Sinkron : Jurnal dan Penelitian Teknik Informatika
Volume 7, Number 4, October 2023
DOI : https://doi.org/10.33395/sinkron.v8i4.12787
e-ISSN : 2541-2019
p-ISSN : 2541-044X

Vehicle Detection and Identification Using Computer Vision Technology with the Utilization of the YOLOv8 Deep Learning Method

Agustritus Pasrah Hati Telaumbanua[1)*], Tri Putra Larosa[2)], Panji Dika Pratama[3)], Ra'uf Harris Fauza[4)], Amir Mahmud Husein[5)]
[1,2,3)]Universitas Prima Indonesia, Fakultas Sains dan Teknologi, Kota Medan
agustritustelaumbanua@gmail.com , amirmahmud@unprimdn.ac.id

Submitted : Jul 26, 2023 | **Accepted** : Aug 8, 2023 | **Published** : Oct 1, 2023

Abstract: Vehicle identification and detection is an important part of building intelligent transportation. Various methods have been proposed in this field, but recently the YOLOv8 model has been proven to be one of the most accurate methods applied in various fields. In this study, we propose a YOLOv8 model approach for the identification and detection of 9 vehicle classes in a reprocessed image data set. The steps are carried out by adding labels to the dataset which consists of 2,042 image data for training, 204 validation images and 612 test data. From the results of the training, it produces an accuracy value of 77% with the setting of epoch = 100, batch = 8 and image size of 640. For testing, the YOLOv8 model can detect the type of vehicle on video assets recorded by vehicle activity at intersections with. However, the occlusion problem overlapping vehicle objects has a significant impact on the accuracy value, so it needs to be improved. In addition, the addition of image datasets and data augmentation processes need to be considered in the future.

Keywords: Vehicle Detection, Vehicle identification, YOLOv8, Intelligent Transport

INTRODUCTION

With the rapid growth of vehicles in urban areas, the Intelligent Transportation System (ITS) is an effective solution to address traffic problems. Many big cities in Indonesia have developed ITS as traffic monitoring, and the city of Medan is no exception. One of the basic components of ITS is traffic monitoring with CCTV (Closed Circuit Television) cameras. However, in reality traffic monitoring currently still uses human power to respond to various incidents of traffic violations, vehicle accidents and vehicle theft which must be monitored using CCTV cameras.

One very important component of ITS is vehicle identification and detection, which aims to collect information about a certain type of vehicle from images or videos in the vehicle (Wang et al., 2021). Vehicle detection in real time is a crucial and challenging task. Currently, the available real-time vehicle detection is still not accurate and fast enough (Hasan et al., 2018; Xiao, 2019). Therefore, real-time systems must be able to detect and identify vehicles with a high degree of accuracy, especially in crime situations such as vehicle theft and road traffic violations. What's more, vehicle detection in complex and obstructed scenes is also a challenge (Phan et al., 2019).

Identification and detection of vehicles using computer vision technology with the use of deep learning methods is one of the research focuses that has attracted the attention of researchers in recent years, especially in the application of intelligent transportation systems (Al-Smadi et al., 2016). Many

*name of corresponding author

Sinkron : Jurnal dan Penelitian Teknik Informatika
Volume 7, Number 4, October 2023
DOI : https://doi.org/10.33395/sinkron.v8i4.12787
e-ISSN : 2541-2019
p-ISSN : 2541-044X

new models regarding deep learning technology have been reported in various studies with promising performance results for different purposes, including vehicle detection (Bugeja et al., 2020; Koay et al., 2021; Li et al., 2018; Pham & Yoo , 2020), vehicle re-identification (Hsu et al., 2020; Peng et al., 2020; Zheng et al., 2020a, 2020b), vehicle plate detection (Gupta et al., 2019), vehicle counting (Kadim et al ., 2020; Liang et al., 2020; Phan et al., 2019; Umair et al., 2021), vehicle tracking (Liang et al., 2020; López-Sastre et al., 2019; Phan et al., 2019), vehicle classification (Phan et al., 2019; Xiao, 2019; Yu et al., 2017) and others.

Various techniques have been proposed for intelligent traffic surveillance systems, but deep learning techniques are more accurate than traditional techniques (Zhang et al., 2019), because they have the ability to learn image or video features and classification and regression tasks (Wu et al., 2019). Therefore, in this study it is proposed to use one stage YOLOv8 for vehicle identification and detection, then the Deep Sort algorithm for vehicle tracking. The performance analysis of the proposed model will be trained on images of vehicle activity extracted from CCTV video recordings. An accurate model will be applied directly in live streaming by utilizing streamlit for making applications (Aboah et al., 2023; Ju & Cai, 2023; Terven & Cordova-Esparza, 2023).

LITERATURE REVIEW

Object detection is a computer vision task applied to predicting the presence of one or more objects, together with their classes and bounding boxes. YOLO (You Only Look Once) is a sophisticated object detector that can carry out real-time object detection with good accuracy. The success of YOLO in the previous version was applied in many fields, so that many developers and the community were very interested in trying the latest technology and the results were sure to be faster and more accurate than the previous version. YOLOv8 introduces new features and enhancements to further increase performance and flexibility. YOLOv8 is designed to be fast, accurate, and easy to use, making it an excellent choice for a wide variety of object detection and tracking, instance segmentation, image classification, and pose estimation tasks.

METHOD

The research stage proposed in this study is illustrated in Figure 1 below. Overall, the authors propose a computer vision approach with different deep learning techniques to understand how it works and analyze the performance of the most optimal model applied.

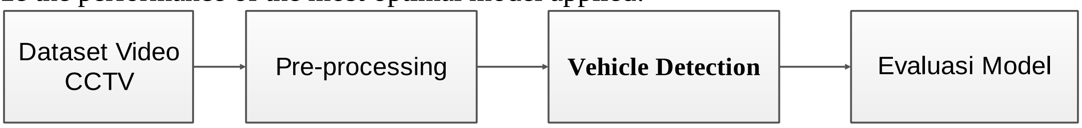


Fig 1 Research Stage

The model architecture proposed in the study is based on the deep learning technique of the YOLOv8 model for vehicle detection. The performance of the model will be analyzed for vehicle detection with the aim of obtaining the most optimal model applied when implementing a traffic control system in the city of Medan. then for direct vehicle tracking, the author adopts the Deep Sort algorithm which is one of the most optimal algorithms currently used for multi-object tracking compared to the ORB and Kalman Filter algorithms. YOLOv8 is the newest state-of-the-art YOLO model that can be used for object detection, image classification, and instance segmentation tasks. YOLOv8 was developed by Ultralytics, who also created the influential and industry-defining YOLOv5 model. YOLOv8 includes numerous architectural and developer experience changes and improvements over YOLOv5.

YOLOv8 is under active development as of writing this post, as Ultralytics work on new features and respond to feedback from the community. Indeed, when Ultralytics releases a model, it enjoys long-term support: the organization works with the community to make the model the best it can be. YOLOv8 achieves strong accuracy on COCO. For example, the YOLOv8m model -- the medium model -- achieves a 50.2% mAP when measured on COCO. When evaluated against Roboflow 100, a dataset that specifically evaluates model performance on various task-specific domains, YOLOv8 scored substantially better than YOLOv5. More information on this is provided in our performance analysis later in the article.

*name of corresponding author

Sinkron : Jurnal dan Penelitian Teknik Informatika
Volume 7, Number 4, October 2023
DOI : https://doi.org/10.33395/sinkron.v8i4.12787
e-ISSN : 2541-2019
p-ISSN : 2541-044X

Furthermore, the developer-convenience features in YOLOv8 are significant. As opposed to other models where tasks are split across many different Python files that you can execute, YOLOv8 comes with a CLI that makes training a model more intuitive. This is in addition to a Python package that provides a more seamless coding experience than prior models. The following image created by GitHub user RangeKing shows a detailed visualization of the network architecture. Figure 2 shows a detailed visualization of the YOLOv8 network architecture

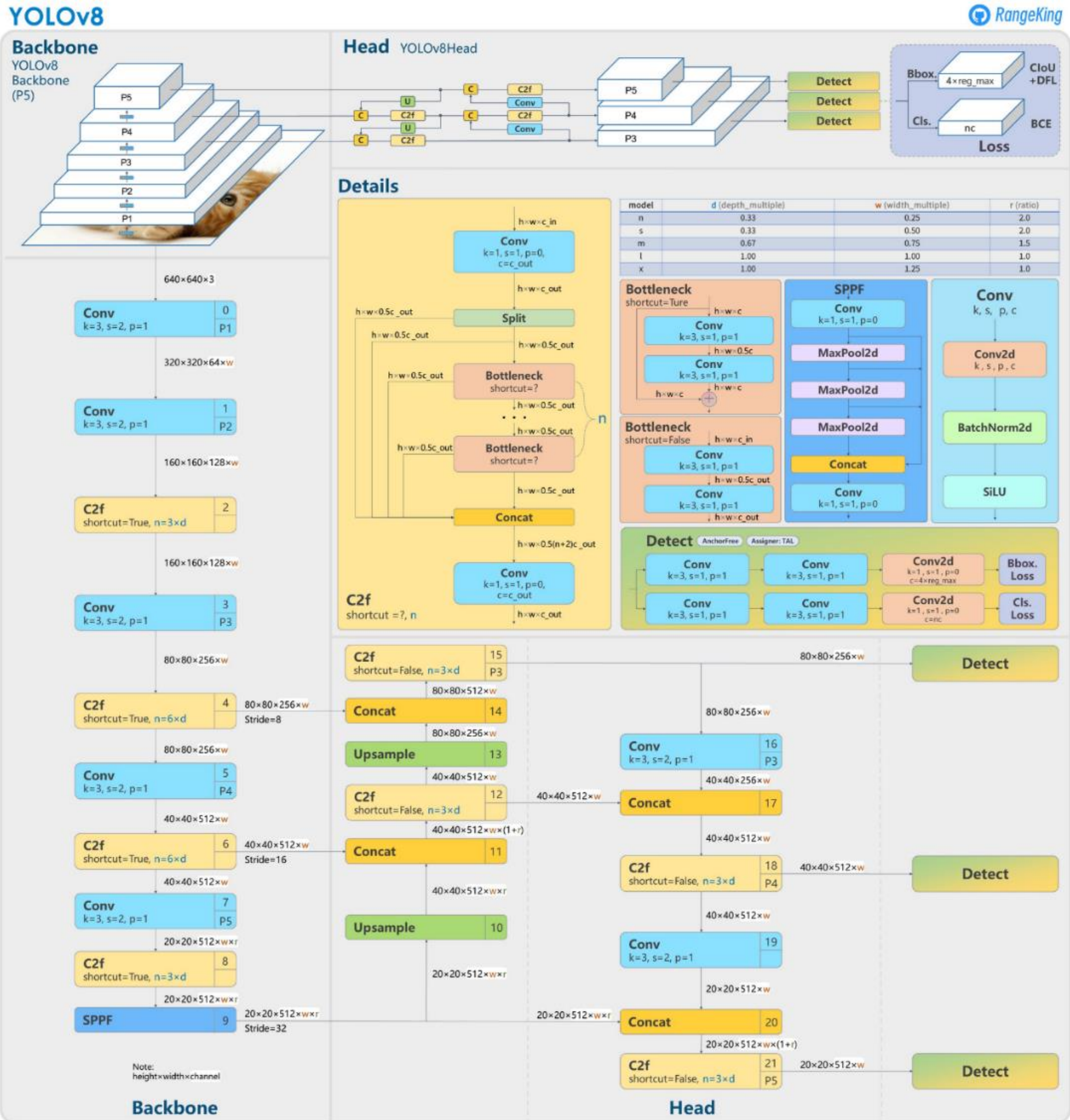


Fig 2 YOLOv8 architecture

The training data set sourced from CCTV video will be extra into several image scenes in jpg format. From the results of image extraction will be used for training needs, validation and testing. The dataset used for training consists of 2,042 images, 204 for validation and 612 for testing as seen in table 1

Table 1 Dataset

train	valid	test
2042	204	612

This dataset is used to train the YOLOv8 model in detecting and identifying 9 classes of vehicles namely 'ambulance', 'bus', 'citycar', 'minibus', 'minitruck', 'car', 'mobilbox', 'pickup', 'truck'. '.

RESULT

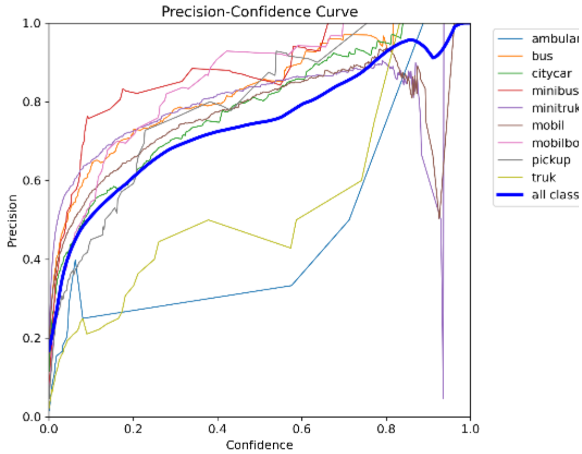
The results of this study consist of 5 (five) steps. First, collect images of vehicles sourced from CCTVV video recordings which are used as an initial stage for model training and evaluation. Second, processing

*name of corresponding author

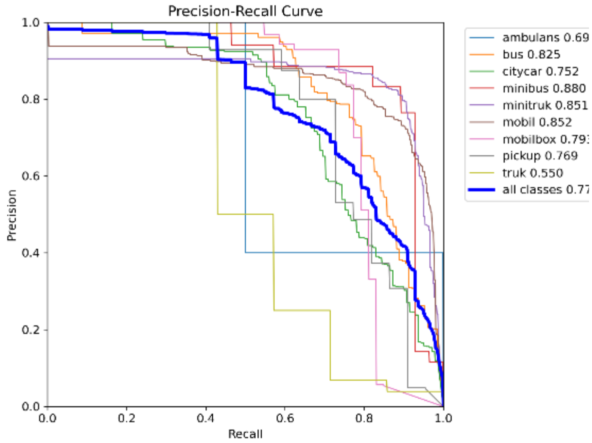
Sinkron : Jurnal dan Penelitian Teknik Informatika
Volume 7, Number 4, October 2023 e-ISSN : 2541-2019
DOI : https://doi.org/10.33395/sinkron.v8i4.12787 p-ISSN : 2541-044X

the entire data set by annotating and labeling vehicle objects in the image. Then, carry out the image data augmentation process using pre-existing photos. Next, we trained the YOLOv8 object recognition model using the vehicle data set. Then, analyze the results and validate the detection performance of the improved model using the validation data set to determine the best model for practical use in field adaptation.

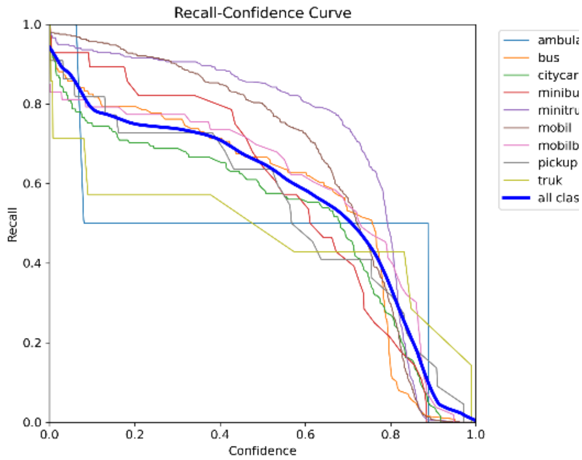
At the training stage we use the YOLOv8 model settings using epoch 100, image size = 640, batch = 8 and optimization using adams. We use the intersection over union (IoU) evaluation metric to evaluate the accuracy of the model detector in the dataset used. Figure 3 is an illustration of the training results based on the error value of the model in training. There are 4 illustrated images, namely Precision-Confidence, Recall, Recall Confidence and F1-Confidence. From the results of the image the accuracy value for the detection of 9 vehicle classes is 77%. Furthermore, in Figure 4 is an illustration of the model training matrix and Figure 5 is the result of the Confusion matrix model used. Confusion matrix is a method for summarizing the performance of classification algorithms. The diagonal lines show the importance of the predicted results in the confusion matrix; horizontal and vertical lines represent false negatives and false positives, respectively.



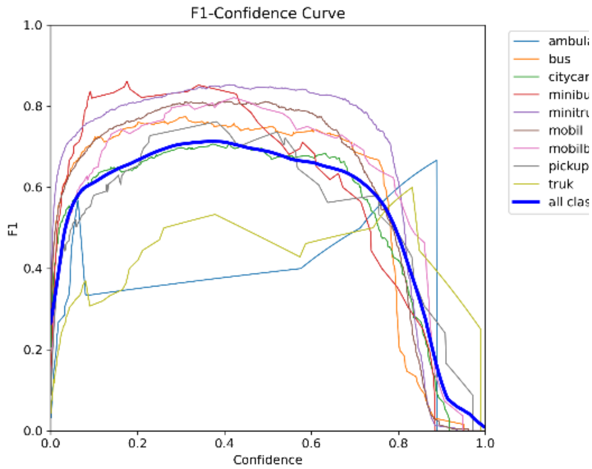
(a)



(b)



(c)



(d)

Fig 3 (a) Precision Confidence, (b) Recall, (c) Recall Confidence dan (d) F1-Confidence

*name of corresponding author

Sinkron : Jurnal dan Penelitian Teknik Informatika
Volume 7, Number 4, October 2023
DOI : https://doi.org/10.33395/sinkron.v8i4.12787

e-ISSN : 2541-2019
p-ISSN : 2541-044X

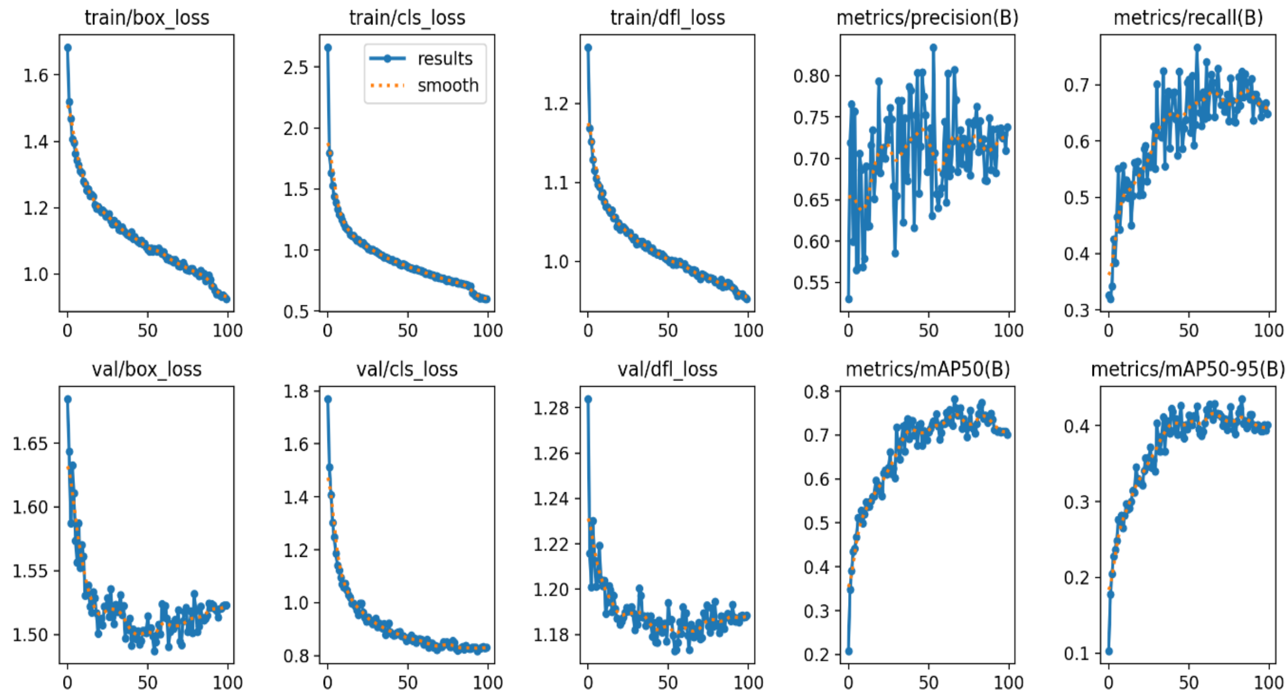


Fig 4 Training matrix

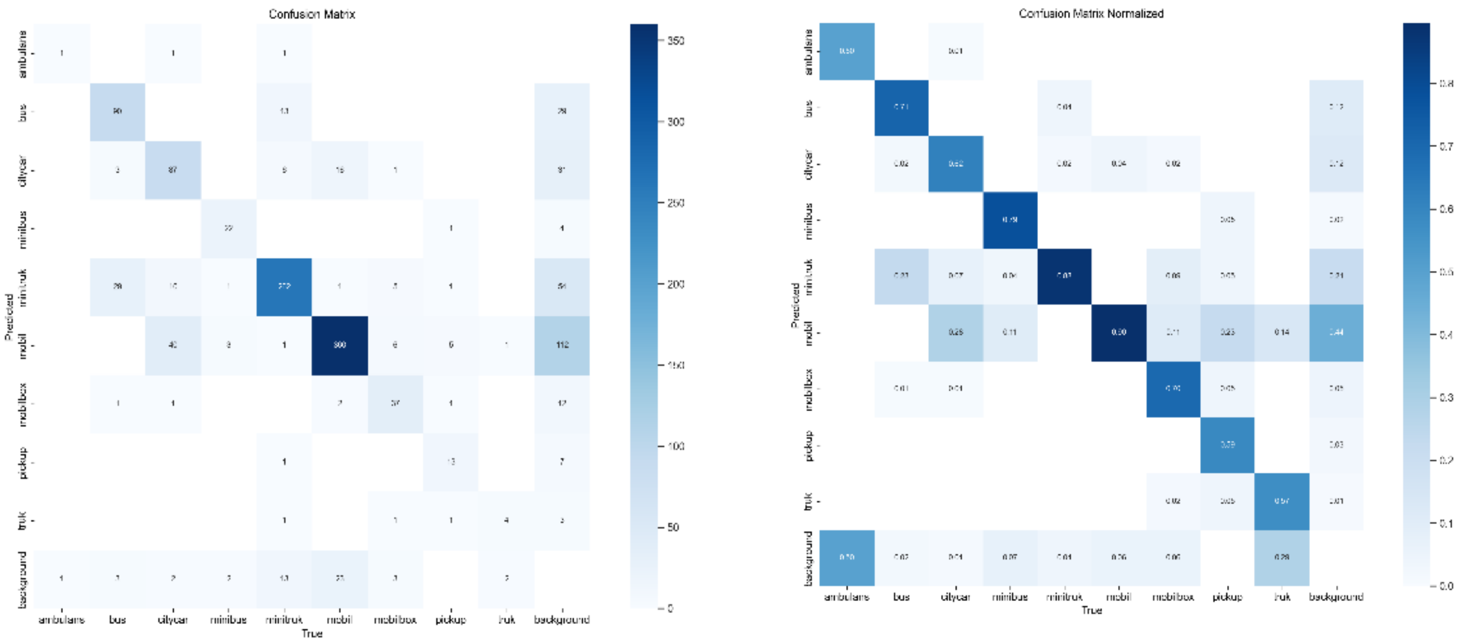


Fig 5 Confusion matrix

The results of the YOLOv8 model training used for detection by identifying 9 vehicle classes will produce a new model with the name best.pt. Then, this model is used to perform vehicle detection on test images consisting of 612 images which produce an accuracy rate of 96% and some of the detection results are shown in the image below.

*name of corresponding author

Sinkron : Jurnal dan Penelitian Teknik Informatika
Volume 7, Number 4, October 2023
DOI : https://doi.org/10.33395/sinkron.v8i4.12787
e-ISSN : 2541-2019
p-ISSN : 2541-044X

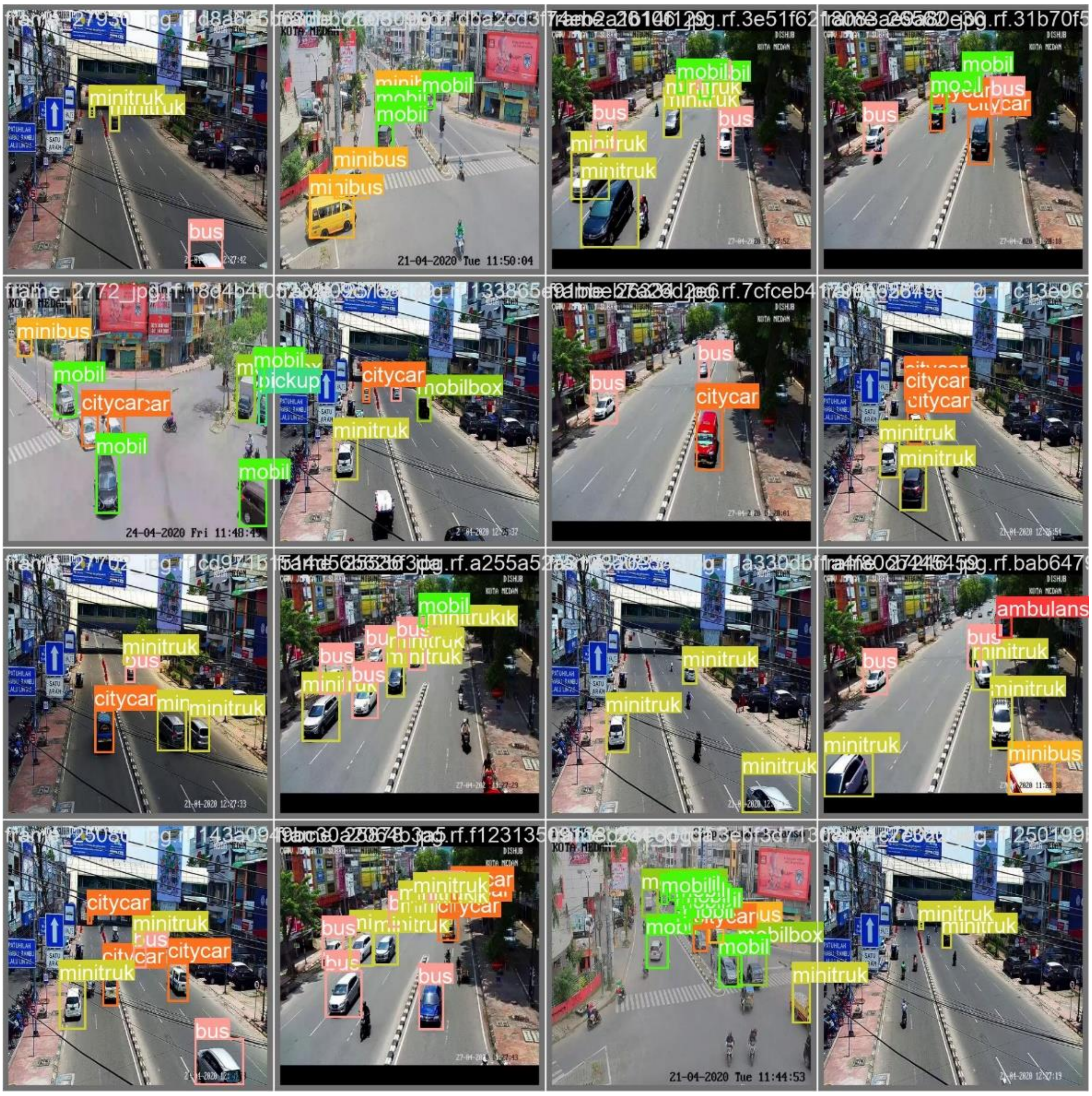


Fig 6 Test result visualization

DISCUSSIONS

Based on the test results of the YOLOv8 model used for the detection and identification of 9 classes of vehicles, the highest training accuracy value of 77% is one of the values that still needs to be improved. In addition, the YOLOv8 model training process still needs further research, especially in the application of the image augmentation process. Several research results found that the image augmentation process can improve detection accuracy, as proposed by [25] who applied a combination of MixUp, Mosaic, and traditional methods in data augmentation. The results of their model testing resulted in increased accuracy. In addition, the detection of small objects in the scene is also an important part; application of SAHI [26] is an option to be used in future research.

CONCLUSION

The application of the YOLOv8 model for vehicle detection and identification proposed in this study still needs to be developed, but the YOLOv8 model for the vehicle dataset used yields an accuracy of 77% for training and 96% for testing. The data augmentation process used in this study uses standard settings from the YOLOv8 model. Detection and identification of 9 vehicle classes still needs to be developed for the future.

e-ISSN : 2541-2019
p-ISSN : 2541-044X

REFERENCES

Aboah, A., Wang, B., Bagci, U., & Adu-Gyamfi, Y. (2023). *Real-time Multi-Class Helmet Violation Detection Using Few-Shot Data Sampling Technique and YOLOv8*. https://arxiv.org/abs/2304.08256v1

Al-Smadi, M., Abdulrahim, K., & Salam, R. A. (2016). Traffic surveillance: A review of vision based vehicle detection, recognition and tracking. *International Journal of Applied Engineering Research*, *11*(1), 713–726.

Bugeja, M., Dingli, A., Attard, M., & Seychell, D. (2020). Comparison of Vehicle Detection Techniques applied to IP Camera Video Feeds for use in Intelligent Transport Systems. *Transportation Research Procedia*, *45*, 971–978. https://doi.org/10.1016/j.trpro.2020.02.069

Gupta, M., Kumar, A., & Madhvanath, S. (2019). Zero shot license plate re-identification. *Proceedings - 2019 IEEE Winter Conference on Applications of Computer Vision, WACV 2019*, 773–781. https://doi.org/10.1109/WACV.2019.00087

Hasan, A. F., Che Husin, M. F., Rosli, K. A., Hashim, M. N., & Zainal Abidin, A. F. (2018). Multiple Vehicle Detection and Segmentation in Malaysia Traffic Flow. *IOP Conference Series: Materials Science and Engineering*, *318*(1), 0–9. https://doi.org/10.1088/1757-899X/318/1/012041

Hsu, H. M., Wang, Y., & Hwang, J. N. (2020). Traffic-aware multi-camera tracking of vehicles based on ReID and camera link model. *ArXiv*, 416–424. https://doi.org/10.1145/3394171.3413863

Ju, R.-Y., & Cai, W. (2023). *Fracture Detection in Pediatric Wrist Trauma X-ray Images Using YOLOv8 Algorithm*. https://arxiv.org/abs/2304.05071v2

Kadim, Z., Johari, K. M., Samaon, D. F., Li, Y. S., & Hon, H. W. (2020). Real-Time Deep-Learning Based Traffic Volume Count for High-Traffic Urban Arterial Roads. *ISCAIE 2020 - IEEE 10th Symposium on Computer Applications and Industrial Electronics*, *C*, 53–58. https://doi.org/10.1109/ISCAIE47305.2020.9108799

Koay, H. V., Chuah, J. H., Chow, C. O., Chang, Y. L., & Yong, K. K. (2021). Yolo-rtuav: Towards real-time vehicle detection through aerial images with low-cost edge devices. *Remote Sensing*, *13*(21), 4196. https://doi.org/10.3390/rs13214196

Li, X., Liu, Y., Zhao, Z., Zhang, Y., & He, L. (2018). A deep learning approach of vehicle multitarget detection from traffic video. *Journal of Advanced Transportation*, *2018*. https://doi.org/10.1155/2018/7075814

Liang, H., Song, H., Li, H., & Dai, Z. (2020). Vehicle Counting System using Deep Learning and Multi-Object Tracking Methods. *Transportation Research Record*, *2674*(4), 114–128. https://doi.org/10.1177/0361198120912742

López-Sastre, R. J., Herranz-Perdiguero, C., Guerrero-Gómez-olmedo, R., Oñoro-Rubio, D., & Maldonado-Bascón, S. (2019). Boosting multi-vehicle tracking with a joint object detection and viewpoint estimation sensor. *Sensors (Switzerland)*, *19*(19), 4062. https://doi.org/10.3390/s19194062

Peng, J., Jiang, G., Chen, D., Zhao, T., Wang, H., & Fu, X. (2020). Eliminating cross-camera bias for vehicle re-identification. *Multimedia Tools and Applications*. https://doi.org/10.1007/s11042-020-09987-z

Pham, T. A., & Yoo, M. (2020). Nighttime vehicle detection and tracking with occlusion handling by pairing headlights and taillights. *Applied Sciences (Switzerland)*, *10*(11), 3986. https://doi.org/10.3390/app10113986

Phan, H. N., Pham, L. H., Thai, T. T., Chung, N. M., & Ha, S. V. U. (2019). A real-time vehicle detection for traffic surveillance system using a neural decision tree. *Proceedings of 2019 25th Asia-Pacific Conference on Communications, APCC 2019*, 256–261. https://doi.org/10.1109/APCC47188.2019.9026506

Terven, J., & Cordova-Esparza, D. (2023). *A Comprehensive Review of YOLO: From YOLOv1 to YOLOv8 and Beyond*. https://arxiv.org/abs/2304.00501v1

Umair, M., Farooq, M. U., Raza, R. H., Chen, Q., & Abdulhai, B. (2021). Efficient Video-Based Vehicle Queue Length Estimation Using Computer Vision and Deep Learning for an Urban Traffic Scenario. *Processes*, *9*(10), 1786. https://doi.org/10.3390/pr9101786

Wang, R., Wang, Z., Xu, Z., Wang, C., Li, Q., Zhang, Y., & Li, H. (2021). A Real-Time Object Detector

e-ISSN : 2541-2019
p-ISSN : 2541-044X

for Autonomous Vehicles Based on YOLOv4. *Computational Intelligence and Neuroscience*, *2021*, 1–11. https://doi.org/10.1155/2021/9218137

Wu, C. E., Chan, Y. M., Chen, C. H., Chen, W. C., & Chen, C. S. (2019). IMMVP: An efficient daytime and nighttime on-road object detector. In *arXiv*.

Xiao, Y. (2019). *Vehicle Detection in Deep Learning*. http://arxiv.org/abs/1905.13390

Yu, S., Wu, Y., Li, W., Song, Z., & Zeng, W. (2017). A model for fine-grained vehicle classification based on deep learning. *Neurocomputing*, *257*, 97–103. https://doi.org/10.1016/j.neucom.2016.09.116

Zhang, P., Zhong, Y., & Li, X. (2019). SlimYOLOv3: Narrower, faster and better for real-time UAV applications. In *arXiv*.

Zheng, Z., Ruan, T., Wei, Y., Yang, Y., & Mei, T. (2020a). VehicleNet: Learning robust visual representation for vehicle re-identification. *ArXiv*, 1–4. https://doi.org/10.1109/tmm.2020.3014488

Zheng, Z., Ruan, T., Wei, Y., Yang, Y., & Mei, T. (2020b). VehicleNet: Learning robust visual representation for vehicle re-identification. In *arXiv*. arXiv. https://doi.org/10.1109/tmm.2020.3014488

*name of corresponding author