# DEEP LEARNING FOR MOTORCYCLE DETECTION IN TRANSPORTATION

Doan Phuoc Mien
*Tra Vinh University*
*No. 126 Nguyen Thien Thanh Street, Ward 5, Tra Vinh City, Viet Nam*
*Email: phuocmien@tvu.edu.vn*

Le Minh Tu
*Tra Vinh University*
*No. 126 Nguyen Thien Thanh Street, Ward 5, Tra Vinh City, Viet Nam*
*Email: lmtu@tvu.edu.vn*

Nguyen Ngoc Dan Thanh
*Tra Vinh University*
*No. 126 Nguyen Thien Thanh Street, Ward 5, Tra Vinh City, Viet Nam*
*Email: ngocdanthanhdt@tvu.edu.vn*

Duong Mong Thu
*Tra Vinh University*
*No. 126 Nguyen Thien Thanh Street, Ward 5, Tra Vinh City, Viet Nam*
*Email: thu1986daotao@tvu.edu.vn*

**ABSTRACT**

*Traffic safety is one of the issues that many countries are concerned about. Especially in Vietnam, the number of traffic accidents is high, and motorcycles are the main means of transportation. Therefore, motorbike detection is considered as part of traffic monitoring and coordination system. In this paper, we propose a method to recognize motorcycles by using deep learning algorithm. We used Faster R-CNN algorithm to detect motorcycles in videos. The videos used for training and testing were collected by mobile phones from several high-traffic routes in Tra Vinh and Da Nang cities, Vietnam. These videos were split into frames with each 40-second video split into 50 images. The next step was to label the regions in the images that contained motorcycles and convert to Tensorflow's format. The model was tested on 60 images containing 286 motorcycles with different sizes, directions and weather conditions. The rate of motorcycles detected was 81.81%.*

Keywords: Deep learning, motorcycle detection, object detection, Faster R-CNN.

## INTRODUCTION

Traffic safety is a big challenge for many countries around the world, especially in Vietnam, traffic safety issues have been concerned by authorities due to the high number of annual traffic accidents. According to the report of the Vietnamese National Traffic Safety Committee, there were 8,385 traffic accidents nation-wide in the first 6 months of 2019. In order to ensure traffic safety, many studies in the field of transport have been invested.

In addition, Vietnam has a large number of motorcycles circulating every day, resulting in frequent traffic congestion in many places. Therefore, how to analyze and calculate the number of vehicles on the roads, thereby offering solutions to avoid congestion such as adjusting traffic lights based on vehicle flow, is one of the solutions to limit traffic jams. Identifying vehicles is the first step in the solution to this problem.

As the situation just mentioned above, we propose a study "Deep learning for motorcycle detection in transportation". The results of this research can be developed to address traffic issues such as counting the number of vehicles on the roads and calculating traffic densities, etc.

## RELATED WORK

### Traffic Moving Object Detection

There are several challenges in detecting moving objects due to the movement of objects in a series of successive frames in the video streams [1]. Moving objects can be people, animals, vehicles such as cars, motorcycles, airplanes, ships, etc. The positions of these objects usually depends on the sequence of consecutive and related frames. In addition, the position of the object can be presented by moving the pixel or moving through the bounding box on the image sequence of the video [2]. Detecting moving objects from video using a static object called the "background" surveyed by Thierry Bouwmans in 2014 [3].

**Vehicle Detection**

Subtraction of pixels on consecutive frames in the video was used to detect vehicles [4]. Most vehicle detection techniques developed so far were based on background differentiation, so the results were very sensitive to variations in ambient light. In the study [4], the author separated foreground objects from the background by subtracting pixels between frames to create a process for processing it in a series of consecutive frames.

Using colors from images can quickly detect vehicles [5]. In this method, colors were used to differentiate vehicles from background and identify candidate vehicles by combining vehicle characteristics such as edge maps, corners, and wavelet transform coefficients.

Khammari et al. [6] used the gradient based and Adaboost classification for Adaptive Cruise Control (ACC) applications. In this method, the system makes assumptions about object locations based on gradient. Then, validation step will be used to verify the hypothesis by using the algorithm AdaBoost without luminance dependence.

In recent years, there have been many studies on vehicle detection by using deep learning method. Convolutional Neural Network (CNN) is one of the most common techniques used for image recognition as well as vehicle detection [7][8]. Tarmizi et al. [7] used CNN to recognize vehicles in dark environment and various bad weathers.

## PROPOSED METHOD

Nowadays, deep learning has been applied to many practical problems of ma-chine learning such as image recognition, speech recognition, object detection, etc. In some situations, computers can detect objects in images more accurately than humans after being trained by deep learning method. In this paper, we pro-pose to use the deep learning model as in Fig. 1 to detect motorcycles. The input data for the model is videos, which are separated into frames. Each frame has dimensions of 512x512. These frames are then transformed into a 262x144 element matrix. Each element will have a value between 0 and 255. The next step will use the filtering matrices called kernels to create hidden layers. Next, the max pooling layer is used to create a matrix of maximum values based on the size of the window.

After creating the hidden layers, the next step the system will perform two calculation layers. The first layer shortens the number of dimensions based on specific characteristics. The second layer performs calculations based on the number of neurons.

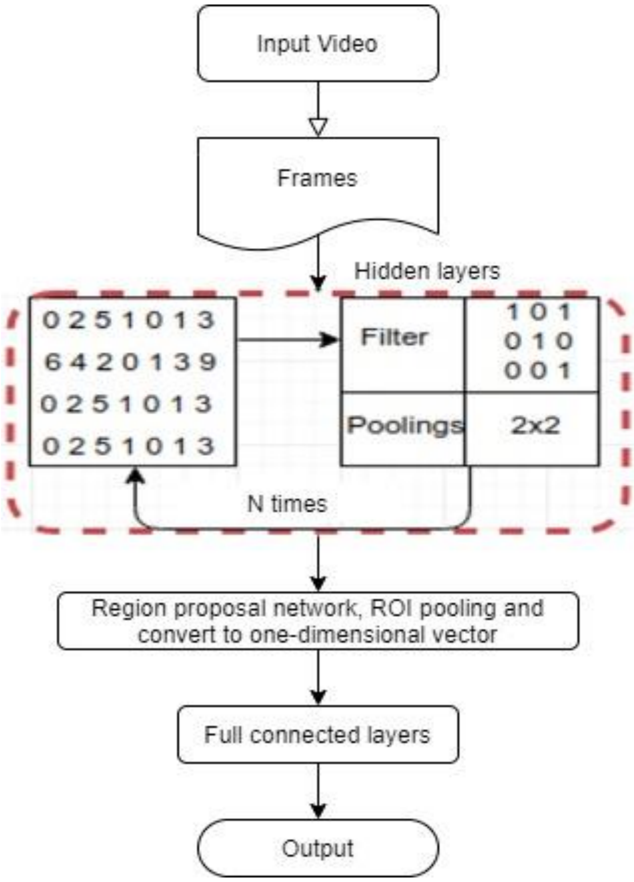


**Figure 1: Proposed method**

Our method is based on Faster R-CNN model for object detection [9]. Faster R-CNN is an algorithm for finding the positions of objects in an image. The algorithm's output is boxes containing objects and their corresponding objects. First, we used convolutional layers and max pooling layers to extract convolutional feature maps from the images. The next step was to predict the region proposals from the feature maps by using the Region Proposal Network. These region proposals were then reshaped to the same size by using Region of Interest (ROI) pooling and then flatten in order input into the classifier to identify the corresponding objects in the region proposals.

## EXPERIMENTS

### Training and Testing Process

Based on the proposed model, the training and testing process is shown in Fig. 2.

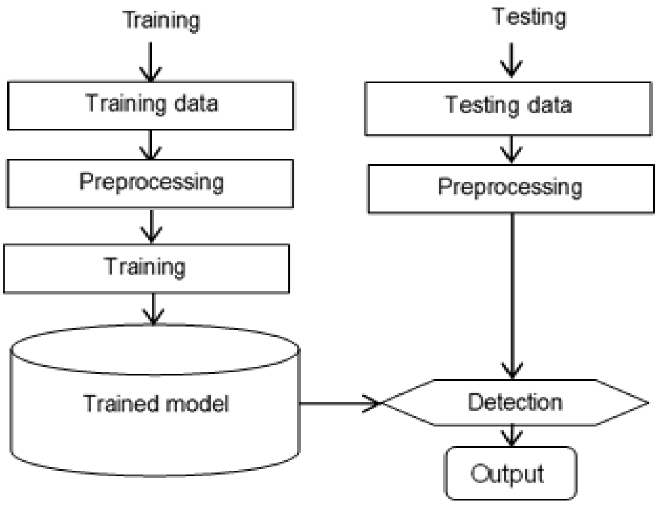


**Figure 2: Training and testing process**

### Dataset

In this study, we used videos recorded at crossroads in Tra Vinh city of Vietnam, internet sources and some videos recorded at the route in Da Nang City of Vietnam. The videos were recorded at the places having lots of transportation.

### Preprocessing

### Preprocessing video frames

From existing videos, we used the OpenCV library to cut these videos into images, on average, for a video of about 40 seconds, we got 50 images. Object detection and tracking algorithms processed these cropped image areas to reduce system processing time.

### Labeled Data

The next step in the preprocessing stage was to label the images. In this study we used labelImage software available at https://github.com/tzutalin/labelImg to label images. Areas containing motorcycles in the images were marked and labeled as "motocycle". Once marked, the data was saved to the annotated .xml files. These XML files would be converted into separate csv files so that they could be further converted into Tensorflow (TFRecord) file format. The rendered data included the positions of the four corners of the rectangle that contained object. Next, the data was divided into two parts which were training and testing data. The training data and testing data were saved to train.record and test.record, respectively.

### Smoothing Images

Due to the low image quality captured from surveillance cameras (320 × 240 pixels), the images need to be smoothed out to eliminate noise. Gaussian operator was used to eliminate noise. The regions of interest (ROI) in each frame were determined using the 2-dimensional Gaussian function and its standard deviation as shown below:

$$G(x,y) = \frac{1}{2\pi\sigma^2} e^{\frac{-x^2+y^2}{2}} = \frac{1}{115}\begin{bmatrix} 2 & 4 & 5 & 4 & 2 \\ 4 & 9 & 12 & 9 & 4 \\ 5 & 12 & 15 & 12 & 5 \\ 4 & 9 & 12 & 9 & 4 \\ 2 & 4 & 5 & 4 & 2 \end{bmatrix}$$

**Color Conversion**
In this step, RGB color images were converted to grayscale images using the following equation:

$$Grayscale\ image = 0.2989 * R + 0.5870 * G + 0.1140 * B \quad (2)$$

**Training**
First, we trained the model by using Faster R-CNN inception with a pre-trained model on the COCO dataset. Input data for the model included TFRecords and configuration file for the pre-trained model.

**Testing**
The input of the testing process was the videos. These videos would be split into frames and identify the motorcycles contained in these frames. The testing process was shown in Fig. 3.

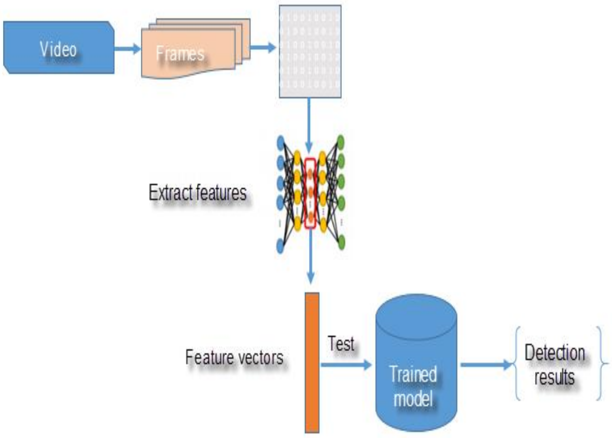


**Figure 3: Testing process**

In this study, we use the Centroid algorithm to track objects. Basically, reading two consecutive frames then calculating the Euclidean distance between objects in the two frames and finding the smallest distance between objects. The smaller the distance between two objects, the two objects are considered one. The results of motorbike identification in video recorded by cameras and mobile phones are shown in figures [4] and [5].

**Figure 4: The result of motorcycle identification using video recorded by the camera.**

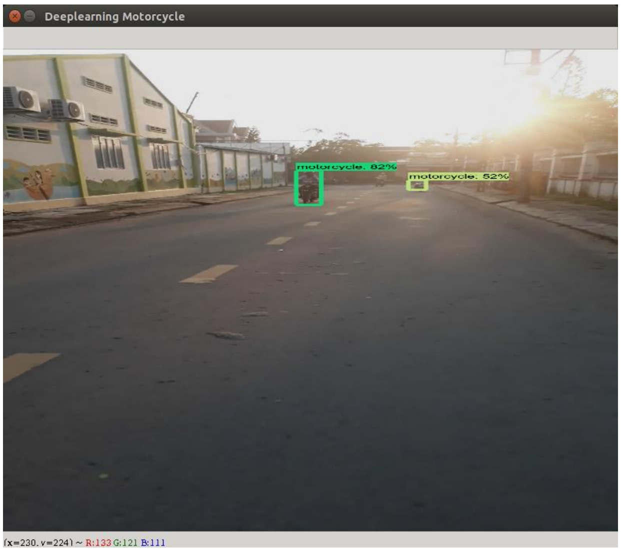


**Figure 5: The result of motorcycle identification using video recorded by mobile phone.**

## RESULTS

After training, we had a model that could recognize motorcycles. The model was used to detect motorcycles in 60 different images contains 286 motorcycles in many directions, large and small sizes. The detection results were shown in Table 1.

| Images | Number of motorcycles | Number of motorcycles detected | Detection rate |
|---|---|---|---|
| 1 – 10 | 54 | 40 | 74.07 % |
| 11 - 20 | 45 | 34 | 75.56 % |
| 21 - 30 | 39 | 33 | 84.61 % |
| 31 - 40 | 65 | 57 | 87.69 % |
| 41 - 50 | 37 | 32 | 86.48 % |
| 51 - 60 | 46 | 38 | 82.61 % |
| Total 60 images | 286 | 234 | 81.81 % |

**Table 1: Detection results.**

## CONCLUSION AND DISCUSSION

The identification results showed that this is a potential method because the data was collected at street locations having a lot of differences in color, size, number of people on motorcycles or goods on vehicles, and many other factors that influenced identification results such as obscurity, tree shade, etc.

Experiments showed that the model had the highest identification results when the vehicles were not obstructed and of moderate size. In most cases, the vehicles were not detected because the vehicles were partially obscured or the vehicles had too small image sizes. Especially, the model was also confused when identifying bicycles and electric bicycles as motorcycles because these two types of vehicles were relatively similar to motorcycles, and the model was not trained to identify these two types of vehicles. In addition, the above results were affected by the small amount of data used for training stage as well.

The algorithm gave the best results in the case of vehicles on one-way roads, vehicles traveling ahead and without obstructions.

**REFERENCES**

Mehran Yazdi and Thierry Bouwmans. New trends on moving object detection in video images captured by a moving camera: A survey. Computer Science Review, 28:157–177, 2018

Di Zang, Junqi Zhang, Dongdong Zhang, Maomao Bao, Jiujun Cheng, and Keshuang Tang. Traffic sign detection based on cascaded convolutional neural networks. In 2016 17th IEEE/ACIS International Conference on Software Engineering, Artificial Intelligence, Networking and Parallel/Distributed Computing (SNPD), pages 201–206. IEEE, 2016.

Bouwmans, Thierry. (2014). Traditional and Recent Approaches in Background Modeling for Foreground Detection: An Overview. Computer Science Review. 11. 10.1016/j.cosrev.2014.04.001.

U. Nagaraj, J. Rathod, P. Patil, S. Thakur, and U. Sharma, “Traffic jam detection using image processing”, International Journal of En-gineering Research and Applications, vol. 3, no. 2, pp. 1087–1091, 2013.

L. Tsai, J. Hsieh and K. Fan, "Vehicle Detection Using Normalized Color and Edge Map," in IEEE Transactions on Image Processing, vol. 16, no. 3, pp. 850-864, March 2007.

A. Khammari, F. Nashashibi, Y. Abramson and C. Laurgeau, "Vehicle detection combining gradient analysis and AdaBoost classification," Proceedings. 2005 IEEE Intelligent Transportation Systems, 2005, Vienna, 2005, pp. 66-71.

I. A. Tarmizi and A. A. Aziz, "Vehicle Detection Using Convolutional Neural Network for Autonomous Vehicles," 2018 International Conference on Intelligent and Advanced System (ICIAS), Kuala Lumpur, 2018, pp. 1-5.

Chen, L., Ye, F., Ruan, Y. et al. An algorithm for highway vehicle detection based on convolutional neural network. J Image Video Proc. 2018, 109 (2018). https://doi.org/10.1186/s13640-018-0350-2

Ren, Shaoqing & He, Kaiming & Girshick, Ross & Sun, Jian. (2015). Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks. IEEE Transactions on Pattern Analysis and Machine Intelligence. 39. 10.1109/TPAMI.2016.2577031.