YOLO-MineSafe: A Vision-Based Abnormal Fall Detection And Emergency Alert Framework For Isolated Mining Workers

Dr. A. Mary Beula[1], Kingston J[2], Kowshik S[3], Vishal J[4], Sameer Ahamed S[5]
[1] Head of Department, Dept of Artificial Intelligence & Data Science Engineering
[2, 3, 4, 5] Dept of Artificial Intelligence & Data Science Engineering
[1, 2, 3, 4, 5] M.A.M College of Engineering and Technology,Tiruchirappalli, Tamil Nadu, India

Abstract- *Mining operations consistently rank among the world's most hazardous occupational environments, with workers stationed in isolated areas facing undetected fall risks, sudden health emergencies, and life-threatening incidents that current safety systems cannot address in real-time. Existing solutions, such as passive closed-circuit television (CCTV), wearable accelerometers, and manual supervision, fail to deliver autonomous, real-time incident detection across the expansive and harsh terrain of active mine sites. This study introduces YOLO-MineSafe, a vision-based fall detection and emergency alert framework purpose-built to close this gap. The system continuously processes surveillance camera videos using a fine-tuned YOLOv8 deep learning model, extracting bounding box geometry, posture orientation, and inter-frame motion vectors to identify anomalous body positions. A temporal classification module employing a 20-frame confirmation window at a 0.4 confidence threshold reliably distinguished genuine fall events from ordinary work postures, such as bending or crouching. Upon confirmed detection, multichannel emergency alerts are dispatched immediately: an annotated incident image via email, an SMS to registered supervisors, and a simultaneous local audio alarm — all without human intervention. The system operates effectively in low-light and dust-heavy environments through dedicated preprocessing, requires no wearable devices, and provides a complete incident audit trail, representing a substantive advance toward reducing preventable fatalities in isolated mining environments.*

Keywords: YOLOv8, Fall Detection, Mining Safety, Deep Learning, Computer Vision, Emergency Alert System, Posture Analysis, Real-Time Monitoring, Worker Safety, Intelligent Surveillance.

I. INTRODUCTION

Among all industrial sectors, mining consistently stands out because of the sheer density and severity of hazards that workers encounter daily. Whether navigating underground tunnels where roof collapses, toxic gas pockets, and complete darkness threaten lives without warning, or working in surface open-pit environments where heavy excavation machinery, unstable slopes, and extreme thermal conditions converge, miners face a uniquely elevated daily risk profile. The International Labour Organization (ILO) has documented that, despite employing a small fraction of the global workforce, the mining sector accounts for a disproportionate share of fatal occupational accidents worldwide.

Beyond purely mechanical hazards, the sustained physical demands of mining labor, such as prolonged shift durations, whole-body vibration exposure, and chronic dust inhalation, substantially elevate the incidence of sudden medical events, such as cardiac episodes, heat-related collapse, and acute loss of consciousness. When these events occur in isolated corners of a mine, they can prove as fatal as any structural accident, particularly when no co-worker or supervisor is present to recognize the situation and summon assistance.

The spatial reality of modern mining considerably amplifies this problem. Workers routinely operate deep within underground tunnel networks or are scattered across sprawling open-pit benches, far beyond any supervisor's direct sightline. Traditional safety protocols, such as regular radio check-ins, buddy system rotations, and manual patrol schedules, depend entirely on the affected worker remaining conscious and communicative. When a miner collapses or loses responsiveness, all communication-reliant safety mechanisms fail simultaneously. Passive closed-circuit television (CCTV) networks document events faithfully but provide zero real-time analytical capability. Wearable sensor platforms have been explored as alternatives; however, the abrasive, high-vibration, and thermally extreme conditions of active mining steadily degrade the sensor fidelity.

The rapid maturation of deep learning-based computer vision has opened a new path. The YOLO family of single-stage object detectors, and specifically YOLOv8, has demonstrated that human postures and body orientations can be identified and classified in real time at practical frame rates on hardware well within typical industrial budget constraints. This study presents YOLO-MineSafe, which combines fine-tuned YOLOv8 detection, posture-motion analysis, temporally robust fall classification, and a multichannel automated notification engine to address this challenge directly. Sections are organized as follows: Section II reviews prior work; Section III states the problem; Section IV details the architecture; Section V covers requirements;

Section VI presents the implementation and results of the study. Finally, Section VII concludes the paper.

II. RELATED WORK

The literature on automated fall detection has expanded substantially, primarily propelled by elderly care and smart home applications. Newaz and Hanada [1] conducted a comprehensive survey classifying methods into sensor-based, vision-based, and hybrid categories, concluding that vision-based approaches suffer pronounced performance degradation under low-light conditions and partial occlusions, which are precisely the conditions most prevalent in underground mining. Gaya-Morey et al. [2] examined CNN, RNN, and transformer architectures for elderly fall detection, reporting strong benchmark accuracy but acknowledging that substantial computational resources and domain-specific annotated datasets are required, both difficult to provision in resource-constrained industrial mining settings.

Federated learning has been proposed to circumvent the need for centralized data. Qi et al. [3] developed FL-FD, a multimodal federated fall detection system. Despite its privacy advantages, its dependence on stable inter-node network connectivity renders it impractical for underground mining, where communications are often intermittent. Cho et al. [4] demonstrated a continuous-wave radar system with binarized neural networks that achieves non-invasive detection without cameras. However, radar sensors cover narrow angular fields and cannot be economically scaled to cover a large mine.

Liu and Zhang [5] developed a self-powered triboelectric nanogenerator wearable that records fall kinematics — innovative, but still subject to the compliance and durability limitations that impede all wearable approaches in harsh mining environments. Floor-based pressure-sensing systems reviewed by Kaur et al. [6] detect body-impact signatures on monitored surfaces; however, their geographic confinement makes full-site mining coverage prohibitive. Sykes [7] combined skeletal pose estimation with Vision Transformer architectures for superior temporal reasoning, but at computational costs exceeding typical mining site hardware. Lupion et al. [8] proposed a depth-and-thermal multi-modal system effective in darkness but requiring significant dual-sensor hardware investment. Rahman et al. [9] released the FallVision benchmark dataset whose controlled domestic content may exhibit domain shift when applied to underground mining conditions. A clear gap persists in the literature: no system has been designed for the specific combination of geographic isolation, degraded visual conditions, wearable-hostile environments, and the absolute requirement for fully autonomous alerting that defines industrial mining. YOLO-MineSafe directly addresses this issue.

III. PROBLEM STATEMENT

Global mining fatality data underscore this urgency. The ILO estimates that, despite employing less than one percent of the world’s workforce, mining is responsible for roughly eight percent of all fatal occupational accidents annually [10]. A substantial proportion of these deaths involve workers who sustained a fall or acute health event in an isolated location and were not discovered before their condition became irreversible. The fundamental issue is not that falls are unpreventable, but that the existing safety infrastructure cannot autonomously detect them and trigger an immediate response.

Each existing solution has disqualifying limitations in the mining context: supervisory coverage cannot span expansive mine workings; wearables degrade under harsh conditions and rely on worker compliance; CCTV records are not analyzed; and radio protocols fail when a worker is incapacitated. Generic fall detection systems from elder care or smart home research are not adapted to the low-light, high-particulate, network-limited, and geometrically constrained conditions of real mining environments. YOLO-MineSafe addresses this directly by automatically detecting falls in real time without wearables and triggering immediate multichannel alerts to supervisors.

IV. PROPOSED SYSTEM — YOLO-MINESAFE

A. System Overview

YOLO-MineSafe is a non-intrusive, wearable-free, fully automated fall-detection and emergency-notification system. Continuous video from fixed cameras was processed frame-by-frame using a fine-tuned YOLOv8 model. A posture motion analysis module extracts the body orientation and inter-frame centroid velocity. A temporal classification module applies a 20-frame confirmation window at a minimum confidence of 0.4 to filter genuine falls from incidental postures. Upon confirmation, SMS, email with annotated frame, and audio alarm are dispatched simultaneously without human intermediary, ensuring supervisors are notified within seconds.

B. System Architecture

1) Data Acquisition Module: Surveillance cameras capture continuous video from key mining locations. Each frame was preprocessed by resizing to 640×640 pixels, histogram equalization to compensate for inadequate lighting, and Gaussian filtering to attenuate particulate visual noise. These steps are essential for reliable detection in the degraded visual conditions of mining.

2) YOLO-Based Human Detection Module: Each preprocessed frame is processed by the fine-tuned YOLOv8n model, producing bounding boxes, class labels (fall detected/normal), and confidence scores. Multiple workers are tracked simultaneously using persistent inter-frame identifiers. The single-stage architecture delivers real-time frames per second (FPS) on mid-range hardware. The trained weights were loaded from best.pt.

3) Fall and Abnormality Detection Module: Three metrics are computed per tracked worker: (i) bounding box height-to-width aspect ratio — a transition from upright to horizontal yields a characteristic ratio increase; (ii) inter-frame centroid displacement velocity — rapid downward movement signals a fall; and (iii) body orientation angle relative to vertical. The anomalous aspect ratio change combined with an elevated downward centroid velocity constitutes the fall signal, distinguishing it from the crouching and forward lean common in mining tasks.

4) Classification Module: Raw fall signals are not immediately processed. Persistence across 20 consecutive frames at a confidence level of ≥ 0.4 is required. At 15–25 fps, this equals approximately one second of sustained anomalous posture, sufficient to eliminate transient false positives while maintaining clinically meaningful response latency. Events that resolved within the window were discarded without notification.

5) Alert and Notification Module: Upon threshold satisfaction, alert.jpg is captured with a bounding box, class label, confidence score, and UTC timestamp. The image is emailed to all registered supervisors via smtplib, an SMS is dispatched via a third-party API, and winsound triggers the local audio alarm. Every confirmed incident is logged in MySQL for auditing and compliance.

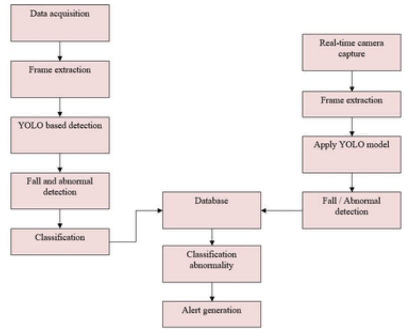

Fig. 1: YOLO-MineSafe System Architecture and Workflow

C. Advantages Over Existing Systems

YOLO-MineSafe offers five key advantages. First, no wearable devices are required, eliminating compliance and durability issues. Second, YOLOv8 single-stage inference delivers real-time performance on economical devices. Third, multi-worker simultaneous tracking does not require per-worker configuration. Fourth, redundant multi-channel alerting (SMS, email, audio) remains effective even if one channel fails. Fifth, mining-specific preprocessing maintains the detection performance under degraded visual conditions that defeat general-purpose systems.

V. SYSTEM REQUIREMENTS

Hardware requirements are deliberately modest: a dual-core processor at a minimum of 2.6 GHz, 4 GB RAM, and 320 GB storage suffice for single-site operation. Standard H.264-encoding IP cameras are fully compatible. The backend is Python-based; Flask provides the web interface; YOLOv8 inference runs via Ultralytics; OpenCV manages video ingestion and preprocessing; MySQL stores incident logs and credentials; and HTML/CSS/JavaScript form the frontend. The development was performed using PyCharm on Windows. The key libraries used were ultralytics, OpenCV, smtplib, requests (SMS API), TensorFlow, and winsound.

VI. IMPLEMENTATION AND RESULTS

A. Dataset and Training

Training used a composite dataset combining the FallVision benchmark [9] — providing diverse fall videos across angles, subject types, and fall scenarios — with supplementary footage from simulated and actual mining environments featuring dusty air, hard-hat-wearing subjects, confined spaces, and poor lighting. The raw video was sampled at 5 fps; near-duplicate frames were filtered; each retained frame was resized to 640×640, Gaussian filtered, and histogram-equalized before bounding box annotation into fall detected and normal classes. Transfer learning from the pretrained YOLOv8n was conducted for 100 epochs, with a batch size of 16 and an initial learning rate of 0.01. The best validation mAP checkpoint was saved as best.pt.

B. Fall Detection Results

The deployed model processes videos at real-time frame rates on the specified hardware. Frames in which the fall-detected confidence reaches ≥0.4 contribute to the temporal confirmation window. The 20-frame threshold was empirically validated: windows below 15 frames yielded excessive false positives from worker bending, whereas windows above 25 frames introduced notification delays incompatible with rapid emergency response. Upon confirmation, alert.jpg is emailed within seconds, SMS reaches supervisors simultaneously, and the winsound alarm is activated immediately. Every confirmed event was logged in MySQL.

C. System Interface

The Flask web application provides role-based access across the following interface screens.

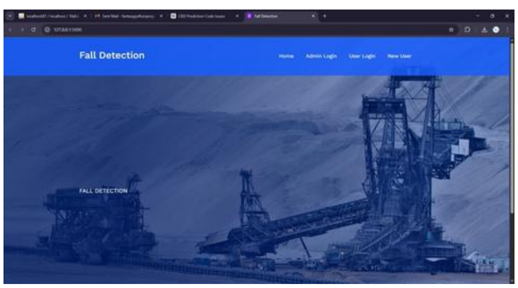

Fig. 2: YOLO-MineSafe Home Page Interface

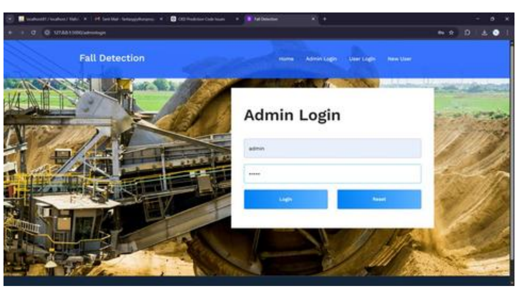

Fig. 3: Admin Login Interface

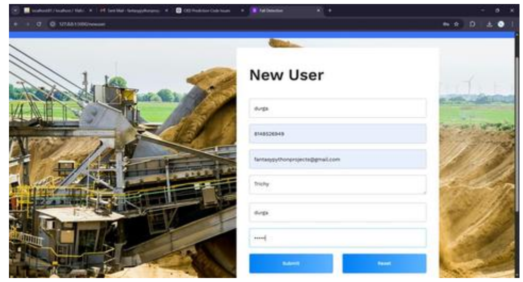

Fig. 4: New Safety Officer Registration Form

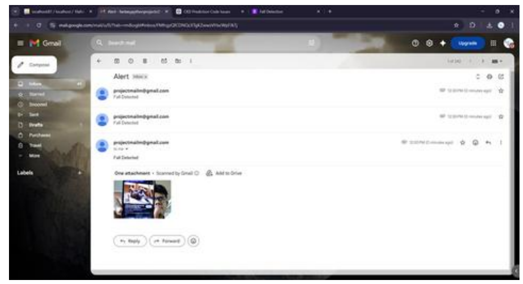

Fig. 5: Automated Email Alert with Annotated Incident Frame

The Home Page (Fig. 2) provides navigation to the Admin Login, User Login, and New User registration. Admin Login (Fig. 3) restricts system configuration to authorized users through secure credential verification. The New User Registration form (Fig. 4) captures the safety officer's contact details used by the automated alert module. Fig. 5 shows the automated email received by the supervisors upon fall confirmation, displaying the annotated alert.jpg with a bounding box, class label, confidence score, and UTC timestamp. Additional interface modules include a User Dashboard for live monitoring and an incident log for historical review.

D. Performance Comparison

TABLE I: COMPARATIVE ANALYSIS OF FALL DETECTION SYSTEMS

System	Det. Type	Wear?	Real-Time	Alert	Mine Suit.
Newaz & Hanada [1]	Hybrid	Yes	Partial	None	Low
Qi et al. FL-FD [3]	Federated	Yes	No	None	Low
Sykes Pose+ViT [7]	Vision	No	Partial	None	Moderate

Lupion et al. [8]	Depth+IR	No	Yes	Partial	Low
YOLO-MineSafe	**YOLOv8**	**No ✓**	**Yes ✓**	**SMS+Email+Audio ✓**	**High ✓**

✓ = Full support Partial = Limited High/Low = Mining suitability

As shown in Table I, YOLO-MineSafe is the only system that combines camera-only detection, wearable-free operation, real-time inference, comprehensive multi-channel alerting, and high mining suitability, uniquely positioning it as a deployable industrial safety solution.

VII. CONCLUSION

This study presented YOLO-MineSafe, a real-time camera-based fall detection and emergency alert framework designed from first principles for isolated mining worker safety. The system closes a gap that wearables, passive closed-circuit television (CCTV), and manual supervision have consistently failed to address: autonomous detection of worker falls with immediate, multi-channel emergency response, requiring no wearables and no human intermediary. The 20-frame confirmation window and 0.4 confidence threshold, validated empirically, balanced false-positive suppression against response latency. The comparative analysis in Table I confirms YOLO-MineSafe's unique positioning across all dimensions relevant to mining safety deployment, representing a meaningful practical step toward reducing preventable fatalities in one of the world's most hazardous working environments.

VIII. FUTURE WORK

Planned extensions include 3D skeletal pose estimation for richer fall characterization beyond 2D bounding box metrics, Vision Transformer (ViT) integration for improved temporal reasoning over body motion sequences, edge computing deployment on camera-embedded processors to improve network-outage resilience underground, predictive analytics to detect pre-fall behavioral signatures enabling proactive intervention, multi-modal sensor fusion with thermal cameras, gas monitors, and vibration sensors, and reinforcement learning for adaptive site-specific detection threshold calibration.

REFERENCES

[1] S. M. R. Newaz and M. Hanada, "The methods of fall detection: A literature review," Sensors, vol. 23, no. 11, p. 5212, 2023.

[2] F. X. Gaya-Morey, C. Manresa-Yee, and J. M. Buades-Rubio, "Deep learning for computer vision-based activity recognition and fall detection of the elderly," arXiv:2401.11790, 2024.

[3] B. Qi, D. Chiaro, and F. Piccialli, "FL-FD: Federated learning-based fall detection with multimodal data fusion," Information Fusion, vol. 99, p. 101890, 2023.

[4] H. Cho, S. Lee, J. Park, and S. Kim, "Fall detection based on continuous wave radar sensor using binarized neural networks," Applied Sciences, vol. 15, no. 2, p. 546, 2025.

[5] Y. Liu and Z. Zhang, "A flexible and stretchable triboelectric nanogenerator for self-powered wearable fall detection," ChemistryOpen, vol. 14, no. 6, 2025.

[6] M. Kaur, R. Singh, and P. Arora, "State-of-the-art fall detection techniques with emphasis on floor-based systems," Biomedical Engineering Advances, vol. 9, p. 100179, 2025.

[7] E. Sykes, "Next-generation fall detection: Harnessing human pose estimation and transformer technology," Health Systems, vol. 14, no. 2, pp. 85-103, 2025.

[8] M. Lupion, N. C. Cruz, J. Ortega, and P. M. Ortigosa, "Privacy-aware fall detection and alert management in smart environments," Internet of Things, vol. 30, p. 101526, 2025.

[9] M. M. Rahman, A. Islam, S. Hossain, and M. N. Uddin, "FallVision: A benchmark video dataset for fall detection," Data in Brief, vol. 59, p. 111440, 2025.

[10] International Labour Organization (ILO), "Safety and health in mining: ILO standards and guidelines," ILO Publications, Geneva, Switzerland