Enhancing Low-Light PPE Violation Detection in Industrial Environments Using Multi-Contrast Image Processing and YOLOv9

Fathorazi Nur Fajri
Department of Informatics, Nurul Jadid University, Indonesia
fathorazi@unuja.ac.id (corresponding author)

Kamil Malik
Department of Informatics, Nurul Jadid University, Indonesia
kamil@unuja.ac.id

Abu Tholib
Department of Informatics, Nurul Jadid University, Indonesia
abu@unuja.ac.id

Received: 23 August 2025 | Revised: 17 September 2025, 30 September 2025, and 6 October 2025 | Accepted: 9 October 2025

 DOI: https://doi.org/10.48084/etasr.14254

ABSTRACT

Ensuring compliance with Personal Protective Equipment (PPE) is critical for workplace safety, yet low-light conditions reduce the accuracy of vision-based monitoring systems. This study introduces a novel violation-aware PPE dataset consisting of 11,407 annotated images across eight categories, explicitly modeling both compliance and non-compliance classes (e.g., helmet/no-helmet, vest/no-vest, gloves/no-gloves, shoes/no-shoes). The dataset differs from prior works by focusing on low-light industrial environments where detection is most challenging. To address these conditions, three multi-contrast enhancement techniques—Contrast Limited Adaptive Histogram Equalization (CLAHE), Auto-CLAHE, and EnlightenGAN—were integrated with an optimized YOLOv9t model. The six You Only Look Once (YOLO) variants (YOLOv5s–YOLOv12n) were benchmarked, with YOLOv9t plus augmentation achieving the best performance, with mAP@50 of 0.915, mAP@50–95 of 0.659, precision of 0.906, recall of 0.844, and inference time of 4.0 ms. The enhancement experiments demonstrated that CLAHE provided the highest detection coverage (79.43%), Auto-CLAHE yielded the greatest detection density (2.29/frame), and EnlightenGAN offered limited benefits due to domain shift. The findings confirm that histogram-based methods consistently improve PPE violation detection under low-light conditions, while Generative Adversarial Network (GAN)-based approaches require domain-specific adaptation. Overall, this study contributes a new dataset, systematic YOLO benchmarking, and the first task-driven comparison of classical, adaptive, and GAN-based enhancement methods for reliable, real-time workplace safety in challenging lighting environments.

Keywords-Personal Protective Equipment (PPE); violation detection; low-light image enhancement; YOLO; CLAHE; Auto-CLAHE; EnlightenGAN

I. INTRODUCTION

Ensuring occupational safety in industrial and construction environments depends on the proper use of PPE, such as helmets, vests, gloves, masks, and boots. The manual inspection of PPE compliance is often unreliable, costly, and prone to human error, which has encouraged the use of computer vision and artificial intelligence for automatic compliance monitoring [1, 2]. Early attempts using handcrafted feature extraction methods, such as HAAR, HOG, and LBP descriptors, yielded acceptable results in constrained scenarios but lacked robustness to variations in illumination, occlusion, and complex environments [3].

To support research progress, several annotated datasets have been introduced. The SH17 dataset [4] provides annotations for multiple PPE categories, while SHEL5K [5] focuses on safety helmet detection. Authors in [6] proposed a large and realistic dataset covering diverse PPE categories, and subsequent works [7, 8] developed video-based surveillance datasets for compliance monitoring. Although valuable, these resources were primarily captured in well-lit conditions and remain insufficient for addressing low-light environments

where detection becomes most critical. Tunnel and underground construction sites, for instance, demand specialized datasets, such as those proposed in [9], but these often cover limited PPE categories.

Parallel to dataset development, deep learning, particularly the YOLO family, has advanced PPE detection significantly. YOLO has evolved from YOLOv1 to YOLOv12, with continuous improvements in backbone architectures, decoupled detection heads, and training strategies [10, 11]. The Convolutional Neural Network (CNN) has been applied for helmet detection in construction [12], while YOLOv5 and YOLOv7 demonstrated superior accuracy for multi-class PPE detection [13, 14]. Domain-specific adaptations, such as PPE Detector [2] and DPPNet [15], further improved robustness under perspective distortions. It has been confirmed that YOLOv9 and YOLOv12 achieve state-of-the-art performance in real-time detection tasks [16].

Other frameworks outside the YOLO family also contributed to PPE monitoring. Authors in [17] proposed a computer vision system for PPE detection using CNNs. Authors in [18] introduced an improved YOLOv5 for detecting protective gear in challenging conditions, while authors in [19] optimized YOLO-based detection for real construction sites, achieving high accuracy and speed. Despite these contributions, most methods still struggle with generalization under complex or low-light conditions. Moreover, compliance verification often remained limited to single PPE types, such as helmets or safety glasses [1, 7, 8].

Low-light Image Enhancement (LLIE) has emerged as a promising solution to address visibility challenges in safety monitoring. Classical methods, such as CLAHE [20] and its optimization with Pelican-inspired algorithms [21], improved local contrast, while Retinex-based correction [9] normalized illumination. Deep learning approaches further advanced LLIE: EnlightenGAN [22] introduced unsupervised light enhancement, authors in [23] demonstrated that EnlightenGAN can still be effective even with small training datasets, and authors in [24] integrated Auto-CLAHE with MobileNet for low-light drowsiness detection. Medical imaging studies and attention-based ViT frameworks for underground helmet detection [25] demonstrated domain-specific gains. It has been confirmed that although LLIE has achieved remarkable progress [26], its integration with PPE compliance detection remains limited.

Despite these advances, significant research gaps remain:

- Existing PPE datasets largely assume well-lit conditions and do not capture violations explicitly.
- Enhancement techniques are rarely evaluated in end-to-end detection pipelines.
- Most prior works focus only on PPE presence rather than modeling unsafe conditions (e.g., no-helmet, no-vest) as distinct classes.

These gaps limit the applicability of existing solutions in real-world industrial environments. Table I shows the datasets used by previous researchers.

TABLE I. EXISTING DATASETS USED IN PRIOR PPE DETECTION STUDIES

Dataset	Year	Number of images	Number of classes	Class name
SH17 [4]	2024	8,099	17	Person, Face, Head, Helmet, Earmuffs, Goggles, Glasses, Vest, Gloves, Masks, Medical Masks, Boots, Leather Boots, Harness Gear
SHEL5K [5]	2022	5,000	6	Person not helmet, Person and helmet, Head and helmet, Head, Helmet, Face
TCRSF [6]	2023	12,373	7	Person, Safety Helmet, Safety Clothing, Other Clothing, Head, Blurred Clothing, Blurred Head
PPE Compliance [7]	2024	386,011	4	Helmet, NoHelmet, Vest, NoVest
Real-time PPE [8]	2021	853	4	Helmet, NoHelmet, Vest, NoVest
YOLOX Tunnel [9]	2023	2371	3	Helmet, Vest, and Person
SHW [12]	2019	7,581	3	Helmet, Head, Person
Pictor PPE [14]	2020	784	3	Worker, Hat, Vest
CHV [19]	2021	1,330	6	Person, Vest, Blue Helmet, Red Helmet, White Helmet, Yellow Helmet
Proposed Dataset	2025	11,407	8	Helmet, No-Helmet, Vest, No-Vest, Gloves, No-Gloves, Shoes, And No-Shoes

To address these issues, this study proposes a novel framework that integrates multi-contrast enhancement with YOLOv9-based PPE detection under low-light conditions. The main objectives are:

- To construct a violation-aware PPE dataset with explicit compliance and violation classes
- To benchmark YOLO variants and identify the most effective model for low-light PPE monitoring
- To evaluate the impact of CLAHE, Auto-CLAHE, and EnlightenGAN on detection performance.

The key contributions of this study are: (1) the creation of a new dataset specifically for low-light PPE violations, (2) the systematic benchmarking of YOLO variants, and (3) the first comparative analysis of classical, adaptive, and GAN-based enhancement methods for violation detection in challenging environments.

II. RESEARCH METHODOLOGY

Figure 1 illustrates the research framework, from data collection to model evaluation.

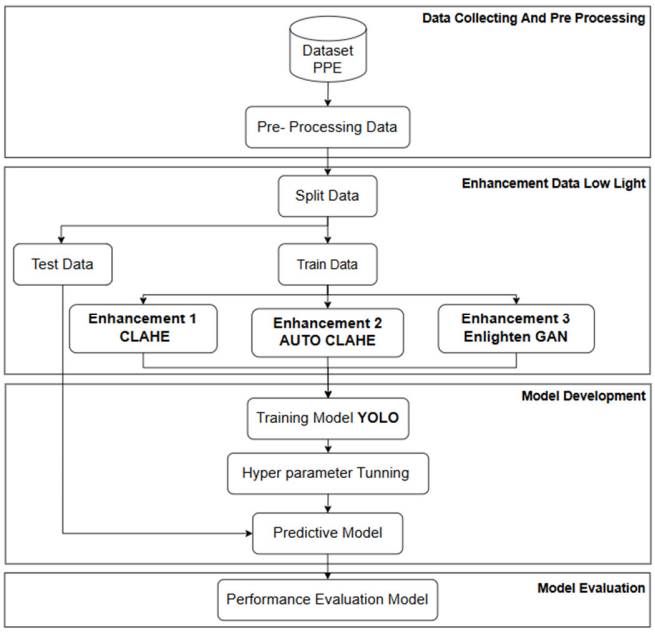

Fig. 1. Research framework of the proposed detection system.

A. *Dataset Construction*

The dataset used in this study extends the previously developed violation-aware PPE dataset, specifically designed to represent both compliance and violation instances under low-light industrial conditions. Images were collected from multiple sources: Google Images (158 samples), Kaggle PPE-related datasets (824 samples) [4], Roboflow repositories (2,300 samples) [5], and real-world CCTV footage from a coal-fired power plant (1,202 samples). This combination ensures both diversity and domain realism.

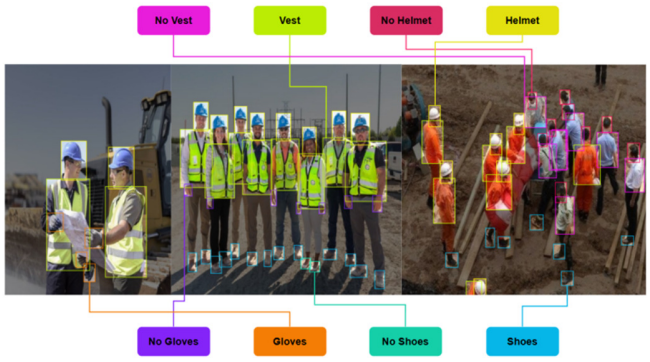

Fig. 2. Annotation data on violations of PPE.

Each image was manually annotated into eight classes: helmet, no-helmet, vest, no-vest, gloves, no-gloves, shoes, and no-shoes. Annotation was conducted using the Roboflow Annotator, applying bounding boxes for each object. Following the COCO standard, annotated objects were categorized into Small (<32×32 pixels), Medium (32×32–96×96 pixels), and Large (>96×96 pixels) to facilitate evaluation across object scales. This categorization highlights the challenges of detecting small objects in low-light settings, which motivated the integration of enhancement methods. The distribution of the dataset annotation is displayed in Table II.

TABLE II. DISTRIBUTION OF ANNOTATED OBJECTS BY CLASS AND SIZE

Class	Size			Total
	Small	Medium	Large	
Gloves	870	2358	2906	6134
No-Gloves	3179	4515	561	8255
Helmet	1109	3793	1998	6900
No-Helmet	1216	3452	2231	6899
Vest	196	2067	4252	6515
No-Vest	199	2404	3889	6492
Shoes	2904	2811	3080	8795
No-Shoes	1201	2298	4207	7706

B. *Enhancement Methods*

To address the degradation of object detection performance in low-light conditions, this study evaluates three contrast enhancement techniques: CLAHE [20], Auto-CLAHE [24], and Enlighten GAN [22].

CLAHE is an image contrast enhancement method that operates on small regions (tiles) of the image, making it adaptive to local variations. An image of size $W \times H$ is divided into tiles of size $M \times N$. $I(x, y)$ represents the pixel intensity at position (x, y), H_k denotes the local histogram of tile k, and the local histogram is calculated as:

$$H_k = \sum_{x=0}^{M-1} \sum_{y=0}^{N-1} \delta(I(x,y) - k) \quad (1)$$

To avoid excessive noise amplification, L is the total number of gray levels, and the histogram is limited by the clip limit C_{max}, defined as:

$$C_{max} = clip\_factor \times \frac{M \times N}{L} \quad (2)$$

The histogram values exceeding this limit are clipped:

$$H'_k = min(H_k, C_{max}) \quad (3)$$

The excess values from clipping are redistributed evenly across all histogram bins:

$$H''_k = H'_k + \frac{\sum(H_k - H'_k)}{L} \quad (4)$$

Next, the cumulative distribution function (CDF) is computed from the redistributed histogram:

$$\mathrm{CDF}(k) = \frac{\sum_{i=0}^{k} H''_i}{M \times N} \quad (5)$$

Finally, the pixel intensities are transformed using:

$$I'(x,y) = \mathrm{round}\left[\mathrm{CDF}\left(I(x,y)\right) \times (L-1)\right] \quad (6)$$

Auto-CLAHE is an adaptive variant of CLAHE that automatically adjusts enhancement parameters based on the luminance histogram of each image. The tile size is adaptively calculated based on the image dimensions:

$$tile_{size} = \alpha \times \sqrt{W \times H} \quad (7)$$

The automatic clip limit is calculated based on the standard deviation of the image intensity:

$$C_{max} = \beta \times \frac{M \times N}{L} \times \left(1 + \frac{\sigma_I}{\sigma_{ref}}\right) \quad (8)$$

This adaptive approach allows it to respond better to varying illumination levels across different surveillance zones, which is common in industrial CCTV setups.

Fig. 3. Comparison of enhancement results.

EnlightenGAN is a deep learning-based low-light enhancement technique that employs a GAN architecture to translate low-light images into enhanced images, with the goal of producing visually brighter images with improved perceptual contrast.

Figure 3 presents the results of applying enhancement techniques to each frame before the YOLO inference process, which produced four sets of evaluations: raw (unenhanced), CLAHE enhanced, Auto-CLAHE enhanced, and EnlightenGAN enhanced.

C. YOLO Model Benchmarking and Selection

YOLO model benchmarking was conducted to determine the optimal variant for violation-aware PPE detection. Six lightweight YOLO architectures were tested: YOLOv5s, YOLOv8n, YOLOv9t, YOLOv10n, YOLOv11n, and YOLOv12n. All models were trained and evaluated on the violation-aware dataset, using identical training configurations to ensure a fair comparison, and the results are outlined in Table III.

The results showed that YOLOv9t with augmentation achieved the best performance. The YOLOv9t architecture adopts a lightweight CSPDarknet-based backbone, a path aggregation neck, and a decoupled detection head. This design balances accuracy and computational efficiency, reducing parameters while maintaining strong feature extraction for small objects. Compared to earlier YOLO versions, YOLOv9 introduces enhanced gradient flow, improved normalization, and streamlined convolutional layers, making it suitable for real-time violation detection in resource-constrained industrial environments.

TABLE III. EXPERIMENTAL RESULTS OF YOLO VARIANTS ON PROPOSED DATASET

Model	Data split	Result			
		mAP@50	mAP@50–95	Precision	Recall
YOLOv5s	80:12:8	0.878	0.611	0.869	0.819
YOLOv8n	80:12:8	0.875	0.605	0.848	0.825
YOLOv9t	80:12:8	0.894	0.632	0.871	0.832
YOLOv10n	80:12:8	0.874	0.611	0.89	0.792
YOLOv11n	80:12:8	0.883	0.618	0.862	0.822
YOLOv12n	80:12:8	0.871	0.602	0.875	0.815
YOLOv5s	80:10:10	0.788	0.489	0.841	0.698
YOLOv8n	80:10:10	0.768	0.478	0.798	0.688
YOLOv9t	80:10:10	0.791	0.504	0.8	0.713
YOLOv10n	80:10:10	0.76	0.477	0.801	0.684
YOLOv11n	80:10:10	0.785	0.495	0.808	0.7
YOLOv12n	80:10:10	0.754	0.465	0.772	0.706
Proposed model	92:4:3 (+augment)	0.915	0.659	0.906	0.844

D. Training and Inference Setup

To optimize YOLOv9t for the proposed dataset, several adjustments were implemented: data augmentation techniques such as mosaic, rotation, and flipping; input size fixed at 640×640 pixels; and learning rate tuning with Stochastic Gradient Descent (SGD) optimizer, with a momentum of 0.937, an initial learning rate of 0.01, and a weight decay of 0.0005 for 100 epochs on an NVIDIA RTX-series GPU to accelerate computation. These optimizations improved the model's generalization ability under varying illumination and object scales, leading to superior performance compared to baseline settings.

Inference was performed on all four evaluation sets (Raw, CLAHE, Auto-CLAHE, and EnlightenGAN), and the results were recorded for further performance analysis.

E. Evaluation Metrics

Model performance was assessed using a combination of frame-level and object-level metrics. At the frame level, the study measured the percentage of frames with valid detections and the total number of detections. At the object level, the study measured the average detections per detected frame and the per-class detection counts for all eight PPE categories.

$$AP = \int_0^1 p(r)dr \quad (9)$$

$$mAP = \frac{1}{N}\sum_{i=1}^{N} AP_i \quad (10)$$

In addition, for a manually annotated subset of low-light frames, the study computed mAP@50, precision, recall, and F1-score. This approach ensures that the evaluation considers not only visual quality improvements but also task-driven

metrics directly relevant to workplace safety monitoring in industrial environments:

$$\text{Precision} = \frac{TP}{TP+FP} \quad (11)$$

$$\text{Recall} = \frac{TP}{TP+FN} \quad (12)$$

III. RESULTS

A. YOLO Model Benchmarking Results

The benchmark results of YOLOv9t for precision and recall with augmented data are shown in Figure 4. The confusion matrix for the proposed model is portrayed in Figure 5.

Before evaluating the enhancement methods, six YOLO variants—YOLOv5s, YOLOv8n, YOLOv9t, YOLOv10n, YOLOv11n, and YOLOv12n—were benchmarked on the violation-aware PPE dataset without enhancement. The results indicated that YOLOv9t with augmentation provided the best trade-off between accuracy and efficiency, achieving mAP@50 of 0.915, mAP@50–95 of 0.659, precision of 0.906, recall of 0.844, and inference time of 4.0 ms per image. With only 2.0M parameters and a model size of 4.7 MB, YOLOv9t was also the most compact among top performers, making it suitable for real-time deployment on edge devices.

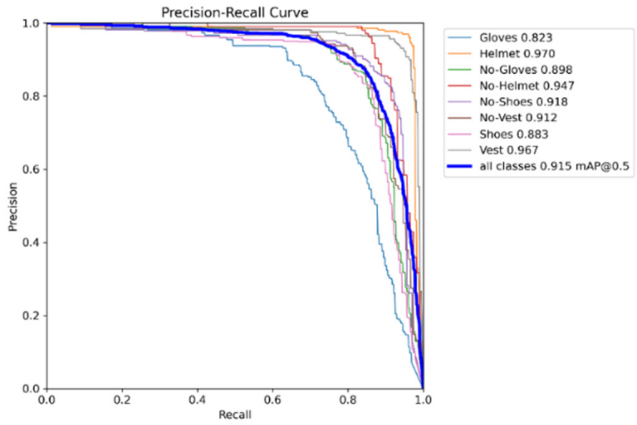

Fig. 4. Precision-recall curve for proposed model.

B. Impact of Enhancement Methods on Detection

Table IV presents the comparative results of the four evaluation sets: Raw (non-enhanced), CLAHE, Auto-CLAHE, and EnlightenGAN. The metrics include frame-level detection coverage, total detections, and detection density per detected frame. The results reveal several key trends: CLAHE achieved the highest detection coverage (79.43%) and the largest total number of detections (6,930), representing a 3% and 10.9% improvement over the Raw baseline, respectively. Auto-CLAHE recorded slightly lower coverage (78.57%) but achieved the highest detection density per detected frame (2.29), suggesting that while it detected slightly fewer frames overall compared to CLAHE, the frames it did detect tended to contain more correctly identified objects. EnlightenGAN significantly underperformed in this task, with detection coverage dropping to 48.39% and total detections reduced by 45% compared to the Raw baseline. This decline is likely due to domain shift—the GAN-based enhancement altered the texture and color distribution in a way that deviated from the model's training data, reducing feature compatibility with YOLOv9t's learned representations.

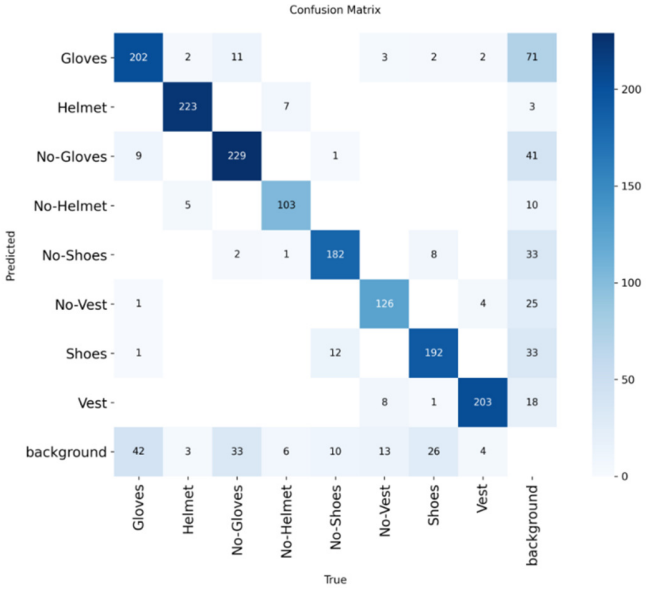

Fig. 5. Confusion matrix for proposed model.

C. Detection Analysis by Class

A closer look at per-class detection counts provides further insight into the effects of enhancement methods: For Helmet detection, CLAHE slightly outperformed Auto-CLAHE (2,877 versus 2,873) and both exceeded the Raw baseline (2,836). For No-Helmet, the Raw baseline produced the highest count (152), while CLAHE and Auto-CLAHE detected fewer violations (83 and 79, respectively). This suggests that the histogram-based enhancements, while improving visibility, may have also increased the visual similarity between compliant and non-compliant cases, leading to more conservative violation detections. For Shoes and No-Shoes, CLAHE recorded the highest counts, indicating improved visibility of footwear features in low-light scenarios. Auto-CLAHE was close behind, while EnlightenGAN struggled significantly, detecting less than half of the Raw baseline counts.

D. Visual Comparison

The qualitative analysis confirmed the quantitative results. CLAHE and Auto-CLAHE enhanced edges and object boundaries without introducing strong artifacts, improving YOLOv9t's detection consistency. EnlightenGAN produced brighter outputs but with oversaturation and color shifts that degraded bounding box accuracy, as shown in Figure 6.

TABLE IV. FRAME- AND CLASS-LEVEL DETECTION PERFORMANCE

File	Frame			Classes				Detect all classes	Avg. time
	All	Detect	Not Detect	No-vest	Helmet	No-Helmet	Shoes		
Raw	3840	2935	905	2836	152	2205	1059	6252	2.13
CLAHE	3840	3050	790	2877	83	2154	1816	6930	2.27
Auto-CLAHE	3840	3017	823	2873	79	2191	1755	6898	2.29
EnlightenGAN	3840	1858	1982	1760	49	900	712	3421	1.84

Fig. 6. Visual comparison result detection.

E. Comparative Study

To contextualize these findings, the proposed YOLOv9t model was compared against the results from prior works, as presented in Table V. Conventional CNN approaches [12] achieved *mAP* of 0.838, while YOLO-based models demonstrated higher performance, with YOLOX [7] at 0.865, YOLOv7 [13] at 0.877, and YOLOX-m [2] at 0.898. However, YOLOv9e [4] produced a lower score of 0.709 under low-light settings. The proposed YOLOv9t achieved the highest *mAP* of 0.915, confirming its robustness for PPE violation detection in challenging environments. These results demonstrate that the integration of multi-contrast enhancement with YOLOv9t offers clear advantages over existing methods.

F. Limitations and Future Work

This study has several limitations. First, the dataset is primarily derived from industrial environments, which may introduce domain bias when applied to other settings. Second, GAN-based enhancement methods exhibited reduced performance due to domain shift, suggesting that domain-adaptive training is required. Third, inference was evaluated using a standard PyTorch pipeline, which, although accurate, may not yet represent the most efficient deployment strategy for real-time applications. Future work will address these issues by expanding dataset diversity, exploring domain adaptation strategies, and applying inference optimizers, such as TensorRT, to reduce latency and computational cost.

TABLE V. COMPARATIVE PERFORMANCE OF PPE DETECTION MODELS

Study	Dataset	Model	*mAP*
[2]	CHVG	YOLOX-m	0.898
[4]	SH17	YOLOv9e	0.709
[7]	PPE Compliance	YOLOX	0.865
[12]	SHW	CNN	0.838
[13]	Safety Gear	YOLOv7	0.877
Proposed model	PPE - Violation	YOLOv9t	0.915

IV. CONCLUSIONS

This study introduced a violation-aware Personal Protective Equipment (PPE) dataset of 11,407 low-light images explicitly labeled into compliance and violation classes. The benchmarking of six You Only Look Once (YOLO) variants demonstrated that the optimized YOLOv9t achieved the best performance (*mAP@50* = 0.915, precision = 0.906, recall = 0.844), surpassing previous Convolutional Neural Network

(CNN)-based and YOLO-based PPE detection studies [2, 4, 7, 12, 13]. The integration of Contrast Limited Adaptive Histogram Equalization (CLAHE) and Auto-CLAHE consistently improved detection in low-light environments, while EnlightenGAN showed limited generalization due to domain shifts.

The uniqueness of this study lies in the combination of a violation-aware dataset with multi-level contrast enhancement in an end-to-end detection framework, complementing previous research that mostly focused on well-lit conditions or single PPE categories. These results confirm that YOLO-based optimization can extend the application of existing approaches to more challenging industrial scenarios. Future research on inference optimization (e.g., TensorRT) for real-time applications is proposed.

DATA AVAILABILITY STATEMENT

The datasets used in this study are publicly available.

- The SH17 dataset ("A Dataset for Human Safety and Personal Protective Equipment Detection in Manufacturing Industry") can be accessed at https://github.com/ahmadmughees/sh17dataset
- The SHEL5K dataset ("An Extended Dataset and Benchmarking for Safety Helmet Detection") is available at https://github.com/MoyoG/SHEL5K

ACKNOWLEDGMENT

This research was supported by the Ministry of Education and Culture, Research, and Technology of Indonesia through the Fundamental Research Grant No. 070/LL7/DT.05.00/PL/2025.

REFERENCES

[1] B. Balakreshnan, G. Richards, G. Nanda, H. Mao, R. Athinarayanan, and J. Zaccaria, "PPE Compliance Detection using Artificial Intelligence in Learning Factories," *Procedia Manufacturing*, vol. 45, pp. 277–282, Apr. 2020, https://doi.org/10.1016/j.promfg.2020.04.017.

[2] Md. Ferdous and Sk. Md. M. Ahsan, "PPE Detector: a YOLO-based Architecture to Detect Personal Protective Equipment (PPE) for Construction Sites," *PeerJ Computer Science*, vol. 8, June 2022, Art. no. e999, https://doi.org/10.7717/peerj-cs.999.

[3] Sutikno, A. Sugiharto, R. Kusumaningrum, and H. A. Wibawa, "Combination of HAAR, HOG, and LBP Descriptors for Enhanced Classification of Moving Objects and Motorcyclists Wearing Helmets," *Engineering, Technology & Applied Science Research*, vol. 15, no. 3, pp. 23283–23289, June 2025, https://doi.org/10.48084/etasr.10677.

[4] H. M. Ahmad and A. Rahimi, "SH17: A Dataset for Human Safety and Personal Protective Equipment Detection in Manufacturing Industry," *Journal of Safety Science and Resilience*, vol. 6, no. 2, pp. 175–185, June 2025, https://doi.org/10.1016/j.jnlssr.2024.09.002.

[5] M.-E. Otgonbold *et al.*, "SHEL5K: An Extended Dataset and Benchmarking for Safety Helmet Detection," *Sensors*, vol. 22, no. 6, Mar. 2022, Art. no. 2315, https://doi.org/10.3390/s22062315.

[6] F. Yu, J. Li, X. Wang, S. Wu, J. Zhang, and Z. Zeng, "Large, Complex, and Realistic Safety Clothing and Helmet Detection: Dataset and Method." arXiv, June 2023, https://doi.org/10.48550/ARXIV.2306.02098.

[7] S. Al-Azani, H. Luqman, M. Alfarraj, A. A. I. Sidig, A. H. Khan, and D. Al-Hamed, "Real-Time Monitoring of Personal Protective Equipment Compliance in Surveillance Cameras," *IEEE Access*, vol. 12, pp. 121882–121895, 2024, https://doi.org/10.1109/ACCESS.2024.3451117.

[8] A. A. Protik, A. H. Rafi, and S. Siddique, "Real-time Personal Protective Equipment (PPE) Detection Using YOLOv4 and TensorFlow," in *2021 IEEE Region 10 Symposium (TENSYMP)*, Jeju, Korea, Republic of, Aug. 2021, pp. 1–6, https://doi.org/10.1109/TENSYMP52854.2021.9550808.

[9] Z. Wang, Z. Cai, and Y. Wu, "An Improved YOLOX Approach for Low-light and Small Object Detection: PPE on Tunnel Construction Sites," *Journal of Computational Design and Engineering*, vol. 10, no. 3, pp. 1158–1175, Apr. 2023, https://doi.org/10.1093/jcde/qwad042.

[10] T. Diwan, G. Anirudh, and J. V. Tembhurne, "Object Detection Using YOLO: Challenges, Architectural Successors, Datasets and Applications," *Multimedia Tools and Applications*, vol. 82, no. 6, pp. 9243–9275, Mar. 2023, https://doi.org/10.1007/s11042-022-13644-y.

[11] N. Jegham, C. Y. Koh, M. Abdelatti, and A. Hendawi, "Yolo Evolution: A Comprehensive Benchmark and Architectural Review of Yolov12, Yolo11, and Their Previous Versions." SSRN, 2025, https://doi.org/10.2139/ssrn.5175639.

[12] J. Wu, N. Cai, W. Chen, H. Wang, and G. Wang, "Automatic Detection of Hardhats Worn by Construction Personnel: A Deep Learning Approach and Benchmark Dataset," *Automation in Construction*, vol. 106, Oct. 2019, Art. no. 102894, https://doi.org/10.1016/j.autcon.2019.102894.

[13] Md. S. Islam, S. Shaqib, S. S. Ramit, S. A. Khushbu, A. Sattar, and S. R. H. Noori, "A Deep Learning Approach to Detect Complete Safety Equipment for Construction Workers Based on YOLOv7." arXiv, 2024, https://doi.org/10.48550/ARXIV.2406.07707.

[14] 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, Apr. 2020, Art. no. 103085, https://doi.org/10.1016/j.autcon.2020.103085.

[15] Y. Alassaf and Y. Said, "DPPNet: A Deformable-Perspective-Perception Network for Safety Helmet Violation Detection," *Engineering, Technology & Applied Science Research*, vol. 14, no. 1, pp. 12659–12669, Feb. 2024, https://doi.org/10.48084/etasr.6633.

[16] R. Sapkota, Z. Meng, M. Churuvija, X. Du, Z. Ma, and M. Karkee, "Comprehensive Performance Evaluation of Yolov12, Yolo11, Yolov10, Yolov9 and Yolov8 on Detecting and Counting Fruitlet in Complex Orchard Environments." SSRN, 2025, https://doi.org/10.2139/ssrn.5201592.

[17] F. Zhafran, E. S. Ningrum, M. N. Tamara, and E. Kusumawati, "Computer Vision System Based for Personal Protective Equipment Detection, by Using Convolutional Neural Network," in *2019 International Electronics Symposium (IES)*, Surabaya, Indonesia, Sept. 2019, pp. 516–521, https://doi.org/10.1109/ELECSYM.2019.8901664.

[18] H. Liu and X. Qin, "Target Detection of Safety Protective Gear Using the Improved YOLOv5," in *2024 5th International Conference on Computers and Artificial Intelligence Technology (CAIT)*, Hangzhou, China, Dec. 2024, pp. 6–13, https://doi.org/10.1109/CAIT64506.2024.10962947.

[19] 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, May 2021, Art. no. 3478, https://doi.org/10.3390/s21103478.

[20] I. Majid Mohammed and N. Ashidi Mat Isa, "Contrast Limited Adaptive Local Histogram Equalization Method for Poor Contrast Image Enhancement," *IEEE Access*, vol. 13, pp. 62600–62632, 2025, https://doi.org/10.1109/ACCESS.2025.3558506.

[21] Y. R. Haddadi, B. Mansouri, and F. Z. I. Khodja, "A Novel Medical Image Enhancement Algorithm Based on CLAHE and Pelican Optimization," *Multimedia Tools and Applications*, vol. 83, no. 42, pp. 90069–90088, Apr. 2024, https://doi.org/10.1007/s11042-024-19070-6.

[22] Y. Jiang *et al.*, "EnlightenGAN: Deep Light Enhancement Without Paired Supervision," *IEEE Transactions on Image Processing*, vol. 30, pp. 2340–2349, 2021, https://doi.org/10.1109/TIP.2021.3051462.

[23] R. Y. Mahendra, W. Anggraeni, and M. H. Purnomo, "Low Light Image Enhancement with Small Training Dataset Using EnlightenGAN," in *2022 International Seminar on Intelligent Technology and Its Applications (ISITIA)*, Surabaya, Indonesia, July 2022, pp. 121–126, https://doi.org/10.1109/ISITIA56226.2022.9855360.

[24] F. Alzami, M. Naufal, H. A. Azies, S. Winarno, and M. A. Soeleman, "Time Distributed MobileNetV2 with Auto-CLAHE for Eye Region Drowsiness Detection in Low Light Conditions," *International Journal of Advanced Computer Science and Applications*, vol. 15, no. 11, 2024, https://doi.org/10.14569/IJACSA.2024.0151146.

[25] M. Yasin, F. Smarandache, M. Waheed Sabir, F. Arslan, and M. Waqas, "AI-driven Automated Helmet Detection in Underground Coal Mines using Attention-Enhanced Vision Transformer," *Engineering, Technology & Applied Science Research*, vol. 15, no. 3, pp. 23736–23741, June 2025, https://doi.org/10.48084/etasr.10868.

[26] Z. Jingchun, G. Eg Su, and M. Shahrizal Sunar, "Low-light Image Enhancement: A Comprehensive Review on Methods, Datasets and Evaluation Metrics," *Journal of King Saud University - Computer and Information Sciences*, vol. 36, no. 10, Dec. 2024, Art. no. 102234, https://doi.org/10.1016/j.jksuci.2024.102234.