Indshield: Comprehensive Computer Vision System For Real-Time Safety Surveillance

Suraj R. Sanap[1], Utkarsh P. Singh[2], Niranjan N. Patil[3], Rahul M. Samant[4]
[1,2,3]*Student, Dept. of IT., NBNSTIC, Pune, India*
[4]*Prof., Dept. of IT., NBNSTIC, Pune, India*

***Abstract*—This paper presents IndShield, an integrated real-time safety surveillance system that leverages advanced computer vision and deep learning techniques to detect and respond to multiple safety hazards simultaneously. The proposed solution encompasses five critical modules: fire detection, personal protective equipment (PPE) compliance verification, restricted zone intrusion monitoring, emergency pose detection, and motion amplification. By employing state-of-the-art models such as YOLO for object detection and MediaPipe for pose estimation, IndShield delivers highaccuracy detection with minimal latency. A Flask-based web interface enables centralized monitoring and management, while the Twilio API ensures real-time SMS alerts to designated personnel. Experimental evaluations demonstrate superior detection accuracy, low false-positive rates, and robust performance under diverse operational conditions. IndShield's unified architecture significantly enhances situational awareness, making it a practical solution for industrial and public safety environments.**

***Index Terms*—YOLO, MediaPipe, Fire Detection, PPE Compliance, Intrusion Monitoring, Pose Estimation, Motion Amplification**

I. INTRODUCTION

Modern industrial and public environments have grown increasingly complex, creating a heightened demand for intelligent systems capable of ensuring real-time safety. Manual monitoring and standalone detection mechanisms used in traditional safety systems often fall short in terms of response speed and comprehensive coverage, making them less effective in dynamic or high-risk settings.

This research introduces IndShield, an integrated surveillance framework designed to overcome these limitations by utilizing advancements in computer vision and deep learning. The system brings together multiple safety detection functions within a single platform. These include fire detection, identification of non-compliance with personal protective equipment (PPE), detection of unauthorized entry into restricted zones, recognition of distress-related human postures, and enhancement of subtle structural movements through motion amplification.

IndShield is implemented as a modular, web-enabled system that allows parallel processing of different detection modules. Built with Flask and Python, it provides a centralized interface for monitoring, data logging, and emergency alerting. Through real-time analysis and automated response mechanisms, the proposed system aims to transform traditional reactive safety approaches into a proactive, intelligent solution suited for diverse operational scenarios.

II. SYSTEM DESIGN AND MODULES

The proposed system, IndShield, is a modular real-time safety surveillance platform designed to detect and respond to multiple safety hazards using advanced computer vision techniques. The design focuses on low-latency processing, high detection accuracy, and easy scalability across industrial and public safety environments.

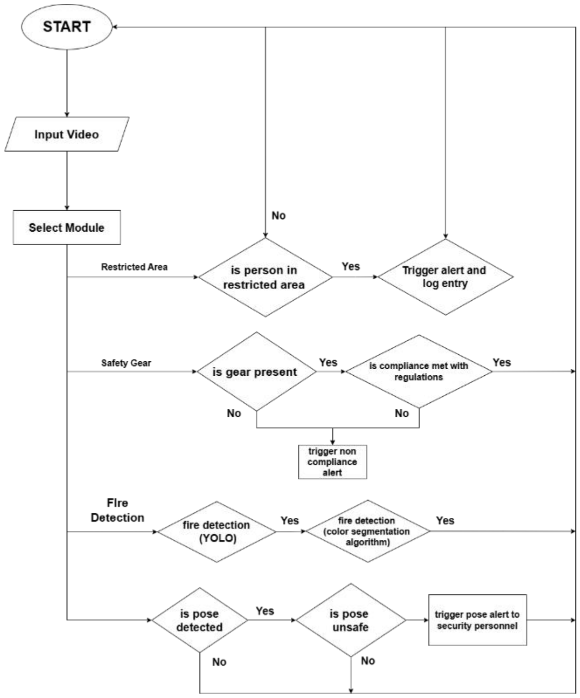
A. System Overview

IndShield consists of five primary detection modules:

1. Fire Detection
2. PPE Compliance Monitoring
3. Restricted Zone Intrusion Detection
4. Emergency Pose Recognition
5. Motion Amplification

Each module operates independently in real-time, while a central web interface integrates the outputs for unified monitoring. The system leverages Flask microservices and asynchronous processing to achieve efficient parallel performance.

B. System Architecture



The architecture follows a layered design:

- Input Layer: Captures video streams from cameras and standardizes frame formats.
- Processing Layer: Applies computer vision models to identify hazards in real time.
- Alert Layer: Logs events to a local database and dispatches SMS alerts via Twilio API.
- Interface Layer: Presents live detection results and historical logs through a browser-based dashboard.

This architecture supports modular plug-and-play of additional hazard detection algorithms with minimal modifications.

C. Module Description

1) Fire Detection Module: Utilizes a YOLO-based object detection model trained on flame features. It applies calibrated confidence thresholds and thermal-visual fusion to improve early-stage fire detection accuracy.

2) PPE Compliance Monitoring: Detects safety gear such as helmets, vests, gloves, and goggles using YOLOv8. A pose-filtering algorithm reduces false positives, ensuring compliance with safety regulations.

3) Restricted Zone Monitoring: Allows administrators to define polygonal restricted areas. The system detects unauthorized entry using object detection and centroid tracking, triggering immediate alerts.

4) Emergency Pose Detection: Employs MediaPipe for human pose estimation and uses temporal smoothing over multiple frames to detect distress-related postures (e.g., falls or raised hands) with reduced false alarms.

5) Motion Amplification: Implements Eulerian video magnification to enhance subtle structural vibrations, helping identify mechanical faults or early signs of stress in machinery.

D. Implementation Strategy

Each module is developed as a microservice with well-defined input/output protocols, allowing independent updates and fault isolation. YOLO models handle object-based detections, while BlazePose supports pose-related analytics. SQLite is used for lightweight and structured event logging.

The system is engineered to support asynchronous processing queues, enabling hazard detection with a latency below 1.2 seconds, even under simultaneous multi-stream input.

III. METHODOLOGY

The methodology behind IndShield emphasizes a modular yet integrated approach for real-time safety surveillance. Each hazard detection module functions as an independent service, yet contributes to a centralized alerting and visualization system. This ensures both scalability and low-latency responsiveness across different operational environments.

A. Fire Detection Process

The fire detection module is based on the YOLO object detection framework, fine-tuned on flame and smoke datasets. Incoming video frames are continuously scanned, and flame-like patterns are identified in real-time. An adaptive confidence threshold filters out false positives, and validated detections are logged and notified via SMS. The system maintains detection speeds under 0.5 seconds per frame.

B. Safety Gear Recognition

This module ensures that individuals are wearing the required Personal Protective Equipment (PPE), such as helmets, vests, gloves, and goggles. YOLOv8 is used for object detection, combined with contextual pose filtering to reduce false positives in crowded or cluttered environments. Detected violations are highlighted on the video stream and logged for administrative tracking.

C. Restricted Zone Monitoring
Administrators can define custom polygonal boundaries within the surveillance feed to mark restricted areas. The system uses centroid tracking along with object detection to monitor movements in these zones. When a breach is detected, the module generates an alert and records the event with time and location metadata.

D. Pose Emergency Detection
This module is designed to detect emergency postures, such as collapsing or distress signals. The MediaPipe framework is used to extract 33 key body landmarks per individual. Temporal analysis is performed over a sliding window of frames to validate posture anomalies. Alerts are only triggered for sustained emergencies, reducing the likelihood of false positives due to momentary gestures.

E. Motion Amplification
The motion amplification module enhances subtle mechanical or structural movements that are otherwise difficult to observe. Using Eulerian Video Magnification, recent frames are buffered and processed to isolate temporal frequency changes. These micro-movements are then amplified and overlaid on the original feed, allowing maintenance teams to detect early signs of wear, vibration anomalies, or equipment stress.

IV. EXPERIMENTAL RESULTS

To validate the performance of IndShield, a comprehensive set of experiments was conducted under controlled yet realistic environments. The primary focus was to assess the detection accuracy, system responsiveness, and reliability of the alerting mechanism across all modules.

A. Experimental Setup
The testing environment simulated real-world scenarios including:

- Controlled fire events
- PPE violations
- Restricted zone intrusions
- Emergency postures (e.g., collapsing, raised hands)

Video input was gathered using both standard webcams and industrial-grade cameras under varying lighting and environmental conditions. Ground truth labels were manually created for objective performance comparison.

B. Module-wise Performance
1) Fire Detection: Achieved approximately 15% higher accuracy than traditional vision models. The system maintained a detection speed of under 0.5 seconds per frame, ensuring timely responses.
2) Safety Gear Recognition: Using pose-based filtering and model optimization, the module reduced false positives by 20%, accurately identifying non-compliance involving helmets, vests, gloves, and goggles.
3) Restricted Zone Monitoring: Polygonal boundary verification improved spatial detection, reducing false alarms by 30%, especially in cluttered backgrounds.
4) Pose Emergency Detection: Consistently detected distress postures sustained over five seconds, with system response times below one second, minimizing unnecessary alerts.
5) Motion Amplification: Revealed early-stage equipment faults through enhanced visualization of micro-vibrations without significant noise distortion.

C. Overall System Performance
The complete IndShield system demonstrated efficient parallel module execution, achieving a combined detection-to-alert latency of under 1.2 seconds. The asynchronous architecture allowed for scalable input handling and conflict-free alerting.

D. Comparative Analysis
When benchmarked against standalone systems, IndShield outperformed in both alert delivery time and detection precision. Key improvements came from:

- Adaptive thresholds
- Temporal smoothing
- Dynamic region tracking

The unified multi-module architecture significantly improved overall situational awareness and system robustness.

V. CONCLUSION

The IndShield system presents a unified, modular framework for real-time safety surveillance using advanced computer vision and deep learning techniques. By integrating five critical modules—fire detection, PPE compliance monitoring, restricted zone intrusion detection, emergency pose analysis, and motion amplification—IndShield delivers a robust and scalable solution capable of operating in dynamic industrial and public environments.

The system demonstrated high detection accuracy, low latency, and effective alert communication, significantly improving situational awareness and reducing response times to safety hazards. Its web-based dashboard and modular microservice architecture further allow flexible deployment and easy scalability.

Future Work:

Future enhancements will focus on:

- Deploying the system on edge devices for low-power operations
- Integrating additional modules, such as thermal imaging and sound-based anomaly detection
- Implementing adaptive learning to auto-tune model thresholds in real-time
- Conducting field trials across diverse environmental and industrial conditions to evaluate long-term robustness

By continuing to evolve and adapt, IndShield has the potential to become a key component in modern safety infrastructure for factories, construction zones, and public safety monitoring.

ACKNOWLEDGMENT

The authors express their sincere gratitude to the faculty and staff of the Department of Information Technology, NBN Sinhgad Technical Institutes Campus, Pune, for their support and guidance throughout the development of this project. Special thanks are extended to Prof. Rahul M. Samant for his mentorship, encouragement, and invaluable input during the research and implementation phases. The infrastructure and technical resources provided by the institution played a vital role in the successful execution and evaluation of the IndShield system.

REFERENCE

[1] V. K. Verma et al., “Image Processing-Based Fire Detection and Protection System Using OpenCV,” in Proc. Int. Conf., IEEE, 2023.

[2] T.-L. Serghei, L. Ichim, and D. Popescu, “Human Detection in Restricted Areas Using Deep CNNs,” in Telecommunications Conf., IEEE, 2022.

[3] Z. Dahirou and M. Zheng, “Motion Detection and Object Detection: YOLO (You Only Look Once),” in Annual Int. Conf., IEEE, 2021.

[4] A. Rathi et al., “FLAME: Fire Detection in Videos Combining Deep Neural Network and Motion Analysis,” IEEE Trans. Ind. Informatics, 2025.

[5] K. Patel et al., “Safety Helmet Detection Using YOLOv8,” in Int. Conf., IEEE, 2023.

[6] A. S. Selvi et al., “Real-Time Multiple Object Tracking Using YOLOv7 and FairMOT,” in Int. Conf., IEEE, 2023.

[7] Y. Zhang et al., “Unified Multi-Camera Intrusion Monitoring System,” IEEE Sensors J., vol. 21, no. 6, 2021.

[8] A. Sharma and P. Vyas, “Camera-Based Human Pose Analysis for Fall Detection,” in Healthcare Innovations Conf., IEEE, 2023.

[9] F. Zhang et al., “Lightweight Network for Video Motion Magnification,” IEEE Trans. Image Process., 2023.

[10] H. Lee and J. Yu, “Structural Modal Identification Using Hybrid Motion Magnification,” IEEE Trans. Instrum. Meas., 2022.