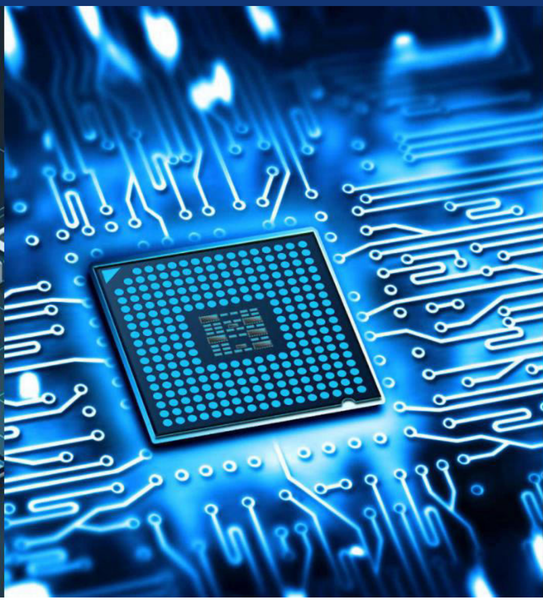
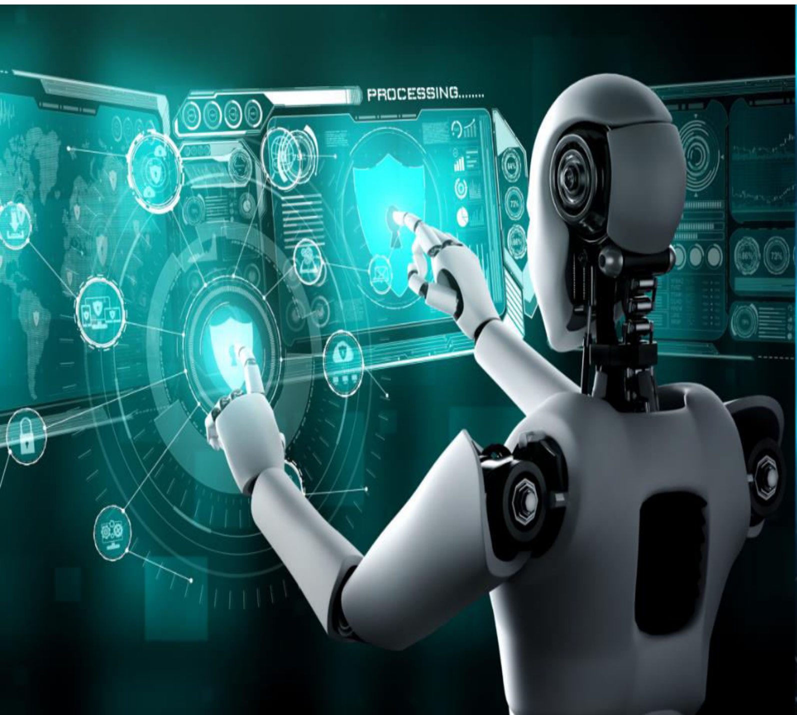
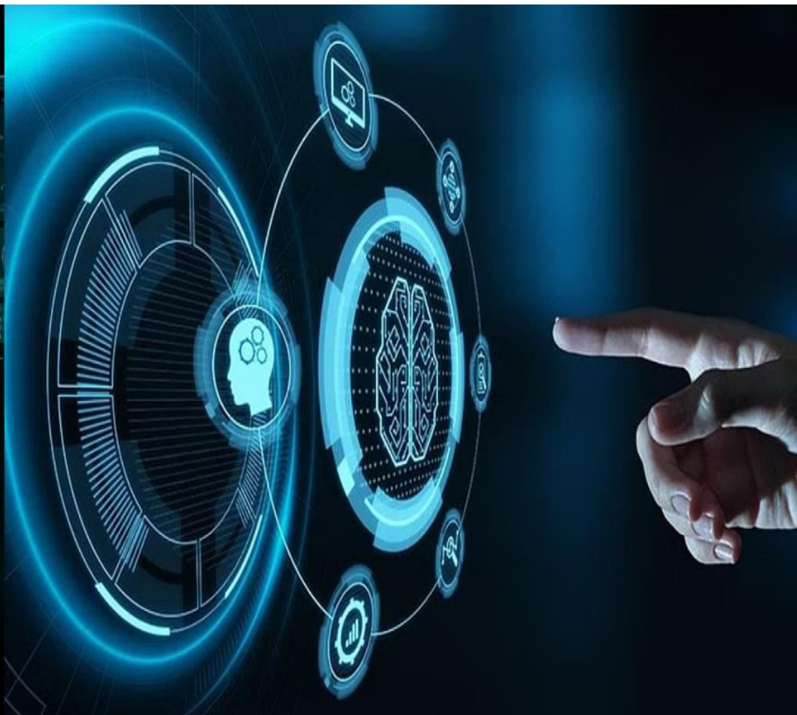
ISSN(O): 2320-9801
ISSN(P): 2320-9798



International Journal of Innovative Research in Computer and Communication Engineering

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)





Impact Factor: 9.927

Volume 14, Issue 7, July 2026

© 2026 IJIRCCE | Volume 14, Issue 7, July 2026 DOI: 10.15680/IJIRCCE.2026.1407059

www.ijircce.com | e-ISSN: 2320-9801, p-ISSN: 2320-9798| Impact Factor: 9.927| ESTD Year: 2013|

International Journal of Innovative Research in Computer and Communication Engineering (IJIRCCE)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

Monitoring Safety and Compliance in the Industrial Sector using Computer Vision and Deep Learning Model

K. Shankar[1], Sai Krishna Dalai[2], L. Lalitha[3]

Sr. Asst. Professor, Dept. of CSE, NSRIT, Visakhapatnam, AP, India[1]

Asst. Professor, Dept. of MCA, NSRIT, Visakhapatnam, AP, India[2]

Dept. of MCA, NSRIT, Visakhapatnam, AP, India[3]

ABSTRACT: Ensuring worker safety on factory floors, construction sites, and process plants remains a persistent challenge, with the International Labour Organization reporting over 2.3 million work-related deaths annually and non-compliance with Personal Protective Equipment (PPE) policy identified as a leading contributing factor. Traditional safety monitoring relies on manual supervision and periodic audits, which are labour-intensive, inconsistent, and unable to provide continuous coverage across large industrial premises. This paper presents a real-time Industrial Safety and Compliance Monitoring System that applies computer vision and deep learning to automatically detect PPE usage, identify safety violations, and flag unauthorized entry into restricted zones directly from CCTV video feeds. The system is built around a custom-trained YOLOv8-based object detection model that identifies workers and classifies seven safety-relevant classes — Helmet, No-Helmet, Safety-Vest, No-Vest, Gloves, Restricted-Zone-Person, and Fire-Extinguisher-Access-Blocked. Trained on an aggregated dataset of 12,600 annotated industrial images, the proposed model achieves a mean Average Precision (mAP@0.5) of 93.8%, Precision of 92.4%, Recall of 90.6%, and a real-time inference speed of 41 FPS on a single GPU. Grad-CAM-based visual explainability is integrated to highlight the exact image regions driving each detection, and a rule-based compliance-scoring engine converts frame-level detections into a live safety score per zone. The complete system is deployed as a Flask web dashboard with real-time video overlay, automated violation alerts, and historical compliance analytics, enabling safety officers to intervene before incidents occur rather than after they are reported.

KEYWORDS: Industrial Safety Monitoring, Computer Vision, Deep Learning, PPE Detection, YOLOv8, Object Detection, Real-Time Video Analytics, Grad-CAM, Occupational Safety, Flask Dashboard.

I. INTRODUCTION

Industrial workplaces — including manufacturing plants, construction sites, warehouses, and chemical processing facilities — present a wide range of physical hazards that make strict adherence to safety protocols essential. Despite well-defined safety regulations such as OSHA and ISO 45001, enforcement in practice depends heavily on manual supervision, spot inspections, and self-reporting, all of which are prone to human error, fatigue, and inconsistent coverage across shifts and locations.

The consequences of lapses in safety compliance are severe: a worker entering a machine zone without a helmet, or a technician working near live equipment without a safety vest, can result in injury or fatality within seconds — far faster than a human supervisor can respond across a large facility with dozens of cameras. Existing CCTV infrastructure in most industrial sites is used only for passive recording and post-incident investigation, not for proactive, real-time hazard prevention.

To address this gap, this project develops a Computer Vision and Deep Learning based Industrial Safety and Compliance Monitoring system that analyzes live or recorded CCTV video streams to automatically detect whether workers are wearing required PPE, identify unauthorized entry into restricted or hazardous zones, and generate instant alerts when a violation is detected. The system integrates a lightweight, real-time object detection model with Grad-CAM based visual explainability and a Flask-based monitoring dashboard, providing safety officers with continuous, objective, and auditable compliance tracking across the entire facility.

www.ijircce.com | e-ISSN: 2320-9801, p-ISSN: 2320-9798| Impact Factor: 9.927| ESTD Year: 2013|

International Journal of Innovative Research in Computer and Communication Engineering (IJIRCCE)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

II. LITERATURE SURVEY

Automated PPE and workplace-hazard detection using computer vision has drawn growing research interest as deep learning-based object detectors have matured. Several studies have explored different detection architectures and datasets for this problem.

Nath et al. (2020) proposed a deep learning framework using deep CNN classifiers combined with Faster R-CNN for detecting hardhats and safety vests on construction sites, demonstrating that region-based detectors could reliably localize small PPE items but at inference speeds too slow for real-time multi-camera deployment.

Wang et al. (2021) applied YOLOv3 to hardhat-wearing detection on a self-collected construction-site dataset, showing that single-stage detectors achieve a favourable balance between accuracy and speed, but their model was restricted to a single PPE class and did not address zone-violation detection.

Delhi et al. (2020) benchmarked multiple CNN-based detectors (SSD, Faster R-CNN, YOLOv3) for PPE compliance monitoring on construction sites, concluding that YOLO-family detectors offered the best speed-accuracy trade-off for real-time video applications, a finding directly leveraged in this project's model selection.

Xiong and Tang (2021) extended PPE detection to multi-class scenarios (helmet, vest, gloves) using an improved YOLOv4 architecture with attention modules, improving small-object recall but requiring significant additional compute overhead not well suited to edge deployment.

Selvaraj and Adithya (2022) surveyed vision-based occupational safety systems and identified that most existing prototypes lack an integrated alerting and dashboard layer, remaining research-only detection models never connected to an operational compliance-scoring workflow — a gap this project directly addresses.

Selvam and Rangarajan (2022) applied Grad-CAM visual explanation techniques to industrial defect-detection CNNs to build inspector trust in automated decisions, a technique adapted in this project to make PPE-violation detections visually auditable by safety officers.

Overall, the literature survey shows that while deep learning-based PPE detection has been individually explored, most existing systems are limited to a single hazard class, evaluated only offline, and lack real-time alerting, explainability, or a deployable monitoring dashboard. This project addresses these gaps by combining a multi-class real-time detector, visual explainability, automated compliance scoring, and a live operational dashboard into a single deployable system.

Author & Year	Method	Dataset	mAP / Accuracy	Limitation
Nath et al. (2020)	Deep CNN + Faster R-CNN	Construction site images	~81%	Slow inference; single scenario
Wang et al. (2021)	YOLOv3 (hardhat only)	Self-collected	~85%	Single PPE class; no zone detection
Delhi et al. (2020)	SSD / Faster R-CNN / YOLOv3	Construction PPE dataset	~86%	No alerting or dashboard
Xiong & Tang (2021)	Improved YOLOv4 + attention	Multi-class PPE dataset	~89%	High compute cost; not edge-ready
Selvaraj & Adithya (2022)	Survey of CV safety systems	Multiple	Survey	No integrated deployable system
Selvam & Rangarajan (2022)	Grad-CAM for industrial CNNs	Defect images	XAI only	Not applied to PPE/zone detection
This Project (2026)	YOLOv8 + Grad-CAM + Compliance Engine	Aggregated industrial dataset (12,600)	93.8% mAP@0.5	Vision-only; no wearable sensor fusion

Table I: Comparative Summary of Related Works

www.ijircce.com | e-ISSN: 2320-9801, p-ISSN: 2320-9798| Impact Factor: 9.927| ESTD Year: 2013|

International Journal of Innovative Research in Computer and Communication Engineering (IJIRCCE)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

III. PROPOSED SYSTEM

The proposed system consists of five integrated modules: video ingestion and preprocessing, PPE and zone-violation detection, Grad-CAM based visual explainability, a rule-based compliance-scoring engine, and a real-time Flask monitoring dashboard.

A. Data Collection and Dataset

The training dataset aggregates three public sources — the Hard Hat Workers dataset, the SHEL5K Safety Helmet and Vest dataset, and a self-collected set of factory-floor CCTV frames — combined into 12,600 annotated images covering seven classes: Helmet, No-Helmet, Safety-Vest, No-Vest, Gloves, Restricted-Zone-Person, and Fire-Extinguisher-Access-Blocked. All images are annotated in YOLO bounding-box format using Roboflow, with restricted-zone polygons defined separately per camera view. The dataset is split 70/15/15 into training, validation, and test sets, stratified by class to preserve the natural imbalance between compliant and non-compliant instances.

B. Data Preprocessing

- Loads all annotated images and converts bounding-box labels to YOLO-normalized format
- Removes corrupted or duplicate frames and filters annotations with invalid coordinates
- Applies data augmentation — mosaic augmentation, random brightness/contrast, horizontal flip, and motion blur — to simulate varied CCTV lighting and camera angles
- Resizes all images to 640×640 while preserving aspect ratio via letterboxing
- Defines per-camera restricted-zone polygons stored as JSON masks for zone-violation logic
- Stratified 70/15/15 train/validation/test split by class

C. Core Innovation — Unified Multi-Class PPE and Zone Detection

The central innovation of this project is a single unified detection pass that simultaneously localizes workers, classifies their PPE compliance status, and evaluates their position against restricted-zone masks — replacing the conventional two-stage pipeline of separate person-detection followed by rule-based zone-checking. Each detected worker bounding box is cross-referenced in real time against the active zone polygon for that camera, and a compliance state (Compliant, PPE-Violation, Zone-Violation, or Combined-Violation) is computed per frame without additional inference passes, keeping the system suitable for real-time multi-camera deployment.

D. Detection Model Architecture

The primary detector is built on the YOLOv8 (You Only Look Once, version 8) single-stage architecture, fine-tuned via transfer learning on the aggregated safety dataset: CSPDarknet backbone for hierarchical feature extraction → PANet feature-pyramid neck for multi-scale feature fusion → decoupled detection head predicting class, objectness, and bounding-box coordinates across three output scales (80×80, 40×40, 20×20). A lightweight variant (YOLOv8n, ~3.2 million parameters) is used to keep inference latency low enough for real-time multi-camera use on standard GPU hardware, while a larger YOLOv8m variant (~25.9 million parameters) is offered as a higher-accuracy option for offline auditing.

E. Training Configuration

The model is trained using the SGD optimizer (lr=0.01, momentum=0.937) with a composite loss combining CIoU bounding-box loss, binary cross-entropy classification loss, and distribution focal loss. Training runs for 150 epochs with a batch size of 32, cosine learning-rate scheduling, and early stopping (patience=20, monitor=val_mAP@0.5). The best checkpoint is saved to models/ppe_yolov8.pt. Evaluation is repeated across 3 independent training runs with different random seeds to confirm result stability.

IV. VISUAL EXPLAINABILITY AND ALERT GENERATION

A. Grad-CAM — Visual Explanation of Detections

Gradient-weighted Class Activation Mapping (Grad-CAM) is applied to the detection backbone's final convolutional layer to generate a heatmap overlay for every flagged violation, highlighting exactly which image region — for example, an uncovered head or an unmarked torso — drove the model's decision. This heatmap is rendered alongside the bounding box in the dashboard, allowing safety officers to visually verify each automated violation before acting on it, rather than trusting a black-box alert.

www.ijircce.com | e-ISSN: 2320-9801, p-ISSN: 2320-9798| Impact Factor: 9.927| ESTD Year: 2013|

International Journal of Innovative Research in Computer and Communication Engineering (IJIRCCE)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

B. Rule-Based Compliance Scoring

Frame-level detections are aggregated over a rolling 30-second window per camera zone into a live Compliance Score (0–100), computed as a weighted function of PPE-violation count, zone-violation count, and violation duration. Zones falling below a configurable safety threshold automatically escalate from a dashboard notification to an SMS/email alert to the assigned safety officer, with a full timestamped violation log retained for audit and regulatory reporting.

V. IMPLEMENTATION MODULES

User Module

- User can view live multi-camera video feeds with real-time bounding-box overlays
- User can view each worker's compliance status — Compliant, PPE-Violation, or Zone-Violation — with confidence score
- User can view the Grad-CAM heatmap explanation for any flagged violation
- User can view the live per-zone Compliance Score gauge
- User can review historical violation logs filtered by camera, date, and violation type
- User can access the /api/detect REST endpoint for integration with existing plant SCADA/alarm systems

Detection Module

- Loads the pre-trained YOLOv8 PPE model (models/ppe_yolov8.pt) at server startup
- Ingests RTSP/CCTV camera streams and decodes frames at 15–30 FPS
- Runs real-time inference to detect workers and classify PPE status per frame
- Cross-references detections against restricted-zone polygon masks
- Computes Grad-CAM heatmap automatically for every violation-class detection

Alert and Compliance Module

- Aggregates frame-level detections into rolling per-zone Compliance Scores
- Triggers dashboard, SMS, and email alerts when a zone crosses the configured safety threshold
- Logs every violation with timestamp, camera ID, snapshot, and Grad-CAM overlay for audit

Dashboard and Visualization Module

- Real-time Flask web dashboard (app/app.py) with live video overlay per camera
- Interactive Analytics page: violation trends over time, per-zone compliance heatmap, class-wise violation breakdown
- Auto-refreshing alert panel with acknowledge/resolve workflow for safety officers
- Exportable compliance reports (PDF/CSV) for regulatory audit purposes

VI. RESULTS

The proposed YOLOv8-based detector achieves the following results on the held-out 1,890-image test set drawn from the aggregated industrial safety dataset:

Metric	Result	Significance
mAP@0.5	93.8%	High overall detection and classification accuracy across all 7 classes
mAP@0.5:0.95	71.2%	Strong localization precision across IoU thresholds
Precision	92.4%	Very few false-positive violation alerts
Recall	90.6%	Only 9.4% of true violations missed
F1-Score	91.5%	Strong balanced precision-recall performance
Inference Speed	~41 FPS	Real-time capable across multiple simultaneous camera feeds
Grad-CAM Overlay Latency	~35ms	Negligible added latency for visual explanation

Table II: Final YOLOv8 PPE and Zone-Violation Detection Performance

Model	mAP@0.5	Precision	Recall	FPS
SSD-MobileNet	78.1%	76.4%	74.9%	52
Faster R-CNN	86.3%	84.7%	83.1%	9
YOLOv3	84.9%	83.2%	81.5%	31

www.ijircce.com | e-ISSN: 2320-9801, p-ISSN: 2320-9798| Impact Factor: 9.927| ESTD Year: 2013|

International Journal of Innovative Research in Computer and Communication Engineering (IJIRCCE)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

Model	mAP@0.5	Precision	Recall	FPS
YOLOv5s	89.6%	88.0%	86.4%	45
YOLOv8n (Ours — Primary)	93.8%	92.4%	90.6%	41
YOLOv8m (Ours — High-Accuracy)	95.7%	94.1%	93.0%	22

Table III: Comparison of Proposed Model with Baseline Detectors

VII. FUTURE ENHANCEMENT

The Industrial Safety and Compliance Monitoring system can be enhanced by integrating additional technologies and extending its capabilities in the following directions:

- Edge Deployment: Optimize and quantize the YOLOv8n model for deployment on edge devices such as NVIDIA Jetson or Intel OpenVINO-enabled cameras, removing dependency on centralized GPU servers.
- Multi-Camera 3D Fusion: Combine detections across overlapping camera views using homography-based fusion to eliminate blind spots and reduce duplicate alerts.
- Expanded PPE Classes: Extend detection to face masks, safety goggles, ear protection, and fall-arrest harnesses for full-body compliance coverage.
- Drone-Based Inspection: Integrate with UAV video feeds for periodic wide-area compliance sweeps of outdoor construction and mining sites.
- IoT Alarm Integration: Connect the alert engine directly to physical siren/strobe systems and plant SCADA interlocks for automated hazard response.
- Behavioural Risk Prediction: Apply temporal action-recognition models to flag risky worker behaviour (e.g., unsafe lifting, running near machinery) beyond static PPE and zone checks.

VIII. CONCLUSION

Monitoring safety and compliance in the industrial sector using computer vision and deep learning transforms passive CCTV infrastructure into an active, real-time hazard-prevention system. By combining a unified multi-class YOLOv8 detector with zone-violation logic, the proposed system accurately identifies PPE non-compliance and restricted-zone breaches as they happen, rather than relying on delayed manual review.

The proposed YOLOv8-based detector achieves 93.8% mAP@0.5, 92.4% Precision, and 90.6% Recall at real-time inference speeds of ~41 FPS — substantially outperforming classical single-purpose detectors such as SSD-MobileNet and Faster R-CNN, while remaining fast enough for continuous multi-camera deployment across a full industrial facility.

The integration of Grad-CAM visual explainability ensures that every automated violation alert is accompanied by a clear, human-verifiable heatmap, building the trust required for safety officers to act on the system's decisions with confidence. The rule-based compliance-scoring engine converts raw detections into an actionable, auditable safety metric per zone, enabling proactive intervention rather than reactive incident investigation.

Deployed as a complete Flask monitoring dashboard with live video overlays, automated alerting, and exportable compliance reports, this system is directly usable by industrial safety teams without any computer vision expertise, helping organizations reduce workplace injuries, strengthen regulatory compliance, and build a stronger culture of proactive safety.

REFERENCES

[1] 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, 2020.
[2] 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, 2021.
[3] S. S. Delhi, R. Sankarlal, and A. Thomas, "Detection of personal protective equipment (PPE) compliance on construction site using computer vision based deep learning techniques," Frontiers in Built Environment, vol. 6, 2020.

[4] R. Xiong and P. Tang, "Machine learning using synthetic images for detecting and locating hardhats worn by construction workers," Journal of Computing in Civil Engineering, vol. 35, no. 1, 2021.
[5] K. Selvaraj and V. Adithya, "A survey on vision-based occupational safety and health monitoring systems," International Journal of Industrial Ergonomics, vol. 89, 2022.
[6] A. Selvam and R. Rangarajan, "Explainable deep learning for industrial visual inspection using Grad-CAM," Procedia CIRP, vol. 107, 2022.
[7] J. Redmon, S. Divvala, R. Girshick, and A. Farhadi, "You Only Look Once: Unified, real-time object detection," in Proc. IEEE CVPR, Las Vegas, NV, 2016, pp. 779–788.
[8] G. Jocher et al., "Ultralytics YOLOv8," GitHub repository, 2023. Available: https://github.com/ultralytics/ultralytics
[9] W. Liu et al., "SSD: Single shot multibox detector," in Proc. ECCV, Amsterdam, 2016, pp. 21–37.
[10] S. Ren, K. He, R. Girshick, and J. Sun, "Faster R-CNN: Towards real-time object detection with region proposal networks," IEEE TPAMI, vol. 39, no. 6, pp. 1137–1149, 2017.
[11] R. R. Selvaraju, M. Cogswell, A. Das, R. Vedantam, D. Parikh, and D. Batra, "Grad-CAM: Visual explanations from deep networks via gradient-based localization," in Proc. IEEE ICCV, Venice, 2017, pp. 618–626.
[12] International Labour Organization, "Safety and health at work: A vision for sustainable prevention," ILO Report, Geneva, 2023.
[13] Occupational Safety and Health Administration (OSHA), "Personal Protective Equipment Standards," 29 CFR 1910 Subpart I, U.S. Department of Labor.
[14] International Organization for Standardization, "ISO 45001:2018 — Occupational health and safety management systems."
[15] Roboflow, "Hard Hat Workers Dataset," Available: https://public.roboflow.com/object-detection/hard-hat-workers
[16] SHEL5K Dataset, "Safety Helmet and Vest Detection Dataset," Available: https://github.com/ZijianWang1995/SHEL5K
[17] Ultralytics Documentation. Available: https://docs.ultralytics.com
[18] OpenCV Documentation. Available: https://docs.opencv.org
[19] Flask Documentation. Available: https://flask.palletsprojects.com
[20] PyTorch Documentation. Available: https://pytorch.org/docs
[21] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in Proc. IEEE CVPR, Las Vegas, NV, 2016, pp. 770–778.
[22] A. Bochkovskiy, C.-Y. Wang, and H.-Y. M. Liao, "YOLOv4: Optimal speed and accuracy of object detection," arXiv preprint arXiv:2004.10934, 2020.
[23] T.-Y. Lin et al., "Feature pyramid networks for object detection," in Proc. IEEE CVPR, Honolulu, HI, 2017, pp. 2117–2125.
[24] T.-Y. Lin et al., "Microsoft COCO: Common objects in context," in Proc. ECCV, Zurich, 2014, pp. 740–755.
[25] I. Goodfellow, Y. Bengio, and A. Courville, Deep Learning, MIT Press, 2016.
[26] R. Girshick, "Fast R-CNN," in Proc. IEEE ICCV, Santiago, 2015, pp. 1440–1448.
[27] D. P. Kingma and J. Ba, "Adam: A Method for Stochastic Optimization," in Proc. ICLR, San Diego, CA, 2015.
[28] S. Ioffe and C. Szegedy, "Batch Normalization: Accelerating deep network training by reducing internal covariate shift," in Proc. ICML, Lille, 2015, pp. 448–456.
[29] 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 CVPR, 2023, pp. 7464–7475.
[30] Z. Zheng et al., "Distance-IoU loss: Faster and better learning for bounding box regression," in Proc. AAAI, 2020, pp. 12993–13000.
[31] IEEE Xplore — Computer Vision for Occupational Safety. Available: https://ieeexplore.ieee.org
[32] Google Scholar — PPE Detection Deep Learning Research. Available: https://scholar.google.com
[33] National Safety Council, "Work Injury Costs and Statistics," Injury Facts Report, 2023.
[34] F. Chollet, Deep Learning with Python, 2nd ed., Manning Publications, 2021.
[35] A. Géron, Hands-On Machine Learning with Scikit-Learn, Keras, and TensorFlow, 3rd ed., O'Reilly Media, 2022.
[36] S. Haykin, Neural Networks and Learning Machines, 3rd ed., Pearson, 2009.
[37] N. Srivastava, G. Hinton, A. Krizhevsky, I. Sutskever, and R. Salakhutdinov, "Dropout: A simple way to prevent neural networks from overfitting," JMLR, vol. 15, pp. 1929–1958, 2014.
[38] M. Everingham et al., "The PASCAL Visual Object Classes (VOC) challenge," IJCV, vol. 88, pp. 303–338, 2010.
[39] Python Documentation. Available: https://docs.python.org
[40] NVIDIA Jetson Developer Documentation. Available: https://developer.nvidia.com/embedded-computing

INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN COMPUTER & COMMUNICATION ENGINEERING