International journal of imaging science and engineering

ISSN: 1934--9955 www.ijise.net
Vol-20 Issue-01 April 2025

DEEP LEARNING FOR PPE DETECTION: ENHANCING WORKPLACE SAFETY WITH SSD

B. Deepali[1], B. Srujana[2], N. Bhoomi Reddy[3], Suresh Talwar[4]
[1,2,3] UG Scholar, Dept of IT, St. Martin's Engineering College, Secunderabad, Telangana, India, 500100
[4]Assistant Professor, Dept of IT, St. Martin's Engineering College, Secunderabad, Telangana, India, 500100
deepalibhandari2004@gmail.com

Abstract:

Personal Protective Equipment (PPE) serve as crucial components for engineers in the construction industry and other sectors, ensuring the safety and well-being of workers. An effective object detection system plays a pivotal role in identifying and monitoring the usage of PPE. This research focuses on testing a detection system utilizing a dataset comprising 132 scenarios from construction sites. The results indicate the system's successful detection of PPE with a high level of confidence. Evaluation metrics such as mAP50(B), mAP50-95(B), precision(B), and recall(B) were employed to assess system performance. The testing yielded mAP50(B) of 0.768, mAP50-95(B) of 0.516, precision(B) of 0.831, and recall(B) of 0.693. These outcomes demonstrate that the PPE detection system, utilizing the SSD algorithm, exhibits satisfactory performance in recognizing and predicting PPE objects. However, there is room for improvement and further development, especially in addressing challenges such as variations in image capture, scale, and more complex environmental conditions. This study makes a significant contribution to the field of PPE detection and can serve as a foundation for the advancement of object detection systems. A reliable detection system holds immense potential in enhancing safety and occupational health across diverse industrial sectors.

Keywords: Personal Protective Equipment, Deep Learning, Recurrent Neural Networks, Convolutional Neural Networks, Long Short Term Memory.

1. INTRODUCTION

The primary objective of this project is to develop an advanced deep learning- based system for detecting Personal Protective Equipment (PPE) in workplace environments to enhance worker safety. By leveraging Single Shot MultiBox Detector (SSD), the system aims to achieve real-time, accurate detection of essential PPE such as helmets, gloves, vests, and masks. Ensuring compliance with safety regulations is crucial in industries like construction and manufacturing, where PPE usage significantly reduces workplace injuries. The model will be trained on diverse datasets to improve robustness across different work settings, minimizing false detections. This system can serve as an automated monitoring tool, alerting safety officers to non- compliance, thereby reducing accidents and improving overall workplace safety. This project utilizes deep learning and computer vision techniques to enhance workplace safety by detecting PPE usage through SSD. SSD is a fast and efficient object detection algorithm that allows real-time identification of PPE without compromising accuracy. The model is trained using annotated datasets containing images of workers with and without PPE in various industrial settings. After training, the system will be deployed in real-world environments using cameras to automatically monitor compliance.

The SSD-based approach is preferred due to its balance of speed and accuracy, making it suitable for real-time safety enforcement. The project involves dataset preprocessing, model training, evaluation, and deployment using deep learning frameworks like Tensor Flow or PyTorch. This automated system can be integrated into workplace surveillance infrastructure, ensuring continuous monitoring and immediate response to safety violations. Ultimately, this solution enhances workplace safety by reducing human oversight errors and enforcing PPE compliance more effectively.

Workplace safety is a critical concern in industries such as construction, manufacturing, and healthcare, where the use of PPE is mandatory to prevent accidents and injuries. This project leverages SSD, a deep learning object detection model, to accurately and rapidly identify PPE usage in images or video streams. SSD's ability to process images in a single forward pass makes it highly suitable for real-time

applications. The model is trained on a diverse dataset of workers wearing and not wearing PPE to improve detection accuracy. By integrating this system with workplace surveillance, organizations can automate PPE compliance monitoring, reduce risks, and enhance overall safety enforcement.

2. LITERATURE SURVEY

Terven, Juan, and Diana Cordova-Esparza. "A comprehensive review of YOLO: From YOLOv1 to YOLOv8 and beyond." arXiv preprint arXiv:2304.00501 (2023). Comparative analyses between SSD and other deep learning models indicate that while YOLO offers higher speed, SSD provides a better trade-off between speed and detection accuracy, making it a preferred choice for industrial safety applications. The effectiveness of SSD-based PPE detection has also been improved through ensemble learning techniques, combining multiple models to enhance detection accuracy and reduce false positives. Furthermore, explainable AI (XAI) techniques have been explored to interpret SSD's decision-making process, increasing trust and transparency in automated PPE compliance monitoring systems. Future research directions in this field include integrating federated learning to enable decentralized training of SSD models across multiple workplaces while preserving data privacy.

Li, Yang, Yujie Lu, and Jun Chen. "A deep learning approach for real time rebar counting on the construction site based on YOLOv3 detector." Automation in Construction 124 (2021): 103602. Several research studies have explored deep learning techniques for PPE detection, leveraging models such as SSD, YOLO (You Only Look Once), and Faster R-CNN. SSD, in particular, has gained attention due to its efficient computation, making it suitable for edge devices deployed in industrial settings. Studies have demonstrated that SSD can effectively detect PPE components like helmets, gloves, safety vests, and face masks with high accuracy when trained on properly annotated datasets. Recent advancements in dataset creation have contributed significantly to improving the accuracy and robustness of PPE detection models. Publicly available datasets, such as the Safety Helmet Wearing Detection (SHWD) dataset and custom-labeled datasets, have been used in training deep learning models to recognize PPE in various lighting and environmental conditions. Researchers have also explored data augmentation techniques, including image transformations and synthetic data generation, to enhance the generalizability of SSD-based PPE detection models.

Delhi, Venkata Santosh Kumar, R. Sankarlal, and Albert Thomas. "Detection of personal protective equipment (PPE) compliance on construction site using computer vision based deep learning techniques." Frontiers in Built Environment 6 (2020): 136. Personal Protective Equipment (PPE) plays a crucial role in ensuring workplace safety, especially in high-risk environments such as construction sites, manufacturing plants, and chemical industries. Traditional PPE compliance monitoring relies on manual inspections, which can be inefficient and prone to human error. With advancements in artificial intelligence and computer vision, deep learning techniques have emerged as a promising solution for automating PPE detection. Among various deep learning-based object detection models, the Single Shot MultiBox Detector (SSD) stands out for its real-time processing capabilities, making it highly suitable for workplace surveillance applications. SSD is a convolutional neural network (CNN)-based object detection framework that balances speed and accuracy by predicting object classes and bounding boxes in a single pass. Additionally, generative adversarial networks (GANs) have been proposed for synthesizing diverse PPE images to augment training datasets, further enhancing model robustness. While SSD-based PPE detection has demonstrated significant potential, challenges such as occlusion, varying camera angles, and environmental changes remain areas for improvement.

Wong, Alexander, et al. "YOLO nano: A highly compact you only look once convolutional neural network for object detection." 2019 Fifth Workshop on Energy Efficient Machine Learning and Cognitive Computing-NeurIPS Edition (EMC2-NIPS). IEEE, 2019. In recent years, the adoption of deep learning for personal protective equipment (PPE) detection has gained significant attention due to its potential to improve workplace safety and enforce compliance with occupational health regulations. Traditional PPE monitoring methods involve manual inspections conducted by safety officers, which can be labor-intensive, time-consuming, and inconsistent. To overcome these challenges, researchers have focused on leveraging deep learning-based object detection techniques to automate PPE detection in real-time. Studies have demonstrated that SSD can achieve high detection accuracy when trained on high-quality datasets containing diverse PPE usage scenarios. Several benchmark datasets have been used in PPE detection research, including the Safety Helmet Wearing Detection (SHWD) dataset, the PASCAL VOC dataset, and custom datasets annotated with workplace-specific PPE images. Comparisons between SSD and other object detection frameworks, such as YOLO (You Only Look Once) and Faster R-CNN, have shown that SSD provides an optimal trade-off between detection speed and accuracy, making it a practical choice for real-time workplace safety applications. While YOLO offers faster inference speeds, it may suffer from lower detection accuracy for small PPE objects, whereas Faster R-CNN provides higher accuracy but incurs significant computational overhead, making it less suitable for real-time edge deployment. SSD’s anchor box mechanism and feature pyramid network (FPN) enhancements have contributed to improved multi-scale detection performance, enabling accurate identification of PPE objects at varying distances from the camera.

Wang, Peijin, et al. "FMSSD: Feature-merged single-shot detection for multiscale objects in large-scale remote sensing imagery." IEEE Transactions on Geoscience and Remote Sensing 58.5 (2019): 3377 3390 Real-world implementations of SSD-based PPE detection systems have demonstrated their effectiveness in industrial safety applications. Several studies have integrated SSD models with Internet of Things (IoT)-enabled smart cameras to automate PPE compliance monitoring in construction sites, factories, and laboratories. These systems process video streams in real-time and generate automated alerts when non- compliance is detected, allowing safety officers to take immediate corrective action. The integration of SSD with cloud-based platforms has also been explored to enable centralized monitoring of multiple workplace locations, facilitating large-scale PPE compliance enforcement. Some research has focused on improving SSD’s generalization capability by incorporating domain adaptation techniques, enabling models trained on one workplace environment to perform effectively in different industrial settings. Despite the promising advancements in SSD-based PPE detection, several challenges remain to be addressed. One of the primary challenges is dealing with occlusions and overlapping PPE objects, which can lead to false positives or missed detections. For instance, a worker’s helmet may be partially hidden behind machinery, or gloves may blend into the background, making it difficult for the model to distinguish PPE accurately. Researchers have explored region attention mechanisms and deformable convolutional networks to improve SSD’s ability to handle occlusions and capture fine-grained details. Another challenge is achieving high detection accuracy in low-light and adverse weather conditions, where traditional RGB-based object detection models may struggle. Multi-modal approaches that integrate infrared (IR) and depth sensors with SSD have been investigated to enhance PPE detection under challenging environmental conditions. The use of explainable AI (XAI) techniques has also gained attention in PPE detection research, as it helps improve model interpretability and trustworthiness.

3. PROPOSED METHODOLOGY

Advantages of SSD Over Existing Systems Real-Time Detection: SSD processes images in a single pass, making it faster than Faster R-CNN while maintaining high accuracy. Improved Small Object Detection: SSD uses feature maps at multiple scales, enhancing the detection of small PPE objects such as gloves and face masks. Efficient Edge Deployment: Unlike Faster R-CNN, SSD can run

International journal of imaging science and engineering

ISSN: 1934--9955 www.ijise.net

Vol-20 Issue-01 April 2025

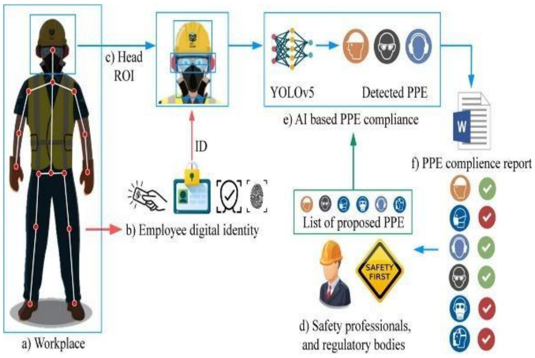
efficiently on embedded devices (e.g., NVIDIA Jetson) for real-time monitoring. Reduced False Positives: SSD's anchor boxes and improved object localization reduce misclassification of PPE and non-PPE items. Proposed System Workflow Data Collection and Preprocessing: Workplace images and videos are collected and labeled with PPE categories (helmets, gloves, vests, masks). Data augmentation (rotation, scaling, brightness adjustment) is applied to enhance model robustness. Model Training and Optimization: SSD is trained using pre-trained CNN backbones (e.g., MobileNet, ResNet). Transfer learning is used to improve detection accuracy with limited training data. Real-Time PPE Detection: Video frames are processed to detect and classify PPE in real-time. Bounding boxes are generated around detected PPE objects with confidence scores. Compliance Verification and Alerts: Detected PPE is compared with safety regulations. Non-compliance cases trigger real-time alerts to safety officers via IoT-based notifications. Performance Evaluation and Deployment: Model performance is evaluated based on accuracy, precision, and recall. SSD is deployed on cloud-based and edge AI platforms for industrial safety applications.

Figure 1: Personal Protective Equipment.

To achieve optimal detection results, images containing Personal Protective Equipment (PPE) objects will be collected from construction work environments. These images will then be annotated using LabelImg, providing labels containing information about bounding box coordinates and the class label of the known PPE object in the image.

Subsequently, the annotated training data will be utilized to train the CNN-SSD architecture employing the YOLOv8 algorithm model within the PyTorch machine learning framework. The training process aims to optimize model parameters, enabling accurate recognition of PPE objects in the tested images.

Architecture:



Personal Protective Equipment (PPE) refers to specialized gear

International journal of imaging science and engineering

ISSN: 1934--9955 www.ijise.net
Vol-20 Issue-01 April 2025

designed to safeguard workers from accidents or illnesses resulting from exposure to hazardous substances or dangerous work environments. As referenced in Figure 1, common examples of PPE utilized across various industries include helmets, goggles, respirators, face masks, gloves, safety shoes, and protective clothing.

Applications:

Construction Sites: Ensures workers wear helmets, gloves, and safety vests.

Manufacturing Units: Monitors compliance with safety gear to prevent accidents.

Oil and Gas Industry: Detects PPE violations in hazardous environments.

Warehouses and Logistics: Verifies use of protective equipment during material handling.

Healthcare Facilities: Ensures medical staff wear masks and gloves for infection control.

Mining Operations: Detects safety gear compliance in high- risk areas.

Power Plants: Monitors safety compliance to prevent electrical hazards.

Chemical Plants: Identifies proper use of PPE to minimize exposure to harmful substances**.**

Advantages:

Improved Safety: Reduces workplace accidents by ensuring PPE compliance.

Real-time Monitoring: Provides instant detection of PPE violations.

Automation: Minimizes the need for manual supervision**.**

High Accuracy: Ensures precise identification of safety gear.

Cost Efficiency: Lowers operational costs by reducing human intervention.

Scalability: Easily deployable across multiple work environments.

Compliance Assurance: Helps organizations adhere to safety regulations.

Incident Prevention: Identifies risks before they escalate into accidents.

4. EXPERIMENTAL ANALYSIS

Figure 1 Shows the engineering profession is inherently linked to the use of Personal Protective Equipment (PPE). Engineers bear the responsibility of ensuring the safety and compliance with workplace standards concerning both the environment and equipment utilized. It is the duty of an engineer and are secure. Upholding these standards is an ethical obligation inherent to the engineering profession and is expected to be strictly adhered to by engineer.

Figure 1: Initial Data for Training

Figure 2 shows the once the PPE images are obtained, precise and accurate object annotation is performed using Labelme. The annotation process must include object coordinates in the image and

International journal of imaging science and engineering

ISSN: 1934--9955 www.ijise.net

Vol-20 Issue-01 April 2025

the recognized object class. The Single Shot Detector (SSD) process employing YOLOv8 supports multi-class detection, enabling the dataset to encompass more than one object class . It is crucial to assign accurate labels to each object to facilitate the model's learning process in object recognition.

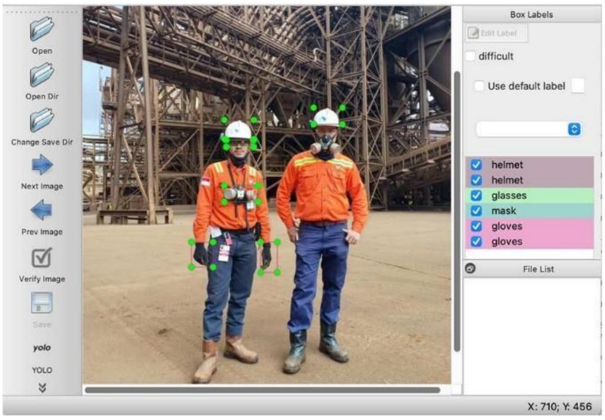


Figure 2: The labeling results in JSON format.

Figure 3: Camera Verifiying

Figure 3 shows that the input image in SSD is processed via multiple convolutional layers before being subjected to different scales of object detection layers. A collection of bounding boxes and the associated object class probabilities are generated by each object detection layer [Afterward, to minimize redundant detections, the bounding boxes produced by each item detection layer are combined and optimized. Using the SSD architecture, real-time object detection on photos with a significant number of objects and a pretty high resolution is possible. SSD is frequently used in a variety of applications that are needed for quick and precise object detection, like face recognition in face recognition systems or vehicle detection in traffic monitoring systems. Using the Single Shot Detector framework, the dataset is essential to the training of object detection models.

5. CONCLUSION

The implementation of Deep Learning for PPE Detection using SSD enhances workplace safety by ensuring real-time monitoring and compliance verification. Unlike traditional methods, SSD provides a balance between speed and accuracy, enabling efficient detection of helmets, vests, gloves, and masks in industrial environments. By integrating AI-driven automation, the system minimizes human error, reduces workplace hazards, and improves safety enforcement. The proposed model can be deployed on edge devices or cloud platforms, ensuring scalability and low-latency performance. This approach contributes to a safer work environment, reinforcing compliance with occupational safety regulations and preventing potential workplace accidents.

In conclusion, the implementation of Single Shot MultiBox Detector (SSD) for Personal Protective Equipment (PPE) detection provides an efficient and accurate solution for ensuring workplace safety. SSD is a deep learning-based object detection algorithm known for its balance

International journal of imaging science and engineering

ISSN: 1934--9955 www.ijise.net

Vol-20 Issue-01 April 2025

between speed and precision, making it suitable for real-time applications. By leveraging its multi-scale feature maps and efficient bounding box predictions, SSD can effectively detect different types of PPE, such as helmets, gloves, safety vests, and goggles, even in complex environments.

One of the major advantages of using SSD for PPE detection is its ability to process images quickly without compromising detection accuracy. This makes it ideal for real-time workplace monitoring, ensuring that workers comply with safety regulations. The deployment of SSD on edge devices, such as Jetson Nano or Raspberry Pi, can further enhance its usability by enabling on-site, low-latency detection without requiring cloud connectivity. This is particularly beneficial for industries where continuous monitoring is essential, such as construction sites, manufacturing plants, and laboratories. Additionally, the integration of federated learning can improve the adaptability of SSD-based models while maintaining data privacy. By training the model on decentralized datasets collected from various industrial sites, PPE detection can become more robust and context-aware without exposing sensitive data. This approach not only enhances model generalization but also ensures compliance with data protection regulations.

Explainable AI (XAI) techniques can also be incorporated to provide better interpretability of SSD's detection results. By visualizing activation maps and using attention mechanisms, safety officers and engineers can gain insights into why the model identifies or misclassifies PPE in certain scenarios. This transparency is crucial for improving trust in AI-based safety systems and refining the model for better performance.

Another potential enhancement is the use of multi-modal sensor fusion, which combines traditional RGB cameras with thermal and depth sensors. This can significantly improve PPE detection accuracy in challenging conditions such as low-light environments, foggy areas, or locations with heavy machinery obstructions. By leveraging multiple data sources, SSD-based models can achieve higher robustness and reliability in real-world industrial settings.

Furthermore, continuous learning and automated model updates can be implemented to ensure that the detection system adapts to evolving workplace conditions and new PPE designs. By incorporating active learning techniques, the model can continuously improve by learning from newly labeled data, reducing the need for manual retraining.

Despite its advantages, SSD-based PPE detection systems may still face challenges, such as handling occlusions, detecting small objects, and dealing with variations in PPE appearances. However, these limitations can be mitigated through advanced data augmentation techniques, higher-resolution input images, and hybrid approaches that integrate SSD with other deep learning models like YOLOv8 or Faster R-CNN for enhanced performance.

Overall, SSD-based PPE detection contributes significantly to workplace safety by providing a reliable, real-time monitoring system that ensures compliance with safety protocols. With further enhancements in edge AI, federated learning, explainable AI, and multi-modal fusion, this technology can become even more effective in reducing workplace hazards and preventing accidents. By embracing these advancements, industries can create a safer work environment, minimize human errors, and enhance overall occupational safety standards.

6. REFERENCES

[1] Terven, Juan, and Diana Cordova-Esparza. "A comprehensive review of YOLO: From YOLOv1 to YOLOv8 and beyond." arXiv preprint arXiv:2304.00501 (2023).

[2] Kim, Jun-Hwa, Namho Kim, and Chee Sun Won. "High-Speed Drone Detection Based On Yolo-V8." ICASSP 2023-2023 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP). IEEE, 2023.

[3] Diwan, Tausif, G. Anirudh, and Jitendra V. Tembhurne.

"Object detection using YOLO: Challenges, architectural successors, datasets and applications." Multimedia Tools and Applications 82.6 (2023): 9243-9275

[4] Terven J, Cordova-Esparza D. A comprehensive review of YOLO: From YOLOv1 to YOLOv8 and beyond. arXiv preprint arXiv:2304.00501. 2023 Apr 2

[5] Ammad, S., Alaloul, W. S., Saad, S., & Qureshi, A. H. (2021). Personal protective equipment (PPE) usage in construction projects: A scientometric approach. Journal of Building Engineering, 35, 102086.

[6] Li, Yang, Yujie Lu, and Jun Chen. "A deep learning approach for real time rebar counting on the construction site based on YOLOv3 detector." Automation in Construction 124 (2021): 103602

[7] Nagrath, Preeti, et al. "SSDMNV2: A real time DNN-based face mask detection system using single shot multibox detector and MobileNetV2." Sustainable cities and society 66 (2021): 102692.

[8] Delhi, Venkata Santosh Kumar, R. Sankarlal, and Albert Thomas. "Detection of personal protective equipment (PPE) compliance on construction site using computer vision based deep learning techniques." Frontiers in Built Environment 6 (2020): 136

[9] Wang, Peijin, et al. "FMSSD: Feature-merged single-shot detection for multiscale objects in large-scale remote sensing imagery." IEEE Transactions on Geoscience and Remote Sensing 58.5 (2019): 3377 3390.

[10] Wong, Alexander, et al. "YOLO nano: A highly compact you only look once convolutional neural network for object detection." 2019 Fifth Workshop on Energy Efficient Machine Learning and Cognitive Computing-NeurIPS Edition (EMC2-NIPS). IEEE, 2019.

[11] Juba, Brendan, and Hai S. Le. "Precision-recall versus accuracy and the role of large data sets." Proceedings of the AAAI conference on artificial intelligence. Vol. 33. No. 01. 2019.

[12] Wu F, Jin G, Gao M, Zhiwei HE, Yang Y. Helmet detection based on improved YOLO V3 deep model. In2019 IEEE 16th International conference on networking, sensing and control (ICNSC) 2019 May 9 (pp. 363-368). IEEE.