Tuijin Jishu/Journal of Propulsion Technology
ISSN:1001-4055
Vol.46 No. 2 (2025)

AI-Powered Safety Monitoring on Metro Platforms with Computer Vision Integration

Paveen Paul*1 , Jaya Christiyan K G2 , A.Ajina3

1&2 Department of Mechanical Engineering, Ramaiah Institute of Technology, Bengaluru - 54

3 Department of AI&ML, Ramaiah Institute of Technology,, Bengaluru-54

**Corresponding: Paveen Paul*

Abstract: The urgent need for improved safety measures arises from the rising number of accidents and deaths on metro platforms, particularly in view of crowding in recent years. This paper examines how computer vision (CV) and artificial intelligence (AI) might assist metro stations' real-time safety monitoring to be improved. It centres on how advanced artificial intelligence models—such as deep learning and pose estimation models—may identify safety concerns, including people crossing safety lines (yellow lines), falling by accident, or exhibiting indicators of suicide attempts. According to the studies gathered for this review, computer vision can monitor platform activity constantly, identify dangerous behaviour, and alert others to stop mishaps. Human posture and motions are also studied using machine learning, which produces better safety predictions than conventional cameras. AI and Internet of Things (IoT) technologies together simplify data processing and site-based quick decision making. This paper also examines how robots might support safety by means of automated barriers, crisis assistance, and adherence to AI-based safety guidelines. These technologies taken together indicate a time when public transport will have smarter and more automatic safety systems. The paper summarises the most recent advancements, addresses present issues in the field, and offers ideas for next studies on artificial intelligence-based safety systems for metro platforms.

Keywords: Artificial Intelligence (AI), Computer Vision, Deep Learning, YOLO, Metro Platform Safety, Fall Detection, Crowd Monitoring, Real-time Surveillance

1. Introduction

Every day millions of people use metro platforms, which are vital components of city transportation. Safety on these platforms is still a major issue, though, particularly in metropolitan areas where accidental falls and even suicide attempts can occur. Because of slow response times, blind spots, and human errors, conventional safety measures, including CCTV and hand monitoring, sometimes fail sufficiently.

Particularly in deep learning and computer vision, new artificial intelligence (AI) advancements have opened better means of real-time monitoring of metro platforms.

These days, artificial intelligence systems can view video footage and identify odd behaviour or dangerous activity as well as rapidly and precisely predict possible collisions [1], [2]. These systems can identify problems including falling, crossing the safety line or sudden crowd movements and send quick alarms to stop injury [3], [4].

Strong outcomes in real-time safety monitoring have come from technologies including pose estimation methods [8], object detection with YOLO models [12], [18], and crowd behaviour analysis [31], [43]. Combining these instruments with artificial intelligence and Internet of Things (IoT) systems greatly increases risk detection accuracy and speed [22],[24], [28]. Moreover, using smaller and faster models like YOLO with Transformers helps the system to function even in very crowded metropolitan stations [20], [46]. Still, some problems remain to be resolved. These comprise the great demand for computer power, blocked views in videos, and faster reactions in various environments [47], [48]. Important issues that need to be addressed are privacy concerns, the capacity of artificial intelligence to operate in novel environments, and knowledge of how AI forms decisions [16], [49].

Focussing on how computer vision is applied on metro platforms, this review gathers recent advancements in AI-based safety systems. Deep learning to forecast accidents, computer vision for live monitoring, connecting artificial intelligence with the Internet of Things, crowd behaviour analysis, and future directions and current challenges identification cover five main domains.

Looking at these subjects, this paper demonstrates how intelligent surveillance might increase metro system safety.

2. Literature Review

The lietrature review investigates several spheres related to the intended work.It covers crowds behaviour analysis and anomaly detection and computer vision for real-time safety monitoring, artificial intelligence and deep learning for accident prediction and prevention.

2.1 AI and Deep Learning for Accident Prediction and Prevention

Using cutting-edge approaches to enhance risk assessment and mitigation, the combination of artificial intelligence and deep learning for accident prediction and prevention in metropolitan platform safety [1] emphasises the need of examining accident precursors in metropolitan railway systems and suggests a safety maturity model (SMM) assessing operational, behavioural, and technical aspects to raise safety performance. Showing that deep learning overcomes traditional statistical models with accuracy, [2] present a convolutional neural network (CNN) model enhanced with a genetic algorithm to predict and classify degree of accident. Using IoT and real-time data from roadside units and weather conditions, [3] presents Spatio-Temporal Conv-Long Short-Term Memory Autoencoder (STCLA) framework achieving an AUROC score of 0.94, so indicating great predictive capability.

With the Word2Vec-BiGRU-CNN model attaining an F1-score of 0.3648, [4] propose an ensemble deep learning model combining BiGRU and CNN for traf-based accident duration prediction, so demonstrating the benefits of multi-modal data and feature fusion. Using EfficientNetB1 and MobileNetV2 CNN architectures for accident detection on synthetic image datasets, employ transfer learning to achieve a Mean Average Preci-sion (mAP) of 0.89, so quite appropriate for real-time metro platform surveillance.

[6] investigate AI-powered decision-making for road safety optimisation by means of proba-bilistic language Multi-Criteria Decision Making (MCDM), so integrating Sugeno-Weber aggregation models for predictive traffic management. Emphasising the need of cross-modal knowledge sharing to improve AI-driven safety solutions, [7] do a comparative study across transportation modes identifying Artificial Neural Networks (ANN), Support Vector Machines (SVM), Hidden Markov Models, and Bayesian models as the most effective in transportation safety.

These results show that deep learning models—including CNNs, BiGRU, LSTM, and hybrid architectures—much enhance accident prediction and classi-fication. IoT and real-time monitoring—including traffic unit data and weather—along with roadside unit data and environmental conditions improve the accuracy of AI-based safety systems. In real-time metro platform surveillance, transfer learning with synthetic datasets shows promise; probabilistic language artificial intelligence models help to prevent accidents by besting multi-criteria traffic decisions. Metro systems can increase safety monitoring, accident prediction and prevention by using these AI-driven technologies, so strengthening passenger security and operational effectiveness.

2.2 Computer Vision for Real-Time Safety Monitoring

Using deep learning models and advanced detection techniques, several approaches have been investigated extensively on the application of computer vision for real-time safety monitoring in metro platforms. Combining vision transformers for enhanced accuracy with transformer-based pose estimating models including OpenPose, BlazePose, and HRNet, [8] proposed a human fall detection frame-work using pose estimation techniques. By using skeleton-based techniques instead of RGB video feeds, their method greatly reduces computational complexity while yet preserving high detection accuracy, so prioritising privacy. Using TensorFlow Lite's PoseNet model, [9] also presented a fall detection system extracting 17 key joint points from video frames and LSTM model analysis. Key-point detection combined with sequential modelling reduced the effects of environmental elements including shadows and lighting and improved inference times. Temporal Convolutional Networks (TCN) combined with Transformer Encoders (TE) developed TCNTE, a real-time skeleton-based fall detection algorithm aiming at increasing robustness and real-time efficiency. Appropriate for metro safety uses, their model was optimised for edge computing on NVIDIA Jetson Orin NX devices and achieved accuracies exceeding 97% when tested on publicly available fall detection datasets. Using PTZ cameras and convolutional neural networks (CNNs), [11] presented a deep learning-based method for pedestrian detection and suspicious activity identification in pedestrian monitoring. Their system detects unusual activities in real-time, so improving metro safety monitoring.

[12], who suggested a railway obstacle intrusion warning mechanism combining YOLOv5-based object detection with an enhanced risk assessment framework, also significantly contribute to metro platform safety. Their approach guarantees more dependable safety monitoring in metropolitan settings by efficiently separating hazardous objects and classifying risks depending on degree of intensity. [13] further investigated human activity recognition using a combination of CNN, ConvLSTM, and LRCN models, obtaining high accuracy in detecting movement patterns vital for identifying risky behaviours on metro platforms. [14] presented a two-stage pose estimation model for enhanced human action recognition

combining a single-shot detector (SSD) with convolutional pose machines (CPM). Their approach extracts features from ResNet and implements multi-scale transformations to increase the accuracy of identifying human activity depending on body movement in real-time video surveillance. In a similar vein, [15] investigated crowd behaviour using a technique known as pattern virtualisation, which clarifies individuals movement and gathering behaviour in metropolitan stations. This helps to prevent potentially dangerous events including crowding and stampedes.

Combining computer vision tools with reasoning models, [16] developed a system to identify harmful human actions, such as damaging protected areas, so handling security threats. Their approach uses explainable artificial intelligence to make the decisions of the system clearer and more accurate in metro station surveillance, and it analyses how people interact with objects. At last, [17] demonstrated how fast people cross by using YOLOv4 and Deep SORT algo-rithms. Their studies show that pedestrian tracking driven by artificial intelligence can enhance metro area crowd control and safety.

Drawing on earlier research, more modern real-time safety systems for metro platforms now use more sophisticated AI technologies to better detect unusual activity, enhance security, and safeguard pedestrians.Using a YOLOv8x model with BOT-SORT for object tracking, [18] presented a video-based safety methodology for pedestrian crosswalk monitoring. Their work concentrated on computing post-encroachment time (PET) and time-to-- collision (TTC) measures from CCTV footage, so illustrating how artificial intelligence-driven risk assessment might be used in high-density settings such as metro platforms.

With a precision of 0.899, the suggested system attained a mAP of 0.861, so improving real-time risk prediction for cars and pedestrians.

Further enhancing AI applications in safety monitoring, [19] suggested a dynamic risk management system using YOLOv4 and MobileNetV2 for spotting dangerous behaviours, especially in staircase use within office workspaces. Their system, with a 98.1% detection accuracy, showed how convolutional neural networks (CNNs) could be used in settings needing continuous real-time monitoring—a notion that fits rather well to metro platform surveillance. [20] also made a significant contribution by developing a real-time surveillance system with an attention-based Transformer-YOLOv8. With a mean average precision (mAP) of 89.67%, this model uses Transformer atten-uation techniques to enhance how features are detected, so attaining a precision of 96.78% and a recall of 96.89%. It is quite handy for maintaining metro stations safe since it performs well in crowded and complicated surroundings. In a related vein, [21] created an intelligent system using YOLO to track railway platforms. When a train arrives, their system automatically monitors platform crowding, modulates lighting depending on the number of people, and sends safety alarms should someone cross the safety line. The YOLO model lets quick and accurate real-time monitoring since it can manage up to 2000 bounding boxes per second, so helping to prevent mishaps on metropolitan platforms.

These fresh advances demonstrate how deep learning and artificial intelligence are supporting metropolitan platform safety. Combining several models including CNNs, LSTMs, TCNs, and Trans-formers with live surveillance systems has helped researchers to detect falls, track individuals, and identify barriers. Including risk assessment instruments such as PET (Post-Encroachment Time) and TTC (Time to Collision) improves the system's pre-dicting risk control ability even more. Additionally useful for object recognition in challenging environments are transformer-based YOLO models. More metro systems using artificial intelligence-powered monitoring makes it abundantly evident how crucial these technologies are for early risk detection, crowd management, and general safety enhancement.

2.3 Integration of IoT and AI in Metro Safety Systems

IoT and AI combined in metro safety systems has been greatly advancing real-time data collecting, edge computing, and machine learning models to improve operational efficiency, accident prevention, and safety monitoring. Emphasising real-time facial recognition technologies (FRT) and AI-driven behavioural monitoring for metro security, [22] address the function of AI-powered public surveillance systems. Their studies show how artificial intelligence can evaluate enormous volumes of data to identify anomalies and forecast security risks, so enabling proactive surveillance systems. Concurrently, [23] suggests a fresh method for crowd monitoring in metropolitan platforms based on an enhanced Mix Gaussian background model for segmentation and Adaptive Boost classi-fiers for density estimate. Their approach provides consistent crowd density estimate to prevent overcrowding and possible safety risks, so addressing the blocked problems in crowded stations.

By means of real-time monitoring and risk prediction, [24] present an AI-powered railway control and dispatching system that inte-grates multi-modal AI models for optimising metro scheduling, so cutting delays, and so enhancing passenger safety. Their system greatly increases transportation efficiency by using automated decision-making based on big-scale machine learning algorithms. [25] Based on real-time telemetry and data consistency analysis, present an intelligent monitoring framework for public transport systems.

Their research underlines how artificial intelligence can help to maximise traffic flow, minimise running delays, and enhance passenger safety in metropolitan stations by means of vehicle scheduling. By means of computer vision and artificial intelligence to examine vehicle trajectories and traffic conflicts, [26] further extend this research by suggesting surrogate safety measures (SSMs) that improve predictive analytics for metro and urban railway safety. [27] develops a system for identifying high-risk pedestrian zones in metro stations by means of pedestrian behaviour analysis utilising street view images and transfer learning approaches. Their AI-based system offers important information for the design of metro infrastructure since it achieves 87.1% accuracy in spotting older pedestrians and evaluating walking surroundings.

[28] train object detection models for safety compliance using game engine simulations, so introducing synthetic data generation for AI-based safety monitoring. Their method helps metro monitoring systems driven by artificial intelligence become more resilient, so lowering false alarms and increasing the accuracy of hazardous behaviour detection. Generally, [29] indicates an AI-driven system ensuring safety compliance by combining Building Information Modelling (BIM) with computer vision for real-time metro workforce safety monitoring.

[30] analyse the effect of characteristics and luggage presence on movement efficiency by statistical modelling of pedestrian behaviour at metro platforms. Their results show that using automated staircases boosts pedestrian flow by 14.5% for those carrying bags and 25.3% for elderly people, so emphasising how artificial intelligence-driven infrastructure enhancements might increase metro safety. From real-time surveillance and predictive risk analysis to pedestrian monitoring and automated safety interventions, these studies taken collectively show the transforming power of IoT and artificial intelligence in urban safety systems. Edge computing, computer vision, and advanced machine learning combined guarantees a more proactive and intelligent approach to metro safety, so decreasing accidents while improving overall passenger security.

2.4 Crowd Behavior Analysis and Anomaly Detection

By applying deep learning, computer vision, and simulation-based approaches to improve real-time monitoring and risk assessment, artificial intelligence-driven methods have evolved the study of crowd behaviour analysis and anomaly detection in metro platform safety.With convolutional deep neural networks and data mining techniques, [31] presented a deep learning-based framework for group trajectory outlier detection with an 88% accuracy rate in identifying collective aberrant behaviours in pedestrian environments. Their model surpassed conventional statistical models in capturing spatiotemporal dependencies. Similarly, [32] created a real-time surveillance system using a Convolutional Neural Network (CNN) with Modi-fied Threshold-Centric K-means Clustering, so enhancing activity classification and tracking accuracy in metro stations. Their approach effectively identified aberrant pedestrian behaviour including running, abrupt stops, and other anomalies.

Further on this work, [33] developed RMPD-DCNN-EL, a multi-modal pedestrian detection framework combining SimAM EfficientNet with Nested Long Short-Term Memory (NRSTM), Deep Belief Networks (DBNs), and Extreme Learning Machines (ELM). Tested on the INRIA dataset, their model obtained an accuracy of 99.3%, so greatly enhancing pedestrian tracking capacity under various environmental conditions. [34] investigated fall detection in high-density metro stations by means of depth image processing to reduce occlu-sion and illumination variations, so offering an efficient way of safety monitoring. Likewise, [35] presented a multi-spectral pedestrian detection system leveraging deep neural networks for improved pedestrian recognition, especially in low-light metropolitan settings.

Further developments in the field include [36], who suggested a multispectral pedestrian detection system employing R-FCN and Network-in- Network (NIN) fusion.

Their approach was quite useful for metro surveillance systems since it was quite good in spotting small and blocked pedestrian events. [37] presented a lightweight pedestrian detection model based on YOLOv8s, optimising small-scale feature extraction via BiFPN and VoVGSCSP, so greatly enhancing detection accuracy in congested metropolitan platforms. Concurrent with this, [38] created GC-YOLOv9, a novel smart city traffic monitoring system with direct relevance to metro platform crowd monitoring that integrates Ghost Convolution and shows exceptional performance on BDD100K and Cityscapes datasets.Using RepViT and C2f improvements into the YOLOv8s framework, [39] presented the Z-YOLOv8s model.

In complex, high-density environments, this model enhanced pedestrian tracking accuracy, so supporting metro station safety precautions. Finally, [40] suggested a high accuracy in aberrant pedes-trian movement detection by means of an efficient data-driven behaviour identification system employing vision transformers, so reducing dependency on big datasets. In addition to these results, [41] presented RS-YOLO, a high-accuracy object detection system specifically for intricate metropolitan settings.

To improve feature extraction, the model combines a Multi-Scale Path Aggregation Feature Pyramid Network (MS-PAFPN) with SPPFormer, so raising detection accuracy by 3.7% over YOLOv8s. Particularly in high-traffic areas, their results showed great applicability for metro station crowd monitoring. [42] investigated visually impaired pedestrian safety at metro stations, noting important risk factors including platform layout, inconsistent tactile paving, and fall prevention barriers. Their research highlights how much improved environmental adaptations are needed to reduce fall hazards. [43] using cellular automaton (CA)-based simulations, investigated small-group pedestrian behaviour and evacua-tion dynamics at metropolitan stations.

Their results showed that close-contact pedestrian groups have longer evacuation times, slower movement speeds, and more crowd interference, which calls for tailored station design to maximise safety. Using smart card data, [44] suggested a seven-phase risk model to examine crowd-gathering risk in metropolitan platforms. Their model provided a strong basis for predictive safety monitoring by clearly pointing up areas of high risk within metropolitan networks. [45] developed a robust multi-modal pedestrian detection framework, RMPD-DCNN-EL, which combines Extreme Learning Machines (ELM), Deep Belief Networks (DBNs), SimAM EfficientNet with Nested Long Short-Term Memory (NLSTM). Tested on the INRIA dataset, their model attained an outstanding 99.3% detection accuracy, so greatly improving pedestrian tracking capability under different environmental conditions.

These advances taken together show the increasing importance of artificial intelligence and deep learning in anomaly detection for metro platform safety and crowd behaviour analysis.

CNNs, vision transformers, multispectral imaging, and advanced tracking models taken collectively greatly improves metro platform real-time monitoring capacity. Moreover, the application of smart card data and simulation-driven risk mod-els gives metro operators predictive instruments to control safety hazards, avoid events of crowding, and maximise passenger management methods.

2.5 Challenges and Future Directions in AI-Powered SafetyMonitoring

Although revealing thrilling future directions, the AI-powered safety monitoring in metro platforms presents many difficulties including computational efficiency, real-time data processing, and model interpretabl-ity. Inspired on an enhanced YOLOv8 framework, [46] presented PV-YOLO, a lightweight pedestrian and vehicle detection model. Suitable for real-time metro safety monitoring, this model reduces computational complexity by using a bidirectional feature pyramid network (BiFPN) and receptive-field attention convolution (RFAConv), so improving small object detection accuracy. Adapting such light-weight models for big-scale urban networks with heavy passenger density still presents difficulties though. [47] investigated deep learning methods for pedestrian tracking and detection, so stressing the limits of CNN-based models in environments with high occlusion. Their work proposes hybrid CNN-RNN models and 3D vision systems' integration to raise metro station object recognition accuracy.

[48] investigated using YOLOv8 for intelligent transport systems (ITS), worrying the need of effective aerial monitoring solutions for vehicle and pedestrian tracking. Although YOLOv8 exceeded its pre-decessors in precision and inference speed, the study found that variations in lighting conditions and object occlusions compromise its performance in real-time applications including metro safety monitoring. [49] presented the Human Behaviour Detection Dataset (HBDset), a fresh dataset particularly meant for emergency management and evacuation safety. Their study shows how object detection algorithms driven by artificial intelligence (AI) can improve public safety when taught on various datasets. The study emphasises, nevertheless, the need of improved model generalisation to identify various pedes-trian behaviours, particularly in unsure metropolitan environments.Based on convolutional neural networks (CNNs) with a mixed attention mechanism, [1] presented a fast and light-weight train fault detection model, FL-Tinet. By 96.37%, their model improved detection speed over state-of- the-art techniques, so highlighting the possibilities of lightweight deep learning architectures for real-time safety monitoring in urban layouts. Notwithstanding these advances, the difficulty in managing detection accuracy with computational economy remains. [51] looked into accident causes in urban railroads and developed a safety maturity model (SMM) to evaluate operational and managerial-related risk factors. Their research highlights the need of combining artificial intelligence with safety systems to develop improved plans for risk reduction and prediction. Future studies in AI-based metro safety monitoring concentrate on creating hybrid models combining CNNs, Transformers, and reinforcement learning to provide anomaly detection more accuracy. By means of synthetic datasets such as HBDet, one can strengthen these models and increase their ability for understanding various forms of pedestrian behaviour.

Edge AI and transfer learning can also help real-time monitoring be improved, so decreasing response times and safeguarding privacy in surveillance systems. Future research has to address issues including blocked views (occlusion), scaling to larger systems, and high computing requirements even if cur-rent AI models show great progress in recognising people and spotting odd behaviour.

3. Comparative Analysis and Discussion

Results from recent research on AI-based safety monitoring for metro platforms are presented in this section, with an emphasis on performance comparisons, typical uses, and the technologies' applicability. According to the research, this is a rapidly expanding field where computer vision and deep learning—particularly YOLO-based object detection models—play a crucial part.

3.1 Dominance of YOLO in Real-Time Detection

The most commonly employed frameworks for actual time pedestrian detection, object recognition, and safety monitoring tasks were found to be YOLO (You Only Look Once) variants across the reviewed works [12], [18], [20], [21], [42], and [44]. Because of their capacity for finding a balance between accuracy and inference speed, they are especially well-suited for urban settings where prompt decision-making is essential. With the help of attention mechanisms, more recent models like YOLOv8 and Transformer-YOLO hybrids [20], [42] have enhanced small object detection and feature representation, accomplishing mean average precision (mAP) scores of up to 89.67%.

3.2 Comparative Effectiveness of AI Architectures

Other AI architectures also perform well for specific applications, yet YOLO dominates detection tasks:

- Activity recognition and recognition of falls are common uses for LSTM and ConvLSTM models, especially when temporal sequence knowledge is present [9], [23].
- Bringing together transformer-based models with pose estimation or TCNs improves recognition robustness in crowded and varied lighting [8, 33, 42].
- In heterogeneous metro data environments, ensemble deep learning methods (e.g., BiGRU-CNN, EfficientNet + DBN + ELM) provide better generalisation and accuracy [4], [9], [45].

The findings indicate that hybrid architectures, which combine temporal and spatial modelling (e.g., CNN + LSTM), are especially useful for identifying anomalies such as suspicious activity, unexpected falls, or erratic behaviour.

Table.1 *Comparative analysis of AI models used in metro platform safety monitoring system*

Model	***Use Case***	***Accuracy / mAP***	***Edge Deployment***	***Reference***
YOLOv4	*Object Detection*	*mAP ~89%*	*Yes*	*[35]*
YOLOv8	*Real-time Monitoring*	*mAP ~89.67%*	*Yes*	*[21]*
Transformer-YOLOv8	*Object Detection with Attention*	*Precision 96.78%, Recall 96.89%*	*Yes*	*[42]*
PV-YOLO	*Lightweight Detection*	*High accuracy on edge devices*	*Yes*	*[15]*
RS-YOLO	*High Accuracy Detection*	*3.7% better than YOLOv8s*	*Yes*	*[44]*

EfficientNet + MobileNetV2	*Accident Detection*	*mAP 0.89*	*Yes*	*[5]*
OpenPose + LSTM	*Fall Detection*	*>97% accuracy*	*Yes*	*[6]*
TCN + Transformer	*Skeleton-based Fall Detection*	*High accuracy (exact not provided)*	*Yes*	*[8]*

3.3 Trends in Application-Specific Deployment

The literature exposes important areas of concentration in metro safety monitoring:

- Fall Detection: Made possible by OpenPose and BlazePose merged with LSTM or Vision Transformers using pose estimation models [6] [8].
- Applied using virtual simulation models and object tracking to evaluate congestion and forecast stampedes, crowd monitoring and density analysis [15] [31], [43].
- Using YOLO-based models and edge-optimized neural networks [12], [21], [20], we addressed obstacle and line intrusion detection.

Crucially, several systems were tuned for real-time deployment on edge devices, especially NVIDIA Jetson modules [10], [20], [42], so reflecting a move towards embedded artificial intelligence solutions that lower latency and infrastructure reliance.

3.4 Gaps and Limitations in Existing Literature

Notwithstanding great progress, several difficulties still exist:

- Many models show good performance on benchmark datasets but lack validation in actual metro systems.
- Heavally packed or low-light environments still compromise detection accuracy for regular CNNs [47], [48].
- Few studies specifically address data security, surveillance ethics, or explainability in artificial intelligence predictions [16], [19].

In addition restricted by reliance on proprietary or synthetic datasets is consistency and model comparison between them. Open benchmark datasets particular to metro platform safety scenarios would help the field.

3.5 Toward Integrated and Predictive Safety Systems

Studies frequently show the drive towards integrated safety ecosystems whereby IoT sensors, real-time telemetry, and dynamic risk modelling [22], [24], [26] AI is coupled with. modifying from reactive to proactive safety management, these systems not only detect hazards but also predict and prevent them. Moreover, the use of robotics and automated interventions have grown more and more important in order to either help with emergency response or physically prevent errors (by means of barriers).

4. Future Directions

As metro systems keep updating, integration of artificial intelligence and computer vision into safety monitoring systems is likely to become routine. To ensure a strong, scalable, and ethically aligned deployment, several technical and operational issues still need to be resolved though. For research and development in AI-powered metro platform safety systems, the following main areas show the most satisfying future directions.

4.1 Advancements in Lightweight and Edge-Optimized Models

Safety monitoring systems deployment in real-world urban areas calls for models that not only are accurate but also effective in terms of computation and utilisation of energy. Strong prospective for real-time inference on edge devices like NVIDIA Jetson is shown by lightweight variants including PV-YOLO [15], RS-YOLO [44], and YOLOv8s [16].

Emphasising low-latency response, small memory footprints, and high frame-rate processing in high-density environments, further research should keep enhancing models for embedded systems.

4.2 Fusion of Vision Transformers and Hybrid Architectures

In video-based safety systems, Vision Transformers (ViTs) have proven rather successful in capturing long-range dependencies and contextual understanding [33], [42]. Particularly in dynamic and occlusion-heavy settings like metro platforms, combining ViTs with convolutional backbones or recurrent layers (e.g., ConvLSTM, TCN) may improve activity recognition. Investigating hybrid and ensemble models could help to improve the generality and dependability of safety measures.

4.3 Privacy-Preserving and Explainable AI

Future systems have to include privacy-preserving technologies like

- skeleton-based analysis instead of raw video given ethical questions regarding public surveillance[8]
- On-device processing to stop data transfer
- outputs of anonymised detection

Integrating explainable artificial intelligence (XAI) models will also help to guarantee open decision-making. Particularly in high-stakes situations involving human safety [16], [34], these techniques can help system operators interpret model outputs and boost public confidence.

4.4 Synthetic Data Generation and Augmented Training

One of the main limitations still is the lack of differed and annotated datasets unique for metro environments. Several studies suggest synthetic data creation with game engines or simulation tools [28], which can generate diverse environmental conditions and rare hazard scenarios. This strategy not only improves model training but also supports safer validation of artificial intelligence systems under stress-test scenarios impractical or unethical to replicate in real life.

4.5 Real-Time Predictive Risk Modeling

Reactive detection to proactive predictive analytics should be the shift in future safety monitoring tools. Early warning systems can be created by including computer vision with metrics like time-to-collision (TTC) and post-encroachment time (PET) [18], crowd flow modelling [44], and behavioural precursors [50], so anticipating risks before events start. These features will be quite important in avoiding mistakes both during busy times and when many people are on board.

4.6 Integration of Robotics and Autonomous Intervention

The long-term vision of the project describes how the integration of robotic systems provides a transforming degree of active safety. Future studies should investigate:

- Robotic fences stopping dangerous platform access
- Mobile robots or drones for instant surveillance
- AI analytics driven autonomous emergency response teams

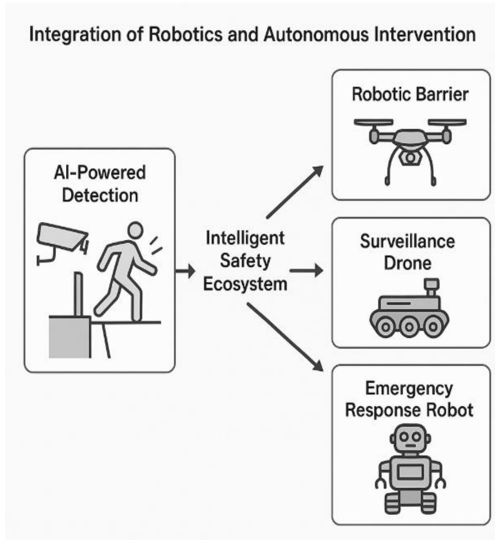

Fig. 4.1 Integration of AI in Robotics for safety monitoring

When combined with AI-powered detection systems, these technologies could create an intelligent, autonomous safety ecosystem able to both instantly recognise and address hazards.

5. Conclusion

Demand for intelligent and proactive safety solutions has grown as metro systems take front stage in contemporary urban mobility. Emphasising the increasing dominance of YOLO-based object detection models and hybrid AI architectures including CNNs, LSTMs, and vision transformers, this review underlines the critical role of artificial intelligence and computer vision in improving real-time safety monitoring on metro platforms. While edge computing systems further support real-time responsiveness and infrastructure independence, these technologies have shown great ability for recognising safety-critical events including falls, intrusions, and crowd congestion. Predictive analytics and artificial intelligence combined with IoT systems indicate a change towards anticipatory safety management. Still, there are difficulties including visual occlusion, poor model generalisation, ethical questions about monitoring, and a dearth of strong, varied datasets. interacting with these problems calls for multidisciplinary invention in fields including robotics integration, data-generated augmentation, and privacy-preserving artificial intelligence. This study helps to build intelligent, flexible, and ethically based safety systems for next-generation metro environments by aggregating present developments and pointing strategic directions ahead.

References

[1] rez-Sala, L., Curado, M., Tortosa, L. and Vicent, J.F. (2023). Deep learning model of convolutional ural networks powered by a genetic algorithm for prevention of traffic accidents severity. Chaos, Solitons Fractals, [online] 169, p.113245. doi:https://doi.org/10.1016/j.chaos.2023.113245.

[2] ei, X. (2024). Enhancing road safety in internet of vehicles using deep learning approach for real-time cident prediction and prevention. International Journal of Intelligent Networks, 5, pp.212–223. i:https://doi.org/10.1016/j.ijin.2024.05.002.

[3] ien, J., Tao, W., Jing, Z., Wang, P. and Jin, Y. (2024). Traffic accident duration prediction using multi-ode data and ensemble deep learning. Heliyon, 10(4), pp.e25957–e25957. i:https://doi.org/10.1016/j.heliyon.2024.e25957.

[4] lhadi, A., Djenouri, Y., Srivastava, G., Djenouri, D., Lin, J.C.-W. and Fortino, G. (2021). Deep learning pedestrian collective behavior analysis in smart cities: A model of group trajectory outlier detection. ormation Fusion, 65, pp.13–20. doi:https://doi.org/10.1016/j.inffus.2020.08.003.

[5] ana, L. and Seldev Christopher, C. (2025). A deep learning behavior analysis model for efficient video :veillance using multi pose features. Ain Shams Engineering Journal, 16(2), p.103245. i:https://doi.org/10.1016/j.asej.2024.103245.

[6] za, A., Yousaf, M.H., Ahmad, W., Velastin, S.A. and Viriri, S. (2025). Human fall detection using pose imation: From traditional machine learning to vision transformers. Engineering Applications of Artificial elligence, [online] 143, p.109809. doi:https://doi.org/10.1016/j.engappai.2024.109809.

[7] amchur, N., Shakhovska, N. and Gregus ml., M. (2022). Person Fall Detection System Based on Video eam Analysis. Procedia Computer Science, 198, pp.676–681. i:https://doi.org/10.1016/j.procs.2021.12.305.

[8] ı, X., Wang, C., Wu, W. and Xiong, S. (2025). A Real-time skeleton-based fall detection algorithm based temporal convolutional networks and transformer encoder. Pervasive and Mobile Computing, [online] 7, p.102016. doi:https://doi.org/10.1016/j.pmcj.2025.102016.

[9] epak Kumar Jain, Zhao, X., Garcia, S. and Subramani Neelakandan (2024). Robust multi-modal destrian detection using deep convolutional neural network with ensemble learning model. Expert stems with applications, 249, pp.123527–123527. doi:https://doi.org/10.1016/j.eswa.2024.123527.

[10] ıng, S.-W. and Lin, S.-K. (2014). Fall detection for multiple pedestrians using depth image processing hnique. Computer Methods and Programs in Biomedicine, 114(2), pp.172–182. i:https://doi.org/10.1016/j.cmpb.2014.02.001.

[11] o, Y., Guan, D., Wu, Y., Yang, J., Cao, Y. and Yang, M.Y. (2019). Box-level segmentation supervised ep neural networks for accurate and real-time multispectral pedestrian detection. ISPRS Journal of otogrammetry and Remote Sensing, 150, pp.70–79. doi:https://doi.org/10.1016/j.isprsjprs.2019.02.005.

[12] ng, L., Wang, Y., Laganière, R., Huang, D. and Fu, S. (2020). Convolutional neural networks for ıltispectral pedestrian detection. Signal Processing: Image Communication, 82, p.115764. i:https://doi.org/10.1016/j.image.2019.115764.

[13] n, D.K., Zhao, X., González-Almagro, G., Gan, C. and Kotecha, K. (2023). Multimodal pedestrian tection using metaheuristics with deep convolutional neural network in crowded scenes. Information sion, 95, pp.401–414. doi:https://doi.org/10.1016/j.inffus.2023.02.014.

[14] ıwande, U., Hajari, K. and Golhar, Y. (2023). Real-Time Deep Learning Approach for Pedestrian tection and Suspicious Activity Recognition. Procedia Computer Science, 218, pp.2438–2447. i:https://doi.org/10.1016/j.procs.2023.01.219.

[15] ı, Y., Huang, Z., Song, Q. and Bai, K. (2025). PV-YOLO: A lightweight pedestrian and vehicle detection odel based on improved YOLOv8. Digital Signal Processing, 156, p.104857. i:https://doi.org/10.1016/j.dsp.2024.104857.

[16] ang, F., Leong, L.V., Yen, K.S. and Zhang, Y. (2024). An enhanced lightweight model for small-scale destrian detection based on YOLOv8s. Digital Signal Processing, [online] 156, p.104866. i:https://doi.org/10.1016/j.dsp.2024.104866.

[17] unetti, A., Buongiorno, D., Trotta, G.F. and Bevilacqua, V. (2018). Computer vision and deep learning hniques for pedestrian detection and tracking: A survey. Neurocomputing, 300, pp.17–33. i:https://doi.org/10.1016/j.neucom.2018.01.092.

[18] ı, R., Zhang, X., Sun, M. and Wang, G. (2024). GC-YOLOv9: Innovative smart city traffic monitoring lution. Alexandria Engineering Journal, 106, pp.277–287. doi:https://doi.org/10.1016/j.aej.2024.07.004.

[19] ntes, C., Hohma, E., Corrigan, C.C. and Lütge, C. (2022). AI-powered public surveillance systems: why (might) need them and how we want them. Technology in Society, [online] 71(0160-791X), p.102137.

i:https://doi.org/10.1016/j.techsoc.2022.102137.

[20] ang, Z., Chen, P., Huang, Y., Dai, L., Xu, F. and Hu, H. (2024). Railway obstacle intrusion warning echanism integrating YOLO-based detection and risk assessment. Journal of Industrial Information egration, [online] 38, p.100571. doi:https://doi.org/10.1016/j.jii.2024.100571.

[21] urat Bakirci (2024). UTILIZING YOLOv8 for ENHANCED TRAFFIC MONITORING in TELLIGENT TRANSPORTATION SYSTEMS (ITS) APPLICATIONS. Digital Signal Processing, 2, pp.104594–104594. doi:https://doi.org/10.1016/j.dsp.2024.104594.

[22] ao, R., Tang, S.H., Bin Supeni, E.E., Rahim, S.A. and Fan, L. (2024). Z-YOLOv8s-based approach for ad object recognition in complex traffic scenarios. Alexandria Engineering Journal, 106, pp.298–311. i:https://doi.org/10.1016/j.aej.2024.07.011.

[23] d. Ashraf Uddin, Md. Alamin Talukder, Muhammad Sajib Uzzaman, Debnath, C., Chanda, M., Paul, S., d. Manowarul Islam, Ansam Khraisat, Ammar Alazab and Aryal, S. (2024). Deep learning-based human tivity recognition using CNN, ConvLSTM, and LRCN. International Journal of Cognitive Computing in gineering. doi:https://doi.org/10.1016/j.ijcce.2024.06.004.

[24] n, R., Zhang, Q., Luo, C., Guo, J. and Chai, H. (2022). Human action recognition using a convolutional ural network based on skeleton heatmaps from two-stage pose estimation. Biomimetic Intelligence and botics, 2(3), p.100062. doi:https://doi.org/10.1016/j.birob.2022.100062.

[25] ago Tamagusko, Matheus Gomes Correia, Minh Anh Huynh and Ferreira, A. (2022). Deep Learning plied to Road Accident Detection with Transfer Learning and Synthetic Images. Transportation Research ocedia, 64, pp.90–97. doi:https://doi.org/10.1016/j.trpro.2022.09.012.

[26] hraf, S., Shahid, T., Kim, J., Hameed, M.S., Hezam, I.M. and Jana, C. (2024). AI-powered decision aking for road safety optimization under probabilistic linguistic Sugeno-Weber aggregation information. eliyon, [online] 10(19), p.e38594. doi:https://doi.org/10.1016/j.heliyon.2024.e38594.

[27] loria, A., Pineda Lezama, O.B. and Vargas, J. (2020). Analysis of crowd behavior through pattern tualization. Procedia Computer Science, [online] 175, pp.102–107. i:https://doi.org/10.1016/j.procs.2020.07.017.

[28] ı, X., Zheng, H., Wang, W. and Li, X. (2013). A novel approach for crowd video monitoring of subway atforms. Optik, 124(22), pp.5301–5306. doi:https://doi.org/10.1016/j.ijleo.2013.03.057.

[29] ı, J., Liu, G., Wang, Y. and Zhang, W. (2024). Artificial-intelligent-powered safety and efficiency provement for controlling and scheduling in integrated railway systems. High-speed Railway. i:https://doi.org/10.1016/j.hspr.2024.06.002.

[30] elentis, D.I., Papadimitriou, E. and van Gelder, P. (2023). The usefulness of artificial intelligence for fety assessment of different transport modes. Accident Analysis & Prevention, [online] 186, p.107034. i:https://doi.org/10.1016/j.aap.2023.107034.

[31] ınha, A., J. Bulas-Cruz and Monteiro, J.L. (1997). Intelligent Monitoring of Public Transport Systems - ıta Consistency Analysis -. IFAC Proceedings Volumes, [online] 30(19), pp.143–147. i:https://doi.org/10.1016/s1474-6670(17)42290-7.

[32] odel-Aty, M., Wang, Z., Zheng, O. and Abdelraouf, A. (2023). Advances and applications of computer sion techniques in vehicle trajectory generation and surrogate traffic safety indicators. Accident Analysis Prevention, [online] 191, p.107191. doi:https://doi.org/10.1016/j.aap.2023.107191.

[33] ang, J., Zhang, Z., Xiao, S., Ma, S., Li, Y., Lu, W. and Gao, X. (2023). Efficient data-driven behavior entification based on vision transformers for human activity understanding. Neurocomputing, 530,

.104–115. doi:https://doi.org/10.1016/j.neucom.2023.01.067.

[34] . Z., Song, X., Chen, S. and Demachi, K. (2024). Armed boundary sabotage: A case study of human ılicious behaviors identification with computer vision and explainable reasoning methods. Computers d Electrical Engineering, [online] 121, p.109924. i:https://doi.org/10.1016/j.compeleceng.2024.109924.

[35] ımed Ghomashchi, Paterson, J., Novak, A.C. and Dutta, T. (2024). Estimating pedestrian walking speed street crossings using the YOLOv4 and deep SORT algorithms: Proof of principle. Applied gonomics/Applied ergonomics, 119, pp.104292–104292. i:https://doi.org/10.1016/j.apergo.2024.104292.

[36] ıoohom, N., Chotivatunyu, P., Maitrichit, N., Nilsumrit, C. and Iamtrakul, P. (2024). The video-based ́ety methodology for pedestrian crosswalk safety measured: The case of Thammasat University, ailand. Transportation Research Interdisciplinary Perspectives, [online] 24, p.101036. i:https://doi.org/10.1016/j.trip.2024.101036.

[37] . D., Wang, R., Grekousis, G., Liu, Y. and Lü, Y. (2023). Detecting older pedestrians and aging-friendly ılkability using computer vision technology and street view imagery. Computers, Environment and Urban stems, 105, pp.102027–102027. doi:https://doi.org/10.1016/j.compenvurbsys.2023.102027.

[38] e, H., Jeon, J., Lee, D., Park, C., Kim, J. and Lee, D. (2023). Game engine-driven synthetic data neration for computer vision-based safety monitoring of construction workers. Automation in ınstruction, 155, pp.105060–105060. doi:https://doi.org/10.1016/j.autcon.2023.105060.

[39] ng, Y., Chen, X., Wang, Z., Zhang, Y. and Huang, X. (2024). Human Behaviour Detection Dataset BDset) Using Computer Vision for Evacuation Safety and Emergency Management. Journal of Safety ience and Resilience, 5(3), pp.355–364. doi:https://doi.org/10.1016/j.jnlssr.2024.04.002.

[40] rbosa, F.G.O., Mourao, G.L., Erazo, J.J.M., Ghedini, G.M. and Valverde, J.A. (2024). Dynamic risk ınagement in office workspaces: Real-time analysis of staircase safety using computer vision and AI. ımputers and Electrical Engineering, [online] 122, p.109902. i:https://doi.org/10.1016/j.compeleceng.2024.109902.

[41] ılinan, A.S., Park, M., Aung, P.P.W., Cha, G. and Park, S. (2024). Advancing construction site workforce ́ety monitoring through BIM and computer vision integration. Automation in Construction, [online] 158, .05227. doi:https://doi.org/10.1016/j.autcon.2023.105227.

[42] vya Nimma, Omaia Al-Omari, Pradhan, R., Zoirov Ulmas, Krishna, R.V.V., Yousef, T. and Rao, V.S.)25). Object detection in real-time video surveillance using attention based transformer-YOLOv8 model. exandria Engineering Journal, 118, pp.482–495. doi:https://doi.org/10.1016/j.aej.2025.01.032.

[43] . Kasiselvanathan, Teresa, V.V., K. Sangeetha and Arunraja A (2022). Intelligent System for Monitoring ıe of Safety in Railway Platforms. 2022 4th International Conference on Inventive Research in ımputing Applications (ICIRCA), [online] pp.1254–1258. i:https://doi.org/10.1109/icirca54612.2022.9985632.

[44] o, B., Wang, Y., Wang, P., Wang, H. and Yue, H. (2024). RS-YOLO: An efficient object detection ;orithm for road scenes. Digital Signal Processing, [online] 157, p.104889. i:https://doi.org/10.1016/j.dsp.2024.104889.

[45] ıgh, R.K., Raj, A. and Chattaraj, U. (2024). Statistical Analysis and Modelling of Pedestrian Behaviour Railway Platform. IFAC-PapersOnLine, [online] 58(10), pp.285–289. i:https://doi.org/10.1016/j.ifacol.2024.07.354.

[46] ızuno, T. and Tokuda, K. (2023). Reducing falls among visually impaired individuals on railway ıtforms: Field research on environmental challenges and solutions. Heliyon, 9(3), p.e14666. i:https://doi.org/10.1016/j.heliyon.2023.e14666.

[47] ang, L., Zeng, W., Zhou, P., Deng, X., Wu, J. and Wen, H. (2024). A fast and lightweight train image ılt detection model based on convolutional neural networks. Image and Vision Computing, [online] 154, .05380. doi:https://doi.org/10.1016/j.imavis.2024.105380.

[48] ang, Q., Qu, J. and Han, Y. (2023). Pedestrian small group behaviour and evacuation dynamics on metro tion platform. Journal of rail transport planning & management, 26, pp.100387–100387. i:https://doi.org/10.1016/j.jrtpm.2023.100387.

[49] ou, Y., Chen, J., Zhong, M., Li, Z., Zhou, W. and Zhou, Z. (2023). Risk analysis of crowd gathering on ɛtro platforms during large passenger flow. Tunnelling and Underground Space Technology, [online] 2, p.105421. doi:https://doi.org/10.1016/j.tust.2023.105421.

[50] ʳriakidis, M., Hirsch, R. and Majumdar, A. (2012). Metro railway safety: An analysis of accident ɛcursors. Safety Science, 50(7), pp.1535–1548. doi:https://doi.org/10.1016/j.ssci.2012.03.004.