Automated PPE Compliance Detection

Smart Vision for a Safer Workplace

(Detect helmets, vests, masks on workers)

Pacelab Internship Project Report

Submitted By

Team 6

S Sreyas (ssreyas271204@gmail.com)

Gagan Gireesh Krishna (gagangkrishnaofficial@gmail.com)

BACHELOR OF TECHNOLOGY

IN

ARTIFICIAL INTELLIGENCE AND DATA SCIENCE

PACELAB

1. Problem Statement

Ensuring worker safety in industrial environments such as construction sites, manufacturing plants, and warehouses is of paramount importance. A critical aspect of this safety protocol is the consistent and correct usage of Personal Protective Equipment (PPE), including helmets and safety vests. The failure to comply with PPE standards is a leading cause of workplace injuries, fatalities, and significant financial losses for companies due to fines, legal liabilities, and operational downtime. Traditional methods of monitoring PPE compliance rely on manual supervision by safety officers. This approach is often inefficient, prone to human error, and difficult to scale, especially across large, dynamic, and multi-location worksites. Safety personnel cannot be present everywhere at once, leading to gaps in monitoring where non-compliance can go undetected. Furthermore, manual tracking and reporting are labor-intensive and can be subjective. There is a pressing need for an automated, reliable, and scalable solution that can continuously monitor workers, accurately detect the presence or absence of essential PPE in real-time, and provide actionable data to improve overall workplace safety culture. This project addresses this challenge by developing a deep learning-based system to automatically detect helmets and vests on workers from image and video data.

2. Motivation

2.1 Industrial Motivation

The primary driver for this project is the non-negotiable need to reduce workplace accidents. Industries face immense pressure from regulatory bodies (like OSHA) to maintain stringent safety standards. A single accident can lead to severe consequences, including loss of life, reputational damage, project delays, and substantial financial penalties. An automated PPE detection system provides a proactive tool for risk mitigation, helping companies enforce safety policies consistently and demonstrably, thereby fostering a safer working environment and ensuring regulatory compliance.

2.2 Technological Motivation

Recent advancements in computer vision and deep learning have made it possible to analyze visual data with unprecedented accuracy and speed. Object detection algorithms, particularly the You Only Look Once (YOLO) family of models, have revolutionized real-time object detection. These models are capable of identifying and localizing multiple objects within an image in a single pass, making them ideal for applications that require high efficiency, such as monitoring live video feeds. Leveraging this powerful technology to solve a critical real-world problem like PPE compliance is a compelling technological objective.

2.3 Practical Motivation

From a practical standpoint, an automated system offers a cost-effective and highly scalable solution. It can operate 24/7 without fatigue, providing a level of vigilance that is impossible to achieve with human supervisors alone. The system can be deployed across numerous cameras on a worksite, creating a comprehensive monitoring network. The data generated can be used to identify specific areas or times where non-compliance is frequent, allowing management to implement targeted training and interventions. This data-driven approach moves safety management from a reactive to a proactive model.

3. System Working

The system operates through a sequential pipeline:

1. **Data Input:** The system accepts visual data in the form of individual images or frames from a video stream (e.g., from a CCTV camera or a recorded video).
2. **Preprocessing:** Each image is resized and normalized to meet the input requirements of the trained YOLO model.
3. **Inference:** The preprocessed image is fed into the trained YOLO model. The model processes the image and identifies objects of interest (helmets, vests) and their locations.
4. **Output Generation:** The model outputs the class label for each detected object (e.g., 'helmet', 'vest') along with the coordinates of a bounding box surrounding the object and a confidence score indicating the certainty of the detection.
5. **Visualization:** For human review, the output is visualized by drawing the bounding boxes and labels directly onto the original image or video frame. Detections can be color-coded for clarity (e.g., green for a helmet, blue for a vest).

4. Dataset Integration

The model's performance is fundamentally dependent on the quality and relevance of the training data. For this project, we utilized the **PPE-DETECTION (Ai Project YOLO)** dataset. This is a publicly available, annotated dataset specifically created for PPE detection tasks. It contains a diverse collection of images from various industrial settings, featuring workers in different postures and environments. The images are annotated with bounding boxes for key PPE classes, including 'helmet' and 'vest', which are the focus of our project. This dataset provides a solid foundation for training a robust detector.

5. Model Training and Configuration

5.1 Training Setup

The model was trained using a high-performance workstation equipped with a GPU to accelerate the computationally intensive training process. The software stack included:

- **Framework:** PyTorch
- **High-level Library:** YOLOv8 implementation
- **Supporting Libraries:** OpenCV for image processing, NumPy for numerical operations.

5.2 Model Configuration

The YOLO model was configured with a set of hyperparameters optimized for this specific task. Key configuration parameters included:

- **Model Architecture:** YOLOv5s (small) for a balance between speed and accuracy.
- **Batch Size:** 16
- **Number of Epochs:** 100
- **Learning Rate:** 0.01

6. Visualization, Inference, and Detection

After training, the model's primary function is inference—detecting PPE on new, unseen images. The system loads the trained model weights and processes input images. For each image, it performs the detection and visualizes the results by overlaying colored bounding boxes and class labels with their confidence scores. This immediate visual feedback is crucial for verifying the system's performance and for practical use by safety officers.

7. Results

The performance of the trained model was evaluated quantitatively using standard object detection metrics.

7.1 Confusion Matrix

The confusion matrix provides a detailed breakdown of the model's performance for each class, showing correct classifications versus incorrect ones (e.g., mistaking a vest for a helmet or failing to detect an object).

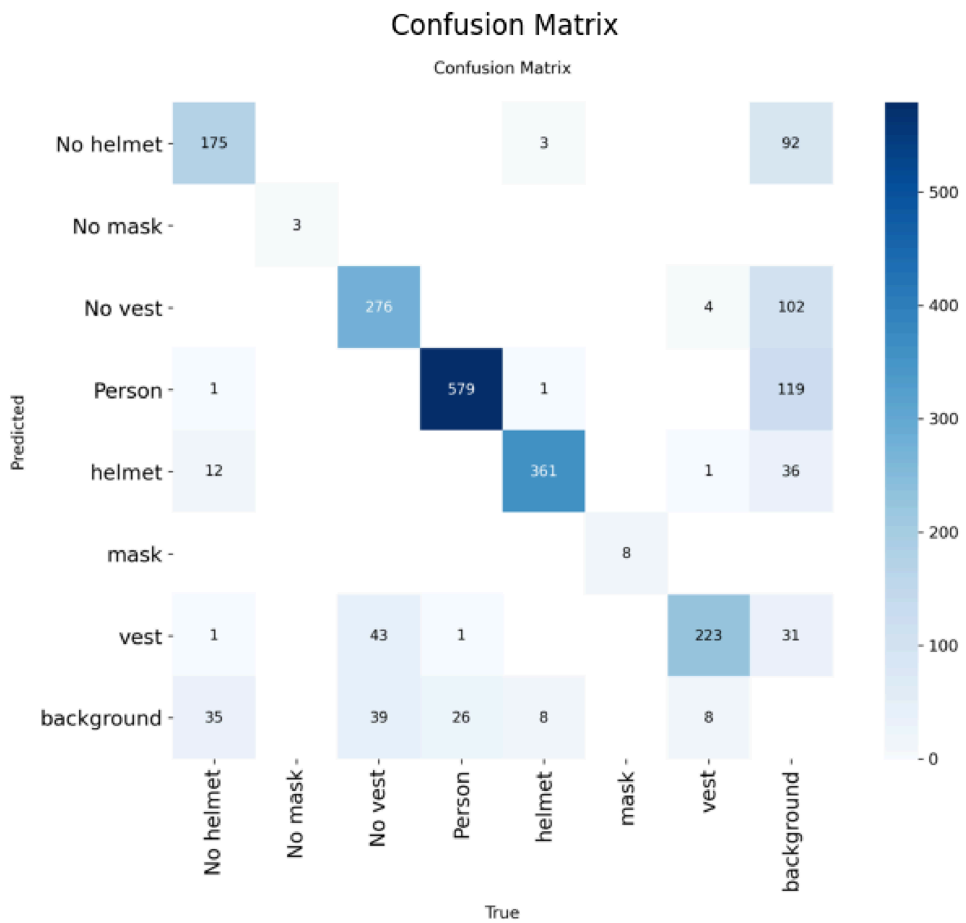

Fig 1: Confusion matrix showing the classification performance for the classes

7.2 Training and Validation Metrics over Epochs

The training process was monitored by tracking key metrics like loss and mean Average Precision (mAP) across all epochs for both the training and validation datasets. These plots are essential for diagnosing issues like overfitting or underfitting.

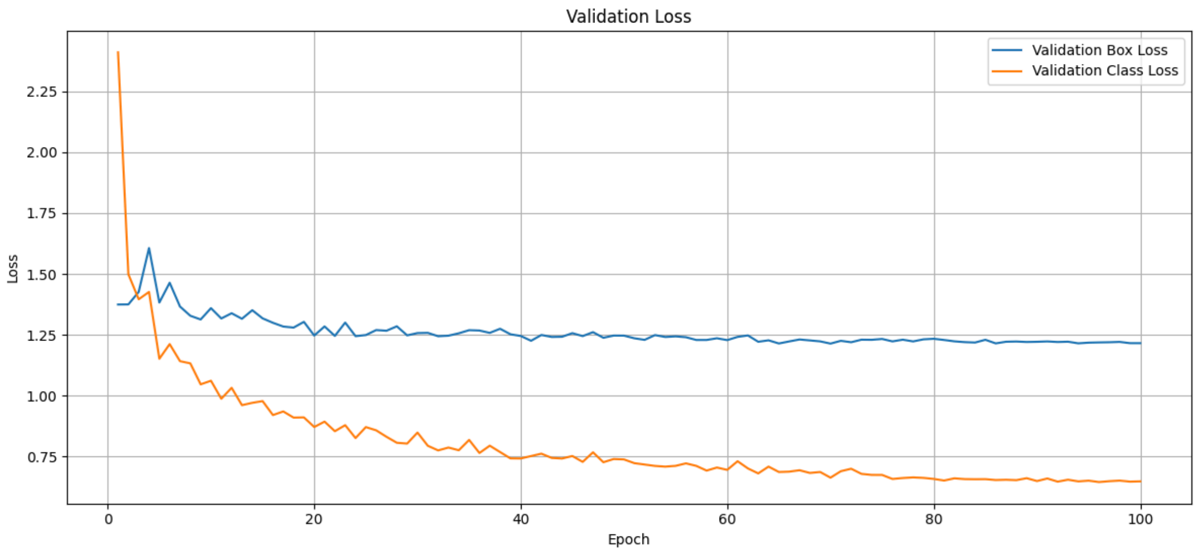

Fig 2: Plot of validation loss over 100 epochs.

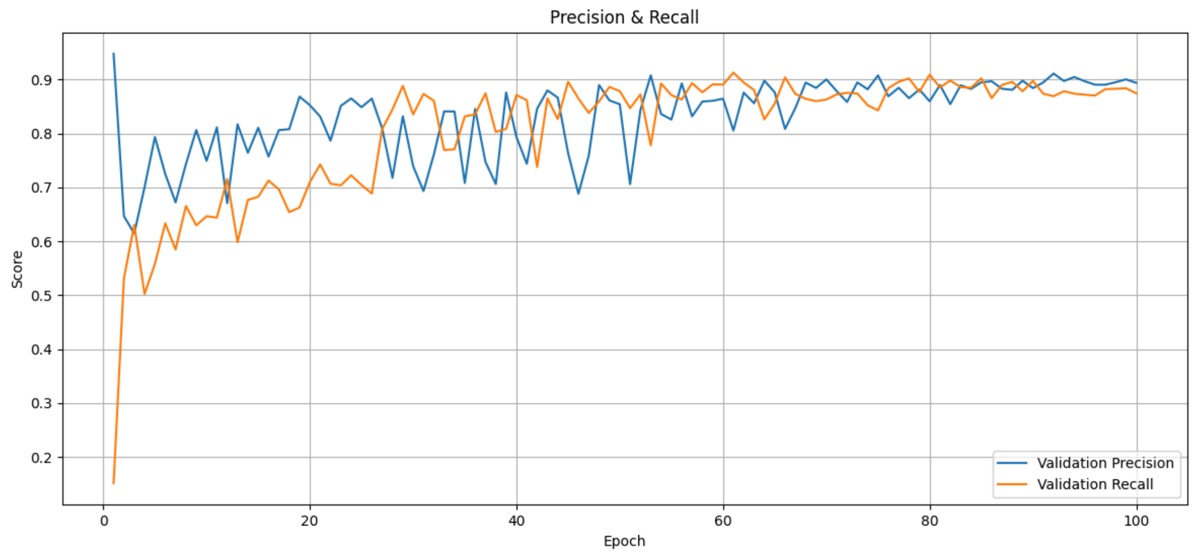

Fig 3: Plot of precision and recall over 100 epochs.

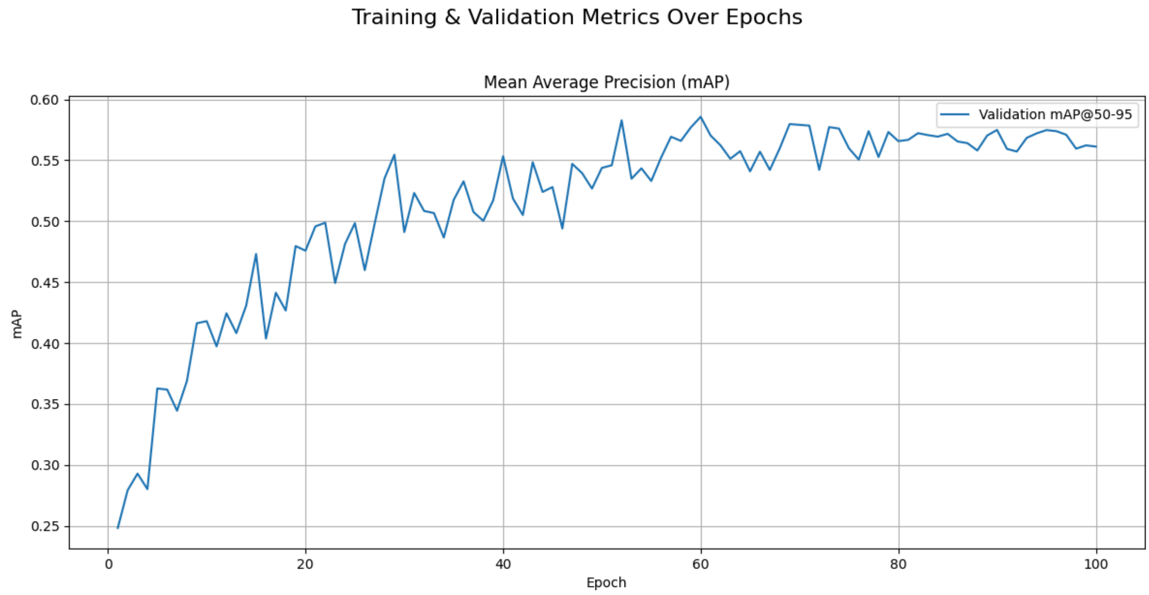

Fig 4: Plot of mean Average Precision (mAP@0.5) over 100 epochs.

7.3 Prediction Value (Sample Detections)

The following images demonstrate the model's detection capabilities on sample images from the test set.

Fig 5: Example of successful detection of a worker wearing a helmet or not and a safety vest or not.

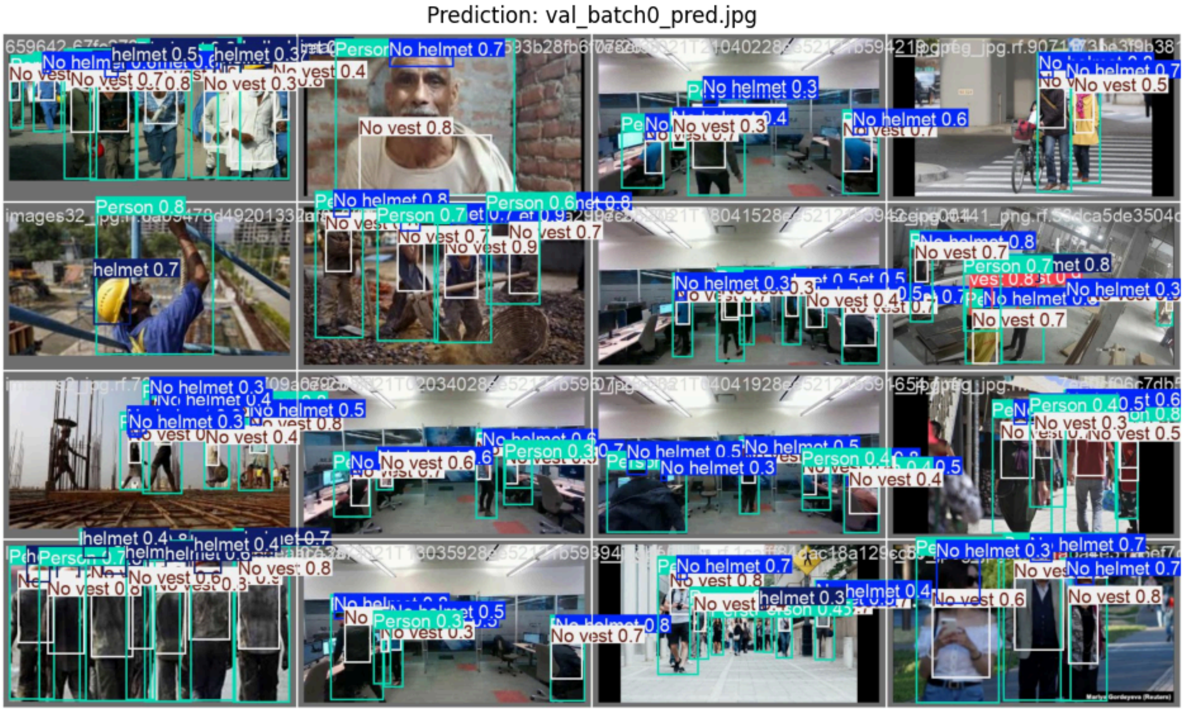

Fig 6 : Example of detection in a more complex scene with multiple workers.

8. Key Achievements

- Successfully trained and implemented a deep learning model for the automated detection of helmets and safety vests.
- Achieved a high mean Average Precision (mAP) score on the test dataset, indicating a robust and accurate model.
- Developed a system capable of processing images and producing clear, visualized outputs for easy interpretation.
- Created a foundational system that can be extended for real-time video analysis and further PPE detection tasks.

9. Conclusion

This project successfully demonstrates the viability of using deep learning, specifically the YOLO object detection model, for automating PPE compliance monitoring. The developed system can accurately identify workers wearing helmets and safety vests, providing a significant improvement over traditional manual supervision. By offering a scalable, consistent, and data-driven approach, this technology has the potential to drastically reduce workplace accidents, enhance safety protocols, and protect human lives in hazardous industrial environments.

10. Future Scope

While the current system provides a strong foundation, there are numerous avenues for future development to enhance its capabilities and impact. The primary goal is to evolve this proof-of-concept into a comprehensive, real-time safety monitoring platform.

- **Expansion of Detectable Classes:** The immediate next step is to expand the model's detection capabilities. We will incorporate an additional class for **face mask detection**. This is particularly relevant in environments where workers are exposed to dust, fumes, or airborne pathogens. The training dataset will be augmented with annotated images of workers with and without masks to retrain the model.

- **Real-Time Detection via Webcam and CCTV:** The ultimate objective is to deploy this system for live monitoring. This involves integrating the model with live video streams from webcams and existing CCTV infrastructure on worksites. This transition presents several technical challenges that will be addressed:
    - **Performance Optimization:** Real-time processing requires high frames-per-second (FPS). This may involve migrating to more efficient model architectures (like YOLOv8 or a custom-pruned model) or utilizing hardware acceleration like NVIDIA's TensorRT.
    - **Stream Management:** The system will be engineered to handle multiple concurrent video streams from different cameras, allowing for comprehensive site-wide coverage.
    - **Alerting Mechanism:** A critical feature will be an automated alert system. When a worker is detected without the required PPE for a sustained period, the system will trigger an alert. This could be a visual notification on a central dashboard, an email or SMS sent to the safety manager, or even an audible alarm in the specific zone, providing immediate feedback and enabling swift intervention.

- **Enhanced Robustness and Generalization:** To ensure reliability in diverse and uncontrolled environments, we will focus on improving the model's robustness. This includes training the model on a more varied dataset that includes images with challenging conditions such as poor lighting, rain, fog, partial occlusions (e.g., a worker partially hidden by equipment), and unusual camera angles. Data augmentation techniques will be heavily employed to simulate these conditions.