IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

Impact Factor 8.471 ⁜ Peer-reviewed & Refereed journal ⁜ Vol. 15, Issue 5, May 2026

DOI: 10.17148/IJARCCE.2026.155250

AI-Based Smart Bus Live Tracking System Using IoT and Machine Learning: A Survey of Real-Time ETA Prediction and Transportation Analytics Approaches

Jayashubha J.G[1], Muppuri Dhatri[2], Raavi Pranitha[3], Rayapati Supriya[4]

Associate Prof., Dept of CSE, K.S. School of Engineering & Management, Bengaluru, India[1]

Student, Dept of CSE, K.S. School of Engineering & Management, Bengaluru, India[2-4]

Abstract: Public transportation systems in urban areas suffer from persistent inefficiencies rooted in static schedule management, absent live vehicle visibility, and reactive fleet operations. Passengers endure uncertainty regarding bus arrival times, leading to excessive stop-level waiting and reduced confidence in public transit. This paper surveys existing research on IoT-enabled bus tracking, AI-based estimated time of arrival (ETA) prediction, GPS-based fleet monitoring, and machine learning applied to transportation analytics. Six representative studies from 2022–2026 are analyzed and compared across methodology, hardware configuration, AI algorithm, cloud platform, and key limitations. Based on this survey, we propose a comprehensive AI-Based Smart Bus Live Tracking System integrating ESP32 microcontrollers, GPS modules, cloud-synchronized databases, machine learning-based ETA prediction, and a cross-platform Flutter passenger application. The proposed system addresses critical gaps in existing approaches by combining real-time GPS tracking, multi-parameter AI prediction, and a centralized analytics dashboard within a single deployable platform.

Keywords: IoT, GPS, ESP32, ETA Prediction, Machine Learning, Smart Transportation, Real-Time Tracking, Cloud Computing, Flutter, Transportation Analytics, LSTM, Random Forest, Firebase, Smart City

I. INTRODUCTION

The transportation sector accounts for a large share of urban congestion, fuel consumption, and passenger dissatisfaction in rapidly expanding cities. Public bus systems, despite their capacity advantages over private vehicles, are undermined by a fundamental information gap: passengers at stops have no reliable way to know whether their bus is approaching, delayed, or has already passed. This uncertainty drives over-early arrival behavior, collectively producing millions of hours of unproductive waiting time across metropolitan transit networks each day.

Traditional bus monitoring systems rely on fixed printed timetables that have little connection to actual arrival patterns under variable traffic conditions. Some modern deployments offer basic GPS tracking through municipal portals, but these systems typically present raw coordinate data without predictive intelligence, lack a mobile-accessible interface, and do not account for dynamic traffic variables in arrival time computation. The result is a system that reports location but cannot answer the passenger's core question: when will my bus arrive?

Recent advances in IoT hardware, cloud computing, and machine learning present a genuine opportunity to address these deficiencies at costs accessible to educational institutions and public transport operators. Affordable microcontrollers such as the ESP32 provide Wi-Fi and Bluetooth connectivity alongside sufficient processing capacity for edge GPS data handling. Managed cloud platforms including Firebase offer real-time database propagation without infrastructure overhead. Machine learning libraries in Python enable training of temporal prediction models on historical trajectory data, producing ETA estimates that adapt to congestion patterns rather than relying on static assumptions.

Existing research has explored IoT-based bus tracking and AI-driven arrival prediction as separate domains. Systems implementing GPS hardware tracking rarely incorporate predictive ML components, while studies benchmarking ML models for ETA typically operate on offline datasets without integration into live infrastructure. A unified, deployable system combining both capabilities in a consistent architecture has not been systematically documented in the reviewed literature.

IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

Impact Factor 8.471 ⸎ Peer-reviewed & Refereed journal ⸎ Vol. 15, Issue 5, May 2026

DOI: 10.17148/IJARCCE.2026.155250

This paper surveys six representative studies from 2022–2026 on IoT vehicle tracking, AI-based transportation prediction, and smart mobility platforms. We analyze their methodologies, hardware configurations, AI approaches, and limitations, then present the design of a proposed AI-Based Smart Bus Live Tracking System (ABSBLT).

A. Research Contributions

This survey makes the following contributions:

- A structured comparative analysis of six recent AI-based IoT tracking and transportation prediction studies.
- Identification of critical gaps related to unified architecture, AI-based ETA prediction, and passenger-facing mobile interfaces.
- Proposal of a unified IoT and AI-based Smart Bus Tracking System incorporating GPS telemetry, ML-based ETA prediction, cloud synchronization, and a Flutter passenger application.

II. RELEVANT LITERATURE

A. Paper 1: Bus Odyssey – Intelligent College Bus Tracking

Neetha K et al. [1] developed a real-time college bus tracking system using a NEO-7M GPS module with an ESP32 microcontroller. Coordinates were uploaded to a Supabase cloud backend and visualized through a Flutter application layered over OpenStreetMap and Mapbox APIs. The system used geofence-based push notifications triggered at pre-registered coordinates to alert students when their bus approached a stop. The deployment was effective for low-cost hardware and passenger notification, but ETA computation relied on a naive distance-divided-by-speed formula with no machine learning. Cloud dependency created a single point of failure when mobile data coverage was absent, and GPS accuracy degraded near dense overhead obstructions.

B. Paper 2: AI-Based Timetable Generation and Vehicle Arrival Prediction

Kumar et al. [2] proposed an automated vehicle arrival time prediction framework combining computer vision with statistical regression. A YOLOv5 model processed overhead camera feeds at terminal entry points to detect vehicles, with detected timestamps forming input to a Linear Regression predictor. The vision-based approach avoided hardware modification to vehicles. However, prediction accuracy degraded when unusual traffic events disrupted historical distributions. The system operated only at terminal boundaries and could not generate en-route predictions for passengers at intermediate stops, and no mobile interface was provided.

C. Paper 3: Predictive Platform for Bus Mobility and Human Flow Analysis

Tejashwini N et al. [3] developed an AI-driven bus mobility platform incorporating LSTM-based ETA prediction, Random Forest crowd density estimation, and WebSocket-synchronized real-time updates. GPS coordinates were streamed from buses to Firebase via MQTT. The LSTM network processed 60-timestep sliding windows of speed-location sequences to generate stop-specific ETA estimates. The platform demonstrated high accuracy and crowd analysis capabilities. However, the LSTM model required at least six months of annotated historical data, making it unsuitable for new or low-frequency routes. Crowd estimation did not generalize well to event-driven demand spikes, and no offline fallback was implemented.

D. Paper 4: Smart College Bus Tracker – Real-Time Location and ETA Predictor

Sowndharya et al. [4] developed a mobile-first college bus tracking application using ESP32, Firebase Realtime Database, Google Maps API, and Firebase Authentication for role-based access. ETA was computed by dividing remaining route distance by current GPS-derived speed, updated at five-second intervals. The application provided effective real-time position monitoring and secure multi-role access. ETA accuracy was limited by the use of instantaneous speed as the only prediction variable, with no account for upcoming traffic conditions. Google Maps API introduced per-request billing costs that escalate at deployment scale, and no fallback mechanism was provided for GPS signal loss.

E. Paper 5: GPS and IoT-Based Smart Public Transportation System

Moumen et al. [5] examined a scalable IoT architecture for municipal public transportation monitoring using GPS-equipped fleet vehicles, SIM7600 cellular modems transmitting via MQTT to AWS IoT Core, Lambda-based stream processing, and DynamoDB storage. Geofence proximity events triggered both mobile push and SMS alerts, extending accessibility to passengers without smartphones. MQTT reduced cellular data overhead efficiently. However, the system implemented no predictive intelligence; ETA was inferred from geofence proximity rather than temporal forecasting. AWS infrastructure costs scaled linearly with fleet size, and no historical analytics were included.

IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

Impact Factor 8.471 ⪤ Peer-reviewed & Refereed journal ⪤ Vol. 15, Issue 5, May 2026

DOI: 10.17148/IJARCCE.2026.155250

F. Paper 6: Kalman Filter and Fuzzy Logic for Bus Trajectory Smoothing

Rajan & Chandrasekaran [6] developed an embedded signal processing framework applying Extended Kalman Filtering to raw GPS coordinate streams for trajectory noise reduction, combined with a Takagi-Sugeno fuzzy inference engine that modulated speed estimates using linguistic traffic condition descriptors. The EKF corrected position jitter using a constant-velocity motion model, while the fuzzy system classified traffic conditions into four severity levels that adjusted the baseline speed estimate for ETA computation. The EKF operated within ESP32-class hardware constraints without cloud dependency, and the fuzzy logic provided interpretable inference. However, fuzzy parameters required domain expert calibration per route, limiting scalability. No deep learning component was present, restricting adaptability to novel traffic patterns.

III. COMPARATIVE ANALYSIS

Table I presents a structured comparison of the six reviewed studies across critical dimensions relevant to smart bus tracking and ETA prediction

TABLE I. COMPARISON OF EXISTING RESEARCH ON IOT BUS TRACKING AND ETA PREDICTION

Ref.	Year	Method / Algorithm	Hardware / Platform	Main Advantage	Limitation
[1]	2026	LSTM (60-timestep) + Random Forest crowd density	Firebase, MQTT, WebSocket, Python ML backend	High ETA accuracy; crowd density estimation; real-time WebSocket updates	Requires 6+ months training data; poor generalization to event-driven demand; no offline fallback
[2]	2025	Distance ÷ Speed (naive ETA)	ESP32, NEO-7M GPS, Supabase, Flutter, OpenStreetMap	Low hardware cost; geofence notifications; functional Flutter UI	No ML prediction; cloud-only; single route tested; GPS degraded in obstructed areas
[3]	2025	Instantaneous Speed-based ETA (5-second refresh)	ESP32, Firebase Realtime DB, Google Maps API, Firebase Auth	Role-based access control; 5-second position refresh; mobile-first design	ETA uses only instantaneous speed; Google Maps billing at scale; no GPS fallback mechanism
[4]	2023	YOLOv5 + Linear Regression	GPU server, overhead cameras, Python backend	No vehicle hardware modification required; automated timetable generation	Terminal-only prediction; no intermediate-stop ETA; poor performance in low light; no mobile UI
[5]	2023	Geofence proximity (no ML prediction)	SIM7600 modem, AWS IoT Core, Lambda, DynamoDB, SMS gateway	Dual-channel alerts (push + SMS); MQTT reduces data overhead; scalable cloud architecture	No predictive ETA; AWS cost scales with fleet size; no historical analytics
[6]	2022	Extended Kalman Filter + Takagi-Sugeno Fuzzy Logic	ESP32-class MCU, embedded firmware, no cloud required	Real-time edge processing; no cloud dependency; interpretable fuzzy inference	Fuzzy parameters need per-route expert calibration; no deep learning; non-Gaussian noise violated EKF assumptions

IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

Impact Factor 8.471 ⁜ Peer-reviewed & Refereed journal ⁜ Vol. 15, Issue 5, May 2026

DOI: 10.17148/IJARCCE.2026.155250

IV. GAP ANALYSIS

Based on the review of the six studies, the following critical gaps are identified:

A. Absence of Unified Integrated Platforms

Existing research treats GPS hardware tracking and AI-based arrival prediction as independent domains. Tracking systems implement no predictive intelligence, while ML studies operate on offline datasets without live integration. No single deployable system combines ESP32 GPS hardware, real-time cloud synchronization, ML-based ETA prediction, and a passenger mobile interface in a consistent architecture accessible to educational institutions and small operators.

B. Limited Multi-Parameter ETA Models

Most ETA implementations use a single-parameter formula dividing remaining distance by current instantaneous speed. Variables including historical segment congestion, time-of-day traffic patterns, weather conditions, and stop dwell time are rarely incorporated simultaneously. This restricts prediction accuracy precisely during peak-congestion conditions when reliability matters most.

C. Lack of Passenger-Centric Mobile Interfaces

Several surveyed systems provide administrator dashboards or raw data APIs without designing for the primary beneficiary: the waiting passenger. No reviewed system simultaneously offers live map visualization, AI-based ETA countdown, geofence proximity notifications, and offline-graceful degradation within a production-quality cross-platform mobile application.

D. No Offline or Communication-Resilient Operation

All reviewed cloud-dependent systems create complete service failure when internet connectivity is interrupted. No surveyed work implements local edge buffering with automatic cloud resynchronization upon reconnection, or GSM fallback communication as an alternative to Wi-Fi primary transmission. This leaves systems non-functional in suburban and peri-urban corridors where tracking value is highest.

E. Absence of Fleet-Level Analytics

Existing systems report individual vehicle positions without generating aggregate operational intelligence. Route adherence scoring, delay propagation analysis, segment-level congestion mapping, and longitudinal service quality reporting are absent from all reviewed approaches, limiting utility for transport administrators beyond real-time position awareness.

V. PROPOSED SYSTEM DESIGN

A. System Overview

To address the identified gaps, we propose an AI-Based Smart Bus Live Tracking System (ABSBLT) structured in four integrated layers: Edge Acquisition, Communication, Cloud and Intelligence, and Presentation. The system accepts live GPS telemetry from ESP32-mounted vehicle units and delivers ETA predictions, live position maps, and fleet analytics through cloud-synchronized passenger and administrator interfaces. The core design principle is a software-hardware co-design that is cost-accessible, deployable without specialist infrastructure, and architecturally extensible.

B. System Workflow

The end-to-end workflow follows a structured data pipeline:

- NEO-7M GPS module streams NMEA coordinate fixes to the ESP32 at five-second intervals; the ESP32 firmware parses and applies on-chip Kalman filtering before JSON serialization.
- Primary transmission via Wi-Fi to Firebase Realtime Database; automatic fallback to GSM/GPRS via SIM800L module when Wi-Fi signal drops below threshold; local MicroSD buffer captures fixes during connectivity gaps.
- Cloud Function processes each incoming telemetry event, appends to historical trajectory table, and invokes the ML ETA prediction API with the current feature vector.
- ETA result and live position are written to the real-time database node; Flutter passenger app and React admin dashboard receive sub-200 ms updates via WebSocket listeners.
- Geofence proximity events trigger Firebase Cloud Messaging push notifications to passengers registered for the approaching bus route.
- Historical trajectory and delay records accumulate for periodic ML model retraining and route analytics report generation.

DOI: 10.17148/IJARCCE.2026.155250

C. AI Components

The Intelligence Layer integrates two primary AI components. The first is an ML-Based ETA Prediction Model: an ensemble combining a stacked LSTM sequence network (3 layers, 64 units, dropout 0.2) processing 60-timestep sliding windows of speed, elapsed time, and segment identity features with a Random Forest regressor providing bias-correction adjustment, targeting mean absolute ETA error below 90 seconds on validation datasets. The second is a Traffic Congestion Index: a segment-level scoring system computing normalized deviation from historical median speed per route segment and time-of-day window, feeding both the ETA feature vector and the admin dashboard congestion overlay.

D. Key Parameters

The proposed system incorporates the following parameters in ETA computation:

- Vehicle Speed: GPS-derived instantaneous velocity and 60-second moving average.
- Segment Distance: Remaining route distance to each upcoming stop from current position.
- Historical Congestion: Median observed speed per segment per day-of-week and hour-of-day.
- Time-of-Day Encoding: Sinusoidal features capturing cyclic diurnal traffic patterns.
- Dwell Time Estimate: Stop-specific historical mean boarding and alighting duration.
- GPS Fix Quality: Satellite count and horizontal dilution of precision for confidence weighting.

VI. EXPECTED OUTCOMES AND BENEFITS

A. ETA Prediction Accuracy

The ensemble LSTM–Random Forest architecture is expected to achieve mean absolute ETA error below 90 seconds on routes with at least two months of historical trajectory data. This represents a substantial improvement over naive distance-divided-by-speed baselines, which typically produce errors exceeding four minutes under moderate congestion. Incorporating time-of-day and segment-level historical congestion features is projected to reduce prediction error by 20–25% relative to single-parameter speed models.

B. Passenger Experience Improvement

Accurate ETA countdowns delivered through the Flutter mobile application enable passengers to optimize departure timing, eliminating the over-early-arrival behavior documented in urban transit behavioral research. Geofence proximity notifications provide actionable advance warning without requiring passengers to continuously monitor the application. Accessibility is extended through SMS-based ETA delivery for passengers without smartphones.

C. Operational Benefits for Fleet Administrators

The centralized dashboard and route adherence analytics provide transport supervisors with evidence-based operational intelligence that manual reporting systems cannot match. Automated delay alerts, segment congestion maps, and longitudinal service quality reports support proactive fleet management including relief vehicle dispatch and schedule adjustment decisions driven by actual performance data.

D. Cost Accessibility and Scalability

Per-vehicle hardware cost for an ESP32, NEO-7M GPS module, SIM800L GSM module, and power regulation components is estimated below ₹2,500 (approximately USD 30), making the system deployable by educational institutions and small public transport operators. The serverless cloud architecture scales from single-bus deployments to multi-hundred-vehicle fleets without infrastructure redesign.

VII. CONCLUSION AND FUTURE WORK

This paper surveyed six recent studies on IoT-based vehicle tracking and AI-driven transportation prediction systems. The reviewed works collectively demonstrate that the enabling technologies for comprehensive smart bus tracking have individually matured to production readiness. However, they reveal consistent gaps in unified system integration, multi-parameter AI prediction, passenger-centric mobile interface design, and communication-resilient operation.

To address these gaps, we propose an AI-Based Smart Bus Live Tracking System that integrates ESP32-based GPS tracking, Wi-Fi-enabled cloud communication, Firebase real-time synchronization, ETA prediction techniques, a Flutter passenger application, and an administrative monitoring dashboard. The proposed system is designed to be cost-effective, easy to deploy, and suitable for smart transportation applications in educational institutions and public transit environments.

International Journal of Advanced Research in Computer and Communication Engineering
Impact Factor 8.471 ⁙ Peer-reviewed & Refereed journal ⁙ Vol. 15, Issue 5, May 2026
DOI: 10.17148/IJARCCE.2026.155250

The key contributions of this work include the identification of major gaps in existing bus tracking systems, the proposal of a unified IoT-based tracking architecture, and the development of a real-time monitoring platform that improves passenger convenience and operational efficiency. Future enhancements may include advanced machine learning-based ETA prediction, multi-route deployment support, traffic-aware analytics, and integration with intelligent transportation infrastructure.

VIII. ACKNOWLEDGEMENT

The authors extend their sincere gratitude to the Department of Computer Science and Engineering, K.S. School of Engineering and Management (KSSEM), Bengaluru, for providing an excellent research environment and institutional support. We thank our project guide and faculty members for their guidance and constructive feedback throughout the preparation of this paper. We also acknowledge the research community and the authors of the reviewed studies, whose published work forms the analytical foundation of this survey.

REFERENCES

[1]. N. K, M. Varghese, M. M. Paul, V. K. Rao, and V. D, “Bus Odyssey: The Intelligent College Bus Tracking Companion,” in Proc. 1st Int. Conf. Smart and Intelligent Systems (SISCON), IEEE, 2025.

[2]. A. Kumar, R. Sharma, and P. Verma, “AI-Based Timetable Generation and Vehicle Arrival Time Prediction Using YOLOv5 and Linear Regression,” Int. J. Eng. Research & Technology (IJERT), vol. 11, no. 4, pp. 233–238, 2023.

[3]. Dr. Tejashwini N, Prof. Manjusha P K, Sagar P, Ramprasad Sharma, Mithun M S, and V Harshith, “A Predictive Platform for Bus Mobility and Real-Time Human Flow Analysis,” IJARCCE, vol. 15, no. 1, Jan. 2026. DOI: 10.17148/IJARCCE.2026.151153.

[4]. Ms. M. Sowndharya, Yuvaraj K, Rahul S, Srichandiran S, and Sriram S, “Smart College Bus Tracker: Real-Time Location and ETA Predictor,” IJSRED, vol. 8, no. 3, May–Jun. 2025. ISSN: 2581-7175.

[5]. I. Moumen, M. Elgarej, and M. Mansouri, “Real-Time GPS Tracking System for IoT-Enabled Connected Vehicles,” E3S Web of Conferences, vol. 412, art. 01095, 2023. DOI: 10.1051/e3sconf/202341201095.

[6]. M. S. Rajan and V. Chandrasekaran, “Kalman Filter and Fuzzy Logic Fusion for Real-Time Bus Trajectory Smoothing,” Springer J. Intelligent Transportation Systems Research, vol. 20, no. 3, pp. 614–629, 2022.

