Jurnal E-Komtek
Vol. 8, No. 2 (2024) pp. 428-437
https://jurnal.politeknik-kebumen.ac.id/index.php/E-KOMTEK
p-ISSN : 2580-3719 e-ISSN : 2622-3066

Application of Deep Learning YOLO in IoT System for Personal Protective Equipment Detection

Waluyo Nugroho[1], Rifdah Zahabiyah[2], Afianto[3], Mada Jimmy Fonda Arifianto[4]
[1,3,4] Mechatronics Department, Politeknik Astra, Indonesia, 17530
[2] Logistics Engineering Technology, Politeknik Astra, Indonesia, 17530

nugroho.research@gmail.com

https://doi.org/10.37339/e-komtek.v8i2.2187

Published by Politeknik Piksi Ganesha Indonesia

Artikel Info
Submitted:
10-12-2024
Revised:
23-12-2024
Accepted:
24-12-2024
Online first :
31-12-2024

Abstract

The use of Personal Protective Equipment (PPE) is a critical step in ensuring worker safety in various sectors, including industry, construction, and health. However, violations in using PPE often occur, which can increase the risk of work accidents. This study aims to develop a deep learning-based PPE detection system using the YOLOv8 algorithm. This method was chosen because of its superior ability to detect objects in real time with high accuracy. The training data consists of various images of workers in different work environments, label to recognize types of PPE such as helmets, masks, and safety vests. The developed system was tested on a test dataset to evaluate model performance based on metrics such as confusion matrix, inference speed, and detection error rate. The experimental results show that the YOLOv8 model can detect PPE with an accuracy level of up to 95%. The implementation of this system is expected to be an effective solution in increasing compliance with the use of PPE and preventing work accidents.

Keywords: *Personal Protective Equipment, Object Detection, Deep Learning, Occupational Safety, YOLOv8*

Abstrak

Penggunaan Alat Pelindung Diri (APD) merupakan salah satu langkah kritis untuk memastikan keselamatan pekerja di berbagai sektor, termasuk industri, konstruksi, dan kesehatan. Namun, pelanggaran dalam penggunaan APD sering kali terjadi, yang dapat meningkatkan risiko kecelakaan kerja. Penelitian ini bertujuan untuk mengembangkan sistem deteksi APD berbasis deep learning menggunakan algoritma YOLOv8. Metode ini dipilih karena kemampuannya yang unggul dalam mendeteksi objek secara real-time dengan akurasi tinggi. Data pelatihan terdiri dari berbagai gambar pekerja dalam lingkungan kerja yang berbeda, dilabeli untuk mengenali jenis APD seperti helm, masker, dan rompi keselamatan. Sistem yang dikembangkan diuji pada dataset uji untuk mengevaluasi kinerja model berdasarkan metrik seperti confusion matrix, kecepatan inferensi, dan tingkat kesalahan deteksi. Hasil eksperimen menunjukkan bahwa model YOLOv8 mampu mendeteksi APD dengan tingkat akurasi mencapai 95%. Implementasi sistem ini diharapkan dapat menjadi solusi efektif dalam meningkatkan kepatuhan penggunaan APD dan mencegah kecelakaan kerja.

Keywords: *Alat Pelindung Diri, Deteksi Objek, Deep Learning, Keselamatan Kerja, YOLOv8*

1. Introduction

Occupational safety and health (OHS) is a fundamental aspect in various industrial sectors, especially in high-risk work environments such as construction, manufacturing, and mining **[1]**. The use of Personal Protective Equipment (PPE) is one of the main ways to protect workers from potential hazards that can result in serious injury or even death. However, compliance with the use of PPE is often ignored, either due to worker negligence or lack of effective supervision . In this context, the development of technology that can automatically detect the use of PPE is a potential solution to improve real-time OHS supervision.

Computer Vision technology, combined with Deep Learning methods, has shown significant success in various applications such as object detection, facial recognition, and video analysis. One of the most popular object detection algorithms is YOLO (You Only Look Once), which is known for its speed and accuracy **[2]** . The latest version, YOLOv8, offers improved performance in more complex object detection and efficient data processing **[3]**. This technology has great potential to be applied in PPE usage detection, which involves recognizing attributes such as helmets, masks, gloves, and protective shoes from images or videos in real time.

Although this technology is promising, implementing PPE detection faces several challenges. Variations in lighting conditions, camera angles, and worker attributes such as PPE color or body size can affect the accuracy of the detection model. In addition, the need to process data quickly and on a large scale, especially in dynamic work environments, is a technical challenge that requires an efficient computational approach **[4]**. Therefore, this study leverages the power of YOLOv8 in detecting objects with high accuracy to overcome these challenges **[5]**. This study aims to develop a Computer Vision-based PPE detection system using the YOLOv8 Deep Learning algorithm. This system is expected to detect the use of PPE automatically and in real time while identifying compliance violations. Thus, this study not only contributes to the development of AI-based technology in the field of OHS but also provides a practical solution that can be integrated with existing occupational safety monitoring systems.

2. Method

The Personal Protective Equipment (PPE) usage detection project in the production area consists of five main stages: data collection, data labeling, model training, system implementation, and performance evaluation and testing [6]. A detailed explanation of each stage is as follows:

a. Data Collection

The first stage is the collection of datasets in the form of images of workers in the production area. This process involves taking images from various sources, both from CCTV footage, cameras installed in the production area, and direct capture using smartphones or other devices. The data collected includes various environmental conditions, such as bright lighting, dim lighting, and different shooting angles. This aims to ensure data diversity so that the model can recognize patterns of PPE use in various work situations.

In addition, the dataset is also designed to include variations in the types of PPE, such as helmets, safety vests, masks, and protective shoes, as well as situations where workers do not use complete PPE [7]. This variation is important for training the model to be able to effectively distinguish between appropriate and inappropriate use of PPE. Data collection is carried out with attention to worker privacy and following applicable ethical guidelines.

Figure 1. Photo Dataset Using PPE and Not Using PPE

b. YOLO Algorithm

Our model structure is made up of a backbone, neck, and head, as illustrated in **Figure 2**. In the upcoming subsections, we will discuss the design principles of each component of the model architecture, along with the modules associated with each part.

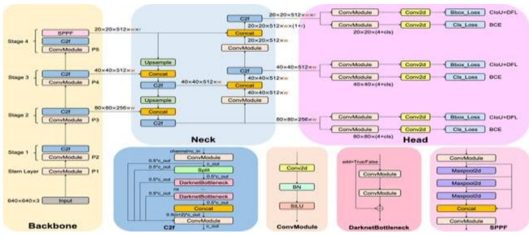

Figure 2. YOLOv8 Model Architecture

Backbone is the main component in YOLOv8 architecture responsible for extracting features from input images. In YOLOv8, the backbone is designed using modern efficient convolutional neural networks (CNN) architectures, such as CSPDarknet or other variants, depending on the model configuration. The backbone processes the image through a series of convolutional layers to produce high-level feature representations, such as edges, textures, and complex patterns [5]. These features are crucial for understanding the image content in subsequent stages.

The backbone structure is designed to strike a balance between speed and accuracy. YOLOv8 uses techniques such as residual connections and spatial pyramid pooling (SPP) to improve efficiency. Residual connections help mitigate the vanishing gradient problem in deep networks, while SPP allows for multi-scale information retrieval without performance loss. The result is a rich feature representation ready to be passed on to the next stage, the neck.

Neck

Neck is a component that connects the backbone with the head in YOLOv8. Its main function is to combine and enrich the features extracted by the backbone to make them more informative and relevant for object detection. The neck often uses a feature pyramid network (FPN) or path aggregation network (PAN) approach to integrate information from different resolution levels. This approach ensures that detailed information from low levels and global information from high levels are optimally utilized [8].

The neck structure is designed to address the challenges of object detection at various scales, from small to large objects. By utilizing techniques such as top-down and bottom-up pathways, YOLOv8 can increase the model's sensitivity to small objects without sacrificing the detection of large objects [9]. The neck plays an important role in ensuring that the model can provide precise predictions in various image conditions [10].

Head is the final component in YOLOv8 that is responsible for generating the final predictions in the form of bounding boxes, confidence scores, and object classifications. The head receives features from the neck and applies them to the output layer to detect the position and type of objects [11]. In YOLOv8, the head is designed to process multi-scale information, ensuring accurate object detection at various sizes and positions.

YOLOv8 utilizes anchor-free or anchor-based techniques, depending on the setting, to predict the bounding box. The model also uses activation functions such as sigmoid to calculate class probabilities. The head ensures that each candidate object has a confidence score that reflects the probability of its existence and classification, resulting in predictions that are ready for further processing.

Loss is a function used to calculate the error between the model prediction and the actual label during the training process. In YOLOv8, loss consists of several components, including bounding box regression loss, objectness loss, and classification loss. Bounding box regression loss is used to measure how well the model predicts the location of an object, while objectness loss measures how confident the model is about the existence of an object in an area. A good loss function is essential to train the model efficiently and guide the parameters towards the optimal configuration. YOLOv8 uses CIoU or GIoU-based loss to improve the accuracy of bounding box prediction, and cross-entropy or binary focal loss for classification and objectness. With effective loss optimization, YOLOv8 can achieve high performance in detecting objects in various environments **[10]**.

c. Data Labeling

After the data is collected, a manual labeling process is carried out to mark the presence and type of PPE in each image. This labeling is done using the Roboflow platform, which provides an intuitive user interface for annotating data. In this process, each PPE object is given a bounding box and categorized into one of two classes: Safety, for workers who use complete PPE according to safety standards, and Not Safety, for workers who do not use PPE or use PPE incompletely.

The labeling process is an important step because this labeled data is the basis for training the detection model. Accurate and consistent labeling will improve the model's ability to recognize patterns and make accurate predictions **[12]**. To ensure quality, a re-check of the labeled dataset is carried out to avoid annotation errors, such as double labeling or mismatches between categories and objects in the image.

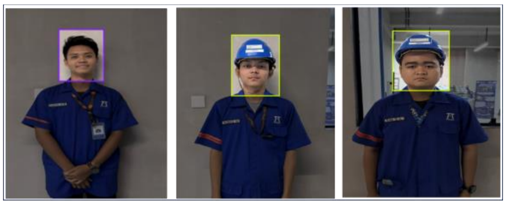

Figure 3. Labeling Dataset

d. Training Model

The next stage is training the object detection model using YOLO (You Only Look Once). This model was chosen because of its efficiency in detecting objects quickly and accurately, even on devices with limited resources. The training was done on Google Colab, which provides access to free GPUs to speed up the computing process. The labeled dataset was uploaded to the training environment, and data augmentation was performed to expand the dataset variations, such as image rotation, lighting changes, and flipping.

During training and testing, 40 epochs were used, and the batch size was optimized to achieve maximum accuracy. The training results are a detection model that can recognize patterns of PPE use automatically. The resulting model is evaluated using metrics such as mean confusion matrix to measure the level of accuracy of its predictions. If it is not adequate, the training process is repeated by adjusting the parameters or adding training data.

e. Node-red

Node-RED is a workflow-based tool for connecting hardware, APIs, and online services that is ideal for IoT-based applications. In the context of Personal Protective Equipment (PPE) detection systems using IoT-based image processing, Node-RED can play an important role in managing data flow and information processing. Node-RED can receive image or video streaming data from cameras used to detect the use of PPE. Node-RED is used to display the results of the number of safety and unsafe detection conditions.

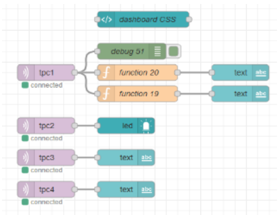

Figure 4. Node-red Flow System

f. System Implementation

After the model is trained, the system is implemented for real-time detection. The program is run using Python Idle with a webcam connected as a visual input device. The webcam captures live video from the production area, and the trained detection model is used to analyze the images frame-by-frame. Each PPE detection is immediately categorized as Safety or Not Safety according to the model's prediction.

The detection results are sent to the Node-RED platform for further monitoring. Node-RED allows integration with other systems, such as dashboards for data visualization, automatic alarms for alerts, or data logs for deeper analysis. The system is designed to run in real-time, allowing for quick action if violations of PPE usage are found. In addition, the system is flexible to be expanded, for example by adding new PPE types or automatic reporting features.

g. Evaluation and Testing

The final stage is the evaluation and testing of the system to ensure its performance meets the needs. Testing is carried out in a production environment with various scenarios, such as moving workers, changing camera angles, and non-ideal lighting conditions. Test results are measured using metrics such as accuracy, precision, recall, and inference time [13]. The detection results are compared with ground truth data to identify model weaknesses, such as missed detections or false positives.

If obstacles are found during testing, such as detection errors under certain conditions, the model is re-evaluated. This evaluation includes an analysis of the dataset, training parameters, and model performance. Based on the evaluation results, improvements are made, for example by adding more specific training data or adjusting model parameters. This process continues until the system reaches an optimal level of accuracy for effective use in the production area [14].

The final stage of this method is to analyze and validate the results of object counting by the system. Data taken from the IoT platform is used to assess the accuracy and performance of the system in detecting objects under various conditions. The evaluation process is carried out by comparing the number of objects detected by the system with the actual number counted manually. The results of this comparison are used to calculate parameters from confusion matrix [15]. Calculate parameter such as accuracy, precision, recall, and F1 Score, according to equations (1), (2), (3), and (4).

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN} \quad (1)$$

$$Precision = \frac{TP}{TP+FP} \quad (2)$$

$$Recall = \frac{TP}{TP+FN} \quad (3)$$

$$F1\ Score = \frac{2 \times Precision \times Recall}{Precision + Recall} \quad (4)$$

3. Results and Discussion

a. Results

This section presents the results obtained from testing the Personal Protective Equipment (PPE) usage detection system using the YOLO model, as well as an analysis of the system's performance and weaknesses. Based on the evaluation results, the YOLO model achieved an accuracy level of 95%. The real-time system performance showed a detection speed of up to 30 frames per second (FPS), so this system can operate optimally in a dynamic work environment. These results indicate that the deep learning-based approach is very effective in detecting PPE usage patterns. The discussion focuses on factors that affect model performance, such as dataset quality, data augmentation, and training parameters. Although the analysis results have quite high accuracy, they show several detection errors, such as false positives that occur in images with backgrounds resembling PPE or miss detections in low lighting conditions. Compared to similar studies, this approach is superior in terms of detection speed but shows limitations in generalization under extreme conditions. Therefore, recommendations for improvement include enriching the dataset with more variations, as well as optimizing the model to improve performance under complex environmental conditions.

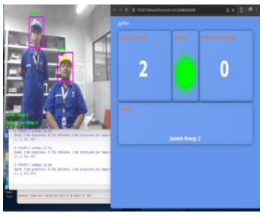

Figure 5. Safety Conditions and Dashboard Display Node-Red

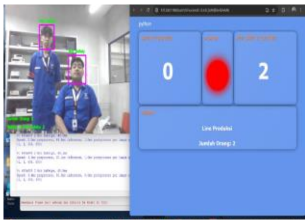

Figure 7. Not safety condition and dashboard display

The results of this study indicate that the YOLO-based PPE detection system can detect the use of PPE with a sufficient level of accuracy in various scenarios. The model trained with various datasets successfully classifies workers into two main classes, namely Safety for workers who use complete PPE, and Not Safety for workers who do not use PPE or use it inappropriately. In trials in a real work environment, the system showed consistent performance, even in dynamic work conditions. Based on the results of the experiment, an accuracy value of 95%, precision of 95%, recall of 99%, and f1 score of 97% were obtained. Visualization of the research results is displayed in the form of a confusion matrix table and a graph of detection results. This test supports the conclusion that the system has great potential to be implemented in the work environment to ensure worker compliance with work safety standards.

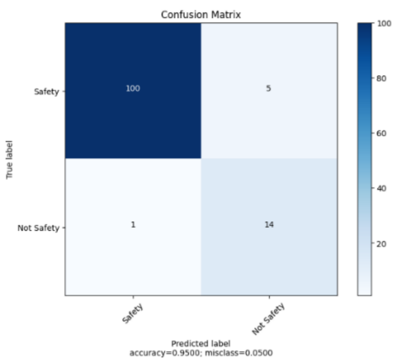

Figure 6. Confusion Matrix

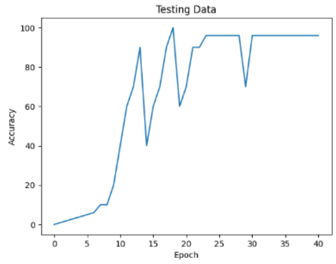

Figure 7. System test graph

3.2. Discussion

An in-depth discussion is conducted to analyze the factors that influence the success and weaknesses of the developed PPE detection system. The success of the YOLO model is largely due to its efficient architecture, which allows object detection in a short time, as well as the use of datasets with a variety of representative conditions. Data augmentation during training also contributes to the model's ability to recognize PPE usage patterns in various scenarios. However, several limitations of the system were identified during testing, such as decreased accuracy in less than optimal lighting conditions and fast worker movements. This suggests that the model requires further improvement, for example through the use of transfer learning techniques to improve generalization to more complex data. The discussion also includes a comparison with other methods used in similar studies, showing that although YOLO excels in terms of speed, there is still room for improvement in terms of robustness under extreme conditions.

4. Conclusion

This study successfully developed a YOLO-based PPE usage detection system that can detect and classify workers into the Safety and Not Safety categories with 95% accuracy, 95% precision, 99% recall, and 97% f1 score. The system shows reliable performance for real-time applications in the work environment, with a detection speed of up to 30 FPS. With this implementation, the system can play a significant role in improving work safety and ensuring worker compliance with safety protocols. However, several limitations need to be overcome, such as decreased performance in low lighting conditions or on images with fast movement. For future research, it is recommended to improve the quality of the dataset, adopt fine-tuning techniques, and explore integration with other supporting technologies, such as object tracking systems or multi-camera detection.

References

[1] P. Visconti, G. Rausa, C. Del-Valle-Soto, R. Velázquez, D. Cafagna, and R. De Fazio, "Machine Learning and IoT-Based Solutions in Industrial Applications for Smart Manufacturing: A Critical Review," Nov. 01, 2024, *Multidisciplinary Digital Publishing Institute (MDPI)*. doi: 10.3390/fi16110394.

[2] N. D. Nath, A. H. Behzadan, and S. G. Paal, "Deep learning for site safety: Real-time detection of personal protective equipment," *Autom Constr*, vol. 112, Apr. 2020, doi: 10.1016/j.autcon.2020.103085.

[3] O. Wosner, G. Farjon, and A. Bar-Hillel, "Object detection in agricultural contexts: A multiple resolution benchmark and comparison to human," *Comput Electron Agric*, vol. 189, Oct. 2021, doi: 10.1016/j.compag.2021.106404.

[4] G. Gallo, F. Di Rienzo, F. Garzelli, P. Ducange, and C. Vallati, "A Smart System for Personal Protective Equipment Detection in Industrial Environments Based on Deep Learning at the Edge," *IEEE Access*, vol. 10, pp. 110862–110878, 2022, doi: 10.1109/ACCESS.2022.3215148.

[5] P. Jiang, D. Ergu, F. Liu, Y. Cai, and B. Ma, "A Review of Yolo Algorithm Developments," in *Procedia Computer Science*, Elsevier B.V., 2021, pp. 1066–1073. doi: 10.1016/j.procs.2022.01.135.

[6] V. S. K. Delhi, R. Sankarlal, and A. Thomas, "Detection of Personal Protective Equipment (PPE) Compliance on Construction Site Using Computer Vision Based Deep Learning Techniques," *Front Built Environ*, vol. 6, Sep. 2020, doi: 10.3389/fbuil.2020.00136.

[7] S. Misra, C. Roy, T. Sauter, A. Mukherjee, and J. Maiti, "Industrial Internet of Things for Safety Management Applications: A Survey," *IEEE Access*, vol. 10, pp. 83415–83439, 2022, doi: 10.1109/ACCESS.2022.3194166.

[8] B. Di, L. Xiang, Y. Daoqing, and P. Kaimin, "MARA-YOLO: An Efficient Method for Multiclass Personal Protective Equipment Detection," *IEEE Access*, vol. 12, pp. 24866–24878, 2024, doi: 10.1109/ACCESS.2024.3365504.

[9] J. Redmon and A. Farhadi, "YOLOv3: An Incremental Improvement," Apr. 2018, [Online]. Available: http://arxiv.org/abs/1804.02767

[10] Q. Zhou, Z. Wang, Y. Zhong, F. Zhong, and L. Wang, "Efficient Optimized YOLOv8 Model with Extended Vision," *Sensors*, vol. 24, no. 20, Oct. 2024, doi: 10.3390/s24206506.

[11] R. Y. Ju and W. Cai, "Fracture detection in pediatric wrist trauma X-ray images using YOLOv8 algorithm," *Sci Rep*, vol. 13, no. 1, Dec. 2023, doi: 10.1038/s41598-023-47460-7.

[12] E. Guney, H. Altin, A. Esra Asci, O. U. Bayilmis, and C. Bayilmis, "YOLO-Based Personal Protective Equipment Monitoring System for Workplace Safety," *JITSI : Jurnal Ilmiah Teknologi Sistem Informasi*, vol. 5, no. 2, pp. 77–85, Jun. 2024, doi: 10.62527/jitsi.5.2.238.

[13] J. Valente, J. António, C. Mora, and S. Jardim, "Developments in Image Processing Using Deep Learning and Reinforcement Learning," Oct. 01, 2023, *Multidisciplinary Digital Publishing Institute (MDPI)*. doi: 10.3390/jimaging9100207.

[14] J. Häikiö, J. Kallio, S.-M. Mäkelä, and J. Keränen, "IoT-based safety monitoring from the perspective of construction site workers," *International Journal of Occupational and Environmental Safety*, vol. 4, no. 1, pp. 1–14, Apr. 2020, doi: 10.24840/2184-0954_004.001_0001.

[15] T. Wu and Y. Dong, "YOLO-SE: Improved YOLOv8 for Remote Sensing Object Detection and Recognition," *Applied Sciences (Switzerland)*, vol. 13, no. 24, Dec. 2023, doi: 10.3390/app132412977.