IJCSE International Journal of Computer Sciences and Engineering Open Access
Research Paper Vol.-7, Issue-5, May 2019 E-ISSN: 2347-2693

Image-Based Vehicle Recognition using Neural Network

Md. Golam Moazzam[1], Mohammad Reduanul Haque[2], Mohammad Shorif Uddin[3*]

[1,3] Department of Computer Science and Engineering, Jahangirnagar University, Dhaka, Bangladesh
[2] Department of Computer Science and Engineering, Daffodil International University, Dhaka, Bangladesh

*Corresponding Author: shorifuddin@gmail.com

DOI: https://doi.org/10.26438/ijcse/v7i5.948954 | Available online at: www.ijcseonline.org

Accepted: 18/May/2019, Published: 31/May/2019

***Abstract*—** Vehicle recognition finds wide-spread applications in analyzing traffic data, collecting electronic tolls and identifying unauthorized vehicles on roads, etc. Diverse methods have been developed for vehicle recognition and these methods give good results in controlled environment. However, variations of illumination, vehicle geometry and occlusion are frequent phenomena in real-world scenarios. Neural network proves effective in handling such variations. In this paper, we have investigated the effectiveness of single-layer neural network, multi-layer neural network and convolutional neural network (CNN) and deep CNN for vehicle detection using a standard Madrid dataset.

***Keywords*—** Vehicle recognition, Neural network, Convolutional neural network, Deep network

I. INTRODUCTION

Computer vision system shows tremendous success in various object detection and recognition. Vehicle is a common object whose recognition is very important in many applications, such as localization of license plate, toll collection, computer assisted driving system to reduce accidents on the roads. Due to some difficulties, such as noise, variations in illumination and geometrical shapes, vehicle recognition process is somehow challenging. Diverse methods [1-9] have been proposed. However, these algorithms have some flaws, such as low recognition accuracy and computational efficiency. Recently, neural-based methods have been proposed to overcome the drawbacks of traditional approaches. This paper thoroughly investigates four types of neural network-based methods using a bench-marked dataset.

The rest of the paper is described as follows. The working methodology is given in section 2. Experimental results are shown in section 3. Finally, section 4 concludes the paper.

II. VEHICLE RECOGNITION TECHNIQUES

There are different neural approaches for vehicle recognition, such as single-layer neural network, multi-layer neural network, convolutional neural network (CNN) and deep CNN. A brief description of these techniques is given below.

A. Single Layer Neural Network

A single-layer neural network (SLNN) [10] consists of a single neuron with adjustable synaptic weights connected to input features and a thresholding activation function, such as step, bipolar and sigmoid functions. This neural network can be treated as a linear neural network, as it linearly classifies a two-class object. Usually, MSE (Minimum Squared Error) by steepest descent minimization procedure is used for training a single layer neural network.

B. Multi-Layer Neural Network

For solving non-linear recognition problems, we use multi-layer neural network (MLNN). Back propagation (BP) algorithm is used for training such a MLNN. In this neural network, the layers are fully connected, that is, every neuron in each layer is connected to every other neuron in the adjacent forward layer [11]. There are two sweeps for MLNN in BP algorithm for weight updating: forward sweep and backward sweep.

In the forward sweep, the input stimuli are given to the network, the network computes the weighted average from all the input units and then passes the average through an activation function. The activated outputs will go the intermediate (hidden) layers and finally, go to the output layer. In backward sweep, the error is calculated at the output layer and then that error backs to the hidden layers (prior to input layer).

C. Convolutional Neural Network (CNN) and Deep CNN

A convolutional neural network [12, 13] consists of basically three layers, which are described below.

i. Convolutional Layer

In convolutional layer, the convolution operation is executed to produce a feature map on the input data with the use of a filter or kernel. To calculate the convolution, the kernel is swept on the image and at every single location calculated the output. Thus, the convolutional layer just performs a two-dimensional image convolution operation using a weighted convolution kernel. Figure 1. describes a convolution operation of an image with a convolutional mask/kernel. A mathematical expression of convolution operation is given in Eq. (1).

The max-pooling converts a $k \times k$ region into a single value output, which is the maximum in that region. For instance, if the input layer has $N \times N$ regions, then the output will be $\frac{N}{k} \times \frac{N}{k}$ regions, that means each $k \times k$ block is converted to just a single value via the max-pooling operation.

iii. Fully Connected Layer

After performing operations on convolutional and max-pooling layers, the high-level reasoning for recognition is done in the fully connected layer. A fully connected layer has full connections with all neurons in the previous layer.

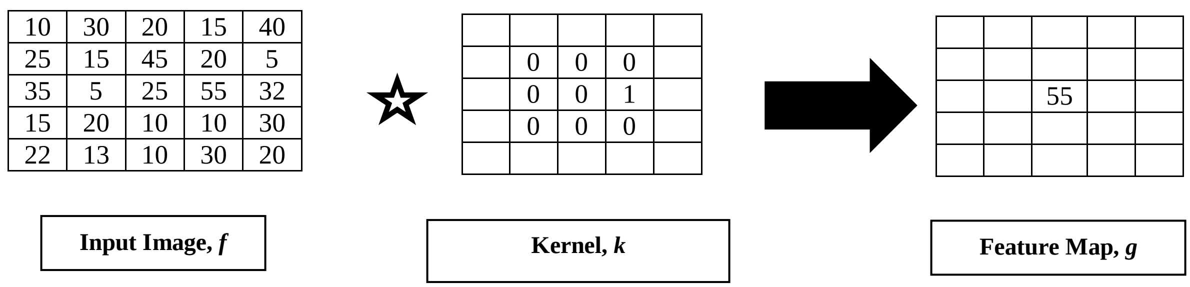

Figure 1. Convolution operation in CNN.

$$g(x,y) = f(x,y) \otimes h(x,y) = \sum_{k=-1}^{1}\sum_{j=-1}^{1} h(j,k)\, f(x-j, y-k) \qquad (1)$$

TABLE 1. A BRIEF ARCHITECTURES OF SLNN, MLNN, CNN AND DEEP CNN

	SLNN	MLNN	CNN	Deep CNN
Input Size	224×224	227×227	224×224	224×224
Conv. Layers	0	1	1	8
Filter Size	3	3	3	3
Stride	1, 2	1, 2	1, 2	2
FC Layer	1	1	1	3

ii. Pooling Layer

After performing a convolutional operation, we have a pooling layer. This layer takes input (feature map) from the convolutional layer and reduces it to a single output based on pooling kernel. The pooling operation is done for shortening training time and controlling the overfitting situation. Generally, 3 mechanisms are used for pooling operation, which are average or maximum (max) or minimum (min). In this research, max-pooling is used that takes the maximum of the block region.

In deep CNN, there are many layers (number of layers ≥ 3) [14] instead of a single convolutional-layer and a pooling-layer.

The architectures of SLNN, MLNN, CNN and deep CNN are shown in Table 1.

III. Experimental Results

To develop and test the experiment of our research, we have used The Polytechnic University of Madrid vehicle database [15].

The University of Madrid vehicle database consists of 7325 images of which 3425 images are vehicle (class 0) and 3900 images are road sequences (class 1) not containing vehicles. The vehicle images are basically the taken from rear and captured from different points of view. The images of the database consist of four different regions (according to the pose): middle/close range in front of the camera, left, middle and far respectively. In addition, the images are extracted in such a way that they do not perfectly fit the contour of the vehicle in order to make the classifier more robust to offsets in the hypothesis generation stage.

The images are recorded in highways of Madrid, Brussels and Turin with 64×64 and 360×256 pixels resolutions. Among the images of this Madrid database 6225 images are used for training and 1100 are used for testing. Figure 2 shows some sample vehicle images of Madrid database.

In deep CNN, we use input image size 64×64, convolution mask size 5×5, max pooling mask size 2×2, so in the fully connected layer the input size becomes 32×32. Tables 1 to 5 show the confusion matrix for SLNN, MLNN, CNN and deep CNN, respectively, for Madrid databases. From these tables we find that deep CNN shows the best recognition among the investigated methods. For clarity, accuracies of different methods are shown in Table 6. Figures 3 to 6 show the visual results of some sample images for SLNN, MLNN, CNN and deep CNN, respectively, for Madrid databases.

Figure 2. Sample images from Madrid database

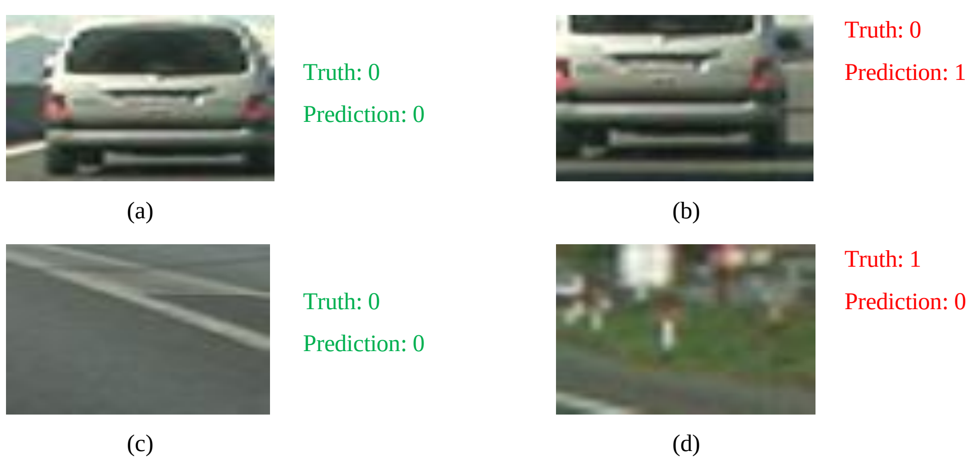

Figure 3. Recognition result of some sample vehicles of Madrid DB by using SLNN. The images of (a) and (c) are correctly recognized while (b) and (d) are not recognized correctly.

TABLE 2. CONFUSION MATRIX OF MADRID DB USING SLNN. RED COLORED FIGURES INDICATE TRUE RECOGNITION.

		Actual Class	
Predicted Class		Vehicle	Non-vehicle
	Vehicle	517	33
	Non-vehicle	44	506

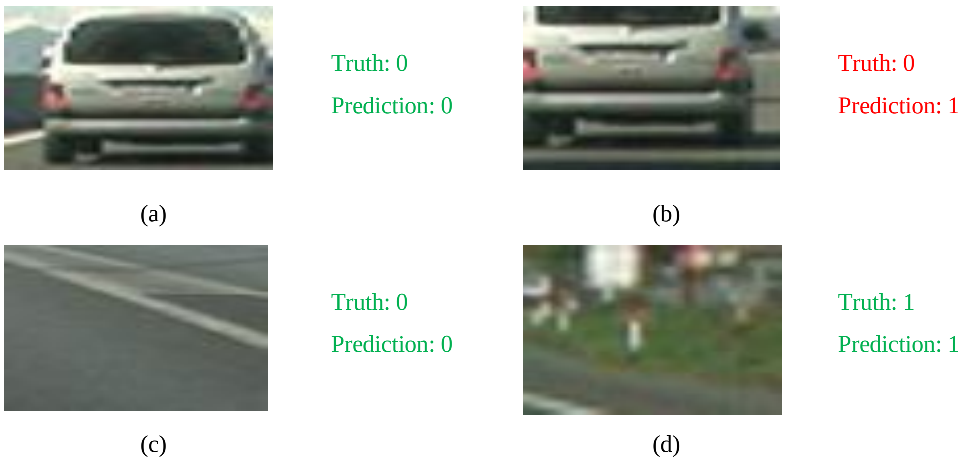

Figure 4. Recognition result of some sample vehicles of Madrid DB by using MLNN. The images of (a), (c) and (d) are correctly recognized while (b) is not recognized correctly.

TABLE 3. CONFUSION MATRIX OF MADRID DB USING MLNN NEURAL NETWORK. RED COLORED FIGURES INDICATE TRUE RECOGNITION

		Actual Class	
Predicted Class		Vehicle	Non-vehicle
	Vehicle	544	6
	Non-vehicle	28	522

TABLE 4. CONFUSION MATRIX OF MADRID DB USING CNN. RED COLORED FIGURES INDICATE TRUE RECOGNITION.

		Actual Class	
Predicted Class		Vehicle	Non-vehicle
	Vehicle	550	0
	Non-vehicle	14	536

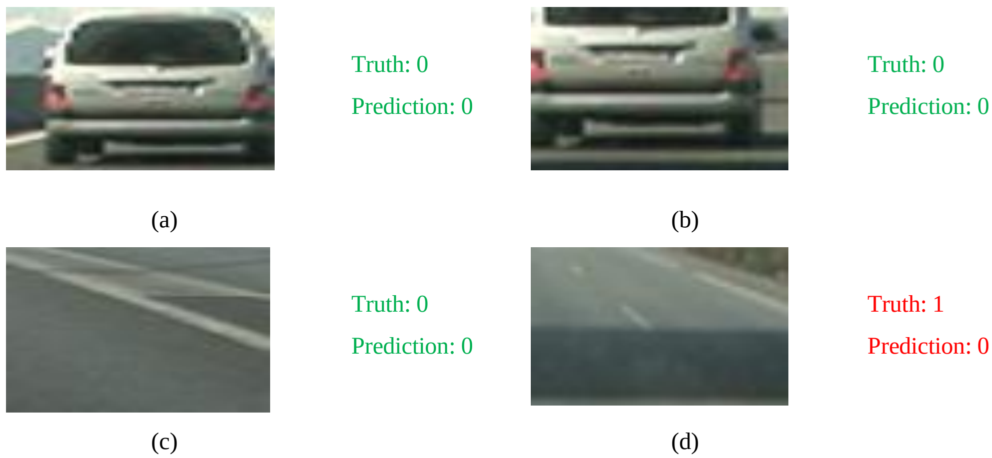

Figure 5. Recognition result of some sample vehicles of Madrid DB by using CNN. The images of (a), (b) and (c) are correctly recognized while (d) can't be recognized by CNN (Vehicle = 0 and Non-vehicle = 1).

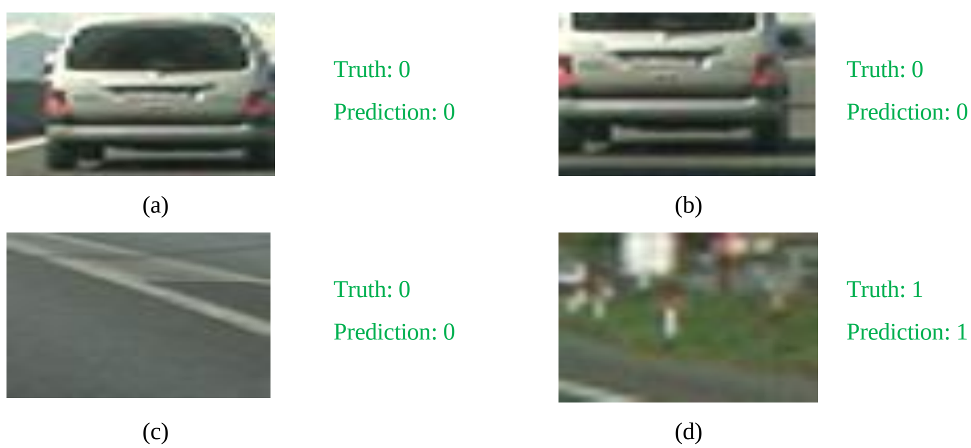

Figure 6. Recognition result of some sample vehicles of Madrid DB by using deep CNN. All images are correctly recognized by Deep CNN (Vehicle = 0 and Non-vehicle = 1).

TABLE 5. CONFUSION MATRIX OF MADRID DB USING DEEP CNN. RED COLORED FIGURES INDICATE TRUE RECOGNITION.

	Actual Class		
Predicted Class		Vehicle	Non-vehicle
	Vehicle	550	0
	Non-vehicle	6	544

TABLE 6. RECOGNITION ACCURACY UNDER DIFFERENT DATABASES AND METHODS

Database	Recognized by			
	Neural Network based Methods			
	SLNN	BPNN	CNN	Deep CNN
University of Madrid Dataset	86%	95.5%	97.5%	99.4%

IV. CONCLUSION

Vehicle recognition is a challenging task that has grabbed great attention for its overgrowing demands in a variety of applications in several domains recently. In this paper we have investigated the effectiveness of four neural network variations. Among these, Deep CNN shows the highest accuracy 99.4%.

REFERENCES

[1] Z. Chen, N. Pears, M. Freeman, and J. Austin, "*Road vehicle classification using Support Vector Machines*," IEEE International Conference on Intelligent Computing and Intelligent Systems, vol. 4, pp. 214–218, 2009.

[2] N. S. Tadayuki Okada, "*Development of preceding vehicle recognition algorithm for lead vehicle of autonomous platooning system based on multi sensor fusion and digital map*," SICE Annual Conference, Tokyo, Japan, September 13-18, 2011.

[3] G. Z. Sun and R. Miller, "*On-road vehicle detection using Gabor filters and support vector machines*," IEEE 14th International Conference on Digital Signal Processing, Santorini, Greece, 1-3 July 2002.

[4] Z. Xiao, B. Dai, T. Wu, L. Xiao, T. Chen, "*Dense Scene Flow Based Coarse-to-Fine Rigid Moving Object Detection for Autonomous Vehicle*, "IEEE Access, vol. 5, pp. 23492-23501, 2017.

[5] A. Petrovskaya and S. Thrun, "*Model based vehicle detection and tracking for autonomous urban driving*," Autonomous Robot, vol. 26, pp. 123-139, 2009.

[6] X. Zhang, N. Zheng, Y. He, F. Wang, "*Vehicle detection using an extended hidden random field model*," 14th IEEE International Conference on Intelligent Transportation Systems (ITSC), pp. 1555-1559, Washington DC, USA, 5-7 October 2011.

[7] F. M. Kazemi, "*Vehicle recognition using curvelet transform and SVM*, "IEEE 4th International Conference on Information Technology, pp. 516-521, Las Vegas, NV, USA, 2-4 April 2007.

[8] F. M. Kazemi, "*Vehicle recognition based on Fourier, wavelet and curvelet transforms - a comparative study*," IEEE Fourth International Conference on Information Technology, Las Vegas, NV, USA, 2-4 April 2007.

[9] W. Wang, "*Research on Sobel operator for vehicle recognition*," 1st IITA International Joint Conference on Artificial Intelligence, Hainan Island, China, 25-26 April 2009.

[10] Y. Freund, R. E. Schapire, *"Large margin classification using the perceptron algorithm,"* Machine Learning. ***37** (3): 277–296, 1999.*

[11] D. Gwala, V. Shrivastava and P. S. Parihar , "*Secure Images for Copyright by Watermarking in Transform Sphere using Bidirectional Neural Network*", International Journal of Scientific Research in Computer Science and Engineering, Vol.1 , Issue.5 , pp.21-25, Sep-2013 (ISSN: 2320-7639).

[12] D. He, C. Lang, S. Feng, X. Du, and C. Zhang, "*Vehicle detection and classification based on convolutional neural network*," 7th International Conference on Internet Multimedia Computing and Service, Hunan, China, August 19-21, 2015.

[13] N. S. Lele, "*Image Classification Using Convolutional Neural Network*", International Journal of Scientific Research in Computer Sciences and Engineering, Vol. 6, Issue 3 , pp.22-26, Jun-2018 (ISSN: 2320-7639).

[14] Alex Krizhevsky, I. Sutskever, G.E. Hinton, "*ImageNet classification with deep convolutional neural networks*," Twenty-sixth Annual Conference on Neural Information Processing Systems (NIPS), pp. 1097-1105, Nevada, USA, 2012.

[15] "*Vehicle image database*," Polytechnic University of Madrid, [Online]. Available: http://www.gti.ssr.upm.es/data/Vehicle_database.html

Authors Profile

Md. Golam Moazzam is pursuing Ph. D. in Computer Science and Engineering from Jahangirnagar University, Dhaka, Bangladesh. Currently, he is working in the area of Image Processing, Artificial Intelligence, Computer Vision, Deep Learning and so on.

Mohammad Reduanul Haque received his Master of Science and Bachelor of Science degrees in Computer Science and Engineering from Jahangirnagar University, Dhaka, Bangladesh in 2012 and 2011, respectively. He is currently served as a Lecturer in the Department of Computer Science and Engineering at Daffodil International University, Dhaka, Bangladesh. His research interests include Computer Vision, Deep Learning and Image Processing.

Mohammad Shorif Uddin received his Doctor of Engineering degree in information Science from Kyoto Institute of Technology in 2002, Japan, Master of Technology Education degree from Shiga University, Japan in 1999, Bachelor of Electrical and Electronic Engineering degree from Bangladesh University of Engineering and Technology in 1991 and also MBA in from Jahangirnagar University in 2013. He began his teaching career as a Lecturer in 1991 at the Bangladesh Institute of Technology, Chittagong (Renamed as CUET). He joined in the Department of Computer Science and Engineering of Jahangirnagar University in 1992 and currently, he is a Professor of this department. He undertook postdoctoral researches at Bioinformatics Institute, Singapore, Toyota Technological Institute, Japan and Kyoto Institute of Technology, Japan, Chiba University, Japan, Bonn University, Germany, Institute of Automation, Chinese Academy of Sciences, China. His research is motivated by applications in the fields of imaging informatics and computer vision. Mohammad Uddin is an IEEE Senior Member and a Fellow of Bangladesh Computer Society and The Institution of Engineers Bangladesh. He wrote more than 100 journal and received the Best Paper award in the International Conference on Informatics, Electronics \& Vision (ICIEV2013), Dhaka, Bangladesh and Best Presenter Award from the International Conference on Computer Vision and Graphics (ICCVG 2004), Warsaw, Poland. He holds two patents for his scientific inventions. Currently, he is the Chair of IEEE CS Bangladesh Chapter and an Associate Editor of IEEE Access.