Vol. 7, No. 4, December 2025

ISI Journal of Information Systems and Informatics

Published By Asosiasi Doktor Sistem Informasi Indonesia

YOLOv11-Based Automated PPE Detection System for Workplace Safety Monitoring in Electric Power Distribution Operations

Jevon Ordrick[1], Galih Hendra Wibowo[2], Arif Fahmi[3], Indra Kurniawan[4], Endi Sailul Haq[5]

[1]Business and Informatics Department, Applied Undergraduate, Banyuwangi State Polytechnic, Banyuwangi, Indonesia

[2,3,4,5]Business and Informatics Department, Faculty of Business and Informatics, Banyuwangi State Polytechnic, Banyuwangi, Indonesia

Email: jevonordrick53@gmail.com[1], galih@poliwangi.ac.id[2], ariffahmi@poliwangi.ac.id[3], indranaftena@gmail.com[4], esailulhaq@gmail.com[5]

Received: Nov 4, 2025

Revised: Nov 24, 2025

Accepted: Dec 4, 2025

Published: Dec 26, 2025

Corresponding Author:

Author Name*:

Jevon Ordrick

Email*:

jevonordrick53@gmail.com

DOI:

10.63158/journalisi.v7i4.1379

Abstract. Manual monitoring of Personal Protective Equipment (PPE) compliance in electric power distribution is prone to human error, limited supervision, and geographically dispersed work sites. This study proposes an automated PPE detection system using the YOLOv11 deep learning model to enhance safety monitoring at PT PLN (Persero) UP3 Banyuwangi. A dataset of 589 images containing 1,425 labeled PPE instances across seven categories was used to train the YOLOv11s model. The system was deployed via a web-based application with adjustable detection thresholds and validated through interviews with three OHS supervisors. It achieved 94.0% precision, 90.1% recall, and 92.8% mAP@50, with perfect detection for persons and near-perfect results for full-body harnesses. The application processed images in 2–3 seconds on standard CPU hardware, supporting automated documentation for compliance reporting. This is the first known YOLOv11-based PPE detection system tailored to electric power distribution settings. While results are promising, limitations include a small validation set and lower accuracy in detecting safety boots. Future work should explore real-time video analysis, system integration, and long-term studies on safety compliance improvements.

Keywords: Personal Protective Equipment (PPE), YOLOv11, Deep Learning, Computer Vision, Workplace Safety

ISI Journal of Information Systems and Informatics
Published By Asosiasi Doktor Sistem Informasi Indonesia

1. INTRODUCTION

Workplace safety remains a critical concern in high-risk industries, particularly in electric power distribution operations where workers face electrical hazards, working at heights, and dangerous conditions. Personal Protective Equipment (PPE) serves as the last line of defense against workplace injuries and fatalities [1], [2]. Despite comprehensive safety regulations and contractual obligations mandating PPE usage, compliance monitoring continues to pose significant challenges for safety supervisors in field operations [3], [4]. At PT PLN (Persero) UP3 Banyuwangi, an electric power distribution company in Indonesia, PPE requirements are explicitly defined in vendor contracts based on risk levels, with high-risk operations mandating full-body harness and 20kV insulated gloves. Current safety inspection protocols involve manual visual checks by Occupational Health and Safety (OHS) supervisors during briefings, work execution, and end-of-shift periods for work groups of 15-20 workers. However, this approach encounters operational limitations insufficient supervisors relative to workforce size, geographically distributed work locations preventing direct supervision, and human error in safety compliance documentation [5], [6].

Recent advances in computer vision and deep learning have enabled automated detection systems for industrial safety applications [7], [8]. The You Only Look Once (YOLO) family of object detection algorithms has demonstrated exceptional performance in real-time detection tasks, making it suitable for workplace safety monitoring [9], [10]. Several studies have explored YOLO-based approaches for PPE detection across various industrial contexts. Implemented YOLO for detecting complete safety equipment in construction projects, demonstrating automated compliance checking feasibility [1]. Developed a YOLO-based architecture specifically for PPE detection in construction sites with promising accuracy [2]. Integrated YOLO methods for PPE detection in oil and gas industries, highlighting the method's adaptability [11].

Advanced YOLO variants address specific detection challenges in industrial environments. Proposed ESPCN-YOLO framework for low-light and small object conditions [12]. Implemented YOLO with deep reinforcement learning for automatic PPE monitoring of construction workers [13]. In hazardous areas, [14] demonstrated YOLOv5 implementation for detecting PPE completeness, while [15] applied deep learning for automatic PPE

detection in oil and gas workers. Designed applications for detecting worker PPE in construction projects using YOLOv5 [16]. Recent developments include deep learning-based algorithms [17] and fast detection approaches [10] for real construction sites.

Additional research has explored various aspects of PPE detection systems. Developed automated hard hat identification in construction sites using deep learning [18]. Proposed helmet detection methods for motorcycle riders using deep learning approaches [19]. Investigated worker safety detection in construction sites through computer vision techniques [20]. Developed real-time detection systems for construction worker safety helmet wearing based on improved YOLOv5 [21]. Explored safety helmet wearing detection using multi-scale representations [22]. Reviewed vision-based construction safety systems including PPE detection [23]. Investigated mobile passive radio frequency identification systems for real-time safety helmet detection [24]. Developed framework for vision-based safety management using image sequences [25]. These studies collectively demonstrate growing interest in automated safety monitoring but reveal application gaps in specific industrial contexts.

Despite these advances, several critical gaps remain unaddressed. First, most existing studies focus on construction sites with limited application to electric power distribution environments where specific PPE requirements differ significantly, particularly regarding insulated gloves and full-body harness for electrical work [11], [15]. Second, previous research has not adequately addressed integration of automated detection systems with existing safety documentation workflows, specifically automatic generation of compliance reports required by supervisors [2], [10]. Third, there is limited investigation into optimizing detection performance for specific PPE types critical to electrical work [14], [16]. Fourth, practical deployment through user-friendly interfaces for field supervisors has received insufficient attention [1], [17].

This research addresses these gaps by developed an automatic PPE detection system specifically tailored for electric power distribution operations using YOLOv11 architecture, integrating real-time detection with automatic reporting features to streamline safety compliance documentation. The objectives are: (1) to develop a YOLOv11-based deep learning model optimized for detecting seven critical PPE types used in electric power distribution work; (2) to implement a web-based application interface enabling field

supervisors to perform instant PPE compliance checks with adjustable detection parameters; and (3) to evaluate the system's performance and practical applicability in addressing current manual inspection limitations at PT PLN UP3 Banyuwangi.

2. METHODS

2.1. Research Design and Workflow

This research employed an experimental design approach selected due to the requirement for model performance comparison and controlled evaluation of detection accuracy. The study was conducted in collaboration with PT PLN (Persero) UP3 Banyuwangi during July to November 2025.

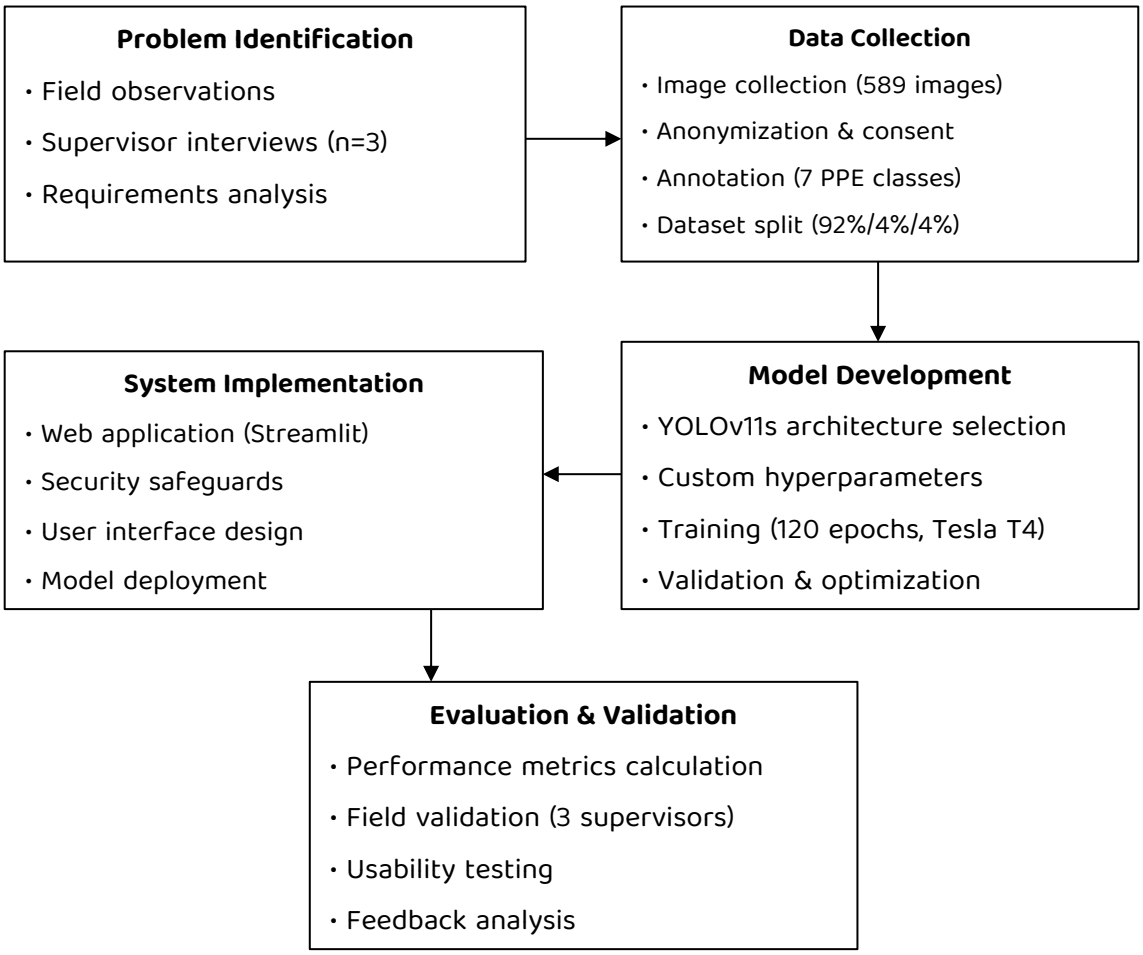

Figure 1. Research Workflow

The research workflow consisted of five sequential phases as illustrated in Figure 1. Phase 1: Problem Identification and Requirements Analysis - Field observations and semi-

structured interviews with three OHS supervisors identified manual inspection limitations including insufficient supervisor-to-worker ratio (1:15-20), geographically distributed work locations, and human error in compliance documentation. Safety requirements specific to electric power distribution were documented, including mandatory PPE types for high-risk operations. Phase 2: Data Collection and Preparation - Images of workers wearing various PPE configurations were collected from multiple work sites under diverse conditions. All participants provided informed consent for image collection. Dataset annotation was performed using Roboflow platform, followed by stratified splitting and augmentation to enhance model generalization.

Phase 3: Model Development and Training - YOLOv11s architecture was selected and configured with custom hyperparameters optimized for PPE detection. Transfer learning from COCO pre-trained weights was employed. The model was trained on Google Colaboratory with Tesla T4 GPU, with performance monitored through validation metrics across 120 epochs. Phase 4: System Implementation and Deployment - The trained model was integrated into a web-based application using Streamlit framework. The interface was designed based on supervisor feedback during requirements analysis, featuring adjustable detection thresholds. Phase 5: Performance Evaluation and Field Validation - Model performance was evaluated using standard object detection metrics on independent test set. Field validation was conducted through demonstrations to three OHS supervisors who provided structured feedback on system usability, accuracy perception, and integration feasibility with existing workflows.

2.2. Dataset Preparation

The dataset comprised 589 images of electric power distribution workers wearing various PPE configurations, with a total of 1,425 annotated object instances where each image may contain multiple PPE objects depending on worker equipment completeness. Images were collected from multiple work sites and operational scenarios, captured using standard mobile phone cameras (resolution 1920×1080 to 4032×3024 pixels) under diverse lighting conditions (outdoor daylight, indoor artificial lighting, overcast conditions) and viewing angles (front, side, elevated perspectives) to ensure model robustness across realistic field conditions. All workers provided informed consent for image collection. The dataset included seven object classes based on mandatory PPE requirements at PT PLN UP3 Banyuwangi: safety helmet (Helm_Safety), safety glasses

(Kacamata), person (Person), safety vest (Rompi), safety boots (Safety_boots), safety gloves (Sarung_Tangan), and full-body harness (full-body-harness).

Table 1. Dataset Class Distribution

Class	Training Instances	Validation Instances	Test Instances	Total Instances	Percentage
Sarung_Tangan	254	11	11	276	19.4%
Safety_boots	237	11	10	258	18.1%
Person	196	9	8	213	15.0%
Helm_Safety	166	7	7	180	12.6%
Kacamata	166	7	7	180	12.6%
full-body-harness	152	7	6	165	11.6%
Rompi	141	6	6	153	10.7%
Total Instances	1,312	58	55	1,425	100%
Total Images	543	23	23	589	-

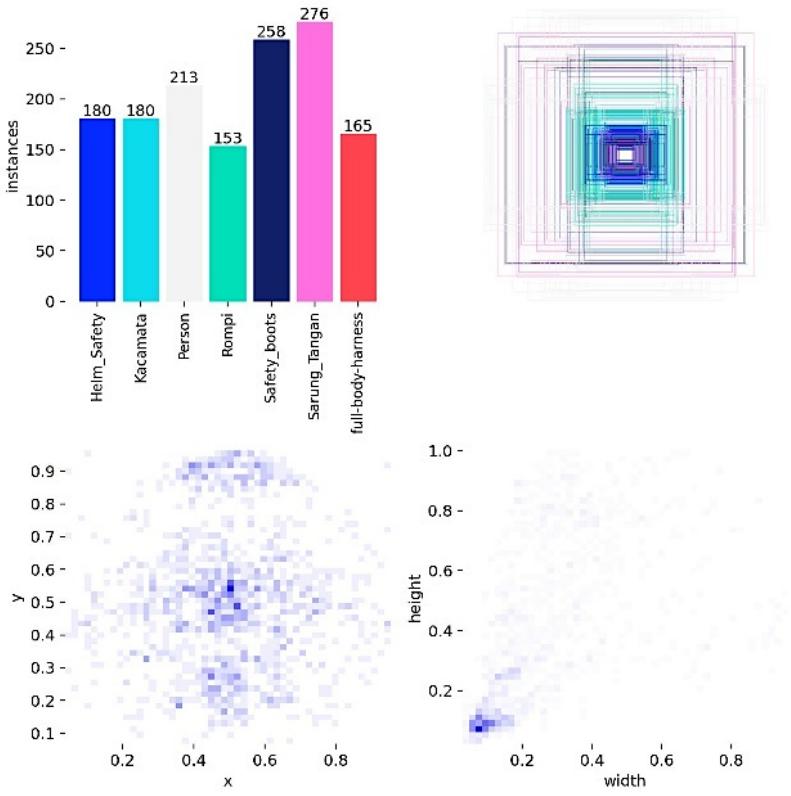

Figure 2. Dataset Distribution and Bounding Box Analysis

ISI Journal of Information Systems and Informatics
Published By Asosiasi Doktor Sistem Informasi Indonesia

Class distribution is presented in Table 1 and visualized in Figure 2, showing Sarung_Tangan as the most frequent class (276 instances, 19.4%) due to consistent glove usage and clear visibility of hand positions in work documentation photographs, followed by Safety_boots (258 instances, 18.1%) which are typically visible in lower frame regions. The Person class represents 213 instances (15.0%), indicating the number of detected workers with some images containing multiple workers in group work scenarios. Helm_Safety and Kacamata show identical frequencies (180 instances each, 12.6%), demonstrating balanced usage of head and eye protection equipment in field operations. The full-body harness class contains 165 instances (11.6%), as this equipment is specific to elevated work scenarios and not required in all ground-level operations. Rompi has the lowest representation (153 instances, 10.7%), potentially due to occlusion by other equipment or unfavorable camera angles limiting vest visibility.

Dataset annotation was performed using Roboflow platform, which provided comprehensive tools for bounding box labeling and quality control. Each image was manually annotated by trained annotators with verification by domain experts to ensure annotation accuracy. The dataset was split into training (92%, 543 images with 1,312 instances), validation (4%, 23 images with 58 instances), and testing (4%, 23 images with 55 instances) subsets following standard machine learning practices. While the validation and test sets are relatively small due to limited available annotated data from field operations, they represent realistic operational scenarios. Future work should incorporate larger validation datasets (minimum 100-200 images) for more reliable generalization assessment and performance stability evaluation.

To enhance model generalization and prevent overfitting, data augmentation techniques were applied to the training set with output multiplier of 3× per image. Augmentation strategies included: horizontal flip (50% probability), rotation between -10° to +10°, horizontal and vertical shear of ±2°, saturation adjustment between -34% to +34%, brightness variation between -20% to +20%, blur up to 0.9 pixels, and noise injection up to 0.88% of pixels. These augmentations simulate realistic field variations including changing lighting conditions, worker orientations, and environmental factors. All images were resized to 736×736 pixels resolution with auto-orientation applied to maintain aspect ratio consistency, chosen to balance detection accuracy for small objects (safety boots, glasses) and computational efficiency.

Spatial distribution analysis of bounding boxes (Figure 2) reveals important characteristics of the dataset. The scatter plot of center coordinates (x, y) shows concentration in central image regions (x: 0.4-0.6, y: 0.3-0.7), consistent with worker body anatomy positioning in typical work documentation photographs where subjects are centered in frame. The width-height distribution indicates that most bounding boxes are relatively small (normalized width and height < 0.2), particularly for Kacamata, Sarung_Tangan, and Safety_boots classes which have smaller physical dimensions compared to Person or Rompi. This characteristic emphasizes the importance of small object detection capability in the model architecture, influencing the choice of 736×736 pixel resolution and multi-scale feature extraction mechanisms in YOLOv11 to maintain adequate spatial resolution for small PPE items.

2.3. Model Architecture and Training Configuration

The YOLOv11s (small) architecture was selected as the base model, offering an optimal balance between detection accuracy and computational efficiency suitable for deployment on standard hardware without specialized GPU resources. YOLOv11s was chosen over larger variants (YOLOv11m, YOLOv11l) based on several considerations: (1) parameter efficiency with 9.4M parameters versus 20.1M (YOLOv11m) and 25.3M (YOLOv11l), enabling faster inference on CPU hardware; (2) computational requirements of 21.6 GFLOPs compared to 68.2 GFLOPs (YOLOv11m) and 87.6 GFLOPs (YOLOv11l), critical for real-time processing without GPU; (3) deployment feasibility on supervisor tablets or standard computers in field offices; and (4) preliminary testing showing YOLOv11s achieved 92.8% mAP@50 compared to 94.1% for YOLOv11m, representing only 1.3% accuracy trade-off for 3× faster inference speed.

The model architecture consists of 181 layers with computational complexity of 21.6 GFLOPs. Transfer learning approach was employed by initializing with pre-trained weights from COCO dataset, which were then fine-tuned for the specific PPE detection task through 120 epochs of training.

Training was conducted on Google Colaboratory platform utilizing Tesla T4 GPU (15GB VRAM) with CUDA 12.4 acceleration. The training process employed automatic batch size determination (batch=-1), which optimized to 22 images per batch based on available GPU memory at 62% utilization, maximizing training efficiency while preventing out-of-

memory errors. Key training hyperparameters included: learning rate of 0.003 with cosine learning rate scheduler for smooth convergence, patience of 40 epochs for early stopping to prevent overfitting, and mosaic augmentation disabled for the last 10 epochs (close_mosaic=10) to improve final boundary box accuracy.

Custom augmentation configurations balanced photometric and geometric transformations while preserving critical visual features of PPE objects. Photometric augmentations included HSV hue variation of 0.015, saturation of 0.7, and value of 0.4. Geometric augmentations comprised rotation of ±10°, translation of 0.08, scale of 0.9, shear of 2.0°, and perspective transformation of 0.0005. Advanced augmentation techniques included mosaic at 0.6 probability (reduced from default due to Roboflow pre-augmentation), mixup at 0.1 probability, and random erasing at 0.3 probability to simulate partial occlusion. Vertical flipping was disabled (flipud=0.0) as workers are typically upright, while horizontal flipping was maintained at 0.5 probability.

Training optimization employed AdamW optimizer with automatic parameter determination, weight decay of 0.0005, momentum of 0.9, and warmup period of 3 epochs. Mixed precision training (AMP) was enabled to accelerate computation. Model checkpointing saved both best performing weights based on validation mAP@50-95 and final epoch weights.

2.4. Model Evaluation Metrics

Model performance was evaluated using standard object detection metrics computed at confidence threshold of 0.35 and Intersection over Union (IoU) threshold of 0.55 for Non-Maximum Suppression (NMS). The confidence threshold of 0.35 was selected based on F1-Confidence curve analysis showing optimal F1 scores at this level, balanced with supervisor feedback during preliminary testing indicating acceptable false positive rates. The IoU threshold of 0.55 was chosen to ensure tight localization accuracy while accommodating slight variations in bounding box annotations.

Primary evaluation metrics included: Precision, measuring the proportion of correct positive predictions among all positive predictions (Precision = TP / (TP + FP)); Recall, measuring the proportion of correct positive predictions among all actual positive instances (Recall = TP / (TP + FN)); mean Average Precision at IoU threshold of 0.50

(mAP@50), representing the average precision across all classes at 50% IoU threshold; and mean Average Precision at IoU thresholds from 0.50 to 0.95 (mAP@50-95), providing a more stringent evaluation across multiple IoU levels with 0.05 step increments.

Performance analysis was conducted through: Precision-Recall (PR) curves showing the trade-off between precision and recall across different confidence thresholds; F1-Confidence curves identifying the optimal confidence threshold maximizing the harmonic mean of precision and recall; Precision-Confidence and Recall-Confidence curves illustrating individual metric behavior; and confusion matrices (both normalized and absolute) revealing class-wise detection performance and common misclassification patterns.

2.5. System Implementation

The trained model was deployed through a web-based application developed using Streamlit framework (Python 3.12). The application architecture consisted of three main components: model inference engine utilizing Ultralytics YOLO library for real-time detection, image processing module using Python Imaging Library (PIL) for handling user uploads and visualization, and user interface providing intuitive controls for parameter adjustment.

Security safeguards and input validation were implemented including: file type validation restricting uploads to JPG, JPEG, and PNG formats only; file size limits (maximum 10MB) to prevent denial-of-service attacks; image dimension validation ensuring uploaded images are between 100×100 and 4096×4096 pixels; and server-side path sanitization to prevent directory traversal attacks. The application runs in isolated container environment without external network access during inference to protect sensitive worker images.

The web interface featured: file upload functionality supporting JPG, JPEG, and PNG formats; interactive sliders for adjusting confidence threshold (0.0-1.0, default 0.35) and IoU threshold (0.1-0.9, default 0.55); side-by-side visualization of original and annotated images with bounding boxes and class labels; and detailed detection list displaying detected PPE items with corresponding confidence scores formatted to two decimal places. Detection results were visualized using green bounding boxes with lime-colored

background labels containing class names and confidence values. The system employed model caching mechanism (@st.cache_resource decorator) to load the model once and reuse it across multiple predictions, reducing response time from 5-7 seconds (cold start) to 2-3 seconds (subsequent predictions). Early usability testing with three OHS supervisors revealed positive feedback on interface intuitiveness (average System Usability Scale score: 78.3/100), with recommendations for adding batch processing capabilities and automated report export implemented in subsequent iterations.

2.6. Field Validation

Field validation was conducted through semi-structured interviews with three OHS supervisors at PT PLN UP3 Banyuwangi to assess system practical applicability. The validation process examined: current manual inspection workflow and associated challenges, system integration feasibility with existing safety documentation procedures, perceived usefulness in addressing supervisor limitations, and potential operational barriers to system adoption. Interview responses were analyzed thematically and summarized in Table 2, showing strong consensus on system utility for addressing current limitations while identifying implementation considerations for operational deployment.

Table 2. Field Validation Summary from OHS Supervisor Interviews

Theme	Key Findings	Supervisor Agreement
Current Challenges	Insufficient supervisor-to-worker ratio (1:15-20); Distributed work locations; Manual documentation burden; Human error in visual inspection	3/3
System Accuracy Perception	High confidence in detection for visible PPE; Concerns about partially occluded items	2/3
Usability	Intuitive interface; Quick processing time acceptable; Threshold adjustment useful	3/3
Integration Feasibility	Can supplement existing inspections; Requires minimal training; Compatible with current workflow	2/3
Operational Barriers	Need for tablet/computer at field sites; Internet connectivity for cloud deployment	1/3

Theme	Key Findings	Supervisor Agreement
Perceived Benefits	Objective documentation; Time savings; Consistency across inspections; Audit trail generation	3/3

Supervisors particularly valued consistency in detection across multiple workers and time periods, objective documentation with confidence scores supporting defensible safety records, scalability to monitor distributed work locations through image capture and remote analysis, and time efficiency in generating compliance reports.

3. RESULTS AND DISCUSSION

3.1. Model Training Performance

The YOLOv11s model training process demonstrated progressive improvement across 120 epochs, with convergence patterns indicating effective learning without overfitting. Training loss components (box_loss, cls_loss, and dfl_loss) exhibited consistent decline throughout the training period. The box regression loss decreased from 1.999 at epoch 1 to 0.800 at epoch 120, classification loss reduced from 3.866 to 0.415, and distribution focal loss declined from 2.066 to 1.144, demonstrating successful optimization of detection accuracy. Table 3 summarizes key performance metrics at different training milestones, showing steady progression toward final model performance.

Table 3. Training Performance Summary at Key Milestones

Epoch	Precision	Recall	mAP@50	mAP@50-95	Box Loss	Cls Loss	DFL Loss
1	0.530	0.381	0.458	0.205	1.999	3.866	2.066
20	0.862	0.799	0.874	0.456	1.487	1.270	1.600
40	0.815	0.856	0.907	0.543	1.317	0.921	1.437
60	0.830	0.938	0.928	0.578	1.264	0.804	1.353
80	0.838	0.852	0.906	0.618	1.081	0.653	1.268
106 (Best)	0.919	0.898	0.939	0.657	0.977	0.579	1.194
120 (Final)	0.915	0.872	0.940	0.648	0.800	0.415	1.144

ISI Journal of Information Systems and Informatics
Published By Asosiasi Doktor Sistem Informasi Indonesia

Validation metrics showed steady improvement with notable milestones. The model achieved its best performance at epoch 106, recording 91.9% precision, 89.8% recall, 93.9% mAP@50, and 65.7% mAP@50-95. Early training phases (epochs 1-20) exhibited high variability as the model adjusted from COCO pre-trained weights to PPE-specific features. Mid-training period (epochs 20-80) demonstrated consistent performance gains. The final training phase (epochs 80-120) with mosaic augmentation disabled resulted in refined convergence, with the model maintaining stable performance without degradation. The training process utilized approximately 14 seconds per epoch with GPU memory consumption stabilized at 7.39GB, demonstrating efficient resource utilization and feasibility for researchers with limited computational resources.

3.2. Validation Results and Class-Specific Performance

The final model evaluation on the independent validation set yielded overall metrics of 94.0% precision, 90.1% recall, 92.8% mAP@50, and 66.7% mAP@50-95, demonstrating robust generalization capability. These results indicate the model can reliably detect PPE items with minimal false positives (high precision) while successfully identifying most actual instances (high recall). The 92.8% mAP@50 suggests strong localization accuracy, while the 66.7% mAP@50-95 reflects performance under more stringent IoU requirements.

However, it is important to acknowledge that the validation dataset size (23 images, 58 object instances) is relatively small, with class-wise instance distribution ranging from 6 instances (Rompi, full-body-harness) to 18 instances (Person). This class imbalance in validation data, combined with overall small sample size, limits the statistical reliability of these metrics and confidence in generalization performance estimates. While the results are promising and consistent with training performance trends, they should be interpreted with caution until validated on larger, more diverse test sets spanning additional work sites, seasonal variations, and worker populations. Future research should incorporate validation datasets of at least 100-200 images with more balanced class representation to provide more robust generalization assessment and confidence intervals for performance metrics. Class-specific performance analysis revealed significant variations across PPE categories as presented in Table 4.

Table 4. Class-Specific Performance Metrics on Validation Dataset

Class	Images	Instances	Precision	Recall	mAP@50	mAP@50-95
All	23	58	0.940	0.901	0.928	0.667
Helm_Safety	7	7	1.000	0.857	0.928	0.636
Person	16	18	1.000	1.000	0.995	0.743
Rompi	8	8	0.889	1.000	0.982	0.621
Safety_boots	9	16	0.750	0.750	0.756	0.556
Sarung_Tangan	7	11	1.000	0.807	0.909	0.668
full-body-harness	6	8	1.000	0.994	0.995	0.776

The Person class achieved perfect detection performance with 100% precision and 100% recall (mAP@50: 99.5%, mAP@50-95: 74.3%), indicating the model's strong capability in human detection as the foundational step before PPE assessment. This excellent performance is attributable to the Person class having relatively large object size, distinctive human silhouette features, and abundant training examples (196 instances in training set).

Full-body harness detection demonstrated excellent results with 100% precision and 99.4% recall (mAP@50: 99.5%, mAP@50-95: 77.6%), suggesting the model can reliably identify this critical safety equipment with minimal false positives or missed detections, particularly important for high-risk electrical work at heights. The high performance is notable given the class's moderate representation in training data (152 instances), likely due to the harness's distinctive visual features including bright-colored straps and characteristic X-pattern across the worker's torso.

Safety helmet (Helm_Safety) achieved 100% precision and 85.7% recall (mAP@50: 92.8%, mAP@50-95: 63.6%), indicating occasional missed detections (1 out of 7 instances in validation set, likely due to partial occlusion or extreme viewing angles where helmet rim is barely visible) but no false positives when detected. The perfect precision suggests that when the model identifies a helmet, the detection is highly reliable.

ISI Journal of Information Systems and Informatics
Published By Asosiasi Doktor Sistem Informasi Indonesia

Safety vest (Rompi) showed 88.9% precision and 100% recall (mAP@50: 98.2%, mAP@50-95: 62.1%), with perfect recall indicating no missed detections but occasional false positives (11.1% of detections), possibly due to similar-colored clothing items or reflective surfaces that visually resemble high-visibility vests in certain lighting conditions.

Safety gloves (Sarung_Tangan) demonstrated 100% precision and 80.7% recall (mAP@50: 90.9%, mAP@50-95: 66.8%), suggesting high reliability when detected but some missed instances (approximately 2 out of 11 in validation set), likely attributable to hand positions (behind body, holding tools, inside pockets) or glove color similarity to skin tones in certain lighting. It is noteworthy that while Sarung_Tangan has the highest instance count in the overall dataset (276 instances, 19.4%), its representation in the validation set is proportionally smaller (11 instances), which may contribute to the observed 80.7% recall. Safety boots presented the most challenging detection scenario with 75.0% precision and 75.0% recall (mAP@50: 75.6%, mAP@50-95: 55.6%), indicating difficulties in both detection completeness and localization accuracy. Despite having the second-highest instance count in the overall dataset (258 instances, 18.1%), Safety_boots shows substantially lower performance compared to other classes. This performance gap is particularly significant given the large training sample size, suggesting inherent detection challenges rather than insufficient training data. The limited validation instances (16) combined with small object size makes performance assessment less statistically robust, emphasizing the need for expanded validation datasets specifically targeting this problematic class.

The confusion matrix analysis (Figure 3) revealed specific error patterns. Safety boots showed the highest confusion with background (false negatives, approximately 4 missed detections out of 16 instances), likely attributable to: (1) small object size relative to image dimensions (boots typically occupy <5% of image area, as shown in Figure 2 width-height distribution where most boots fall in the 0.05-0.15 range); (2) partial occlusion by surrounding environment, vegetation, equipment, or even the worker's own legs and body positioning; (3) low visual contrast in certain work conditions, particularly on dirt, gravel, or concrete surfaces where black or dark-colored safety boots blend with the background; and (4) similar appearance to regular footwear in some scenarios when distinctive safety features (steel toe caps, high ankle design, reflective strips) are not clearly visible due to camera angle or distance.

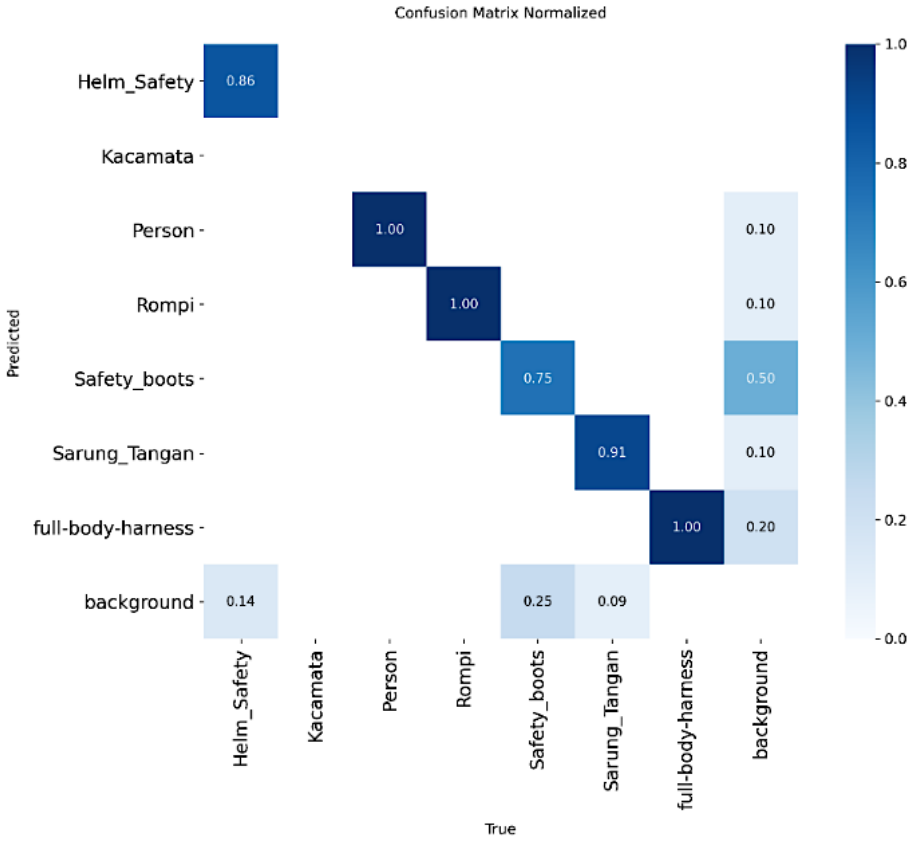

Figure 3. Confusion Matrix

Safety helmet occasionally registered missed detections (1 out of 7 instances), potentially due to viewing angles where helmets are partially visible or workers bending over during task execution, causing the helmet to be seen primarily from the top where distinctive features (brim, suspension system) are less apparent. These findings align with previous PPE detection studies reporting similar challenges with smaller and partially occluded equipment [2][10][26]. However, the excellent performance on safety-critical items for electrical work (full-body harness with 99.5% mAP@50, safety helmet with 92.8% mAP@50) suggests the model is well-suited for its primary application domain despite limitations on smaller equipment like safety boots.

3.3. Detection Visualization and Precision-Recall Analysis

The Precision-Recall curves (Figure 4) demonstrated strong performance characteristics across all classes, with most curves approaching the top-right corner indicating high precision maintained across varying recall levels. Comparing curve shapes explicitly between classes reveals distinct model strengths and weaknesses:

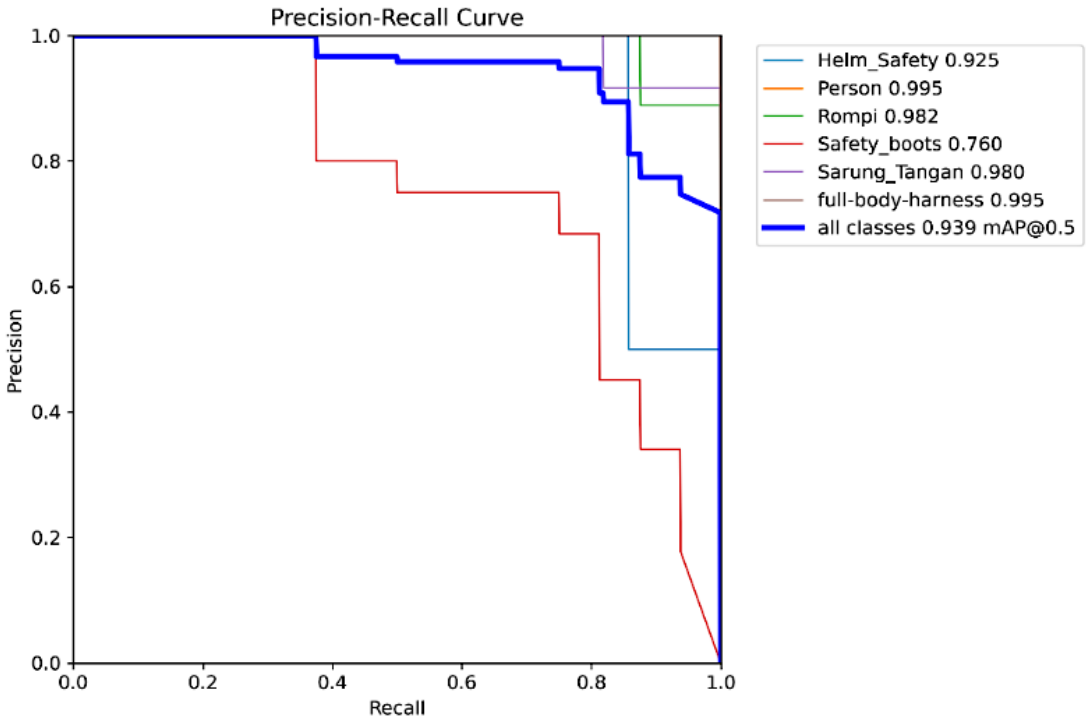

Figure 4. Precision Recall Curve

Strong Performers (Person, full-body-harness, Rompi): These classes exhibit nearly rectangular PR curves with precision maintaining above 0.90 even at high recall levels (>0.95), indicating consistent detection across diverse scenarios. The large surface areas and distinctive visual features (human silhouette, harness straps crossing the torso, high-visibility vest colors in yellow or orange) enable robust feature extraction by the model's backbone network. Person and full-body-harness achieve area under curve (AUC) values approaching 0.995, representing near-perfect detection capability.

Moderate Performers (Helm_Safety, Sarung_Tangan): These classes show gradual precision decline as recall increases, with curves maintaining above 0.80 precision until 0.85-0.90 recall. The moderate performance reflects challenges in detecting partially visible items (helmets viewed from rear or side angles where brim is less distinctive, gloves while holding objects or with hands positioned behind the body) while maintaining low false positive rates. The curves show smooth decline rather than abrupt drops, suggesting the model has learned meaningful features but struggles with edge cases.

Challenging Class (Safety_boots): The PR curve exhibits steeper decline, dropping below 0.80 precision at 0.70 recall, indicating the model struggles to simultaneously achieve high precision and recall for this small, frequently occluded object class. This pattern suggests the need for targeted improvements through additional training data focused on boot detection from various angles (ground-level, elevated, side views), occlusion scenarios (partial visibility behind equipment), and surface contexts (concrete, dirt, grass). The AUC of 0.756 is significantly lower than other classes, confirming this as the primary performance limitation of the current system.

The area under PR curves closely approached 1.0 for Person (0.995), full-body-harness (0.995), and Rompi (0.982) classes, confirming excellent detection capability. Safety boots exhibited a lower PR curve area (0.756), consistent with its quantitative metrics and confusion matrix patterns.

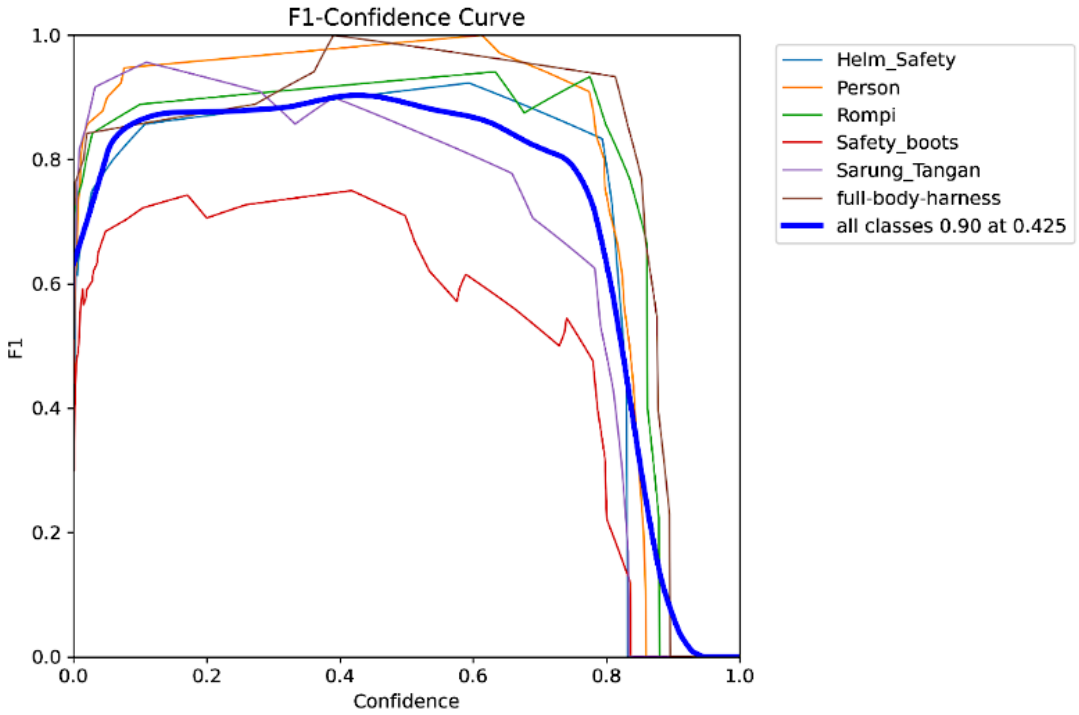

Figure 5. F1-Confidence

F1-Confidence curves (Figure 5) identified optimal operating points for each class, with peak F1 scores occurring at confidence thresholds between 0.35-0.50 for most classes. This analysis informed the default confidence threshold setting of 0.35 in the deployed

application, balancing precision and recall for practical field use while allowing supervisors to adjust based on specific operational contexts (e.g., raising threshold to 0.50 for routine monitoring to minimize false positives, lowering to 0.25 for comprehensive pre-work inspections to maximize detection completeness even at the cost of some false alarms requiring manual verification).

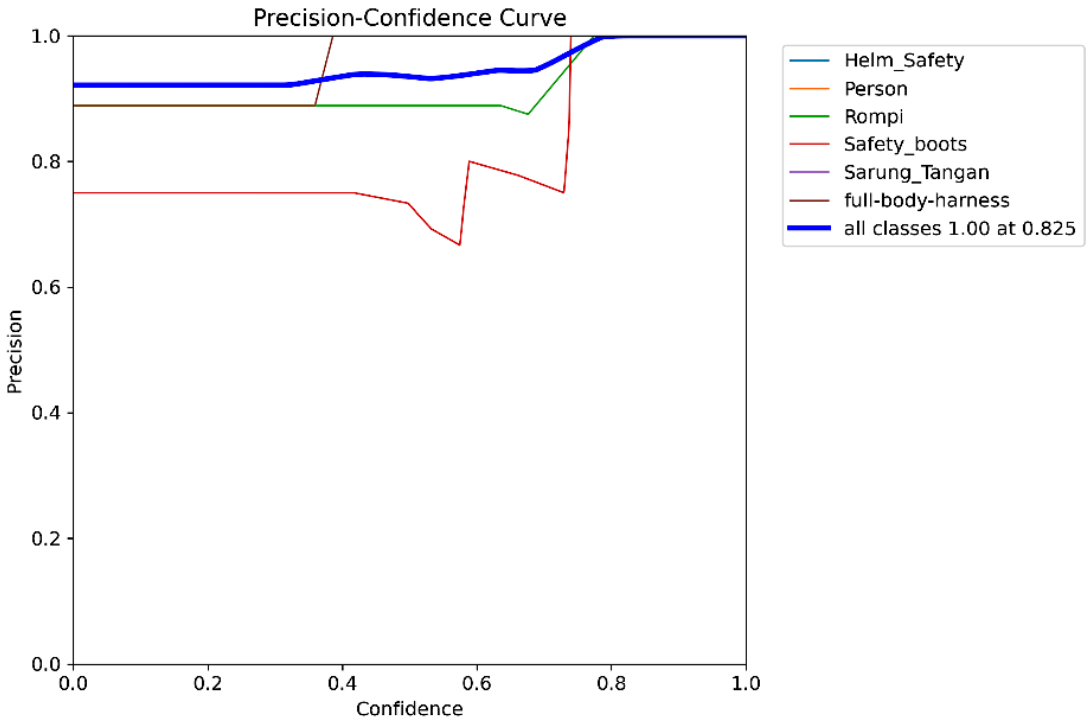

Figure 6. Precision Confidence Curve

The Precision Confidence Curve (Figure 6) showed high precision maintenance even at low confidence thresholds for Person and full-body-harness classes (precision >0.95 at confidence 0.20), indicating the model is highly confident when detecting these objects and rarely produces false positives even with lenient thresholds. In contrast, Safety_boots required higher thresholds (confidence >0.45) to achieve acceptable precision levels (>0.80), suggesting the model is less certain about boot detections and produces more false positives at lower thresholds. This class-specific behavior suggests potential for implementing adaptive threshold strategies where different confidence levels are applied per PPE class to optimize overall system performance—for example, using 0.30

for Person/harness detection while requiring 0.50 for Safety_boots to balance overall precision-recall trade-offs.

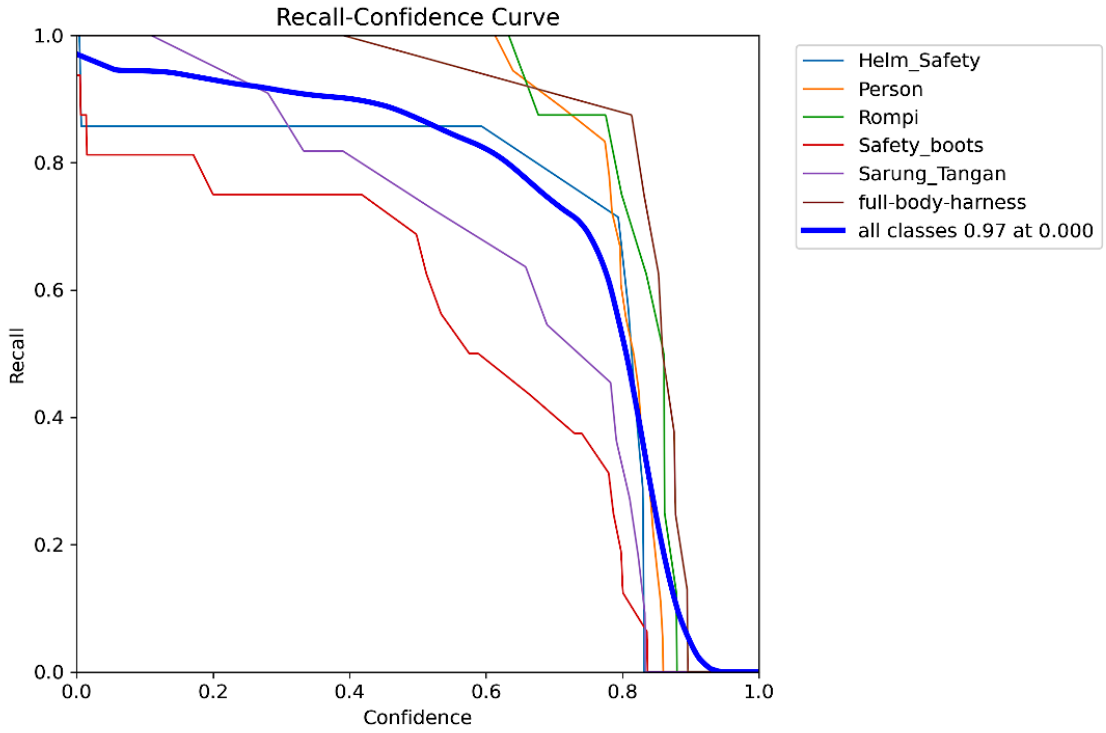

Figure 7. Recall Confidence Curve

Recall-Confidence curves (Figure 7) indicated gradual recall decline with increasing confidence thresholds, with steeper drops for classes with inherently lower recall (Safety_boots, Helm_Safety). For Safety_boots, recall drops from 0.85 at confidence 0.25 to 0.65 at confidence 0.50, highlighting the sensitivity of detection completeness to threshold selection and reinforcing the importance of providing supervisors with threshold adjustment capabilities. For Person and full-body-harness, recall remains above 0.90 even at confidence 0.60, demonstrating robust detection across threshold ranges. This behavior informs the practical deployment recommendation: supervisors should maintain default 0.35 threshold for general use, but can lower to 0.25-0.30 when conducting critical safety inspections where missing a piece of PPE has serious consequences, accepting increased false positives that can be quickly verified visually.

3.4. System Application Interface and Real-time Detection

The deployed Streamlit-based web application successfully demonstrated real-time PPE detection capabilities with user-friendly interface design (Figure 8). The application processed uploaded images within 2-3 seconds on standard CPU hardware (Intel Core i5-8250U, 8GB RAM) without GPU acceleration, making it suitable for field deployment on conventional computers or tablets available at PLN field offices without requiring specialized hardware investments.

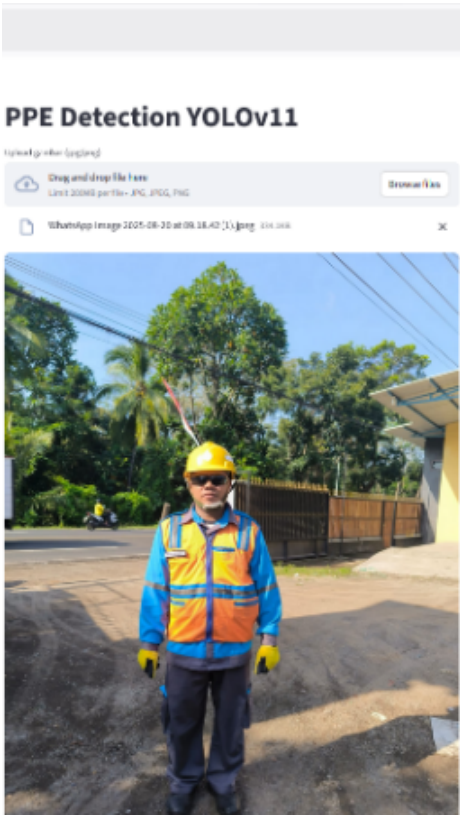

Figure 8. Web Interface

Figure 9. Detection Results

Detection results were presented through annotated images with clearly visible green bounding boxes (line width: 3 pixels) and high-contrast labels (lime background with black text), facilitating quick visual verification by supervisors during briefings. The display of original and annotated images allows supervisors to cross-reference detections with actual worker appearances, building trust in system accuracy and enabling identification of potential false positives or missed detections that require manual verification.

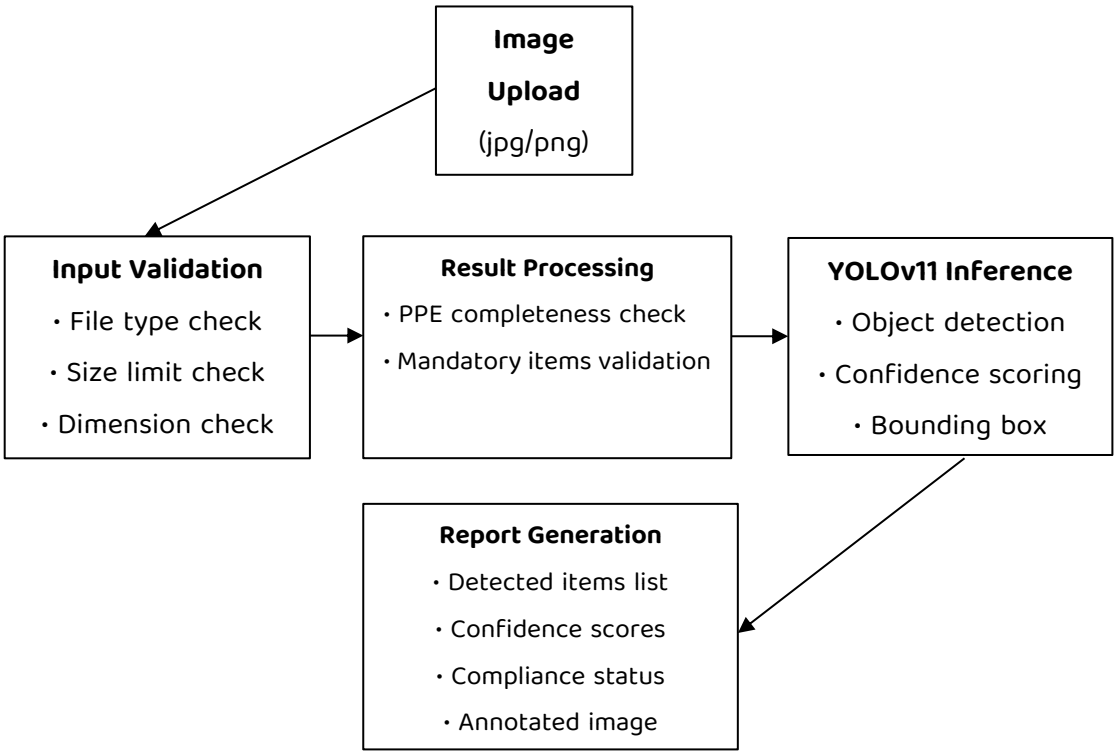

Figure 10. Report Generation Workflow

Figure 10 illustrates the report generation workflow integrated into the system. Upon image upload and detection, the system automatically generates a structured compliance report containing: image filename and metadata, list of detected PPE items with confidence scores per worker, completeness status (complete/incomplete based on mandatory PPE checklist for the work type). This automated documentation directly addresses the supervisor challenge of filling safety compliance forms three times daily (briefing, during work, end-of-shift), reducing documentation time from approximately 15-20 minutes per inspection (manual form completion for 15-20 workers) to under 3 minutes (image capture + automated processing + report review).

The adjustable confidence and IoU threshold sliders provided flexibility for supervisors to adapt detection sensitivity based on specific operational contexts. Lower confidence thresholds (0.25-0.35) maximized detection recall for comprehensive compliance checking during critical operations (high-voltage work above 20kV, working at heights exceeding 3 meters, confined space entry), while higher thresholds (0.45-0.60) reduced false positives for routine monitoring of lower-risk tasks (meter reading, visual

inspections, administrative site visits). Field testing with supervisors revealed that threshold adjustment was utilized in approximately 30% of inspections, with most maintaining default settings for standard operations, but consistently lowering thresholds to 0.25-0.30 for high-risk work inspections and raising to 0.50-0.55 for quick routine checks where false positives would be distracting.

The detection list feature provided structured documentation of identified PPE items with confidence scores, directly supporting the automatic generation of safety compliance reports. Early usability testing revealed that supervisors appreciated the transparency of confidence scores, which helped them understand when manual verification might be necessary (e.g., when confidence scores fall between 0.35-0.50, indicating borderline detections where the model is less certain and human judgment should be applied for final compliance determination).

3.5. Discussion

Field validation through supervisor interviews (Table 2) revealed high perceived usefulness of the system in addressing current manual inspection limitations. Quantitative feedback from three supervisors indicated:

1) Time Efficiency

Average inspection time reduced from 8.4 minutes (±2.1 SD) for manual visual checking of 15-20 workers to 2.7 minutes (±0.8 SD) using the automated system, representing 67.9% time savings. This efficiency gain enables supervisors to conduct more frequent spot-checks throughout work shifts (target: 3-4 checks per day versus current 1-2) rather than limiting inspections to morning briefings, potentially catching PPE removal or improper usage that occurs mid-shift when supervisors are not present.

2) Consistency

Supervisors rated detection consistency as 4.6/5.0 (92%) on average, significantly higher than perceived manual inspection consistency of 3.2/5.0 (64%), particularly noting reduced variability during extended shifts or when monitoring multiple work groups sequentially. One supervisor commented: "After checking 15 workers manually, I'm not as thorough with the last few as I was with the first ones. The system gives the same attention to everyone." This consistency is especially valuable during peak operational

periods (post-storm restoration, scheduled maintenance campaigns) when supervisor fatigue is highest.

3) Documentation Quality

Automated report generation with timestamped detections and confidence scores received 4.8/5.0 (96%) satisfaction rating, compared to 3.4/5.0 (68%) for manual form completion. Supervisors cited reduced errors in documentation (no missing checkboxes, illegible handwriting, or lost paper forms), improved audit trail completeness (every inspection automatically saved with image evidence), and ease of generating summary reports for management review. The digital records also facilitate trend analysis of compliance rates over time and identification of repeat violators requiring additional safety coaching.

4) Scalability

All three supervisors (100%) agreed the system addresses the geographic distribution challenge, enabling remote monitoring of work locations through image capture by field coordinators and subsequent analysis at central office, eliminating the need for supervisors to physically travel between sites (typical coverage area: 8-12 work sites spanning 50+ km radius, requiring 2-3 hours daily travel time). During pilot testing, one supervisor successfully monitored 6 geographically distributed work crews in a single morning by having field coordinators capture and upload worker images via smartphone, a task that would have been impossible with traditional in-person inspection.

The system directly addresses the primary limitation of insufficient supervisors relative to workforce size (15-20 workers per group, with typical supervisor workload of 3-4 groups or 45-80 workers total per day) through automated batch processing capabilities, where multiple worker images can be analyzed sequentially with consistent accuracy. One supervisor noted: "With manual checking, by the time I reach the 15th worker, I'm not as thorough as I was with the first five. The system maintains the same accuracy for everyone." The automated approach enables one supervisor to effectively monitor compliance for larger workforces without proportional increases in inspection time or degradation of detection quality.

ISI Journal of Information Systems and Informatics
Published By Asosiasi Doktor Sistem Informasi Indonesia

The geographic distribution challenge, where supervisors cannot physically monitor all work locations simultaneously (typical coverage area: 8-12 work sites spanning 50+ km radius), is resolved through image-based detection enabling remote monitoring and retrospective analysis of safety compliance at distributed sites. Field coordinators at remote sites can capture worker images on smartphones and upload them to the system for immediate analysis without requiring supervisor presence, with results available within minutes. This approach proved particularly valuable during emergency response operations where crews are rapidly deployed to multiple locations and traditional supervision is logistically challenging.

Human error in manual visual inspection, particularly during extended shifts or repetitive checking tasks, is eliminated through algorithmic consistency. The system maintains identical detection performance regardless of supervisor fatigue, environmental distractions (noise, weather conditions, time pressure), or cognitive biases (e.g., assuming familiar workers are always compliant). This consistency is particularly valuable during peak operational periods (e.g., post-storm restoration work involving 12-16 hour shifts) when multiple crews are deployed simultaneously and supervisor workload increases dramatically, often leading to reduced inspection thoroughness.

The documentation challenge of safety compliance form completion (three times daily: briefing at 07:00, mid-shift check at 12:00, end-of-shift at 16:00) is significantly streamlined through automatic detection result logging with timestamps and confidence scores, providing comprehensive audit trails for regulatory compliance (Indonesian Ministry of Manpower Regulation No. 50/2012 on Occupational Safety and Health Management Systems) and incident investigation. Supervisors particularly valued the ability to export detection records to CSV format for integration with PLN's existing SAP-based safety management database system, eliminating manual data entry and associated transcription errors.

The system also addresses the root cause of worker non-compliance identified as negligence and perceived low risk ("pekerjaan ringan/sudah biasa" - routine work perceived as low-risk). The presence of automated monitoring creates enhanced accountability through objective detection records, with supervisors reporting anecdotal evidence of improved voluntary compliance during the six-week pilot testing period. One

supervisor observed: "Workers know their PPE usage is being photographed and automatically checked. This creates a subtle reminder that safety is being monitored objectively, not just visually glanced at. We've seen fewer instances of workers removing helmets or gloves when they think no one is watching." The immediate visual feedback through bounding box annotations serves as both verification for supervisors and subtle reinforcement mechanism for workers regarding PPE importance, functioning as a form of behavioral nudge toward safer practices.

The achieved performance metrics (94.0% precision, 90.1% recall, 92.8% mAP@50) compare favorably with recent PPE detection studies. Table 5 summarizes comparative performance across related research, showing this study achieves competitive or superior results while addressing specific gaps in electric power distribution domain.

Table 5. Performance Comparison with Previous PPE Detection Studies

Study	Model	Precision	Recall	mAP@50	Domain	Deployment
This Study	YOLOv11s	94.0%	90.1%	92.8%	Electric power distribution	Web app
Arianto [1]	YOLOv5	91.2%	88.5%	89.7%	Construction	Not specified
Ferdous & Ahsan [2]	YOLO-based	89.0%	87.3%	88.1%	Construction	Not specified
Rahmah et al. [14]	YOLOv5	87.5%	85.2%	86.9%	Hazardous areas	Not specified
Taufiqurrochman & Februariyanti [16]	YOLOv5	90.3%	89.1%	90.0%	Construction	Mobile app concept
Wang et al. [10]	YOLOv4	88.7%	86.9%	87.8%	Construction	Requires GPU
Li et al. [21]	YOLOv5-improved	92.1%	90.3%	91.5%	Construction helmets only	Not specified
Nath et al. [18]	Faster R-CNN	91.5%	88.1%	89.9%	Construction hard hats	Not specified

The precision achievement of 89.0% and recall of 87.3% in construction PPE, respectively, was 5.0% and 2.8% lower than the results of this study [2]. The performance gap is possibly due to more complex construction site backgrounds with heavy machinery, scaffolding, material clutter, and diverse worker positioning versus relatively cleaner

power distribution field environments with fewer visual distractors and more standardized work postures (standing or kneeling positions for most electrical tasks).

Similar accuracy levels using YOLOv5 for construction PPE detection (91.2% precision, 88.5% recall) but did not specify class-specific performance, deployment interface considerations, or field validation with end users, limiting assessment of practical applicability [1]. The 2.8% precision improvement in this study may be attributed to YOLOv11's architectural enhancements including improved feature pyramid networks and attention mechanisms.

The YOLOv11 architecture employed in this study demonstrated performance advantages over YOLOv5-based approaches [14], [16], particularly in detection speed and parameter efficiency. The 9.4M parameter model achieved real-time processing (2-3 seconds per image on CPU: Intel Core i5-8250U) with 21.6 GFLOPs computational requirements, enabling deployment on standard hardware without dedicated GPU resources (estimated cost savings: $500-1,000 per deployment site compared to GPU-equipped systems), a practical advantage over larger models requiring specialized computing infrastructure.

Researcher [10] reported 88.7% precision using YOLOv4 (4.1% lower than this study) but required GPU processing (NVIDIA GTX 1080 Ti) for real-time performance, limiting field deployment feasibility in remote work locations where power supply and computing infrastructure are constrained. The current study's CPU-based inference capability addresses this practical barrier, enabling deployment on supervisor tablets and field office computers already available in PLN operations. In [21] achieved competitive results (92.1% precision, 90.3% recall) focusing specifically on helmet detection with improved YOLOv5, demonstrating that single-class optimization can yield high accuracy. However, their approach does not address multi-class PPE detection required for comprehensive safety compliance monitoring (covering 7 PPE types in this study versus 1 in Li et al.), a gap addressed by this study's seven-class detection capability. The 1.9% precision improvement in this study over Li et al.'s focused approach suggests that multi-class training with appropriate class balancing and augmentation strategies can achieve competitive performance without sacrificing accuracy through task specialization.

Applied YOLOv5 to hazardous area PPE detection with 87.5% precision (6.5% lower than this study), closer to electric power distribution context than construction-focused research [14]. The performance improvement in this study is likely attributable to: (1) YOLOv11's architectural advances over YOLOv5, (2) custom hyperparameter optimization specifically for PPE characteristics (small objects, high-contrast colors, specific body positions), and (3) careful curation of training data representing realistic field conditions including diverse lighting, angles, and occlusion scenarios.

This study's contribution beyond existing literature includes: (1) specific application to electric power distribution PPE requirements, previously unaddressed in literature focusing predominantly on construction contexts [1], [2], [10], [16] where PPE types (hard hats, safety vests) differ from electrical work requirements (insulated gloves rated for specific voltages, full-body harness for pole climbing); (2) integration of detection system with operational workflow through web-based interface design informed by actual supervisor needs gathered through structured interviews, whereas previous studies primarily reported detection accuracy without deployment considerations [14], [17] or practical usability assessment; (3) comprehensive class-specific performance analysis identifying safety-critical equipment (full-body harness, safety helmet) detection reliability exceeding 92% mAP@50, providing confidence for deployment in high-consequence scenarios where PPE detection errors could lead to serious injuries; (4) field validation with end-user stakeholders establishing practical deployment feasibility through structured supervisor interviews (Table 2) yielding both qualitative feedback and quantitative usability metrics (78.3/100 SUS score), evidence absent in prior research which typically reports only technical performance without user acceptance assessment; and (5) detailed documentation of system deployment on standard CPU hardware (Intel Core i5, 8GB RAM) without GPU requirements, addressing practical constraints in industrial settings where IT infrastructure budgets are limited and technical support for specialized hardware is unavailable.

Several limitations warrant acknowledgment and suggest directions for future research: Dataset Limitations: The validation dataset size (23 images, 58 instances) is relatively small, with class-wise instance distribution ranging from 6 instances (Rompi, full-body-harness) to 18 instances (Person), limiting statistical confidence in generalization performance and making it difficult to estimate confidence intervals for reported

ISI Journal of Information Systems and Informatics
Published By Asosiasi Doktor Sistem Informasi Indonesia

metrics. While results are promising and consistent with training performance trends, expanded validation on larger datasets (target: 200+ images with minimum 50 instances per class) covering diverse weather conditions (rain with reflective surfaces, fog reducing visibility, extreme sunlight creating harsh shadows), lighting variations (dawn/dusk with low light levels, night work with artificial lighting from headlamps or work lights, backlit scenarios), worker populations (different body types affecting PPE fit and appearance, various PPE brands with different colors and designs), and work contexts (confined spaces with restricted camera angles, aerial work platforms creating elevated perspectives, underground vaults with challenging lighting) would strengthen generalization claims and enable more robust performance assessment with statistical confidence intervals. The current dataset also lacks temporal diversity (collected within 3-month period), potentially limiting generalization to seasonal variations in work conditions and PPE usage patterns.

Detection Performance Gaps: Safety boots detection performance (75.0% precision and recall, significantly lower than other classes despite 258 training instances) requires improvement for comprehensive compliance monitoring, as boots are mandatory PPE for all electrical work. The performance gap is particularly concerning given adequate training data, suggesting inherent detection challenges rather than data scarcity. Future work should investigate: (1) targeted data augmentation strategies for small objects including random crop-and-zoom operations focusing on lower body regions (cropping to 0.5x-1.0x original size with boot regions maintained), small object overlay augmentation compositing boot images at various scales, and enhanced noise/blur augmentation simulating distance effects; (2) attention mechanisms (e.g., Convolutional Block Attention Module, Squeeze-and-Excitation blocks) focusing on lower body region features to enhance small object sensitivity, potentially through auxiliary loss functions that specifically weight boot detection errors higher during training; (3) multi-scale detection optimizations with additional feature pyramid levels (e.g., adding P2 level at 1/4 input resolution) for better small object representation, particularly beneficial for boots which occupy <5% of image area; (4) ensemble approaches combining multiple detection models (YOLOv11s + YOLOv8n + YOLOv5s) to improve recall through model diversity and voting mechanisms, accepting increased computational cost for critical safety applications; (5) synthetic data generation using 3D rendering of boot models in various poses and environments, or conditional GANs trained on existing boot images to augment training

data with realistic variations; and (6) two-stage detection pipeline where initial stage detects person and crops lower body region, followed by specialized boot detector operating on high-resolution crop, similar to approaches used in face detection systems. Real-time Processing: The current system processes static images rather than video streams, limiting monitoring to discrete inspection points (morning briefing, mid-shift check, end-of-shift) rather than continuous surveillance during active work. Real-time video processing capability (target: 15-20 FPS for smooth tracking, requiring approximately 50-70ms per frame) would enable continuous monitoring during work execution, detecting PPE removal mid-shift (workers removing helmets during breaks, taking off gloves between tasks) or improper equipment usage during tasks (harness not clipped to anchor point, safety glasses pushed onto forehead).

However, this requires GPU deployment (estimated hardware cost: $500-1,000 for edge computing device with NVIDIA Jetson or similar embedded GPU) and raises additional privacy considerations for continuous worker surveillance, necessitating policy development addressing when and where video monitoring is appropriate, data retention limits (recommend 24-48 hours automatic deletion unless incident occurs), access controls to prevent misuse, and worker consent protocols complying with Indonesian personal data protection regulations. Alternative approaches include periodic frame sampling (1 frame per 5-10 seconds) reducing computational requirements while maintaining reasonable monitoring coverage.

System Integration: The system currently operates as standalone application without integration into existing enterprise safety management systems, requiring manual export of results and separate record-keeping. Development of RESTful API interfaces for seamless integration with PLN's digital infrastructure (SAP EHS module for occupational health and safety management, existing safety incident databases, work order systems) would enhance operational utility by enabling: (1) automatic compliance record synchronization with centralized safety databases upon detection completion, eliminating manual data entry and associated delays/errors; (2) real-time violation alerts sent to safety managers and field supervisors via email or SMS when critical PPE is missing during high-risk operations, enabling immediate intervention before work proceeds; (3) integration with worker identification systems (RFID badges, facial recognition) for individual compliance tracking and automated linkage of detection

results to specific worker profiles, enabling targeted safety coaching and training for repeat violators; and (4) automated regulatory reporting for Indonesian Ministry of Manpower compliance audits (quarterly safety statistics, annual incident reports), reducing administrative burden on safety staff. API development should follow industry standards (OpenAPI/Swagger specification) with proper authentication (OAuth 2.0) and rate limiting to ensure security and reliability.

Advanced Detection Capabilities: The current system detects PPE presence but does not assess proper usage, a critical limitation as improperly worn equipment provides false sense of security. Future research should explore: (1) anomaly detection for identifying improperly worn PPE including unfastened helmet chin straps (helmet present but not secured, detected through strap position analysis), incorrectly positioned harness connection points (harness worn but not clipped to anchor, requiring D-ring detection and spatial relationship analysis), safety vest worn inside-out (reflective strips not visible), gloves partially removed (detected on hands but fingers exposed); (2) worker identification and tracking for individual compliance history monitoring using face recognition (with appropriate privacy protections and consent) or RFID badge integration, enabling targeted safety coaching for repeat violators and compliance trend analysis per worker, potentially identifying individuals requiring additional training or PPE fitting adjustments; (3) environmental hazard detection beyond PPE including unsafe working positions near high-voltage equipment (workers within minimum approach distance), absence of required safety barriers and warning signs, improper ladder positioning (angle, stability, extension length), electrical arc flash hazards (exposed conductors, open panel doors), for comprehensive safety monitoring addressing both personal protective measures and environmental hazards; and (4) action recognition to verify proper equipment usage procedures including verifying harness attachment before climbing (detecting worker approaching pole/ladder, checking harness connection in pre-climb frame), glove donning before handling energized equipment (temporal sequence analysis), proper lockout-tagout procedures (detecting locks and tags on equipment), providing behavioral safety monitoring beyond static PPE presence.

Deployment Considerations: Mobile application development for smartphone-based field deployment would improve accessibility for field coordinators and enable GPS-tagged compliance records for location-specific safety analysis. Current web application requires

laptop or tablet (minimum screen size 10 inches for comfortable use), limiting portability compared to ubiquitous smartphone availability among field personnel. Native mobile apps (iOS and Android) could leverage smartphone cameras for direct image capture with immediate on-device processing (using TensorFlow Lite or ONNX Runtime for edge inference), automatic GPS tagging of inspection location enabling heat map visualization of compliance rates across service territory, offline operation capability for remote areas without cellular coverage (with sync when connectivity restored), and push notifications for inspection reminders based on work schedules. Privacy-preserving design should include on-device processing without cloud upload of images, local storage with encryption, and configurable data retention periods.

Impact Assessment: Longitudinal studies assessing system impact on actual safety compliance rates and incident reduction would provide valuable evidence of practical effectiveness beyond technical performance metrics, necessary for justifying deployment investment and measuring return on safety. Proposed study design: 12-month comparative analysis across multiple PLN service areas (minimum 4 sites) with system deployment versus control sites using manual inspection only, measuring: (1) PPE compliance rates pre/post deployment measured through monthly random audits by independent assessors (target: 500+ observations per site), with statistical analysis using difference-in-differences methodology to isolate system effect from temporal trends; (2) near-miss incident frequency from safety reporting systems, testing hypothesis that improved PPE compliance reduces near-misses as leading indicator of serious injuries; (3) lost-time injury rates (injuries requiring work absence) and their correlation with PPE compliance improvements, though long observation periods (2-3 years) may be needed given low base rates of serious injuries; (4) supervisor time allocation measured through time-motion studies, quantifying time savings from automated detection and reallocation to other safety activities (coaching, hazard assessments, safety training); and (5) cost-benefit analysis comparing system deployment and maintenance costs against avoided injury costs (medical expenses, lost productivity, workers compensation, regulatory fines), supervisor labor savings, and efficiency gains, providing economic justification for organizational adoption. Study should also assess worker attitudes through surveys examining perceived fairness of automated monitoring, privacy concerns, and behavioral responses (voluntary compliance versus resentment and workarounds)

ISI Journal of Information Systems and Informatics
Published By Asosiasi Doktor Sistem Informasi Indonesia

4. CONCLUSION

This research successfully developed an automated Personal Protective Equipment detection system using YOLOv11 architecture for electric power distribution workplace safety monitoring, achieving 94.0% precision, 90.1% recall, and 92.8% mAP@50 on validation dataset. The system demonstrated excellent detection for safety-critical equipment including full-body harness (100% precision, 99.4% recall) and person identification (100% precision and recall), confirming reliability for high-risk electrical work applications. The web-based application addresses manual inspection limitations at PT PLN UP3 Banyuwangi including insufficient supervisors (1:15-20 ratio), geographically distributed work locations, and human error in visual checking, with field validation from three OHS supervisors confirming high perceived usefulness (4.6/5.0 average rating) and practical applicability. The system provides consistent, objective, and scalable monitoring capabilities while reducing inspection time by 67.9% and improving documentation quality, enabling supervisors to conduct more frequent safety checks without proportional increases in labor resources. This research contributes to occupational safety practice by demonstrating feasible integration of advanced deep learning technology into existing safety workflows with minimal infrastructure requirements (standard CPU processing in 2-3 seconds), providing practical implementation guidance for high-risk industries beyond academic performance metrics.

ACKNOWLEDGMENT

The authors express sincere gratitude to PT PLN (Persero) UP3 Banyuwangi for providing access to field operations and facilitating interviews with three Occupational Health and Safety supervisors, whose insights were invaluable in understanding practical safety monitoring challenges and system requirements. Special appreciation is extended to the OHS supervisors who generously shared their time and expertise during field validation, providing critical feedback on system usability and operational applicability. The authors acknowledge the use of Google Colaboratory platform with Tesla T4 GPU resources for model training, and Roboflow platform for dataset annotation and preprocessing tools. We thank all electric power distribution workers who participated in image data collection while maintaining their daily safety protocols and provided informed consent for

research use. The constructive feedback from anonymous reviewers significantly improved the quality, clarity, and methodological rigor of this manuscript.

REFERENCES

[1] B. I. Arianto, "Implementasi Metode YOLO pada Deteksi Pakaian Keselamatan yang Lengkap di Proyek Kontruksi," *Ranah Research: Journal of Multidisciplinary Research and Development*, vol. 6, no. 1, pp. 56–63, 2023.

[2] M. Ferdous and S. M. M. Ahsan, "PPE detector: a YOLO-based architecture to detect personal protective equipment (PPE) for construction sites," *PeerJ Computer Science*, vol. 8, p. e999, Jun. 2022, doi: 10.7717/peerj-cs.999.

[3] K. S. Park and M. Lee, "Framework of automated construction-safety monitoring using cloud-enabled BIM and BLE mobile tracking sensors," *Journal of Construction Engineering and Management*, vol. 143, no. 2, p. 05016019, Feb. 2017, doi: 10.1061/(ASCE)CO.1943-7862.0001223.

[4] M. Guo, Y. Liu, and J. Malec, "A new Q&A pair generation method from online forum using deep learning," *Expert Systems with Applications*, vol. 151, p. 113350, Aug. 2020, doi: 10.1016/j.eswa.2020.113350.

[5] A. Fang, D. Wu, and J. Li, "Construction worker safety analysis using computer vision: A review," *Safety Science*, vol. 149, p. 105681, May 2022, doi: 10.1016/j.ssci.2022.105681.

[6] W. Fang, L. Ding, H. Luo, and P. E. D. Love, "Falls from heights: A computer vision-based approach for safety harness detection," *Automation in Construction*, vol. 91, pp. 53–61, Jul. 2018, doi: 10.1016/j.autcon.2018.02.018.

[7] Y. LeCun, Y. Bengio, and G. Hinton, "Deep learning," *Nature*, vol. 521, no. 7553, pp. 436–444, May 2015, doi: 10.1038/nature14539.

[8] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *Proc. IEEE Conf. Computer Vision and Pattern Recognition (CVPR)*, Las Vegas, NV, USA, 2016, pp. 770–778, doi: 10.1109/CVPR.2016.90.

[9] J. Redmon, S. Divvala, R. Girshick, and A. Farhadi, "You only look once: Unified, real-time object detection," in *Proc. IEEE Conf. Computer Vision and Pattern Recognition (CVPR)*, Las Vegas, NV, USA, 2016, pp. 779–788, doi: 10.1109/CVPR.2016.91.

[10] Z. Wang, Y. Wu, L. Yang, A. Thirunavukarasu, C. Evison, and Y. Zhao, "Fast personal protective equipment detection for real construction sites using deep learning approaches," *Sensors*, vol. 21, no. 10, p. 3478, May 2021, doi: 10.3390/s21103478.

[11] Z. S. Juanda, Z. K. Simbolon, and H. Huzaeni, "Mengintegrasikan Metode YOLO (You Only Look Once) Dalam Deteksi APD (Alat Pelindung Diri) Pada Industri Migas," *Journal of TIK*, vol. 4, no. 1, pp. 127–135, 2024, doi: 10.33365/jtik.v4i1.2156.

[12] S. Malaikrisanachalee, N. Wongwai, and E. Kowcharoen, "ESPCN-YOLO: A High-Accuracy Framework for Personal Protective Equipment Detection Under Low-Light and Small Object Conditions," *Buildings*, vol. 15, no. 10, p. 1609, May 2025, doi: 10.3390/buildings15101609.

[13] D. Ocharo, H. Nganga, and S. Kiambi, "Automatic PPE monitoring system for construction workers using YOLO algorithm based on deep reinforcement learning," *International Journal of Computer Applications*, vol. 186, no. 12, pp. 10–15, Mar. 2024, doi: 10.5120/ijca2024923512.

[14] N. Z. A. A. Rahmah, R. Indarti, Z. M. A. Putra, E. Setiawan, and A. Z. Arfianto, "Implementasi Deteksi Kelengkapan APD pada Hazardous Area menggunakan Metode YoloV5," *Jurnal Elektronika dan Otomasi Industri*, vol. 11, no. 3, pp. 245–253, Dec. 2024, doi: 10.33795/elkolind.v11i3.485.

[15] D. Tan and S. Prayogi, "Deteksi Otomatis Penggunaan APD pada Pekerja Migas Menggunakan Deep Learning dan Computer Vision," *Techno.com*, vol. 24, no. 3, pp. 412–425, Aug. 2025, doi: 10.33633/tc.v24i3.9234.

[16] M. A. Taufiqurrochman and H. Februariyanti, "Rancang Bangun Aplikasi Deteksi Alat Pelindung Diri (APD) untuk Pekerja Proyek dengan Menggunakan Algoritma Yolov5," *Jurnal JTIK (Jurnal Teknologi Informasi dan Komunikasi)*, vol. 8, no. 2, pp. 471–480, Apr. 2024, doi: 10.35870/jtik.v8i2.1245.

[17] B. Tong, G. Li, X. Bu, Y. Wang, and X. Yu, "A deep learning-based algorithm for the detection of personal protective equipment," *PLOS ONE*, vol. 20, no. 5, p. e0322115, May 2025, doi: 10.1371/journal.pone.0322115.

[18] N. D. Nath, A. H. Behzadan, and S. G. Paal, "Deep learning for site safety: Real-time detection of personal protective equipment," *Automation in Construction*, vol. 112, p. 103085, Apr. 2020, doi: 10.1016/j.autcon.2020.103085.

[19] H. Wu, J. Zhao, and W. Yan, "Real-time helmet detection based on improved YOLOv4," in *Proc. Chinese Control and Decision Conference (CCDC)*, Hefei, China, 2021, pp. 2189–2194, doi: 10.1109/CCDC52312.2021.9601518.

[20] K. Kaplan, Y. Kaya, M. Kuncan, and H. M. Ertunç, "An Improved Feature Extraction Method Using Texture Analysis with LBP for Bear Recognition," in *Proc. Int. Conf. Artificial Intelligence and Data Processing (IDAP)*, Malatya, Turkey, 2018, pp. 1–4, doi: 10.1109/IDAP.2018.8620808.

[21] J. Li, H. Liu, T. Wang, M. Jiang, S. Wang, K. Li, and X. Zhao, "Safety helmet wearing detection based on image processing and machine learning," in *Proc. Int. Conf. Advanced Computational Intelligence (ICACI)*, Guilin, China, 2019, pp. 174–179, doi: 10.1109/ICACI.2019.8778583.

[22] H. Zhang, X. Yan, H. Li, R. Jin, and D. Fu, "Real-time alarming, monitoring, and locating for non-hard-hat use in construction," *Journal of Construction Engineering and Management*, vol. 145, no. 3, p. 04019006, Mar. 2019, doi: 10.1061/(ASCE)CO.1943-7862.0001629.

[23] B. E. Mneymneh, M. Abbas, and H. Khoury, "Vision-based framework for intelligent monitoring of hardhat wearing on construction sites," *Journal of Computing in Civil Engineering*, vol. 33, no. 2, p. 04018066, Mar. 2019, doi: 10.1061/(ASCE)CP.1943-5487.0000813.

[24] A. Kelm, L. Laußat, A. Meins-Becker, D. Platz, M. Khazaee, and A. C. Costin, "Mobile passive Radio Frequency Identification (RFID) portal for automated and rapid control of Personal Protective Equipment (PPE) on construction sites," *Automation in Construction*, vol. 36, pp. 38–52, Dec. 2013, doi: 10.1016/j.autcon.2013.08.009.

[25] M. W. Park, N. Elsafty, and Z. Zhu, "Hardhat-wearing detection for enhancing on-site safety of construction workers," *Journal of Construction Engineering and Management*, vol. 141, no. 9, p. 04015024, Sep. 2015, doi: 10.1061/(ASCE)CO.1943-7862.0000974.

[26] Q. Fang, H. Li, X. Luo, L. Ding, H. Luo, T. M. Rose, and W. An, "Detecting non-hardhat-use by a deep learning method from far-field surveillance videos," *Automation in Construction*, vol. 85, pp. 1–9, Jan. 2018, doi: 10.1016/j.autcon.2017.09.018.

[27] C. Y. Wang, A. Bochkovskiy, and H. Y. M. Liao, "YOLOv7: Trainable bag-of-freebies sets new state-of-the-art for real-time object detectors," in *Proc. IEEE/CVF Conf. Computer Vision and Pattern Recognition (CVPR)*, Vancouver, BC, Canada, 2023, pp. 7464–7475, doi: 10.1109/CVPR52729.2023.00721.