Optimization of Vehicle Detection at Intersections Using the YOLOv5 Model

I Wayan Adi Artha Wiguna[1], Roy Rudolf Huizen[2], Gede Angga Pradipta[2]
[1]Magister Program, Departement of Magister Information Systems, Institut Teknologi dan Bisnis STIKOM Bali, Indonesisa,
[2]Departement of Magister Information Systems, Institut Teknologi dan Bisnis STIKOM Bali, Indonesisa

ARTICLE INFO

Article history:

Received July 02, 2024
Revised September 02, 2024
Published January 15, 2025

Keywords:

Traffic Jam;
Object Detection;
YOLOv5;
Model Optimization

ABSTRACT

This study aims to analyze and evaluate the performance of the YOLOv5 model in detecting vehicles at intersections to optimize traffic flow. The methods used in this research include training the YOLOv5 model with traffic datasets collected from various intersections and optimizing hyperparameters to achieve the best detection accuracy. The study results show that the optimized YOLOv5 model can detect multiple types of vehicles with high accuracy. The model achieved a detection accuracy of 85.47% for trucks, 87.12% for pedestrians, 86.54% for buses, 77.20% for cars, 80.48% for motorcycles, and 78.80% for bicycles. Significant improvements in detection performance were achieved compared to the default model. This research concludes that the optimization of the YOLOv5 model is effective in improving vehicle detection accuracy at intersections. Implementing this optimized model can significantly contribute to traffic management, reduce congestion, and improve road safety. It is expected that the implementation of this technology can be more widely applied for more efficient traffic management in various major cities.

Corresponding Author:

I Wayan Adi Artha Wiguna, Magister Program, Departement of Magister Information Systems, Institut Teknologi dan Bisnis STIKOM Bali, Indonesia
Email: 222012018@stikom-bali.ac.id

1. INTRODUCTION

Traffic congestion is one of the most challenging urban problems in the modern era. In many major cities around the world, streets are often clogged with long lines of vehicles, causing not only delays and discomfort for drivers but also negative impacts on the economy and the environment. Several cities in Indonesia have been using CCTV (Closed Circuit Television) to monitor traffic flow. In this monitoring process, CCTV extracts information from images such as vehicle speed, congestion levels, vehicle types, and license plates, as well as incidents of violations or accidents [1], [2].

In recent years, research on artificial intelligence, such as object detection, has become a major focus because it can help researchers identify objects in images more efficiently [3]. Object detection is a crucial field within computer vision, which is the science of how computers can see and analyze objects in images. [4], [5]. Artificial intelligence technology, especially in the fields of pattern recognition and digital image processing, offers innovative solutions to this problem [6]. YOLOv5, as one of the object detection models, has demonstrated its ability to quickly and accurately identify vehicles and other objects under various traffic conditions [7].

Research conducted by A. Heri, S. Budi, M. A. Baiquni, B. Mulyanti, and M. Fadli developed a system that detects vehicle speed and number plates based on recorded video using YOLOv5, DeepSORT for tracking, and HyperLPR for number plate recognition. The results show that this system is effective in video recording-based traffic monitoring. The focus of this research is on speed and number plate detection, which is similar to your thesis goal of optimizing vehicle detection performance at intersections, although aspects of tracking and number plate recognition are also their focus [8]. R. Illmawati and Hustinawati implemented YOLOv5 to detect vehicle numbers in the DKI Jakarta area, with a focus on detection accuracy in dense urban environments. This research shows that YOLOv5 can detect vehicle numbers with a satisfactory level of accuracy. The similarity with your thesis is the use of YOLOv5 for vehicle detection, but this research focuses more on vehicle number recognition, while your thesis focuses on optimizing vehicle detection at intersections [5]. Research by M. R. Rais, F. Utaminingrum, and H. Fitriyah developed a vehicle number plate recognition system for residential access using YOLOv5 and Pytesseract, running on a Jetson Nano device. The system was tested in a residential scenario and showed good performance in recognizing vehicle number plates. The difference is that the focus of this research is on residential access and number plate recognition, while your thesis focuses more on general vehicle detection at traffic intersections [9]. M. Dio Riza Pratama, B. Priyatna, S. S. Hilabi, and A. L. Hananto focused their research on accident object detection in four-wheeled vehicles using YOLOv5. The model was trained to detect various types of accidents and tested on a dataset of vehicle accidents. The research results show that YOLOv5 is effective in detecting accident objects with a high level of accuracy. Even though the application context is different, namely accident detection, the methods used and the focus on optimizing object detection remain relevant to the objectives of the thesis [10]. Research by R. Dwiyanto, D. W. Widodo, and P. Kasih discusses the implementation of YOLOv5 for vehicle classification using CCTV recordings in Tulungagung Regency. The research results show that YOLOv5 can classify various types of vehicles with good accuracy, and can be integrated with CCTV-based traffic monitoring systems. A similarity to your thesis is the use of YOLOv5 for vehicle detection in the context of traffic surveillance, although this research places more emphasis on vehicle type classification rather than optimization of detection at intersections [11].

Optimizers are algorithms used to adjust model weights during training to minimize prediction errors. The use of appropriate optimizers can speed up model convergence and improve detection accuracy. Various studies have demonstrated the effectiveness of optimizers in improving YOLOv5 performance [12]. Several optimizers such as Adam, SGD, and RMSprop have been used to improve detection accuracy and speed in various object detection applications. For example, Liu *et al.* applied Adam Optimizer in the development of the YOLOv5s-MobileNetV2 model for traffic sign detection, resulting in significant improvements in accuracy and computational efficiency [13]. Similarly, Isa *et al.* show the improved performance of YOLOv5 in underwater detection with the use of SGD Optimizer [14].

In a previous study conducted by R. Dwiyanto *et al.*, entitled "Implementation of the You Only Look Once (YOLOv5) Method for Vehicle Classification on CCTV in Tulungagung Regency," the finding was that the YOLOv5 model succeeded in achieving an accuracy of 79.8% in classifying vehicles. Although these results have been modeled optimally, there is considerable room for improvement, especially in terms of accuracy [11]. Considering the importance of achieving a higher level of accuracy to increase efficiency in traffic control, this research proposes an innovation through the integration of an optimizer into the YOLOv5 model structure. With this strategic adaptation, this research aims to not only improve vehicle detection accuracy beyond the predefined 79.8% threshold but also to make a significant contribution to advancing object detection technology. This is expected to bring significant progress in efforts to create a more efficient and safer traffic control system.

This research contributes to the field of traffic management in two significant ways: Firstly, by providing an in-depth analysis of how different optimizers affect the performance of the YOLOv5 model, highlighting the importance of optimizer tuning for accuracy improvement. Secondly, by offering a real-world application example that shows substantial enhancements in traffic system efficiency through optimized vehicle detection at intersections.

2. METHODS

2.1. Literature Review on the YOLO Model

This research is inspired by various works that have explored the use of the YOLOv5 model in the context of traffic monitoring and its application in urban environments.

A. Related Research

The table provides a comparative analysis of various studies and their methodologies related to traffic congestion, object detection, and the application of AI in traffic management. Each reference is evaluated based on its main focus, methodology, key findings, similarities to the topic of vehicle detection at intersections using YOLOv5, and differences in terms of methodology or application context.

For instance, references [1] and [15] both explore factors contributing to traffic congestion in urban areas, but their focus is on empirical and accident data analysis, respectively, rather than object detection. Reference [16] highlights a dataset useful for training object detection models like YOLOv5, while [17] examines the environmental impacts of traffic, showing a broader context of urban traffic without delving into object detection specifics. Studies [18] and [21] discuss traffic monitoring and control systems using different methodologies such as system development and fuzzy logic, demonstrating the varied approaches to traffic management. Reference [19] combines YOLOv5 with BiFPN for ship detection, showcasing the versatility of YOLOv5 in different application contexts, while [20] focuses on using YOLO for aerial object detection with UAVs.

References [22] and [23] discuss AI and machine learning applications in improving traffic light efficiency and handling disruptions, respectively, indicating how AI can optimize traffic flow. In contrast, reference [11] examines the use of YOLOv5 for vehicle classification on CCTV, closely aligning with the vehicle detection focus but differing in the application scope. References [24] and [25] demonstrate the use of YOLOv5 and AI in different contexts, such as face mask detection and managing traffic during the Hajj pilgrimage, highlighting the adaptability of these technologies to various detection and traffic scenarios. This comparison can be seen in Table 1, which succinctly summarizes the main points of each reference, providing a clear overview of their contributions and relevance to the study of vehicle detection using YOLOv5 at intersections.

Table 1. Comparison of Previous Research with Current Research

References	Main Focus	Methodology	Finding	Similarities	Differences
[1]	Urban Scale Factors in Traffic Congestion	Empirical Analysis	The main factors causing traffic jams in urban areas	Urban traffic context	Focus on factor analysis, not object detection
[15]	Causes of Congestion Due to Traffic Accidents	Accident Data Analysis	The main determinants of traffic jams are accidents	Urban traffic context	Focus on accident analysis, not object detection
[16]	Street View Dataset for Object Detection	Data collection	A rich dataset for training object detection models	Data source for YOLOv5 model training	Focus on data collection, not detection methodology
[17]	The Effect of Traffic on Urban Microclimate	IoT Research	Impact of traffic on urban heat islands	Urban traffic context	Focus on environmental impact, not object detection
[18]	Automatic Traffic Monitoring System	System Development	Effectiveness in heavy traffic monitoring	Technology application in traffic	Focus on monitoring not specific object detection
[19]	Ship Detection with YOLOv5 and BiFPN	Combining YOLOv5 and BiFPN	Improved object detection at sea	Use of YOLOv5 for object detection	Application context (sea vs. traffic)
[20]	Object Detection from UAV with YOLO	Using YOLO on UAV	Effectiveness of YOLO in Aerial Detection	Use of the YOLO model for object detection	Different platforms (UAV) and detection context
[21]	Traffic Light Control with Fuzzy Logic	Fuzzy Logic Implementation	Optimization of traffic flow at intersections	Traffic control context	Fuzzy logic-based methodology, not YOLOv5
[22]	Traffic Flow Prediction for Smart Traffic Lights	Machine Learning Algorithms	Improved traffic light efficiency	AI application in traffic	Focus on traffic prediction, not object detection
[23]	Using Deep Reinforcement Learning for Traffic Light Control	Deep Reinforcement Learning	Handling disruptions in traffic control	AI application in traffic	Focus on reinforcement learning, not object detection
[11]	Vehicle Classification with YOLOv5 on CCTV	YOLOv5 Implementation on CCTV	Effectiveness in vehicle classification	Use of YOLOv5 in traffic and CCTV	Focus on vehicle classification, not traffic control
[24]	Mask Detection with YOLOv5	YOLOv5 Implementation	High accuracy in face mask detection	Use of YOLOv5 for object detection	Context of face mask detection, not traffic
[25]	Using AI to Reduce Traffic Congestion During Hajj	AI Application in Traffic	Reduction of congestion during Hajj	AI application in traffic management	Specific Hajj context and general AI solutions

B. YOLOv5

This research develops a YOLOv3-based system to collect traffic data such as flow intensity, driving direction, and vehicle speed in real time. This system shows a vehicle counting accuracy of more than 92% and a maximum speed error of 1.5 km/h. Contribution: Provides perspective on how object detection technology can be used to comprehensively collect traffic data, which supports my research goal of developing more efficient traffic light management systems [26].

The provided image illustrates the architecture of a YOLOv5 model, divided into three primary components: Backbone, Head (PA-Net), and Detection, as can be seen in Fig. 1.

In the Backbone section, the model starts with the Focus module, which includes slicing, convolution operations, and the Leaky ReLU activation function. This is followed by several BottleNeckCSP modules, designed for efficient feature extraction by combining cross-stage partial connections. The final part of the Backbone is the Spatial Pyramid Pooling (SPP) module, which helps aggregate contextual information by pooling features at multiple scales [27], [28].

The Head section, known as the Path Aggregation Network (PA-Net), begins by concatenating features from different stages. It then performs upsampling to increase the resolution of these features, followed by 1x1 convolution operations to reduce the number of channels. This process is repeated, and interspersed with additional BottleNeckCSP modules for further feature refinement. The PA-Net also involves downsampling using 3×3 convolutions with stride 2, followed by concatenation to combine features from different scales [29], [30]. Finally, the Detection layer employs 1×1 convolution operations to produce the final output predictions, which include object classes, bounding boxes, and confidence scores [31].

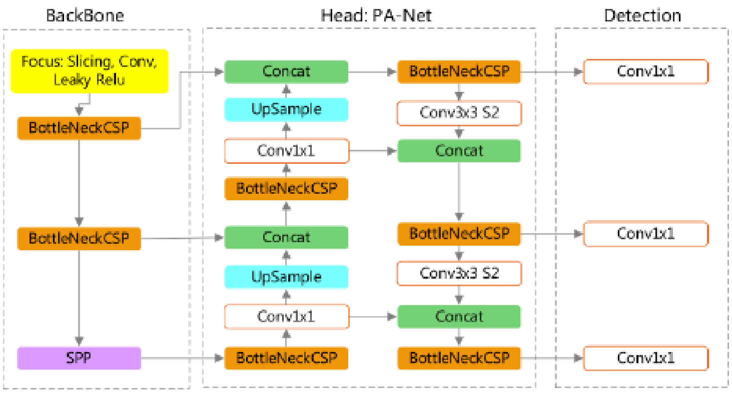

Fig. 1. YOLOv5 Architecture

YOLOv5 uses the Sigmoid-weighted Linear Units (SiLU) activation function in the following equation.

$$\alpha_k(z_k) = z_k \sigma (z_k) \tag{1}$$

$$z_k = \sum_i \omega_{ik} S_i + b_k \tag{2}$$

$$\sigma(x) = \frac{1}{1+e^{-x}} \tag{3}$$

The Sigmoid-weighted Linear Units (SiLU) activation function used in YOLOv5 can be described by the equation $\alpha_k(z_k) = z_k \sigma (z_k)$ where $\alpha_k(z_k)$ is the output of the SiLU function for the k-th neuron This can be achieved by method (1). The input z_k is calculated as the summation of the product of weights ω_{ik} and inputs S_i, plus a bias $b_k(z_k = \sum_i \omega_{ik} S_i + b_k)$ this can be achieved by method (2). The sigmoid function, expressed as $\sigma(x) = \frac{1}{1+e^{-x}}$, transforms the input x into a value within the range of 0 to 1 this can be achieved by method (3). The combination of weights ω_{ik} inputs S_i, plus a bias b_k in the calculation of z_k, along with the application of the sigmoid function, enhances convergence and model performance, resulting in the final output $\alpha_k(z_k)$ that reflects the collective influence of all these factors [31], [32].

The main focus of the research "OPTIMIZATION OF VEHICLE DETECTION AT INTERSECTIONS USING THE YOLOV5 MODEL" is to optimize the YOLOv5 object detection model to overcome the problem of traffic congestion at road intersections. This research aims to increase the accuracy and speed of vehicle detection using the YOLOv5 model so that the information obtained can be used to optimize traffic flow and reduce congestion.

The key difference between this research and previous studies is the emphasis on improving vehicle detection accuracy through proper optimizer tuning in the YOLOv5 model. By providing appropriate optimizer tuning, this research aims to significantly improve detection accuracy compared to previous approaches. This will produce a smarter and more efficient solution for detecting vehicles so that it can be used to optimize traffic management at road intersections and reduce congestion.

2.2. YOLOv5 Model Development

At this stage, a vehicle classification model is developed using YOLOv5, involving a process of design, implementation, and adjustment to suit the research objectives.

The diagram illustrates a structured workflow for processing a dataset, training a model, and testing the results. The process begins with loading the dataset, followed by the dataset preparation phase, which includes data augmentation, image resizing, data annotation, and organization. The prepared data is then used in the training process, which involves defining a learning rate, tuning hyperparameters, selecting an optimizer, and training the model. Finally, the testing process evaluates the trained model's performance by analyzing the results and generating evaluation metrics. This comprehensive workflow ensures a systematic approach to machine learning model development and evaluation, as can be seen in Fig. 2.

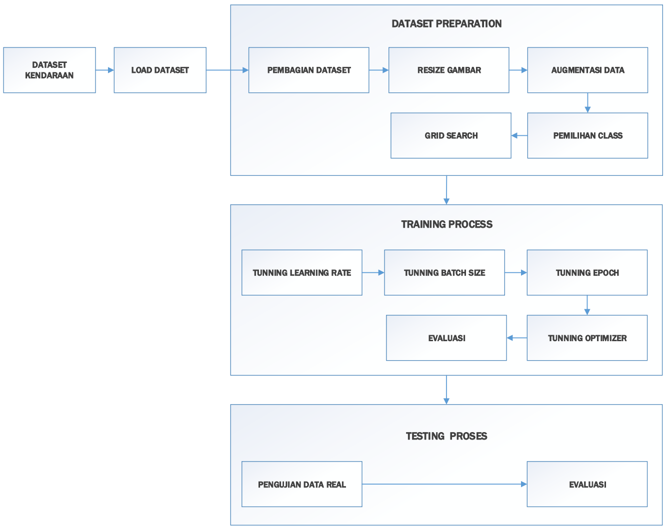

Fig. 2. Model Overview

2.3. Data source

Use of datasets from Roboflow Universe for model training and testing purposes, utilizing datasets collected by FSMVU and available in Roboflow Universe. This dataset, which consists of 8,693 vehicle images that have been annotated in detail, is in .jpg format with a resolution of 512x512 pixels. The high annotation quality and appropriate resolution of this dataset ensure rich data for YOLOv5 model development and validation. Examples of images in the dataset that will be used in this research are in Fig. 3 [16].

Fig. 3. Example of Dataset Image

2.4. Preprocessing

The preprocessing process includes adjusting the image size to 640×640 pixels and dividing the dataset into three main parts, namely training, validation, and testing sets. The dataset-sharing ratio between the main parts is 87% for the training set, 9% for validation, and 4% for testing [33], [34].

2.5. Training Phase

In the training stage, the YOLOv5 model is applied with modifications to the existing layers. This process includes adding, removing, or adjusting layers to improve the quality of the model. Fig. 3 shows that an input image measuring 448×448 pixels with 3 color channels (RGB) is processed through a convolutional network

layer, producing an output with dimensions of 7×7×30. Zero-padding and stride padding are used during the convolution process. Optimizer Adam was also added to increase the efficiency and effectiveness of the model, by adjusting the learning rate adaptively [35].
YOLO uses a linear activation function for the final layer and Leaky ReLU activation for all other layers. The Leaky ReLU activation function can be achieved by method (4).

$$\sigma(x) = \frac{1}{1+e^{-x}} \tag{4}$$

This activation function is used to provide non-linearity to the network and prevent the "dying ReLU" problem, where neurons are inactive and unable to learn [36].

2.6. Testing Phase

In the testing stage, the model is used to predict test data that the model has never seen before. The test data serves as a final test to test the model's performance in classifying the number of vehicles at road intersections that are not included in the training and validation datasets. After the model makes predictions on the test data, an evaluation is carried out [37], [38].

Evaluation of model performance is carried out by measuring precision, which is the ratio of correctly predicted objects to the overall model prediction results. In addition, recall is measured, namely the ratio of correctly predicted objects to the overall actual results. Mean Average Precision (mAP) is calculated using precision, recall, and F1-Score values. The formulas for calculating Precision (P), Recall (R), Average Precision (AP), and Mean Average Precision (mAP) can be found below [39], [40].

$$AP = \sum (recall_{n+1} - recall_n)\ x\ precision_{interp}\ x\ (recall_{n+1}) \tag{5}$$

$$f1 - score = \frac{2\ x\ precision\ x\ recall}{precision + recall} \tag{6}$$

$$precision = \frac{true\ positives}{true\ positives + false\ positives} \tag{7}$$

$$recall = \frac{true\ positives}{true\ positives + false\ negatives} \tag{8}$$

The formula for calculating Average Precision (AP) in object detection evaluation, describes how AP is calculated based on changes in the interpolated recall, multiplied by interpolated precision, and then multiplied by the recall at each different threshold point this can be achieved by method (5). The formula for F1-Score combines precision and recall in one metric by calculating the double of the product of precision and recall, divided by the sum of precision and recall this can be achieved by method (6). The formulas for calculating precision and recall, respectively, where precision measures the accuracy of the model's positive predictions and recall measures the model's ability to identify all true positive objects can be achieved by methods (7), (8).
In the context of model evaluation, some important terms are:
True Positive (TP): It is a situation when the model correctly identifies or predicts the positive class. In this case, the model correctly estimates the presence of vehicles at the road intersection.
True Negative (TN): Represents the situation when the model correctly identifies or predicts the negative class. In this context, the model correctly predicts the absence of vehicles at road intersections.

3. RESULTS AND DISCUSSION

3.1. Performance Evaluation of Train Model YOLOv5

3.1.1. Confusion Matrix

It can be seen in Fig. 4 that it illustrates the performance of the YOLOv5 model in detecting various objects. Each cell shows the proportion of predictions compared to the true label. High values along the diagonal indicate good performance for the corresponding classes, such as bicycles (0.76), cars (0.91), and trucks (0.84). However, there are some classification errors, especially in the detection of people who are often misidentified as motorbikes (0.12).

3.1.2. Precision and Recall

It can be seen in Fig. 5 that this curve, we can see that the model has a high level of accuracy during training for all classes tested. Accuracy for each class is as follows: truck 99%, bus 93%, car 95%, motorbike

83%, person 72%, and bicycle 83%. This indicates strong performance in detecting and classifying these objects.

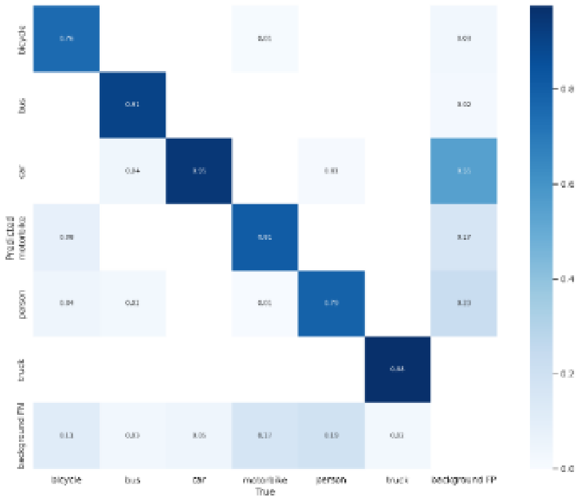

Fig. 4. Confusion Matrix

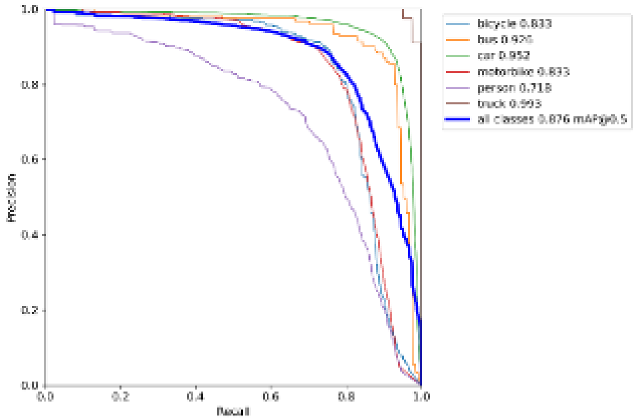

Fig. 5. Precision and Recall

3.1.3. YOLOv5 Train Model Prediction Results

Fig. 6 shows the results of object detection using the YOLOv5 model in several traffic scenarios at intersections. Each detected object is assigned a bounding box and a class label, such as a truck, car, motorcycle, or person, with a confidence score indicated next to the label. Detection performed well for most objects, but there were some detection errors, such as people being misidentified as other objects. This figure illustrates the model's ability to detect various types of vehicles and people in complex traffic environments, although there is still room for improvement in accuracy.

3.2. Performance Evaluation of Train Model YOLOv5

3.2.1. Real Data Testing

Table 2 displays data collected from various intersections on the Ngurah Rai Sanur Bypass over four different days. This data includes the number of vehicles and other objects detected by the CCTV system at these intersections. Observations were carried out at three-time intervals every day, namely morning, afternoon, and evening.

3.2.2. Model Reading from Real Data

Table 3 of the model reading below shows the detection results carried out by the YOLOv5 model after being trained using datasets collected from intersections on the Ngurah Rai Sanur Bypass. This data includes the number of vehicles and other objects detected by the model at those intersections. Observations were carried out at three-time intervals every day, namely morning, afternoon, and evening.

Fig. 6. Train Prediction Results

Table 2. Factual Data on the Number of Vehicles Counted

No.	Truck	Person	Bus	Car	Motorbike	Bicycle
1	10	15	5	50	100	20
2	8	10	3	60	90	15
3	12	20	7	55	80	25
4	11	12	4	45	110	18
5	9	14	6	52	95	22
6	10	17	5	48	85	20
7	13	11	3	50	105	19
8	7	16	6	58	78	21
9	11	19	7	62	90	24
10	9	13	4	49	92	17
11	8	15	5	51	88	23
12	10	14	6	53	97	20

Table 3. Model Reading Results

No.	Truck	Person	Bus	Car	Motorbike	Bicycle
1	8	13	4	46	85	18
2	7	8	2	54	76	12
3	10	16	6	49	68	22
4	9	10	3	40	94	16
5	7	12	5	45	81	19
6	8	15	4	42	72	17
7	11	9	2	44	90	15
8	6	14	5	51	66	18
9	9	17	6	54	77	21
10	7	11	3	44	78	14
11	7	13	4	45	75	19
12	8	12	5	47	82	18

3.2.3. Confusion Matrix

Fig. 7 shows the performance of the classification model in six classes: Trucks, People, Buses, Cars, Motorbikes, and Bicycles. In the first row, out of 117 actual trucks, the model managed to correctly identify 100 trucks, while 4 trucks were classified as people, 3 as buses, 6 as cars, 3 as motorbikes, and 1 as a bicycle. In the second row, of the 163 actual people, the model correctly classifies 142, but incorrectly classifies 6 as trucks, 5 as buses, 3 as cars, 5 as motorbikes, and 2 as bicycles. For buses, of the 56 actual buses, 45 were classified correctly, while 4 were classified as trucks, 3 as people, 1 as a car, and 3 as a motorbike. In the case of cars, of the 482 actual cars, the model identified 464 correctly, with some misclassifications into trucks (5), people (5), buses (2), motorbikes (4), and bicycles (2). For motorcycles, of the actual 872, the model correctly identified 845, but incorrectly classified 8 as trucks, 6 as people, 3 as buses, 7 as cars, and 5 as bicycles. Ultimately, of the 187 actual bicycles, the model managed to correctly classify 171, with misclassifications of 1 to trucks, 4 to people, 1 to buses, 4 to cars, and 6 to motorbikes.

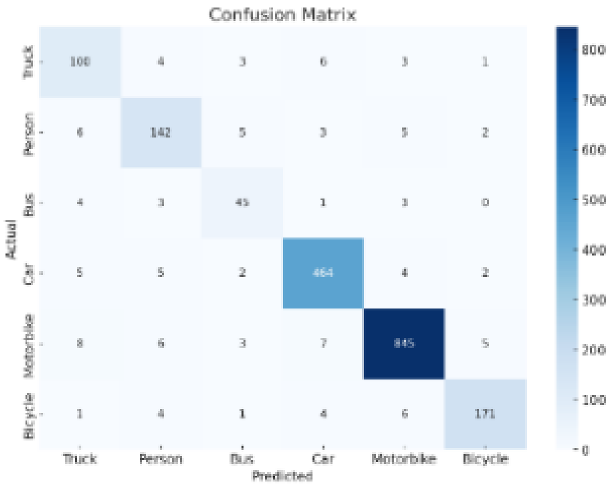

Fig. 7. Real Data Confusion Matrix

3.2.4. Precision and Recall

Table 4 shows the percentage of correct detections for each class. The below gives an idea of how accurate the YOLOv5 model is in classifying six categories of objects: Truck, Person, Bus, Car, Motorbike, and Bicycle. This percentage is calculated based on the number of correct detections compared to the total number of actual objects in each class.

Table 4. Accuracy of Each Class

Class	Percentage of Correct Detections	Precision	Recall	F1 Score
Truck	85.47%	1	0.967	0.983
Person	87.12%	1	0.750	0.857
Bus	86.54%	0.909	1	0.952
Car	77.20%	1	0.967	0.983
Motorbike	80.48%	1	0.917	0.957
Bicycle	78.80%	0.889	1	0.941

3.2.5. YOLOv5 Model Prediction Results

Fig. 8 shows the results of object detection using the YOLOv5 model in several traffic scenarios at intersections. Each detected object is assigned a bounding box and a class label, such as a truck, car, motorcycle, or person, with a confidence score indicated next to the label. Detection performed well for most objects, but there were some detection errors, such as people being misidentified as other objects. This figure illustrates the model's ability to detect various types of vehicles and people in complex traffic environments, although there is still room for improvement in accuracy

Fig. 8. Prediction Results

4. CONCLUSION

This research concludes that the implementation of YOLOv5 significantly increases the efficiency and intelligence of the traffic management system. YOLOv5 can detect various types of vehicles with high accuracy, enabling more dynamic and responsive traffic light settings. The model achieves impressive detection accuracy: truck 99%, bus 93%, car 95%, motorbike 83%, person 72%, and bicycle 83%. The use of Adam Optimizer proved effective in improving detection accuracy, and the quality dataset from Roboflow Universe supports the development of reliable models. Despite some misclassification errors, these results demonstrate the great potential of YOLOv5 in traffic management systems. Implementation of this model is expected to increase traffic efficiency, reduce congestion, and improve road safety. With further optimization, YOLOv5 can make a greater contribution to traffic management in the future.

REFERENCES

[1] M. M. Rahman, P. Najaf, M. G. Fields, and J. C. Thill, "Traffic congestion and its urban scale factors: Empirical evidence from American urban areas," *Int. J. Sustain. Transp.*, vol. 16, no. 5, pp. 406–421, 2022, https://doi.org/10.1080/15568318.2021.1885085.

[2] D. Iskandar Mulyana and M. A. Rofik, "Implementasi Deteksi Real Time Klasifikasi Jenis Kendaraan Di Indonesia Menggunakan Metode YOLOV5," *J. Pendidik. Tambusai*, vol. 6, no. 3, pp. 13971–13982, 2022, https://doi.org/10.31004/jptam.v6i3.4825.

[3] R. Yang, Y. Yu, "Artificial convolutional neural network in object detection and semantic segmentation for medical imaging analysis," *Frontiers in oncology*, vol. 11, p. 638182, 2021, https://www.frontiersin.org/journals/oncology/articles/10.3389/fonc.2021.638182/full.

[4] A. D. Nugroho and W. M. Baihaqi, "Improved YOLOv5 with Backbone Replacement to MobileNet V3s for School Attribute Detection," *SinkrOn*, vol. 8, no. 3, pp. 1944–1954, 2023, https://doi.org/10.33395/sinkron.v8i3.12702.

[5] R. Illmawati and Hustinawati, "YOLO v5 untuk Deteksi Nomor Kendaraan di DKI Jakarta," *J. Ilmu Komput. Agri-Informatika*, vol. 10, pp. 32–43, 2022, [Online]. Available: https://jurnal.ipb.ac.id/index.php/jika.

[6] A. A. Ouallane, A. Bahnasse, A. Bakali, and M. Talea, "Overview of Road Traffic Management Solutions based on IoT and AI," *Procedia Comput. Sci.*, vol. 198, no. 2021, pp. 518–523, 2021, https://doi.org/10.1016/j.procs.2021.12.279.

[7] S. Jin and L. Sun, "Application of Enhanced Feature Fusion Applied to YOLOv5 for Ship Detection," in *2021 33rd Chinese Control and Decision Conference (CCDC)*, pp. 7242–7246, 2021, https://doi.org/10.1109/CCDC52312.2021.9602100.

[8] A. Heri, S. Budi, M. A. Baiquni, B. Mulyanti, and M. Fadli, "Sistem Deteksi Laju dan Plat Nomor Kendaraan Berbasis Video Rekaman Menggunakan YOLOv5 - DeepSORT dan HyperLPR Video Recording Based Vehicle Number Plate and Speed Detection System Using YOLOv5-DeepSORT and HyperLPR," vol. 11, no. 2, 2023.

[9] M. R. Rais, F. Utaminingrum, and H. Fitriyah, "Sistem Pengenalan Plat Nomor Kendaraan untuk Akses Perumahan menggunakan YOLOv5 dan Pytesseract berbasis Jetson Nano," *J. Pengemb. Teknol. Inf. dan Ilmu Komput.*, vol. 7, no. 2, pp. 681–685, 2023, [Online]. Available: https://j-ptiik.ub.ac.id/index.php/j-ptiik/article/view/12282.

[10] M. Dio Riza Pratama, B. Priyatna, S. S. Hilabi, and A. L. Hananto, "Deteksi Objek Kecelakaan Pada Kendaraan Roda Empat Menggunakan Algoritma YOLOv5," *J. Ilm. Sist. Informas*, vol. 12, no. 2, pp. 15–26, 2022.

[11] R. Dwiyanto, D. W. Widodo, and P. Kasih, "Implementasi Metode You Only Look Once (YOLOv5) Untuk Klasifikasi Kendaraan Pada CCTV Kabupaten Tulungagung," *Semin. Nas. Inov. Teknol.*, vol. 1, no. 1, pp. 102–104, 2022.

[12] J. A. A. Silva, J. C. López, C. P. Guzman, N. B. Arias, M. J. Rider, and L. C. P. da Silva, "An IoT-based energy management system for AC microgrids with grid and security constraints," *Appl. Energy*, vol. 337, p. 120904, 2023, https://doi.org/10.1016/j.apenergy.2023.120904.

[13] X. Liu and others, "Traffic sign recognition algorithm based on improved YOLOv5," in *2021 International Conference on Control, Automation and Information Sciences (ICCAIS)*, pp. 980–985, 2021, https://doi.org/10.1109/ICCAIS52680.2021.9624657.

[14] I. S. Isa, M. S. A. Rosli, U. K. Yusof, M. I. F. Maruzuki, and S. N. Sulaiman, "Optimizing the Hyperparameter Tuning of YOLOv5 for Underwater Detection," *IEEE Access*, vol. 10, pp. 52818–52831, 2022, https://doi.org/10.1109/ACCESS.2022.3174583.

[15] Z. Zheng, Z. Wang, L. Zhu, and H. Jiang, "Determinants of the congestion caused by a traffic accident in urban road networks," *Accid. Anal. Prev.*, vol. 136, no. May 2019, p. 105327, 2020, https://doi.org/10.1016/j.aap.2019.105327.

[16] FSMVU, "Street View Dataset," *Roboflow Universe*. Roboflow, Sep. 2023. [Online]. Available: https://universe.roboflow.com/fsmvu/street-view-gdogo.

[17] E. Husni *et al.*, "Microclimate investigation of vehicular traffic on the urban heat island through IoT-Based device," *Heliyon*, vol. 8, no. 11, p. e11739, 2022, https://doi.org/10.1016/j.heliyon.2022.e11739.

[18] D. M. Tan and L. M. Kieu, "TRAMON: An automated traffic monitoring system for high density, mixed and lane-free traffic," *IATSS Res.*, vol. 47, no. 4, pp. 468–481, 2023, https://doi.org/10.1016/j.iatssr.2023.10.001.

[19] C. Yu and Y. Shin, "SAR ship detection based on improved YOLOv5 and BiFPN," *ICT Express*, vol. 10, no. 1, pp. 28-33, 2023, https://doi.org/10.1016/j.icte.2023.03.009.

[20] C. Jiang *et al.*, "Object detection from UAV thermal infrared images and videos using YOLO models," *Int. J. Appl. Earth Obs. Geoinf.*, vol. 112, p. 102912, 2022, https://doi.org/10.1016/j.jag.2022.102912.

[21] S. Jafari, Z. Shahbazi, Y. C. Byun, "Improving the road and traffic control prediction based on fuzzy logic approach in multiple intersections," *Mathematics*, vol. 10, no. 16, p. 2832, 2022, https://doi.org/10.3390/math10162832.
[22] A. Navarro-Espinoza *et al.*, "Traffic Flow Prediction for Smart Traffic Lights Using Machine Learning Algorithms," *Technologies*, vol. 10, no. 1, pp. 1–11, 2022, https://doi.org/10.3390/technologies10010005.
[23] F. Rasheed, K. L. A. Yau, R. M. Noor, and Y. W. Chong, "Deep reinforcement learning for addressing disruptions in traffic light control," *Comput. Mater. Contin.*, vol. 71, no. 2, pp. 2225–2247, 2022, https://doi.org/10.32604/cmc.2022.022952.
[24] H. Dawami, E. Rachmawati, and M. D. Sulistiyo, "Deteksi Penggunaan Masker Wajah Menggunakan YOLOv5," *e-Proceeding Eng.*, vol. 10, no. 2, pp. 1746–1764, 2023.
[25] F. Gazzawe and M. Albahar, "Reducing traffic congestion in Makkah during Hajj through the use of AI technology," *Heliyon*, vol. 10, no. 1, p. e23304, 2024, https://doi.org/10.1016/j.heliyon.2023.e23304.
[26] J. Ieamsaard, S. N. Charoensook, and S. Yammen, "Deep Learning-based Face Mask Detection Using YoloV5," *Proceeding 2021 9th Int. Electr. Eng. Congr. iEECON*, pp. 428–431, 2021, https://doi.org/10.1109/iEECON51072.2021.9440346.
[27] F. Xu, H. Wang, X. Sun, X. Fu, "Refined marine object detector with attention-based spatial pyramid pooling networks and bidirectional feature fusion strategy," *Neural Computing and Applications*, vol. 34, no. 17, pp. 14881-14894, 2022, https://doi.org/10.1007/s00521-022-07264-8.
[28] C.-Y. Wang, A. Bochkovskiy, and H.-Y. M. Liao, "CSPNet: A new backbone that can enhance the learning capability of CNN," *Proc. IEEE/CVF Conf. Comput. Vis. pattern Recognit. Work.*, pp. 390–391, 2020, https://doi.org/10.1109/CVPRW50498.2020.00203.
[29] Z. He *et al.*, "Structure-aware residual pyramid network for monocular depth estimation," *IEEE Trans. Intell. Transp. Syst.*, vol. 22, no. 10, pp. 6584–6594, 2021, https://doi.org/10.48550/arXiv.1907.06023.
[30] J. Liu, X. Zhang, Z. Li, T. Mao, "Multi-scale residual pyramid attention network for monocular depth estimation," *In 2020 25th International Conference on Pattern Recognition (ICPR)*, pp. 5137-5144, 2021, https://doi.org/10.1109/ICPR48806.2021.9412670.
[31] G. Jocher, A. Chaurasia, J. Qiu, and L. Stoken, "YOLOv5," *GitHub Repos.*, 2020, [Online]. Available: https://github.com/ultralytics/yolov5.
[32] S. Elfwing, E. Uchibe, and K. Doya, "Sigmoid-weighted linear units for neural network function approximation in reinforcement learning," *Neural Networks*, vol. 107, pp. 3–11, 2021, https://doi.org/10.1016/j.neunet.2017.12.012.
[33] A. Bochkovskiy, C.-Y. Wang, and H.-Y. M. Liao, "YOLOv4: Optimal speed and accuracy of object detection," *arXiv Prepr. arXiv2004.10934*, 2020, https://doi.org/10.48550/arXiv.2004.10934.
[34] A. M. Roy, R. Bose, J. Bhaduri, "A fast accurate fine-grain object detection model based on YOLOv4 deep neural network," *Neural Computing and Applications,* vol. 34, no. 5, pp. 3895-3921, 2022, https://doi.org/10.1007/s00521-021-06651-x.
[35] J. Redmon, S. Divvala, R. Girshick, and A. Farhadi, "You only look once: Unified, real-time object detection," *Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit.*, vol. 2016-Decem, pp. 779–788, 2019, https://doi.org/10.1109/CVPR.2016.91.
[36] M. H. Ashar and D. Suarna, "KLIK: Kajian Ilmiah Informatika dan Komputer Implementasi Algoritma YOLOv5 dalam Mendeteksi Penggunaan Masker Pada Kantor Biro Umum Gubernur Sulawesi Barat," *Media Online*, vol. 3, no. 3, pp. 298–302, 2022, [Online]. Available: https://djournals.com/klik.
[37] C. Ren, Y. Zhang, F. Xu, F. Yuan, and F.-Y. Wang, "Benchmarking deep learning models for traffic sign detection and classification under challenging weather conditions," *IEEE Trans. Intell. Transp. Syst.*, vol. 21, no. 11, pp. 4627–4636, 2020.
[38] C. Wu, S. Zhang, J. Yu, M. Wang, and Q. Li, "A comprehensive review on neural network-based approaches for image captioning," *IEEE Trans. Neural Networks Learn. Syst.*, vol. 31, no. 12, pp. 4999–5015, 2020, https://doi.org/10.1109/TNNLS.2019.2955165.
[39] X. Weng and K. Kitani, "Beyond Online Tracking: A Survey of Online Object Tracking," *IEEE Trans. Intell. Transp. Syst.*, vol. 21, no. 7, pp. 2896–2915, 2020.
[40] A. Dosovitskiy *et al.*, "An image is worth 16x16 words: Transformers for image recognition at scale," *Int. Conf. Learn. Represent.*, 2021, https://cir.nii.ac.jp/crid/1370580229800306183.

BIOGRAPHY OF AUTHORS

I Wayan Adi Artha Wiguna, am an undergraduate student of the Department of Magister Information Systems, Institut Teknologi dan Bisnis STIKOM Bali., 222012018@stikom-bali.ac.id and 0000-0001-6122-7901.

Vol. 10, No. 4, December 2024, pp. 885-896

Roy Rudolf Huizen graduated with a Doctor of Computer Science (2018) from Universitas Gadjah Mada (UGM) Yogyakarta, Indonesia. Lecturer and researcher at the Department of Magister Information System at the Institut Teknologi dan Bisnis STIKOM Bali, with research interests in the fields of Object Identification, Signal Processing, Cyber Security Forensics and Artificial Intelligence., roy@stikom-bali.ac.id and 0000-0002-3671-6030.

Gede Angga Pradipta, holds a Doctor of Computer Science from the Department of Computer Science and Electronics, Faculty of Natural Sciences, Universitas Gadjah Mada (UGM), Yogyakarta, Indonesia., in 2021. He also received a bachelor's degree in computer informatics from Universitas Atma Jaya (UAJY), Yogyakarta, Indonesia, in 2012 and a master's degree in information technology from Universitas Gadjah Mada (UGM), Yogyakarta, Indonesia, in 2014. His research interests include Machine Learning, Pattern Recognition, And Image Processing. He is currently lecturing with the Department of Magister Information Systems, Institut Teknologi dan Bisnis STIKOM Bali, Indonesia., angga_pradipta@stikom-bali.ac.id and 0000-0002-6087-619X.