2023 International Conference Advancement in Data Science, E-learning and Information Systems (ICADEIS)
ISBN: 979-8-3503-0341-4

Implementation of YOLOv7 (You Only Look Once v7) Method for Traffic Density Detection

*Note: Sub-titles are not captured in Xplore and should not be used

1st Muhamad Nizam Azmi
dept. Informatics Enggineering, Mataram University
Jl. Majapahit 62, Mataram, Lombok, West Nusa Tenggara, INDONESIA
muhamadnizamazmi@gmail.com

2nd Fitri Bimantoro
dept. Informatics Enggineering, Mataram University
Jl. Majapahit 62, Mataram, Lombok, West Nusa Tenggara, INDONESIA
bimo@unram.ac.id

3th Ario Yudo Husodo
dept. Informatics Enggineering, Mataram University
Jl. Majapahit 62, Mataram, Lombok, West Nusa Tenggara, INDONESIA
ario@unram.ac.id

***Abstract*—Indonesia ranks 46th in the world congestion index. Various efforts have been made by the government to overcome this problem, especially in big cities in Indonesia. These problems can be in the form of accidents, pollution, and carbon emissions. The occurrence of this accident also resulted in congestion and many were caused by traffic violations and violating traffic lights. One solution to this congestion is by implementing smart traffic lights or Smart and Economical Traffic Lights (SMILE) with the implementation of computer vision to control the volume of vehicles at each intersection. The YOLOv7 method is used to detect vehicles automatically in real time. This method is one of the fastest and most accurate methods for object detection and can even exceed 2 times the ability of other object detection algorithms. This study tested the accuracy and detection results of YOLOv7 and YOLOv7-tiny in the morning, afternoon, evening, and night conditions at the 4 Pagesangan intersection, the comparison in the morning was around 49.5%, around 55% in the afternoon, around 38.1% in the evening and around 55.3% at night. The test results show that YOLOv7 provides a higher percentage of accuracy and detection results compared to YOLOv7-tiny in all-time conditions. The traffic density obtained in the morning is 42%, 29.3% during the day, 67% in the afternoon, and 41% at night. In conclusion, the YOLOv7 method can be applied to detect traffic density in real time and help overcome traffic congestion problems in big cities in Indonesia.**

Keywords—Computer Vision, YOLOv7, Real time, Vehichel Detection, Smart and Economical Traffic Light (SMILE)

I. Introduction

The development number of vehicles in Indonesia based on the type of vehicle in the form of passenger cars, buses, freight cars, and motorcycles by the end of 2022 is 151,424,276 units, where the data comes from the Indonesian Police Traffic Corps[1]. In 2021 the increase in the number of motorized vehicles in the Mataram area will increase from the previous year by 113 thousand where the number of motorized vehicles in 2021 is 530,570 units where the data is sourced from website Mataram City Transportation Service [2]. Traffic lights according to Law no. 22/2009 concerning road traffic and transportation: a traffic signaling device or (APILL) is a lamp that controls the flow of traffic installed at crossroads. One of the benefits of using traffic lights is to reduce congestion. This traffic jam can have a series of negative effects, starting with wasted waiting time, causing increased fuel consumption, which can then lead to increased gas emissions.[3]. Setting traffic lights is one of the important parameters in regulating and controlling the flow of vehicles in urban areas. The current traffic control system uses a time system, but there are deficiencies in this system, such as the time that does not match the density of vehicles at the intersection. For example, the green light is still on when the flow of vehicles at one of the intersections is empty, or when a dense traffic flow is only given a little green light [4]. With the increase in vehicles each year, especially in the city of Mataram, there will be an increase of around 27% in the number of vehicles in 2021, while the road width has not increased significantly, causing traffic jams. To overcome this problem, it is necessary to formulate an engineering strategy for traffic so that traffic jams don't get worse. Manipulation traffic requires data related to road and vehicle conditions, including knowing the number of vehicles on a road section from time to time [5]. YOLO method (You Only Look Once) is the fastest and most accurate method for object detection and can even exceed up to 2 times the ability of other object detection algorithms. YOLO has many versions that are often used, starting from YOLO, YOLOv2, and YOLOv3 to the latest, YOLOv7.[5]. Therefore, the authors propose the newest method for now, namely the YOLOv7 method because it is considered better than the previous version. The models to be used in this study are the YOLOv7 and YOLOv7-tiny to compare their accuracy and speed. YOLOv7 and YOLOv7-tiny are two object detection models (object detection) in computer vision developed by ultralytics.

A. Formulation of the Problem

- How to implement YOLOv7 and YOLOv7-tinyfor the vehicle detection process real time?.
- How to find out the traffic density using computer vision?

B. Research Objectives

- Implement the YOLOv7 and YOLOv7- methods to detect vehicles in real-time at traffic light intersections by considering the speed factor and detection accuracy.

- Applying computer vision to calculate traffic density using vehicle detection data obtained from the implementation of the YOLOv7 and YOLOv7-methods at the traffic light intersection.

II. LITERATUR REVIEW

A. Related Research

There are several studies related to image detection using the method You Only Look Once(YOLO). The first research made by Mawaddah Hope et al. has completed research with the title "Intelligent Monitoring System Traffic Flow Using the YOLOv3 Method "where in this study, the dataset is taken from 4 conditions, namely morning, afternoon, evening, and night. The result of object detection obtained from the system is sometimes inaccurate because it is influenced by several factors such as the distance of the camera and the size of the object captured by the camera. However, the system remains useful in monitoring traffic flow at crossroads with fairly high accuracy [6]. The second study made by Faqih Rofii et al. completed research entitled "Increasing the Accuracy of Calculation of the Number of Vehicles With Generating YOLOv4-Based Identity Detection SequencesDeep Neural Networks" Where in The study, the dataset is taken from 4 conditions namely morning, afternoon, evening, and Evening. Testing is done indirectly in real-time by using an algorithm YOLOv4. The results of this study indicate that in daytime conditions with a camera height of 1.5 meters and a 90° angle, total accuracy, precision, and recall obtained are 77%, 84%, and 85%. This shows that the method developed by Faqih Rofii et al. (2021) can improve the accuracy of calculations for a sufficient number of vehicles [5]. The several studies described previously explained that the method You Only Look Once (YOLO) is the most accurate method used for object detect and object classification in a manner real-time. However, the studies described previously, most of them used YOLO version 1 to version 4. In this study YOLOv7 was used for object detection a manner real-time then it will predict the estimated time based on the density of the vehicles detected by the system. YOLOv7 is the latest version of YOLO with higher processing speed and accuracy than other YOLO versions.

B. Supporting Theory Following

- SMILE (Small and Economical Traffic Light)

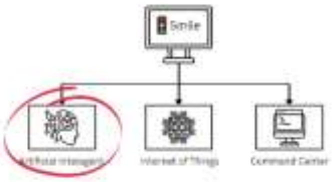

Fig. 1. Image of SMILE Field Distribution

SMILE (Smart and Economical Traffic Light) is an intelligent traffic light system that uses computer vision to see the volume of vehicles on the road will then be adjusted according to the duration of the traffic lights that will be given, the greater the volume of a vehicle in a lane, the longer the duration of the green light will be given to avoid congestion. SMILE is a final project product proposed by Informatics Engineering students at the University of Mataram. Based on Figure 1, the author takes part in artificial intelligence, namely the application of computer vision to detect traffic density real-time using method You Only Look Once(YOLO).

- YOLO (You Only Look Once)

You Only Look Once(YOLO) is one object detection single-stage other than SSD (Single Shot Detector). By using YOLO, object detection is done by looking at the problem as a regression problem to separate the box and the probability class associated with the bounding box such, as a neural network predicting the bounding box and class direct prediction of the entire image from one evaluation [16]. YOLO has several advantages over based systems classifiers. YOLO will generate prediction results with a single evaluation network, unlike the RCNN system which requires multiple networks to perform 10 processing of 1 image or image. Therefore, YOLO is said to be thousands of times faster than R-CNN and hundreds of times faster than fast R-CNN [16].

Fig. 2. Comparison of YOLOv7 Performance by Other Model Deep Learning

Based on Figure 2 is a working illustration of YOLO where initially the image will be split into several grids dimensions S x S. Then on each cell on grids it will be predicted by N possibilities bounding boxes and the probability value. Most bounding boxes it has a very low probability, so this algorithm will delete boxes that have a probability below a certain limit.

Fig. 3. Architerture of the YOLOv7 Model

Based on Figure 3 there are several divisions of the YOLOv7 model architecture including backbone, neck, and heads. Backbone In this layer there is a network of nerves convolutional which carries out the process of merging and forming image features on various types of image details. The layer backbone can be seen in Figure 3 where it consists of layers input, CBS layer, ELAN layer, and MPConv layer.

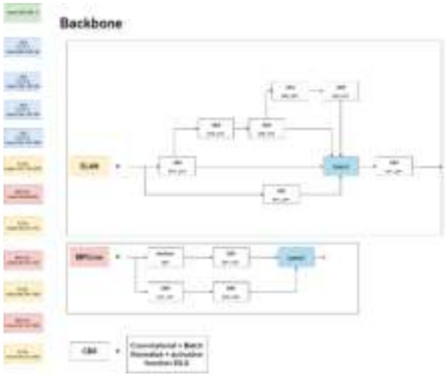

Fig. 4. Architerture of the YOLOv7 Model Backbone

Based on Figure 4 there is a CBS layer and a Concat layer where on CBS layer consists of 1 layer convolutions,1 layerBatch Normalization and the activation used is activation Sigmoid Linear Units(SiLU). Concat is a slicing layer and is used to slice the previous layer. Neck, In this layer there are a series of network layers that mix and combine image features and pass image features to the prediction layer.

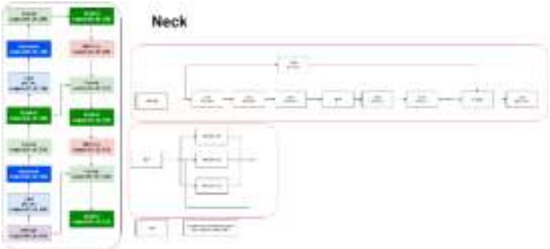

Fig. 5. Architerture of the YOLOv7 Model Neck

Based on Figure 5 there is layers Spatial Pyramid Pooling(SPP) which is the layer pooling which is used to remove the fixed size limitation on the network. SPPCPS is the development layer of the SPP layer.

III. Research Methodology

A. Tools and Materials

The tools used consist of hardware and software. Those tools are contained in Table II and Table III.

TABLE I. HARDWARE REQUIREMENTS IN RESEARCH

No	Hardware	Specification
1	Laptop	HP Pavilion Gaming Laptop 15 ec2010AX with AMD Ryzen 5 processor, NVIDIA GeForce RTX 30, and RAM 16GB

TABLE II. SOFTWARE REQUIREMENTS IN RESEARCH

No	Software	Specification
1	Operating system	Windows 11 64-bit
2	Programming Language	Python 3.8.0
3	Environment Programming	Visual Studio Code

B. System Design

- Download model YOLOv7

  The second step is to download the YOLOv7 Model along with the deployment code from GitHub (https://github.com/WongKinYiu/yolov7). Downloaded models are YOLOv7 and YOLOv7-tiny models for comparison of accuracy.

- Get Vehicle

  The third step is to separate objects with the vehicle and non-vehicle categories by determining the center of each vehicle as shown in Figure 6.

Fig. 6. Example Figure Get Vehicles

In Figure 3.4 there are 4 vehicles detected by YOLOv7 which can be seen in the figure green center on every detected vehicle. Midpoint or enter the red one is an object that is categorized as a vehicle and its position is on the right-hand side while the midpoint center those in green are objects that are not categorized as vehicles and are not vehicles that are not on the right side or facing the camera.

Fig. 7. Example Bounding Box

In the image used, it is known that the length and width of the image are 576 and 704. The equation for determining center x and y can be seen in equation (6) and equation (7).

$$centerX = x1 + \frac{(x2-x1)}{2} \qquad (6)$$

$$centerY = y1 + \frac{(y2-y1)}{2} \qquad (7)$$

So that the results obtained from the calculation of the 4 vehicles detected using equation (6) and equation (7).

- Set Condition

  The fourth step is to set the conditions to see the density of traffic. Based on previous research, the conditions that are regulated amount to 3 conditions, namely low, medium, and high. This process still uses sensors to detect these 3 conditions. [20].

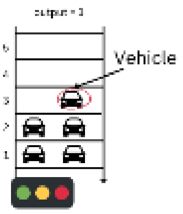

Fig. 8. Example Figure Dataset

Then this study used 5 conditions to determine traffic density where each condition must be filled by objects or vehicles. When condition 1 is not filled and condition 2 is filled, the output density is 1 because the object only fills 1 condition. The process of setting these conditions is done by taking the center of each detected vehicle. An illustration of managing vehicle density conditions in traffic can be seen in Figure 3.7

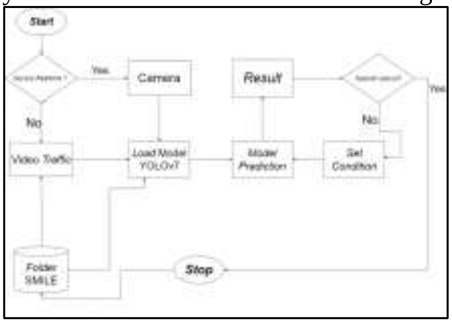

Fig. 9. Process Flow Diagram Testing

- Load YOLOv7 Model

The second step is to select the model used for the testing process that has been downloaded via the storage local computer. The models used for testing are 2 models from YOLOv7, namely the YOLOv7 model and the YOLOv7- model. tiny.

TABLE III. DIFFERENCES BETWEEN YOLOV7 AND YOLOV7-TINY

Model	*Parameters* (*million*)	FPS	AP *test* (%)
YOLOv7	36.9	161	51.4
YOLOv7-*tiny*	6.2	286	38.7

- Folder SMILE

Folder SMILE is somewhere to save files used from the start of the process of preparing until process testing. Folders it contains dataset used the YOLOv7 model and the YOLOv7-tiny modeling, the program used to build the SMILE system and the results are in the form of videos and images

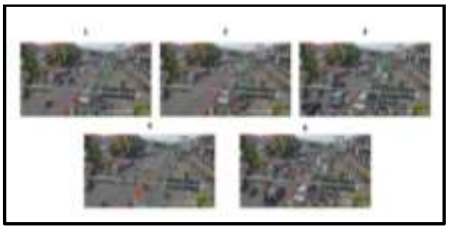

Fig. 10. Figure Example Condition 1-5

IV. Result and Discussion

- Sample Datasets

Dataset used is video data from Simpang 4 Pagesangan with non-static camera conditions (changing camera positions), where at this point 4 conditions were taken morning, afternoon, evening, and night with the camera type hkvision ptz 360 and each In this case, testing done by taking 10 frames of every video with the same range, Each video is about 1 minute long and has a height and width frame which is 1280x720. With total frames around 1400 frames, taking 10 frames can provide a sufficiently representative sample to perform tests on the object detection model. The purpose of using this dataset is to test the performance of the object detection model under non-static camera conditions and at different times. Thus, it can provide a better picture of model performance in more realistic situations and make it possible to better evaluate the accuracy of object detection.

- Test Result at Intersection 4 Pagesangan Morning Condition

Table IV consists of the density levels predicted by the YOLOv7 model and the supposed density conditions where the hypothesis of the prediction results and the accuracy of the predictions is the percentage accuracy of the density predicted by the model.

TABLE IV. ACCURACY OF PREDICTION OF SIMPANG 4 PAGESANGAN MORNING CONDITION

Number frames	Density Level	Condition should density	Prediction Accuracy (%)
1	4	5	80
2	1	5	20
3	1	5	20
4	1	5	20
5	2	5	40
6	3	5	60
7	3	5	60
8	3	5	60
9	3	5	60
10	0	5	0
	Total		42%

Based on Table IV it was found that the prediction accuracy of the 10 images produced a prediction accuracy value of 42% of the 10 samples used as experiments which were man every year. frames were taken randomly from a range of 1 minute.

TABLE V. ACCURACY COMPARISON OF YOLOV7 AND YOLOV7-TINY MORNING CONDITION

Condition	location	YOLOv7		YOLOv7-*tiny*	
		Percentage Accuracy (%) Average	Detection results (%) Average	Percentage Accuracy (%) Average	Detection results (%) Average
Morning	Intersection 4 Pagesangan	26,5%	76,3%	4,9%	15,9%

Based on Table V above, the results show that the average percentage of accuracy and average detection between the YOLOv7 and YOLOv7 models-tiny.

Fig. 11. Example of the Detection Results for the Comparison of YOLOv7 and YOLOv7 tiny Morning

- Test Result at Intersection 4 Pagesangan Daytime Condition

TABLE VI. ACCURACY OF PREDICTION OF SIMPANG 4 PAGESANGAN MORNING CONDITION

Number frames	Density Level	Condition should density	Prediction Accuracy (%)
1	1	4	25
2	0	4	0
3	0	4	0
4	1	4	25
5	1	3	33
6	2	4	50
7	2	4	50
8	2	4	50
9	3	5	60
10	0	5	0
	Total		29,3%

Based on Table VI, the prediction accuracy of the 10 images yielded a prediction accuracy value of 29.3% of the 10 samples used as an experiment. The accuracy of predicting traffic density at Simpang 4 yards this afternoon is not correct because the camera used is of low quality and the light received by the camera gives a bias that affects the accuracy of vehicle detection.

TABLE VII. ACCURACY COMPARISON OF YOLOV7 AND YOLOV7-TINY DAYTIME CONDITION

Condition	location	YOLOv7		YOLOv7-*tiny*	
		Percentage Accuracy (%) Average	Detection results (%) Average	Percentage Accuracy (%) Average	Detection results (%) Average
Morning	Intersection 4 Pagesangan	20,1%	75,1%	5,6%	23,14%

Based on Table X above, the results show that the average percentage of accuracy and average detection between the YOLOv7 and YOLOv7 models-tiny.

Fig. 12. Example of the Detection Results for the Comparison of YOLOv7 and YOLOv7 tiny Daytime

- Test Result at Intersection 4 Pagesangan Afternoon Condition

TABLE VIII. ACCURACY OF PREDICTION OF INTERSECTION 4 PAGESANGAN AFTERNOON CONDITION

Number frames	Density Level	Condition should density	Prediction Accuracy (%)
1	5	5	100
2	3	5	60
3	3	4	75
4	3	4	75
5	3	5	60
6	3	5	60
7	3	5	60
8	3	5	60
9	4	5	80
10	2	5	40
	Total		67%

Based on Table VIII, it was found that the prediction accuracy of the 10 images resulted in a prediction accuracy value of 67% of the 10 samples used as an experiment.

TABLE IX. ACCURACY COMPARISON OF YOLOV7 AND YOLOV7-TINY MORNING CONDITION

Condition	location	YOLOv7		YOLOv7-*tiny*	
		Percentage Accuracy (%) Average	Detection results (%) Average	Percentage Accuracy (%) Average	Detection results (%) Average
Morning	Intersection 4 Pagesangan	41,1%	79,2%	9,4%	16,4%

Based on Table XIII above, the results show that the average percentage of accuracy and average detection between the YOLOv7 and YOLOv7 models-tiny. The data used is a sample taken from 1 point dataset provided in the evening conditions. Accuracy comparison of YOLOv7 and YoloV7-tiny31.7% for the average percentage of traffic detection accuracy.

Fig. 13. Example of the Detection Results for the Comparison of YOLOv7 and YOLOv7 tiny Afternoon

- Test Result at Intersection 4 Pagesangan Night Condition

TABLE X. ACCURACY OF PREDICTION OF INTERSECTION 4 PAGESANGAN MORNING CONDITION

Number frames	Density Level	Condition should density	Prediction Accuracy (%)
1	2	4	50
2	2	5	40
3	2	5	40
4	2	5	40
5	2	5	40
6	2	5	40
7	3	5	60
8	3	5	60
9	2	5	40
10	0	5	0
	Total		41%

Based on Table X, it was found that the prediction accuracy of the 10 images resulted in a prediction accuracy value of 41% of the 10 samples used as an experiment.

TABLE XI. ACCURACY COMPARISON OF YOLOV7 AND YOLOV7-TINY NIGHT CONDIITION

Condition	location	YOLOv7		YOLOv7-*tiny*	
		Percentage Accuracy (%) Average	**Detection results (%) Average**	**Percentage Accuracy (%) Average**	**Detection results (%) Average**
Morning	Intersection 4 Pagesangan	21,1%	76,4%	3,4%	12,4%

Fig. 14. Example of the Detection Results for the Comparison of YOLOv7 and YOLOv7 tiny Night

V. CONCLUTIONS AND RECOMMENDATION

A. *Conclution*

Based on the problems described, 3 conclusions can be drawn:

- Research on CCTV detection testing at Simpang 4 Pagesangan shows that YOLOv7 and YOLOv7-tiny can be implemented for real-time vehicle detection using video data from CCTV taken under different conditions. However, detection accuracy can be affected by several factors such as camera quality and non-static camera position.
- Success in knowing traffic density using computer vision can be influenced by various factors, including camera quality, camera position, and light received by the camera. Therefore, careful testing and proper shooting conditions are required to ensure the accuracy of the traffic density prediction results.
- Research on CCTV detection testing at Simpang 4 Pagesangan shows that the YOLOv7 model has an average percentage of accuracy higher than the YOLOv7- model. tiny in morning, afternoon, evening, and night conditions, with a significant average percentage difference in accuracy

B. *Suggestion*

Based on the problems described, 3 can be drawn:

- Finetune the YOLOv7 model with a dataset that matches the problem to improve the detection accuracy of the YOLOv7 model.
- Comparing the performance of the model with other object detection models. In this study, only the performance of the YOLOv7 model with YOLOv7-tiny. It is better to do a comparison with other object detection models to see which one is more suitable for use in the case of object detection on the highway.

REFERENCES

[1] “Korlantas Polri - Informasi Lalu Lintas Indonesia.” https://korlantas.polri.go.id/ (diakses 21 Desember 2022).

[2] “Dinas Perhubungan Kota Mataram.” https://web.mataramkota.go.id/dinas-perhubungan (diakses 22 Desember 2022).

[3] N. Nurdjanah, “Emisi Co 2 Akibat Kendaraan Bermotor di Kota Denpasar,” Denpasar, 2015.

[4] M. I. Hermawan., I. I. Tritoasmoro., dan N. Ibrahim., “Pengaturan Lampu Lalu Lintas Berdasarkan Kepadatan Kendaraan Menggunakan Metode YOLO,” Feb 2021.

[5] F. Rofii, G. Priyandoko, M. I. Fanani, dan A. Suraji, “Peningkatan Akurasi Perhitungan Jumlah Kendaraan dengan Membangkitkan Urutan Identitas Deteksi Berbasis Yolov4 Deep Neural Networks,” *TEKNIK*, vol. 42, no. 2, hlm. 169–177, Agu 2021, doi: 10.14710/teknik.v42i2.37019.

[6] M. Harahap *dkk.*, “Sistem Cerdas Pemantauan Arus Lalu Lintas Dengan YOLO (You Only Look Once v3),” *Seminar Nasional APTIKOM*, hlm. 367–376, 2019.

[7] Y. Pratama dan E. Rasywir, “Eksperimen Penerapan Sistem Traffic Counting dengan Algoritma YOLO (You Only Look Once) V.4.,” *JURNAL MEDIA INFORMATIKA BUDIDARMA*, vol. 5, no. 4, hlm. 1438, Okt 2021, doi: 10.30865/mib.v5i4.3309.

[8] D. Ariyoga, R. Rahmadi, dan R. A. Rajagede, “Penelitian Terkini Tentang Sistem Pendeteksi Pelanggaran Lalu Lintas Berbasis Deep Learning: Sebuah Kajian Pustaka,” 2021.

[9] V. N. Sichkar dan S. A. Kolyubin, “Real time detection and classification of traffic signs based on YOLO Version 3 algorithm,” *Scientific and Technical Journal of Information Technologies, Mechanics and Optics*, vol. 20, no. 3, hlm. 418–424, Mei 2020, doi: 10.17586/2226-1494-2020-20-3-418-424.

[10] D. Bandukwala, M. Momin, A. Khan, A. Khan, dan Dr. L. Islam, “Object Detection using YOLO,” *Int J Res Appl Sci Eng Technol*, vol. 10, no. 5, hlm. 823–829, Mei 2022, doi: 10.22214/ijraset.2022.42088.

[11] S. Journal dan U. Journal, “International Journal of Innovative Technology and Exploring Engineering (IJITEE),” 2020, doi: 10.35940/ijitee.E3038.049620.

[12] N. Fadlia dan R. Kosasih, “Klasifikasi Jenis Kendaraan Menggunakan Metode Convolutional Neural Network (CNN),” *Jurnal Ilmiah Teknologi dan Rekayasa*, vol. 24, no. 3, hlm. 207–215, 2019, doi: 10.35760/tr.2019.v24i3.2397.

[13] M. T. Audina, F. Utaminingrum, dan D. Syauqi, “Sistem Deteksi dan Klasifikasi Jenis Kendaraan berbasis Citra dengan menggunakan Metode Faster-RCNN pada Raspberry Pi 4B,” 2021. [Daring]. Available: http://j-ptiik.ub.ac.id

[14] A. A. Oni dan N. Kajoh, “Video-Based Vehicle Counting System for Urban Roads in Nigeria Using Yolo and DCF-CSR Algorithms,” 2019.

[15] X. Li, Y. Liu, Z. Zhao, Y. Zhang, dan L. He, “A deep learning approach of vehicle multitarget detection from traffic video,” *J Adv Transp*, vol. 2018, 2018, doi: 10.1155/2018/7075814.

[16] P. Shinde, S. Yadav, S. Rudrake, dan P. Kumbhar, “Smart Traffic Control System using YOLO,” *International Research Journal of Engineering and Technology*, 2019.

[17] C.-Y. Wang, A. Bochkovskiy, dan H.-Y. M. Liao, “YOLOv7: Trainable bag-of-freebies sets new state-of-the-art for real-time object detectors,” 2021.

[18] Dr. Priyanto Hidayatullah, “Ebook YOLOv7,” hlm. 1–62, 2022.

[19] R. A. Hamzah., C. Setianingsih. R. A. Nugrahaeni, “Deteksi Pelanggaran Parkir Pada Bahu Jalan Tol Dengan Intelligent Transportation System Menggunakan Algoritma Faster R-Cnn,” *e-Proceeding of Engineering*, vol. 9, 2022.

[20] M. Reski., K. Budayawan, “Smart Traffic Light Berbasis Arduino,” vol. 9, 2021.