Evolution in Electrical and Electronic Engineering Vol. 2 No. 2 (2021) 826-833
© Universiti Tun Hussein Onn Malaysia Publisher's Office

EEEE

Homepage: http://publisher.uthm.edu.my/periodicals/index.php/eeee
e-ISSN : 2756-8458

Deep Learning of Traffic Volume (Vehicle Recognition & Counting)

Iskandar Naqiuddin Mohd Hisham[1], Mohd Fadzli Abd Shaib[1]*

[1]Faculty of Electrical and Electronic Engineering,
University Tun Hussein Onn Malaysia, Batu Pahat, 86400, Johor, Malaysia

*Corresponding Author Designation

DOI: https://doi.org/10.30880/eeee.2021.02.02.099
Received 04 July 2021; Accepted 07 October 2021; Available online 30 October 2021

Abstract: Neural Network (NN) and Deep Learning are considered to learn abstract representations through their hierarchical architecture, which is inspired by the brain. Today, however, it is not widely understood how this happens. This is the technique through which a moving vehicle is recognized with a camera. Capturing automobile from the monitoring camera in video sequence is a challenging application to improve tracking performance. It increases the number of applications like traffic control, traffic monitoring, traffic flow, security etc. This technology. The expected costs will be much lower with this technique. These data provide three comparisons of the prominent method of recognition by vehicles: Local Binary Pattern (LBP) and Transform Scale-Invariant Features (SIFT) and for my research I'm using Neural Network (NN). For the Local Binary Pattern (LBP), the primary finding is 64.6%, for the Scale-Invariant Feature Transform (SIFT) 78.3%, and with the Deep Neural Network (DNN) 88.4 percent. In addition, the Convolutional Neural Network approach is compatible in the Neural Network method. This comparison is based on the compatibility, accuracy (percentage). These discoveries may also explain how the human brain learns abstract. Proof here that Neural Network (NN) learns abstract representations through a demodulation process. The activation function is introduced and it is used to indicate that the neural network (NN) is best studied and performed for demodulation.

Keywords: Deep Learning, Neural Network, Vehicle Recognition

1. Introduction

This Deep learning is an area of artificial intelligence that is concerned with emulating the approach to learning used by individuals to increase specific types of knowledge. Deep Learning can be thought of as an approach to mechanizing prescient investigation at its least complex. Although traditional AI calculations are straightforward, calculations of deep learning are stacked in a progression of multifaceted expansion of existence and reflection. Deep learning works when a similar procedure is experienced by PC programs that use deep learning [1].

**Corresponding author: fadzli@uthm.edu.my*

Capturing vehicle from a surveillance camera in a video sequence requires an application to increase the precision of monitoring. This technology has increased the number of applications such as traffic control, traffic management, traffic flow, security, etc. The estimated costs would be very low for the use of this technology. Compared to the manual method, there have been a few inconvenient which will cause errors such as counting errors and classification errors [2].

Next is that existing image processing techniques applied to traffic data collection rely solely on low- level image interpretation of sequences of images. Today, the video sequence capture of the vehicle from the surveillance camera is a challenging application to enhance the efficiency of monitoring [3].

The goal is to create an automated system that can accurately locate and monitor the total number of cars shown in aerial video frames.

In this project, we will develop Deep Learning algorithms to recognize vehicles, which will be used in the future. The goal of this project is to develop a system that can count vehicles using a deep learning algorithm. To assess the precision with which the proposed system's identification and classification are performed.

The comparison of parameter and recognition method is shown in Table 1 [4]. This table reveals that three methods have been investigated. The first approach, Local Binary Pattern, has the lowest detection accuracy of 64.6 percent. Then there's Scale Invariant Feature Transform (SIFT), which has a detection accuracy of 78.3 percent, followed by Neural Network (NN), which has a detection accuracy of 84.13 percent.

Table 1: Comparison table between parameter and method of recognition

Method	(LBP)	(SIFT)	(NN)
Accuracy (%)	64.6	78.3	88.4

1.1 Comparison table between parameter and method of recognition

The comparison of parameter and recognition method is shown in Table 2. There are six (6) methods in Neural Network that have been researched, as shown in the table. Recurrent Neural Networks, Convolutional Neural Networks, Modular Neural Networks, Feedforward Neural Networks, Radial Basic Function Neural Networks, and Multilayer Perception are some of the methods [5].

Table 2: Comparison between Neural Network

Parameter / Neural Network Method	Accuracies (%)	Compatibility
RNN	High accuracy for Binary (75%- 80%)	• Language processing • Speech Signal
CNN	88.4%	• Images • Video
MNN	44%-74%	• Images
FFNN	85%	• Images
RBFNN	-	• Restoring system
Multilayer Perception	-	• Non-linear activation function • Using hyperbolic Tangent and Logistic Function

2. Methodology

Using the situation of the vehicle at each edge, the aggregate of the vehicle at each edge is determined, so the next stage is to locate and work out the spots of the vehicle. The location of the car is necessary to understand the moving vehicle in the following frames and hence the frame rate is known for the capture and measurement. This data must be captured in a non-stop cluster in the same size as the camera, provided that the separation between the center and the vehicle is supposed to be interpreted. Figure 1 provides more specifics of the operation of this unit. For the first level, the vehicle video or vehicle example must be entered. The machine would then recognize the video. After a training method, use other videos and measure the input as well as the other videos. The machine would then identify the video of the vehicle. It compares the model of vehicle recognition and gets the result.

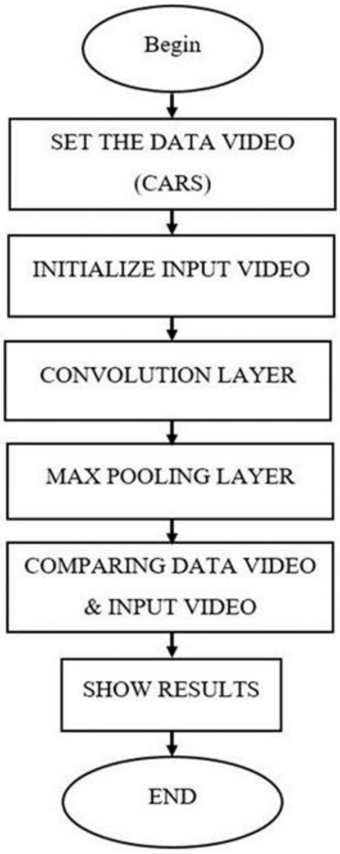

Figure 1: Project flow chart

1) SET THE DATA VIDEO

Set the videos to data in the software code. This device will store the data videos from the code of the programs.

2) INITIALIZE INPUT VIDEO

This step is to enter the input video for convolution layer.

3) CONVOLUTIONAL LAYER

Convolution layers greatly decrease the sum of those layers, a much simpler problem of learning. To minimize overfitting of the parameters, use convolutional layers.

4) MAX POOLING LAYER

After convolutional layers, pooling layers are approached. The pooling layers simplify the data from the convolution layer. Each component map that is derived from the convolution layer is taken by the pooling layer and a merged element map is ready.

5) ANALYZE THE INPUT AND OUTPUT VIDEO

Analyze and classify all videos (input data video and output video).

6) SHOW RESULTS

The result would show that the data we want in the video will be detected the same or nearly the same.

2.1 Software and Hardware

The essential software and hardware used in this project are shown in this section. Python, darknet framework, Yolo OpenCV library and mp4 dataset are the core software and hardware. Figure 2 displays the software and hardware process flow used in this study

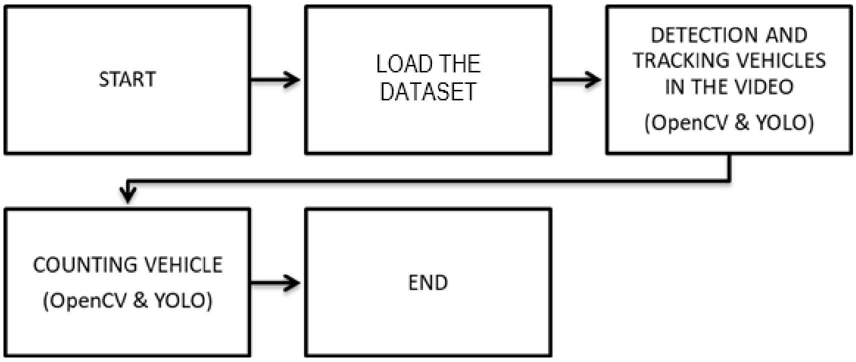

Figure 2: Project Flow for software and hardware

2.2 Project setup

The proposed algorithm is implemented using OpenCV 4.2.2 and Python 3.8 in a laptop with intel i7- 4750HQ, CPU of 2.60 GHz processor with 12GB RAM. The parameters of the YOLO are selected as provided by the respective papers.

2.3 Dataset

There is a scarcity of standard road video datasets for vehicle counting. The video datasets were created using pre-recorded footage from the internet. As a result, videos are taken in a roadway facing downwards with various lighting conditions and shadow effects to assess the efficacy of the suggested method as shown in Figure 3. All the videos were created in RGB.mp4 format with a resolution of 680 x 480 pixels. Additionally, the video contains complex background with moving object and some obstacle. To initiate the counting process, a small road section is manually in the first frame. Figure 4 shows the total vehicle counted parameter interface. A simple background subtraction algorithm often fails to extract the moving vehicle accurately. In addition, shadows and illumination variation degrade the performance of the background subtraction algorithm. As a result, a convolutional neural network is utilized in conjunction with the darknet framework and YOLO to detect and classify moving cars in the entry window where the tracking trajectory begins. The frame's boundary, on the other hand, serves as an exit line for all vehicles whose tracking paths are terminated by incrementing the count. The display of the bounding boxes and classes names with percentage of vehicles confidence is shown Figure 5.

Figure 3: Sample Road of video used for vehicle counting

Figure 4: Total vehicle counted parameter interface

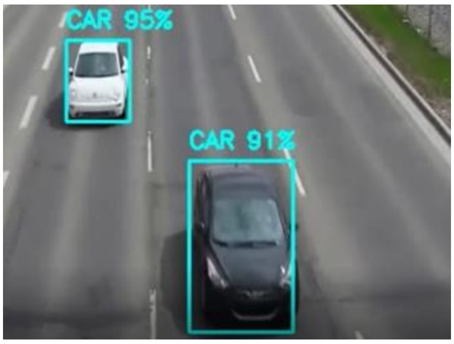

Figure 5: The bounding boxes and classes names with percentage of vehicles confidence

3. Results and Analysis

This chapter will describe the results of deep learning of volume traffic, as well as demonstrate the functionality of this system, which can recognize and counting the presence of vehicles. The findings will be analyzed based on the accuracy with which the vehicles were detected and counted, as well as the total number of vehicles detected.

3.1 Analysis

The experimental results include a comparison of the suggested method's vehicle count with a manual count. The precision and recall assessment measures used for object detection are taken into consideration when measuring quantitatively [6]. The following are the definitions of precision (P) and recall (R):

P = Number of correctly detected bounding boxes ÷ Number of ground-truth bounding boxes [7]

R = Number of correctly detected bounding boxes ÷ Number of detected bounding boxes [8]

Thus, R indicates how many of the detected bounding boxes are correct, whereas P indicates false alarms. Ideal systems should have a high value, near to one. If the overlap of the detected bounding box and the ground-truth bounding is larger than 0.5, the vehicle is detected. F-scores are used to calculate the harmonic mean of precision and recall [9].

$$\mathbf{F} = 2\mathbf{PR} \div \mathbf{P} + \mathbf{R} \qquad \textit{Eq.1}$$

Counting accuracy is calculated in the same way as:

Accuracy = Number of bounding boxes detections ÷ Number of ground-truth detections

Table 3 shows the videos, recall, precision, F-measure, ground-truth count, and vehicle count acquired with the proposed method.

Table 3: Result deep learning of volume traffic counting and recognition

Video (mp4)	Total of vehicles in the ground- truth	Total of vehicles detected by the system	Missing/ Error	Precision (%)	Recall (%)	F-score	Counting accuracy (%)
Output 1	11	11	0/1	91	91	91	100
Output 2	27	26	1/0	96.3	100	98.1	96.2
Output 3	14	14	0/2	86	86	86	100
Output 4	16	14	2/1	100	93	96.3	87.5
Output 5	11	11	1/1	91	91	91	100

3.2 Discussions

Five videos were used to evaluate the efficiency of the proposed method, each with a distinct volume of traffic. Table 3 shows that using the proposed method, the convolutional neural network achieves better accuracy. As a result, an average accuracy of 96.2 percent was achieved. In addition, during the development of this project, some problems developed which required adjustment of the project. While this project used a CPU to process the data, there was a significant delay in object recognition and detection. The best solution for this problem is to process the data using the GPU. Furthermore, this study considers a single lane for the number of the vehicle, which in future will be enlarged to 2 lanes. Total of vehicle counted are display at the top corner in video output 2 and output as shown in Figure 6 and 7, respectively.

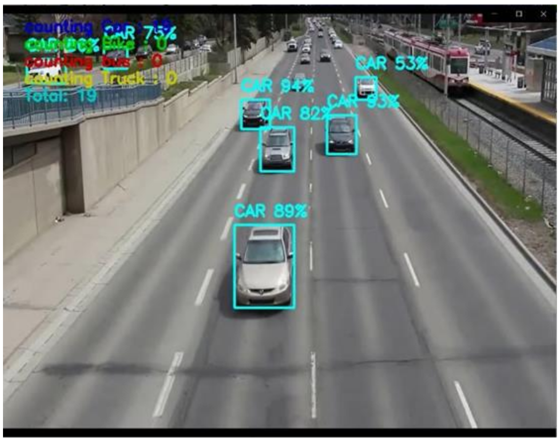

Figure 6: Total of vehicle counted are display at the top corner in video output 2

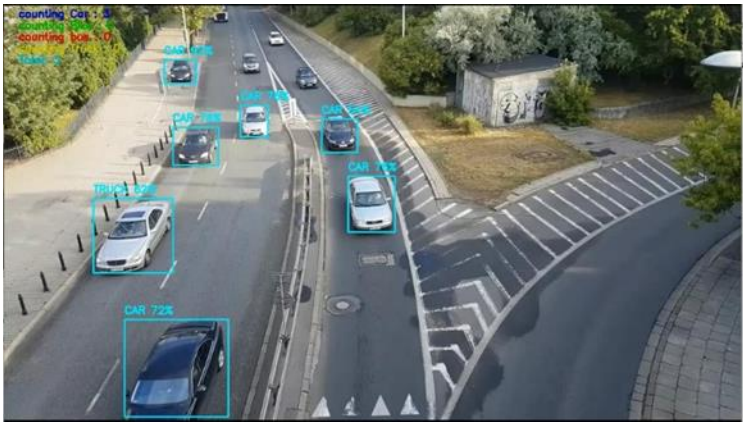

Figure 7: Total of vehicle counted are display at the top corner in video output 3

4. Conclusion

This Deep Learning System is already applicable to the prior systems, Local Binary Pattern, Scale-Invariant Transform Feature, and Neural Network. However, research shows that this Deep Learning System is more trustworthy to apply to the Neural Network approach. However, there are several different types of neural networks, including Recurrent Neural Networks, Convolutional Neural Networks, Modular Neural Networks, Feedforward Neural Networks, Radial Basis Function Neural Networks, and Multilayer Perceptron. This system can classify different types of vehicles as well as counting the number of vehicles that pass through the route. Similarly, stated in chapter 1, the primary goal of this project is to construct a system for vehicle recognition employing the provided algorithm in Deep Learning algorithm moving. To make the algorithm more compatible, there are still a few modifications that can be made in the future. This study considers a single lane for the number of the vehicle, which in future will be enlarged to 2 lanes. Furthermore, an automatic counter extraction will in the future be considered by lane detection. The accuracy of the procedure determines how many vehicles can be counted in this system. However, according to these tests, the Convolutional Neural Network approach has a 96.2 percent accuracy but still has a detecting error.

Acknowledgement

The authors would like to thank the Faculty of Electrical and Electronic Engineering, Universiti Tun Hussein Onn Malaysia for its support.

References

[1] Ingcheng Luoa, Ruihan Shenb, Jian Huc, Jianhua Dengd, Linji Hue and Qing Guanfa Xingcheng Luo, Ruihan Shen, (2012), A Deep Convolution Neural Network Model for Vehicle Recognition

[2] P. Zheng and M. Mike, "An Investigation on the Manual Traffic Count Accuracy," *Procedia - Social and Behavioral Sciences*, vol. 43, pp. 226–231, 2012, doi: 10.1016/j.sbspro.2012.04.095

[3] C. L. Wan and K. W. Dickinson, "Road Traffic Monitoring Using Image Processing—A Survey of Systems, Techniques and Applications," IFAC Proceedings Volumes, vol. 23, no. 2, pp. 27– 34, Sep. 1990, doi: 10.1016/s1474-6670(17)52646-4

[4] Şaban Öztürk, Bayram Akdemir, (31st April 2018), Comparison of HOG, MSER, SIFT, FAST, LBP and CANNY features for cell detection in histopathological images

[5] [James Le, (Aug 2, 2018), The 10 Neural Network Architectures Machine Learning Researchers Need to Learn

[6] Analytics Vidhya, “Precision vs Recall | Precision and Recall Machine Learning,” Analytics Vidhya, Sep. 03, 2020. https://www.analyticsvidhya.com/blog/2020/09/precision-recall- machine-learning/ (accessed Jun. 24, 2021)

[7] “Classification: Precision and Recall | Machine Learning Crash Course,” Google Developers, 2020. https://developers.google.com/machine-learning/crash-course/classification/precision- and-recall (accessed Jun. 24, 2021)

[8] “Classification: Precision and Recall | Machine Learning Crash Course,” Google Developers, 2020. https://developers.google.com/machine-learning/crash-course/classification/precision- and-recall (accessed Jun. 24, 2021)

[9] “Classification: Precision and Recall | Machine Learning Crash Course,” Google Developers, 2020. https://developers.google.com/machine-learning/crash-course/classification/precision- and-recall (accessed Jun. 24, 2021)