Open Access International Journal of Science and Engineering

ISSN: 2456-3293

Volume 9 Issue 1s (2026) | Pages: 15-19

Missing Person Detection in Railway Stations Using AI-Based YOLOv8 and Clothing Attribute Recognition

Shruti Vishwakarma[1], Deepak Ray[2], Vedant Kakade[3], Pallavi Deshpande[4], Sakshi Muknak[5], Payal Kadam[6]

[1,2,3,4,5]Electronics & Telecommunication, Bharati Vidyapeeth (Deemed to be) University, Pune, India
[6]Symbiosis Institute of Technology, Pune Campus, Symbiosis International (Deemed University), Pune, India
[1]vishwakarmashruti51@gmail.com, [2]dkray@bvucoep.edu.in, [3]vedantkakade22@gmail.com, [4]psdeshpande@bvucoep.edu.in, [5]sakshimuknak3@gmail.com, [6]payal.kadam@sitpune.edu.in

Peer Review Information

Type: *Article*
Received: *29 March 2026*
Revised: *26 April 2026*
Accepted: *30 May 2026*
Published: *19 June 2026*

Abstract

Identifying missing persons in crowded public environments such as railway stations remains a difficult task for law enforcement agencies due to heavy passenger flow, frequent visual obstructions, and the limitations of manual CCTV monitoring. To address these challenges, this work presents an automated missing person detection approach based on artificial intelligence (AI) techniques, combining the YOLOv8 deep learning model with clothing attribute recognition [1]. The system detects individuals from surveillance video streams and extracts distinguishing clothing attributes, such as garment type and color, which assist in identification when facial information is partially visible or unavailable. The proposed system architecture consists of video preprocessing, YOLOv8-based person detection, clothing attribute analysis, and an alert generation mechanism. Experimental evaluation conducted on railway-like surveillance data achieves a mean Average Precision (mAP@0.5) of 94.4% while maintaining efficient processing suitable for real-time monitoring. The results demonstrate that incorporating clothing-based cues enhances identification reliability in crowded scenes and supports practical deployment in large-scale public surveillance infrastructures.

Keywords: Missing Person Detection; YOLOv8; CCTV Surveillance; Clothing Attribute Recognition; Computer Vision; Artificial Intelligence.

How to Cite This Article

Vishwakarma, S., Ray, D., Kakade, V., Deshpande, P., Muknak, S., & Kadam, P. (2026). Missing Person Detection in Railway Stations Using AI-Based YOLOv8 and Clothing Attribute Recognition. *Open Access International Journal of Science and Engineering*, 9(1s), 15-19.

Introduction

Railway stations are also some of the most congested open spaces, and countless masses of people pass across the platforms, waiting bays and entry points throughout the day. In these dynamic environment, missing person cases involving children, elderly and other vulnerable members are common cases reported. Current surveillance procedures require a substantial amount of manpower to be used in the observational process of CCTV footage, which is time- consuming, bulky, and prone to human errors, especially in situations where there are large groups of people and poor visibility. The latest advancements in artificial intelligence and deep learning have made it possible to process video streams using an auto-mated technique, which offers a nice solution to intelligent surveillance systems. Detectors that belong to the YOLO family are single-stage and are optimal with real- time surveillance due to their desirable trade-off between speed and accuracy [2]. Object detection models based on real- time have demonstrated a high potential of recognizing people in the congested and multifaceted scenes. In spite of these improvements, facial recognition systems fail to work well in railways because of high rate of occlusions, low image quality, different camera positions, and issues to do with privacy. To address these issues, a real-time person detection system of missing persons is introduced that combines person detection with the help of YOLOv8 and the identification of clothing attributes. Visual features in clothing such as color and type of garment are good indicators which are not affected even in cases where the faces are partially or entirely covered. The availability of such features in the detection pipeline increases the accuracy of the identification and adds robustness to the system in high-population railway stations.

Related Work

The intelligence surveillance systems have centered much attention on object detection and person identification. Previously, the approach usually used was based on manual features, including Haar cascades and Histogram of Oriented Gradi-ents (HOG) with conventional classification algorithms. As the field of deep learning progresses, the convolutional neural network detectors are the detectors of choice because they are more accurate and stronger. Two stage detection systems, like Faster R-CNN, are very accurate in detecting, but their computational requirements are too complex, making them unsuitable in real-time surveillance systems [3]. Single- stage detectors such as SSD and YOLO, in turn, provide a more reasonable trade-off between the speed and precision of detection and can be used as appropriate in CCTV-based surveillance in a crowded area. More recent versions of YOLO, e.g. YOLOv5 and YOLOv8, have also enhanced detection by refining the architecture and training regimen. Simultaneously, attribute-based person search and query-driven identification methods have also been studied, with semantic attribute such as clothing color and type of garment employed to find people in large-scale surveillance data [4]. Person re-identification and surveillance Person recognition Clothing attribute recognition has also been explored in person re-identification and surveillance, mainly in situations where facial recognition fails because of occlusions, low resolution images or privacy issues related limitations [5][6]. Proper literature suggests that clothing color and type of garments are good visual features used in tracking individuals over various frames and camera shots. Nevertheless, there is little literature that has dealt with combining real-time person detection with clothing attribute analysis to identify the missing person in the railway station environments. The given approach is based on this research gap.

Proposed System

- System Architecture: The system architecture consists of five primary parts, namely: (i) video input and preprocessing, (ii) person detection based on YOLOv8, (iii) clothing attribute extraction, (iv) matching and alert generation, and (v) visualization. Video streams obtained from CCTV cameras are preprocessed to balance resolution and lighting before being sent to the detection module. This modular architecture allows each component to be optimized independently and supports real-time deployment. Query-based identification is also supported, where a search can be performed using clothing attributes such as color and garment type, and the matching results are highlighted in the output stream.
- Person Detection and Clothing Attribute Extraction:YOLOv8 is used as the primary detection engine because of its single-stage architecture and ability to perform real- time inference. Bounding boxes and confidence scores are generated for each detected individual. The detected regions are cropped and passed to the clothing attribute extraction module. HSV (Hue, Saturation, Value) color-space analysis is used to determine clothing color, while garment types are classified using predefined visual categories. These extracted features serve as auxiliary identifiers when facial information is unavailable.

Methodology

- Dataset and Annotation: The dataset was generated through the Roboflow platform and consists of annotated images that were gathered on the CCTV scenes which looked more like railway station scenes [7]. It will comprise several lines of clothing, e.g., jacket, kurta, pant, saree, shirt, and t-shirt, which can be seen as the variety that is generally present in the open areas. The Roboflow tools were used to annotate bounding-boxes and version control of the data mass. To have an unbiased performance measure, the dataset was split into training, validation, and test groups.

- Model Training: The use of the Google Colab, which supports the use of a GPU, was used to train the model and make the learning process faster. The annotated dataset was fine-tuned on a pre- trained YOLOv8 model to adapt it to the conditions of a railway station surveillance better. In order to increase the model robustness, the data augmentation methods such as random scaling, varying brightness, and blur were used in the process of the model training. Standard loss curves and validation metrics were used to monitor model performance and early stopping was used to minimize the chance of overfitting.
- Image Processing and Attribute Extraction: The extraction of image preprocessing and clothing at- tributes was done using the OpenCV library [8]. The identified person regions were subjected to color-space transformation and HSV-based analysis was used to find the dominant color of clothing. The visual attributes were then extracted and acted as the second identifiers to aid in recognizing missing people in a crowded surveillance scene.
- Query Based Person Identification: Besides automated detection, the system can also be used to query people by name like the operators can use semantic description of clothes like red t-shirt or blue kurta. The queries that are user-defined are handled by mapping the given attributes to pre-determined clothing color and garment categories that are received in the detection stage. The people the user is interested in, which match the query criteria in their clothing properties, are shown in real-time with bounding box-es, and the non-matching detections are suppressed. The rationale behind this attribute-based querying model is that the previous studies on the concept of per-son re-identification and attribute-based surveillance infrastructures are based on this attribute and is especially applicable in highly populated settings where facial recognition cannot be done. The resultant visual emphasis allows quick localization of the target persons in the crowded CCTV scenes and therefore saves on the manual search effort.

Experimental Results

Evaluation Metrics

System performance was assessed using widely recognized object detection metrics, including Precision, Recall, and mean Average Precision (mAP). The evaluation emphasized both detection accuracy and real-time feasibility, which are essential requirements for practical surveillance applications.

Quantitative Results

The trained YOLOv8 model demonstrated strong detection performance, achieving a Precision of 89.8%, a Recall of 89.3%, and an mAP@0.5 value of 94.4%. An mAP@0.5:0.95 score of 66.2% indicates stable performance under more stringent overlap thresholds. Class-wise evaluation of clothing attributes reveals higher accuracy for jacket, kurta, and pant categories, while comparatively lower performance was observed for the saree category due to limited training samples.

Table 1. *Performance Metrics of YOLOv8 Model*

Metric	**Description**	**Value**
Precision	Ratio of correctly detected per-sons to total detections	0.898
Recall	Ratio of correctly detected per-sons to actual persons	0.893
mAP@0.5	Mean Average Precision at IoU threshold 0.5	0.944
mAP@0.5:0.95	Mean Average Precision across IoU 0.5–0.95	0.662

Graphical Results

Graphical analysis of the proposed system performance is provided to illustrate training convergence, class-wise classification behavior, and characteristics of the dataset. Fig. 1 illustrates the training convergence of the YOLOv8 model, showing the variation of box loss and mAP@0.5 over training epochs.

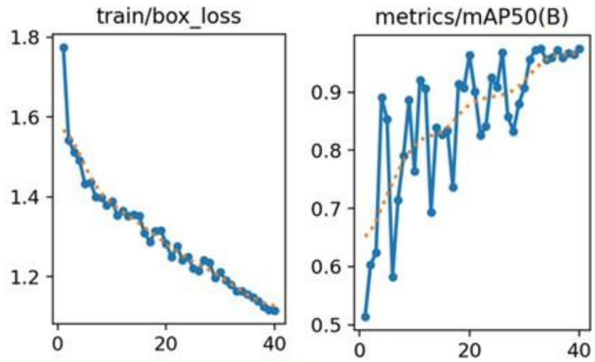

Fig. 1. *Training box loss and mAP@0.5 progression of the YOLOv8 model across epochs.*

The gradual reduction in box loss, along with a consistent increase in mAP@0.5, reflects stable training behavior and effective learning without notable overfitting. Class-wise classification performance of clothing attributes is further illustrated using a normalized confusion matrix, as shown in Fig. 2.

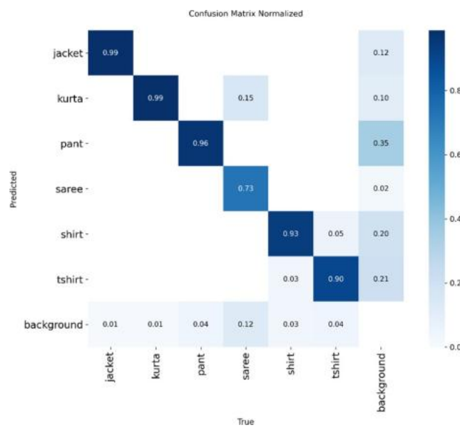

Fig. 2. *Normalized confusion matrix for clothing attribute classification.*

The strong diagonal dominance indicates high classification accuracy across most clothing categories, with minimal con- fusion among visually similar classes. Figure 3 illustrates the distribution of training samples across the different clothing categories included in the dataset.

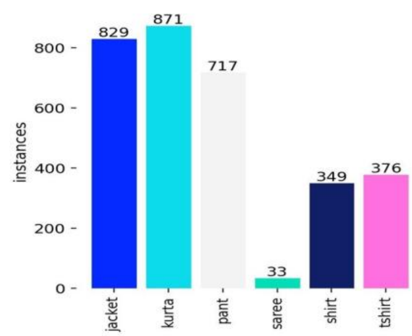

Fig. 3. *Distribution of training samples across clothing attribute classes.*

The imbalance in class distribution, particularly for the saree category, explains the relatively lower performance observed for this class in the confusion matrix.

Discussion

The findings show that integrating YOLOv8-based detection with clothing attribute recognition improves identification reliability in scenarios where facial recognition is ineffective. The system achieves high detection accuracy while maintaining real-time performance, making it appropriate for deployment in high-traffic public environments. Nonetheless, reduced performance was noted for underrepresented clothing categories, underscoring the need for greater dataset diversity.

Conclusion and Future Work

A real-time missing person detection system for railway stations based on YOLOv8 and clothing attribute recognition has been presented. The proposed approach addresses challenges associated with crowded environments and frequent occlusions by utilizing clothing-based visual cues. Experimental results confirm high detection accuracy along with real- time feasibility on railway surveillance data. Future research directions include expanding the dataset, incorporating multi- object tracking for cross-camera identification, and optimizing the system for edge-based deployment.

Acknowledgment

The authors would like to thank the Department of Electronics and Telecommunication Engineering, Bharati Vidyapeeth (Deemed to be University) College of Engineering, Pune, for providing the necessary facilities and support to carry out this research.

References

1. G. Jocher, A. Chaurasia, and J. Qiu, “YOLOv8 by Ultralytics,” 2023.
2. J. . Redmon, S. Divvala, R. Girshick and A. Farhadi, ”You Only Look Once: Unified, Real-Time Object Detection,” 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Las Vegas, NV, USA, 2016, pp. 779-788, doi: 10.1109/CVPR.2016.91.
3. S. Ren, K. He, R. Girshick and J. Sun, ”Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks,” in IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 39, no. 6, pp. 1137-1149, 1 June 2017, doi: 10.1109/TPAMI.2016.2577031.
4. Yutian Lin, Liang Zheng, Zhedong Zheng, Yu Wu, Zhilan Hu, Cheng- gang Yan, Yi Yang,Improving person re-identification by attribute and identity learning, Pattern Recognition, Volume 95, 2019, Pages 151-161, ISSN 0031-3203.
5. Gray Douglas & Tao Hai. (2008). Viewpoint Invariant Pedestrian Recognition with an Ensemble of Localized Features. 5302. 262-275. 10.1007/978-3-540-88682-2 21.
6. Zheng, Liang & Yang, Yi. (2016). Person Re-identification: Past, Present and Future. 10.48550/arXiv.1610.02984.
7. Roboflow, “Roboflow: Dataset Management and Annotation Tool,” 2023.
8. OpenCV Development Team, “Open-Source Computer Vision Library,” 2023.