Optimization of YOLOv4-Tiny Algorithm for Vehicle Detection and Vehicle Count Detection Embedded System

Rachmat Muwardi[1], Ivan Prasetyo Nugroho[1], Ketty Siti Salamah[1], Mirna Yunita[2], Rizky Rahmatullah[3], Gregorius Justin Chung[2]
[1]Department of Electrical Engineering, Universitas Mercu Buana, Jakarta, Indonesia
[2]School of Computer Science and Technology, Beijing Institute of Technology, China
[3]School of Integrated Circuits and Electronics, Beijing Institute of Technology, Beijing, China

ARTICLE INFO

Article history:

Received August 02, 2024
Revised September 26, 2024
Published November 20, 2024

Keywords:

ARM Processor;
YOLO;
Median Filter;
Grayscale;
Object Detection

ABSTRACT

Currently, the implementation of object detection systems in the traffic sector is minimal. CCTV cameras on highways and toll roads are primarily used to monitor traffic conditions and document violations. However, the data recorded by these cameras can be further utilized to enhance traffic management systems. The author proposes a vehicle detection and counting system using YOLOv4-Tiny. The research aims to improve vehicle detection and counting accuracy by employing a median filter and grayscale processing, which simplify object detection. The proposed YOLOv4-Tiny algorithm has shown impressive results on various datasets, including MAVD, GRAM-RTM, and author dataset. The system achieved a detection accuracy of 98.95% on the MAVD dataset, 99.5% on the GRAM-RTM dataset (comparable to YOLOv4), and 99.1% on the author dataset. Furthermore, the system operates at 25 frames per second (FPS), a notably high rate compared to other methods. While the system demonstrates excellent accuracy in counting cars, it encounters some accuracy loss with other vehicle classifications. The author concludes that the system is highly suitable for real-world applications but notes that inaccurate labeling can lead to vehicle counting errors.

Corresponding Author:

Rachmat Muwardi, Department of Electrical Engineering, Universitas Mercu Buana, Jakarta, Indonesia
Email: rachmat.muwardi@mercubuana.ac.id

1. INTRODUCTION

The rapid advancement of technology has permeated various aspects of life, including city management, particularly in regulating traffic on urban toll roads. The increasing number of motor vehicle owners has led to higher traffic volumes on toll roads. Traffic congestion is especially prevalent at intersections following toll exits, where vehicles must take turns crossing the intersection, regulated by traffic lights. This scenario often results in long queues and traffic jams at intersections. The primary issue is that the traffic control systems at these intersections are usually ineffective, contributing to congestion on one or more sides of the intersection. The inefficiency in traffic management systems exacerbates the problem, causing significant delays and traffic bottlenecks [1] – [4], [40] – [42].

Yu Zhang *et al.* [5] propose using the Flip-Mosaic algorithm to enhance the network's perception of small targets. They established a multi-type vehicle target dataset collected in various scenarios and trained the detection model on this dataset. Experimental results indicate that the Flip-Mosaic data enhancement algorithm improves vehicle detection accuracy and reduces the false detection rate.

Gabriel Oltean *et al.* [6] present an approach for real-time vehicle counting using Tiny YOLO for detection and fast motion estimation for tracking. Their application runs on Ubuntu with GPU processing, and they plan to test it on low-budget devices like the Jetson Nano. Experimental results show their approach achieves high accuracy at real-time speeds (33.5 FPS) on real-traffic videos.

Madhusri Maity *et al.* [7] review numerous research projects focused on detecting and tracking vehicles, which have applications such as reducing fatal accidents caused by driver negligence, poor visibility during adverse weather conditions, or inadequate illumination. The paper comprehensively reviews existing methods based on Faster Region-based Convolutional Neural Networks (Faster R-CNN), and You Only Look Once (YOLO) architectures. The survey categorizes existing vehicle detection methods according to the architecture used (Faster R-CNN/YOLO) and is organized chronologically to highlight interrelations between methods. In addition to analyzing existing methods, the paper details the architectures of Faster R-CNN, YOLO, and their proposed variants. The paper concludes by listing the limitations of current works, unexplored aspects of this research area, and future research directions.

To tackle the identified traffic issues, a system that counts vehicles approaching intersections using digital image processing and YOLOv4-Tiny is required [8] – [11], [43] – [46]. Cameras installed on highways will provide the necessary data. This system aims to improve the efficiency and effectiveness of current traffic management systems [12] – [14], [47] – [49]. It will generate data on the number of vehicles passing through intersections, which can then be used to optimize traffic light durations for red, yellow, and green signals on each side of the intersection. Additionally, the collected data can be used by government agencies to manage and control vehicle flow in specific areas. The author's contribution ensures that YOLOv4-Tiny runs efficiently on an ARM Processor, which is suitable for vehicle detection on roads.

The primary contributions of this research are the integration of YOLOv4-Tiny with a median filter and grayscale preprocessing to improve vehicle object detection. By leveraging YOLOv4-Tiny, the system achieves a significant increase in frames per second (FPS), enhancing vehicle counting accuracy and reducing error rates. Moreover, the system enables simultaneous vehicle counting through bounding boxes that cross virtual lines, allowing for precise calculation of vehicle inputs and outputs. These improvements offer a reliable solution for traffic management and vehicle detection on urban toll roads.

2. METHODS

The author will offer a detailed and concise description of the methodologies intended for use in the proposed system.

2.1. Preprocess for Traffic System

The median filter is a non-linear digital filtering technique that removes noise from an image or signal [15] –[18], [50]. It works by moving a window (or kernel) across the image, replacing the center pixel in the window with the median value of all the pixels. The window size is typically a square of an odd number of pixels. The equation for the median filter can be expressed as follows:

$$\tilde{f}(x,y) = m\{f(a,b)\} \quad (1)$$

where $\tilde{f}(x,y)$is the output (filtered) image at position (x, y), $f(a,b)$ are the pixel values in the neighborhood of the pixel at (x,y), $m\{\}$ denotes the median value of the set of pixel values in the window centered at (x,y).

The grayscale conversion algorithm transforms a color image into a grayscale image by combining the color channels (typically red, green, and blue) into a single intensity value [19]-[21], [51]. This process involves calculating the luminance of each pixel based on its color components. Considering the human eye's sensitivity to different colors, the weighted sum method is the most commonly used formula for this conversion. The equation for converting a color image to grayscale is:

$$I = 0.299R + 0.587G + 0.114B \quad (2)$$

where I is the intensity value of the grayscale image, R, G and B are the red, green, and blue color channel values of the pixel, respectively. The coefficients 0.299, 0.587, and 0.114 reflect the human eye's sensitivity to red, green, and blue light, respectively.

2.2. Vehicle Detection

In vehicle detection, the author identifies vehicles to facilitate counting their numbers and minimize traffic congestion. The detection process employs the YOLOv4-Tiny algorithm [22] – [26]. Fig. 1 illustrates the structure of the YOLOv4-Tiny algorithm used in this system. The author will elaborate on the implementation of vehicle detection within the system to facilitate the accurate counting of vehicles. The detection results are optimized by utilizing two lines as input and output. Fig. 2 presents the framework for vehicle detection.

The preprocessing steps for vehicle detection include applying a median filter algorithm and a Grayscale algorithm. Given the high speed at which vehicles move, the median filter is crucial for minimizing interference and producing a more precise image than the original. Following this, the conversion to grayscale increases the detection accuracy further. Grayscale conversion simplifies feature detection, significantly boosting accuracy compared to no preprocessing. Together, these preprocessing techniques ensure the vehicle detection system is more reliable and efficient, leading to better performance in dynamic environments.

In this research, vehicle detection was enhanced through a Median filter, clarifying unclear images, and the application of Greyscale, simplifying the conversion of RGB images to black and white. This approach facilitates more straightforward and accurate color calculations. As a result, the system achieves a higher frames per second (FPS) rate, positively impacting the accuracy of vehicle count calculations.

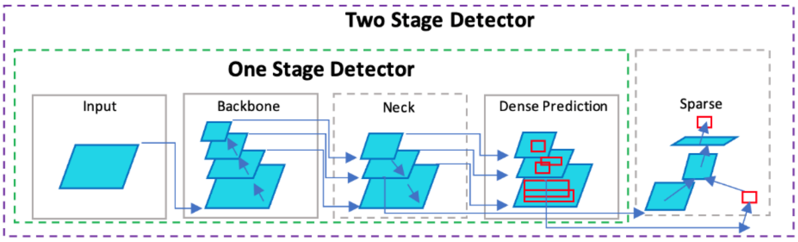

Fig. 1. YOLOv4- Tiny Architectur

Fig. 2. Vehicle Detection Framework

2.3. Vehicle Count Detection

In this section, the author will describe the vehicle count detection process, which employs two virtual lines positioned at the top and bottom of the detection area. These lines act as sensors to detect vehicles as they cross [27] - [29]. The upper line, or input line, initially processes vehicles that pass through it. Subsequently, the lower line, or output line, verifies these vehicles through classification in the vehicle detection system. For example, if the input line detects a car, the output line will verify and count the same vehicle. This method ensures better data storage, improving accuracy and effectiveness, particularly when multiple vehicles pass simultaneously. The bounding box feature significantly aids in the accurate counting of vehicles. Fig. 3. illustrates the two lines serving as input and output sensors.

Fig. 3. Vehicle Count Detection

2.4. Propose Method

The author outlines the proposed Traffic System, designed to enhance vehicle detection efficiency and streamline the system's detection processes. Fig. 4 provides a detailed illustration of the Traffic System framework.

Fig. 4. Traffic System Framework

The camera will be mounted at a height of 4 meters, which is considered optimal for ensuring precise vehicle detection. The system processes the camera feed and classifies vehicles using bounding boxes, allowing the vehicle count sensor to operate with high efficiency and accuracy. When an object is detected by the input sensor, the bounding box is triggered to determine if the object has crossed the input line. The process then continues to the output line, verifying whether the vehicle has successfully passed through it. The vehicle count will be displayed in the top left corner of the image, streamlining the counting process and improving accuracy.

3. RESULTS AND DISCUSSION

In the results and discussion section, the author will elaborate on the evaluation process of using the dataset for vehicle detection and compare the proposed algorithm against other algorithms. The author will use three datasets: MAVD [30], GRAM-RTM [31], and the author's dataset. Four algorithms will be compared: YOLOv3 [32], YOLOv4 [33], Faster RCNN [34], and YOLOv4-Tiny. A detailed and comprehensive explanation of the results will be provided.

For testing, the author will employ the Confusion Matrix Method, which is highly suitable for comparing image processing outcomes [35] – [37]. This method will yield metrics such as Recall, Precision, Accuracy, and MAP (Mean Average Precision). These results serve as standards to demonstrate the efficacy and applicability of the proposed algorithm for the intended system. Additionally, the author will explain several equations used in the Confusion Matrix Method.

$$Recall = \frac{TP}{TP + FN} \tag{3}$$

$$Precision = \frac{TP}{TP + FP} \tag{4}$$

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \tag{5}$$

3.1. Dataset and Training

The dataset includes four class classifications: Motorcycles, Cars, Trucks, and Buses, as shown in the classification table [38] – [39]. All images will be resized to standard YOLOv4-Tiny to facilitate more effective implementation of the training data in Table 1. This model was trained and evaluated on an Nvidia RTX 3090Ti GPU.

Table 1. Classification Dataset

No	Vehicle Class	Image
1	Motorcycle	
2	Car	
3	Truck	
4	Bus	

The next phase of the Traffic System experiment will involve testing on the Jetson Nano Microprocessor. The model underwent training for 10,000 steps, with an increasing learning rate at each iteration to aid loss convergence. Weight reduction techniques were employed to prevent overfitting. Data augmentation enhanced the dataset by modifying the saturation, exposure, and hue of the images and rotating the images at various angles of Vehicle Detection.

3.2. Vehicle Detection

In this experiment, the author will perform a comparative analysis using several datasets, including MAVD, GRAM-RTM, and an author's dataset. These datasets will be renamed according to the author's specifications. The experiment involves conducting ten trials to ensure the reliability and consistency of the detection results and accuracy. To obtain the results, the author employs a confusion matrix, as illustrated in Table 2. This matrix details the calculations and comparisons using four different algorithms. Fig. 5 presents the vehicle detection results.

Table 2. Vehicle Detection Result

Data Set	Algorithm	Recall	Precision	Accuracy	FPS
MAVD [30]	YOLOv3 [32]	0.9566	0.9818	0.9766	15
	YOLOv4 [33]	0.9833	0.9916	0.9891	15
	Faster RCNN [34]	0.9600	0.9756	0.9721	5
	YOLOv4-Tiny	0.9790	0.9850	0.9895	25
GRAM-RTM [31]	YOLOv3 [32]	0.9700	0.9770	0.9802	15
	YOLOv4 [33]	0.9930	0.9940	0.9950	15
	Faster RCNN [34]	0.9090	0.8330	0.9154	5
	YOLOv4-Tiny	0.9917	0.9917	0.9950	25
Author Dataset	YOLOv3 [32]	0.9700	0.9770	0.9800	15
	YOLOv4 [33]	0.9850	0.9860	0.9910	15
	Faster RCNN [34]	0.9710	0.9710	0.9770	5
	YOLOv4-Tiny	0.9860	0.9930	0.9915	25

The results indicate that YOLOv4-Tiny achieves excellent accuracy. Specifically, the GRAM-RTM dataset performs comparably to YOLOv4, scoring 99.5%. Both algorithms yield excellent results. For the author's dataset, YOLOv4-Tiny outperforms the other algorithms with an accuracy of 99.15%. The primary advantage of YOLOv4-Tiny lies in its more straightforward structure, which results in very high FPS and significantly enhances the accuracy of vehicle count detection. Fig. 5 explains the results from the Traffic System. The author presents the findings from applying the YOLOv4-Tiny algorithm to the proprietary dataset. These results are detailed in the confusion matrix in Table 3, which includes four classifications. The model was rigorously tested over 1,000 epochs in a single test, yielding highly accurate and reliable results.

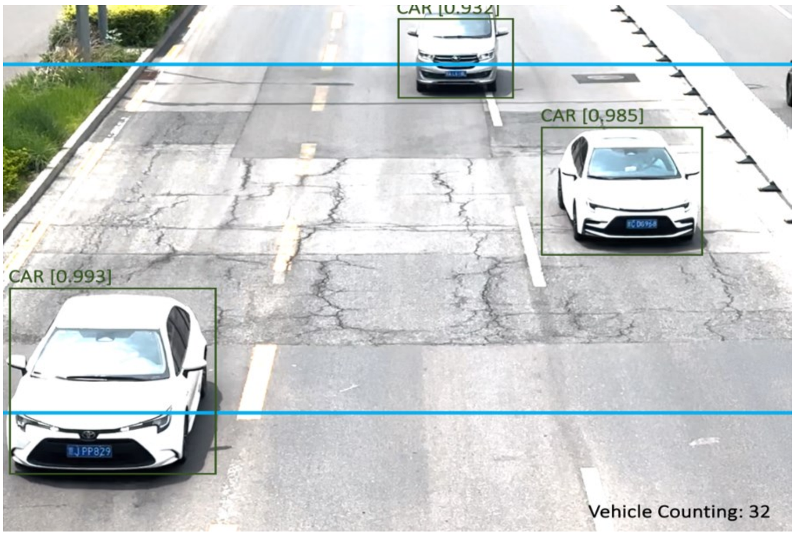

Fig. 5. Result Traffic System

Table 3. Confusion Matrix Result

Actual\Predicted	MC	C	T	B
MC	19720	108	84	88
C	161	29580	139	120
T	58	120	5916	66
B	52	71	66	3944

3.3. Vehicle Count Detection

The author conducted this experiment by creating five videos to ensure the results were appropriate and accurate. Each video features 50-60 vehicles, and the author will analyze the accuracy of the detection calculations. Table 4 illustrates the vehicle count detection results.

Table 4. Vehicle Count Detection Result

Video	On Video					On Vehicle Count Detection				
	MC	C	T	B	Total	MC	C	T	B	Total
1	15	30	3	5	53	13	30	2	5	50
2	20	30	4	6	60	19	30	4	4	57
3	18	25	2	6	51	15	25	2	4	46
4	8	35	2	5	50	7	35	2	4	48
5	13	32	3	7	55	13	32	3	6	54

The author achieved excellent results in this vehicle count detection experiment. However, a challenge arose due to errors in vehicle class naming. The input and output line processes experienced changes in classification within the bounding box, leading to miscalculations and inaccuracies in vehicle detection. While car accuracy remained consistent and satisfactory, other classifications encountered errors, resulting in less accurate calculations.

4. CONCLUSION

The author developed a Traffic System using the YOLOv4-Tiny algorithm in this research. The system achieved excellent results due to the implementation of the Median Filter and Grayscale Method, which facilitated the detection process with YOLOv4-Tiny. The author also significantly improved FPS by employing the NVIDIA JASTON NANO as the microprocessor for running the Traffic System. The author conducted experiments to count vehicles using a virtual line sensor, incorporating two lines to predict the number of passing vehicles. The vehicle detection results were highly satisfactory across three different datasets. Using the author's dataset, an accuracy of 99.11% and an FPS of 25 were achieved, which are outstanding results for vehicle detection. With the MAVD dataset, the accuracy was 98.95%, further demonstrating the effectiveness of the YOLOv4-Tiny algorithm in creating a Traffic System.

Regarding vehicle classification, the system achieved 100% accuracy for cars. However, other classifications were less accurate due to errors in labeling from input to output. The author concludes that the accuracy of vehicle count detection can be improved with more precise label naming and by integrating additional methods to enhance YOLOv4-Tiny. The author hopes that this system can be further developed and refined in the future.

Acknowledgments

Special thanks are extended to Universitas Mercu Buana for funding this research. Additionally, I am grateful to all my colleagues who contributed to the collaborative research efforts between UMB and BIT. I look forward to the opportunity to participate in future joint research endeavors.

REFERENCES

[1] N. K. Jain, R. K. Saini and P. Mittal, "A Review on Traffic Monitoring System Techniques," in *Soft Computing: Theories and Applications*, pp. 569-577, 2018, https://doi.org/10.1007/978-981-13-0589-4_53.

[2] A. S. Putra and H. L. H. S. Warnars, "Intelligent Traffic Monitoring System (ITMS) for Smart City Based on IoT Monitoring," in *2018 Indonesian Association for Pattern Recognition International Conference (INAPR)*, pp. 161-165), 2018, https://doi.org/10.1109/INAPR.2018.8626855.

[3] R. Muwardi, H. Gao, H. U. Ghifarsyam, M. Yunita, A. Arrizki and J. Andika, "Network Security Monitoring System Via Notification Alert," *Journal of Integrated and Advanced Engineering (JIAE),* vol. 1, no. 2, pp. 113-122, 2021, https://doi.org/10.51662/jiae.v1i2.22.

[4] S. Budiyanto, F. A. Silaban, L. M. Silalahi, T. Pangaribowo, M. H. I. Hajar, A. Sepbrian, R. Muwardi and G. Hongmin, "The automatic and manual railroad door systems based on IoT," *Indonesian Journal of Electrical Engineering and Computer Science,* vol. 21, no. 3, pp. 1847-1855, 2021, https://doi.org/10.11591/ijeecs.v21.i3.pp1847-1855.

[5] Y. Zhang, Z. Guo, J. Wu, Y. Tian, H. Tang and X. Guo, "Real-time vehicle detection based on improved yolo v5," *Sustainability,* vol. 14, no. 9, p. 12274, 2022, https://doi.org/10.3390/su141912274.

[6] G. Oltean, C. Florea, R. Orghidan and V. Oltean, "Towards Real Time Vehicle Counting using YOLO-Tiny and Fast Motion Estimation," in *2019 IEEE 25th International Symposium for Design and Technology in Electronic Packaging (SIITME)*, pp. 240-243, 2019, https://doi.org/10.1109/SIITME47687.2019.8990708.

[7] M. Maity, S. Banerjee and S. S. Chaudhuri, "Faster r-cnn and yolo based vehicle detection: A survey," in *2021 5th International Conference on Computing Methodologies and Communication (ICCMC)*, pp. 1442-1447, 2021, https://doi.org/10.1109/ICCMC51019.2021.9418274.

[8] R. Muwardi, H. Qin, H. Gao, H. U. Ghifarsyam, M. H. I. Hajar and M. Yunita, "Research and Design of Fast Special Human Face Recognition System," in *2020 2nd International Conference on Broadband Communications, Wireless Sensors and Powering (BCWSP)*, pp. 68-73, 2020, https://doi.org/10.1109/BCWSP50066.2020.9249452.

[9] R. Muwardi, M. Yunita, H. U. Ghifarsyam and H. Juliyanto, "Optimize Image Processing Algorithm on ARM Cortex-A72 and A53," *Jurnal Ilmiah Teknik Elektro Komputer dan Informatika (JITEKI),* vol. 8, no. 3, pp. 399-409, 2022, http://dx.doi.org/10.26555/jiteki.v8i3.24457.

[10] Yuliza, R. Muwardi, M. Rhozaly, Lenni, M. Yunita and G. E. Yehezkiel, "Fast Human Recognition System on Real-Time Camera," *Jurnal Ilmiah Teknik Elektro Komputer dan Informatika (JITEKI),* vol. 9, no. 4, pp. 895-903, 2023, https://doi.org/10.26555/jiteki.v9i4.27009.

[11] D. P. Carrasco, H. A. Rashwan, M. Á. García and D. Puig, "T-YOLO: Tiny Vehicle Detection Based on YOLO and Multi-Scale Convolutional Neural Networks," *IEEE Access,* pp. 22430-22440, 2021, https://doi.org/10.1109/ACCESS.2021.3137638.

[12] S. Javaid, A. Sufian, S. Pervaiz and M. Tanveer, "Smart traffic management system using Internet of Things," in *2018 20th International Conference on Advanced Communication Technology (ICACT)*, pp. 393-398, 2018, https://doi.org/10.23919/ICACT.2018.8323770.

[13] O. Avatefipour and F. Sadry, "Traffic Management System Using IoT Technology - A Comparative Review," in *2018 IEEE International Conference on Electro/Information Technology (EIT)*, pp. 1041-1047, 2018, https://doi.org/10.1109/EIT.2018.8500246.

[14] H. Hamidi and A. Kamankesh, "An Approach to Intelligent Traffic Management System Using a Multi-agent System," *International Journal of Intelligent Transportation Systems Research,* vol. 16, pp. 112-124, 2018, https://doi.org/10.1007/s13177-017-0142-6.

[15] V. Arulkumar, S. J. Prakash, E. Subramanian and N. Thangadurai, "An Intelligent Face Detection by Corner Detection using Special Morphological Masking System and Fast Algorithm," in *2021 2nd International Conference on Smart Electronics and Communication (ICOSEC)*, pp. 1556-1561, 2021, https://doi.org/10.1109/ICOSEC51865.2021.9591857.

[16] F. Orujov, R. Maskeliūnas, R. Damaševičius and W. Wei, "Fuzzy based image edge detection algorithm for blood vessel detection in retinal images," *Applied Soft Computing,* vol. 94, p. 106452, 2020, https://doi.org/10.1016/j.asoc.2020.106452.

[17] G. Yang, W. Feng, J. Jin, Q. Lei, X. Li, G. Gui and W. Wang, "Face Mask Recognition System with YOLOV5 Based on Image Recognition," in *2020 IEEE 6th International Conference on Computer and Communications (ICCC)*, pp. 1398-1404, 2020, https://doi.org/10.1109/ICCC51575.2020.9345042.

[18] V. H. D. Priya and A. V. Juliet, "Automatic Detection of Covid-19 based on Xception Network with Optimized CNN," *Computers and Computing,* vol. 70, no. 1, pp. 372-380, 2022, https://doi.org/10.1080/03772063.2022.2138583.

[19] C.-M. Hsu, C.-C. Hsu, Z.-M. Hsu, F.-Y. Shih, M.-L. Chang and T.-H. Chen, "Colorectal Polyp Image Detection and Classification through Grayscale Images and Deep Learning," *Sensors,* vol. 21, no. 18, p. 5995, 2021, https://doi.org/10.3390/s21185995.

[20] M. Hagara, R. Stojanović, T. Bagala, P. Kubinec and O. Ondráček, "Grayscale image formats for edge detection and for its FPGA implementation," *Microprocessors and Microsystems,* vol. 75, p. 103056, 2020, https://doi.org/10.1016/j.micpro.2020.103056.

[21] A. Singh, G. Sethi and G. S. Kalra, "Spatially Adaptive Image Denoising via Enhanced Noise Detection Method for Grayscale and Color Images," *IEEE Access,* vol. 8, pp. 112985-113002, 2020, https://doi.org/10.1109/ACCESS.2020.3003874.

[22] Z. Jiang, L. Zhao, S. Li and Y. Jia, "Real-time object detection method based on improved YOLOv4-tiny," *arXiv preprint arXiv:2011.04244,* 2020, https://doi.org/10.48550/arXiv.2011.04244.

[23] C. Guo, X.-l. Lv, Y. Zhang and M.-l. Zhang, "Improved YOLOv4-tiny network for real-time electronic component detection," *Scientific Report,* vol. 11, no. 1, p. 22744, 2021, https://doi.org/10.1038/s41598-021-02225-y.

[24] M. Cui, Y. Lou, Y. Ge and K. Wang, "LES-YOLO: A lightweight pinecone detection algorithm based on improved YOLOv4-Tiny network," *Computers and Electronics in Agriculture,* vol. 205, p. 107613, 2023, https://doi.org/10.1016/j.compag.2023.107613.

[25] L. Wang, K. Zhou, A. Chu, G. Wang and L. Wang, "An Improved Light-Weight Traffic Sign Recognition Algorithm Based on YOLOv4-Tiny," *EEE Access,* vol. 9, pp. 124963 - 124971, 2021, https://doi.org/10.1109/ACCESS.2021.3109798.

[26] R. Muwardi, A. Faizin, P. D. P. Adi, R. Rahmatullah, Y. Wang, M. Yunita and D. Mahabror, "Design Human Object Detection Yolov4-Tiny Algorithm on ARM Cortex-A72 and A53," *Jurnal Ilmiah Teknik Elektro Komputer dan Informatika (JITEKI),* vol. 9, no. 4, pp. 1168-1178, 2023, https://doi.org/10.26555/jiteki.v9i4.27402.

[27] V. Mandal and Y. Adu-Gyamfi, "Object Detection and Tracking Algorithms for Vehicle Counting: A Comparative Analysis," *Journal of Big Data Analytics in Transportation,* vol. 2, p. 251–261, 2020, https://doi.org/10.1007/s42421-020-00025-w.

[28] C. Liu, D. Q. Huynh, Y. Sun, M. Reynolds and S. Atkinson, "A Vision-Based Pipeline for Vehicle Counting, Speed Estimation, and Classification," *IEEE Transactions on Intelligent Transportation Systems ,* vol. 22, no. 12, pp. 7547 - 7560, 2021, https://doi.org/10.1109/TITS.2020.3004066.

[29] J. Azimjonov and A. Özmen, "A real-time vehicle detection and a novel vehicle tracking systems for estimating and monitoring traffic flow on highways," *Advanced Engineering Informatics,* vol. 50, p. 101393, 2021, https://doi.org/10.1016/j.aei.2021.101393.

[30] B. Citamak, O. Caglayan, M. Kuyu, E. Erdem, A. Erdem, P. Madhyastha and L. Specia, "MSVD-Turkish: a comprehensive multimodal video dataset for integrated vision and language research in Turkish," *Machine Translation ,* vol. 35, p. 265–288, 2021, https://doi.org/10.1109/SIU.2019.8806555.

[31] R. J. López-Sastre, C. Herranz-Perdiguero, R. Guerrero-Gómez-Olmedo, D. Oñoro-Rubio and S. Maldonado-Bascón, "Boosting Multi-Vehicle Tracking with a Joint Object Detection and Viewpoint Estimation Sensor," *Sensors,* vol. 19, no. 19, p. 4062, 2019, https://doi.org/10.3390/s19194062.

[32] A. T. Tajar, A. Ramazani and M. Mansoorizadeh, "A lightweight Tiny-YOLOv3 vehicle detection approach," *Journal of Real-Time Image Processing,* vol. 18, p. 2389–2401, 2021, https://doi.org/10.1007/s11554-021-01131-w.

[33] C.-J. Lin, S.-Y. Jeng and o.-W. Lioa, "A Real-Time Vehicle Counting, Speed Estimation, and Classification System Based on Virtual Detection Zone and YOLO," *Lin, C. J., Jeng, S. Y., & Lioa, H. W. (2021). A Real-Time Vehicle Counting, Speed Estimation, and Classification System Based on Virtual Detection Zone and YOLO. Mathematical Problems in Engineering,* vol. 1, p. 1577614, 2021, https://doi.org/10.1155/2021/1577614.

[34] S. Ren, K. He, R. Girshick and J. Sun, "Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks," *IEEE Transactions on Pattern Analysis and Machine Intelligence,* vol. 39, no. 6, pp. 1137-1149, 2017, https://doi.org/10.1109/TPAMI.2016.2577031.

[35] M. Grandini, E. Bagli and G. Visani, "Metrics for Multi-Class Classification: an Overview," *arXiv preprint arXiv:2008.05756,* 2020, https://doi.org/10.48550/arXiv.2008.05756.

[36] M. Heydarian, T. E. Doyle and R. Samavi, "MLCM: Multi-Label Confusion Matrix," *IEEE Access ,* vol. 10, pp. 19083-19095, 2022, https://doi.org/10.5121/csit.2020.100801.

[37] A. Theissler, M. Thomas, M. Burch and F. Gerschner, "ConfusionVis: Comparative evaluation and selection of multi-class classifiers based on confusion matrices," *Knowledge-Based Systems,* vol. 247, p. 108651, 2022, https://doi.org/10.1016/j.knosys.2022.108651.

[38] M. Yunita, R. Muwardi and Z. Iklima, "Implementation of Bayesian inference MCMC algorithm in phylogenetic analysis of Dipterocarpaceae family," *SINERGI,* vol. 27, no. 1, pp. 23-30, 2023, https://doi.org/10.26555/jiteki.v9i4.27009.

[39] Z. Iklima, B. N. Rohman, R. Muwardi, A. Khan and Z. Arifiansyah, "Defect classification of radius shaping in the tire curing process using Fine-Tuned Deep Neural Network," *SINERGI,* vol. 26, no. 3, pp. 335-342, 2022, http://dx.doi.org/10.22441/sinergi.2022.3.009.

[40] I. Bisio, C. Garibotto, H. Haleem, F. Lavagetto and A. Sciarrone, "A Systematic Review of Drone Based Road Traffic Monitoring System," *IEEE Access*, vol. 10, pp. 101537-101555, 2022. https://doi.org/10.1109/ACCESS.2022.3207282.

[41] V. Mandal, A. R. Mussah, P. Jin and a. Adu-Gyamfi, "Artificial Intelligence-Enabled Traffic Monitoring System," *Sustainability*, vol. 12, no. 21, p. 9177, 2020. https://doi.org/10.3390/su12219177.

[42] M. A. A. Al-qaness, A. A. Abbasi, H. Fan, R. A. Ibrahim, S. H. Alsamhi and A. Hawbani, "An improved YOLO-based road traffic monitoring system," *Computing*, vol. 103, no. 2, pp. 211-230, 2021. https://doi.org/10.1007/s00607-020-00869-8.

[43] R. Muwardi, A. Faizin, P. D. P. Adi, R. Rahmatullah, Y. Wang, M. Yunita and D. Mahabror, "Design Human Object Detection Yolov4-Tiny Algorithm on ARM Cortex-A72 and A53," *Jurnal Ilmiah Teknik Elektro Komputer dan Informatika (JITEKI),* vol. 9, no. 4, pp. 1168-1178, 2023. https://doi.org/10.26555/jiteki.v9i4.27402.

[44] A. Amrouche, Y. Bentrcia, A. Abed and N. Hezil, "Vehicle Detection and Tracking in Real-time using YOLOv4-tiny," *in 2022 7th International Conference on Image and Signal Processing and their Applications (ISPA)*, pp. 1-5, 2022. https://doi.org/10.1109/ISPA54004.2022.9786330.

[45] X. Xiang, N. L. Fanda Meng and H. Yin, "Engineering Vehicles Detection for Warehouse Surveillance System Based on Modified YOLOv4-Tiny," *Neural Process Lett*, vol. 3, pp. 2743-2759, 2023. https://doi.org/10.1007/s11063-022-10982-8.

[46] A. A. Alsanabani, S. A. Saeed, M. Al-Mkhlafi and M. Albishari, "A Low Cost and Real Time Vehicle Detection Using Enhanced YOLOv4-Tiny," *in 2021 IEEE International Conference on Artificial Intelligence and Computer Applications (ICAICA)*, pp. 372-377, 2021. https://doi.org/10.1109/ICAICA52286.2021.9498188.

[47] N. Nigam, D. P. Singh and J. Choudhary, "A Review of Different Components of the Intelligent Traffic Management System (ITMS)," *Symmetry*, vol. 15, no. 3, p. 583, 2023. https://doi.org/10.3390/sym15030583.

[48] A. Khanna, R. Goyal, M. Verma and D. Joshi, "Intelligent Traffic Management System for Smart Cities," in *Intelligent traffic management system for smart cities. In Futuristic Trends in Network and Communication Technologies: First International Conference, FTNCT* 2018, pp. 152-164, 2019. https://doi.org/10.1007/978-981-13-3804-5_12.

[49] M. Rath, "Smart Traffic Management System for Traffic Control using Automated Mechanical and Electronic Devices," in *IOP Conference Series: Materials Science and Engineering*, vol. 377, no. 1, p. 012201, 2018. https://doi.org/10.1088/1757-899X/377/1/012201.

[50] S. Mammeri, M. Amroune, M.-Y. Haouam, I. Bendib and A. C. Silva, "Early detection and diagnosis of lung cancer using YOLO v7, and transfer learning," *Multimedia Tools and Applications,* vol. 83, no. 10, pp. 30965-30980, 2024. https://doi.org/10.1007/s11042-023-16864-y.

[51] A. L. R. Siriani, V. Kodaira, S. A. Mehdizadeh, I. d. A. Nääs, D. J. d. Moura and D. F. Pereira, "Detection and tracking of chickens in low-light images using YOLO network and Kalman filter," *Neural Computing and Applications*, vol. 34, no. 24, pp. 21987-21997, 2022. https://doi.org/10.1007/s00521-022-07664-w.

BIOGRAPHY OF AUTHORS

Rachmat Muwardi, is currently a Lecturer in the Department of Electrical Engineering, Universitas Mercu Buana, Jakarta, Indonesia. He graduated from the Beijing Institute of Technology in 2020 with a Master's in Electronic Science and Technology. Currently, He declared as a recipient of a China Scholarship Council (CSC) to continue his doctoral program at the Beijing Institute of Technology in September 2022, majoring in Optical Engineering. During his undergraduate, he received a double degree scholarship from Universitas Mercu Buana and Beijing Institute of Technology in Electrical Engineering and Computer Science. His research interest is Object Detection, Target Detection, and Embedded System. He can be contacted at email: rachmat.muwardi@mercubuana.ac.id.

Ivan Prasetyo Nugroho, is currently an iOS Developer in Ubersnap Pte Ltd. He graduated from Universitas Mercu Buana Jakarta in 2023 with Bachelor's degree in Electrical Engineering at Control System major. He finished his Bachelor's study by his research interested in Object Detection using YOLOv4-tiny, OpenCV and python. Before He continue his study in Universitas Mercu Buana Jakarta, He graduated from Akademi Teknologi Warga Surakarta in 2019 with Diploma's degree in Electrical Engineering at Mechatronics major. He completes his study by creating Gimbal Camera using Inertial Measurement Unit (IMU) sensor, servo motor, and Arduino as the controller. He has interest in software development, and currently He focuses his career in Mobile Application development majoring in iOS platform. He experienced developing 4 iOS applications that required to communicate to hardware such as printer, scanner, and DSLR camera. He can be contacted at email: ivanprasetyo559@gmail.com.

Mirna Yunita, received a Master's in Computer Science and Technology from Beijing Institute of Technology, Beijing, China. Currently, as a Ph.D. student at the School of Computer Science and Technology, Beijing Institute of Technology, China. She is interested in related topics in Machine Learning, Web Development, Data Mining, and Bioinformatics. She was a Frontend and Mobile Application Developer in a Logistics & Supply Chain company in Jakarta, Indonesia. She can be contacted at email: mirnayunita@bit.edu.cn.

Rizky Rahmatullah, he is currently a master's degree at the 3School of Integrated Circuits and Electronics, Beijing Institute of Technology (BIT), Beijing, China and works as a researcher at the Indonesian National Research and Innovation Agency. He is interested in research on the Internet of Things, Wireless Communication and Antennas. He can be contacted at email: rizk032@brin.go.id/3820231111@bit.edu.cn.

Ketty Siti Salamah, is currently a Lecturer in the Department of Electrical Engineering, Universitas Mercu Buana, Jakarta, Indonesia. She graduated from the Universitas Mercu Buana in 2016 with a Master's in Electrical Engineering. Her research interest is Wireless Sensor Network, Internet of Things, Image Processing. She can be contaced at email: kettysitisalamah@mercubuana.ac.id.

Gregorius Justin Chung, is currently pursuing master program at Beijing Institute of Technology majoring in Computer Science and Technology. He graduated from Beijing Institute of Technology majoring in Mathematics and Applied Mathematics, obtaining Bachelor of Science. His research interest is Data Science Application in Economic and Financial Fields, Outlier Detection and Quantitative Strategy. He can be contacted at email: gregoriusjustin@bit.edu.cn.