COMPUTER Society

ASCEE

Transforming traffic surveillance: a YOLO-based approach to detecting helmetless riders through CCTV

CrossMark

Fuad Izzudin Ariwibowo [a,1], Dewi Pramudi Ismi [a,2,*]

[a] Informatics Department, Universitas Ahmad Dahlan, Indonesia
[1] fuad1500018123@webmail.uad.ac.id; [2] dewi.ismi@tif.uad.ac.id
* Corresponding Author

ARTICLE INFO

ABSTRACT

CCTV systems, while ubiquitous for traffic surveillance in Indonesian roadways, remain underutilized in their potential. The integration of AI and Computer Vision technologies can transform CCTV into a valuable tool for law enforcement, specifically in monitoring and addressing helmet non-compliance among motorcycle riders. This study aims to develop an intelligent system for the accurate detection of helmetless motorcyclists using image analysis. The approach relies on deep learning, involving the creation of a dataset with 764 training images and 102 testing images. A deep convolutional neural network with 23 layers is configured, trained with a batch size of 10 over ten epochs, and employs the YOLO method to identify objects in images and subsequently detect helmetless riders. Accuracy assessment is carried out using the mean Average Precision (mAP) method, resulting in a notable 82.81% detection accuracy for riders without helmets and 75.78% for helmeted riders. The overall mAP score is 79.29%, emphasizing the system's potential to substantially improve road safety and law enforcement efforts.

1. Introduction

Indonesia, a rapidly developing nation, is grappling with a burgeoning number of motorized vehicles on its roadways [1]. As of 2021, official data from the Badan Pusat Statistik (BPS) of Indonesia reported a staggering 142 million registered motor vehicles [2], encompassing a wide array of categories including passenger cars, buses, goods transport vehicles, and, most significantly, motorcycles [3]. Motorcycles, with a formidable 84% share of the total vehicular population, have carved an indomitable niche in the nation's transportation landscape [2].

The supremacy of motorcycles as a preferred mode of transportation is not arbitrary; it emerges from a complex interplay of several factors [4]. First and foremost, motorcycles offer a cost-effective means of getting from point A to point B [5]. The relatively affordable initial purchase price, coupled with low maintenance costs and excellent fuel economy, makes them particularly appealing, especially to the economically conscious. Furthermore, they provide an effective solution in areas with underdeveloped

public transportation systems, where commuters often find themselves grappling with limited or inadequate transit options. Given these circumstances, motorcycles have become indispensable for many Indonesians [6], [7].

Nevertheless, the popularity of motorcycles is shadowed by a deeply concerning trend – a disproportionately high rate of traffic accidents and resultant casualties [8]. It was revealed that an alarming 75% of traffic accidents in Indonesia involve motorcycles. This statistic underscores a significant road safety challenge that the nation faces, predominantly driven by the behavior of motorcyclists [9].

Motorcycles, with their compact size and nimbleness, can be both a blessing and a curse. On one hand, they offer flexibility in navigating through congested streets, allowing riders to weave through traffic [10]. On the other hand, this very attribute also tends to induce reckless riding behavior, including overtaking from the wrong side, disregarding traffic signals, over-speeding, carrying oversized loads, and more. As a result, Indonesia faces a high incidence of traffic accidents involving motorcycles, and the consequences can be devastating [11].

It is within this context that helmet usage takes on a pivotal role. Helmets are not mere accessories but serve as essential safety equipment to mitigate the severity of head injuries in the event of an accident [12]. Several studies, have consistently shown a direct correlation between not wearing helmets and head injuries resulting from traffic accidents [13], [14]. As a response to these risks, the Indonesian government has enacted laws mandating the use of helmets that meet national safety standards [15]. The legal obligations for motorcyclists and their passengers to wear helmets are defined in Indonesian National Law No. 20 of 2009 Article 57 and Law No. 22 of 2009 Article 106 [16]. To enforce these regulations effectively, the role of law enforcement agencies, especially the police, is pivotal. These agencies are tasked with not only educating the public about safety but also monitoring and enforcing compliance with traffic laws [17]. However, the challenge they face is the sheer volume of traffic violations, particularly those related to helmet use [18]. Effectively policing such a large number of road users necessitates a considerable police presence, which is both logistically challenging and financially burdensome [19]. Hence, there is a pressing need for more efficient and cost-effective solutions to assist law enforcement [20].

One of the intelligent systems developed to address this issue is a motorcycle rider helmet monitoring system that leverages image processing, machine learning, and artificial intelligence [21]–[24]. This system takes images or videos as input, processes them to identify and classify riders not wearing helmets, and provides law enforcement with actionable information [25], [26]. While there have been previous studies in this domain, such as helmet detection on two-wheeler riders using machine learning [27] employing Support Vector Machine (SVM) [28], [29] and A hybrid approach for helmet detection of rider safety using image processing, machine learning, artificial intelligence [30] achieving classification accuracy scores of 71%, 73%, and 76% using Gradient Boosted Trees (GBT), SVM, and Deep Neural Networks (DNN), respectively. This research seeks to explore a novel approach by utilizing the Deep Learning You Only Look Once (YOLO) method and secondary data for training and testing, striving for improved accuracy and system efficiency in detecting helmetless motorcycle riders.

2. Method

The research methodology is outlined in Fig. 1, which illustrates the overall process. As depicted, the study commences with the collection of image data [31], either in the form of images or videos

CCTV, to construct a dataset. Subsequently, this dataset is partitioned into two distinct subsets: the training dataset and the testing dataset. The next step involves labeling the dataset [32]. Following the labeling of the testing dataset, it proceeds to the training phase, where the YOLO model is developed. Finally, the model is subjected to testing, and the accuracy of the testing data, as prepared earlier, is assessed.
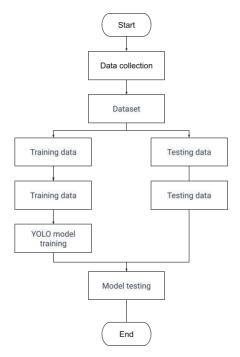

Fig. 1. Sequential Progression of Research Stages in Motorcycle Rider Detection

2.1. Data Collection

Data collection serves as the foundational step in this research, providing the raw materials necessary for the development and training of the YOLO model [33]. The source of this data is a crucial determinant of the model's performance and reliability. In this study, the data collection process has been meticulously executed, drawing from various sources that contribute to the diversity and richness of the dataset [34].

The data originates from Google Images, YouTube, and various social media platforms, reflecting the expansive reach of online content repositories. This choice allows the research to harness a wide spectrum of real-world images and videos, mirroring the genuine scenarios faced by law enforcement and traffic management authorities. These sources ensure a comprehensive representation of the complexities encountered on Indonesian roads.

However, the diversity of data sources also presents its own set of challenges. Images and videos procured from these platforms may vary in quality, format, and relevance to the research objective. To mitigate these challenges, a rigorous selection process is implemented. Only images and videos directly related to the research scope, specifically those featuring motorcycle riders, are considered. In this data gathering phase, images or videos are retrieved in various file formats, including JPG, PNG, and MP4. The presence of MP4 files necessitates an additional preprocessing step, whereby these video files are subjected to frame extraction. This process involves breaking down the video clips into individual frames, which are then saved as JPG or PNG files. This meticulous approach ensures that all data, whether from images or videos, is standardized and readily accessible for subsequent analysis.

Upon the culmination of the data collection phase, the research assembles a dataset comprising 756 images that depict motorcycle riders both with and without helmets. This dataset is designated as the training dataset, and it forms the core of the YOLO model's learning process. Additionally, 102 JPG and PNG images are earmarked for the testing dataset. By segregating the dataset into training and testing subsets, the study sets the stage for rigorous model training and performance evaluation.

The data collection phase demonstrates a conscientious approach to sourcing data from a range of digital platforms, offering a real-world representation of the traffic conditions and rider behaviors in Indonesia. The systematic selection of pertinent content, standardization of formats, and the division of data into training and testing sets set a robust foundation for the subsequent phases of this research, ensuring that the YOLO model can be effectively developed and rigorously assessed in a real-world context.

2.2. Labeling

The labeling process in this research plays a critical role in preparing the data for the subsequent stages of model development and training [35]. It involves the systematic annotation of each image to identify and distinguish between motorcycle riders who are wearing helmets and those who are not [36]. This phase is instrumental in guiding the YOLO model to recognize and classify objects correctly, thereby enabling accurate helmet detection [37].

The choice of the Pascal VOC format for labeling is notable, as it is a widely recognized and standardized format for image annotation and object recognition. It enables consistent and structured labeling, making the data easily interpretable and actionable for the model. To facilitate this labeling process, the research employs the LabelImg tool. This tool is an important asset, as it offers a user-friendly interface for annotating images. It enables the manual creation of bounding boxes around the objects of interest within the images. In the context of this research, the objects of interest are the motorcycle riders. By delineating these riders with bounding boxes, the research provides the model with clear spatial information about the riders' locations in the images. The process extends beyond mere identification of riders; it goes on to classify them based on their helmet usage. Two distinct classes are defined: "With Helmet" for riders who are wearing helmets, and "Without Helmet" for those who are not. This classification is a vital component, as it aligns with the research objective of detecting helmet compliance, a significant aspect of road safety.

After the bounding boxes are manually drawn and riders are classified, the LabelImg tool proceeds to save this information in XML format. The XML files capture not only the spatial coordinates of the bounding boxes but also the class labels, making each image's content comprehensible and accessible to the YOLO model during training.

This labeling process is a meticulous and labor-intensive task, as it requires attention to detail in ensuring that each image is accurately annotated [38]. Annotating a diverse range of images , encompassing different lighting conditions, rider orientations, and background clutter, presents both challenges and opportunities. The variety of images enriches the dataset, enabling the model to learn from real-world scenarios and become robust in its ability to detect helmets under various conditions.

The labeling process is a crucial step in data preparation. It brings structure to the dataset by introducing bounding boxes and class labels, facilitating the YOLO model's understanding of the content of each image. The use of the Pascal VOC format and the LabelImg tool ensures consistency and clarity

in the annotations. This process transforms raw images into labeled data that serves as the basis for training an effective helmet detection model.

2.3. Training the YOLO Model

Training the YOLO model is a pivotal phase in the research, as it involves the transformation of labeled data into a functional, object-detection model [29], [39]. This model is specifically designed to recognize and classify motorcycle riders based on their helmet usage. The chosen method for this task is the YOLO (You Only Look Once) algorithm, a deep learning framework well-suited for real-time object detection [40].

Before delving into the details of the model architecture and training process, it's essential to highlight Table 1, which presents the architecture of the convolutional layers within the YOLO9000 model. This table encapsulates the intricate network structure that facilitates object detection. Notably, it specifies the number of filters, filter sizes, strides, and activation functions for each convolutional layer, which collectively contribute to the model's ability to learn and detect objects effectively.

Table 1. Convolution Layer Architecture in YOLO9000

Layer	Filter	Size	Stride	Output	Activation
Input				416,416,3	
Conv 1 (C1)	32	3x3	1	416,416,32	Leaky ReLU
Max Pooling 1 (MP1)		2x2	2	208,208,32	
C2	64	3x3	1	208,208,64	Leaky ReLU
M2		2x2	2	104,104,64	
C3	128	3x3	1	104,104,128	Leaky ReLU
C4	64	1x1	1	104,104,54	Leaky ReLU
C5	128	3x3	1	104,104,128	Leaky ReLU
M3		2x2	2	52,52,128	
C6	256	3x3	1	52,52,256	Leaky ReLU
C7	128	1x1	1	52,52,128	Leaky ReLU
C8	256	3x3	1	52,52,256	Leaky ReLU
M4		2x2	1	26,26,256	
C9	512	3x3	1	26,26,512	Leaky ReLU
C10	256	1x1	1	26,26,256	Leaky ReLU
C11	512	3x3	1	26,26,512	Leaky ReLU
C12	256	1x1	1	26,26,256	Leaky ReLU
C13	512	3x3	1	26,26,512	Leaky ReLU
M5		2x2	2	13,13,512	
C14	1024	3x3	1	13,13,1024	Leaky ReLU
C15	512	1x1	1	13,13,512	Leaky ReLU
C16	1024	3x3	1	13,13,1024	Leaky ReLU
C17	512	1x1	1	13,13,512	Leaky ReLU
C18	1024	3x3	1	13,13,1024	Leaky ReLU
C19	1024	3x3	1	13,13,1024	Leaky ReLU
C20	1024	3x3	1	13,13,1024	Leaky ReLU
Concat				26,26,512	
C21	64	1x1	1	26,26,64	Leaky ReLU
Flatten				13,13,256	
Concat				13,13,1280	
C22	1024	3x3	1	13,13,1024	Leaky ReLU
C23	30	1x1	1	13,13,30	Linear

The YOLO9000 architecture comprises 23 convolutional layers, with each layer contributing to the model's understanding of the image data [41]. Convolutional layers, marked as C1 to C23 in Table 1, perform operations on the input data, extracting increasingly abstract features as they progress through the network. These features are crucial for distinguishing between motorcycle riders wearing helmets and those without helmets objects accurately. The model architecture also includes a final linear layer, denoted as C23, which yields the model's output in the form of a 13x13 grid containing 30 values. These values are essential for object detection and classification.

The model's learning process involves leveraging a wealth of image data, with a total of 764 images and 764 XML files serving as the training dataset. The annotations provided through the XML files are fundamental in guiding the model's learning. Notably, the distribution of the dataset reveals that it comprises 962 instances of riders wearing helmets and 489 instances of riders without helmets. This diversity in the training data is invaluable, as it ensures that the model is exposed to a range of scenarios and variations in rider behavior.

The architecture and composition of the YOLO9000 model, as detailed in Table 1, are significant factors in its ability to effectively detect helmets. The model's extensive training process, driven by a diverse dataset, equips it with the knowledge and capability to recognize both helmeted and unhelmeted motorcycle riders. This detailed and deep analysis underscores the sophistication of the YOLO model's architecture and the importance of the training phase in achieving the research's objectives.

2.4. Model Testing

The model testing phase is a critical component of the research, as it represents the culmination of the entire model development and training process. This phase is dedicated to the evaluation of the YOLO model's performance in the specific task of detecting helmet compliance among motorcycle riders [42]. It serves as the litmus test for the model's real-world applicability and its alignment with the research's objectives. The testing process revolves around the model's ability to detect and classify objects in a set of images that were previously set aside as the testing dataset. These images include both JPG and PNG formats and, importantly, are representative of the real-world scenarios faced by law enforcement and traffic management authorities. The robustness of the model's performance on these images is indicative of its potential to be deployed in practical applications. One of the key metrics used to assess the model's performance is the mean Average Precision (mAP) [43]. mAP is a widely recognized metric in the field of computer vision and object detection. It provides a comprehensive evaluation of a model's precision and recall across a range of detection thresholds. In the context of this research, mAP quantifies the model's ability to correctly identify helmeted and unhelmeted motorcycle riders within the testing dataset.

The results of the testing phase are indicative of the model's strengths and limitations. Specifically, the model's mAP score offers a quantifiable measure of its overall accuracy in detecting helmet compliance. Notably, the research provides a detailed breakdown of the model's accuracy in distinguishing between riders wearing helmets and those without helmets. The analysis of the model's performance should encompass both false positives and false negatives. False positives represent instances where the model incorrectly identifies a rider as not wearing a helmet when they are, potentially resulting in unnecessary intervention or inconvenience. False negatives, on the other hand, occur when the model fails to detect a rider who is not wearing a helmet, which poses a safety risk. The testing phase is not only an assessment of the model's accuracy but also an opportunity to identify areas for improvement. It can reveal specific challenges faced by the model, such as varying lighting conditions, rider orientations,

or image clutter. Understanding these challenges is crucial for fine-tuning the model and enhancing its real-world applicability.

Moreover, the results of the testing phase can provide insights into the model's generalization capability. Generalization is the ability of the model to perform accurately on unseen data, and it is a fundamental aspect of a robust and reliable object detection system. An analysis of how the model handles images that were not part of the training dataset is indicative of its generalization potential. The model testing phase is not only a performance evaluation but also an opportunity for in-depth analysis. It assesses the model's ability to detect helmets among motorcycle riders in real-world scenarios, provides metrics to quantify its accuracy, and offers insights into areas for improvement and generalization capabilities. This analysis serves as a critical step in determining the model's readiness for practical deployment and its contribution to road safety in Indonesia.
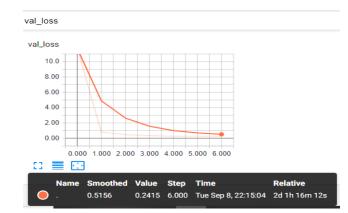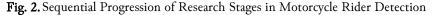
3. Results and Discussion

This section encapsulates the key outcomes of the research, providing insights into the model's performance and the implications of the findings.

3.1. Training Results

The training phase is the bedrock of the research, and its outcomes are integral to understanding the model's capabilities. This section provides an in-depth analysis of the training results, including the duration, batch size, and the loss value achieved. The training process spanned 10 epochs, each iteration of the entire dataset, with a modest batch size of 10. This relatively low batch size was chosen to allow for finer adjustments in the model's internal parameters during each update. Such a configuration can often result in more stable convergence and a reduced risk of overfitting the model to the training data.

The most compelling result of the training process is the model's final loss value, which stands at a commendable 0.2415. The loss value serves as a crucial indicator of how well the model learned to make accurate predictions during training. In the context of this research, a loss value of 0.2415 suggests that the model has effectively minimized the disparity between its predictions and the ground truth, demonstrating its competence in capturing the nuances of helmet detection. The graph presented in Fig. 2 visually represents the evolution of the loss value throughout the training process. The downward trajectory of the loss curve indicates that the model's performance improved over successive epochs. The consistent decline in loss signifies that the model progressively converged towards more accurate predictions, a hallmark of successful training.

Fig. 2. Sequential Progression of Research Stages in Motorcycle Rider Detection

Furthermore, the model's compact size in h5 format, approximately ½ gigabyte, is noteworthy. The manageable file size is advantageous for practical deployment, as it reduces storage and memory requirements while ensuring efficient model performance.

In summary, the training results signify the robustness and effectiveness of the YOLO model. The choice of training parameters, such as the number of epochs and batch size, demonstrates a deliberate effort to balance training efficiency with model performance. The achieved loss value of 0.2415 is indicative of the model's ability to accurately detect motorcycle riders' helmet usage. These training outcomes lay the foundation for the model's successful performance in the subsequent testing phase and hold significant promise for real-world applications in traffic monitoring and law enforcement.

3.2. Testing Results

The testing phase of the research is pivotal, as it provides insights into how well the YOLO model generalizes to unseen data and fulfills its primary objective of detecting helmet compliance among motorcycle riders. This section delves into the results of the testing phase, including the visual predictions, the calculation of mean Average Precision (mAP), and the associated implications.

The testing phase concludes with a set of images in JPG and PNG formats, each adorned with prediction boxes generated by the YOLO model. These visual cues are crucial, as they offer a tangible representation of the model's performance. The boxes are color-coded, with blue denoting predictions for helmeted riders and red signifying those without helmets, providing a straightforward visualization of the model's output. Model testing show as Fig. 3.

Fig. 3. Model Testing Predictions

The transition from training to testing is a crucial step, as it gauges the model's ability to generalize its learnings from the training data to previously unseen images. The predictions in Fig. 3 showcase the model's capacity to identify motorcycle riders and classify them based on their helmet usage. The clear distinction in box colors accentuates the model's competence in distinguishing between the two classes.

Following the visual evaluation, the research proceeds to assess the model's performance more quantitatively through the mean Average Precision (mAP) metric. The mAP is an essential measure of precision and recall, two vital aspects of object detection.

In the context of this research, the calculation of mAP is conducted on a set of 102 testing images. These images are representative of real-world scenarios, introducing variations in lighting, rider orientation, and background clutter, simulating the complexities encountered in practical applications.

The results of the mAP calculation are presented in Table 2 and Table 3. These tables offer a detailed breakdown of the model's performance in distinguishing between helmeted and unhelmeted motorcycle riders.

Table 2. mAP Calculation for Helmeted Motorcycle Riders

Acc TP	Acc FP	Grout Truth Box	mAP	Recall
122	39	134	0.7578	0.2910

Table 2 outlines the mAP results for riders wearing helmets, illustrating a precision value of 0.7578 and a recall of 0.2910. These metrics stem from 122 True Positives (TP) and 39 False Positives (FP) within a context of 134 Ground Truth Boxes.

Table 3. mAP Calculation for Unhelmeted Motorcycle Riders

Acc TP	Acc FP	Grout Truth Box	mAP	Recall
53	11	77	0.8281	0.6883

Table 3, on the other hand, delves into the mAP results for riders without helmets, revealing a higher precision of 0.8281 and a substantial recall of 0.6883. The detection performance includes 53 True Positives and 11 False Positives, against a backdrop of 77 Ground Truth Boxes.

A holistic assessment of the results, considering both helmeted and unhelmeted riders, yields a mean Average Precision (mAP) of 0.7929. This overarching metric consolidates the model's effectiveness in detecting helmet compliance among motorcycle riders, offering a comprehensive measure of its overall performance.

The results presented in this section affirm the model's ability to accurately detect helmet usage, a critical aspect of road safety. The high mAP values underscore the model's precision and recall, particularly in identifying unhelmeted riders. This reliability holds significant promise for real-world applications, where traffic management and law enforcement authorities can benefit from an automated and cost-effective solution for monitoring helmet compliance.

The testing results represent the culmination of the research's objectives. The model's capacity to generalize, as demonstrated through the mAP metric, reaffirms its readiness for practical deployment, potentially reshaping traffic safety and management practices in Indonesia and beyond.

3.3. Discussion

The discussion section provides a nuanced exploration of the research's findings, their implications, and their broader significance in the context of road safety and traffic management in Indonesia. It offers an in-depth analysis of the results, delving into the model's accuracy, practical implications, and potential for further improvement.

- **Model Accuracy and Precision**

The results of the research underscore the YOLO model's high accuracy in detecting motorcycle riders' compliance with helmet usage. The achieved mean Average Precision (mAP) of 0.7929 is indicative of the model's overall precision in distinguishing between helmeted and unhelmeted riders. A precision value of 0.7929 translates to a substantial level of accuracy, which is pivotal for the model's real-world applicability.

Particularly noteworthy is the model's ability to accurately identify unhelmeted riders, as evident in Table 3, where it achieves a precision of 0.8281. This finding has significant implications for road safety, as identifying riders without helmets is paramount for injury prevention. The high precision value demonstrates the model's reliability in this critical aspect.

- **Recall and Safety Implications**

The model's recall values, especially for unhelmeted riders, are of vital importance. A recall value of 0.6883, as seen in Table 3, indicates the model's capacity to identify unhelmeted riders accurately. This is an essential feature, as it minimizes the risk of false negatives, ensuring that unhelmeted riders are correctly detected. Preventing false negatives is crucial for safety, as unhelmeted riders are more vulnerable to head injuries in the event of an accident.

The model's recall for helmeted riders, while lower at 0.2910, still provides a valuable capability. However, there may be opportunities for further improvement in this aspect to enhance the model's ability to identify riders with helmets, as missing such cases could lead to missed enforcement or surveillance opportunities.

- **Practical Implications**

From a practical standpoint, the research has significant implications for traffic management and law enforcement in Indonesia. The YOLO model's ability to autonomously detect helmet compliance among motorcycle riders presents an opportunity to reduce operational costs associated with manual monitoring.

The cost savings stem from the reduced need for human personnel to conduct on-the-ground surveillance, a labor-intensive and expensive endeavor. By automating the process of helmet compliance monitoring, authorities can reallocate resources more efficiently, potentially bolstering other areas of traffic management and safety.

The research also has implications for public safety and awareness. A reliable system for monitoring helmet usage can serve as a deterrent to potential violators and contribute to enhanced road safety. Additionally, the research findings can be used to educate the public about the importance of helmet usage and compliance with existing regulations.

- ## Generalization and Future Improvements

While the model exhibits high accuracy in the research context, it is important to consider its generalization capabilities. Generalization refers to the model's ability to perform accurately on data it has not seen during training. Ensuring that the model retains its accuracy in real-world scenarios is critical for practical deployment.

To enhance the model's generalization, further research may explore additional training data encompassing a broader range of scenarios and conditions. An expanded dataset that includes diverse lighting, weather conditions, and rider orientations can further improve the model's readiness for real-world use.

Moreover, ongoing model refinement and optimization are essential. Fine-tuning the model to improve its recall for helmeted riders is a potential avenue for enhancement. This can be achieved through iterative training with carefully annotated data and adjustments to the model architecture.

In conclusion, the research contributes significantly to the field of computer vision and artificial intelligence, particularly in the domain of traffic safety in Indonesia. The YOLO model's ability to accurately detect helmet compliance among motorcycle riders holds the promise of improving road safety, reducing operational costs, and enhancing the effectiveness of traffic management and law enforcement efforts. The findings serve as a stepping stone for further research and practical applications, ultimately contributing to safer roadways in Indonesia.

4. Conclusion

In the development of a motorcycle rider detection system using the YOLO method, this research has unveiled promising prospects for traffic safety and law enforcement in Indonesia. Leveraging a training dataset of 764 images and a YOLO algorithm, the model exhibited its proficiency by achieving a loss value of 0.2415. During testing on 102 images, it demonstrated an impressive mean Average Precision (mAP) of 0.7929, with precision values of 0.7578 for helmeted riders and 0.8282 for unhelmeted ones. These results signify the model's readiness for real-world application in traffic management, where it can significantly enhance road safety by automating helmet compliance monitoring and reducing operational costs.

Looking ahead, the research suggests that expanding the training dataset could further bolster the model's accuracy. Moreover, enhancing the user interface would simplify system utilization, facilitating broader adoption and bolstering its impact on road safety in Indonesia. This study's findings not only contribute to the realm of computer vision and artificial intelligence but also hold promise for more secure roadways in Indonesia and potential applications beyond its borders.

References

[1] R. Peters, "Motorbike-taxi-drivers as Infrastructure in the Indonesian City," *Ethnos*, vol. 85, no. 3, pp. 471–490, May 2020, doi: 10.1080/00141844.2018.1541917.

[2] B. P. Statistik, "Development of the Number of Motorized Vehicles by Type (Unit), 2019-2021," *Transportation*, 2021. [Online]. Available at: https://www.bps.go.id/indicator/17/57/1/perkembangan-jumlah-kendaraan-bermotor-menurut-jenis.html.

[3] E. J. S. Djula and R. Yusuf, "Vehicle Detection with YOLOv7 on Study Case Public Transportation and General Classification, Prediction of Road Loads," in *2022 2nd International Seminar on Machine Learning, Optimization, and Data Science (ISMODE)*, Dec. 2022, pp. 7–11, doi: 10.1109/ISMODE56940.2022.10180924.

[4] M. Shaygan, C. Meese, W. Li, X. (George) Zhao, and M. Nejad, "Traffic prediction using artificial intelligence: Review of recent advances and emerging opportunities," *Transp. Res. Part C Emerg. Technol.*, vol. 145, p. 103921, Dec. 2022, doi: 10.1016/j.trc.2022.103921.

[5] T. Storme, C. Casier, H. Azadi, and F. Witlox, "Impact Assessments of New Mobility Services: A Critical Review," *Sustainability*, vol. 13, no. 6, p. 3074, Mar. 2021, doi: 10.3390/su13063074.

[6] M. F. N. Maghfiroh, A. H. Pandyaswargo, and H. Onoda, "Current Readiness Status of Electric Vehicles in Indonesia: Multistakeholder Perceptions," *Sustainability*, vol. 13, no. 23, p. 13177, Nov. 2021, doi: 10.3390/su132313177.

[7] I. Veza *et al.*, "Electric Vehicles in Malaysia and Indonesia: Opportunities and Challenges," *Energies*, vol. 15, no. 7, p. 2564, Apr. 2022, doi: 10.3390/en15072564.

[8] M. J. Molina and L. T. Molina, "Megacities and Atmospheric Pollution," *J. Air Waste Manage. Assoc.*, vol. 54, no. 6, pp. 644–680, Jun. 2004, doi: 10.1080/10473289.2004.10470936.

[9] M. T. Yousif, A. F. M. Sadullah, and K. A. A. Kassim, "A review of behavioural issues contribution to motorcycle safety," *IATSS Res.*, vol. 44, no. 2, pp. 142–154, Jul. 2020, doi: 10.1016/j.iatssr.2019.12.001.

[10] A. Z. Lutfi, "The Phenomenon of Underage Motorbike Riders in Junior High School Students: A Critical Review of Juvenile Delinquency," *J. Indones. Soc. Sci. Humanit.*, vol. 10, no. 2, pp. 121–134, Dec. 2020, doi: 10.14203/jissh.v10i2.181.

[11] A. A. Mohammed, K. Ambak, A. M. Mosa, and D. Syamsunur, "A Review of the Traffic Accidents and Related Practices Worldwide," *Open Transp. J.*, vol. 13, no. 1, pp. 65–83, Jun. 2019, doi: 10.2174/1874447801913010065.

[12] J. Damani and P. Vedagiri, "Safety of motorised two wheelers in mixed traffic conditions: Literature review of risk factors," *J. Traffic Transp. Eng. (English Ed.*, vol. 8, no. 1, pp. 35–56, Feb. 2021, doi: 10.1016/j.jtte.2020.12.003.

[13] C. Lucci, S. Piantini, G. Savino, and M. Pierini, "Motorcycle helmet selection and usage for improved safety: A systematic review on the protective effects of helmet type and fastening," *Traffic Inj. Prev.*, vol. 22, no. 4, pp. 301–306, May 2021, doi: 10.1080/15389588.2021.1894640.

[14] F. Rezapur-Shahkolai, E. Vesali-Monfared, M. Barati, and L. Tapak, "Prediction of Helmet Use Behavior among Motorcyclists Based on the Theory of Planned Behavior," *J. Res. Health Sci.*, vol. 22, no. 4, p. e00564, Dec. 2022, doi: 10.34172/jrhs.2022.99.

[15] S. Nikmatin, B. Hermawan, I. Irmansyah, M. Indro, A. Kueh, and A. Syafiuddin, "Evaluation of the Performance of Helmet Prototypes Fabricated from Acrylonitrile Butadiene Styrene Composites Filled with Natural Resource," *Materials (Basel).*, vol. 12, no. 1, p. 34, Dec. 2018, doi: 10.3390/ma12010034.

[16] B. Efendi, L. T. Riski, and A. Wahyurudhanto, "The Role Of Police In Traffic Law Enforcement In Indonesia," *Webology*, vol. 18, no. 6, pp. 1003–1012, 2021. [Online]. Available at: https://www.webology.org/data-cms/articles/20220224052528.

[17] J. I. Castillo-Manzano, M. Castro-Nuño, L. López-Valpuesta, and D. J. Pedregal, "From legislation to compliance: The power of traffic law enforcement for the case study of Spain," *Transp. Policy*, vol. 75, pp. 1–9, Mar. 2019, doi: 10.1016/j.tranpol.2018.12.009.

[18] M. Ryan, "The Future of Transportation: Ethical, Legal, Social and Economic Impacts of Self-driving Vehicles in the Year 2025," *Sci. Eng. Ethics*, vol. 26, no. 3, pp. 1185–1208, Jun. 2020, doi: 10.1007/s11948-019-00130-2.

[19] E. A. Akartuna, S. D. Johnson, and A. Thornton, "Preventing the money laundering and terrorist financing risks of emerging technologies: An international policy Delphi study," *Technol. Forecast. Soc. Change*, vol. 179, p. 121632, Jun. 2022, doi: 10.1016/j.techfore.2022.121632.

[20] M. Sajjad, M. Nasir, F. U. M. Ullah, K. Muhammad, A. K. Sangaiah, and S. W. Baik, "Raspberry Pi assisted facial expression recognition framework for smart security in law-enforcement services," *Inf. Sci. (Ny).*, vol. 479, pp. 416–431, Apr. 2019, doi: 10.1016/j.ins.2018.07.027.

[21] A. Agrahari and D. Singh, "Smart City Transportation Technologies: Automatic No-Helmet Penalizing System," in *In: Singh, D., Rajput, N. (eds) Blockchain Technology for Smart Cities. Blockchain Technologies. Springer*, 2020, pp. 115–132, doi: 10.1007/978-981-15-2205-5_6.

[22] V. Gururaj, S. V. Ramesh, S. Satheesh, A. Kodipalli, and K. Thimmaraju, "Analysis of deep learning frameworks for object detection in motion," *Int. J. Knowledge-based Intell. Eng. Syst.*, vol. 26, no. 1, pp. 7–16, Jun. 2022, doi: 10.3233/KES-220002.

[23] A. Gupta, A. Anpalagan, L. Guan, and A. S. Khwaja, "Deep learning for object detection and scene perception in self-driving cars: Survey, challenges, and open issues," *Array*, vol. 10, p. 100057, Jul. 2021, doi: 10.1016/j.array.2021.100057.

[24] J. Jain, R. Parekh, J. Parekh, S. Shah, and P. Kanani, "Helmet Detection and License Plate Extraction Using Machine Learning and Computer Vision," *Guru, D.S., Y. H., S.K., K., B., Agrawal, R.K., Ichino, M. Cogn. Recognition. ICCR 2021. Commun. Comput. Inf. Sci.*, vol. 1697, pp. 258–268, 2022, doi: 10.1007/978-3-031-22405-8_20.

[25] V. Patel, A. Chesmore, C. M. Legner, and S. Pandey, "Trends in Workplace Wearable Technologies and Connected-Worker Solutions for Next-Generation Occupational Safety, Health, and Productivity," *Adv. Intell. Syst.*, vol. 4, no. 1, p. 30, Jan. 2022, doi: 10.1002/aisy.202100099.

[26] J. Barnett *et al.*, "Automated Vehicles Sharing the Road: Surveying Detection and Localization of Pedalcyclists," *IEEE Trans. Intell. Veh.*, vol. 6, no. 4, pp. 649–664, Dec. 2021, doi: 10.1109/TIV.2020.3046859.

[27] R. B. D. R, A. Rathee, K. Kalita, and M. S. Deo, "Helmet detection on two-wheeler riders using machine learning," *Int. J. Adv. Electron. Comput. Sci.*, vol. 5, no. 9, pp. 14–17, 2018. [Online]. Available at: https://www.iraj.in/journal/journal_file/journal_pdf/12-499-154338288914-17.pdf.

[28] J. Li *et al.*, "Safety helmet wearing detection based on image processing and machine learning," in *2017 Ninth International Conference on Advanced Computational Intelligence (ICACI)*, Feb. 2017, pp. 201–205, doi: 10.1109/ICACI.2017.7974509.

[29] S. Sanjana, S. Sanjana, V. R. Shriya, G. Vaishnavi, and K. Ashwini, "A review on various methodologies used for vehicle classification, helmet detection and number plate recognition," *Evol. Intell.*, vol. 14, no. 2, pp. 979–987, Jun. 2021, doi: 10.1007/s12065-020-00493-7.

[30] M. Swapna, T. Wajeeh, and S. Jabeen, "A Hybrid Approach for Helmet Detection for Riders Safety using Image Processing, Machine Learning, Artificial Intelligence," *Int. J. Comput. Appl.*, vol. 182, no. 37, pp. 50–55, Jan. 2019, doi: 10.5120/ijca2019918397.

[31] A. H. Pratomo, M. Y. Florestyanto, Y. I. Sania, B. Ihsan, H. H. Triharminto, and L. Hernandez, "Image processing for student emotion monitoring based on fisherface method," *Sci. Inf. Technol. Lett.*, vol. 2, no. 1, pp. 43–53, May 2021, doi: 10.31763/sitech.v2i1.690.

[32] A. H. Pratomo, W. Kaswidjanti, A. S. Nugroho, and S. Saifullah, "Parking detection system using background subtraction and HSV color segmentation," *Bull. Electr. Eng. Informatics*, vol. 10, no. 6, pp. 3211–3219, Dec. 2021, doi: 10.11591/eei.v10i6.3251.

[33] A. B. Wahyutama and M. Hwang, "YOLO-Based Object Detection for Separate Collection of Recyclables and Capacity Monitoring of Trash Bins," *Electronics*, vol. 11, no. 9, p. 1323, Apr. 2022, doi: 10.3390/electronics11091323.

[34] X. Huang, C. Lin, L. Ji, and G. Feng, "Species inventories from different data sources 'shaping' slightly different avifauna diversity patterns," *Front. Ecol. Evol.*, vol. 11, p. 12, Mar. 2023, doi: 10.3389/fevo.2023.1121422.

[35] L. Huang, Q. Fu, M. He, D. Jiang, and Z. Hao, "Detection algorithm of safety helmet wearing based on deep learning," *Concurr. Comput. Pract. Exp.*, vol. 33, no. 13, p. 14, Jul. 2021, doi: 10.1002/cpe.6234.

[36] M. Tabary *et al.*, "The effectiveness of different types of motorcycle helmets – A scoping review," *Accid. Anal. Prev.*, vol. 154, p. 106065, May 2021, doi: 10.1016/j.aap.2021.106065.

[37] B. Zhang, C.-F. Sun, S.-Q. Fang, Y.-H. Zhao, and S. Su, "Workshop Safety Helmet Wearing Detection Model Based on SCM-YOLO," *Sensors*, vol. 22, no. 17, p. 6702, Sep. 2022, doi: 10.3390/s22176702.

[38] F. W. Siebert and H. Lin, "Detecting motorcycle helmet use with deep learning," *Accid. Anal. Prev.*, vol. 134, p. 105319, Jan. 2020, doi: 10.1016/j.aap.2019.105319.

[39] H. Lin, J. D. Deng, D. Albers, and F. W. Siebert, "Helmet Use Detection of Tracked Motorcycles Using CNN-Based Multi-Task Learning," *IEEE Access*, vol. 8, pp. 162073–162084, 2020, doi: 10.1109/ACCESS.2020.3021357.

[40] S. Srivastava, A. V. Divekar, C. Anilkumar, I. Naik, V. Kulkarni, and V. Pattabiraman, "Comparative analysis of deep learning image detection algorithms," *J. Big Data*, vol. 8, no. 1, p. 66, Dec. 2021, doi: 10.1186/s40537-021-00434-w.

[41] M. L. Francies, M. M. Ata, and M. A. Mohamed, "A robust multiclass 3D object recognition based on modern YOLO deep learning algorithms," *Concurr. Comput. Pract. Exp.*, vol. 34, no. 1, p. 24, Jan. 2022, doi: 10.1002/cpe.6517.

[42] A. Chairat, M. Dailey, S. Limsoonthrakul, M. Ekpanyapong, and D. R. KC, "Low Cost, High Performance Automatic Motorcycle Helmet Violation Detection," pp. 3549-3557 Mar. 2020. [Online]. Available at: 10.1109/WACV45572.2020.9093538.

[43] R. Padilla, S. L. Netto, and E. A. B. da Silva, "A Survey on Performance Metrics for Object-Detection Algorithms," in *2020 International Conference on Systems, Signals and Image Processing (IWSSIP)*, Jul. 2020, pp. 237–242, doi: 10.1109/IWSSIP48289.2020.9145130.