



# International Journal of Innovative Research in Science Engineering and Technology (IJIRSET)

*(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)*



**Impact Factor: 8.699**

**Volume 14, Issue 7, July 2025**

# **Automatic Billing System for Public Urban Transportation Vehicles Using Computer Vision and Deep Learning**

**Rashida Farsath<sup>1</sup>, Hishal Shahad K<sup>2</sup>, Irfan Ahmed<sup>3</sup>, Mohammed Farseen K. K<sup>4</sup>,**

**Mohammed Hafnaz P<sup>5</sup>, Shanid Malayil<sup>6</sup>, Mubeena A. K<sup>7</sup>.**

Assistant Professor, Department of Computer Science & Engineering, MEA Engineering College, Kerala, India<sup>1</sup>

Student, Department of Computer Science & Engineering, MEA Engineering College, Kerala, India<sup>2,3,4,5</sup>

Assistant Professor, Department of Computer Science & Engineering, MEA Engineering College, Kerala, India<sup>6,7</sup>

**ABSTRACT:** The accuracy of fare calculation and passenger identification is crucial for the efficiency of public transportation systems. This article introduces an advanced camera-based system that automates passenger identification and fare estimation. The system utilizes OpenCV-Python for image processing, YOLOv3 for passenger detection, and dlib for robust face recognition. A combination of Convolutional Neural Networks (CNN) and feature extraction techniques enables precise boarding and alighting recognition. Passenger and fare details are securely stored in an SQL database, ensuring accurate transaction records. A web interface, developed using Django, HTML, CSS, and JavaScript, facilitates system management and real-time monitoring. Experimental evaluations on diverse datasets demonstrate the system's high accuracy, optimized memory usage, and improved processing speed. This scalable solution enhances fare collection mechanisms, strengthens passenger tracking, and ensures revenue integrity in public transportation networks.

**KEYWORDS:** Passenger Detection, Fare Estimation, Camera based Identification, Deep Learning, Computer Vision, Web based Monitoring.

## **I. INTRODUCTION**

Urban transportation systems serve as the backbone of daily commuting for millions, offering a sustainable and efficient mode of travel. However, accurate passenger identification and fare calculation remain significant challenges, often plagued by inefficiencies, inaccuracies, and revenue losses. Traditional fare collection methods, such as manual ticketing and fixed route tariffs, are labour-intensive and susceptible to errors and fraud. Advances in computer vision and deep learning present a transformative opportunity to automate passenger identification and enable real-time fare computation based on travel distance.

This paper proposes a camera-based detection and tracking system that enhances passenger identification, streamlines fare collection, and provides valuable insights for optimizing transit operations. The system employs OpenCV-Python for image processing, YOLOv3 for real-time passenger detection, and dlib for face recognition. CNN-based deep learning models are integrated to accurately track passenger movements, while an SQL database securely stores passenger and fare details. A web-based interface, developed using Django, HTML, CSS, and JavaScript, facilitates real-time system monitoring and management.

The main contributions of this paper include:

1. A novel passenger detection and tracking system utilizing YOLOv3 for real-time identification within the confined space of a bus.
2. Integration of dlib-based facial recognition to accurately identify passengers and enhance system security.

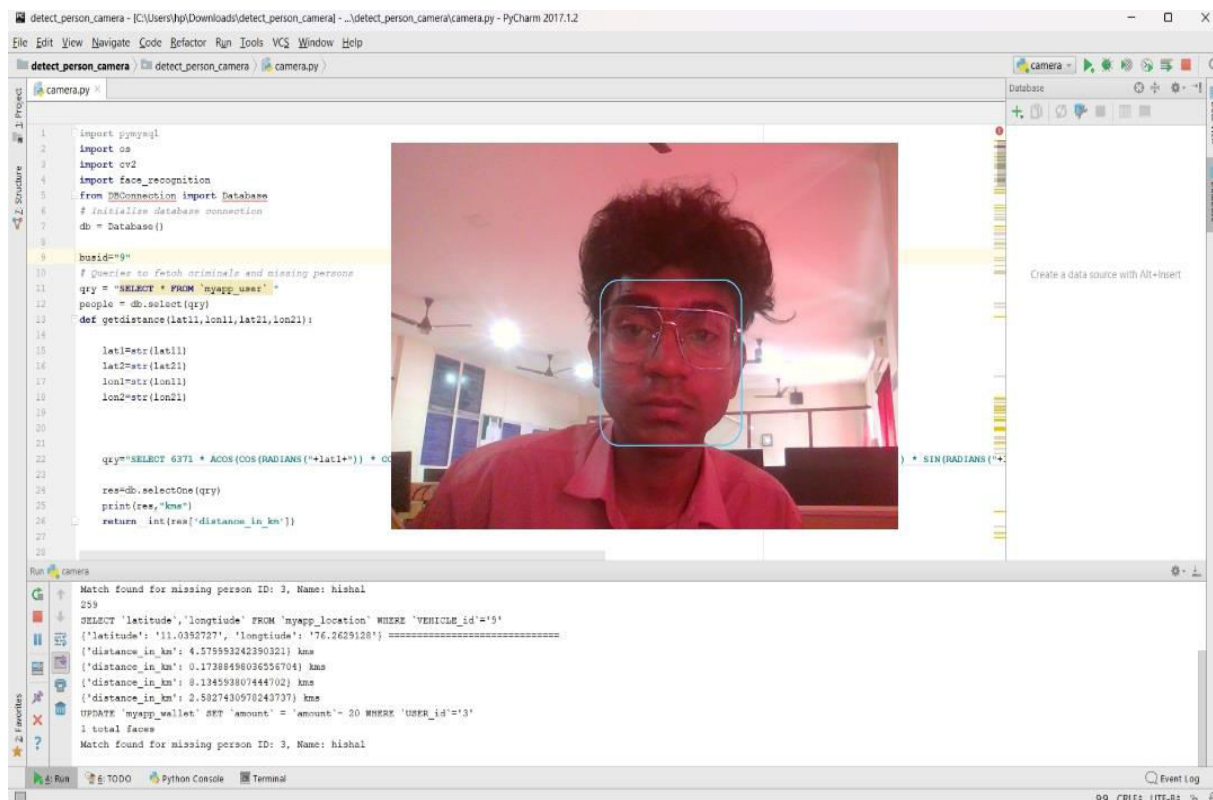


3. Implementation of an SQL-backed dynamic fare calculation module that computes fares based on actual travel distance, ensuring fairness and transparency.
4. Development of a web-based monitoring platform using Django, enabling operators to track and manage the system efficiently.
5. This work demonstrates a scalable and high-performance solution for modernizing fare collection and improving revenue integrity in public transportation systems.

## II. DESCRIPTION MODEL

The Passenger Identification and Fare Calculation System aims to bring together multiple advanced technologies to ensure accurate human detection, identification, tracking, and automated fare processing within public transportation environments. By leveraging tools like deep learning-based computer vision models and real-time data analytics, the system is capable of recognizing individual passengers, estimating their travel distance or time, and calculating fares accordingly. The framework also ensures a seamless billing process, potentially integrating with digital payment systems to provide a fully automated and contactless transit experience. Special emphasis is placed on capturing side-angle views, which are often missed in traditional detection setups, thereby increasing reliability across varied environmental and positional conditions.

Designed with real-time performance in mind, the architecture of the system ensures compatibility with existing public transport infrastructure, such as buses, metro systems, and even ferries. The design supports high-speed data processing, fast image recognition, and instantaneous response, allowing for quick and effective decision-making during peak transit hours. Scalability is another key aspect—future iterations could incorporate multi-camera inputs, behaviour analysis, and privacy-preserving techniques to strengthen both performance and ethical standards. The upcoming sections of this report will provide a detailed explanation of each functional component, outlining how they work individually and interact collectively to form a smart, efficient, and scalable solution for modern urban transit challenges.



### ***A. Structure of the System***

The system includes the following crucial modules:

- **Image Acquisition:** High-resolution cameras are strategically placed on the bus to capture facial images of passengers getting on and off. These cameras work nonstop, delivering a continuous flow of visual information for analysis.
- **Image Preprocessing:** Captured images undergo preprocessing steps to enhance their quality and prepare them for feature extraction. This includes resizing, normalization, noise reduction, and illumination correction.
- **Passenger Identification:** YOLOv3 and Faster R-CNN are used to detect and isolate passengers in real-time, ensuring accurate identification.
- **Feature Extraction:** Convolutional Neural Networks (CNN) and dlib perform feature extraction and face recognition to accurately categorize and recognize passengers.
- **Estimation of Boarding and Alighting Points:** The system monitors passenger movement to estimate boarding and exit points, which are crucial for fare calculation.
- **Fare Calculation:** The system calculates fares based on the distance between boarding and alighting points, ensuring accurate billing.
- **Storage and Control of Data:** A SQL database securely stores passenger details, boarding and alighting points, and fare records for real-time updates and auditing.

### ***B. Preprocessing of Images***

Effective image preprocessing improves detection and identification accuracy. The preprocessing steps include:

- **Resizing:** Standardizing image dimensions to optimize computational efficiency.
- **Normalization:** Adjusting pixel intensity values to a uniform scale for consistent lighting conditions.
- **Noise Reduction:** Applying filters such as Gaussian blur to remove unwanted noise.
- **Illumination Correction:** Using histogram equalization to enhance visibility under different lighting conditions.

### ***C. Passenger Detection***

Passenger detection is performed using:

- Background subtraction to distinguish passengers from the stationary background.
- YOLOv3 and Faster R-CNN to accurately detect passengers within frames and generate bounding boxes.

### ***D. Passenger Identification with CNN and dlib***

CNN and dlib are employed for face recognition, ensuring accurate passenger identification through:

- Optimized CNN architecture with convolutional layers, activation functions (ReLU), pooling layers, and fully connected layers.
- Training on a labelled dataset containing diverse passenger images.
- Classification using extracted facial features to match passengers accurately.

### ***E. Tracking and Alighting Point Estimation***

Passenger tracking ensures accurate journey logging using:

- Object tracking algorithms like Kalman Filters or SORT to maintain consistent identification across frames.
- Alighting point estimation based on movement patterns and last known positions before exiting.

***I. System Flow: Passenger Journey and Billing*****1) Passenger Boards**

- Camera captures entry image.
- YOLOv3 detects the passenger.
- dlib and CNN perform face recognition and emotion analysis.
- System registers passenger entry.

**2) Passenger Exits**

- Camera captures exit image.
- YOLOv3 detects the passenger's exit.
- dlib and CNN verify the face again.
- System records passenger's exit point.

**3) Fare Calculation**

- Distance travelled is computed as (Exit Point Boarding Point).
- System calculates fare based on distance.
- Fare is deducted from the passenger's wallet (SQL database).

**4) Data Update**

- Passenger's wallet is updated in the database.
- Journey and payment history are stored.

Admin dashboard displays updated vehicle load and fare collection.

**III. LITERATURE REVIEW**

[1] Vandit Gajjar et al. (2016) proposed a human detection method using HOG, Visual Saliency, and a Deep Multi-Level Network. It offers high accuracy and real-time performance but struggles with cluttered scenes, lighting changes, and has slow initial processing.

[2] Noppadol Chadil et al. (2008) developed "Goo-Tracking," a GPS-GPRS based real-time tracking system with Google Earth integration. It is cost-effective and easy to integrate but limited by GPRS coverage, hardware, and possible data delays.

[3] Bindu Sivakoti et al. (2023) introduced a CCTV-based people counting system using Haar Cascade and Dlib CNN. It is cost-efficient and user-friendly but limited to entrances and raises privacy issues.

[4] D. M. Gavrilu (1999) explored human activity recognition in computer vision, enabling intelligent interaction. While innovative, it lacked real-time application and was limited by outdated tech and small datasets.

[5] Navneet Dalal and Bill Triggs proposed a human detection method using HOG and SVM. It is robust to lighting changes and interpretable but sensitive to pose, occlusion, and has fixed window limitations.

[6] Bindu Sivakoti et al. (2023) presented a similar system to [3], combining CCTV, Haar Cascade, and Dlib CNN for face-based tracking. Its strengths and limitations mirror those in [3].

#### IV. FUTURE SCOPE

- └ **Multi-Camera Integration:** Using multiple cameras can enhance tracking accuracy and provide broader coverage across public spaces.
- └ **3D Human Detection Models:** Future work can integrate 3D models to improve accuracy, especially in side-angle or complex scenes.
- └ **Real-Time Performance Boost:** As hardware advances, further optimization could support faster, high-volume real-time processing without losing accuracy.
- └ **AI-Based Behaviour Detection:** Adding behaviour analysis can help detect unusual actions in crowded public areas, boosting security beyond ID tracking.
- └ **Support for All Transport Modes:** Expanding to trains, ferries, and subways offers a unified solution for fare and passenger tracking.
- └ **Mobile & Edge Compatibility:** Adapting the system for mobile and edge devices ensures efficient performance in low-resource or remote areas.
- └ **Privacy-Focused Identification:** Future work should include privacy-preserving algorithms for secure yet respectful public surveillance.

Future improvements may involve multi-camera integration and 3D detection to boost accuracy and coverage in public spaces. Enhancing real-time performance and incorporating AI-driven behaviour analysis can strengthen security measures. Expanding the system to include various transport modes like trains and subways increases its versatility. Additionally, making the platform compatible with mobile and edge devices, along with implementing privacy-preserving identification methods, ensures efficient operation in diverse settings while addressing ethical concerns around surveillance and data protection.

#### V. CONCLUSION

In conclusion, this project effectively demonstrates the development and implementation of a robust system for detecting and identifying humans from side-angle photographs by leveraging advanced deep learning techniques such as YOLO, Faster R-CNN, and Open Pose. These models allow for high-precision human detection and pose estimation, even under challenging conditions like varied lighting, occlusions, and complex backgrounds. The integration of these technologies supports the system's suitability for real-time applications, making it particularly valuable in domains that demand quick and accurate identification, such as public security, intelligent surveillance, and crowd management.

Furthermore, the system's ability to operate in real-time opens up practical applications in public transportation, especially in scenarios involving automated fare collection and passenger tracking in buses. This functionality not only enhances the operational efficiency of transit systems but also contributes to better service planning and safety monitoring.

The project lays a strong foundation for future developments, including expansion to multi-camera setups, behaviour analysis, and deployment on mobile or edge devices, ultimately aiming for broader adoption in smart city infrastructure and advanced surveillance ecosystems.

#### REFERENCES

- [1] M.Cornia, L. Baraldi, G. Serra and R. Cucchiara. A deep multi-level network for saliency prediction. In 23rd International Conference on Pattern Recognition (ICPR), 2016, pages 3488-3493. IEEE, 2016.
- [2] N.Dalal and B. Triggs. Histograms of oriented gradients for human detection. In (CVPR), 2005, pages 886-893.
- [3] N.Dalal, B. Triggs and C. Schmid. Human detection using oriented histograms of flow and appearance. In (ECCV), 2006, pages 428-441.

- [4] J.W.Davis and V. Sharma. Background-subtraction using contour-based fusion of thermal and visible imagery. In Computer Vision and Image Understanding 106, no. 2 (2007): 162-182.
- [5] J.A.Hartigan and M. A. Wong. Algorithm AS 136: A k-means clustering algorithm. In Journal of the Royal Statistical Society. Series C (Applied Statistics) 28, no. 1 (1979): 100-108.
- [6] T.Judd, K. Ehinger, F. Durand and A. Torralba. Learning to predict where humans look. In (ICCV), 2009, pages 2106-2113.
- [7] K.Mikolajczyk, C. Schmid and A. Zisserman. Human detection based on a probabilistic assembly of robust part detectors. In (ECCV), 2004, pages 69-82.
- [8] P.Felzenszwalb and D. Huttenlocher. Efficient matching of pictorial structures. CVPR, Hilton Head Island, South Carolina, USA, pages 66–75, 2000.
- [9] W.T.Freeman and M.Roth. Orientation histograms for hand gesture recognition. Intl. Workshop on Automatic Face and Gesture- Recognition, IEEE Computer Society, Zurich, Switzerland, pages 296–301, June 1995.
- [10] W.T.Freeman, K.Tanaka, J. Ohta, and K. Kyuma. Computer vision for computer games. 2nd International Conference on Automatic Face and Gesture Recognition, Killington, VT, USA, pages 100–105, October 1996.





INTERNATIONAL  
STANDARD  
SERIAL  
NUMBER  
INDIA



# INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN SCIENCE | ENGINEERING | TECHNOLOGY

 9940 572 462  6381 907 438  [ijirset@gmail.com](mailto:ijirset@gmail.com)



[www.ijirset.com](http://www.ijirset.com)

Scan to save the contact details