

AI-DRIVEN REAL-TIME PPE COMPLIANCE MONITORING IN THE PALM OIL INDUSTRY USING YOLOv8

Che Aqil Zulhazim Che Hassan^a, Muhammad Rusydi Muhammad Razif^{a*}, Muhammad Nasim Sulaiman^a, Nurul Hasyimah Mohd Mustapha^a, Muhd Amin Saad^b, Mohd Norazsyaam Azman^c

^aElectrical Engineering Technology Department, Faculty of Engineering Technology, Universiti Tun Hussein Onn Malaysia, 84600 Muar, Johor, Malaysia

^bElectrical Engineering Technology and Multimedia Cluster, Laboratory Management Office, Universiti Tun Hussein Onn Malaysia, 84600 Muar, Johor, Malaysia

^cVictory Enghoe Plantations Sdn. Bhd., Southern Malay Palm Oil Mill Miles 41, Jln Johor Bahru 86200 Simpang Renggam, Kluang, Johor, Malaysia

Article history

Received

13 February 2025

Received in revised form

21 June 2025

Accepted

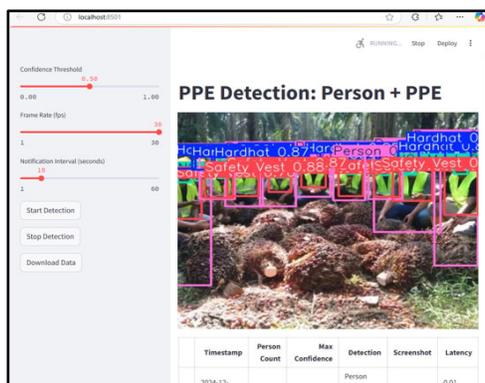
13 August 2025

Published Online

27 February 2026

*Corresponding author
rusydi@uthm.edu.my

Graphical abstract



Abstract

Ensuring personal protective equipment (PPE) compliance in hazardous work environments, such as the palm oil industry, remains a significant challenge due to the limitations of manual inspections. Current monitoring methods are mostly labor-intensive, prone to human error, and lack real-time capabilities. This study proposes an AI-driven PPE detection system utilizing YOLOv8, an object detection model, to enhance safety compliance through real-time monitoring. A comprehensive dataset of PPE usage scenarios was developed, and the YOLOv8 model was trained to recognize essential safety gear, including hard hats, safety vests, masks, and safety shoes. Based on experimental results it was found that the proposed system achieves a mean average precision (mAP) of 67.3% for 200 epochs of training, with a precision of 96.5% and a recall of 85.4%, significantly improving PPE detection accuracy compared to previous models. Furthermore, the system integrates Streamlit for an interactive interface and a Telegram application-based notifications for real-time compliance alerts. These findings suggested that AI-based monitoring can provide an effective, scalable, and automated solution for enforcing PPE regulations, reducing workplace hazards, and enhancing operational efficiency in the palm oil industry.

Keywords: PPE compliance, YOLOv8, palm oil industry, AI-based monitoring, real-time detection

Abstrak

Memastikan pematuhan peralatan perlindungan diri (PPE) dalam persekitaran kerja berisiko tinggi seperti industri kelapa sawit merupakan satu cabaran utama disebabkan oleh keterbatasan pemeriksaan secara manual. Kaedah pemantauan tradisional memerlukan tenaga kerja yang tinggi, mudah terdedah kepada kesilapan manusia, dan tidak menyediakan pemantauan masa nyata. Kajian ini mencadangkan sistem pengesanan PPE berdasarkan AI menggunakan YOLOv8, iaitu model pengesanan objek, untuk meningkatkan pemantauan keselamatan secara automatik. Satu dataset komprehensif mengenai senario penggunaan PPE telah dibangunkan, dan

model YOLOv8 telah dilatih untuk mengenal pasti peralatan keselamatan utama seperti topi keselamatan, jaket keselamatan, pelitup muka, dan kasut keselamatan. Keputusan eksperimen menunjukkan bahawa sistem ini mencapai ketepatan purata (mAP) sebanyak 67.3% selepas 200 kitaran latihan, dengan ketepatan sebanyak 96.5% dan kebolehsesanan sebanyak 85.4%, sekali gus meningkatkan ketepatan pengesanan PPE berbanding model terdahulu. Selain itu, sistem ini diintegrasikan dengan Streamlit untuk antara muka yang interaktif dan pemberitahuan berasaskan aplikasi Telegram untuk amaran pematuhan masa nyata. Dapatan ini mencadangkan bahawa pemantauan berasaskan AI mampu menjadi penyelesaian yang berkesan, berskala, dan automatik bagi penguatkuasaan peraturan PPE, seterusnya mengurangkan risiko kemalangan di tempat kerja dan meningkatkan kecekapan operasi dalam industri kelapa sawit.

Kata kunci: Pematuhan PPE, YOLOv8, industri kelapa sawit, pemantauan berasaskan AI, pengesanan masa nyata.

© 2026 Penerbit UTM Press. All rights reserved

1.0 INTRODUCTION

Personal protective equipment (PPE) plays a crucial role in ensuring worker safety in high-risk industries such as palm oil plantations. These environments expose workers to various occupational hazards, including heavy machinery, falling objects, chemical exposure, and extreme weather conditions. PPE, such as helmets, safety vests, gloves, face masks, and safety boots, helps significantly in reducing the risk of workplace injuries and illnesses. However, ensuring consistent PPE compliance remains a challenge, primarily due to the limitations of manual inspections, which are labor-intensive, prone to human error, and reactive rather than proactive [1, 2].

Traditional safety monitoring systems rely on periodic supervisory checks and audits, which are often inefficient in large-scale environments such as palm oil plantations, where workers are dispersed across vast areas [3, 4]. The palm oil industry employs millions of workers globally, making safety compliance a critical concern [5]. While regulatory frameworks, such as the Occupational Safety and Health Act (OSHA) and industry-specific certifications like the Malaysian Sustainable Palm Oil (MSPO) and the Roundtable on Sustainable Palm Oil (RSPO), mandate PPE usage, their enforcement mainly depends on manual oversight, which is often inconsistent [6]. Studies have shown that non-compliance with PPE requirements significantly increases workplace accidents, leading to severe injuries and financial losses [7, 8].

To address these challenges, there is a growing interest in leveraging Artificial Intelligence (AI) and computer vision technologies for automated PPE monitoring. Deep learning models, particularly Convolutional Neural Networks (CNNs), have demonstrated high accuracy in detecting objects, including PPE, in real-time scenarios [9]. Among various object detection models, You Only Look Once (YOLO) has emerged as a leading approach for fast and accurate real-time detection, processing entire

images in a single pass to enable rapid detection while maintaining high precision [10]. The iteration of YOLOv8 further enhances detection efficiency, feature extraction, and robustness in complex environments, making it well-suited for applications in dynamic outdoor settings such as palm oil plantations [11]. However, implementing AI-driven PPE detection in such environments presents unique challenges, including variable lighting conditions, occlusions, worker movement, and background clutter, which can affect detection accuracy [12].

Furthermore, the integration of real-time monitoring systems with safety management frameworks requires seamless communication channels, such as automated alerts via platforms like the Telegram application, to ensure immediate corrective action on PPE non-compliance [13]. To complement this, a responsive and user-friendly dashboard is essential for visualizing and interpreting detection results in real time. Streamlit fulfills this role effectively, offering a lightweight, open-source, and Python-based solution that supports rapid development, minimal setup, and highly customizable interfaces, making it particularly suitable for real-time AI monitoring applications.

Although research on AI-driven PPE compliance monitoring is growing, most existing studies focus on structured environments such as construction sites and manufacturing plants, where lighting conditions and worker positioning are relatively controlled. The application of AI in unstructured environments such as palm oil plantations remains underexplored, necessitating further research into developing robust models capable of adapting to real-world conditions [14]. Addressing these limitations is essential for achieving scalable, reliable, and cost-effective PPE monitoring solutions.

This research aims to bridge this gap by developing an AI-based PPE detection system utilizing YOLOv8, specifically optimized for large-scale palm oil plantations. By integrating automated real-time monitoring with notification features, this study aims to

provide a scalable and effective solution for PPE compliance enforcement, contributing to enhanced workplace safety in the palm oil industry. Furthermore, the proposed system has the potential for cross-industry applications, extending its impact beyond plantations to sectors such as construction, healthcare, and manufacturing, where PPE compliance is equally critical [15, 16].

By addressing the limitations of manual PPE enforcement and leveraging AI for real-time safety monitoring, this study contributes to the advancement of AI-integrated occupational safety solutions, setting a foundation for future research in automated compliance enforcement. The findings of this research can serve as a benchmark for developing AI-driven safety monitoring systems in various high-risk industries, promoting a more efficient working environment [17, 18].

2.0 METHODOLOGY

This section explains the procedures for image detection. A detection and localization framework were established using the YOLOv8 model to monitor PPE compliance in palm oil plantations. The YOLOv8 architecture is structured to autonomously extract and assess distinct PPE characteristics from input images. This allows the model to accurately recognize and pinpoint the necessary equipment in real-time, enhancing both safety oversight and operational efficiency. Figure 1 illustrates the process for real-time PPE detection using YOLOv8. The process begins with data collection, where images are gathered from various sources and divided into training datasets and testing. The training process involves creating a label map, configuring the YOLOv8 model and training it to produce the trained model and performance graphs. Data preprocessing, a critical step in AI model development and training, is discussed in Section 2.2. The testing phase, detailed in Section 2.4 on system integration, includes inference, parameter adjustments, and result evaluation to generate outputs with bounding boxes and notifications of the detection via instant messaging (IM) through the Telegram application.

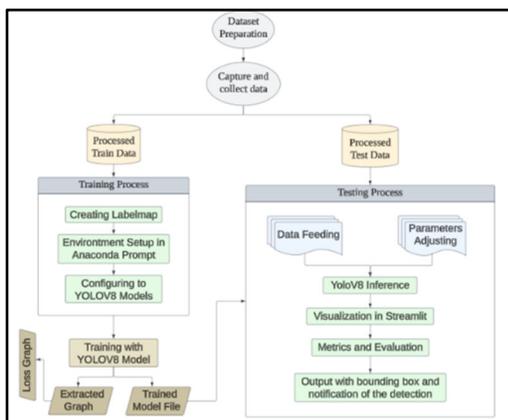


Figure 1 Workflow of the image detection

2.1 Image Acquisition

Data collection is a crucial step in this project, involving the gathering of information from various reliable sources. For this study, the dataset was created by collecting images from manual capture, social media platforms and online databases. To enhance decision making, additional images were also collected from the dataset used in a previous study [19]. The dataset contains images showing various situations related to PPE detection in the palm oil mill section. They were divided into seven main classes which are Person, Hardhat, Mask, Safety Vest, No-Hardhat, No-Mask, and No-Safety Vest. These classes ensure the data set includes examples of both compliance and non-compliance with safety measures.

A well-structured dataset is essential for training machine learning models and achieving accurate results. The dataset was carefully prepared to meet the goals of this project, ensuring it had enough varieties to train the YOLOv8 model effectively. The images were split into three parts which are 60% for training, 30% for validation, and 10% for testing. This distribution helps to evaluate the model fairly and ensure it performs well in real-life conditions. Figure 2 presents sample images from the dataset, whereas Table 1 details the number of images per class from a total of 7,500 images.

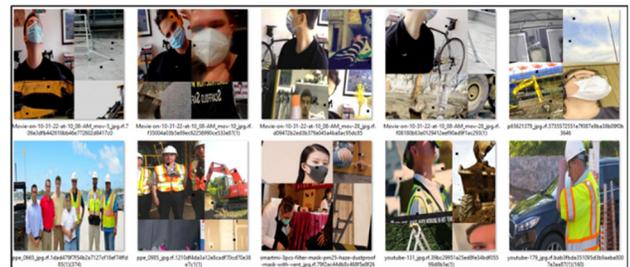


Figure 2 Sample images from the dataset

Table 1 Distribution of images by class in the dataset

Class	Train	Test	Valid	Total
Hardhat	2,628	272	552	3,452
Mask	2,192	156	266	2,614
Safety Vest	3,638	194	352	4,184
No-Hardhat	2,760	238	518	3,516
No-Mask	3,062	268	590	3,920
No-Safety Vest	3,728	340	774	4,842
Person	5,052	542	1128	6,722

2.2 Data Preprocessing and Annotation

This section outlines the data preprocessing and annotation procedures, which include image resizing, labeling, and annotation before generating a plain text file (.txt) for further processing.

The image resizing process ensures that all dataset images are standardized to a consistent size. In this study, all images were resized to 640×640 pixels, as illustrated in Figure 3.



Figure 3 Example of dataset image before and after resizing to 640x640 pixels

For image annotation, the Roboflow labelling software was used. Tools such as the bounding boxes and polygons, were utilized to highlight specific regions within the images and categorize them into seven predefined classes. The annotated images were then automatically saved, and the software further generated text files containing detailed information about the labelled regions. These plain text files facilitated structured data export and import ensuring seamless integration into subsequent processing stages. Figure 4(a) provides a flowchart outlining the labelling process and the preprocessing steps is comprehensively illustrated in Figure 4(b), whereas Figure 5 demonstrates examples of the labelled data.

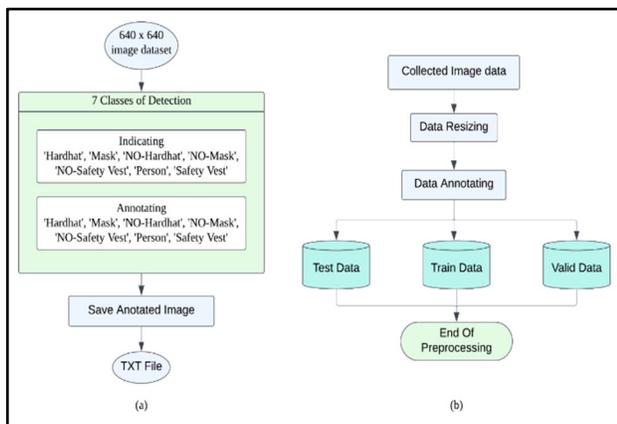


Figure 4 Workflow for (a) Labeling the resized image data and (b) Data preprocessing flow



Figure 5 Sample of labelled image dataset

Once the labelling was completed, the dataset was divided into three subsets namely training, validation, and testing which were 60% for training, 10% for testing, and 30% for validation. This distribution ensured a well-balanced dataset to support effective model training and evaluation. With finalized annotation and preprocessing, the dataset was ready for subsequent data processing and model development phases.

2.3 Model Development and Training

YOLOv8, one of the most advanced iterations in the YOLO family, introduces an anchor-free design, enhanced feature extraction, and improved optimization strategies compared to its predecessors, YOLOv7, resulting in faster inference while maintaining high accuracy. With detection speeds between 5 to 160 FPS, YOLOv8 outperforms its predecessors and other state-of-the-art detectors, making it highly suitable for real-time applications. Its exceptional performance on benchmark datasets like MS COCO further reinforces its dominance in the field [20]. Figure 6 illustrates the YOLOv8 network structure.

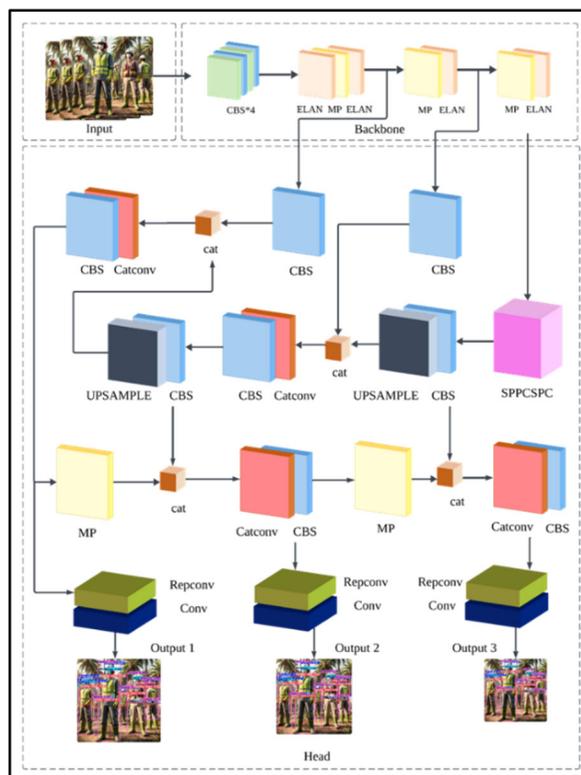


Figure 6 Architecture of YOLOv8 network

The YOLOv8 network consists of three main components which are the input network, the backbone network, and the head network. Initially, the input network preprocesses the image by resizing it to $640 \times 640 \times 3$ before feeding it to the backbone network for feature extraction. The backbone network incorporates advanced modules like the Max-Pooling (MP) module, Convolution, BatchNorm, and SiLU (CBS) composite module, and the Efficient Layer Aggregation Network (ELAN), which were originally introduced in YOLOv7 and refined for YOLOv8.

In the CBS composite module, the processes involved are convolution, batch normalization, and activation using the Sigmoid Linear Unit (SiLU) function, improving the feature representation of input data. Meanwhile, the ELAN module uses expand, shuffle, and merge cardinality techniques to increase the learning capacity without disrupting gradient flow, thereby increasing the accuracy. Additionally, the MP module employs a dual-branch design where one branch reduces the spatial dimensions using max pooling, whereas the other uses convolutions to achieve the same result. These two outputs are then merged, generating a feature map with reduced spatial size and increased channel depth.

To build the three-layer output from the backbone network, the head network needs to perform object detection and prediction across multiple scales. The Spatial Pyramid Pooling - Cross Stage Partial Connections (SPPCSPC) module, inspired by YOLOv5's Spatial Pyramid Pooling - Fast (SPPF), captures multi-scale object information using max pooling operations with kernel sizes of 5, 9, and 13. This approach increases the receptive field while preserving the spatial dimensions of the feature maps. Additionally, the RepConv module introduces a re-parameterization mechanism, where a multi-branch structure is used during training to enhance learning. During inference, these branches are merged into a single convolutional layer, producing a model that balances high efficiency with improved accuracy.

2.4 System Integration

The integration of the YOLOv8 model with Streamlit was implemented to develop an interactive platform for real-time PPE detection monitoring. Streamlit is an open-source Python library designed for building web-based applications for machine learning and data science projects. It provides an intuitive and efficient framework for creating interactive user interfaces with minimal code, making it an ideal choice for this project. By leveraging Streamlit, the YOLOv8-based system enables users to monitor PPE compliance in real time through a web interface, eliminating the need for complex GUI development. Additionally, the lightweight and fast deployment capabilities of Streamlit enable seamless integration with various computing environments, ensuring accessibility across different devices.

During system integration, the YOLOv8 model, trained on a custom dataset, was seamlessly incorporated into the Streamlit application. The system enables users to adjust detection parameters such as confidence threshold and frame rate via a dynamic sidebar as shown in Figure 7 (a), enabling customization to suit various operational environments. Real-time video input is captured through a webcam and processed frame-by-frame using YOLOv8's inference capabilities. Figure 7 (b) illustrates the area where the model identified with PPE compliance and displaying the video feed in real-time.

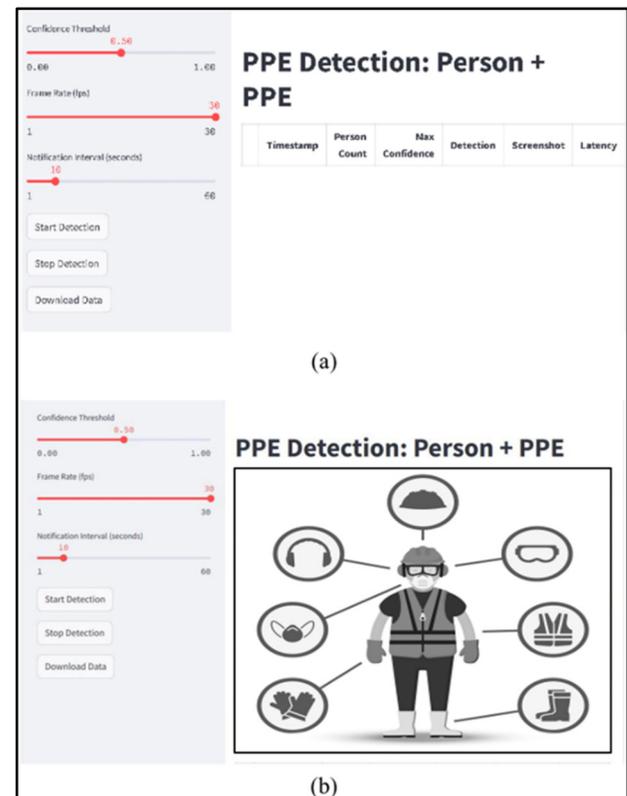


Figure 7 (a) Initializing Streamlit and (b) Detection of a person with PPE compliance

Streamlit's features were further utilized to provide a comprehensive overview of detection results. The main performance metrics, including timestamp, number of person detections, maximum confidence scores, and detection latency, were dynamically updated and tabulated in tables within the application, as shown in Figure 8 (a). These visualizations offer users an intuitive way to understand system performance and detection trends over time. Additionally, the system includes a data export functionality allowing detection results to be saved in a CSV file (.csv) for further analysis, as illustrated in Figure 8 (b).

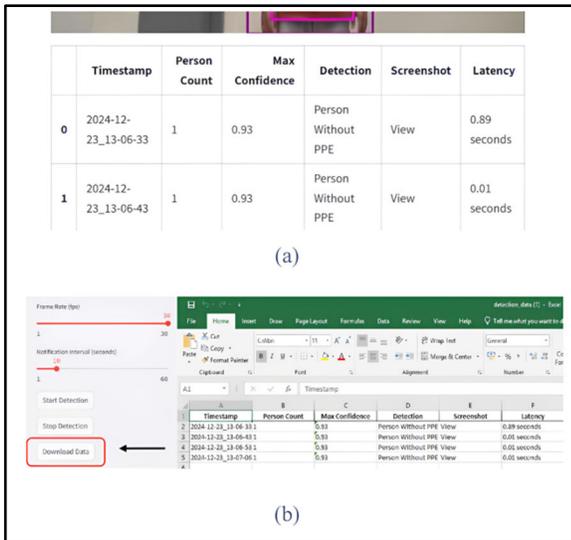


Figure 8 (a) Streamlit detection table and (b) Detection data download function in CSV format

To enhance safety monitoring, the system integrated a notification mechanism via Telegram application. When specific non-compliance safety conditions occur, such as detecting a person without PPE, the monitoring system will capture the image of the person, save a screenshot of the image and send it to a pre-configured Telegram application as shown in Figure 9. The real-time alerting mechanism developed can ensure safety compliance in the workplace, thereby reducing the risk of injury.

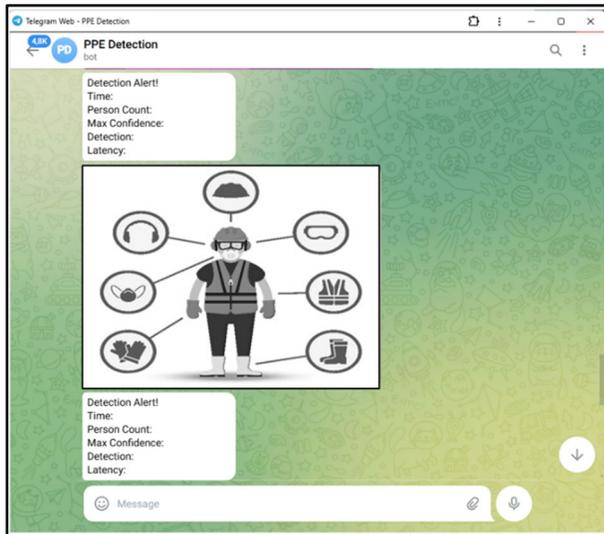


Figure 9 Notification of the detection via Telegram apps

The architecture of the system leveraged Streamlit's capabilities and Python's threading functionalities to manage real-time video processing, data visualization and notifications efficiently. Tasks like saving images and sending Telegram application

notifications were handled concurrently, ensuring smooth operation without interrupting the main detection loop.

2.5 Testing and Evaluation

This study utilized precision, recall, mean average precision (mAP), and F1 score to assess the model's performance. Precision, one of the most commonly used metrics, can be defined as the ratio of correctly identified targets to the total number of detected targets. In general, a higher precision value reflects improved detection accuracy. However, although precision is a straightforward metric, it may not always provide a complete evaluation. To achieve a more comprehensive assessment, mAP, recall, and F1 score were also incorporated. The formulas for calculating precision, recall, mAP, and F1 score are as equation (1) to equation (5) respectively.

$$\text{Precision: } P = \frac{TP}{TP+FP} \times 100\% \tag{1}$$

$$\text{Recall: } R = \frac{TP}{TP+FN} \times 100\% \tag{2}$$

$$\text{Average Precision: } AP = \int_0^1 P(r)dr \tag{3}$$

$$\text{Mean Average Precision: } mAP = \frac{1}{n} \sum_{i=1}^n AP_i \tag{4}$$

$$\text{F1 score: } F1 = 2 \times \frac{P \times R}{P+R} \tag{5}$$

where true positive (TP) refers to the number of correctly detected PPE objects, false positive (FP) denotes the number of other objects mistakenly identified as PPE, and false negative (FN) represents the number of PPE objects that were not detected or missed. These metrics are crucial for understanding the model's reliability in identifying PPE compliance and non-compliance in real-world scenarios.

To evaluate the performance of the YOLOv8 model, a comparison was made with YOLOv5 and YOLOv11 models under the same conditions. All models were trained and tested on a dataset with seven detection classes which are Hardhat, Mask, Safety Vest, No-Hardhat, No-Mask, No-Safety Vest, and Person. The models were trained for 100 and 200 epochs to analyze the impact of longer training durations on performance.

Trainable parameters refer to the weights and biases in the model that are updated during training to improve accuracy. The training process was conducted on a computer using Anaconda as the environment manager. A virtual environment was set up, and essential tools such as PyTorch and the YOLOv8 library were installed. The training scripts were executed in the Anaconda Prompt, with batch size, learning rate, and other configurations specified in the

YOLOv8 file. Checkpoints were saved during training to track progress. Table 2 presents the parameters and testing setup, whereas Section 3 discusses the results in detail.

Table 2 Testing parameters for different models

Model	Epoch	Class	Trainable Parameters
Proposed YOLOv8	100	All	11.2M
Proposed YOLOv8	200	All	11.2M
YOLOv11	100	All	13.5M
YOLOv11	200	All	13.5M
YOLOv5	100	All	7.2M
YOLOv5	200	All	7.2M

2.6 Dataset Collection

To build a robust and diverse dataset, images were collected from various industrial environments, with a particular focus on scenarios in palm oil plantations. The dataset includes a wide range of PPE compliance and non-compliance cases, such as workers wearing safety helmets, face masks, and safety vests, as well as instances where PPE was missing or improperly worn. The dataset collection was conducted using high-resolution cameras in real-environment conditions to ensure the relevance and applicability of the model to safety monitoring tasks. The collected dataset was manually annotated using tools such as Roboflow, where PPE elements were highlighted and labeled for detection.

The dataset collection process accounted for diverse environmental factors to simulate real-world conditions more accurately. Variations in lighting, including bright sunlight, shadows, and low-light scenarios, were incorporated to test the model's adaptability to different illumination levels. Worker positions and postures were also varied to capture PPE usage from multiple angles, ensuring that the model could generalize across perspectives. Additionally, different types of PPE, such as varying colors and designs of helmets, safety vests and face masks, were included to enhance the model's robustness and reduce bias toward specific appearances.

2.7 Experimental Setup

The YOLOv8 model's capability in detecting PPE compliance was assessed using Python on a system equipped with CUDA 12.4, featuring an NVIDIA GeForce RTX 4070 Ti GPU with 12 GB of VRAM and 64 GB of RAM. With a network comprising 218 layers and 25,843,813 parameters, the model exhibited high computational efficiency, achieving a GFLOP value of 78.7.

The experimental setup as illustrated in Figure 10 involved processing the collected dataset through the YOLOv8 model to detect and classify PPE objects. Images were preprocessed by resizing them to 640×640 pixels to standardize input data. Real-time inference was tested on video streams, simulating practical safety monitoring scenarios. The system's performance was analyzed using multiple evaluation metrics, providing a comprehensive assessment of its accuracy, speed, and effectiveness in ensuring PPE compliance.

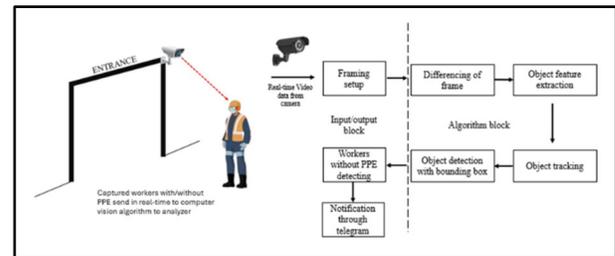


Figure 10 Experimental setup for PPE detection

3.0 RESULTS AND DISCUSSION

The YOLOv8-based model for PPE compliance detection was trained with specific configurations and hyperparameters. It underwent 200 epochs using the Stochastic Gradient Descent (SGD) optimizer and a pre-trained Common Object in Context (COCO) model. To prevent overfitting and improve efficiency, an early stopping mechanism was applied, automatically stopping training if no improvement was observed over the last 50 epochs, reducing unnecessary computations.

Figure 11 illustrates the YOLOv8-based recall-confidence curve and training performance over 200 epochs for PPE detection. In Figure 11 (a), the recall-confidence curve for individual PPE classes, such as Person, Hardhat, Mask, Safety Vest, No-Hardhat, No-Mask, and No-Safety Vest demonstrates that all classes consistently achieve values around 0.90 at a confidence threshold of 0.0. Additionally, the YOLOv8-based training graph, depicted in Figure 11 (b), demonstrates that the best performance for the proposed system was achieved at epoch 100. Therefore, the system was trained for 100 epochs due to this observed trend. The early stopping mechanism remained untriggered since the performance difference between epoch 100 and the final epoch was minimal.

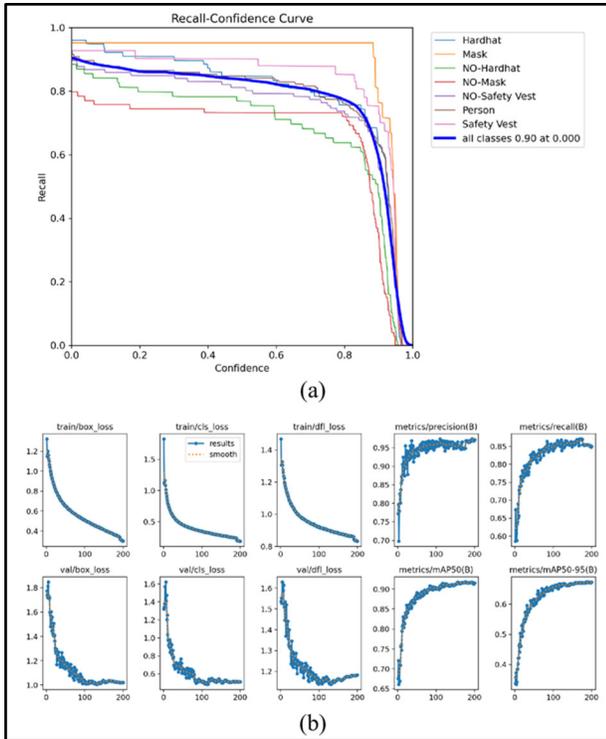


Figure 11 YOLOv8 performance (a) Recall-confidence curve and (b) Training performance with 200 epochs

In addition, the performance of the YOLOv8 model was able to achieved recall, precision and mAP of 85.4%, 96.5%, and 67.3% respectively. This performance was achieved after the YOLOv8 model underwent training for 200 epochs, as shown in Figure 12. For the results of all classes, the recall-confidence threshold value stabilized at 0.90 at a confidence level of 0.000. This illustrates that the recall-confidence converges smoothly and shows its consistent performance across PPE categories.

```

100 epochs completed in 4.403 hours.
optimizer stripped from runs\detect\train\weights\last.pt, 52.1MB
optimizer stripped from runs\detect\train\weights\best.pt, 52.1MB

validating runs\detect\train\weights\best.pt...
analytics 8.3.36 # Python-3.9.20 torch-2.5.1+cu124 CUDA:0 (NVIDIA GeForce RTX 4070 Ti, 12282MiB)
model summary (fused): 218 layers, 25,843,813 parameters, 0 gradients, 78.7 GFLOPs
Class      Images  Instances  Box(P)      R      mAP50  mAP50-95: 100% | 48/48 [00:08]
all         1520     7576      0.965      0.854  0.916  0.673
    
```

Figure 12 Training results of 200 epoch using YOLOv8

The confusion matrix shown in Figure 13 provides detailed information on the YOLOv8 model's performance in PPE detection, covering all seven classes and the background. TP represents correctly classified instances, whereas FP indicates misclassifications. The confusion matrix is measured because of its importance in demonstrating the performance of a classification algorithm and provide insight into the strengths and limitations of the model. In the confusion matrix, the evaluations measured are metrics such as precision, recall, and overall accuracy.

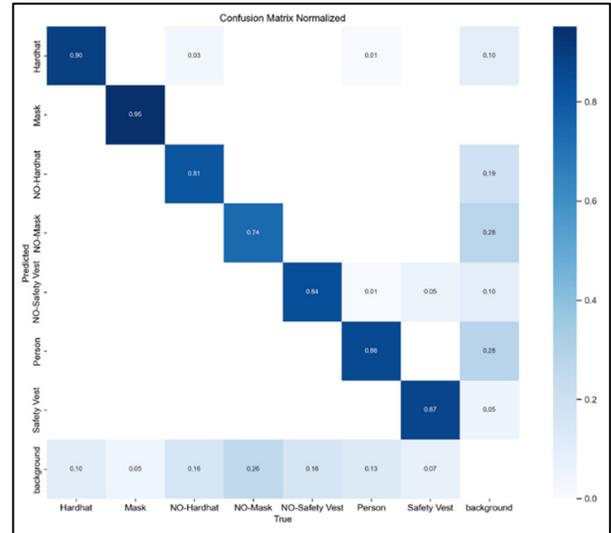


Figure 13 Confusion matrix diagram for 200 epochs

For the Hardhat class, the model achieves 90% TP accuracy, with 10% of instances misclassified as background or false negative (FN). Similarly, Mask detection shows 95% TP accuracy, with 5% FN. For Safety Vest detection shows 86% TP accuracy, with 7% FN. For No-Hardhat, the model detects 81% of instances correctly (TP) but is misclassified as 16% as background (FN). No-Mask detection attains 74% TP accuracy but misclassifies 26% as background (FN). The No-Safety Vest class achieves 84% TP accuracy, with 16% FN. Finally, for the Person class, the model identifies 86% correctly (TP) but confuses 13% as background (FN).

Figure 13 also shows that the range of the confusion matrix is from 0.01 to 0.90. This range gives a positive sign of confidence for each classification. Based on the confusion matrix value, false negatives can be addressed and thus improve the overall detection accuracy.

In addition, Figure 13 shows in detail the confusion matrix at the 200th iteration. At 200th iteration, the model performed accurately of 97.13%. The figure also shows that for the Mask class, the class dissection matrix for the predicted and actual, shows a significant peak accuracy of 95% and a slightly lower performance for the No Mask detection of 74%. Based on this performance, it can be concluded that the YOLOv8-based model shows highly potential for real-time PPE detection across different applications, such as static images, video streams, and live camera feeds. This highlights its effectiveness in practical safety monitoring applications through rigorous testing and evaluation.

3.1 Model Comparison

To evaluate the performance of the proposed YOLOv8 model, a comparison was conducted with YOLOv5 and YOLOv11. Each model was trained on a

dataset consisting of seven PPE classes for 100 epochs and 200 epochs. The performance of the models was compared based on their mAP, F1 Score, precision, and recall, whereas the number of trainable parameters provided insight into each model's computational complexity. Table 3 presents a summary of the models' performance comparison.

Table 3 Comparison of models for performance evaluation

Model	Epoch	mAP (%)	Precision (%)	Recall (%)	Trainable Parameters
Proposed YOLOv8	100	66.4	96.0	83.4	11.2M
Proposed YOLOv8	200	67.3	96.5	85.4	11.2M
YOLOv11	100	57.3	92.0	74.8	13.5M
YOLOv11	200	57.9	93.1	80.2	13.5M
YOLOv5	100	62.8	94.2	82.2	7.2M
YOLOv5	200	66.4	96.4	85.6	7.2M

The decision to select YOLOv8 as the proposed model was based on its balance between computational complexity and performance. While YOLOv8 has 11.2M trainable parameters, it delivers the best detection results in terms of mAP, precision, and recall, making it the most effective model for PPE detection tasks. Additionally, its performance improvements with more epochs indicate strong learning capabilities, ensuring better adaptability to the dataset. Although YOLOv11 is a newer version, its lower performance could be attributed to its architectural modifications not being well-suited for small object detection or the specific characteristics of the PPE dataset used in this study. Furthermore, newer models often require more careful hyperparameter tuning or may be optimized for different benchmarks, which can affect performance in custom applications. Therefore, YOLOv8 is the optimal choice for this study, providing reliable and accurate results while maintaining computational efficiency.

3.2 Integration into the Streamlit Platform

The integration of the YOLOv8 model with Streamlit demonstrated promising results in creating an efficient and interactive platform for real-time PPE detection in palm oil plantations. The custom-trained YOLOv8 model was successfully incorporated into a web-based Streamlit application, allowing real-time video input processing through a connected webcam. Main detection features, such as identifying PPE non-compliance (e.g., absence of Hardhat, Mask or Safety Vest), were accurately annotated on the video feed. The system also enabled users to dynamically adjust operational parameters, including detection confidence thresholds and frame rates, via a user-friendly sidebar interface, as depicted in Figure 14. These capabilities enhanced system adaptability, making it suitable for diverse field conditions. Early tests showed a seamless frame-by-frame processing workflow with low latency, ensuring smooth and accurate detection even in varied operational environments.

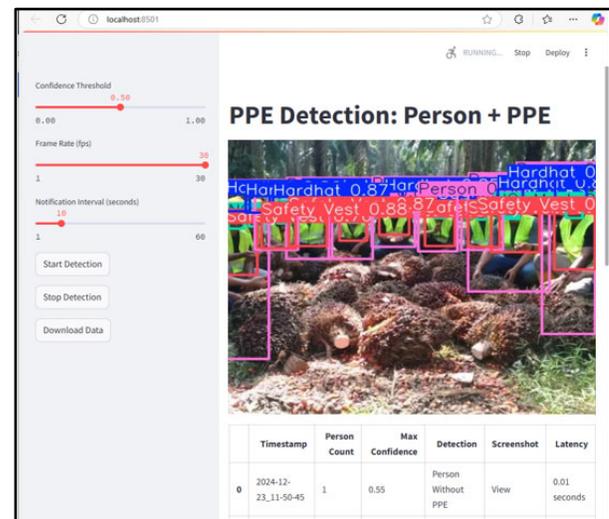


Figure 14 Streamlit platform custom setup for PPE compliance detection

Additionally, the Streamlit platform offered dynamic visualization of detection results, providing actionable insights into system performance and trends. Metrics such as detection counts, confidence scores, and latency were updated in real time. This intuitive presentation of data facilitated a clear understanding of the system's effectiveness. Preliminary tests also highlighted the robustness of the notification mechanism, which leveraged the Telegram application to send real-time alerts of detected non-compliance, as presented in Figure 15. Screenshots of these non-compliance conditions were saved and dispatched without interrupting video processing, confirming the efficacy of concurrent tasks managed by Streamlit's asynchronous features. The combination of efficient video processing, real-time monitoring, and notification capabilities highlights the system's potential to enhance workplace safety monitoring in palm oil plantations.

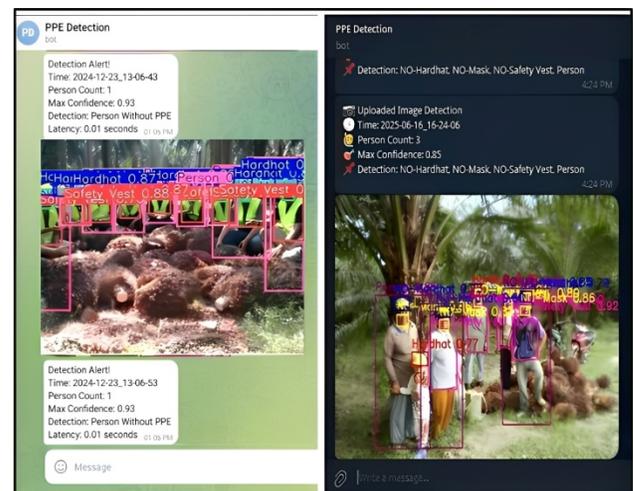


Figure 15 Notification alert for persons without PPE via Telegram

4.0 CONCLUSION

This study developed an AI-driven PPE compliance monitoring system using YOLOv8, achieving high accuracy and efficiency in workplace safety monitoring within palm oil plantations. The proposed YOLOv8 model outperformed YOLOv5 and YOLOv11, recording a precision of 96.5%, recall of 85.4%, and mean average precision (mAP) of 67.3% after 200 epochs. The confusion matrix analysis further highlighted the model's strong classification capabilities, with the highest accuracy observed for Mask detection (95%) and reliable performance across all PPE categories.

Integrated with Streamlit for real-time monitoring and Telegram alerts for instant notification of PPE non-compliance, the system ensures immediate corrective action while maintaining a low-latency processing workflow. Its scalability makes it adaptable to various industries, including construction, manufacturing, and healthcare. Future enhancements should focus on further AI model optimization, IoT integration, and large-scale deployment, improving real-time safety monitoring for enhanced workplace protection and regulatory compliance.

Acknowledgement

This research was supported by the Ministry of Higher Education (MOHE) through the Fundamental Research Grant Scheme (FRGS) (FRGS/1/2022/TK07/UTHM/02/3) and Universiti Tun Hussein Onn Malaysia (UTHM) through GPPS (vot J075).

Conflicts of Interest

The authors declare that there is no conflict of interest regarding the publication of this paper.

References

- [1] Li, Z., W. Xie, L. Zhang, S. Lu, L. Xie, H. Su, W. Du, and W. Hou. 2022. Toward Efficient Safety Helmet Detection Based on YOLOv5 with Hierarchical Positive Sample Selection and Box Density Filtering. *IEEE Transactions on Instrumentation and Measurement*. 71: 1–14. <https://doi.org/10.1109/TIM.2022.3169564>.
- [2] Yipeng, and J. Wu. 2024. Personal Protective Equipment Detection for Construction Workers: A Novel Dataset and Enhanced YOLOv5 Approach. *IEEE Access*. 12: 47338–47358. <https://doi.org/10.1109/ACCESS.2024.3382817>.
- [3] Lin, B. 2024. Safety Helmet Detection Based on Improved YOLOv8. *IEEE Access*. 12: 28260–28272. <https://doi.org/10.1109/ACCESS.2024.3368161>.
- [4] Wang, Z., Y. Wu, L. Yang, A. Thirunavukarasu, C. Evison, and Y. Zhao. 2021. Fast Personal Protective Equipment Detection for Real Construction Sites Using Deep Learning Approaches. *Sensors*. 21(10): 3478. <https://doi.org/10.3390/s21103478>.
- [5] Zou, J., T. Song, S. Cao, B. Zhou, and Q. Jiang. 2024. Dress Code Monitoring Method in Industrial Scene Based on Improved YOLOv8n and DeepSORT. *Sensors*. 24(18): 6063. <https://doi.org/10.3390/s24186063>.
- [6] Liu, G., Y. Zhang, J. Liu, D. Liu, C. Chen, Y. Li, X. Zhang, and P. L. Touko Mbouembe. 2024. An Improved YOLOv7 Model Based on Swin Transformer and Trident Pyramid Networks for Accurate Tomato Detection. *Frontiers in Plant Science*. 15: 1452821. <https://doi.org/10.3389/fpls.2024.1452821>.
- [7] Redmon, J., S. Divvala, R. Girshick, and A. Farhadi. 2016. You Only Look Once: Unified, Real-Time Object Detection. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. 779–788. <https://doi.org/10.1109/CVPR.2016.91>.
- [8] Chen, H., J. Qi, M. Wang, and C. Wu. 2023. Helmet-Wearing Detection Algorithm Based on Improved YOLOv5s. In *Proceedings of the IEEE International Conference on Computer Science and Applications (ICCSA)*, 8564–8569. <https://doi.org/10.1109/ICCSA.2023.8564896>.
- [9] Medioni, M., and M. Cahyandito. 2015. Is Converting Shrub Land to Palm Oil Plantation Worth It? An Ecosystem Services Analysis. *Journal of Academic Business and Retail Management*. 10(1). <https://doi.org/10.2139/ssrn.2676275>.
- [10] Mamun, K. A., et al. 2024. Smart Reception: An Artificial Intelligence-Driven Bangla Language-Based Receptionist System Employing Speech, Speaker, and Face Recognition for Automating Reception Services. *Engineering Applications of Artificial Intelligence*. 136: 108923. <https://doi.org/10.1016/j.engappai.2024.108923>.
- [11] Khatun, R., M. I. H. Reza, M. Moniruzzaman, and Z. Yaakob. 2017. Sustainable Oil Palm Industry: The Possibilities. *Renewable and Sustainable Energy Reviews*. 76: 608–619. <https://doi.org/10.1016/j.rser.2017.03.077>.
- [12] Waqar, A., N. Bheel, and B. A. Tayeh. 2024. Modeling the Effect of Implementation of Artificial Intelligence Powered Image Analysis and Pattern Recognition Algorithms in Concrete Industry. *Developments in the Built Environment*. 17: 100349. <https://doi.org/10.1016/j.dibe.2024.100349>.
- [13] Verma, A., G. Ghajari, K. M. T. Jawad, H. P. Salehi, and F. Amsaad. 2024. Real-Time Automated Donning and Doffing Detection of PPE Based on YOLOv4-Tiny. In *Proceedings of the IEEE National Aerospace Electronics Conference (NAECON)*. 398–402. <https://doi.org/10.1109/NAECON.2024.10670616>.
- [14] Luwes, N., and W. Pretorius. 2024. Development of a Cost-Effective Artificial Intelligence-Based Image Processing Sorting Mechanism for Conveyor Belt System. In *Proceedings of the IEEE International Conference on Human System Interaction (HSI)*. 1–6. <https://doi.org/10.1109/HSI.2024.10613600>.
- [15] Waqar, A., N. Bheel, and B. A. Tayeh. 2024. Modeling the Effect of Implementation of Artificial Intelligence Powered Image Analysis and Pattern Recognition Algorithms in Concrete Industry. *Developments in the Built Environment*. 17: 100349. <https://doi.org/10.1016/j.dibe.2024.100349>.
- [16] Luwes, N., and W. Pretorius. 2024. "Development of a Cost-Effective Artificial Intelligence-Based Image Processing Sorting Mechanism for Conveyor Belt System. In *International Conference on Human System Interaction (HSI)* IEEE Computer Society. 1–6. <https://doi.org/10.1109/HSI61632.2024.10613600>.
- [17] Al-Azani, S., H. Luqman, M. Alfarraj, A. A. I. Sidig, A. H. Khan, and D. Al-Hammed. 2024. Real-Time Monitoring of Personal Protective Equipment Compliance in Surveillance Cameras. *IEEE Access*. <https://doi.org/10.1109/ACCESS.2024.3451117>.
- [18] Li, H., H. Xu, Y. Li, and X. Li. 2024. Application of Artificial Intelligence (AI)-Enhanced Biochemical Sensing in Molecular Diagnosis and Imaging Analysis: Advancing and Challenges. *TrAC - Trends in Analytical Chemistry*. 174: 117700. <https://doi.org/10.1016/j.trac.2024.117700>.
- [19] Arfan, M., S. Sumardi, and H. Huboyo. 2023. Advancing Workplace Safety: A Deep Learning Approach for Personal Protective Equipment Detection Using Single Shot Detector. In *International Workshop on Artificial Intelligence and Image Processing (IWAIP 2023)*, Institute of Electrical

- and Electronics Engineers Inc. 127–132. <https://doi.org/10.1109/IWAIP58158.2023.10462804>.
- [20] Wu, D., et al. 2022. Detection of Camellia Oleifera Fruit in Complex Scenes by Using YOLOv7 and Data Augmentation. *Applied Sciences (Switzerland)* 12 (22): 11318. <https://doi.org/10.3390/app122211318>.