Journal of Advanced Research in Applied Sciences and Engineering Technology

Journal homepage:
https://semarakilmu.com.my/journals/index.php/applied_sciences_eng_tech/index
ISSN: 2462-1943

Machine Learning Techniques for Sustainable Smart Cities Traffic Management

Rudee Roslan[1], Sokchoo Ng[1,*], Lee Chin Yee[2]

[1] Faculty of Acts and Science, International University of Malaya-Wales, Malaysia
[2] Faculty of Business, International University of Malaya-Wales, Malaysia

ARTICLE INFO

Article history:
Received 8 June 2023
Received in revised form 25 September 2023
Accepted 6 October 2023
Available online 19 October 2023

Keywords:
Machine Learning; road traffic; Convolutional Neural Network

ABSTRACT

The issue of traffic congestion is one of the problems that disrupt the smoothness and comfort of the community in carrying out daily affairs. Rapid development, with an expanding population and vehicles in cities, has caused traffic congestion to become one of the most pressing concerns. As a result, this study proposed vehicle tracking method for traffic system management to improve and keep up with the rising demand for better traffic on the road. In the proposed system, cameras were placed at traffic intersections and along the road or highway to gather real-time traffic data. Image processing and machine learning algorithm were applied to calculate the number of vehicles based on the traffic flow and vehicle speed on the road to reflect the current traffic condition so that it helps general public to plan their trip and less pollution is produced in long run. This study focused on the cities of Cyberjaya, Putrajaya and selected Klang Valley areas to study the traffic condition during the peak and off-peak hour. Experimental results showed that the proposed machine learning solution provided high accuracy and effective detection on real-time traffic congestion. Current method used variables such as vehicle speed and number of vehicles to calculate the traffic flow on the road. As future direction, this method can be expanded to incorporate vehicle density as one of the variables for more precise traffic congestion reporting.

1. Introduction

The economies have increased the dependency of working people on transportation as it is reliable and highly efficient [1]. However, traffic congestion is an enduring issue for the sustainability of transportation development [2]. Besides that, transportation is also one of the obstacles which provide negative impact to achieve sustainable smart city [3,4]. The issue of traffic congestion is a global problem to all cities around the world especially those cities which have become the economic centre and attract more population to live in the city leading to the increasing number of private vehicles usage [5].

* *Corresponding author.*
E-mail address: ashleyng@iumw.edu.my

https://doi.org/10.37934/araset.33.1.246255

Report shows that 88% of the total population of Malaysia is highly dependent on private vehicles as it is the most important mode of transportation. This is one of the causes of traffic congestion in the city and it leads to high transportation costs [6,7]. In Malaysia, the traffic congestion issue is directly associated with the increase usage of private vehicle [8]. Main roads in the city centre cannot accommodate the ever-increasing number in vehicles. This is further compounded by the fact that the number of new vehicle registration in Malaysia is around 700,000 a year. Thus, Malaysia is among the countries with tremendous vehicle growth on the road. However, the very slow construction of roads or highways in Malaysia are unable to cope with the vehicle's increase rate. Malaysians waste a lot of time stuck in traffic jam every day with at an average rate of two hours per day. This has resulted in a huge loss to Malaysia's productivity, amounting to RM5.51 billion annually. Hence, forecasting traffic flow is important for traffic management [9].

Reducing traffic congestion is the priority agenda in smart city development [10]. The services of smart traffic in smart cities, especially in big cities, are the most promising and prominently challenging part of the Internet of Things (IoT) [11]. Machine learning has been adopted to predict, plan and analysis the traffic congestion model in order to realize the idea of sustainable smart city [12]. Contemporary cities have encountered many problems including traffic, waste collectors, residential facilities and the environment and studies on related services have increased globally [13]. Many of the most developed countries are implementing smart city technology in their largest cities for the purpose of improving the efficiency of their city traffic management. Cities such as Hong Kong, San Francisco, Taipei, Malta, and Vancouver are some examples where the implementation of smart city technologies have been employed [14].

Technology driven infrastructure determine the smartness of the city [15]. Most of the cities and smart cities appreciable employ the IoT applications for the deployment of the traffic signalling. The traffic light intelligence is controlled by the traffic control centre allowing the control of traffic stipulated by guidelines defined and it was no different with urban traffic management [16]. It seems to be a nightmare for the residents living in high traffic congestion area, and this problem has been unresolved until now, thus disrupting the lives and well-being of the residents. There is a great interest particularly in urban traffic management to rectify the congestion and this issue has remained great challenge for researcher to response and address the impact of the poor traffic management especially on the daily life of the inhabitants [17].

Smart transportation is the key factors in smart environment to improve people's quality of life [21]. Most cities in the world experience the same traffic congestion problem from one continent to another. This problem is also not uncommon in large cities in Southeast Asia. These mentioned cities are affected as the existing traffic system cannot cope with the increasing number of vehicles each year. To solve these social and the traffic congestion problems, each city should be able to improve the quality of life by adding the new features by considering and encouraging civic participation in human capital revitalization and utilisation [22]. By using IoT technology, it is possible to provide the smart transportation such as optimized route suggestion and it brings benefits to the services providers and the end-users [18,19]. Predicting accurate traffic flow is the necessary condition in Intelligent Transportation System [20]. Malaysia has been adopting smart traffic light and edge-based computing for Smart Traffic Analytics and Recognition System (STARS) to optimize the traffic flow through a junction. MATC Intelligent Traffic System (MITS) detects the number of vehicles passing through a junction in a given time window based on a controller time setting. Upon changing the controller time setting, the device conducts before and after traffic analysis by comparing the revised parameters. This study aims to explore an alternative approach for intelligent sensor detection by applying Convolutional Neural Network (CNN) to detect the traffic flow. The proposed module will indicate the traffic condition that the road users can use to avoid traffic congestion on the main road, especially at intersections, highways, and city roads. The study is focusing on how to incorporate machine learning algorithm to detect the increasing number of vehicles passing through city centre, especially during peak hours.

2. Methodology

To understand the feasible of applying IoT technology to solve the traffic congestion, innovative detector system was developed. The proposed detection sensor system aims to provide an efficient solution for reducing road congestion. This study is cyclical research which is depicted as a linear model for clarity in the visualization [23]. Cyberjaya is the smart city district in Malaysia which also the test bed for nurture technological innovations [24]. High volume of traffic flow occurs in this district to its neighbouring districts on weekdays. Therefore, this study selected the cities of Putrajaya, Cyberjaya and selected Klang Valley areas as the case study areas to test the performance of the proposed system. The study area covers an area of approximately 49 square kilometres for Putrajaya, approximately 28 square kilometres for Cyberjaya, and approximately 600 square kilometres for the rest of Klang Valley. The frequency of cars on the road was perceived as an indicator that describes the traffic condition for data collection. During the data collection process, recording video data was collected from pedestrian bridge areas on the Damansara Puchong Highway and the Puchong Sungai Besi Highway (as shown in Figure 1). The video data was collected via manual video recording using an iPhone 12 from 7am to 4pm (from 5th November to 22nd November, 2021). There is no specific location for video recording as the main objective of the experiment is to record the movement of vehicles. Thus, only one video device is used in the data collection.

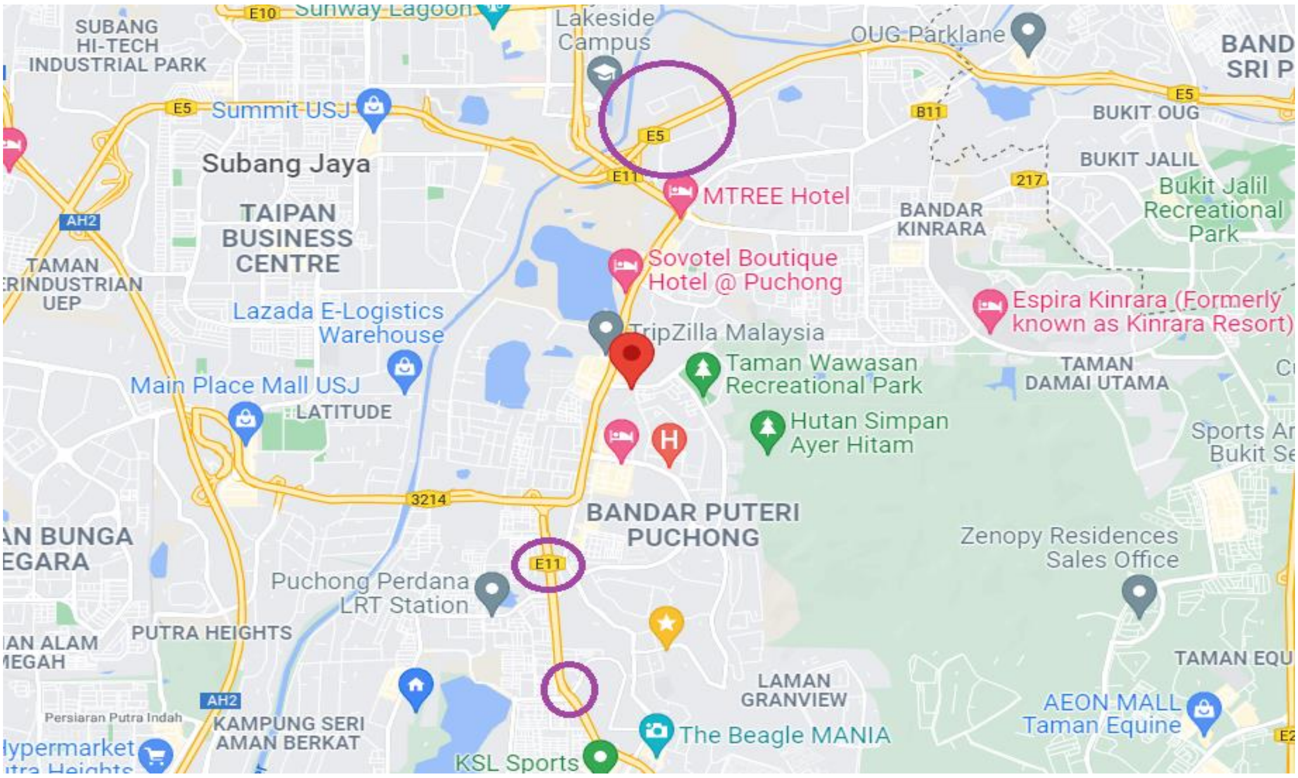

Fig. 1. Circle shows areas for video recording for the study (Google Maps 2019 (95))

Video recordings were gathered throughout the data collection phase. Method to analyse the video data includes the following steps (Figure 2):

i. video data was separated into three categories: training, testing, and evaluation
ii. training data was used to assess YOLO's performance
iii. test data was utilized to evaluate YOLO's performance during the experiment phase.

A modest picture frame is required in traffic analysis applications. In this study, unit of analysis was the speed of the vehicle, the gap between the vehicle and its moving across the road. Other unit analysis included total number of vehicles and public's working time (night shift, Afternoon shift or normal shift).

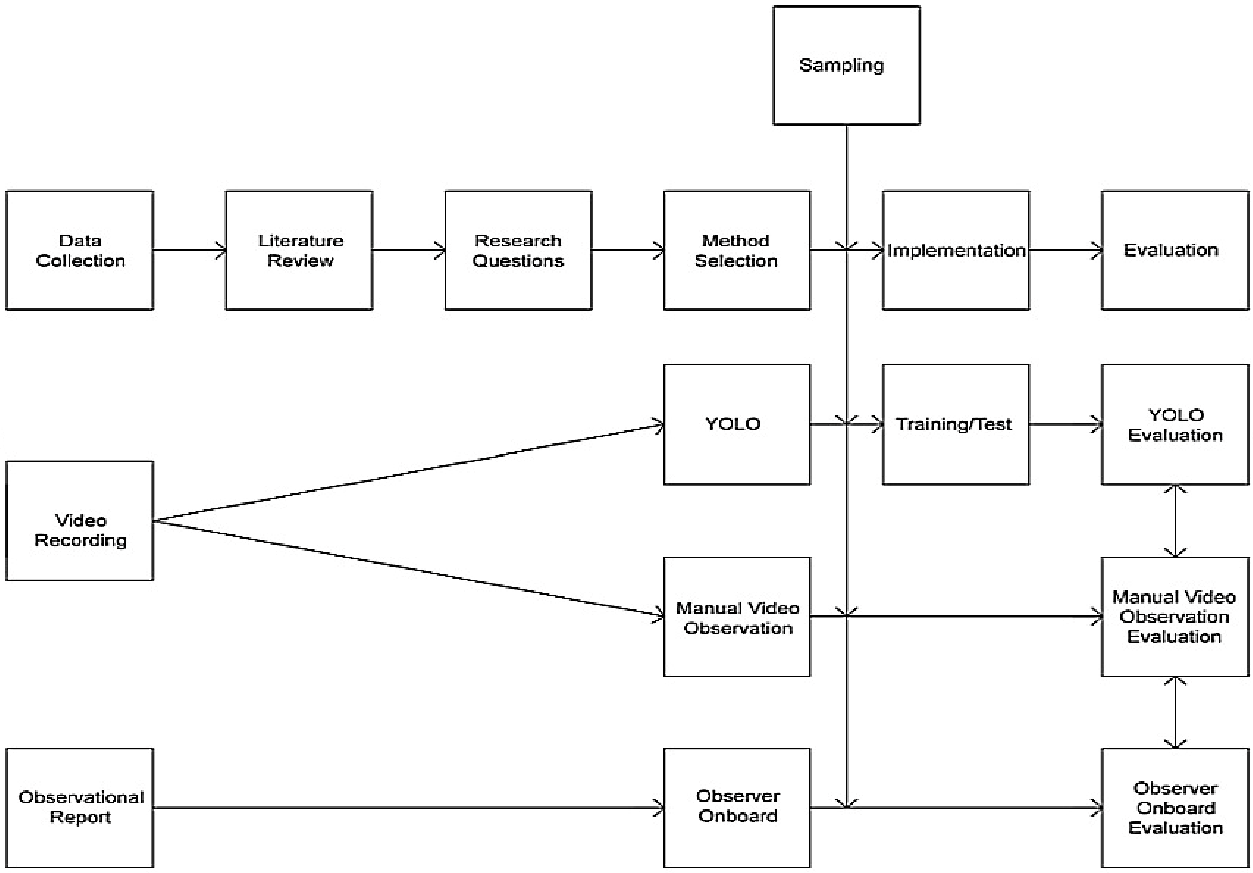

Fig. 2. Overview of Research Design

This research sampled three types of data: training, testing, and evaluation for analysis. In the context of machine learning, dividing data into two sections for training and testing is referred to as the split-test technique [25]. Although the computer vision method separates the data into two distinct training and testing segments, the present study strategy differs from the mentioned machine learning context by adding in the evaluation data set to validate the performance of the detection model. The data ratio utilised in this study is as follows: 70% of the data was utilised to create the YOLO simulation, known as training data set, which is based on the video and coding used. The performance of YOLO during the training phase is evaluated using 20% of the data, dubbed the test data set. The last 10% of data was used to determine the accuracy of YOLO (refer to Table 1).

Table 1
Training, Test and Evaluation Datasets Distribution

Data Type	Percentage
Training	70%
Testing	20%
Evaluation	10%
Total	100%

3. Implementation

The proposed traffic detection system or known as a tracking device is able to count the number of vehicles and detect any image on the road to inform the road congestion. Several software packages were used to develop the system, such as YOLO, Pygame, and Python which uses some metrics and as coding on simulation experiments to detect congestion through CCTV. The tracking device identifies vehicles on the road by combining YOLO, OpenCV with Python and Pygame. The targeted device is positioned at an angle, for as atop a lamppost, to read the movement of passing vehicles. The detector picture showed each car with a green box image and displayed the reading on the screen for data reading output. The data set used in this study was traffic videos recording collected by using handphone cameras of iPhone12 on various roads in the research area. Image data was extracted from the traffic videos using a script and coding was performed using Python, OpenCV, and YOLO. The data set being developed was only for detection purposes without any classification, as the study focuses on detecting the traffic congestion at different hours. The result of this data set is the total number of vehicles on the road which reflects the traffic condition (Figure 3).

The traffic density on the road scanned by the vehicle tracking module was determined by the convolutional neural network (CNN) algorithm. The CNN algorithm first applied a homogeneous neural network to the entire picture. The algorithm then split the picture into parts and calculated the bounding box (window) coordinates and the number of vehicles object detected for each area. The vehicle object value was labelled as a key and produced in JSON format according to the coordinates of the object in the window. Each scan will be updated with the window's coordinates data and readings of each vehicle object that passes by on the road or highway. The vehicle object input from each window was added up and analysed in order to determine the overall number of vehicles on the road. Each scan was captured, calculated, and altered in accordance with the vehicle object detected in the windows. The number of vehicles may be raised and lowered to any number at the coordinate point of the road line. For the detection lane line to be effective, vehicles must be detected and counted as they pass through the detection zone. The tracking device will show the number of vehicles that have passed through the line and the findings of this investigation. Each frame of a traffic video is analysed using the Python programme to identify moving objects as well as objects that are detected. The detector serves as an output that confirms the congestion in the experiment.

Journal of Advanced Research in Applied Sciences and Engineering Technology
Volume 33, Issue 1 (2023) 246-255

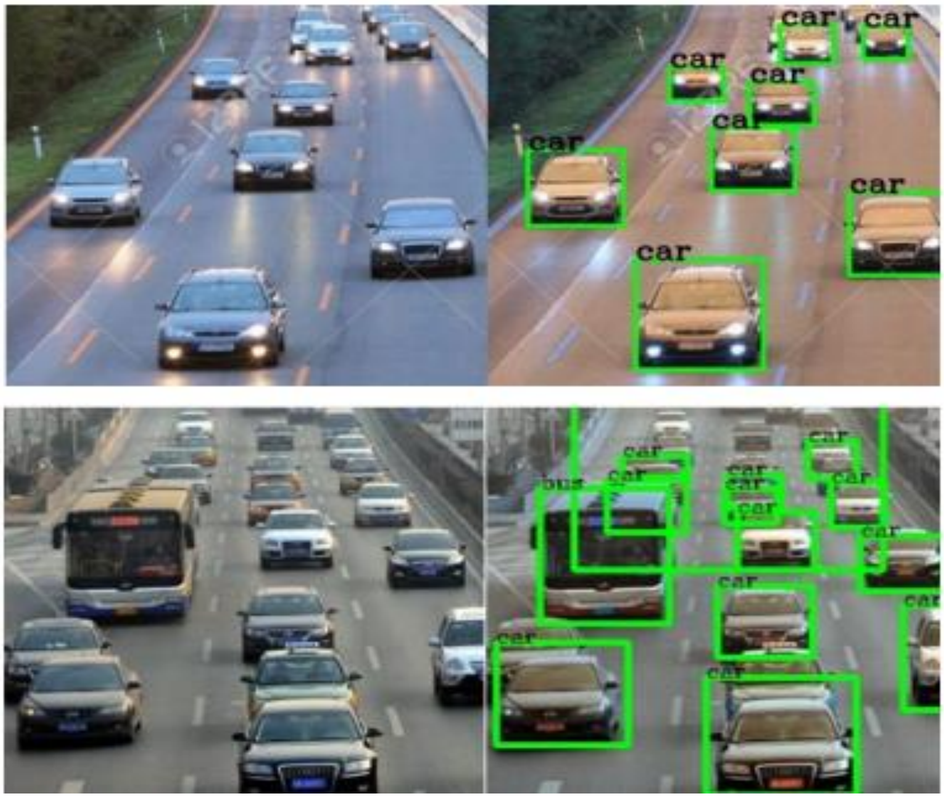

Fig. 3. The proposed traffic detection system detection results

4. Results

To measure the accuracy of the proposed traffic tracking module detection, the detection outcome was compared to some of the traffic recorded in video form. There were 6 videos recorded for the purpose of this experiment, and each set of detector systems tested was based on a video duration of 90 seconds. Each video has a different distribution of traffic on the round trip. The detection system in this experimental video is based on a frame detector as shown in Figure 4, and the vehicles are counted after passing the encoded detection line to count the number of vehicles.

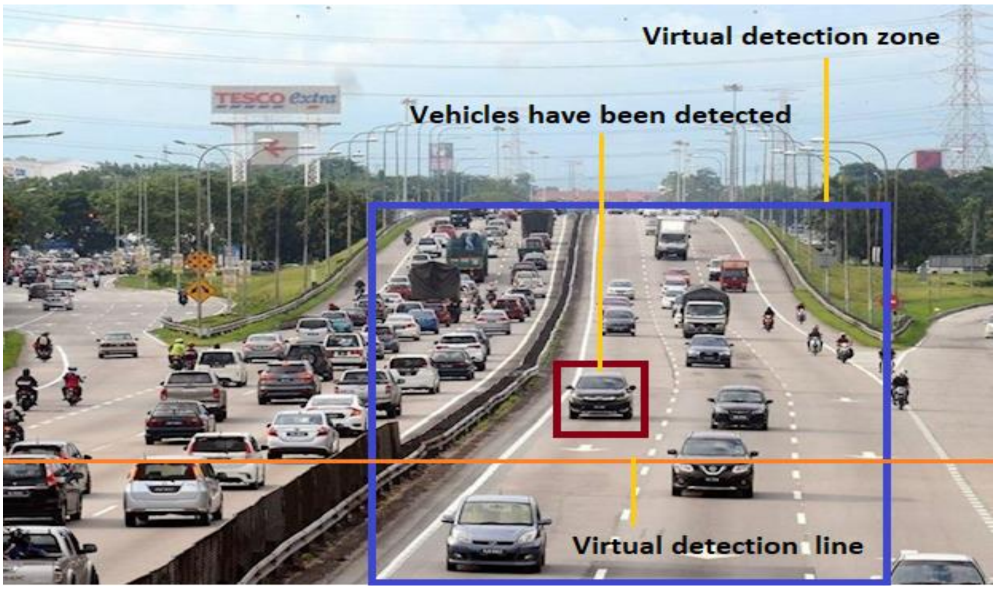

Fig. 4. Frame detector for vehicle object

The test is measured by the total number of vehicles crossing the line and will be compared to other videos in testing road congestion. In other words, each vehicle passing by on the road represents the value of data in tracking and analysing each movement of the vehicle, and this data will be captured if there is significant vehicle movement, which will cause traffic congestion. As an example, the number of vehicles suddenly increases and the detector output shows the vehicle movement is smooth as more vehicles are counted. If the number of vehicles suddenly becomes less, then the detector output shows the vehicle movement is slow and fewer vehicles are counted. This number of vehicles was measured in terms of the number and movement of vehicle that was able to cross the line in the road test found in the video as shown in Figure 5.

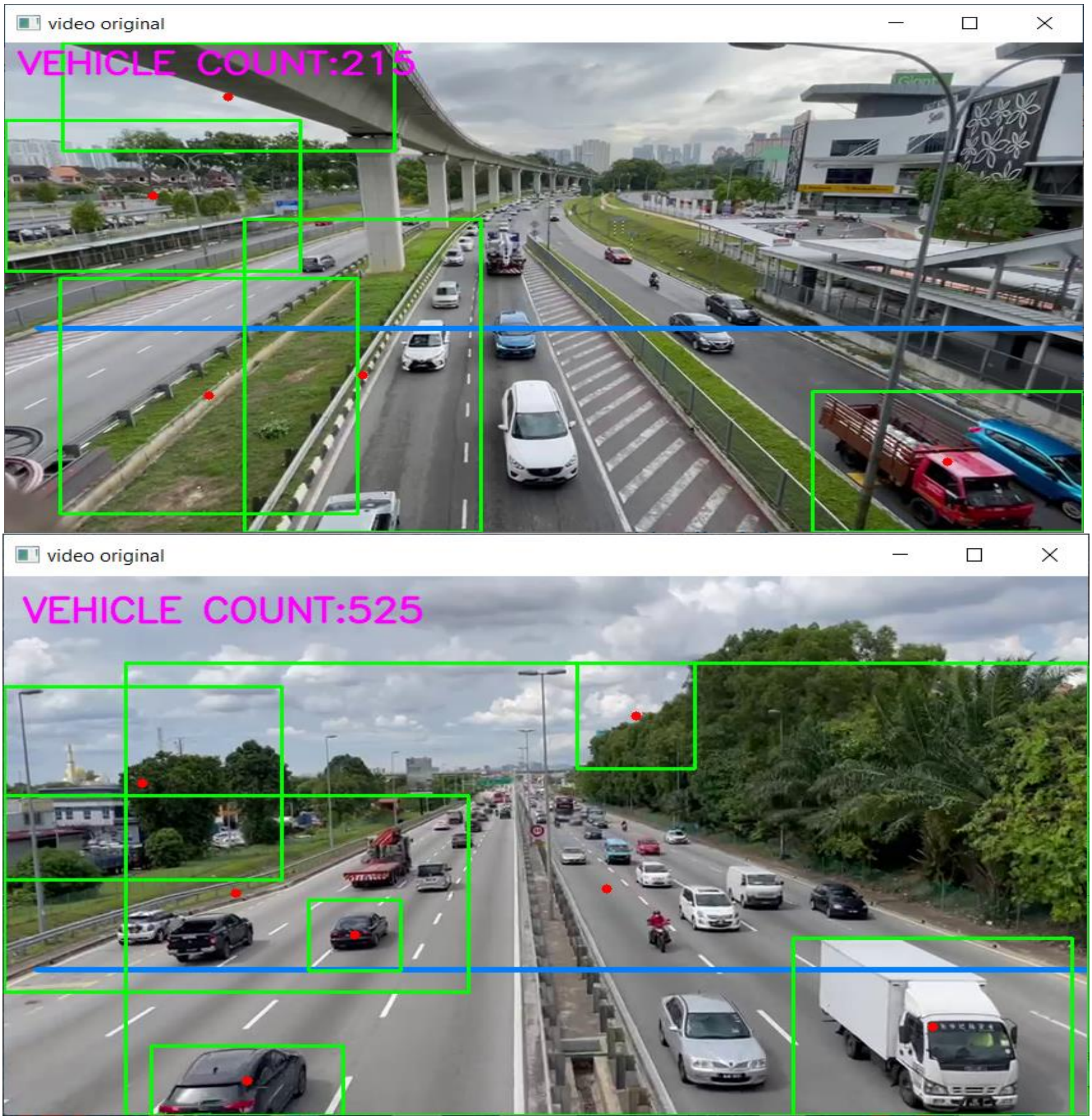

Fig. 5. Count of vehicles on the highway during (a) off-peak hour (b) peak hour

Each test and the test results obtained are recorded and tabulated by the number of vehicles passing through the road. In that case, the essential reading in this experiment is the video output on the detector, which determines whether the vehicle movement increases or vice versa. The series of experiments and the test results are shown in Table 2. The resulting traffic condition was determined based on the vehicle speed index (VSI) and volume of capacity, V/C ratio [26].

Table 2
Total vehicles and traffic condition reported by the vehicle tracking module

Video No.	Total Vehicle	VSI (km/hour)	V/C Ratio	Resulting traffic condition
Video 1	135	50-75	0.61-0.70	Traffic is smooth as fewer vehicles were counted.
Video 2	215	50-75	0.61-0.70	Traffic is smooth as fewer vehicles were counted.
Video 3	400	0-25	0.91-1.00	Traffic is becoming congested as more vehicles are counted.
Video 4	197	50-75	0.61-0.70	Traffic is smooth as fewer vehicles were counted.
Video 5	401	0-25	0.91-1.00	Traffic is becoming congested as more vehicles are counted.
Video 6	249	25-50	0.71-0.90	Traffic is considerable smooth with mild congestion as the total number of vehicles on the road is still reasonable.

5. Discussion

The results from this experiment are, to some extent, satisfactory, but not optimal. This is due to the YOLO detector on the vehicle is not based on the frame specified in the encoding but the setting of OpenCV by an individual Python programmer. Therefore, the YOLO output of vehicle captured in the experiment (Figure 5b) appeared in multiple overlapping image frames. This phenomenon will affect the accuracy of the vehicle captured. To fix this, the code entered should be tested, and libraries such as DLIP and YOLO should be uploaded correctly based on the desired frame image. It is important to train, improve, and keep track on the accuracy of tracking system which can be improved from time to time.

6. Conclusions

This study proposed a vehicle tracking system by using CNN to reflect the traffic condition of selected area in Klang Valley. The vehicle number is calculated using a virtual detection zone created using Python and YOLO. In comparison to the conventional approach, the vehicle detection technique based on YOLO, Python, and OpenCV has a faster detection time and more resilience, which may help minimising the missing and error rates. This method is still be able maintain a high-test rate in a complicated environment, and the fast detection speed is able to satisfy real-time needs while achieving a high level of effectiveness. However, the study could not be conducted comprehensively by area but only involved the main roads in selected urban areas.

The experimental results of the method inferred that the tracking method by using CNN is feasible for the real-time traffic congestion monitoring. The information acquired in this study is used to identify and report on traffic congestion. It may subsequently be expanded to provide an alarm message that can assist the traffic department in taking appropriate action. Practical computer vision results showed that the proposed method can accurately detect, locate, and inspect operating vehicles. This study benefits all the general public and serves as an effective tool to reduce congestion on the roads. The proposed method is able to detect the number of vehicles and their speed to reflect the traffic condition during peak and off-peak hour. Using CNN for the vehicle tracking enhances the current traffic monitoring. Hence, the user can prepare their trip and reduce the traffic congestion. The authority or traffic control centres may utilise this vehicle tracking method to handle the issue of traffic congestion which is the direct factor to carbon emission. As future work, the experiments of current study will be extended to investigate the carbon emission in relation to traffic congestion as the effort to realize low carbon cities and greener urban planning.

Acknowledgement
This research was not funded by any grant.

References
[1] Colon, Célian, Stéphane Hallegatte, and Julie Rozenberg. "Criticality analysis of a country's transport network via an agent-based supply chain model." *Nature Sustainability* 4, no. 3 (2021): 209-215. https://doi.org/10.1038/s41893-020-00649-4

[2] Afrin, Tanzina, and Nita Yodo. "A survey of road traffic congestion measures towards a sustainable and resilient transportation system." *Sustainability* 12, no. 11 (2020): 4660. https://doi.org/10.3390/su12114660

[3] Majumdar, Sharmila, Moeez M. Subhani, Benjamin Roullier, Ashiq Anjum, and Rongbo Zhu. "Congestion prediction for smart sustainable cities using IoT and machine learning approaches." *Sustainable Cities and Society* 64 (2021): 102500. https://doi.org/10.1016/j.scs.2020.102500

[4] Khan, Huma H., Muhammad N. Malik, Raheel Zafar, Feybi A. Goni, Abdoulmohammad G. Chofreh, Jiří J. Klemeš, and Youseef Alotaibi. "Challenges for sustainable smart city development: A conceptual framework." *Sustainable Development* 28, no. 5 (2020): 1507-1518. https://doi.org/10.1002/sd.2090

[5] Olayiwola, K. O., A. M. Olaseni, and O. Fashina. "Traffic congestion problems in central business district (CBD) Ikeja, Lagos metropolis, Nigeria." *Research on Humanities and Social Sciences* 4, no. 1 (2014).

[6] Gianluca Lange. "Overcoming Traffic Congestion in the City Without Building Roads." Utusan Malaysia. 2015. https://www.utusan.com.my/bisnes/korporat/overcome%20traffic%20congestion%20in%20the%20city%20without%20building%20roads-1.80327

[7] Muneera, C. P., and K. Krishnamurthy. "Economic evaluation of traffic congestion at intersection: case study from an Indian City." *Transportation Research Procedia* 48 (2020): 1766-1777. https://doi.org/10.1016/j.trpro.2020.08.212

[8] Ladin, M. A., L. J. Yung, A. S. Darawati, N. Bolong, I. Saad, MR Mat Yazid, A. Abbil, and H. A. M. Yahia. "Development of Transportation Models Based on Students' Interest in a Parking Charging System at Universiti Malaysia Sabah (UMS)." In *Journal of Physics: Conference Series*, vol. 2314, no. 1, p. 012025. IOP Publishing, 2022. https://doi.org/10.1088/1742-6596/2314/1/012025

[9] Chen, Cen, Kenli Li, Sin G. Teo, Xiaofeng Zou, Keqin Li, and Zeng Zeng. "Citywide traffic flow prediction based on multiple gated spatio-temporal convolutional neural networks." *ACM Transactions on Knowledge Discovery from Data (TKDD)* 14, no. 4 (2020): 1-23. https://doi.org/10.1145/3385414

[10] Saleem, Muhammad, Sagheer Abbas, Taher M. Ghazal, Muhammad Adnan Khan, Nizar Sahawneh, and Munir Ahmad. "Smart cities: Fusion-based intelligent traffic congestion control system for vehicular networks using machine learning techniques." *Egyptian Informatics Journal* 23, no. 3 (2022): 417-426. https://doi.org/10.1016/j.eij.2022.03.003

[11] Kim, Jiyeon, Sooyoung Jang, Deockgu Jee, Eunjin Ko, Seung Han Choi, and Mi Kyong Han. "5G based smartcity convergence service platform for data sharing." In *2020 international conference on information and communication technology convergence (ICTC)*, pp. 1522-1524. IEEE, 2020. https://doi.org/10.1109/ICTC49870.2020.9289155

[12] Nosratabadi, Saeed, Amir Mosavi, Ramin Keivani, Sina Ardabili, and Farshid Aram. "State of the art survey of deep learning and machine learning models for smart cities and urban sustainability." In *International conference on global research and education*, pp. 228-238. Cham: Springer International Publishing, 2019. https://doi.org/10.20944/preprints201908.0154.v1

[13] Lee, JooChan, JangHoon Kim, and JungTaek Seo. "Cyber attack scenarios on smart city and their ripple effects." In *2019 international conference on platform technology and service (PlatCon)*, pp. 1-5. IEEE, 2019. https://doi.org/10.1109/PlatCon.2019.8669431

[14] Nam, Taewoo, and Theresa A. Pardo. "Conceptualizing smart city with dimensions of technology, people, and institutions." In *Proceedings of the 12th annual international digital government research conference: digital government innovation in challenging times*, pp. 282-291. 2011. https://doi.org/10.1145/2037556.2037602

[15] Bhattacharya, Sweta, Siva Rama Krishnan Somayaji, Thippa Reddy Gadekallu, Mamoun Alazab, and Praveen Kumar Reddy Maddikunta. "A review on deep learning for future smart cities." *Internet Technology Letters* 5, no. 1 (2022): e187. https://doi.org/10.1002/itl2.187

[16] Basconcillo, Kimberly Ann C., Diuse Josiah B. Benitez, Elfred Alver S. Cantuba, Renz Erwin L. Enriquez, Chester Robert I. Falcon, Kanny Krizzy D. Serrano, Emmanuel C. Guevara, Angelo R. dela Cruz, and Ryan Rhay P. Vicerra. "Development of a vehicle and pedestrian simulation environment with MISO fuzzy logic controlled intelligent traffic light system." In *2017 5th International Conference on Information and Communication Technology (ICoIC7)*, pp. 1-6. IEEE, 2017. https://doi.org/10.1109/ICoICT.2017.8074643

[17] Steenbruggen, John, Peter Nijkamp, and Maarten Van Der Vlist. "Urban traffic incident management in a digital society: An actor–network approach in information technology use in urban Europe." *Technological Forecasting and Social Change* 89 (2014): 245-261. https://doi.org/10.1016/j.techfore.2013.08.038

[18] Zantalis, Fotios, Grigorios Koulouras, Sotiris Karabetsos, and Dionisis Kandris. "A review of machine learning and IoT in smart transportation." *Future Internet* 11, no. 4 (2019): 94. https://doi.org/10.3390/fi11040094

[19] Li, Guofu, Ning Cao, Pengjia Zhu, Yanwu Zhang, Yingying Zhang, Lei Li, Qingyuan Li, and Yu Zhang. "Towards smart transportation system: A case study on the rebalancing problem of bike sharing system based on reinforcement learning." *Journal of Organizational and End User Computing (JOEUC)* 33, no. 3 (2021): 35-49. https://doi.org/10.4018/JOEUC.20210501.oa3

[20] Yang, Di, Songjiang Li, Zhou Peng, Peng Wang, Junhui Wang, and Huamin Yang. "MF-CNN: traffic flow prediction using convolutional neural network and multi-features fusion." *IEICE TRANSACTIONS on Information and Systems* 102, no. 8 (2019): 1526-1536. https://doi.org/10.1587/transinf.2018EDP7330

[21] Babar, Muhammad, and Fahim Arif. "Real-time data processing scheme using big data analytics in internet of things based smart transportation environment." *Journal of Ambient Intelligence and Humanized Computing* 10 (2019): 4167-4177. https://doi.org/10.1007/s12652-018-0820-5

[22] Kim, Jiyeon, Sooyoung Jang, Deockgu Jee, Eunjin Ko, Seung Han Choi, and Mi Kyong Han. "5G based smartcity convergence service platform for data sharing." In *2020 international conference on information and communication technology convergence (ICTC)*, pp. 1522-1524. IEEE, 2020. https://doi.org/10.1109/ICTC49870.2020.9289155

[23] Nachmias, David, and Chava Nachmias. "Research methods in the social sciences." *(No Title)* (1981).

[24] Hamamurad, Qasim Hamakhurshid, Normal Mat Jusoh, and Uznir Ujang. "Factors Affecting Stakeholder Acceptance of a Malaysian Smart City." *Smart Cities* 5, no. 4 (2022): 1508-1535. https://doi.org/10.3390/smartcities5040077

[25] Bengio, Yoshua, Ian Goodfellow, and Aaron Courville. *Deep learning*. Vol. 1. Cambridge, MA, USA: MIT press, 2017.

[26] Yodo, Nita, and Pingfeng Wang. "Resilience analysis for complex supply chain systems using bayesian networks." In *54th AIAA Aerospace Sciences Meeting*, p. 0474. 2016. https://doi.org/10.2514/6.2016-0474