# Personal Protective Equipment Detection System for Workers

Dr. Yogendra V Pati[1], Aditya Aher[2], Dhanashri Dagade[3], Madhumita Mahata[4], Meghashree Pawar[5]
[1,2,3,4,5] Assoc.Prof, Computer Engineering, Marathwada Mitramandal's Institute of Technology

***Abstract*— The safety and well-being of workers across various industries is of paramount importance, with Personal Protective Equipment (PPE) serving as a critical component in safeguarding their health. This abstract introduces a novel Personal Protective Equipment Detection System (PPEDS), designed to enhance workplace safety through advanced computer vision technology. The PPEDS is engineered to identify and monitor the correct usage of PPE among workers in real-time, thereby mitigating workplace accidents and ensuring adherence to safety regulations. Utilizing state-of-the-art image recognition algorithms and machine learning models, the PPEDS analyzes video feeds from strategically placed surveillance cameras to detect and track various types of PPE, such as helmets, safety goggles, face masks, gloves, reflective vests, and ear protection devices. The system is capable of determining whether workers are equipped with the appropriate PPE for their specific tasks and can generate alerts and notifications when non-compliance is detected. Key features of the PPEDS include real-time detection capabilities, customizable rule settings and alerts, comprehensive data logging and reporting, integration with existing safety systems, educational support for workers, and stringent privacy safeguards. By implementing the PPEDS, organizations can significantly reduce the incidence of workplace accidents, improve compliance with safety standards, and ultimately protect workers' lives in hazardous environments.**

Keyword: *Image Recognition, PPE detection, industrial Safety, Real time detection.*

## I. INTRODUCTION

The Personal Protective Equipment (PPE) Detection System for Workers is a cutting-edge technological solution designed to significantly improve workplace safety and address occupational risks. In sectors like construction and heavy industry, where employees are frequently exposed to potential hazards, ensuring proper PPE usage is essential. This system integrates advanced sensor technology, computer vision, and artificial intelligence to monitor and verify the usage of critical PPE items, including helmets, safety goggles, masks, gloves, and protective clothing. By utilizing real-time data analysis, the system can detect non-compliance with PPE requirements and promptly alert both workers and supervisors, thereby helping to prevent accidents and injuries. The system's machine learning algorithms are adept at recognizing various types of PPE and assessing their condition, ensuring that employees not only wear the required equipment but also that it is in effective working order. Moreover, the system provides detailed compliance reports and data insights, enabling management to enhance safety protocols and reduce legal and financial risks. By offering a proactive and automated approach to PPE monitoring, the PPE Detection System represents a significant advancement in workplace safety technology, aiming to protect workers and improve safety standards across diverse industrial settings.

## II. RELATED WORK

Recent advancements in video surveillance technology and the proliferation of extensive image datasets from industrial settings have accelerated the development of Computer Vision (CV) algorithms for critical area monitoring. These algorithms leverage visual data to perform essential tasks such as worker tracking, defect detection, and risk identification within industrial environments. Convolutional Neural Networks (CNNs), particularly in their various forms such as R-CNN, Fast R-CNN, Faster R-CNN, SSD, and YOLO, have become pivotal in object detection due to their exceptional performance. These deep learning models are widely used to enforce workplace safety by monitoring PPE usage, predicting potential collisions, and detecting safety gear like helmets and vests. Transfer learning is frequently applied to adapt pre-trained object detectors for PPE recognition, with recent studies evaluating different YOLO versions and integrating machine learning classifiers and decision trees to enhance accuracy. The incorporation of pose estimation techniques further refines PPE detection by focusing on specific body regions.In addition to traditional cloud-based approaches, there is growing interest in IoT-based solutions and edge computing for real-time PPE monitoring. IoT systems equipped with sensors and data analysis capabilities provide immediate feedback on PPE usage, while edge AI systems offer the advantage of low latency and high privacy by processing data close to the source.

This paper reviews recent research and identifies both the challenges and opportunities in developing effective PPE detection systems. It highlights the need for robust, real-time analysis and explores the potential of edge computing to address computational and latency issues, proposing a framework for deploying PPE detection systems in industrial environments using embedded systems and advanced deep learning techniques.

## III. PROPOSED SYSTEM

### System Overview and Implementation

The PPE Detection System leverages the state-of-the-art YOLOv5s architecture to effectively train a neural network for real-time recognition of personal protective equipment (PPE). To achieve optimal performance, a custom dataset is meticulously created using advanced image segmentation and data augmentation techniques, which triples the original number of images. This robust dataset enables the trained model to accurately analyze webcam frame captures and detect PPE—such as helmets, vests, and gloves—in real time.

**Dataset Creation**

Creating a large, labeled dataset is a time-consuming process involving the careful selection and labeling of images. This task often requires manual annotation, which can introduce errors. For our PPE detection system, the dataset is designed to identify three specific types of PPE: helmets, vests, and gloves. Consequently, the dataset includes six distinct classes: "hand without glove," "hand with glove," "chest without vest," "chest with vest," and "head without helmet." For simplicity, these classes are referred to as head, helmet, chest, vest, hand, and glove.

**YOLO Network Selection**

After a thorough evaluation of various neural network architectures, YOLO v8 was chosen for its superior training efficiency, compatibility with video and webcam sources, and lighter weight compared to previous YOLO versions. The selection of the YOLOv8 variant was influenced by two main factors: its accuracy as measured against the Common Objects in Context (COCO) dataset and its processing speed. Although YOLOv5n is often preferred for its balanced processing time and accuracy, YOLOv8s was selected for this project to prioritize accuracy over processing speed.

Python was chosen for implementing and training the neural network due to its extensive library support and active community. To manage the intensive GPU resources required for training—an often costly endeavor—Google Colab was utilized as the cloud service platform.

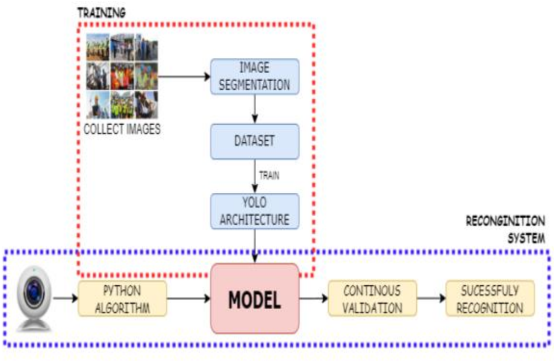



*Fig.1. System Architecture*

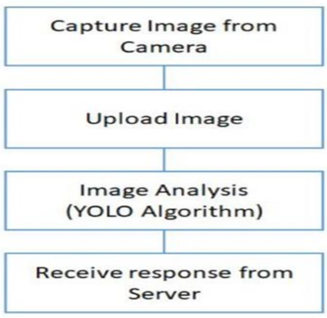



*Fig.2. Yolo Procedure*

## IV. PPE DETECTION TASK TRRAINING AND METHODOLOGY

### A. *DNN's for Detection Task*

When creating a predictive model like Deep Neural Networks (DNNs) for the task of object detection in images, it's essential to have a well-curated set of labeled training samples. For our project, we need RGB images paired with detailed information about the objects in each image, including their location and classification. The ideal training dataset for this purpose should contain images from industrial environments, featuring a wide variety of PPE items that we aim to detect, such as hard hats, safety vests, and protective gloves. This dataset should also include images captured under diverse ambient conditions, with varying backgrounds, distances, and camera angles. This diversity is crucial to ensure that the DNNs can accurately recognize PPE in different scenarios and environments.

For DNNs, particularly in image object detection, the quality of the trained model significantly improves with a larger number of labeled training samples. This is due to the fact that DNNs have a vast number of parameters that require optimization. As a result, having an extensive set of labeled images is vital for training a DNN from scratch. However, collecting and labeling a large dataset can be an arduous and time-consuming task. To address this challenge, researchers and practitioners often use transfer learning, where pre-trained models built for general object detection tasks are fine-tuned for more specific tasks, such as recognizing particular PPE items in our case.

During the pre-training stage of DNNs for image object detection, large datasets with millions of labeled images across thousands of categories are often utilized. There are several highly accurate pre-trained DNNs available in AI and ML software libraries and online repositories. In our study, we evaluated the effectiveness of five well-known DNNs for object detection, including YOLOv8 and its lightweight version YOLOv8-Tiny, SSD MobileNet V2, CenterNet V2, and EfficientDet D0. For each of these networks, we first downloaded the pre-trained models, which were trained using the COCO dataset, from various public code repositories. The pre-trained YOLOv8 and YOLOv8-Tiny models, for instance, can be found in Alexey's GitHub repository, while the others are available in the TensorFlow official repository. We then fine-tuned each pre-trained network to customize it for our specific PPE detection task. This was done using three different datasets, which are detailed below, along with the fine-tuning process and the performance comparison of the five deep learning networks.

*1) Detection Scenario*

Personal protective equipment is critical in safeguarding workers and preventing serious injuries. Various parts of the body require specific protective measures. For instance, hearing protection and safety goggles are vital for shielding the ears and eyes from loud machinery and flying debris. Safety vests enhance the visibility of the chest area, while harnesses provide stability. Gloves and safety shoes are essential for protecting the limbs from burns, scratches, and other injuries.

In this study, we selected one specific type of PPE for each body area: safety helmets for head protection, safety vests for the upper body, and gloves for the arms and hands, as these are commonly used in industrial settings.

The work environment is divided into low-risk and high-risk areas. In high-risk zones, workers are required to wear PPE, including helmets, vests, and gloves, to ensure their safety. The objective of our proposed system is to analyze real-time images captured by surveillance cameras to identify workers who are not wearing the required PPE. If a worker enters a high-risk area without the necessary protection, the system will generate visual or audible alerts to notify the worker. Specifically, an alert will be triggered for each type of PPE that is not worn. Additionally, the system can be integrated with a control mechanism that can shut down potentially hazardous machinery in the high-risk area, thus preventing accidents and enhancing overall safety.


*Fig.3. Detection webapplication*

## V. Evaluation And Result

### *A. Performance on Image Datasets*

The image processing speed ranged between 1 to 2 images per second, regardless of the number of classes or bounding boxes detected within each image. To assess the performance of the trained model, a confusion matrix was employed, which cross-referenced human interpretations with the model's predictions. The key metrics used to evaluate performance included accuracy and the F1 score.

To calculate accuracy, we recorded the True Positives (TP), True Negatives (TN), False Positives (FP), and False Negatives (FN) for each class. Accuracy was then determined by dividing the sum of all True Positives by the total number of data points. Precision, which measures the likelihood that a positive prediction was correct, was calculated by dividing TP by the sum of TP and FP. Recall, which indicates the model's effectiveness in identifying all instances of a particular class, was calculated by dividing TP by the sum of TP and FN.

The F1 score, which provides a balanced measure of precision and recall, was computed as the harmonic mean of these two metrics. Specifically, it was calculated as twice the product of precision and recall, divided by the sum of precision and recall. Together, these metrics offered a comprehensive evaluation of the model's ability to accurately identify the various classes present in the dataset.

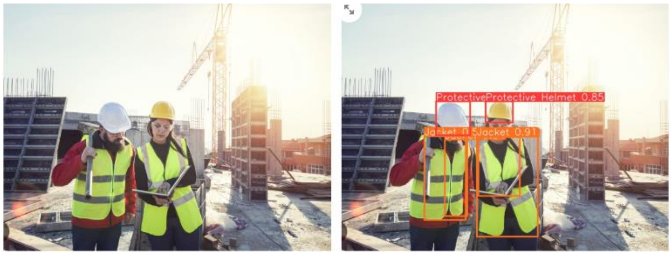


*Fig.4. Performance of image dataset*

*B. Training Dataset*

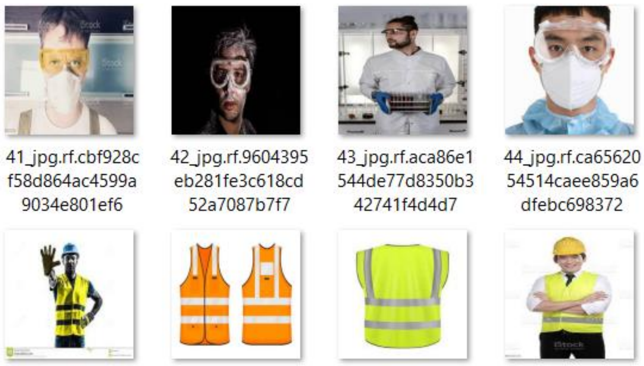



*Fig.4. Train Dataset*

The junctions that are not included in the training dataset are used only for validation of the network. Moreover, the images are resized to 416×416 pixels. The network was trained on a workstation equipped with an Nvidia GPU with mini-batch size of 10, maximum number epochs of 60, a selected initial learning rate of 10−4 and solved by the stochastic gradient descent [32] with momentum optimizer. The trained network provides an accuracy of 100% and 89.2% on training and validation datasets, respectively.

*C. Test Dataset*


*Fig. Test Dataset*

### 1) *Training and Validation Process*

The network was validated using junctions not included in the training dataset. All images were resized to 416×416 pixels before training. The training was carried out on a workstation equipped with an Nvidia GPU, using a mini-batch size of 10 and a maximum of 60 epochs. The initial learning rate was set at 10−410^{-4}10−4, and the training process was optimized using stochastic gradient descent with a momentum optimizer. This rigorous training process yielded an accuracy of 100% on the training dataset and 89.2% on the validation dataset.

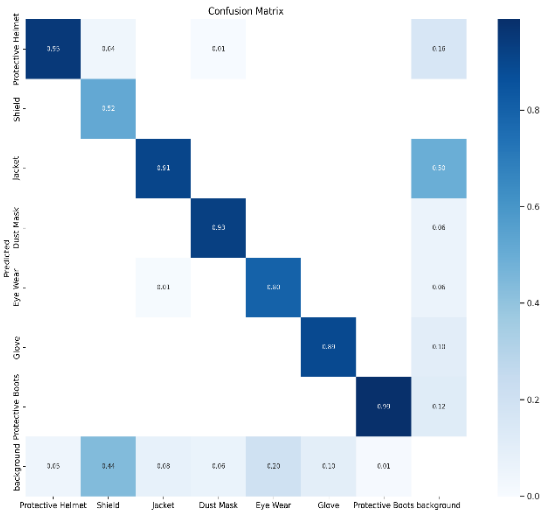
### 2) *Test Dataset and Its Role*

The test dataset serves as an unbiased benchmark for evaluating the final model after it has been trained on the training and validation datasets. The test dataset is used only once the model training is complete. It provides a "gold standard" to gauge the model's performance in real-world scenarios. Typically, the validation set is available alongside the training set during the development phase, while the test set is only used at the very end, often in competition settings to determine the final rankings. It's important to note that using the validation set as a test set is not a best practice, as the test set is generally curated to cover a wide range of data that the model might encounter in real-world applications.

### 3) *Validation Dataset Evaluation*

The model demonstrated strong predictive capabilities on the validation dataset, achieving an impressive accuracy of 96.27%. This high accuracy was supported by the alignment of the model's predictions (indicated by green cells) with the ground truth, affirming the model's effectiveness. Beyond accuracy, the model's performance was further evaluated through a detailed classification report, which included metrics such as True Positives (TP), True Negatives (TN), False Positives (FP), and False Negatives (FN) for each class. These metrics provided insights into the model's precision, recall, and F1 score across different classes, highlighting its strengths and areas for improvement. Overall, the results confirm the model's robustness and reliability in accurately predicting classifications within the validation dataset, demonstrating its potential for real-world application with high confidence.

Confusion Matrix:



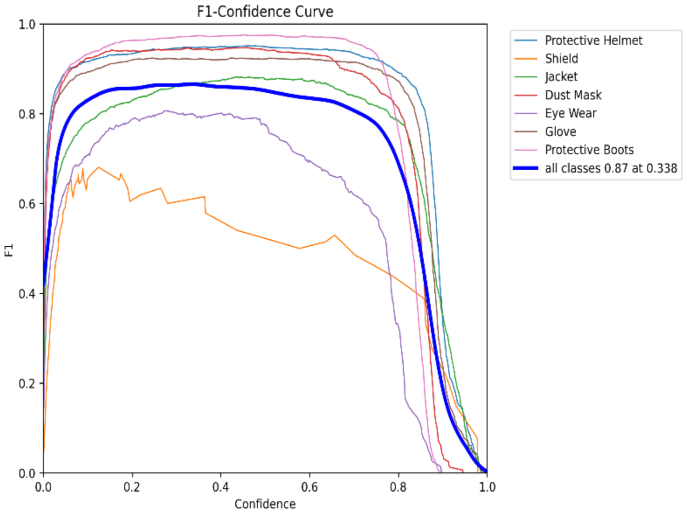



*Fig.4. F1 confidence Curve*

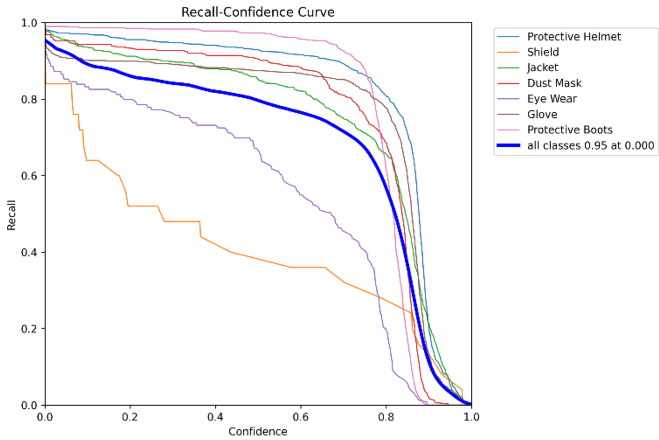



*Fig.5. Recall confidence Curve*

### *D. Performance of the Video File*

To assess the trained model's robustness, it was employed to predict a video file generated from CCTV footage in mp4 format, processed at a rate of 2 frames per second due to limited computational resources. Despite the reduced speed, the model exhibited proficient performance in detecting individuals across frames. Notably, the algorithm accurately classified individuals as unsafe when there was a modification in their personal protective equipment (PPE), such as the removal of a hard hat or jacket. Conversely, individuals initially labeled as unsafe were promptly reclassified as SAFE upon restoring PPE compliance. This dynamic showcases the model's adaptability and reliability in real-world scenarios, where changes in safety equipment usage can impact risk assessment.


*Fig.6. Video detection output*

## VI. CONCLUSION

The development and implementation of the Personal Protective Equipment Detection System represent a noteworthy advancement in the realm of workplace safety technology. By leveraging computer vision and machine learning, the system provides a robust solution for real-time monitoring of PPE compliance. This not only ensures that workers are properly equipped but also helps organizations maintain compliance with safety regulations. The system's ability to integrate with existing safety management frameworks enhances its utility, making it an invaluable addition to any organization's safety infrastructure. The system's effectiveness in reducing workplace accidents and improving safety compliance highlights the importance of continued innovation in this field. Future enhancements could focus on expanding the system's capabilities to detect additional PPE types and improving performance under challenging conditions, such as poor lighting or extreme weather. Additionally, integrating the system with other emerging safety technologies, like IoT devices and wearable sensors, could provide a more comprehensive safety solution for workers in high-risk industries.

## 6. Future Work

The future development of the PPE Detection System offers substantial potential for enhancing its features and expanding its applicability. One area of focus could be on developing more sophisticated algorithms capable of detecting a wider array of PPE items, including those used in specialized industries like mining, chemical processing, and healthcare. Furthermore, efforts could be directed towards improving the system's ability to function effectively in challenging environments, such as low-light conditions, extreme temperatures, or areas with high levels of dust and debris. Another promising avenue for development is the integration of the PPE Detection System with other safety technologies, such as IoT devices and wearable sensors.

### ACKNOWLEDGMENT

*We invested a lot of time and energy into this project. However, a lot of people contributed to and encouraged us as we finished our tasks. We want to express our sincere gratitude to each one of them. We are grateful to Dr. Yogendra V Patil for acting as our mentor and for always being available to supervise us and provide the project with the information and insights we needed. We are appreciative of his guidance. We would like to express our gratitude to our parents and friends at Marathwada Mitra mandal's Institute of Technology for their support and cooperation in helping us finish this job.*

### REFERENCES

[1] C.-C. Hung, G. Ananthanarayanan, P. Bodik, L. Golubchik, M. Yu,P. Bahl, and M.Philipose,''VideoEdge:Processingcamerastreamsusinghierarchicalclusters,''inProc.IEEE/ACMSymp.EdgeComputing.(SEC),Oct.2018,pp.115–131.

[2] L.Ciampi,C.Gennaro,F.Carrara,F.Falchi,C.Vairo,andG.Amato,''Multi-cameravehiclecounting using edge-AI,'' Expert Syst.Appl.,vol. 207,Nov.2022,Art.no.117929B. Smith, "An approach to graphs of linear forms (Unpublished work style)," unpublished.

[3] R.Rajavel,S.K.Ravichandran,K.Harimoorthy,P.Nagappan,andK.R.Gobichettipalayam,''IoT-basedsmarthealthcarevideosurveillancesystemusingedgecomputing,'' J. Ambient Intel. Hum. Compute., vol. 13,no. 6, pp. 3195–3207, Jun.2022.

[4] S.Liu,L.Liu,J.Tang,B.Yu,andW.Shi,''Edgecomputingforautonomousdriving:Opportunitiesandchallenges,''Proc.IEEE,vol.107,no.8,pp.1697–1716,Jun.2019.

[5] A. Ndikumana, N. H. Tran, D. H. Kim, K. T. Kim, and C. S. Hong, ''Deep learningbased caching for self-driving cars in multi-access edge computing,'' IEEE Trans. Intel.Transp.Syst.,vol.22,no.5,pp.2862–2877,May2020.

[6] The Potential of Artificial Intelligence for Tourism: A Review of the Literature by Buhalis and Foerste (2020)

[7] The Impact of Artificial Intelligence on Tourist Behavior: A Case Study of Online Travel Agencies by Xiang et al.