2017 2nd International Conference on Artificial Intelligence and Engineering Applications (AIEA 2017)
ISBN: 978-1-60595-485-1

Vehicle Classification with Convolutional Neural Network on Motion Blurred Images

QIAOJIN GUO, ZHONGYAN LIANG and JIE HU

ABSTRACT

Vehicle classification is an important research area of Intelligent Traffic System. We first collect large set of static and clear vehicle images from internet and split the images into train and validation sets. Then we train the dataset with convolutional neural network (CNN). Experimental results show that CNN achieve high accuracy on the validation dataset. However, the trained model achieves poor performance on motion blurred images captured from videos. This paper proposes a new method for dealing with motion blurred images. Random blurred images are generated during training in order to optimize the network parameters. The final experimental results show that our proposed method achieves better performance than training directly with CNN.

KEYWORDS

CNN, Vehicle Classification, Motion Blur, GoogleNet.

INTRODUCTION

Vehicle classification and identification, also known as Vehicle Make and Model Recognition (MMR), is an important technique for Intelligent Traffic System (ITS), which detects vehicles from images and videos and classify vehicles into different types or brands. With the rapidly increment of vehicles, more and more sensors and computers are employed to monitor traffic conditions. Automatic vehicle classification is an efficient technique for secure access, traffic monitoring applications, accidents prevention and terrorist activities inspection etc. Traditional vehicle identification methods are based on Automatic License Plate Recognition (ALPR). However, ALPR needs sophisticated image capture hardware to produce high quality photographs, and is sensitive to light, dust and occlusion. Vehicle classification is a efficient complement to ALPR, which is able to recognize and identify vehicles with cheap image/video capture equipment.

Qiaojin Guo. guoqiaojin@163.com; Zhongyan Liang. aptivaleung@foxmail.com;
Jie Hu hujieyc@139.com.
Nanjing Research Institute of Electronic Engineering, Nanjing 210000, China.

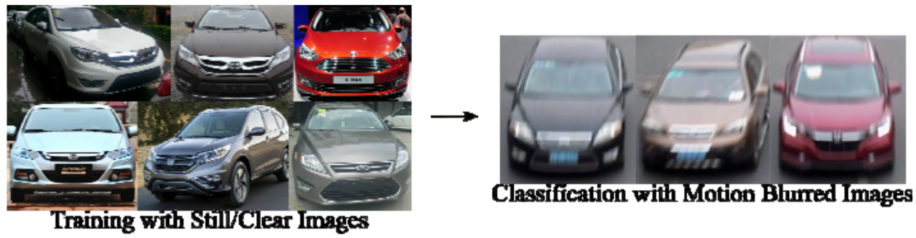

Figure 1. Example images of vehicle classification dataset and motion blurred vehicle images captured from video.

Various of vehicle classification methods have been proposed in the last few years [1, 2, 3, 4]. Petrovi and Cootes [5] classify vehicles into different classes by using the number plate to locate region of interest (ROI) and then compare the extracted features of ROI between test image and training set. Emami et al. [6] detect ROIs based on license plate localization and divide ROI into 4 parts to improve recognition rate. Chen et al. [7] use SVM to determine vehicle types (car, van , HGV). Huang and Liao [8] extract vehicle regions from video frames and then classify side images of vehicles into 7 classes, including sedan, van, pickup, truck, van truck, bus, and trailer.

Recently, convolutional neural network has shown great power in different image classification tasks, and the classification performance exceeds studies that use manually engineered feature extraction pipelines. Gao and Lee [9] detect moving cars from video with background subtraction and identify the detected ROI with three layers of restricted Boltzmann machines. Huttunen et al. [10,11] employ deep learning to classify vehicle images into 4 classes, include Bus, Truck, Van and Small car.

Most of the previous researchers do vehicle classification with static images or non-blurred video frames. However, most of the vehicles captured within video frames are blurred and hard to distinguish, As shown in Fig. 1. In this paper, we propose a new vehicle classification method based convolution neural network, which classifies detected vehicles into 66 brands and is able to deal with motion blurred images.

The rest of this paper is organized as follows. In the next section, we first describe the dataset used in this paper and evaluate the classification accuracy by utilizing CNN. In Section 3, we describe the proposed method in detail, which performs better on small and blur images. Experimental results are provided in Section 4.

DATASET COLLECTION AND CLASSIFICATION WITH CNN

The dataset used in this paper is collected from internet, the initial dataset contains 184226 images of 121 categories. We first employ SSD [12] to detect vehicles from these images and keep the largest detected region for each image. Then we need to remove images which are captured from side because of the purpose of our work is to recognize motion blurred vehicle images as shown in Fig. 1. In order to do this, we manually collect 1000 vehicle images captured from side and another 1000 images captured from the front and the tail of the car.

Then we train a GoogleNet on this dataset and do classification on the entire image set. After manual verification, we remove the classes which contain less than 20 images. The final dataset contains 40749 images, 66 categories.

Figure 2. The original vehicle image and motion blurred images convolved with kernels of different size.

Deep neural networks provide an end-to-end way to do image classification, which is easy to use and also achieves better performance than utilizing traditional manually designed features. In this paper, we use GoogleNet [13] to do vehicle classification, which contains 22 layers, and reaches the state-of-the art of the ImageNet Large-Scale Visual Recognition Challenge 2014 (ILSVRC 2014). We use Caffe [14] platform to propose dataset, train GoogleNet and evaluate the performance.

We random split the dataset and use 80% for training and left 20% for evaluation. GoogleNet achieves 94.99% on the evaluation dataset. However, most of the vehicle images captured from videos are motion blurred. In order to evaluate the classification performance on blurred images, we convolve the evaluation dataset with manually designed motion blur kernels, as shown in Fig. 2. The input images are resized to 256x256 and blurred with kernel size of 5, 11 and 21. Experimental results show that the model trained on the initial dataset achieves very poor performance as shown in Table 2. We also resize the evaluation images to different sizes (64x64, 128x128, 256x256) and compare the classification accuracies. Experimental results show that model trained with images of 256x256 achieves poor performance on small input images as shown in Table 3.

The most probable cause is that the previous model was trained with clear, still and large images. In order to improve the classification accuracy with small and blurred images, the most immediate way is to train CNN on more small and blurred images. However, directly extend the training dataset is time consuming and will cost large size of disk space. Meanwhile, larger dataset consumes more time to optimize the model. In this paper, we modify the data layer of CNN to do random resize and blur on initial images, which costs the same disk space and similar time comparing with training directly on the initial dataset. The details will be described in the next session.

TRAINING WITH MOTION BLURRED IMAGES

During the training process of CNN, batches of images are loaded from disk and feed to following convolutional layers. After images are loaded, the images will be subtracted by the mean image/value and random cropped to 224x224 to fit the following convolutional layer.

In order to train CNN with more small and blur images, We modify the process steps of input data layer as shown in Fig. 3. We first random resize the input image to different sizes in the range of 64x64 and 256x256 and then resize back to 256x256.

This will generate low resolution images which will be used to teach CNN to recognize vehicles with less detail information.

Then we perform random motion blur on the images. The motion blur kernel size is in the range of 1 to 21 pixels, and the blur angle is between -10 to 10 degrees. The input images keep unchanged with kernel size of 1×1. Finally, the result images are cropped to 224x224 and subtract the mean image and feed to the next convolutional layer. Thus, the final CNN will be trained with different images during each batch.

Random resize, random blur and random crop are only used during training. During classification, we use the original image and crop the central part for recognition.

EXPERIMENTS

In this paper, we use Caffe to perform data preprocessing, network training and evaluation. We modify Caffe to do random resize and blur to deal with motion blurred vehicle image classification. We first random split the dataset and use 80% for training and left 20% for evaluation. Then, we use GoogleNet as baseline network to verify the performance improvement of our method. We compare the classification accuracies of GoogleNet trained with and without random resized/blurred data. The experimental results is shown in Table 1. GoogleNet trained with original dataset achieves 94.99% on the evaluation set.

We also evaluate the classification performance on dataset with different size. Experimental results show that standard GoogleNet training with original images achieves poor performance on small input images, the accuracy on evaluation dataset is only 9.5% with input image of 64x64. Our method achieves better performance on smaller input images. The detail experimental results is shown in Table 2.

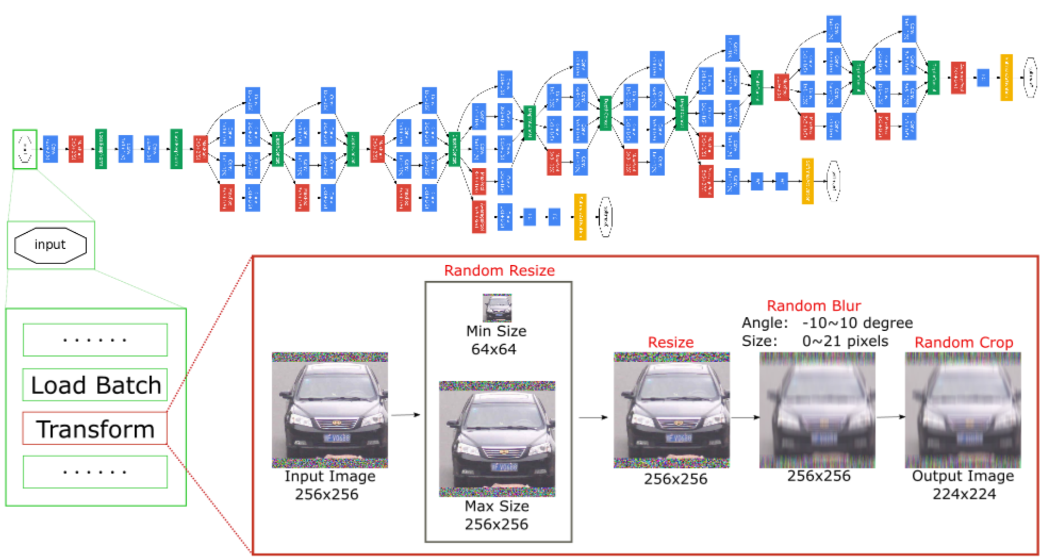

Figure 3. The training flow of proposed method.

TABLE 1. CLASSIFICATION ACCURACIES ON EVALUATION DATASET BLURRED WITH DIFFERENT SIZE OF MOTION KERNELS.

Blur Kernel Size	GoogleNet	GoogleNet with Blur
0	94.99%	98.37%
5	38.77%	85.22%
11	2.1%	84.98%
21	0.93%	86.5%

TABLE 2. CLASSIFICATION ACCURACIES ON EVALUATION DATASET OF DIFFERENT INPUT SIZE.

Input Image Size	GoogleNet	GoogleNet with Blur
64	9.55%	91.85%
128	66.71%	98.02%
256	94.99%	98.37%

Figure 4. Classification performance comparison on motion blurred images captured from video. Column (a) shows the sample blurred images, column (b) shows the top 5 classification results of GoogleNet with blur, column (c) shows the classification results of standard GoogleNet.

We also evaluate the performance of our method with vehicle images captured from video streams. The input video is 720P and captured on the top of road to get front view of vehicles. We use SSD to perform real-time vehicle detections and the detected regions are filled to be square and resized to 256x256. Finally, the optimized CNN model is employed to do classification.

The result vehicle images are motion blurred and standard GoogleNet get poor performance on these images. Sample images and the classification results is shown in Fig. 4. Experimental results show that our method achieves better performance than standard GoogleNet.

CONCLUSION

This paper proposes a vehicle type classification based on convolutional neural network. We first collect large numbers of vehicle images from internet. Then we select and build a new vehicle classification dataset with 66 categories. We train the dataset with CNN. Experimental results show that CNN achieve high accuracy on the validation dataset. However, the trained model achieves poor performance on motion blurred images captured from videos. This paper proposes a new method for dealing with motion blurred images. Random blurred images are generated during training in order to optimize the network parameters. The final experimental results show that our proposed method achieves better performance than training directly with CNN.

REFERENCES

1. Abdelmaseeh, M., Badreldin, I., Abdelkader, M.F., et al. Car make and model recognition combining global and local cues. International Conference on Pattern Recognition. Tsukuba Science City, Japan, 2012, p. 910–913.
2. Pearce, G., Pears, N. Automatic make and model recognition from frontal images of cars. IEEE International Conference on Advanced Video and Signal-Based Surveillance. Klagenfurt, Austria, 2011, p. 373–378.
3. Sorber, L., Barel, M.V., Lathauwer, L.D. Detection and classification of vehicles from video using multiple time-spatial images. IEEE Transactions on Intelligent Transportation Systems. Vol. 13 (2012) No. 3, p. 1215–1225.
4. Yousaf, K., Iftikhar, A., Javed, A. Comparative analysis of automatic vehicle classification techniques: A survey. International Journal of Image Graphics and Signal Processing. Vol. 4 (2012) No. 9, p. 52–59.
5. Petrovi, V.S., Cootes, T.F. Analysis of features for rigid structure vehicle type recognition. British Machine Vision Conference. London, UK, 2004, p. 587–596.
6. Emami, H., Fathi, M., Raahemifar, K. Real time vehicle make and model recognition based on hierarchical classification. International Journal of Machine Learning and Computing. Vol. 4(2014) No. 2, p. 142–145.
7. Chen, Z., Pears, N., Freeman, M., et al. Road vehicle classification using support vector machines. IEEE International Conference on Intelligent Computing and Intelligent Systems. Shanghai, China, Cambridge, UK, 2009, p. 214–218.
8. Huang, C.L., Liao, W.C. A vision-based vehicle identification system. International Conference on Pattern Recognition. Cambridge, UK, 2004, p. 364–367.
9. Yongbin, G., Hyo Jong, L. Real time vehicle make and model recognition based on hierarchical classification. Advanced Science and Technology Letters Vol. 90(2015), p. 57–61.
10. Huttunen, H., Yancheshmeh, F.S., Chen, K. Car type recognition with deep neural networks. arXiv preprint arXiv:1602.07125 (2016)
11. Renjie, X., Heikki, H., Shuoxin, L., et al. Road vehicle classification using support vector machines. IEEE European Signal Processing Conference. Budapest, Hungary, 2016, p. 1862–1866.
12. Liu, W., Anguelov, D., Erhan, D., et al.: SSD: Single shot multibox detector. arXiv preprint arXiv: 1512.02325 (2015).
13. Szegedy, C., Liu, W., Jia, Y., et al. Going deeper with convolutions. International Conference on Computer Vision and Pattern Recognition. Ohio, USA, 2014, p. 1–9.
14. Jia, Y., Shelhamer, E., Donahue, J., et al. Caffe: Convolutional architecture for fast feature embedding. arXiv preprint arXiv: 1408.5093 (2014).