle detection

This network model is divided into two phases, training and testing phase. At training phase, the convolution neural network of initial parameters after pre-training on ImageNet is re-trained to obtain the vehicle detection model; At test phase, the test samples are input into detection model to obtain the test result. There are three kinds of network in the process of the pretraining: CaffeNet, VGG_Cnn_M_1024, VGG-16. All of those are the five Maximum pooling layers and the convolution layers ranging from 5 to 13. This paper chooses CaffeNet to initialize the Fast R-CNN, but needs to modify the three places:

(1) the last pooling layer is RoI pooling layer;

(2) the last fully connection layer and softmax layer are replaced by two parallel layers;

(3) the network input two sets of data: a set of images and the images of a group of RoIs.

The whole calculation process is shown in Figure 1:

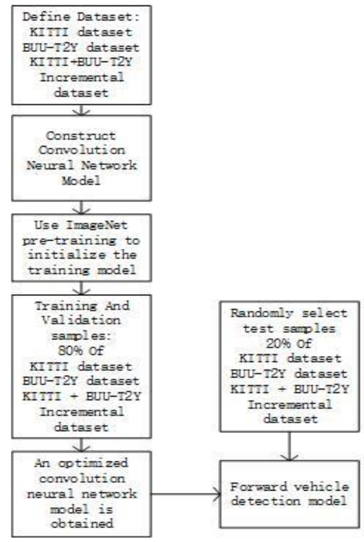

Figure 1. The whole process of forward vehicle detection

B. Datasets

This paper chooses the KITTI public data set of cars derived from the KITTI Vision Benchmark Suite and BUU - T2Y data set collected by Beijing Key Laboratory of Information Service Engineering of Beijing union university. Selected the KITTI data set as the training validation set and test set. In order to research the generalization ability on vehicle detection, the method expands BUU - T2Y data set as the training validation set and test set.

KITTI data set's and BUU - T2Y data set's part of data, respectively, are shown in Figure 2, Figure 3:

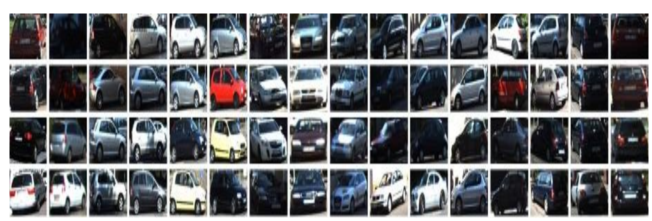

Figure 2.Part of images in KITTI

Figure 3.Part of images in BUU-T2Y

C. Vehicle training process

In this paper, the image is pre-processed and the region of interest is obtained. The network model obtained by training ImageNet is used to initialize all the layers before the RoI layer in the network model. The training samples are subjected to secondary iteration training to obtain the vehicle detection model. The last convolution layer adds an ROI pooling layer. ROI pooling layer can firstly locate the ROI to the feature map in the image, and then use a single layer of the SPP layer to this feature map patch poolling into a fixed size of the feature and then passed into the full connection layer. The loss function uses the multi-task loss function and adds the bounding box regression to the CNN network directly. And the bounding box regression is used to obtain the corrected position coordinates information on detecting the vehicle object. Softmax regression is used to calculation of two types of targets'(vehicle class and background class) probability estimation.

D. Vehicle detection process

Use a trained model to test. Firstly, randomly selecting 20% KITTI data as a test set to test the model ,adjust the parameters and get the best test results. Then, apply the trained model to the BUU-T2Y dataset to test. Finally, the BUU-T2Y data set and the KITTI data set are randomly combined with incremental testing.

III. EXPERIMENTS AND RESULTS ANALYSIS

A. Experiment Platform and Experimental setting

The task test data 1 is derived from the KITTI data set co-founded by the Karlsruhe Institute of Technology and the Toyota American Institute of Technology. It is the largest computerized visual algorithm evaluation data set in the world. The amount of data onto this data set is large. The KITTI data set's data acquisition platform is equipped with two grayscale cameras, two color cameras. Task test data 2 is obtained by Beijing Key Laboratory of Information Service Engineering of Beijing union university in Beijing city road.The BUU-T2Y dataset's data acquisition platform is equipped with a color monocular camera.

The forward vehicle detection includes: small passenger cars, cars, sport utility vehicles, light trucks, large passenger cars, heavy trucks and other common models. Most of the deep learning methods of data sets are processed by generating a set of training sets, validation sets, and test sets at a certain scale. In this experiment, the KITTI dataset was randomly generated with a training validation set and test sets at a ratio of 8: 2. In training validation set, 80% was used as the training data set and the remaining 20% was used as the validation data set. So as the BUU-T2Y dataset. The whole experiment were based on the deep learning framework Caffe, and the results were obtained by several experiments.

The composition and number of training validation set and test set are shown in TABLE I and TABLE II:

TABLE I. SAMPLE COMPOSITION

KITTI Dataset	*Proportion*	*Way*
Training Validation set	80%	Random selection (where the composition of the training set and the verification set is 8: 2)
Test set	20%	Random selection

TABLE II. NUMBER OF SAMPLES

BUU-T2Y Dataset	*Positive samples*	*Negative samples*
Training Validation set	14000	30000
Test set	3500	7500

B. Results and Analysis

The evaluation criteria are as follows (1) (2):

$$DR = \frac{TP}{TP + FP + FN} \quad (1)$$

$$FPR = \frac{FP}{TP + FP} \quad (2)$$

The experimental results are shown in TABLE III:

TABLE III. KITTI AND BUU-T2Y TEST SET TEST RESULTS

Method	*DR(%)*	*FPR(%)*
KITTI->KITTI	84.9	5.8
KITTI->BUU-T2Y	80.5	11.2
BUU-T2Y->BUU-T2Y	69.6	28.2
KITTI+BUU-T2Y->BUU-T2Y	86.2	5.4
Optimal Result	86.2%	5.4%

The experimental results are shown in Figure 4:

Figure 4. BUU-T2Y dataset test result

IV. CONCLUSION

In this section, conclusion is drawn: (1) The experimental results showed that if the KITTI data set was used to train the model, and then the BUU-T2Y dataset was used to adjust to the optimization model. Finally, the BUU-T2Y data set was added to the KITTI dataset to form a new data set for incremental learning. The experimental result can reach 86.2%. The optimal result were selected from several times experiments. It showed the proposed method is effective. To a large extent This method solved the problem of missed detection in the traditional vehicle target detection model. (2) However, compared to other models, it was not the end-to-end network structure. And it was not real time and spent much time. (3) Under the effective vehicle detection accuracy, to solve the above problems is the key issues in the future.

ACKNOWLEDGMENT

The work was supported by the Project of the Major Programs, funded by National Natural Science Foundation of China (Grant No. 91420202). We would like to acknowledge all of our team members, whose contributions were essential for our autonomous vehicle. We also would like to thank Professor Hong Bao* and teacher Nan Ma for their guidance and help in the process of writing our paper.

REFERENCES

[1] Zhang Quanfa, et al. "Engineering vehicle inspection based on HOG feature and machine learning." *Computer system application* ,2013:104-107.

[2] Zhang, Yunzhou, et al. "Real-time vehicle detection in highway based on improved Adaboost and image segmentation." *IEEE International Conference on Cyber Technology in Automation, Control, and Intelligent Systems* IEEE, 2015:2006-2011.

[3] Qiu, Qin Jun, L. Yong, and D. W. Cai. "Vehicle detection based on LBP features of the Haar-like Characteristics." *Intelligent Control and Automation* IEEE, 2015:1225-7.

[4] Girshick, Ross, et al. "Rich Feature Hierarchies for Accurate Object Detection and Semantic Segmentation." *IEEE Conference on Computer Vision and Pattern Recognition* IEEE Computer Society, 2014:580-587.

[5] Girshick,Ross."Fast R-CNN." *IEEE International Conference on Computer Vision* IEEE Computer Society, 2015:1440-1448.

[6] Jia, Yangqing, et al. "Caffe:Convolutional Architecture for Fast Feature Embedding." (2014):675-678.

[7] Feyzabadi, Seyedshams. "Joint Deep Learning for Car Detection." *Computer Science* (2016).

[8] Simonyan,Karen,andA. Zisserman."Very Deep Convolutional Networks for Large-Scale Image Recognition." *Computer Science* (2014).

[9] Ren, S., et al. "Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks." *IEEE Transactions on Pattern Analysis & Machine Intelligence* 39.6(2015): 1137-1149.

[10] Uijlings, J. R., et al."Selective Search for Object Recognition." *International Journal of Computer Vision* 104.2(2013):154-171.

[11] Zitnick, C. Lawrence, and P. Dollár. "Edge Boxes: Locating Object Proposals from Edges." *European Conference on Computer Vision*Springer, Cham, 2014:391-405.

[12] Cai, Zhaowei, et al. "A Unified Multi-scale Deep Convolutional Neural Network for Fast Object Detection." (2016):354-370.

[13] He, K., et al. "Spatial Pyramid Pooling in Deep Convolutional Networks for Visual Recognition. " *IEEE Transactions on Pattern Analysis & Machine Intelligence* 37.9(2015):1904-16.

[14] Sang Jun, et al. "Faster-RCNN model identification analysis." *Journal of Chongqing University7*(2017).

[15] Medera, Dusan, and S. Babinec. "Incremental Learning of Convolutional Neural Networks." *Ijcci 2009 - Proceedings of the International Joint Conference on Computational Intelligence, Funchal, Madeira, Portugal, October* DBLP, 2009:547-550.

[16] Polikar, R, et al. "Learn++: an incremental learning algorithm for supervised neural networks." *Systems Man & Cybernetics Part C Applications & Reviews IEEE Transactions on* 31.4 (2001):497-508.