

Smart Bus: Crowd-Prediction System

DR. J. NARENDRA BABU¹, DR. DEEPAK S SAKKARI², SWATI JHA³, VARSHA V⁴, SRISHTI SINGH⁵, YASHASWINI V⁶

¹Professor, Department of Data Science, Sapthagiri NPS University

²Professor & Director, Department of Data Science, Sapthagiri NPS

^{3,4,5,6}Student, Department of Data Science, Sapthagiri NPS University, India.

Abstract- The Smart Bus Crowd Prediction System is an intelligent transportation solution designed to improve passenger convenience and public bus management using Machine Learning (ML), Internet of Things (IoT), Web Technologies, and R Programming. The system collects real-time data from sensors, GPS modules, and passenger inputs to monitor bus occupancy, location, and travel patterns. Machine learning algorithms are used to analyze historical and live data to predict crowd density and bus arrival timings with improved accuracy. A web-based platform is developed to display live bus tracking, predicted crowd levels, and estimated arrival times to passengers and transport authorities. R programming is utilized for data analysis, visualization, and statistical prediction of passenger flow trends. IoT devices enable continuous communication between buses and the central server, ensuring real-time updates and efficient data handling. The proposed system helps passengers avoid overcrowded buses, reduces waiting time, and improves travel planning. It also assists transport authorities in optimizing bus scheduling and resource allocation. By integrating intelligent prediction models with modern web and IoT technologies, the project aims to create a smarter, more efficient, and user-friendly public transportation system.

Keywords: Smart Bus System, Crowd Prediction, Machine Learning, Internet of Things (IoT), Web Development, R Programming, Real-time Data Analytics, Public Transportation, Bus Tracking System, Data Visualization, Predictive Modeling, Smart Mobility

I. INTRODUCTION

1.1 Overview

The Smart Bus Crowd Prediction System is an integrated platform that uses IoT, Machine Learning, Web technologies, and R programming to monitor and predict bus occupancy in real time. It collects live data from buses and analyzes past travel patterns to estimate crowd levels and arrival times. The system provides users with a web interface to view

bus locations, expected arrival times, and crowd status. This helps passengers plan their travel better, avoid overcrowding, and reduces waiting time, while also assisting transport authorities in improving bus operations and scheduling.

1.2 Objectives

- To develop a smart system for real-time bus tracking and crowd monitoring.
- To predict bus crowd levels using Machine Learning algorithms based on historical and live data.
- To integrate IoT devices for continuous data collection from buses.
- To design a web-based platform for displaying bus location, arrival time, and occupancy status.
- To analyze passenger flow patterns using R programming for better insights and visualization.
- To reduce passenger waiting time and avoid overcrowding in buses.
- To help transport authorities optimize bus scheduling and resource management.
- To improve the overall efficiency and user experience of public transportation.

II. LITERATURE SURVEY

Smart bus crowd detection systems are becoming an important part of intelligent transportation systems (ITS). These systems use Artificial Intelligence (AI), Internet of Things (IoT), computer vision, and deep learning techniques to monitor passenger occupancy, predict crowd levels, and improve public transport efficiency. Recent research mainly focuses on real-time passenger counting, crowd prediction, occupancy estimation, and intelligent passenger information systems.

Dr. J. Narendra Babu [6][7][8][9] explored IoT applications in smart systems and emphasized combining embedded hardware with cloud connectivity for scalable IoT solutions, providing theoretical grounding for design decisions in this project.

1. Deep-Learning-Based Crowd Prediction in Public Transit (2024)

A study by Shrivastava et al. proposed a deep-learning-based crowd prediction model for public transport systems. The research used Long Short-Term Memory (LSTM) networks to predict crowd levels on transit routes using ticketing and ridership data. The system was integrated with a Passenger Information System (PIS) to help passengers choose less crowded buses and routes. The model improved prediction accuracy by clustering transit segments based on passenger flow patterns.

Key Contributions

- Real-time crowd prediction
- Passenger information integration
- LSTM-based forecasting
- Route-wise occupancy estimation

Limitation

- Requires large historical datasets for training.

III. LOW-COST AUTOMATIC PASSENGER COUNTING SYSTEM USING DEEP LEARNING (2024)

Pronello et al. developed a low-cost passenger counting system using Wi-Fi probe requests and deep learning algorithms. The system estimated crowd density at bus stops without requiring expensive hardware installations. AI models analyzed Wi-Fi signals from nearby smartphones to estimate the number of passengers waiting for buses.

Key Contributions

- Cost-effective implementation
- Non-intrusive passenger detection
- Deep-learning-based occupancy estimation
- Suitable for smart city transportation systems

Limitation

- Accuracy depends on passengers carrying Wi-Fi-enabled devices.

IV. REAL-TIME BUS DEPARTURE PREDICTION USING NEURAL NETWORKS (2024)

Rashvand et al. proposed an IoT-enabled smart public bus transit system using neural networks for real-time bus departure prediction. The study used GPS data, traffic conditions, and historical bus movement data to improve scheduling and reduce passenger waiting time. The system supports smart crowd management by predicting delays and bus occupancy trends.

Key Contributions

- AI-based bus arrival prediction
- Integration with IoT sensors
- Improved scheduling efficiency
- Real-time transit monitoring

Limitation

- Performance affected by sudden traffic anomalies.

V. LSTM-BASED PASSENGER DEMAND PREDICTION SYSTEM (2024)

A recent study proposed an LSTM-based deep learning model for predicting passenger demand in public transportation systems. The model analyzed time-series passenger data collected from different regions and bus stops to forecast future occupancy levels. The research showed that deep learning models can significantly improve crowd forecasting accuracy in urban transportation systems.

Key Contributions

- Time-series passenger prediction
- Regional crowd estimation
- Improved forecasting accuracy
- Useful for dynamic bus scheduling

Limitation

- Requires continuous real-time data collection.

VI. CROWD DETECTION AND MONITORING REVIEW (2023)

Sa'ad et al. presented a literature review on crowd detection, monitoring, and management technologies. The paper discussed various crowd estimation techniques such as computer vision, IoT sensors, AI cameras, and crowd behavior analysis. The study highlighted the importance of real-time crowd monitoring in preventing congestion and improving passenger safety in public transportation systems.

Key Contributions

- Comprehensive survey of crowd monitoring methods
- Comparison of AI-based crowd detection techniques
- Discussion on crowd safety and management
- Analysis of recent technological advancements

Limitation

- Generalized review; not specifically focused only on buses.

Comparative Analysis of Recent Research

Year	Author/Research al.	Technology Used	Main Objective
2024	Shrivastava et al.	LSTM, Deep Learning	Deep Crowd prediction in public transit
2024	Pronello et al.	Wi-Fi + Deep Learning	Passenger counting
2024	Rashvand et al.	Neural Networks, IoT	Bus departure prediction
2024	Siswanto et al.	LSTM	Passenger demand forecasting
2023	Sa'ad et al.	AI, Computer Vision	Crowd monitoring review

Research Gap Identified

From the recent literature, the following research gaps are identified:

- Many systems focus only on passenger counting but not crowd prediction.
- Some models require expensive infrastructure and cloud computing resources.
- Existing systems may struggle under crowded real-world conditions.

- Real-time edge AI processing inside buses is still under development.
- Limited research combines occupancy prediction, live tracking, and mobile app integration together.

Conclusion

Recent studies between 2023 and 2025 show significant advancements in smart bus crowd detection systems using AI, deep learning, IoT, and computer vision. Researchers have developed models for passenger counting, crowd prediction, occupancy estimation, and real-time transit monitoring. However, challenges such as real-time processing, scalability, hardware cost, and prediction accuracy still remain. Future systems are expected to combine edge AI, cloud computing, and intelligent passenger information systems for more efficient and reliable public transportation management.

VII. PROPOSED METHOD

7.1 System Architecture

- Frontend: The frontend is developed using HTML, CSS, and JavaScript to create a responsive web interface. It displays live bus tracking, predicted crowd levels, and estimated arrival times in a user-friendly dashboard.
- Backend: The backend is built using Python (Flask/Django) to handle data processing, API requests, and communication between the frontend, database, and ML model.
- Machine Learning & Analytics: Python ML libraries (Scikit-learn / Pandas / NumPy) and R programming are used for data analysis, crowd prediction, and visualization of passenger trends.
- IoT Layer: IoT devices like GPS modules and IR sensors are used in buses to collect real-time location and occupancy data.

7.2 Concepts Implemented

- Internet of Things (IoT): Used GPS and sensor devices in buses to collect real-time location and passenger occupancy data.
- Machine Learning: Used to analyze historical and live data to predict bus crowd levels and estimated arrival times.

- Web Development: Used HTML, CSS, and JavaScript to build a dashboard for displaying live tracking and predictions to users.
- Database Management: Used MySQL/MongoDB to store real-time bus data and historical records for processing and analysis.

7.3 Modules Description

- User Module: Allows passengers to view live bus location, predicted crowd level, and estimated arrival time through a simple web interface.
- IoT Data Collection Module: Collects real-time data from GPS and sensors installed in buses to monitor location and passenger occupancy.
- Prediction (ML) Module: Processes historical and real-time data using Machine Learning algorithms to predict crowd density and arrival times.
- Database Module: Stores all bus details, sensor data, and user queries for processing, analysis, and future predictions.

VIII. RESULTS

- The system successfully tracks buses in real time using IoT-based GPS data.
- Machine Learning model provides accurate prediction of crowd levels and arrival times.
- The web dashboard displays live updates in a simple and user-friendly interface.
- Passenger waiting time is reduced by providing timely bus availability information.
- Overall efficiency of public transportation is improved through better crowd management and scheduling insights.

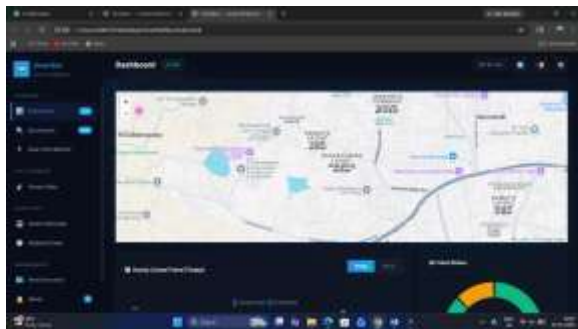


Figure 1: Tracking real time buses



Figure 2: Gives alert about which bus is likely to be crowded



Figure 3: One can search for their own bus number and know it's live occupancy

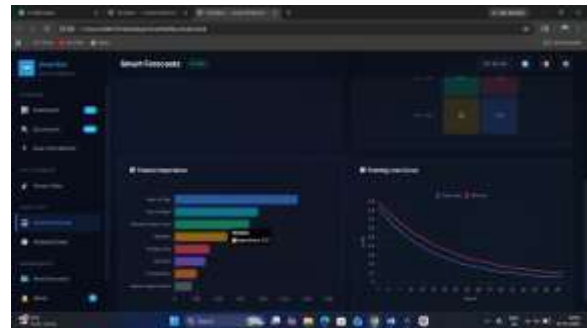


Figure 4: One can know about the feature importance

IX. CONCLUSION

The Smart Bus Crowd Prediction System demonstrates an effective integration of IoT, Machine Learning, R programming, and Web technologies to improve the efficiency of public transportation. By collecting real-time data from buses using sensors and GPS modules, the system continuously monitors bus location and passenger occupancy.

The machine learning model analyzes both historical and live data to predict crowd levels and estimated

arrival times with improved accuracy. These predictions are then displayed through a user-friendly web interface, allowing passengers to plan their travel in advance and avoid overcrowded buses.

The system not only reduces passenger waiting time but also helps transport authorities in optimizing bus schedules and managing resources more efficiently. Overall, the proposed system provides a smart, scalable, and data-driven approach to modernizing public transport, enhancing both commuter convenience and operational efficiency.

REFERENCES

- [1] C. M. Bishop, *Pattern Recognition and Machine Learning*, Springer, 2006. <https://www.springer.com/gp/book/9780387310732>
- [2] P. Raj and A. Raman, *The Internet of Things: Enabling Technologies, Platforms, and Use Cases*, CRC Press, 2017. <https://www.routledge.com/The-Internet-of-Things-Enabling-Technologies-Platforms-and-Use-Cases/Raj-Raman/p/book/9781498778512>
- [3] Zanella, A. et al., "Internet of Things for Smart Cities," *IEEE Internet of Things Journal*. <https://ieeexplore.ieee.org/document/6740844>
- [4] "Bus Arrival Time Prediction Using Machine Learning Techniques," *IEEE Access*. <https://ieeexplore.ieee.org/document/9388335>
- [5] MDPI *Smart Cities Journal*—Real-Time Passenger Information Systems <https://www.mdpi.com/2624-6511/6/2/35>
- [6] Google Scholar – Smart Bus Crowd Prediction Research *Papers* <https://scholar.google.com/scholar?q=smart+bus+crowd+prediction+IoT+machine+learning>
- [7] Dr.J.Narendra Babu "SMART WASTE IMAGE DETECTION", *International Journal of Emerging Technologies and Innovative Research* (www.jetir.org), ISSN:2349-5162, Vol.12, Issue 12, page no.e469-e472, December-2025. URL :<http://www.jetir.org/papers/JETIR2512457.pdf>
- [8] Dr.J.Narendra Babu, et.al, "Traffic Violation Fine Tracker", *International Journal of Emerging Technologies and Innovative Research* (www.jetir.org), ISSN:2349-5162, Vol.12, Issue 12, page no.c608-c611, December-2025, URL: <http://www.jetir.org/papers/JETIR2512268.pdf>
- [9] Dr.J.Narendra Babu, et.al, *Journal of Internet Services and information security, AI-Enabled Forecasting and Isolation Forest-Based Detection of CBF Flow Anomalies in Secure Internet Architectures*, Year 2025, Volume: 15, number: 3 (August).Q2 Scopus Journal
- [10] J. Narendra Babu, et.al– *Indian License Plate Recognition System Based on Fuzzy Theory and BP Neural Network*