ISSN: 2582-7219

# International Journal of Multidisciplinary Research in Science, Engineering and Technology

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

**Impact Factor: 8.206**

**Volume 8, Special Issue 2, November 2025**

International Journal of Multidisciplinary Research in Science, Engineering and Technology (IJMRSET)

| ISSN: 2582-7219 | www.ijmrset.com | Impact Factor: 8.206| | Monthly Peer Reviewed & Refereed Journal |

|| Volume 8, Special Issue 2, November 2025 ||

National Conference on Emerging Trends in Engineering and Technology 2025 (NCETET–2025)

Organized by

Mookambigai College of Engineering, Keeranur, Tamil Nadu, India

# A Lightweight Real-Time PPE Detection in Industrial Environments using RT-DETR

**Jayasri M[1], Madhumitha G[2]**

Department of Artificial Intelligence and Data Science, Mookambigai College of Engineering, Pudukkottai, Tamil Nadu, India[1-2]

**ABSTRACT**: Industrial safety relies heavily on ensuring that workers correctly wear Personal Protective Equipment (PPE). Manual PPE monitoring is prone to human error, inefficiency, and delay. Deep learning models such as YOLO and Faster R-CNN offer high accuracy but are computationally intensive, limiting their real-time usability on edge devices. This paper proposes a lightweight Real-Time Detection Transformer integrated with MobileNetV3 backbone (RT-DETR-MV3) to achieve faster and efficient PPE detection. The MobileNetV3 backbone accelerates feature extraction using depthwise separable convolutions and attention mechanisms, while the RT-DETR head leverages transformer-based multi-head attention for context-aware detection. Experimental results show that RT-DETR-MV3 achieves comparable accuracy to the original RT-DETR model with a 60% reduction in FLOPs and a 3× speed increase, enabling real-time PPE detection for industrial surveillance systems.

**KEYWORDS:** RT-DETR, MobileNetV3, PPE Detection, Industrial Safety, Object detection, Transformer Networks

## I. INTRODUCTION

Industrial environments involve numerous safety risks that necessitate strict PPE compliance to protect workers from injuries. Current manual monitoring systems are inefficient and fail to provide real-time insights. To address this, deep learning-based object detection models have emerged as effective alternatives.
However, many of these models (e.g., YOLOv8, Faster R-CNN) require high computational resources, making them unsuitable for deployment on low-power industrial edge devices. The RTDETR (Real-Time Detection Transformer) provides an end-to-end detection pipeline with superior contextual reasoning, but its heavy ResNet backbone limits its efficiency.

To overcome these limitations, this work integrates MobileNetV3 as the backbone of RT-DETR, forming RT-DETR-MV3, a lightweight architecture capable of real-time PPE detection. The model identifies essential safety gear such as helmets, gloves, and vests under varying lighting and occlusion conditions.

### 1.1 Problem Motivation

Conventional deep learning-based PPE monitoring systems struggle with speed–accuracy tradeoffs. High-speed detectors compromise precision, while accurate models demand high computational power. A lightweight and efficient detection model is thus required to enable realtime PPE monitoring in industrial environments.

### 1.2 Contributions

- Integration of **MobileNetV3** backbone with RT-DETR to reduce computational cost.
- Improved feature extraction using **depthwise separable convolutions** and **SE attention**.
- Retention of transformer-based **contextual detection** for better accuracy.
- Comprehensive evaluation on industrial PPE datasets for real-world applicability.

| ISSN: 2582-7219 | www.ijmrset.com | Impact Factor: 8.206| | Monthly Peer Reviewed & Refereed Journal |

|| Volume 8, Special Issue 2, November 2025 ||

National Conference on Emerging Trends in Engineering and Technology 2025 (NCETET–2025)

Organized by

Mookambigai College of Engineering, Keeranur, Tamil Nadu, India

## II. RELATED WORK

Traditional detectors such as Faster R-CNN and SSD deliver high precision but require heavy computation. YOLOv8 achieves real-time detection but sacrifices some accuracy when deployed on low-power devices. RT-DETR employs transformer-based architecture for object detection, eliminating the need for manual anchor tuning and providing end-to-end learning. However, its backbone (ResNet) is too large for real-time edge deployment. Lightweight networks such as MobileNetV3, designed using neural architecture search and attention mechanisms, reduce computation drastically. Combining MobileNetV3 with RT-DETR ensures both speed and accuracy, making it ideal for industrial PPE detection.

## III. METHODOLOGY

### 3.1 Architecture Overview

The proposed system (Fig. 1) consists of:

1. Input Image: Captures real-time industrial scenes.
2. Feature Extraction: MobileNetV3 backbone extracts low- and high-level features.
3. Transformer Encoder–Decoder: RT-DETR's transformer head processes global relationships among detected objects.
4. Detection Head: Generates class predictions and bounding boxes for PPE components.
5. Set-Based Loss: Uses Hungarian matching to optimize one-to-one prediction accuracy.

### 3.2 Mathematical Components

Depthwise Separable Convolution:

•$FLOPsdw\_pw=H\times W\times(C_{in}K^2+C_{in}C_{out})$

This reduces computation 4–5× compared to standard convolution.

Hard-Swish Activation:

$h\_swish(x)=x\cdot ReLU6(x+3)/6$

Self-Attention (Transformer Head):

$Attention(Q,K,V)=softmax(QK^T/\sqrt{d})V$

Set-Based Detection Loss:

•$L=-\log p^{cls}+\lambda_1\|b-b\hat{}\|1+\lambda_2(1-GIoU(b,b\hat{}))$

## IV. EXPERIMENTAL RESULTS

### 4.1 Dataset

A custom dataset containing 15,000+ annotated PPE images (helmets, vests, gloves) captured under various industrial conditions was used for training and testing.

### 4.2 Evaluation Metrics

The model performance is evaluated using:

- mAP (Mean Average Precision)
- FPS (Frames Per Second)
- Precision and Recall
- FLOPs and Parameter Count

| ISSN: 2582-7219 | www.ijmrset.com | Impact Factor: 8.206| | Monthly Peer Reviewed & Refereed Journal |

|| Volume 8, Special Issue 2, November 2025 ||

National Conference on Emerging Trends in Engineering and Technology 2025 (NCETET–2025)

Organized by

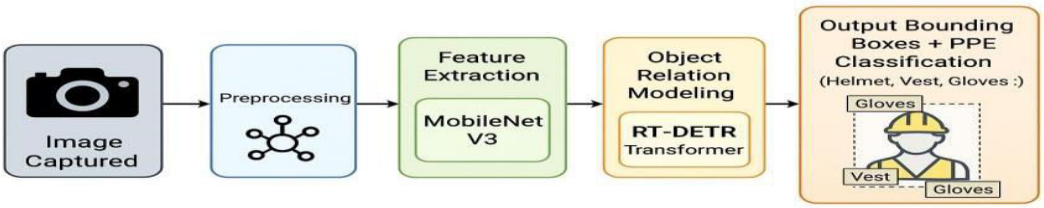
Mookambigai College of Engineering, Keeranur, Tamil Nadu, India

### 4.3 Comparative Table

| Model | mAP (%) | FPS | Parameters(M) | FLOPs (G) |
|---|---|---|---|---|
| YOLOv8-S | 90.7 | 45 | 11.2 | 28 |
| RT-DETR (ResNet-50) | 92.3 | 23 | 42.1 | 120 |
| RT-DETR-MV3 (Proposed) | 91.6 | 65 | 8.9 | 17 |

The proposed RT-DETR-MV3 provides a 3× improvement in speed and 60% reduction in FLOPs with negligible loss in accuracy, achieving real-time performance suitable for industrial deployment.

FIGURES




## V. CONCLUSION

This research presents a **lightweight and real-time PPE detection model** by integrating **MobileNetV3** with **RT-DETR**. The proposed architecture effectively balances speed and accuracy, enabling reliable PPE detection for industrial safety systems. Future work will focus on **edge deployment**, **multi-camera integration**, and expanding detection categories to include masks and boots.

## VI. ACKNOWLEDGEMENTS

The authors acknowledge institutional support and access to industrial datasets.

## REFERENCES

1. Hao Wang et al. (2024). Personal Protective Equipment Detection for Industrial Environments: A Lightweight Model Based on RT-DETR for Small Targets.
2. Howard, A. et al. (2019). Searching for MobileNetV3. Proc. ICCV.
3. Li, Z., & Zhang, Y. (2023). RT-DETR: Real-Time Detection Transformer. Baidu Research.

INTERNATIONAL STANDARD SERIAL NUMBER INDIA

# INTERNATIONAL JOURNAL OF MULTIDISCIPLINARY RESEARCH

## IN SCIENCE, ENGINEERING AND TECHNOLOGY

| Mobile No: +91-6381907438 | Whatsapp: +91-6381907438 | ijmrset@gmail.com |

www.ijmrset.com