© January 2026 | IJIRT | Volume 12 Issue 8 | ISSN: 2349-6002

A Deep Learning Framework for Person Re-Identification in Railway Surveillance Systems

A. Rishika[1], A. Sahithi[1], J. Shreshta[1], M. Aishwarya[1], Surya Narayana[2], Dr. S Shiva Prasad
[1]*Student, Department of CSE(DataScience), MallaReddy Engineering college, Secunderabad*
[2]*Assistant Professor, Department of CSE(DataScience), MallaReddy Engineering college, Secunderabad*
[3] *Professor, Department of CSE(DataScience), MallaReddy Engineering college, Secunderabad*

***Abstract*—The growing concern for public safety in densely populated transportation hubs such as Indian Railways highlights the need for advanced surveillance systems. This project focuses on developing a deep learning-based Person Re-Identification model that can identify, track individuals across multiple surveillance cameras. Traditional surveillance methods depend on handcrafted visual features, which often lack precision and robustness in real-world scenarios involving varying poses, illumination, and camera angles. These limitations hinder effective monitoring and timely detection of suspicious activities. To overcome these challenges, the proposed system utilizes Convolutional Neural Networks to extract rich facial and body features, enabling accurate and consistent identification of individuals across different views. This intelligent Re-ID system enhances security monitoring efficiency and can significantly aid in maintaining safety across public transportation environments.**

***Index Terms*—Deep learning, Convolutional Neural Networks, Support vector machine, Random Forest, Advanced surveillance systems.**

I. INTRODUCTION

Ensuring public safety in a densely populated country like India is a major concern, particularly in railway stations, which serve as critical transportation hubs. Monitoring and identifying suspicious individuals across these crowded environments pose significant challenges due to factors such as varying camera angles, lighting conditions, and changes in a person's appearance. Traditional surveillance systems that depend on manually extracted features often fail to provide consistent accuracy in such complex scenarios.
To address these limitations, this project introduces an intelligent Person Re-Identification (Re-ID) system based on Deep Learning techniques. The proposed system accurately identifies and tracks individuals across multiple CCTV camera feeds by extracting deep visual features using Convolutional Neural Networks (CNNs). These extracted features capture critical facial and pose information, which are then used by classification models such as Random Forest and Support Vector Machine (SVM) to detect and match suspicious persons with existing criminal databases.
The system is specifically designed to enhance surveillance within the Indian Railways network. It includes multiple functional modules such as:
Admin Panel: For uploading datasets, managing employee access, and training machine learning models.
Employee Panel: For monitoring real-time or recorded CCTV footage and identifying potential suspects.
Alert System: For generating and storing alerts when suspicious individuals are detected.
The backend of the system utilizes a Python-based web server and a MySQL database for efficient data storage and management. The integration of these components enables real-time video monitoring, automated alerts, and centralized database access, making the solution both scalable and effective for improving public safety and security in railway environments.

Figure.1 CCTV

II. LITERATURE SURVEY

Initial efforts in person re-identification leaned on manually designed feature extractors. Dalal and Triggs [1] (2005) pioneered the Histogram of Oriented Gradients (HOG) method for detecting humans, establishing it as a key tool for capturing human visuals. While it worked well in stable settings, these engineered features often faltered in dynamic surveillance contexts marked by cluttered scenes and inconsistent illumination. As machine learning progressed, scholars turned to metric learning strategies to boost re-identification precision. Zheng [2] (2015) delivered a landmark contribution with a massive benchmark dataset for person re-identification. This resource standardized testing protocols, spurring rapid advancements by enabling equitable assessments of various Re-ID models.

Deep learning's rise, driven by Convolutional Neural Networks (CNNs), transformed person re-identification. Krizhevsky [3] (2012) showcased CNNs' potential via ImageNet classification, paving the way for end-to-end feature extraction from unprocessed images. This change allowed Re-ID models to discover distinguishing traits automatically, moving beyond manual designs. To refine CNN capabilities, Szegedy [4] (2016) unveiled the Inception model, featuring multi-scale processing in one framework. These designs gained traction in Re-ID for adeptly handling both broad and fine-grained visual details.

Sun [5] (2018) advanced CNN baselines through Refined Part Pooling, focusing on segmented body representations. Their method tackled pose-induced misalignments and set new performance records across key Re-ID datasets. Crafting effective loss functions proved essential for Re-ID gains. Hermans [6] (2017) highlighted Triplet Loss's strengths, which pulls same-identity embeddings closer while pushing different-identity ones apart. This approach sharpened the separation in feature spaces for re-identification.

Building on part-focused techniques, Zhao [7] (2017) developed SpindleNet, incorporating pose estimates to parse the body into zones. By extracting tailored features per zone, it enhanced resilience to pose shifts and camera angles. Metric learning from adjacent areas has also shaped Re-ID. Liu [8] (2017) applied Deep Relative Distance Learning to vehicle re-identification, proving deep metrics excel at differentiating look-alikes. These ideas translate well to person Re-ID in dense settings like railway platforms. Classic machine learning classifiers still support Re-ID workflows. Pedregosa [9] (2011) launched Scikit-learn, a popular Python toolkit offering tools like Support Vector Machines (SVM) and Random Forests for post-extraction classification and matching.

III. PROPOSED METHODOLOGY

To overcome the limitations of traditional person re-identification systems, the proposed system integrates Deep Learning and Machine Learning techniques to achieve high accuracy, automation, and real-time monitoring. It leverages Convolutional Neural Networks (CNNs) to automatically extract discriminative facial and pose features from images of individuals. Unlike handcrafted feature-based methods, CNNs can learn complex visual representations, making the system more robust to changes in lighting, camera angles, and human appearance. The extracted deep features are then used to train Machine Learning models such as Random Forest and Support Vector Machine (SVM). Both models are evaluated for their performance, and the one achieving higher accuracy is selected for deployment. This hybrid training approach ensures reliable and optimized person identification across different surveillance feeds. By integrating a Python-based web interface with a MySQL database, the system supports efficient data storage, real-time monitoring, and centralized management. This deep learning-based solution enhances public safety by enabling automated, intelligent, and scalable surveillance within the Indian Railways environment.

Figure.2 Activity Diagram

IV. SYSTEM WORKFLOW

4.1.1 Data Collection
In this stage, the railway administrator uploads a dataset consisting of images of known suspicious individuals. These images form the reference database that the system uses for training and comparison during surveillance.

4.1.2 Feature Extraction
Once the dataset is uploaded, a Convolutional Neural Network (CNN) processes each image to extract meaningful facial and pose-related features. These deep features capture unique visual patterns that are more robust than traditional handcrafted features.

4.1.3 Model Training
The extracted feature vectors are used to train Machine Learning classifiers, namely Random Forest and Support Vector Machine (SVM). Both models learn to distinguish between different individuals based on the extracted features, and the model with higher accuracy is selected for deployment.

4.1.4 Video Processing
Authorized railway employees upload surveillance videos through the monitoring interface. The system processes the video frame by frame, detecting faces and extracting features using the trained CNN model.

4.1.5 Re-Identification
The features extracted from the video frames are compared with the stored feature database of suspicious individuals. If a match is found within a defined threshold, the person is identified as suspicious.

4.1.6 Alert Generation
When a suspicious individual is successfully re-identified, the system automatically generates an alert and records it in the database. These alerts can be reviewed by the administrator for further security action.

V. PROPOSED METHODS

This section explains the use of machine learning classifiers for identifying individuals based on features derived from Convolutional Neural Networks (CNNs). The deep features extracted by CNNs serve as inputs to supervised classifiers that aid in recognizing and flagging suspicious individuals in surveillance feeds.

5.1.1 Support Vector Machine (SVM)
The Support Vector Machine (SVM) is a supervised classification algorithm designed to find the optimal hyperplane that separates data points belonging to different categories within a high-dimensional feature space. In this project, the SVM model is trained using feature representations obtained from CNNs that capture distinctive facial and pose characteristics.
SVMs are known for their strong performance with high-dimensional data and their ability to generalize well when class boundaries are clearly defined. However, in scenarios involving noisy or highly complex surveillance data with illumination changes or occlusions, the performance may decline, and parameter tuning becomes critical for achieving accuracy.

5.1.2 Random Forest
The Random Forest algorithm operates as an ensemble learning method that builds multiple decision trees during training and combines their outputs through majority voting to reach a final decision. Each decision tree learns patterns from distinct subsets of the CNN-derived features, allowing the model to capture intricate, non-linear relationships within the data.
Due to its robustness, resistance to overfitting, and ability to handle large variations in real-world inputs, Random Forest has shown strong performance in surveillance applications, where environmental changes and varying viewpoints are common.

5.1.3 Monitoring Workflow
Video Upload: Authorized railway personnel upload surveillance videos through a secure monitoring interface. The system validates access to ensure only permitted users can submit surveillance data. Uploaded videos are queued for automated analysis without manual intervention.
Feature Extraction: The system processes video frames and applies a pre-trained CNN to extract facial and pose-based features. These features capture distinctive visual patterns required for reliable identification. Feature extraction is performed efficiently to support near real-time processing.
Feature Matching: Extracted features are compared

against a stored database containing profiles of known suspicious individuals. Similarity matching techniques are used to identify potential matches. This step helps narrow down candidates before final decision-making.
Classification: Based on the matching results and model predictions, each detected individual is classified as suspicious or non-suspicious. The classifier ensures high accuracy while minimizing false alarms. This supports timely and reliable decision-making.
Logging and Alerts: All detections, classifications, and alerts are securely logged in the database. Alerts are generated for suspicious detections and made available for administrative review. This ensures traceability and improves overall security monitoring efficiency.

VI. SYSTEM ARCHITECTURE

The proposed system architecture outlines how data flows between modules responsible for data collection, feature extraction, model training, and alert dissemination. It includes two primary user modules:
Admin Module: Oversees dataset management, model training, and database maintenance.
Employee Module: Conducts live monitoring and analysis of uploaded footage.

Figure.5 System Architecture

System Architecture Description:
Probe Image: The probe image refers to the input frame or image captured by the railway surveillance system that requires identification.
Template Gallery: The gallery contains pre-registered images of individuals of interest, uploaded by the administrator to act as reference data.
Descriptor Generation: Both probe and gallery images are passed through the CNN to extract distinctive feature descriptors representing facial and pose characteristics.
Matching Score Computation: Using the trained Random Forest or SVM classifier, similarity scores are computed between the probe features and every gallery entry.
Ranked List Creation: Based on the computed similarity scores, the gallery images are ranked in descending order of resemblance to the probe.
Decision and Alert Generation: If the highest similarity score surpasses a predefined threshold, the individual is classified as Suspicious, the system generates an alert else, the person is non-suspicious.
This architecture transforms visual inputs into discriminative feature vectors and identifies individuals through similarity-based ranking. It ensures accurate and efficient person re-identification while supporting automated alerting for enhanced railway security.

VII. RESULTS

The implemented system successfully detects and identifies individuals from surveillance video footage in railway environments. The SVM and Random Forest models were trained using the features extracted from CNNs. Their effectiveness was measured using performance metrics such as accuracy, precision, recall, F1-score. Experimental evaluations revealed that Random Forest consistently achieved superior results compared to SVM in terms of both accuracy and robustness. While SVM excelled when feature distributions were distinct, its effectiveness was reduced in real-world conditions with illumination and pose variations. Random Forest's ensemble structure allowed it to capture complex relationships, reduce overfitting, and maintain better generalization.

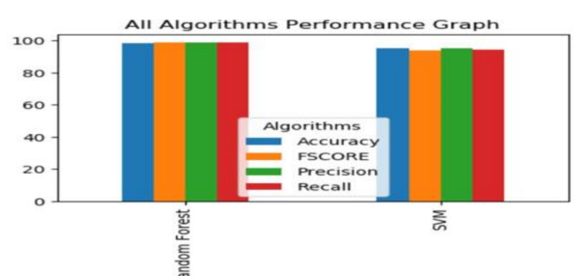

Figure.3 Algorithms performance graph of SVM, Random Forest

As a result, Random Forest was selected as the final model for deployment in the person re-identification system, ensuring more precise and reliable detection during live video monitoring.

Table.1 Performance Comparison of Classification Algorithms

Algorithm Name	Accuracy (%)	Precision (%)	Recall (%)	F1-Score (%)
SVM with CNN Features	95.082	95.143	94.167	94.017
Random Forest with CNN Features	98.361	98.571	98.750	98.564

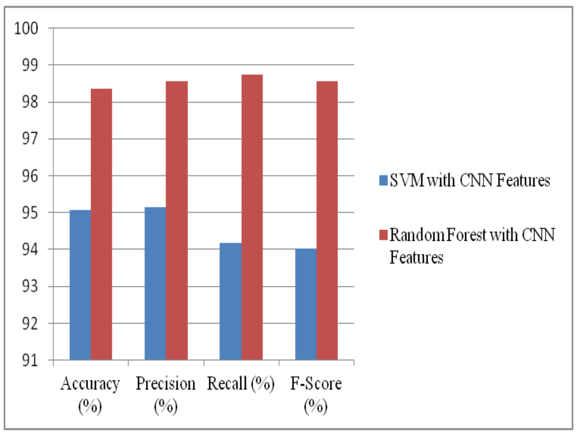

Figure.4 Performance Algorithms metrics

The bar graph compares the performance of SVM with CNN features and Random Forest with CNN features across accuracy, precision, recall, and F-score. The Random Forest model consistently outperforms the SVM model in all evaluation metrics, achieving higher accuracy and better balance between precision and recall. This indicates that Random Forest is more effective in learning discriminative patterns from CNN-extracted features and provides more robust classification performance. We implemented as web-based project. Upon uploading a video, the system processes each frame using deep learning techniques and matches detected faces with the database of known individuals. Based on this comparison, individuals are classified as “suspicious” and “non-suspicious”.

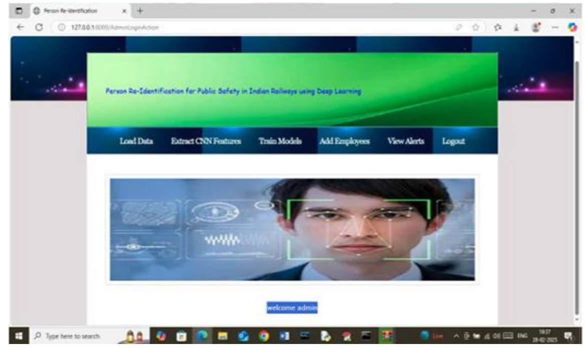

Figure.6 Admin clicks on 'Load data' to load any dataset

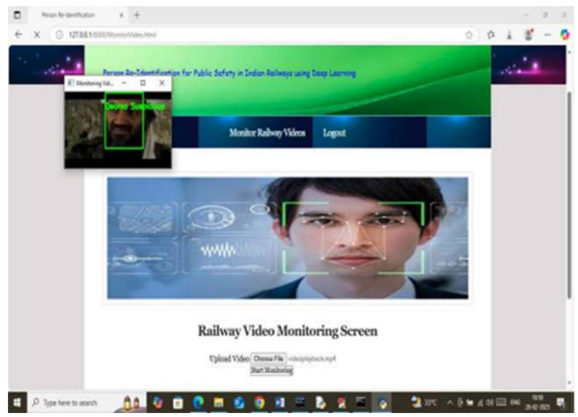

Figure.7 A "Suspicious Person" detected"

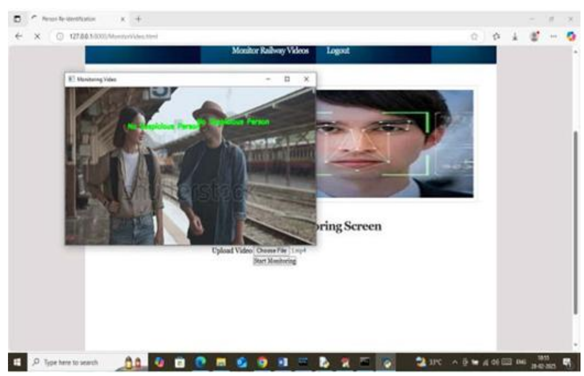

Figure.8 A “No Suspicious Person” detected

The obtained results show that the system performs effective face detection and recognition under varying lighting conditions and background environments commonly observed in railway stations.

When a suspicious individual present in the database is detected, an alert is automatically generated and the detection details are logged into the database. For other individuals, the system correctly identifies them as “non- suspicious,” resulting in minimal false alerts. These results indicate that the proposed Person Re-

Identification system can support railway surveillance by providing automated detection and alert generation, thereby assisting authorities in identifying potential threats efficiently.

The system analyzes the uploaded surveillance video and extracts facial and pose features using the CNN model. Since no significant match is found with the suspicious person database, the individual is correctly classified as non- suspicious, indicating normal activity and preventing unnecessary alerts.

VIII. CONCLUSION

In this, developed Person Re-Identification System for Public Safety in Indian Railways demonstrates the effectiveness of applying deep learning techniques for intelligent video surveillance. The project successfully replaces traditional handcrafted feature methods with Convolutional Neural Networks (CNNs), leading to a significant improvement in accuracy and reliability in identifying suspicious individuals across multiple camera feeds.

The integration of machine learning classifiers such as Random Forest and Support Vector Machine (SVM) has further enhanced recognition efficiency, with Random Forest achieving the highest accuracy during testing. The system's modular design incorporating separate Admin and Employee roles enables smooth operation, data management, and real-time monitoring.

Overall, the project proves to be a robust and scalable solution for improving safety in crowded railway stations, reducing manual monitoring efforts, and providing an automated mechanism for early detection of suspicious activities.

REFERENCES

[1] Dalal, N., & Triggs, B. (2005). Histograms of oriented gradients for human detection. Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 886–893. A classical method often compared with modern deep learning methods.

[2] Zheng, L. (2015). Scalable person re-identification: A benchmark. Proceedings of the IEEE International Conference on Computer Vision (ICCV), 1116–1124.https://doi.org/10.1109/ICCV.2015.133 Introduces a scalable benchmark for evaluating person re- identification models.

[3] Krizhevsky, A. (2012). ImageNet classification with deep convolutional neural networks. Advances in Neural Information Processing Systems, 25, 1097–1105. A foundational paper in deep learning with CNNs.

[4] Szegedy, C. (2016). Rethinking the inception architecture for computer vision. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2818–2826. One of the architectures used in CNN-based feature extraction.

[5] Sun, Y. (2018). Beyond part models: Person retrieval with refined part pooling (and a strong convolutional baseline). Proceedings of the European Conference on Computer Vision (ECCV), 480–496. Proposes a robust CNN baseline model with part-based feature learning for re-ID tasks.

[6] Hermans, A. (2017). In defense of the triplet loss for person re-identification. arXiv preprint arXiv:1703.07737. https://arxiv.org/abs/1703.07737 Demonstrates effectiveness of triplet loss in training deep networks for person re-ID.

[7] Zhao, H. (2017). SpindleNet: Person re-identification with human body region guided feature decomposition and fusion. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 1077–1085. Explores human pose-guided deep feature learning for re-ID.

[8] Liu, H. (2017). Deep relative distance learning: Tell the difference between similar vehicles. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2167–2175. Related approach in deep metric learning which can be adapted to person re-ID.

[9] Pedregosa, F. (2011). Scikit-learn: Machine Learning in Python. Journal of Machine Learning Research, 12, 2825–2830. https://jmlr.org/papers/v12/pedregosa11a.html Library that includes SVM and Random Forest used in your project.