Vehicle Type Classification Using Convolutional Neural Network

Kshitij Kaushik, Kunal Mali, Shrikant Dangi, Tushar Dalal, Pallavi Suradkar

[1] *Student, Computer Department, NBNSSOE, Maharashtra, India*
[2] *Student, Computer Department, NBNSSOE, Maharashtra, India*
[3] *Student, Computer Department, NBNSSOE, Maharashtra, India*
[4] *Student, Computer Department, NBNSSOE, Maharashtra, India*
[5] *Professor, Computer Department, NBNSSOE, Maharashtra, India*

ABSTRACT

Vehicle classifier has important roles in intelligent traffic management and transportation. Identifying one vehicle type from another is a difficult task. Vehicle classification problem is essentially image classification problem for which Convolutional Neural Networks are best to work with as CNN require little image preprocessing as compared to other approaches. Convolutional neural network does not require handcrafted features for classification as opposed to traditional machine learning approaches. In this paper we use supervised approach to train Convolutional Neural Network. The Neural network output a probability vector for identifying vehicle type. A vehicle classifier is very useful in intelligent traffic management. It can classify an Emergency Vehicle like an Ambulance or a Fire Truck from other vehicle types and this information can be used to intelligently change traffic flow as needed.

Keywords: - *Convolutional Neural Network, Vehicle Classifier*

---

1. Introduction

Among the many long-distance traffic detection systems, vision-based systems have received much attention. Convolutional neural networks have been vastly used to track and detect vehicles. This is partly due to cost reductions for cameras and other relevant hardware costs [1]. The installation of camera is not at all intrusive as compared to incorporation of sensors on the road to measure the weight of the vehicle and to classify the dimensions of the vehicle [2]. Other methods for detecting and classifying vehicles are the acoustic signature, the radar signal, the frequency signal [4]. Vision-based systems are discrete and such systems could obtain much more detailed traffic information other than just the vehicle types [2]. In the recent past, deep learning technology has been widely used to detect faces and pedestrians [5]. If in-depth learning can be successfully applied to the tracking and detection of vehicles in the natural science, it will be of great value to build the intelligent traffic system and the driverless system. Intelligent transportation systems include intelligent vehicles, intelligent highway systems and intelligent drivers [5]. Convolutional neural networks are multilayer feed-forward neural networks which are inspired by biological structure of brain and these networks can learn multiple stages of never changing features [7].

2. Architecture

The architecture of convolutional neural network is described in Fig.1. The network contains two stages. Both of these layers has convolutional, absolute value rectification, batch normalization, average pooling and subsampling layers [7].

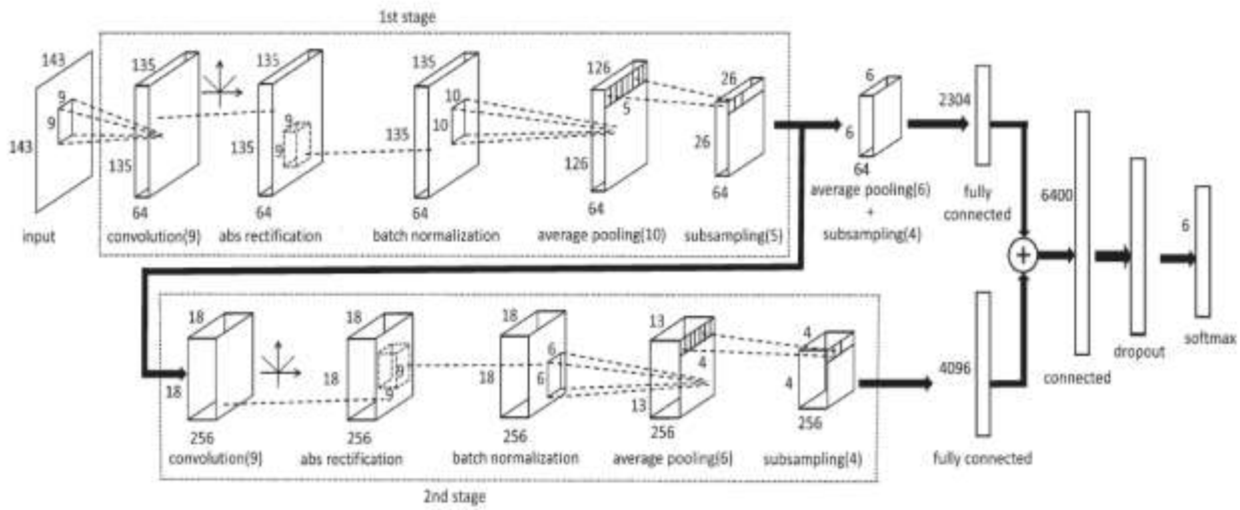

Figure-1: Convolutional Neural Network Architecture

2.1 Convolutional Layer

This layer computes convolutions between input image and filter. Sigmoid function is used as the activation function for this layer. Convolution computed with each filter provides a mapping from input to output where each neuron in output looks at a small region in input. Consider input is x1 * x2 * x3 and output is y1 * y2 * y3. Size of input 2D feature map (x1* x2) and output 2D feature map (y1*y2) is represented by x_i and y_j. Activation function is Sig () then y_j is calculated as

$$y_j = sig\left(\sum_i k_{ij} \otimes x_i\right)$$

Where k_{ij} is filter learned using supervised learning methods. If size of filter is s1 * s2 then size of output will be y1 = x1 -s1 + 1 and y2 = x2 - s2 + 1. As shown in Figure1 size of input feature map is 143 * 143 and size of filters is 9 * 9. This setup yields output feature maps of size 135 * 135 as shown in image [7].

2.2 Rectification Layer

This layer applies absolute value function to its input. Thus output becomes

$$Y_{i,j,k} = |X_{i,j,k}|.$$

Here $x_{i,j,k}$ are input elements and $y_{i,j,k}$ are output elements. Relation between two items in real world is always positive or zero, that's what is inspiration for this layer. This layer does not change size of input [6].

2.3 Normalization Layer

This layer performs batch normalization on input. Batch normalization reduces the internal covariate shift, so the next layer can expect the similar input distribution at each iteration. Batch normalization also has regularization effects which helps in training to prevent overfitting and also prevents the training to get stuck on saturated areas of nonlinearities.

2.4 Pooling and Subsampling Layer

This layer is used to provide robustness against small distortions, this plays the role of complex cells in visual perception. outputs

$$y_{i,j,k} = \sum_{p,q} w_{p,q} \cdot x_{i,j+p,k+q}$$

where $w_{p,q}$ is a uniform weighting windows [6]. Each output feature is subsampled by a factor S horizontally and vertically. Number of output and input feature maps are identical but the resolution is decreased. A filter size of 10 and subsampling rate of 5 in both direction will yield output size 26*26 [6].

2.2 Softmax Classifier Layer

A Softmax classifier is used to calculate probability of each vehicle type and is implemented as the output layer. The input to this layer is features learned by previous layer and the output is probability vector of each type. The mapping between input and output is linear and given by $v = w^T.x + b$ where x represents the input feature and v is a intermediate variable. w is $[w_1, w_2, w_3 \dots \dots . w_c]$ and each column of w is vehicle type parameter C is the total number of vehicles [7].

3.1 Experiment

We constructed a dataset with total of 1496 images. The training images in the Dataset are as follows: 55 bus images, 554 'not passenger vehicle' images, 430 images of sedans, 131 images of vans. The testing images in the Dataset are as follows: 14 images of bus, 139 images of 'not passenger vehicle', 140 images of sedans, 33 images of vans. The images in the Dataset varies in size so all the images are resized to 143 * 143 before feeding to neural network. There are many vehicle images which are occluded by other vehicles.

3.2 Results

The classifier achieves 88% accuracy. The classifier often misclassifies between images of bus and vans, this is mainly because similarity in appearances of the two. The classifier performs well even on the partially occluded vehicle images, this is because network is able to learn discriminative features from the training dataset.

Fig -2: classification Results

4. CONCLUSIONS

The filters of the network are learned via supervised learning. Softmax classifier is used as the classification layer. The Convolutional neural network takes vehicle images as input and outputs a probability vector. The classifier provided works effectively on the Dataset from the training dataset.

5. REFERENCES

[1]. Kangyan Wang, Rui Wang, Yutian Feng, Haiyan Zhang, Qunfeng Huang, Yanliang Jin, and Youzheng Zhang, “Vehicle recognition inacoustic sensor networks via sparse representation,” in IEEE International Conference on Multimedia and Expo Workshops (ICMEW). IEEE,2014, pp. 1–4.

[2]. Heong-tae Kim and Bongsob Song, “Vehicle recognition based on radar and vision sensor fusion for automatic emergency braking,” in 13th International Conference on Control, Automation and Systems (ICCAS). IEEE, 2013, pp. 1342–1346.

[3]. Zezhi Chen, Tim Ellis, and SA Velastin, “Vehicle detection, tracking and classification in urban traffic,” in 15th International IEEE Conference on Intelligent Transportation Systems (ITSC). IEEE, 2012, pp. 951–956.

[4]. Troy R McKay, Carl Salvaggio, Jason W Faulring, Philip S Salvaggio, aDonald M McKeown, Alfred J Garrett, David H Coleman, and Larry D Koffman, “Passive detection of vehicle loading.

[5]. Pradeep Kumar Mishra and Biplab Banerjee, “Multiple kernel based knn classifiers for vehicle classification,” International Journal of Computer Applications, vol. 71, no. 6, 2013.

[6]. Sayanan Sivaraman and Mohan M Trivedi, “Real-time vehicle detection using parts at intersections,” in 15th International IEEE Conference on Intelligent Transportation Systems (ITSC). IEEE, 2012, pp. 1519–1524.