JURNAL UPI

Jurnal Pendidikan Multimedia
(EDSENCE)

Journal homepage: https://ejournal.upi.edu/index.php/Edsence

Development of Object Detection System on Non-Helmed Riders Using YOLOv8

Bima Prihasto[1], Nisa Rizqiya Fadhliana[2], Agustina Hariyani[3], Fauzan M. Alwafi[4], and Tsaqila B. Askarin[5]

[1,2,3,4,5] Informatics Department, Institut Teknologi Kalimantan, Balikpapan, Indonesia
Correspondence: E-mail: bima@lecturer.itk.ac.id

ABSTRACT

Motorcycle accidents are a severe problem, with the number of incidents reaching 66,602 in 2023. Helmets as head protection are mandatory, but awareness of their use is still low. This research utilises Deep Learning, specifically YOLOv8, to detect helmet use violations among motorbike riders. The research results show high accuracy with a Precision of 89.5%, Recall at 78.4%, and mAP50 at 85.7%. YOLOv8 effectively detects violations and provides a solid basis for increasing motorist awareness. Through this innovative approach, it is hoped that a safer driving culture and collective awareness of responsibility in traffic safety will be created.

ARTICLE INFO

Article History:
Submitted/Received 15 August 2023
First Revised 5 December 2023
Accepted 11 December 2023
First Available online 27 December 2023
Publication Date 31 December 2023

Keyword:
YOLOv8,
Object Detection,
Traffic,
Motorcyclists,
Helmets.

1. INTRODUCTION

By 2023, motorcycle accidents had the highest number, with 66,602 accidents (Rizqiyah, 2023). A helmet is a driving accessory that must be used as a head protector to reduce the risk of severe injuries in traffic accidents. Using the Indonesian national standard helmet while driving a motorcycle is an obligation regulated in Law No. 22 of 2009 on Traffic and Road Transport Article 57, paragraphs 1 and 2. Therefore, compliance with these rules is not just an obligation but an essential step in improving the safety of crossings. Nevertheless, the awareness of helmet use is still less optimal. So, it requires innovative technology to raise drivers' awareness in compliance with the rules (Khoiriyah & Armawan, 2023).

Deep Learning is a subfield of machine learning with a layered neural network architecture to process data by generating more accurate outputs. Deep Learning will automatically extract complex and abstract features from existing data (Setyawan et al., 2021). A survey on face recognition using deep learning has been conducted by Prihasto et al. (2016) Deep Learning technology in breach detection does not use helmets as an innovative solution to improve efficiency and accuracy. YOLOv8 can perform high-complexity real-time data processing, even under dense and varied traffic conditions.

This study aims to detect violations and raise awareness among drivers regarding the importance of helmet use. The system is expected to provide information directly to drivers who violate the rules, providing an opportunity to introspect and better understand the risks that will be faced. This approach is expected to support the transition to a safer driving culture and an awareness of collective responsibility in cross-traffic safety issues.

2. METHODS

This research uses one of the deep learning algorithms of You Only Look Once (YOLO), specifically YOLOv8. YOLO is an algorithm that belongs to the category of deep learning and was proposed by Joseph Remond in 2015. The YOLO algorithm uses a neural network to divide the input image into a grid that then performs image analysis and detection of bounding boxes and their probability (Setiyadi et al., 2023). The study focuses on detecting helmet use in motorcyclists to see if the driver is using a helmet and not using a helmet.

DOI: https://doi.org/10.17509/edsence.v5i2.65910
p-ISSN:2685-2489 | e-ISSN:2685-2535

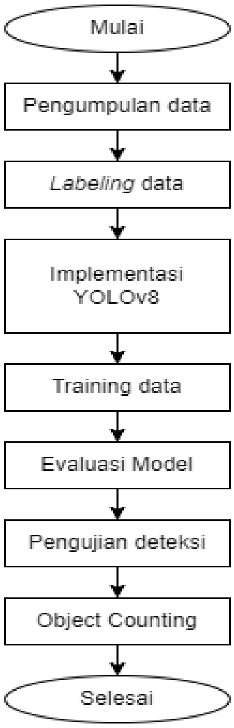


Figure 1. Research flow

From Figure 1, the research flow commenced with the initial stage of the study, where data collection was conducted to gather relevant information from various sources. The collected data encompassed the necessary elements for the research objectives. Subsequently, a meticulous data labelling process was performed, ensuring accurate and consistent annotations for each data point. This step aimed to enhance the dataset's quality and facilitate subsequent analysis. To further optimise object detection capabilities, YOLOv8 complementation was implemented. They integrated YOLOv8, a state-of-the-art object detection algorithm, into the research framework. The detection accuracy and efficiency were significantly improved by leveraging its advanced features. The labelled data and the YOLOv8 complementation were the foundation for data training. This process involved feeding the dataset into the model to allow it to learn and develop the ability to recognise and classify objects accurately. Following training, the model underwent a comprehensive evaluation to assess its performance. Various metrics and performance indicators were employed to gauge the model's effectiveness in meeting the research objectives. Detection testing used diverse scenarios and settings to validate the model's real-world applicability. This step assessed the model's robustness and reliability in different environments. Additionally, object counting was performed to quantify the number of identified objects accurately. This process involved analysing the results generated by the model to obtain reliable object count estimations. Finally, the research flow culminated in the concluding phase, summarising the findings and insights obtained throughout the research process. This final stage aimed to comprehensively understand the research outcomes and their implications within the chosen domain.

DOI: https://doi.org/10.17509/edsence.v5i2.65910
p-ISSN:2685-2489 | e-ISSN:2685-2535

2.1. You Only Look Once (YOLO)

You Only Look Once (YOLO) is a deep learning algorithm with an intelligent neural network that is useful for direct detection. YOLO also typically uses a single neural network in the entire image; then, the network will be divided into several areas, and boundary boxes and probability will be predicted. For each box in the boundary region, the possibility will be considered to do the classification, whether the object or not (Iskandar Mulyana, 2022). The detection network on YOLO itself is known to have several layers, namely 24 layers of convolution or can be said as a convolutional layer and at the same time as a fully connected layer or may be said to be a fully connected layer present in the image.

2.2. YOLOv8

The YOLO model (You Only Look Once) has been an enormous success in computer vision, based on which researchers have improved and added new modules to this method, proposing many classic models. YOLOv8 is an algorithm released by Ultralytics on January 10, 2023. Compared to previous superior models in the YOLO series (such as Yolov5 and YOLOw7), YOLOV8 is a sophisticated and latest model that offers higher detection accuracy and speeds.

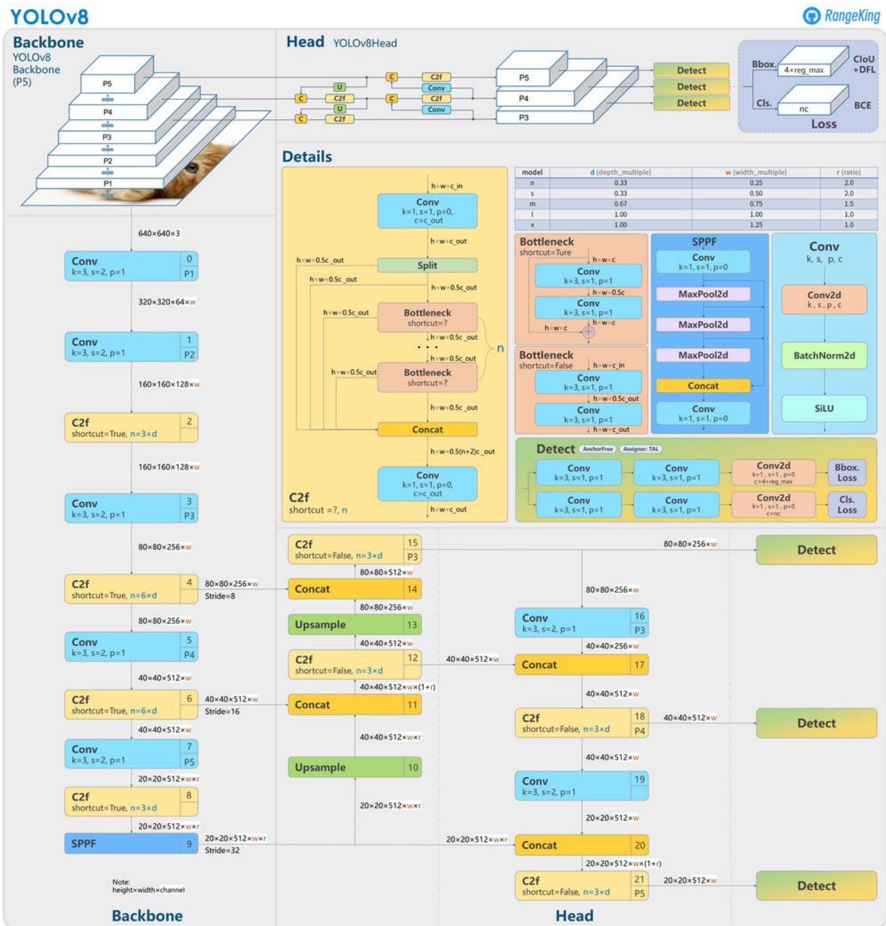


Figure 2. YOLOv8 architecture
Source: (Zhuang et al., 2023)

DOI: https://doi.org/10.17509/edsence.v5i2.65910
p-ISSN:2685-2489 | e-ISSN:2685-2535

In YOLOv8 architecture, like in Figure 2, some elements perform input image processes and generate predictions. Here is the definition of each element:

- Conv (convolutional) is a type of layer in the convolutional nerve network that processes input images and implements convolutional filters that allow the nerve tissue to extract essential features in the image.
- C2F (Cross to Feature) is a type of layer that contains a CSP Bottleneck involving 3 convolution processes.
- SPPF (Spatial Pyramid Pooling with FPN) is the type of Layer integrated with the FPN to assist models in understanding objects with different levels of resolution in a layer.
- Concat is a Layer that combines several outputs from two different paths within a neural network.
- Upsample is the layer that enlarges the size of the input image by increasing the dimension of the feature map to improve the resolution.
- Detect is the layer responsible for predicting the boundary box, object score, and probability of the object detected in the image.
- A bottleneck is a block that functions to extract existing features from the input image

There is a layer structure of YOLOv8 with three main parts: the backbone, neck, and head. CSPDarknet53 was modified as a backbone network with input features taken as many as 5 times to produce 5 different features. PAN_FPN on the neck part supplements short information with excess information to generate more diverse information and complete features. Using two separate head structures, namely object classification and prediction, can improve accuracy and accelerate model convergence (Wang et al., 2023).

2.3. Model Evaluation

Performance metrics are an essential tool for evaluating the accuracy and potential of an object identification algorithm. They explain the ability of the model to find and identify objects in the image and help understand how the model deals with positives and false negatives. This insight is crucial to evaluating and improving the functionality of the model. According to Ultralytics documentation (YOLO Performance Metrics, 2023), performance metrics on object detection are as follows.

- Intersection over Union (IoU): Overlap between predicted boundary boxes and baseline truth boundaries measured using the IoU. It is essential to assess how accurate the location of the object is.
- Average Precision (AP): Count the area below the precision-recall curve, combining model accuracy and recall performance into a single metric.
- Mean Averted Precision(mAP): represents the average AP value in several classes of objects. It is helpful in thoroughly assessing model performance to detect and classify objects with a certain degree of precision.
- Precision and Recall: Precision calculates the proportion of true positives between all positive predictions to assess the ability of the model to avoid false positives. Recall measures how well the model identifies each instance of a class by finding out how much the actual positive proportion is genuinely positive.

DOI: https://doi.org/10.17509/edsence.v5i2.65910
p-ISSN:2685-2489 | e-ISSN:2685-2535

- F1 Score: Performance metrics combine precision and recall information in a single value to measure the performance of a model

2.4. Object Counting

Object Counting is used to count objects passing through the calculation line. This study uses object counting to calculate how many motorcyclists do not wear helmets. ByteTrack & Supervision usage. Zhang et al. Introduced ByteTrack, the real-time object-tracking algorithm. The goal is to efficiently track objects in video series (Zhang et al., 2022). Before performing object calculations, object tracking with ByteTrack is done to help track the movement of objects from each frame in a video or series of images. Supervision helps calculate objects by detecting the use of helmets on motorcyclists and counting objects passing the line of calculation (counting).

3. RESULTS AND DISCUSSION

3.1. Data Preparation

The training process using YOLOv8 could not be started with a ready-to-use dataset. Data preparation at this stage is the process of data collection and data labelling. In the labelling process, each object will be trained in an image with a bounding box or boundary box labelled, and then the coordinates of the bounded box will be stored in a ".xml" file (Lee et al., 2019). Figure 3 shows an example of labelling on the object to be used.

Figure 3. Examples of berhelm (helmet) & tidak berhelm (no helmet) class labeling

This data labelling process involves two classes, namely "Berhelm" and "Tidak Berhelm." In training, each image is required to recognise the classification to be processed. This labelling process produces a total of 1346 objects out of 600 images. The composition of objects on the dataset is as follows.

DOI: https://doi.org/10.17509/edsence.v5i2.65910
p-ISSN:2685-2489 | e-ISSN:2685-2535

Table 1. Object composition

Label Class	Number
Berhelm (Helmet)	651
Tidak Berhelm (No Helmet)	695
Total	**1346**

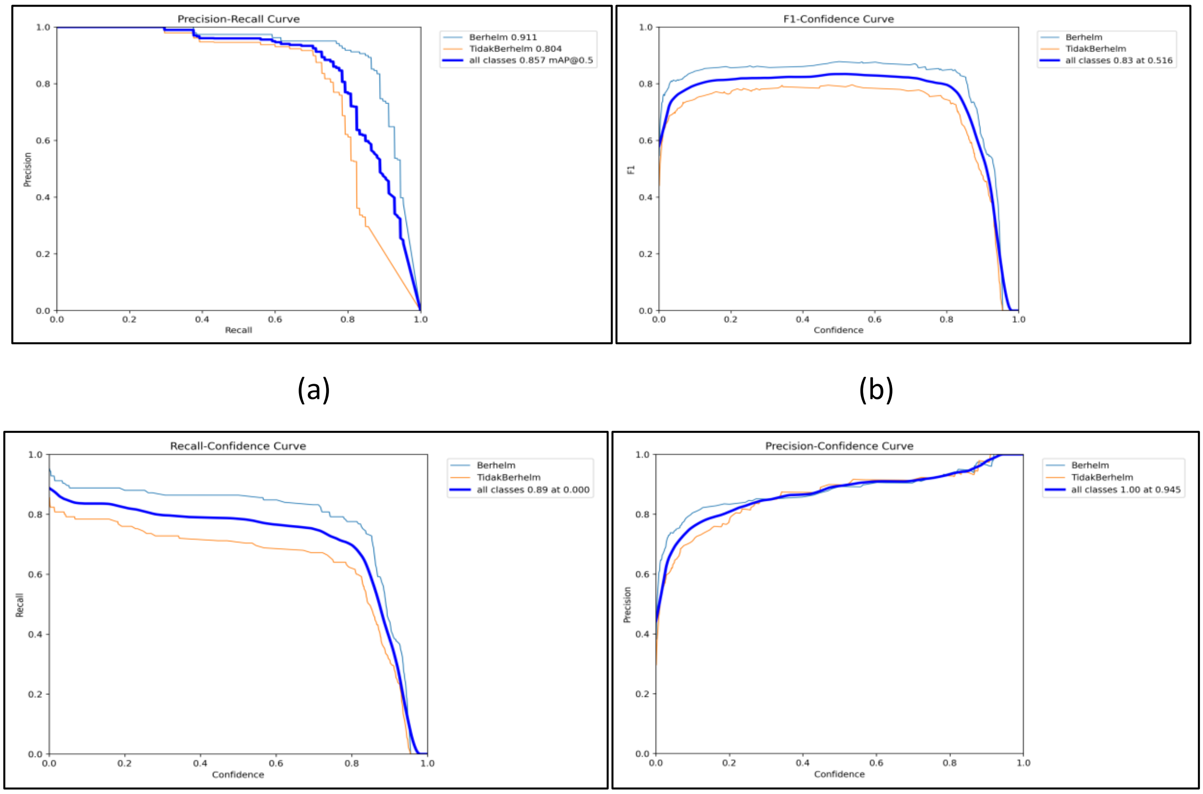
3.2 Model Performance Evaluation

Before the model can predict image or video input, the training is done with the division of training data of 80%, testing data of 10%, and validation data of 10%. Then, the evaluation is carried out with evaluation metrics, namely precision, recall, and mAP50. Table 2 shows the evaluation results of the YOLOv8 model in helmet detection on motorcyclists as follows.

Table 2. Evaluation Results

Class	Precision	Recall	mAP50
All	89.5%	78.4%	85.7%
Berhelm (Helmet)	89.2%	86.3%	91.1%
Tidak Berhelm (No Helmet)	89.8%	70.6%	80.4%

As for model performance, visualisation based on evaluation metrics on validation data is as follows.



(a) (b)

(c) (d)

Figure 4. (a) Precision-Recall Curve (b) F1 Score Curve (c) Recall Curve (d) Precision Curve

DOI: https://doi.org/10.17509/edsence.v5i2.65910
p-ISSN:2685-2489 | e-ISSN:2685-2535

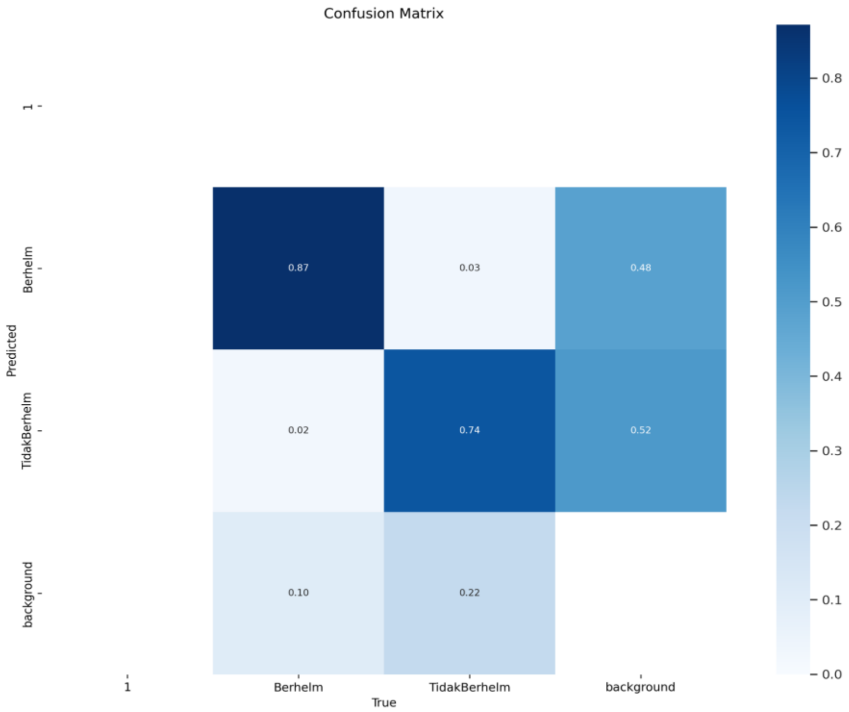


Figure 5. Confusion Matrix

The prediction results on the validation data on the motorcycle driver's helmet detection are as follows.

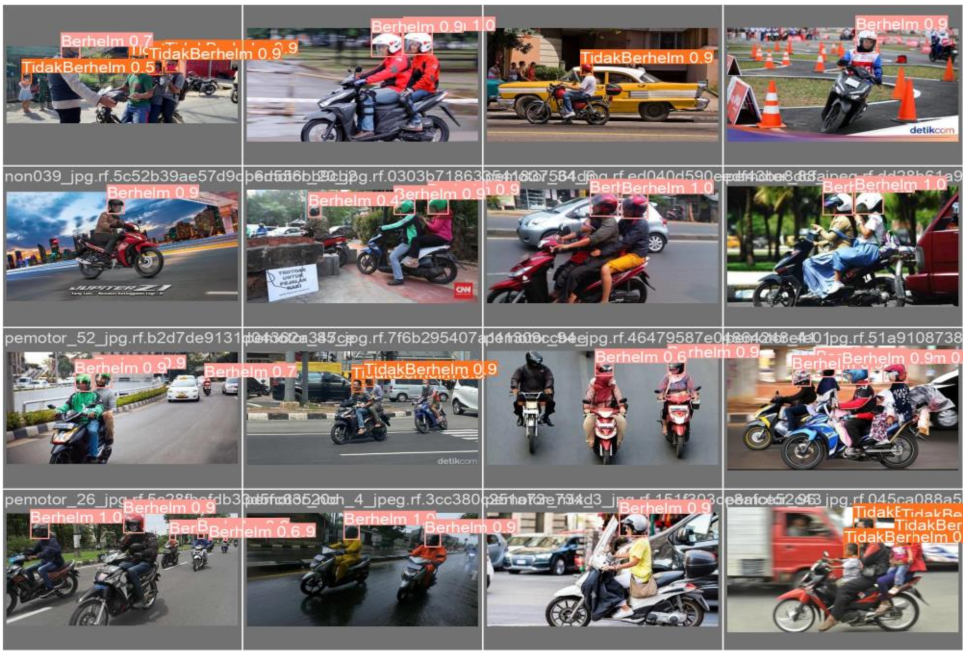


Figure 6. Detection Results on Validation Data

DOI: https://doi.org/10.17509/edsence.v5i2.65910
p-ISSN:2685-2489 | e-ISSN:2685-2535

3.3. Model Implementation

The following process is the implementation of a model on the input data in the form of images or videos as well as video in real-time running in Streamlit. The model implementation on motorist helmet detection is as follows.

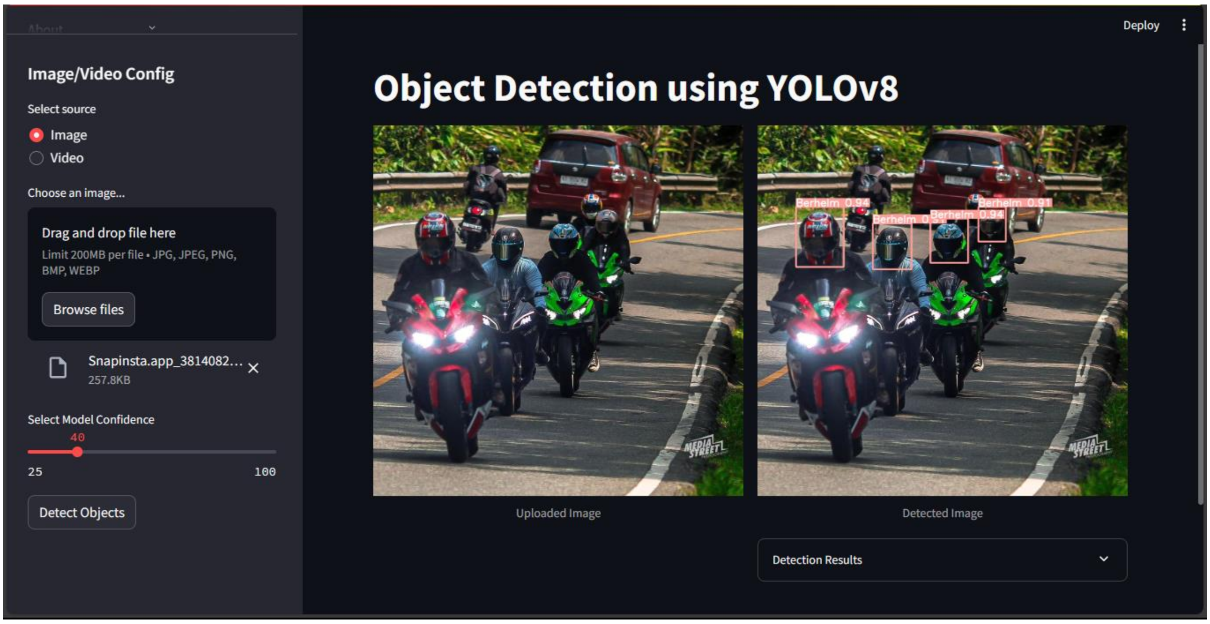


Figure 7. Image Input Data Detection Results

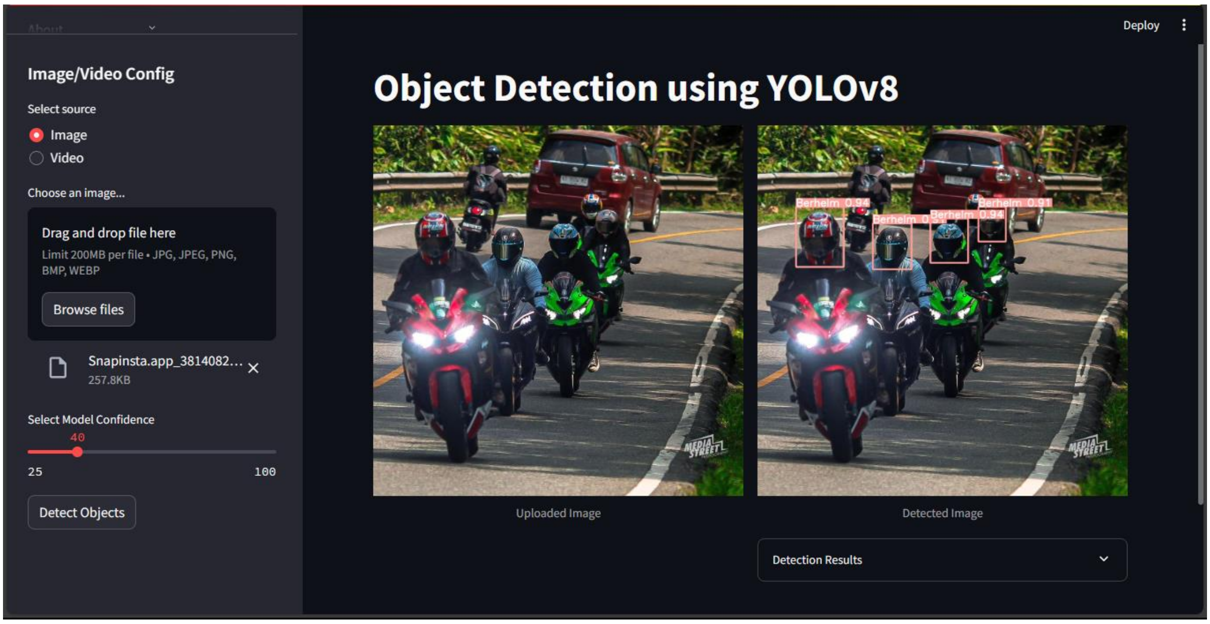


Figure 8. Video Input Data Detection Results

DOI: https://doi.org/10.17509/edsence.v5i2.65910
p-ISSN:2685-2489 | e-ISSN:2685-2535

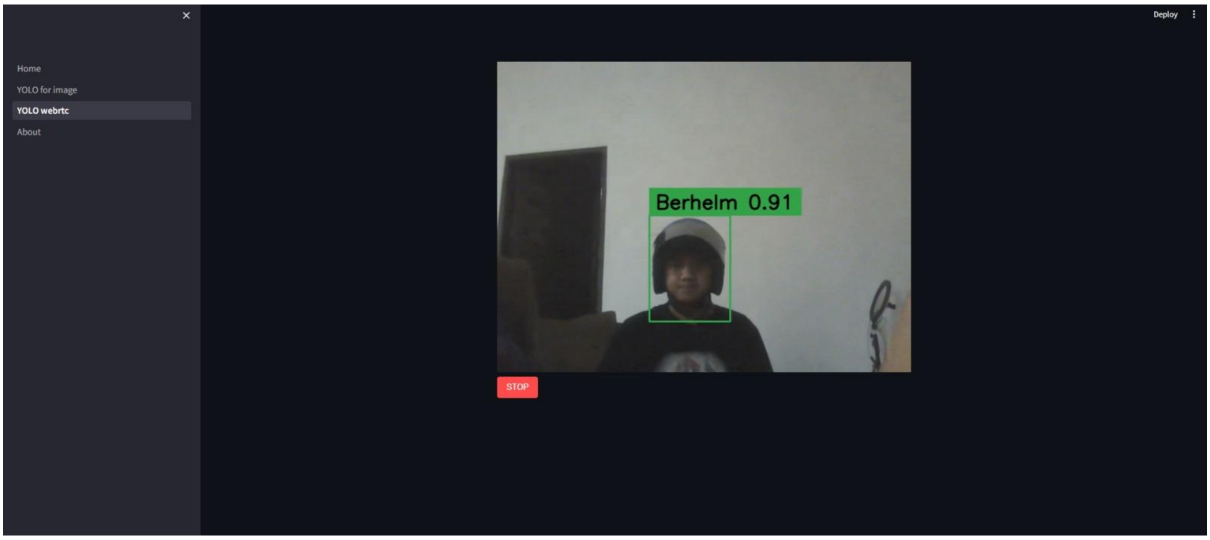


Figure 9. Figure 9. Real-time detection results

Figures 7, 8, and 9 show that the YOLOv8 model can detect helmet usage in motorcyclists based on whether the motorcycle driver uses a helmet. As for the result with video input data, it also produces pretty good detection, but as the object gets farther away, the accuracy will also decrease. The accuracy of the detection also depends on the given video quality

3.4 Object Counting Implementation

Figure 10. Object Counting Detection Results

DOI: https://doi.org/10.17509/edsence.v5i2.65910
p-ISSN:2685-2489 | e-ISSN:2685-2535

Based on the detection results from the input data, it can be seen that the YOLOv8 model can be used to detect helmet usage on motorcyclists and count objects passing the calculation line. (counting). The counting line itself is made into two entrances and exits. (out). Object counting will only work when the bounding box of the detection results crosses or crosses the calculation line. However, detection results sometimes still do not appear on some frames in the video, causing some detection outcomes to be delayed from the calculation line and not counted as objects. It can be caused by several factors, such as objects unclear on some frames, poor input data quality, motion, or model parameters that are not optimal.

4. CONCLUSION

In conclusion, our research has demonstrated the remarkable accuracy of the YOLOv8 model in detecting non-helmet riders. With a Precision (setting) of 89.5%, the model effectively identifies riders who do not wear helmets. Furthermore, it achieves a Recall of 78.4%, successfully detecting the majority of non-helmeted riders. The mAP50 score of 85.7% provides a comprehensive overview of the model's advantages in processing information and extracting it with high confidence. The success of the YOLOv8 model in handling custom problems, such as helmet detection, can be attributed to the complexity of its algorithm, which contributes to its impressive accuracy. However, there is still potential for improvement in the data preprocessing stage, including incorporating additional image processing methods. Furthermore, testing under more complex scenarios, such as adverse weather conditions or low lighting, would provide valuable insights into the model's performance and potential areas for enhancement. In summary, our research underscores the efficacy of the YOLOv8 model for non-helmet rider detection. While there are opportunities for refinement, the achieved performance is commendable. These findings lay the groundwork for advancing rider safety and promoting the development of more sophisticated systems in real-world applications.

REFERENCES

Khoiriyah, K., & Armawan, M. F. A. A. (2023, Juli 2). Deteksi pengendara Motor Tanpa Menggunakan Helm dengan Yolo Algoritma Deep Learning Yolo. Jurnal Elektro & Informatika Swadharma (JEIS), 03(02).

Lee, Y., Im, D., & Shim, J. (2019). Data Labeling Research for Deep Learning Based Fire Detection System. International Conference on Systems of Collaboration Big Data, Internet of Things & Security (SysCoBIoTS), 1-4. 10.1109/SysCoBIoTS48768.2019.9028029.

Mulyana, D. I., & Rowis, M. A. I. (2022). Optimization of Text Mining Detection of Tajweed Reading Laws Using the Yolov8 Method on the Qur'an. QALAMUNA: Jurnal Pendidikan, Sosial, dan Agama, 14(2), 1089-1110.

Prihasto, B., Choirunnisa, S., Nurdiansyah, M. I., Mathulaprangsan, S., Chu, V. C.-M., Chen, S.-H., & Wang, J.-C. (2016). A survey of deep face recognition in the wild. 2016 International Conference on Orange Technologies (ICOT). https://doi.org/10.1109/icot.2016.8278983

DOI: https://doi.org/10.17509/edsence.v5i2.65910
p-ISSN:2685-2489 | e-ISSN:2685-2535

Rizqiyah, A. (2023). Angka Kecelakaan Lalu Lintas Terus Meningkat, Usia Pelajar Mendominasi. (2023, September 1). GoodStats. Retrieved November 23, 2023, from https://goodstats.id/article/angka-kecelakaan-lalu-lintas-terus-meningkat-usia-pelajar-mendominasi-zYuep

Setiyadi, A., Utami, E., & Ariatmanto, D. (2023). Analisa Kemampuan Algoritma YOLOv8 Dalam Deteksi Objek Manusia Dengan Metode Modifikasi Arsitektur. Jurnal Sains Komputer & Informatika (J-SAKTI), 7(2), 891-901.

Setyawan, S. B., Pribadi, W., Arrosida, H., & Nugroho, E. P. (2021). Sistem Deteksi Pengendara Sepeda Motor Tanpa Helm dan Kelebihan Penumpang Pada Dengan Menggunakan YOLOv3. Seminar Nasional Terapan Riset Inovatif (SENTRINOV) ke-VII, 7(1).

Wang, G., Chen, Y., An, P., Hong, H., Hu, J., & Huang, T. (2023, Agustus 15). UAV-YOLOv8: A Small-Object-Detection Model Based on Improved YOLOv8 for UAV Aerial Photography Scenarios. Sensors. https://doi.org/10.3390/s23167190

YOLO Performance Metrics. (2023, November 12). Ultralytics YOLOv8 Docs. Retrieved November 24, 2023, from https://docs.ultralytics.com/guides/yolo-performance-metrics/#introduction

Zhang, Y., Sun, P., Jang, Y., Yu, D., Weng, F., Yuan, Z., Lio, P., Liu, W., & Xinggang Wang. (2022). ByteTrack: Multi-object Tracking by Associating Every Detection Box. In European Conference on Computer Vision, 1-21.

Zhuang, L., Ding, G., Li, C., & Li, D. (2023, July 29). DCF-Yolov8: An Improved Algorithm for Aggregating Low-Level Features to Detect Agricultural Pests and Diseases. Agronomy. https://doi.org/10.3390/ agronomy13082012

DOI: https://doi.org/10.17509/edsence.v5i2.65910
p-ISSN:2685-2489 | e-ISSN:2685-2535