International Journal of Applied Engineering Research
ISSN 0973-4562 Volume 10, Number 11 (2015) pp. 27717-27723

http://www.ripublication.com

Car Manufacturer and Model Recognition based on Deep Network

Yongbin Gao

Division of Computer Science and Engineering,
Chonbuk National University
E-mail: gaoyongbin.sam@gmail.com

Hyo Jong Lee

Division of Computer Science and Engineering,
Center for Advanced Image and Information Technology,
Chonbuk National University, Corresponding author
E-mail: hlee@chonbuk.ac.kr.

Abstract

Vehicle analysis is an important task in many intelligent applications, such as automatic toll collection and driver assistance systems. Among these applications, moving car detection and model recognition is a challenging task due to the close appearance between car models. In this paper, we propose a framework to recognizethe car manufacturer and model based on deep network. We first detect the moving car using frame difference; the resultant binary image is used to detect the frontal view of a car by a symmetry filter. The detected frontal view is used to identify a car based on deep learning with three layers of restricted Boltzmann machines (RBMs). Experiment results show that our proposed framework achieves favorable recognition accuracy.

Keywords: Moving car detection, Car model recognition, Deep Learning

1 Introduction

Vehicle analysis is an essential component in many intelligent applications, such as automatic toll collection, driver assistance systems, self-guided vehicles, intelligent parking systems, and traffic statistics (vehicle count, speed, and flow). Specially, an electronic toll collection system can automatically collect tolls according to the identification of vehicle models. Also, the identification of vehicle models can provide valuable information to the police for searching suspect vehicles. The appearance of a vehicle will change under varying environmental conditions and market requirements. The shapes of vehicles between companies and models are very similar, which results in confusion in vehicle model recognition. This makes the vehicle model recognition a challenging task.

Vehicle detection is a prerequisite of vehicle analysis [2]. Background subtraction [3]–[5] is widely used to extract motion features to detect moving vehicles from videos. However, this motion feature is not available for still images. To address this problem, Wu et al. [6] proposed the use of wavelet transformation to extract texture features to locate possible vehicle candidates. Tzomakas and Seelen [7] found that the shadow of a vehicle is a good cue for detecting vehicles. Ratan et al. [8] localize the possible vehicles based on the detection of vehicle wheels and verify the candidate vehicle by a diverse density method.

There are various applications of vehicle analysis. Chen et al. [9] proposed to use SVM and random forests to classify vehicles on the road into four classes, namely, car, van, bus, and bicycle/motorcycle. Lee et al. [10] used histogram of oriented gradient and SVM to recognize the vehicle models from video. AbdelMaseeh et al. [11] used the combination of global and local cues to recognize car model. Hsieh et al. [12] proposed symmetrical SURF for both vehicle detection and model recognition. Considering the favourable performance of deep learning [13], we’ve applied it to car model recognition. This is the first attempt to use deep learning for car model recognition to our knowledge.

In this paper, we propose a framework to recognize car manufacturer and model based on deep network. We first detect the moving car using frame difference; the resultant binary image is used to detect the frontal view of a car by a symmetry filter. The detected frontal view is used to identify a car based on deep learning. It is noted that frontal view suffers less variances than a full car, thus, we prefer to use frontal view instead of full car in this paper.

The remainder of this paper is organized as follows. Section 2 describes other works related to our paper with regard to neural network and back propagation. The moving car detection and frontal view extraction method are introduced in Section 3. We then introduce the car model recognition based on deep learning in Section 4. Section 5 applies the above algorithm to our car database, and presents the experiment results. Finally, we conclude this paper in Section 6.

2 Related Work

Multilayer neural network is able to learn nonlinearity of the input data, which is powerful in various applications with machine learning task. Neural network consists of neurons and adaptive weights, which mimics a biological neural network capable of approximating nonlinear functions of input data. Fig.1 shows the four layers of neural network, that is, input layer, hidden layers, and output layer. For the training process, feedforward is first conducted from bottom to top to get error signal of outputs and labels. Conversely, the error signal is back propagated from top to bottom to get derivatives of weights.

(1) Feedforward: considering three layers of neural network with *c* output units, one for each of the categories. As a result, each output unit of feedforward from input signal is a discriminant function $g_k(x)$ as follows:

$$g_k(x) = f(\sum_{j=1}^{N_h} w_{kj} f\left(\sum_{i=1}^{d} w_{ji} x_i + w_{j0}\right) + w_{k0}) \quad (1)$$

where N_h and d are the number of hidden units and dimension of input data, respectively, w_{ji} are the weights between input layer and hidden layer, w_{kj} are weights between hidden layer and output layer, w_{j0} and w_{k0} are the corresponding bias. $f(.)$ is a transfer function with nonlinearity property.

(2) Back propagation: training process is based on back propagation with input data and desired output. Error signal is achieved by comparing the feedforward output and desired output as follows:

$$J(w) = 1/2(t - z)^2 \quad (2)$$

where *t* and *z* are the target and output vectors, *w* represents all weights in the network. The aiming of back propagation is to minimize the classification of training data. Gradient descent is the natural selection to achieve minimum error. The weights are initialized with random values, and updated in a way that reduces the error using gradient as follows:

$$\Delta \mathrm{w} = -\eta \frac{\partial J}{\partial w} \quad (3)$$

where η is the learning rate that indicates the change rate in weights.

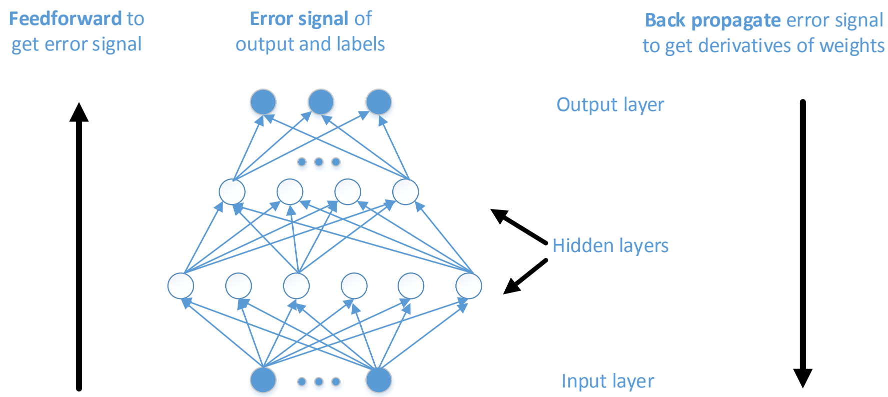

Fig 1. Framework of training procedure of neural network using back propagation.

3Moving Car Detection and Frontal View Extraction

Moving car detection is the prerequisite of car analysis. In this paper, we use frame difference to detect the moving car on account of the fact that our camera is fixed on the street. Frame difference is simple and sufficient in this scenario, which enables real time application.

It is not intuitive to recognize a car by its whole appearance due to its 3D characteristic. The 3D object results in large variance within one class incurred by pose changes. In this paper, we proposed to use the frontal view of a car to recognize the car model. The frontal view of a car offers sufficient features to represent a car model. Also, a portion of the car image has a reduced computation time than an entire car image. Furthermore, the frontal view of a car is basically symmetrical, thus, we use a symmetrical filter to detect the frontal view. After the binary image is calculated from the frame difference, a symmetrical filter is used to extract the symmetrical region of the binary image. As a result, the symmetrical region is regarded as a frontal view of a car.

4Car model recognition based on deep learning

Deep architectureis capable of learning different layers of representations of the input data. Traditional back-propagation method to learn a deep architecture suffers from the poor local optima problem as well as the long learning time. Also, labelled training data is necessary for back-propagation, which is not always satisfied in casesof small dataset. Deep learning method [13] provides a solution to address all these problems, which enables us to use unlabelled data to initialize the deep network. The idea behind the deep learning method is to learn $p(image)$ instead of $p(label|image)$, which, in another words, is to maximize the probability that a generative model would have produced the input data.

As for our car model recognition, we use three layers RBMs as shown in Fig. 2. The binary image of frontal part of a car with size $40{\times}50$ is used as input. This binary image is unrolled into a vector with dimension of 2000. The following three RBM layers are used as pre-training to obtain an initial weight. After the pre-training by RBM, we use traditional back-propagation method to fine-tuning the deep network using the labels of 107 car models.

5Results

To evaluate the performance of our proposed framework, we built a car database, which consists of 3210 car images with varying companies and models. These images composed of 107 car models with 30 images for each model. We are apt to use image instead of video, because it is easy to measure the accuracy. However, to use the frame difference for images, we shifted each image with 10 pixels to generate a neighbouring image. The difference image was generated by an image and its shifted image.

Fig. 2 shows the frontal view detection results, which includes four car models with four companies. The leftcolumn shows the original car image with detected frontal region marked by red box. The rightcolumn shows the binary representation of frontal view. Our system is able to detect the frontal view of each car image accurately. We test the detection algorithm on all 3210 images in the system, and get 100% accuracy with regard to the detection accuracy. Thus, the detection algorithm is fast and effective in our system.

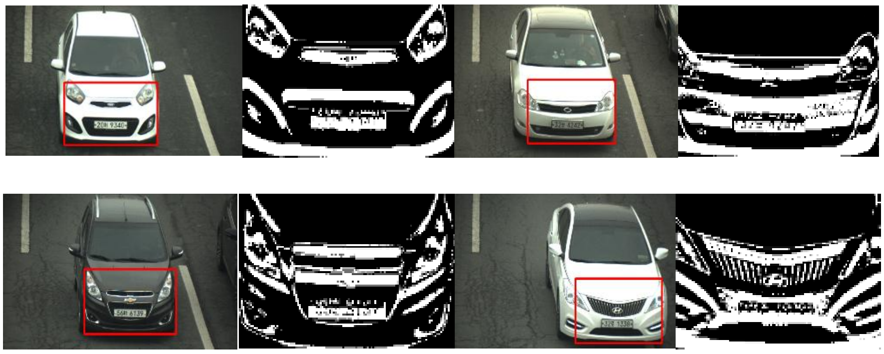

Fig 2. Results of frontal view extraction on five car images with four companies and four models.

Once the frontal view of a car was obtained, we fed it into trained deep model as input; the output label is used to recognize the car model. We compared our deep learning method to the prestigious features as follows: local binary pattern (LBP) [14], local Gabor binary patterns (LGBP) [15], and scale-invariant feature transform (SIFT) [1].

(1) LBP: the LBP operator is widely used in various applications as texture descriptor. It is invariant to monotonic changes and computational efficiency. In our experiment, the binary image of frontal view is transformed into a histogram vector of length 59, and weighted Chi square distance is used to measure the difference of two histogram of LBP.

(2) LGBP: LGBP is the extension of LBP incorporated with Gabor filter;a binary image is first filtered by Gabor and transformed to Gabor Magnitude Pictures (GMPs) in frequency domain. For each GMP, block LBP is performed to obtain LGBP maps; histograms of LGBP maps are concatenated into one vector.In our experiment, five scales and eight orientations are used for Gabor filters; as a result, 40 GMPs are generated for each image. Also, weighted Chi square distance is used to measure the difference of two histogram of LGBP.

(3) SIFT: SIFT is an effective local descriptor with scale, rotation invariant properties. Images are first transformed into scale invariant space to search extrema as candidate keypoints based on difference of Gaussian. These candidate keypoints are refined by excluding those localized along an edge. After that, one or multiple orientations and descriptors are assigned to each keypoints based on histogram of gradient. Training is not required in this algorithm, while 30 car images are collected for each car model. We use some

of them to compare with a test image, and a summation of number of matching keypoints is used to measure the similarity of two images.

The experimental results are shown in Table 1. In our experiments, we used29 images of each model for training, and theone left image for testing. The results show that deep learning can achieve impressive performance compared with other methods.

Table 1. Performance comparison of car model recognition with prestigious methods.

Algorithm	Accuracy (%)
LBP	46.0
LGBP	68.8
SIFT	78.3
Deep Learning	88.2

6Conclusion

In this paper, we proposeda framework to recognize car manufacturer and model based on deep network. We first detected the moving car using frame difference; the resultant binary image is used to detect the frontal view of a car by a symmetry filter. The detected frontal view is used to identify a car based on deep learning. Experiment results showed that our proposed framework achieves favourable recognition accuracy.

Acknowledgments.

This work was supported by the Brain Korea 21 PLUS Project, National Research Foundation of Korea. This work was also supported by Business for Academic-industrial Cooperative establishments funded Korea Small and Medium Business Administration in 2014 (Grants No. C0221114).

References

[1] D. G. Lowe, "Distinctive image features from scale-invariant keypoints," Int. J. Comput. Vis., 60, 2, (2004), 91–110.

[2] W. Li, P. Liu, Y. Wang, H. Ni, C. Wen, and J. Fan, "On-board Robust Vehicle Detection Using Knowledge-based Features and Motion Trajectory,"International Journal of Signal Processing, Image Processing and Pattern Recognition, 8, 2, (2015), 201-212.

[3] A. Faro, D. Giordano, and C. Spampinato, "Adaptive background modeling integrated with luminosity sensors and occlusion processing for reliable vehicle detection," IEEE Trans. Intell. Transp. Syst., 12, 4, (2011), 1398–1412.

[4] A. Jazayeri, H.-Y. Cai, J.-Y. Zheng, and M. Tuceryan, "Vehicle detection and tracking in car video based on motion model," IEEE Trans. Intell. Transp. Syst., 12, 2, (2011), 583–595.

[5] G. L. Foresti, V. Murino, and C. Regazzoni, "Vehicle recognition and tracking from road image sequences," IEEE Trans. Veh. Technol., 48, 1, (1999), 301–318.

[6] J. Wu, X. Zhang, and J. Zhou, "Vehicle detection in static road images with PCA-and-wavelet-based classifier," in Proc. IEEE Conf. Intell. Transp. Syst., Aug. 25–29, (2001), 740–744.

[7] C. Tzomakas and W. Seelen, "Vehicle detection in traffic scenes using shadow," Inst. Neuroinf., Ruhtuniv., Bochum, Germany, Tech. Rep. 98-06, 1998.

[8] A. Lakshmi Ratan,W. E. L. Grimson, andW. M.Wells, "Object detection and localization by dynamic template warping," Int. J. Comput. Vis., 36, 2, (2000), 131–148.

[9] Z. Chen, T. Ellis, and S. A. Velastin, "Vehicle type categorization: A comparison of classification schemes," in Proc. 14th Int. IEEE Conf. Intell. Transp. Syst., (2011), 74–79.

[10] S. Lee, j. Gwak, M. Jeon, "Vehicle Model Recognition in Video," International Journal of Signal Processing, Image Processing and Pattern Recognition, 6, 2, (2013), 175-184.

[11] M. AbdelMaseeh, I. BadreIdin, M. F. Abdelkader and M. EI Saban, "Car Make and Model recognition combining global and local cues," in Proc. IEEE Int. Conf. Pattern Recognition, (2012), 910-913.

[12] J. W. Hsieh, L. C. Chen and D. Y. Chen, "Symmetrical SURF and Its Applications to Vehicle Detection and Vehicle Make and Model Recognition," IEEE Trans. Intell. Transp. Syst., 15, 1, (2014), 6-20.

[13] G. E. Hinton, R. R. Salakhutdinov, "Reducing the Dimensionality of Data with Neural Networks," Science, 313, 5786, (2006), 504-507.

[14] T.Ahonen, A. Hadid,M. Pietikainen,"Face Description with Local Binary Patterns: Application to Face Recognition," IEEE Transactions on Pattern Analysis and Machine Intelligence, 28, (2006), 2037-2041

[15] W. Zhang, S. Shan, W. Gao,X. Chen,H. Zhang, "Local Gabor Binary Pattern Histogram Sequence (LGBPHS): A Novel Non-Statistical Model for Face Representation and Recognition,"in Proc.IEEE International Conference on Computer Vision (ICCV), (2005), 786-791.