2023 2nd International Conference on Computer System, Information Technology, and Electrical Engineering (COSITE) | 979-8-3503-4306-9/23/$31.00 ©2023 IEEE | DOI: 10.1109/COSITE60233.2023.10249521

Real-Time Object Detection Performance of YOLOv8 Models for Self-Driving Cars in a Mixed Traffic Environment

Afdhal Afdhal[1,2,5*], Khairun Saddami[2,5], Sugiarto Sugiarto[3], Zahrul Fuadi[4], Nasaruddin Nasaruddin[2,5]
[1]Doctoral Program, School of Engineering, Universitas Syiah Kuala, Indonesia
[2]Department of Electrical and Computer Engineering, Universitas Syiah Kuala, Indonesia
[3]Department of Civil Engineering, Universitas Syiah Kuala, Indonesia
[4]Department of Mechanical dan Industrial Engineering, Universitas Syiah Kuala, Indonesia
[5]Telematics Research Center (TRC), Indonesia
*Corresponding Author: afdhal@usk.ac.id

***Abstract*— Self-driving cars have gained significant attention in recent years. Real-time object detection is a critical component of their perception system. One of the main challenges in developing safe and efficient self-driving cars lies in accurately and in real-time detecting objects in diverse and complex traffic environments. This paper presents the performance evaluation of real-time object detection of YOLOv8 models, a state-of-the-art deep learning framework for self-driving cars in mixed traffic environments. The objective is to assess the extent of YOLOv8's object detection capabilities can be used for self-driving car within complex real-world traffic scenarios. The experimental results show that the accuracy of YOLOv8 during normal daylight scenarios ranges from 0.60 ~ 0.80. In contrast, the accuracy values obtained in night scenarios fall between 0.15 ~ 0.25. Similarly, the F-Measure of YOLOv8 models under daylight conditions ranges from 0.75 ~ 0.87. Conversely, the F-Measure in night conditions falls between 0.27 ~ 0.46. Based on our thorough evaluation results, YOLOv8 has demonstrated its robustness in detecting objects. However, the algorithm requires improvement to effectively handle the challenges of self-driving cars' object detection systems in mixed traffic, such as diverse object classifications, small-scale objects, fast-moving objects, blur, glare, and low-light illumination, particularly in nighttime scenarios.**

Keywords—Real-Time, Object Detection, YOLOv8, Self-Driving Car, Mixed-Traffic Environments

I. INTRODUCTION

In recent years, the advancement of self-driving car technology has led to the development of robust perception systems capable of accurately detecting and understanding objects in complex real-world environments [1]. Object detection is a fundamental task within the perception pipeline, as it enables self-driving cars to identify and classify various objects in an image or video, such as vehicles, pedestrians, traffic signs, and obstacles. Accurate object detection is essential for the safe and efficient operation of self-driving cars, as it allows them to make informed decisions and take appropriate actions based on their surroundings [2].

However, object detection in mixed traffic environments poses significant challenges due to the presence of diverse objects, varying scales, occlusions, and unpredictable interactions between different entities. Mixed traffic is a transportation model that combines diverse types of vehicles with traffic infrastructure that includes two or more road sections and opposing traffic directions without divided by the road median strip [3]. This environment features a variety of objects, including trucks, vehicles, and motorcycles, as well as bicycles, pedestrians, animals, and other objects that require additional identification. As self-driving cars interact with mixed traffic environments, it is essential to employ object detection algorithms that can handle the challenges associated with these scenarios.

Over the years, various deep learning-based object detection algorithms have been proposed [2], each with its strengths and limitations. One of the pioneering approaches is the You Only Look Once (YOLO) algorithm, which revolutionized real-time object detection [4]. YOLO processes the entire image in a single pass through a convolutional neural network (CNN) and directly predicts the bounding boxes and class probabilities of objects. The original YOLO algorithm achieved impressive speed but suffered from lower accuracy, particularly in detecting small objects and accurately localizing overlapping objects.

To address the limitations of the original YOLO, subsequent versions were introduced, each aiming to improve detection accuracy and efficiency. YOLOv2 introduced anchor boxes and multi-scale training, enhancing localization accuracy and handling objects of various sizes [5]. YOLOv3 further improved object detection by employing a feature pyramid network (FPN) and utilizing multiple detection scales, leading to better detection of objects at different resolutions [6]. These advancements significantly boosted the performance of YOLO in various scenarios. YOLOv4 introduced advanced architecture modifications [7], such as the CSPDarknet-53 backbone [8] and PANet [9], to achieve state-of-the-art accuracy in real-time object detection. YOLOv5 focused on streamlining and optimizing the YOLO architecture for even faster inference times without compromising on accuracy [10]. YOLOv6 [11] and YOLOv7 [12] brought further enhancements, exploring novel approaches to feature extraction and improving detection performance. In 2023, Ultralytics introduced YOLOv8 [13]. YOLOv8 represents the evolution of the YOLO series, building upon the strengths of previous versions and introducing new advancements. YOLOv8 is expected to build upon the strengths of previous versions and explore new approaches to enhance real-time object detection capabilities.

In this paper, we focus on evaluating the performance of the YOLOv8 object detection algorithm for self-driving cars in mixed traffic environments. YOLOv8 represents an evolution of the YOLO architecture, incorporating

Authorized licensed use limited to: VIT University- Chennai Campus. Downloaded on October 14,2024 at 04:08:35 UTC from IEEE Xplore. Restrictions apply.

advancements in feature extraction, training methodology, and performance optimization. YOLOv8 stands out for its real-time processing capability, achieving a balance between accuracy and computational efficiency. Its ability to detect multiple objects simultaneously and its robustness in challenging scenarios make it an appealing choice for self-driving cars operating in mixed traffic environments.

By conducting a comprehensive evaluation of YOLOv8, we aim to assess its accuracy, efficiency, and robustness in detecting objects of interest in realistic driving scenarios. To achieve this, we compare all YOLOv8 models, including YOLOv8n (nano), YOLOv8s (small), YOLOv8m (medium), YOLOv8l (large), and YOLOv8x (extra-large). This comparative evaluation allows us to understand the strengths and weaknesses of YOLOv8 models in relation to other algorithms and provides insights into its applicability in self-driving car environments. Furthermore, we consider the challenges associated with mixed traffic environments, such as varying object scales, complex background clutter, and dynamic scene changes. We investigate how YOLOv8 handles these challenges and assess its ability to provide reliable object detection results in such demanding scenarios. By evaluating the performance of YOLOv8 models in mixed traffic environments, we aim to contribute to the development of more accurate and efficient object detection systems for self-driving cars. The insights gained from this evaluation can guide the improvement of autonomous vehicle perception systems, leading to enhanced safety, reliability, and overall performance.

The rest of this paper will be organized as follows: the literature review and research method is discussed in Section II and III. The result and discussion are presented in Section IV. Finally, the conclusion is provided in Section V.

II. Literature Review

The YOLO series of object detection algorithms has made significant contributions to the advancement of real-time detection capabilities. YOLO algorithms revolutionized the field by introducing a single-shot detection approach, which processes an entire image in a single pass through a convolutional neural network (CNN). This approach eliminates the need for separate region proposal generation and object classification stages, resulting in faster inference times. The YOLO algorithms have been very successful in a variety of applications, including self-driving cars.

The original YOLO algorithm (YOLOv1) was introduced by J. Redmon et al. in 2015 [4]. The architecture of YOLOv1 comprises 24 convolutional layers followed by 2 fully connected layers. YOLOv1 was performed faster than any object detector, such as Fast R-CNN. It was able to achieve real-time performance, but it struggled to detect small objects and localize them precisely. YOLOv2 was released in 2016 [5]. It improved on the YOLOv1 by using a more powerful CNN, better anchor boxes, and a new loss function. YOLOv2 was significantly more accurate than YOLOv1, while still maintaining a fast inference speed. In 2018, YOLOv3 was introduced with the aim of achieving higher accuracy than its predecessors by utilizing larger CNNs [6]. The architecture backbone of YOLOv3 contains 53 convolutional layers called Darknet-53. YOLOv3 incorporates improved data augmentation techniques and employs a new multiscale training approach. Despite its accuracy, YOLOv3 remains capable of real-time execution.

In 2020, Bochkovskiy et al. proposed YOLOv4 [7]. YOLOv4 brought advanced changes to its structure and optimization methods. It used the CSPDarknet-53 [8] backbone to extract features, making it more efficient and accurate. Additionally, YOLOv4 introduced the PANet [9] module to combine information from different scales, resulting in highly accurate real-time object detection. In the same year, Glenn Jocher et al. introduced YOLOv5 [10], which focused on making the YOLO architecture faster without sacrificing accuracy. YOLOv5 is similar to YOLOv4, but it uses a different framework called PyTorch instead of DarkNet. YOLOv5 achieved faster inference times by using a lightweight backbone network and implementing model scaling techniques [14].

C. Li et al. was introduced YOLOv6 in 2020 [11] by making enhancements to the backbone network, neck, detection head, and training strategy. YOLOv6 was designed a customizable network structure using RepVGG-EfficientRep and Rep-PAN as the foundation. It utilized an optimized and efficient Efficient Decoupled Head, which reduces the time required while maintaining accurate object detection [15]. YOLOv7 is a state-of-the-art object detection algorithm that was released in July 2022 [12]. It is based on the YOLOv4 architecture, but it has been improved in a number of ways using Extended Efficient Layer Aggregation Network (E-ELAN). YOLOv7 algorithm surpasses all previous object detection models and YOLO versions in both speed and accuracy.

The newest model in the YOLO series for object detection is YOLOv8 [13]. It was released on January 10th, 2023. In comparison to its predecessor, YOLOv8 demonstrates improved speed, a crucial factor for real-time object detection, without compromising accuracy. YOLOv8 has an anchor-free architecture, which means it doesn't rely on anchor boxes, making it simpler to train with various datasets. It also includes an advanced backbone network and multi-scale prediction to enhance accuracy, especially for detecting small objects. YOLOv8 has shown improved accuracy compared to previous YOLO versions and competes with state-of-the-art object detection models.

The key features of YOLO object detection algorithms as shown in Table I.

TABLE I. The Key Features of YOLO Series

YOLO Series	Backbone	Feature Extraction	Additional Key Features
YOLOv1 (2015)	Darknet	Single scale	Divided input image into a grid for predictions within each grid cell
YOLOv2 (2016)	Darknet-19	Single scale	Introduced anchor boxes and multi-scale training
YOLOv3 (2018)	Darknet-53	FPN	Introduced FPN to capture object features at different scales
YOLOv4 (2020)	CSPDarknet-53	PANet	Introduced advanced architectural modifications, including CSPDarknet53 backbone and the PANet module to integrate information from different scales.
YOLOv5 (2020)	CSPDarknet-53	PB-FPN based on PANet and BiFPN	Optimized and streamlined the YOLO architecture for faster inference times and added new features such as hyperparameter optimization.
YOLOv6 (2022)	RepVGG-EfficientRep	Rep-PAN	Enhanced the backbone network, neck, detection head, and training strategy
YOLOv7 (2022)	E-ELAN	PANet	Improved the YOLOv4 architecture using E-ELAN for increasing speed and accuracy
YOLOv8 (2023)	Anchor-free architecture	Multi scale	Improved the backbone network, faster and high accuracy, anchor-free architecture, and multi-scale prediction

Authorized licensed use limited to: VIT University- Chennai Campus. Downloaded on October 14,2024 at 04:08:35 UTC from IEEE Xplore. Restrictions apply.

III. RESEARCH METHOD

This study evaluated the performance of YOLOv8 object detection on a mixed traffic dataset using a video-based simulation. The research methodology pipeline, as illustrated in Fig. 1.

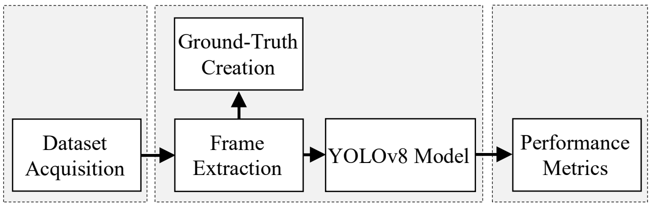


Fig. 1. Experimental Pipeline

The initial stage involved using a dataset of videos captured in mixed-traffic driving environment. The YOLOv8 model was evaluated by extracting frames from the videos and categorizing the images. Ground Truth was generated from the dataset images to evaluate the model. The simulation was conducted during the evaluation stage, employing five YOLOv8 models for object detection. Finally, a comprehensive evaluation was performed to assess the accuracy and robustness of the YOLOv8 model.

A. Mixed-Traffic Dataset

In this study, we created a mixed-traffic driving environment dataset with moving objects on both sides of the road. The dataset was captured in videos from different viewpoints, low-light illumination conditions, and traffic scenarios. It was collected using a moving camera mounted on the observer's vehicle. The dataset includes three distinct scenarios with diverse objects of varying sizes, blur, glare, and other elements of the mixed-traffic scene. The dataset is illustrated in Fig. 2.

(a) normal (daylight) (b) blur (daylight)

(c) normal (night) (d) glare (night)

Fig. 2. Mixed-traffic driving environment dataset

B. Experimental Environment

To evaluate a reliable object detection result in such demanding scenarios, we use the experimental environment shown in Table 2.

TABLE II. EXPERIMENTAL ENVIRONMENT

Parameters	Experimental environment
CPU	Intel i7-10700F @2.90GHz
GPU	GeForce GTX 1660 SUPER, 6144MiB
RAM	64 GB
OS	Windows 10 64-bit
Python	3.9.16
Framework	Ultralytics YOLOv8.0.111, torch-2.0.1+cu118

C. Performance Evaluation Metrics

The evaluation of the results was performed using *Precision*, *Recall*, and *f-measure* metrics. Precision is calculated as:

$$Precision = \frac{TP}{TP + FP} \tag{1}$$

Recall is defined as follows:

$$Recall = \frac{TP}{TP + FN} \tag{2}$$

In equation (1) dan (2), True-Positive (TP) represents the targeted object in the Ground-Truth (GT) images that were correctly detected by the model, while False-Positive (FP) represents the instances of incorrect classification made by the model. Otherwise, False-Negative (FN) refers to the cases where the model failed to detect the targeted object in the GT.

The *f-measure* also known as *f-score,* is a metric that combines *recall* and *precision* into a single score. It is defined as the harmonic mean of recall and precision, given by the formula:

$$f - measure = \frac{2}{\frac{1}{Precision} + \frac{1}{recall}} \tag{3}$$

In addition, we also evaluate the accuracy metric. The accuracy metric is incorporated to provide a measure of the overall correctness of the model's predictions. It is calculated as the ratio of correctly classified prediction to the total number of object predictions. Accuracy is calculated as:

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \tag{4}$$

Accuracy is a widely used performance metric, particularly in balanced datasets where the number of positive and negative prediction is relatively similar. It provides a straightforward measure of the model's correctness in classification tasks.

IV. RESULT AND DISCUSSION

In this section, we aim to demonstrate the robustness and accuracy of YOLOv8 by assessing both qualitative and quantitative measurements in a mixed traffic environment.

A. Qualitative Measurement

For the qualitative measurement of the YOLOv8 model, we examined three distinct scenarios, each presenting its unique challenges. The first scenario involved creating a driving environment with typical daytime traffic conditions. This scenario was free from any disruptive noise or distortion in mixed traffic. In this setting, objects of varying sizes emerged from different directions, exhibiting different speeds. The objects ranged from small to large, and the conditions were unpredictable. The test results for each YOLOv8 model in the first scenario are presented in Fig. 3.

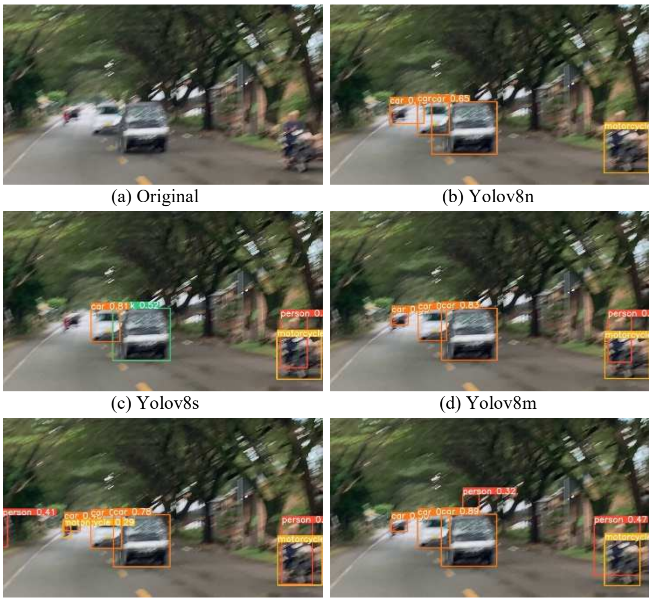
Figure 3 depicts the varying confidence scores and prediction accuracy observed among each YOLOv8 model. Notably, certain models still exhibit errors in their predictions. The results obtained from the first scenario indicate that YOLOv8x demonstrates the highest detection performance and the fewest prediction errors. Interestingly, the outcomes produced by YOLOv8l closely resemble those of YOLOv8x. Conversely, YOLOv8m exhibits some prediction errors. Nevertheless, in terms of confidence score, YOLOv8m still outperforms both YOLOv8s and YOLOv8n.

Authorized licensed use limited to: VIT University- Chennai Campus. Downloaded on October 14,2024 at 04:08:35 UTC from IEEE Xplore. Restrictions apply.

(a) Original (b) Yolov8n (c) Yolov8s (d) Yolov8m (e) Yolov8l (f) Yolov8x

Fig. 3. The comparison results of the YOLOv8 model for object detection in mixed-traffic under normal daylight conditions

In the second scenario, we conducted qualitative measurements by introducing road conditions that were not sufficiently smooth, leading to camera vibrations and resulting in blurry videos and/or images. The remaining conditions were kept consistent with those of the first scenario. The results from the second scenario are depicted in Fig. 4.

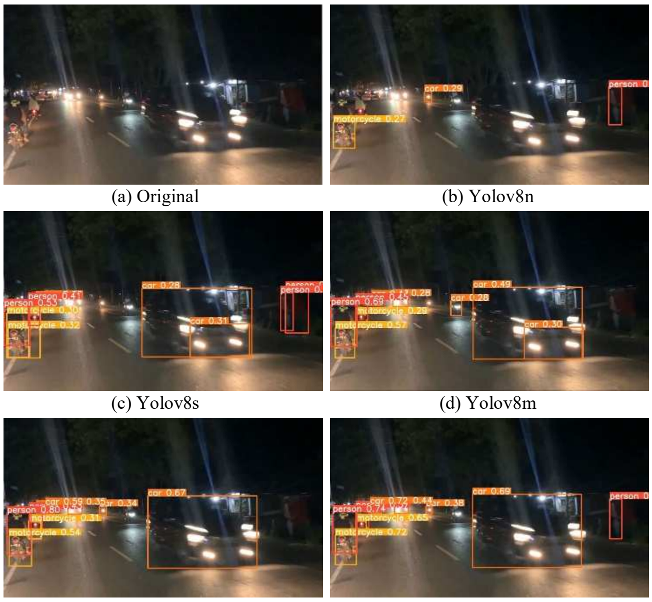


(a) Original (b) Yolov8n (c) Yolov8s (d) Yolov8m (e) Yolov8l (f) Yolov8x

Fig. 4. The comparison results of the YOLOv8 model for object detection in mixed-traffic under daylight with blurriness conditions.

Figure 4 shows that all YOLOv8 models can detect objects, but the predictions are not entirely reliable. Some small objects are still misclassified or even not detected at all. The test results indicate that YOLOv8x still outperforms the other models. However, there are still instances where YOLOv8x misclassifies objects or fails to detect them altogether. The confidence scores also vary significantly across the models. When predicting relatively large objects, YOLOv8x maintains a higher confidence score than the other models. Despite having lower confidence scores, YOLOv8l appears able in detecting almost all objects more precisely. The same pattern was observed in YOLOv8n, which outperformed YOLOv8s in detecting small objects under blurry conditions.

The third scenario aimed to assess the detection capabilities of the YOLOv8 model in a nighttime driving environment with low lighting conditions and blinding light in mixed traffic. In addition, the average speed of surrounding vehicles in this environment ranged from 60 to 80 km/h. The test results for each YOLOv8 model in this scenario are shown in Fig. 5.



(a) Original (b) Yolov8n (c) Yolov8s (d) Yolov8m (e) Yolov8l (f) Yolov8x

Fig. 5. The comparison results of the YOLOv8 model for object detection in mixed-traffic under night condition

Figure 5 shows that YOLOv8n was unable to detect existing objects. Some objects were detected, but they were misclassified. Meanwhile, YOLOv8s and YOLOv8m were able to detect objects, but their predictions were still not accurate. Overall, YOLOv8x had the highest level of confidence in its predictions, but it still made some errors, such as misclassifying objects. The detection of YOLOv8l showed a similar pattern to the second scenario. Although its confidence score was lower than YOLOv8x, it was able to detect almost all objects with slightly more precise predictions than the other models.

B. Quantitative Measurement

Based on the scenarios described previously, we present quantitative measurement of each YOLOv8 model. The Precision, Recall, and F-Measure of the YOLOv8 model for object detection in mixed-traffic under normal daylight conditions as shown in Fig. 6. According to Figure 6, it is confirmed that the YOLOv8x model exhibits the highest performance value among all other models. In terms of precision, the YOLOv8x model achieved a remarkable value of 0.89, indicating a higher proportion of correct positive predictions compared to the other models. YOLOv8l closely followed with a precision value of 0.85, while YOLOv8m, YOLOv8s, and YOLOv8n obtained precision values of 0.82, 0.78, and 0.76, respectively. In regards to recall, YOLOv8x also attained the highest value of 0.85, signifying its effective capture of a substantial proportion of actual positive instances. YOLOv8l achieved a recall value of 0.81, followed by

Authorized licensed use limited to: VIT University- Chennai Campus. Downloaded on October 14,2024 at 04:08:35 UTC from IEEE Xplore. Restrictions apply.

YOLOv8m with 0.78, YOLOv8s with 0.76, and YOLOv8n with 0.74. As for the F-measure, YOLOv8x demonstrated the highest value of 0.87, representing a balanced combination of precision and recall. YOLOv8l achieved an F-measure of 0.83, YOLOv8m scored 0.84, YOLOv8s obtained 0.77, and YOLOv8n had a value of 0.75. These results consistently demonstrate the superior performance of YOLOv8x over the other models in terms of precision, recall, and F-measure, likely attributed to the inclusion of a greater number of convolution layers in YOLOv8x compared to the other models.

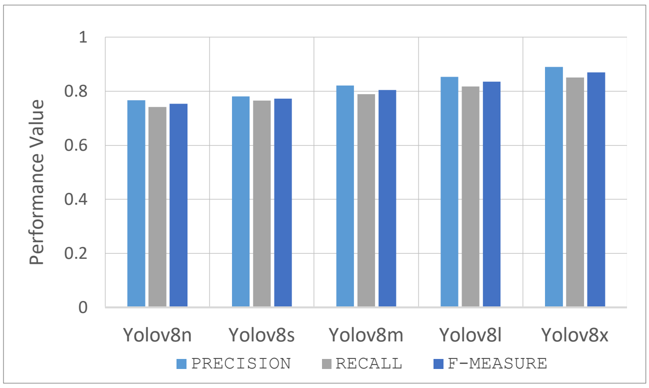


Fig. 6. Precision, Recall, and F-Measure of the YOLOv8 model for object detection in mixed-traffic under normal daylight conditions

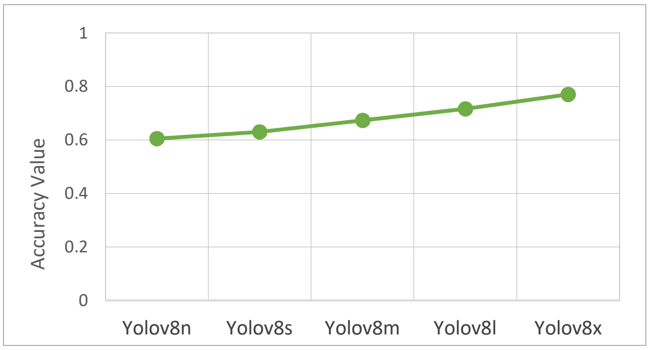


Fig. 7. Object Detection Accuracy under normal daylight conditions

Figure 7 demonstrates that the detection results from YOLOv8x exhibit higher accuracy compared to other YOLOv8 models. It is evident that the YOLOv8x model, which incorporates a larger number of convolutional layers compared to the other YOLOv8 models, consistently outperforms them in terms of accuracy. YOLOv8x achieves the highest accuracy rate of 0.77, followed by YOLOv8l with 0.71, YOLOv8m with 0.67, YOLOv8s with 0.63, and YOLOv8n with 0.61. This indicates that the inclusion of more convolutional layers in the YOLOv8x model contributes to its improved accuracy in object detection. The improved accuracy of YOLOv8 model can lead to more reliable and precise object detection in mixed traffic scenarios. This can enhance the overall safety and efficiency of the traffic environment.

The Precision, Recall, and F-Measure of the YOLOv8 model for object detection in mixed-traffic under night condition as shown in Fig. 8. According to Figure 8, The Precision, Recall, and F-Measure values of YOLOv8x in detecting objects at night within a mixed traffic environment, demonstrating its overall lead the other YOLO models. YOLOv8x achieves a Precision value of 0.59, surpassing YOLOv8l at 0.51, YOLOv8m at 0.45, YOLOv8s at 0.42, and YOLOv8n at 0.34. Similarly, in terms of Recall, YOLOv8x obtains a value of 0.38, outperforming YOLOv8l at 0.30, YOLOv8m at 0.33, YOLOv8s at 0.25, and YOLOv8n at 0.22. Notably, there is an anomaly in the F-Measure values, where YOLOv8n exhibits a higher value compared to YOLOv8s. The F-Measure values indicate that YOLOv8x achieves 0.46, followed by YOLOv8l at 0.38, YOLOv8m at 0.34, YOLOv8n at 0.27, and YOLOv8s at 0.25. These results further reinforce the superior detection performance of YOLOv8x in capturing objects accurately within a mixed traffic environment at nighttime.

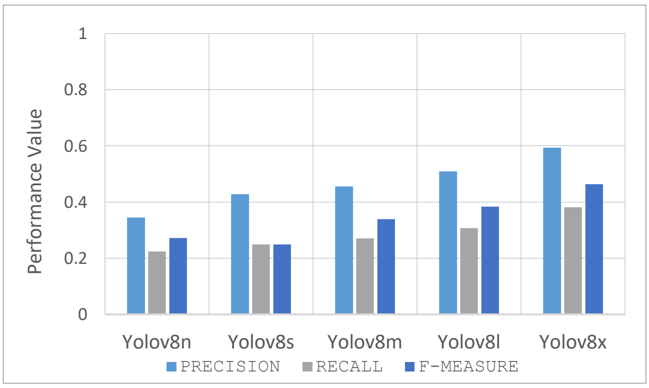


Fig. 8. Precision, Recall, and F-Measure of the YOLOv8 model for object detection in mixed-traffic under night condition

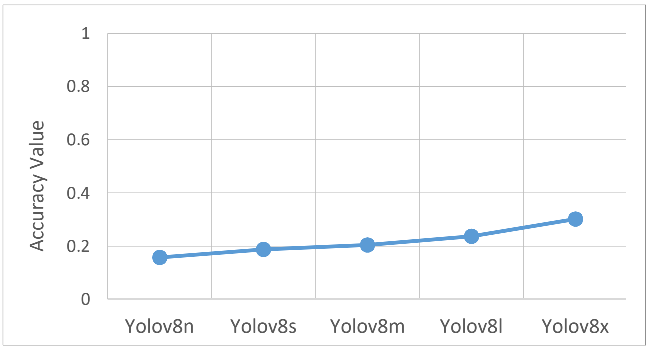


Fig. 9. Object Detection Accuracy under night condition

Figure 9 presents the detection results of YOLOv8x achieves the highest accuracy of 0.30 in object detection in a mixed traffic environment under night conditions, followed by YOLOv8l at 0.24, YOLOv8m at 0.20, YOLOv8s at 0.19, and YOLOv8n at 0.16. These results further highlight the performance of YOLOv8x is better than other YOLOv8 model. However, when assessing the detection ability at night, it becomes apparent that the accuracy level of YOLOv8 is significantly lower compared to daytime.

Figure 10 highlights this disparity, revealing that the accuracy levels of YOLOv8 during the day range between 0.60 ~ 0.80. Contrary, the accuracy values obtained at night fall between 0.15 ~ 0.25. This indicates that the YOLOv8 model still encounters difficulties in handling object detection under low light conditions, glare, and with small objects at night. These challenges are further compounded by the presence of various types of objects. Consequently, YOLOv8's detection ability at night remains inaccurate, both in predicting objects on the road and those located at the side. In this case, the detection results also significantly influenced by object scale and an increase in the speed of object movement. This is evident from the reduced prediction ability when objects move swiftly, either in the same or opposite

Authorized licensed use limited to: VIT University- Chennai Campus. Downloaded on October 14,2024 at 04:08:35 UTC from IEEE Xplore. Restrictions apply.

direction. Mixed-traffic, as a driving environment characterized by high levels of uncertainty, due to its complexity, necessitates further improvements in YOLO's detection capabilities.

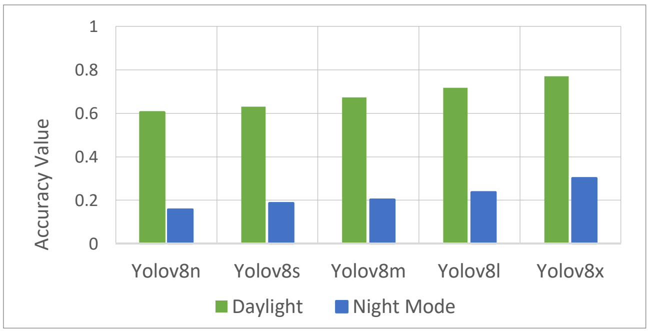


Fig. 10. Daylight vs Night Mode Detection Accuracy

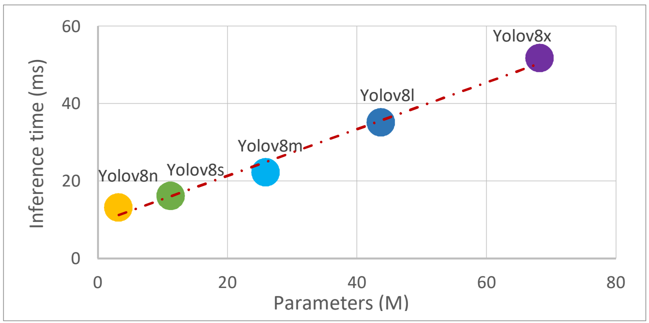


Fig. 11. Inference time (ms) vs Paramaters (M) for each Yolov8 model

In addition, we also provide an assessment of performance metrics for running the YOLOv8 object detection model that includes real-time frame processing, inference speed and the number of parameters used for each model. Figure 11 shows that YOLOv8x has an inference time of 51.7ms, which means it takes approximately 51.7 milliseconds for the model to analyze the image and make predictions using YOLOv8x. It has 6.8 million parameters, which indicates that the model has a moderate level of complexity and capacity. It is followed by YOLOv8l, which has an inference time of 35.1ms with 43.7 million parameters. Next is YOLOv8m with an inference time of 22.2ms and 25.9 million parameters. Then, YOLOv8s has an inference time of 16.1ms with 11.2 million parameters. Finally, YOLOv8n has the fastest inference time of 13.1ms, indicating very quick predictions. It has 3.2 million parameters, suggesting a lower complexity and capacity compared to the other YOLOv8 variants.

In terms of real-time frame processing, YOLOv8n outperforms other models, achieves 76 FPS. YOLOv8s follows with 52 FPS, YOLOv8m with 45 FPS, YOLOv8l with 28 FPS, and YOLOv8x with 19 FPS. Higher FPS values indicate preferable performance in fast and real-time video processing applications. In real-time applications, achieving higher FPS is desirable, as it allows the model to process more frames in a given amount of time, resulting in smoother and more responsive real-time predictions.

V. Conclusion

This paper presents a comprehensive evaluation of the real-time object detection performance of the YOLOv8 model in diverse and complex traffic environments, aiming to advance precise and efficient real-time object detection systems for self-driving cars. The study uncovers a significant disparity in object detection accuracy between daytime and nighttime for the YOLOv8 model. One noteworthy highlight of this study is YOLOv8n's capabilities. Despite lower accuracy, YOLOv8n demonstrates quick predictive ability, as proved by smaller inference time and higher FPS compared to other YOLOv8 models. Overall, challenges like low light conditions, glare, and small objects contribute to decreased accuracy at night, impeding the model's effectiveness. Accurately predicting objects on the road and at the side proves challenging, especially with fast-moving or differently directed objects. Therefore, considering the complexity of mixed-traffic, further improvements in YOLOv8's real-time detection capabilities are necessary for enhancement.

Acknowledgment

This work is funded by Institute for Research and Community Service, Universitas Syiah Kuala, Grant Number: 204/UN11.2.1/PT.01.03/PNBP/2023, May 03, 2023.

References

[1] H. -H. Jebamikyous and R. Kashef, "Autonomous Vehicles Perception (AVP) Using Deep Learning: Modeling, Assessment, and Challenges," in IEEE Access, vol. 10, pp. 10523-10535, 2022.

[2] K. Muhammad, A. Ullah, J. Lloret, J. D. Ser and V. H. C. de Albuquerque, "Deep Learning for Safe Autonomous Driving: Current Challenges and Future Directions," in IEEE Transactions on Intelligent Transportation Systems, vol. 22, no. 7, pp. 4316-4336, July 2021.

[3] A. Afdhal, N. Nasaruddin, Z. Fuadi, S. Sugiarto, H. Riza and K. Saddami, "Evaluation of Benchmarking Pre-Trained CNN Model for Autonomous Vehicles Object Detection in Mixed Traffic," 2022 International Conference on ICT for Smart Society (ICISS), Bandung, Indonesia, 2022, pp. 01-06

[4] J. Redmon, S. Divvala, R. Girshick and A. Farhadi, "You Only Look Once: Unified, Real-Time Object Detection," 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Las Vegas, NV, USA, 2016, pp. 779-788.

[5] J. Redmon and A. Farhadi, "YOLO9000: Better, Faster, Stronger," 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Honolulu, HI, USA, 2017, pp. 6517-6525.

[6] J. Redmon and A. Farhadi, "Yolov3: An incremental improvement," arXiv preprint arXiv:1804.02767, 2018.

[7] A. Bochkovskiy, C.-Y. Wang, and H.-Y. M. Liao, "Yolov4: Optimal speed and accuracy of object detection," arXiv preprint arXiv:2004.10934, 2020.

[8] C. -Y. Wang, H. -Y. Mark Liao, Y. -H. Wu, P. -Y. Chen, J. -W. Hsieh and I. -H. Yeh, "CSPNet: A New Backbone that can Enhance Learning Capability of CNN," 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW), Seattle, WA, USA, 2020, pp. 1571-1580.

[9] S. Liu, L. Qi, H. Qin, J. Shi and J. Jia, "Path Aggregation Network for Instance Segmentation," 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition, Salt Lake City, UT, USA, 2018, pp. 8759-8768.

[10] G. Jocher, YOLOv5 by Ultralytics, 2020, [Online]. Available: https://github.com/ultralytics/yolov5

[11] C. Li, L. Li, H. Jiang, K. Weng, Y. Geng, L. Li, Z. Ke, Q. Li, M. Cheng, W. Nie, et al., "Yolov6: A single-stage object detection framework for industrial applications," arXiv preprint arXiv:2209.02976, 2022.

[12] C.-Y. Wang, A. Bochkovskiy and H.-Y. Mark Liao, "YOLOv7: Trainable bag-of-freebies sets new state-of-the-art for real-time object detectors", arXiv:2207.02696, 2022.

[13] G. Jocher, A. Chaurasia, and J. Qiu, YOLO by Ultralytics (Version 8.0.0) [Computer software], 2023, [Online]. Available: https://github.com/ultralytics/ultralytics

[14] H. Liu, F. Sun, J. Gu, and L. Deng, "SF-YOLOv5: A Lightweight Small Object Detection Algorithm Based on Improved Feature Fusion Mode," Sensors, vol. 22, no. 15, p. 5817, Aug. 2022

[15] C. Zheng, "Stack-YOLO: A Friendly-Hardware Real-Time Object Detection Algorithm," in IEEE Access, vol. 11, pp. 62522-62534, 2023

Authorized licensed use limited to: VIT University- Chennai Campus. Downloaded on October 14,2024 at 04:08:35 UTC from IEEE Xplore. Restrictions apply.