INTERNATIONAL JOURNAL OF ECONOMIC SOCIAL SCIENCE AND MANAGEMENT LAW
Peer Reviewed, Referred & Indexed Journal
E-ISSN: 3070-0558
VOL.7, NO.1 (2026)

SMART AUTONOMOUS VEHICLE PERCEPTION USING REINFORCEMENT LEARNING–DRIVEN V2I SENSOR FUSION

K LAVANYA[1], KUNAM LAVANYA[2], KALLA ROHITH KUMAR[3], KONDKTI HARI KRISHNA[4], KUMMARI RAJU[5], MADDALA SHANMUKA SAIGANESH[6]

ASSISTANT PROFESSOR[1], UG SCHOLAR[2,3,4,5&6]

DEPARTMENT OF CSE, NARSIMHA REDDY ENGINEERING COLLEGE (UGC- AUTONOMOUS) MAISAMMAGUDA (V), KOMPALLY, SECUNDERABAD, TELANGANA-500100

ABSTRACT: Autonomous driving research has evolved from rule-based vehicle systems to data-driven perception using V2I communication and multi-sensor fusion. In India, rapid urbanization has led to severe traffic congestion, with over **1.5 lakh road fatalities annually** and average urban traffic speeds below **25 km/h**. Government initiatives like **Smart Cities Mission and Intelligent Transportation Systems (ITS)** emphasize V2I-enabled safety, making AI-driven perception and traffic intelligence highly relevant. This framework aims to enhance autonomous vehicle perception by fusing infrastructure and vehicle sensor data using machine learning and reinforcement learning to accurately predict traffic density and vehicle priority levels. In traditional systems, traffic monitoring relies on fixed rule-based logic, human-operated traffic signals, roadside cameras, and manual control centers. Traffic density is estimated through periodic surveys or static sensors, while vehicle priority is assigned using predefined rules for emergency or public vehicles without adaptive intelligence. Manual systems lack real-time adaptability, are prone to human error, and cannot effectively handle dynamic traffic patterns. They fail to scale with increasing vehicle density, provide delayed responses, and are unable to learn from historical data, leading to inefficient traffic flow and reduced road safety. The motivation arises from limitations of manual and rule-based systems in handling complex, dynamic traffic environments. This research aims to overcome poor scalability, delayed decision-making, and lack of learning capability by introducing intelligent, data-driven models that adapt in real time and improve perception accuracy. The proposed V2I-MSF framework integrates multi-sensor data from vehicles and roadside infrastructure and applies machine learning and reinforcement learning for intelligent perception and decision-making. SVM and Random Forest models classify vehicle priority and predict traffic density, while stacked RF-HST-LLR ensembles improve robustness and accuracy.

1. INTRODUCTION Autonomous vehicles (AVs) are designed to navigate real-world scenarios with minimal or no human intervention, aiming to make transportation safer and more efficient. Equipped with sensors such as cameras, LiDAR, radar, and the Global Positioning System (GPS), these systems rely on a multi-sensor approach to perceive their environment reliably. This redundancy addresses the leading cause of road accidents—human error—by reducing human involvement in driving, thereby minimizing crashes, injuries, and fatalities [1]. AVs integrate data from multiple sensors, such as cameras and LiDAR, to create a complete and accurate understanding of their surroundings. This process, known as multi-sensor data fusion, significantly enhances perception capabilities [2]. Despite these advancements, AVs face challenges related to sensor limitations and integration. Cameras provide detailed visuals but lack depth perception and struggle in low-light conditions, while LiDAR offers 3D spatial data but is sparse at long distances and misses semantic details such as traffic light states [3]. Multi-sensor fusion enhances object detection, tracking, and situational awareness, which are critical for safe and efficient vehicle navigation. However, traditional sensor fusion methods often focus only on local features, missing the broader context required for complex urban scenarios like busy intersections or interactions with vehicles and pedestrians. Additionally, delays in processing sensor data can result in outdated information, leading to suboptimal real-time decisions [4]. In this study, we address these challenges by proposing a multi-sensor fusion framework that integrates data from vehicle-mounted sensors, such as LiDAR and cameras, with infrastructure-based sensor data. This V2I-MSF approach leverages the strengths of both sources: high-resolution spatial and semantic data from vehicle sensors, and enhanced perception from strategically positioned infrastructure sensors, which can mitigate issues like occlusions and limited sensor range [5]. Additionally, our work specifically targets urban scenarios such as intersections, where obstacle avoidance, optimal lane selection, and reduced congestion are crucial for safety and efficiency. By incorporating advanced techniques to improve 3D visualization and camera-based environmental understanding, our framework addresses obstacles more effectively, ensures better autonomous decision-making, and reduces errors in perception. These enhancements are especially valuable in improving navigation and reducing bottlenecks in complex, high-traffic environments [6]. To achieve these objectives, we utilize a suite of sensors tailored for both vehicle-mounted and infrastructure-based applications

1.1 Overview

Autonomous vehicle perception has evolved from isolated onboard sensing to cooperative intelligence enabled by Vehicle-to-Infrastructure (V2I) communication and multi-sensor fusion. Early traffic systems relied on manual monitoring and fixed signal timings, which proved inadequate for dense urban environments. In India, rapid urbanization and vehicle growth have created critical transportation challenges, with over 150,000 road fatalities reported annually and urban congestion causing average travel speeds below 25 km/h in major cities. Government initiatives such as Intelligent Transportation Systems (ITS), Smart Cities Mission, and Digital India promote infrastructure-assisted vehicular intelligence. The proposed V2I Multi-Sensor Fusion with Reinforcement Learning framework integrates infrastructure sensors, vehicle data, and machine learning models to enhance real-time perception, priority classification, and traffic density prediction, enabling safer and more efficient autonomous transportation systems.

Introduction with Applications (4 lines):

The proposed system enhances autonomous vehicle perception using V2I multi-sensor fusion and machine learning. It enables accurate traffic density estimation and vehicle priority classification. The framework supports intelligent decision-making in dynamic traffic environments. Applications include smart traffic control, emergency vehicle prioritization, and autonomous navigation.

1.2 Problem Definition

Before the adoption of machine learning, traffic management and autonomous perception systems relied on manual control, static rules, and isolated sensing mechanisms. These systems lacked real-time adaptability and failed to handle dynamic traffic patterns. Priority assignment depended on predefined vehicle categories rather than real-time conditions. Traffic density estimation suffered from delayed and inaccurate measurements. The absence of learning mechanisms resulted in inefficient resource utilization and increased road safety risks.

1.3 Research Motivation

The motivation for this research arises from the need to overcome the limitations of manual and rule-based traffic perception systems. Increasing vehicle density, complex urban road networks, and real-time decision requirements demand intelligent and adaptive solutions. Machine learning enables accurate pattern recognition from multi-sensor data. Ensemble learning improves prediction reliability. Reinforcement learning supports continuous optimization of perception and decision processes.

1.4 Objective of the Study

The primary objective of this study is to design and implement a V2I-based multi-sensor fusion framework that accurately classifies vehicle priority levels and predicts traffic density using machine learning and advanced ensemble techniques. The system aims to improve autonomous vehicle perception, reduce traffic congestion, enhance road safety, and support real-time intelligent transportation decisions.

1.5 Applications

The proposed framework supports smart traffic signal control by dynamically adjusting signal timing based on traffic density. It enables emergency vehicle prioritization through real-time priority classification. Autonomous vehicles benefit from improved situational awareness and safer navigation decisions. Urban traffic authorities gain accurate congestion monitoring tools. Edge and cloud computing environments support scalable deployment. Smart city platforms integrate real-time traffic analytics. Logistics and fleet management systems optimize routing decisions. Public transportation systems improve schedule reliability and passenger safety.

1.6 Significance of the Study

This research contributes a robust and scalable intelligent perception framework for next-generation transportation systems. It improves traffic efficiency and safety through accurate data-driven predictions. The integration of V2I communication with ensemble machine learning strengthens autonomous decision-making. The proposed system supports national smart mobility initiatives and enables practical deployment in real-world urban environments

2.LITERATURE SURVEY

- Autonomous driving systems significantly benefit from advancements in multi-sensor fusion and cooperative perception, which address challenges like environmental variability, sensor synchronization, and data heterogeneity. Fusion strategies are broadly categorized into symmetric and asymmetric methods, with subcategories such as data-data and result-result fusion, enhancing perception robustness and real-time processing in complex scenarios [1]. Infrastructure-augmented systems, combining vehicle and roadside sensor data, overcome latency and jitter issues using novel interframe fusion techniques, ensuring timely decision-making and improved driving safety under dynamic network conditions [2]. Roadside monocular cameras, processed into BEV maps, utilize techniques like Kalman filtering and Graph Neural Network (GNN) algorithms to address occlusions and enable reliable vehicle tracking in highway merging scenarios [3]. Advanced frameworks such as Vehicle-Infrastructure Multi-view Intermediate fusion (VIMI) and ViT-FuseNet integrate multimodal data (e.g., LiDAR and cameras) using attention mechanisms and feature compression to enhance 3D object detection accuracy while minimizing data transmission overhead, achieving state-of-the-art performance on datasets like DAIR-V2X [4], [5].
- Collaborative perception frameworks, such as V2X-Sim and OccFusion, leverage simulations and dynamic fusion modules to enhance object detection,

multi-object tracking, and semantic segmentation by mitigating occlusion and environmental challenges through intermediate feature sharing and attention-based integration [7], [15]. Infrastructure-assisted systems like IN2Lab employ global object fusion and collaborative perception strategies to improve detection accuracy, extend vehicle awareness, and predict potential collision scenarios, particularly in adverse conditions [6], [18]. Federated reinforcement learning (FRL) frameworks and deep learning models like TransFuser and ContextualFusion address the complexities of urban environments and adverse conditions by dynamically adjusting data integration and leveraging attention mechanisms for trajectory planning and reduced collision rates **b21** [22], [23]. Probabilistic data fusion frameworks using methods like Covariance Intersection and advanced Kalman filtering enhance data reliability and improve cooperative perception under scenarios with sensor failures or latency, enabling robust decision-making and precise object tracking [10], [18].

- Feature fusion techniques, such as Fusion2comm and ContextualFusion, dynamically prioritize critical features based on environmental context (e.g., occlusion, adverse weather) to optimize bandwidth and detection precision, improving overall perception capabilities [20], [21]. Transformer-based fusion techniques effectively integrate multimodal data (e.g., LiDAR, cameras, and radar) through cross-attention mechanisms, achieving superior performance in 3D object detection and surpassing traditional CNN-based methods [13]. V2I frameworks like V2IViewer employ spatiotemporal fusion methods and robust tracking algorithms to address communication delays and temporal asynchrony, enhancing real-time detection accuracy and robustness [9], [24]. Simulation frameworks, such as CARLA-based V2X systems, simulate diverse scenarios and generate valuable datasets to test and improve autonomous driving algorithms, advancing research in collaborative perception [7], [25].
- RL approaches, including Deep Q-Network (DQN) and Shared Experience Actor-Critic (SEAC), optimize resource allocation in 5G-NR-V2X networks under high-density traffic and dynamic data patterns, improving reliability and reducing collisions by leveraging multi-agent learning and advanced actor-critic networks [14], [17]. These studies collectively emphasize the importance of multi-sensor integration, robust data fusion, and AI-driven methodologies to enhance perception, safety, and efficiency in autonomous driving systems.
- To further optimize sensor fusion and hypothesis testing, recent approaches have introduced autoencoder-based multisensor fusion, improving bandwidth efficiency in distributed inference by eliminating redundant sensor data while retaining critical information, making it highly suitable for V2I applications [26]. Additionally, UAV-based multi-sensor fusion has demonstrated enhanced predictive accuracy by integrating RGB, multispectral, and thermal infrared sensor data, a concept that parallels the need for robust multi-modal fusion in autonomous driving [27]. Moreover, decentralized multiple hypothesis testing has been proposed to improve decision accuracy in heterogeneous sensor networks using total variance regularization and probabilistic clustering, which can enhance V2I-based perception, anomaly detection, and object recognition in autonomous driving [28].
- Realistic communication challenges are essential in autonomous driving system evaluations. [30] considers practical traffic scenarios through microsimulations that account for communication interruptions during vehicle platooning, reflecting real-world delays and bandwidth constraints [29] simulates dynamic network conditions during ADS training under traffic congestion, signal variations, and data loss, ensuring robust system performance under practical V2X communication challenges [31] assumes non-ideal communication by integrating real-time traffic and environmental factors in their VR-enabled platform, ensuring ADS operations are tested under fluctuating network conditions such as varying latencies and intermittent packet loss

3.SYSTEM ANALYSIS

EXISTING SYSTEM

Current autonomous vehicle perception systems mainly rely on **onboard sensors** such as cameras, LiDAR, radar, and ultrasonic sensors to understand the surrounding environment. These sensors collect information about road conditions, nearby vehicles, pedestrians, traffic signals, and obstacles. The data from multiple sensors is usually combined using **sensor fusion techniques** to improve perception accuracy.

Most existing systems perform sensor fusion within the vehicle itself using algorithms such as **Kalman filters, Bayesian fusion, and deep learning-based perception models**. While these methods improve object detection and environmental understanding, they still have several limitations. Autonomous vehicles often face challenges such as **limited sensing range, occlusion of objects, poor weather conditions, and blind spots**. When obstacles are hidden behind buildings or other vehicles, onboard sensors alone may not detect them in time.

Additionally, traditional sensor fusion methods generally rely on **static or rule-based decision mechanisms**, which may not adapt efficiently to dynamic traffic environments. The lack of communication between vehicles and infrastructure limits the availability of external information that could improve perception accuracy.

PROPOSED SYSTEM

The proposed system introduces a **Smart Autonomous Vehicle Perception framework using Reinforcement Learning–Driven Vehicle-to-Infrastructure (V2I) Multi-Sensor Fusion**. This approach enhances vehicle perception by integrating data from both **vehicle sensors and roadside infrastructure sensors**.

In this system, infrastructure units such as **smart traffic cameras, roadside LiDAR sensors, and edge computing nodes** collect environmental data and share it with nearby vehicles through **Vehicle-to-Infrastructure (V2I) communication networks**. The data from vehicle sensors and infrastructure sensors is combined using **multi-sensor fusion techniques** to provide a more comprehensive view of the environment.

A **reinforcement learning (RL) framework** is applied to intelligently manage the fusion process and decision-making. The RL agent learns optimal strategies for selecting, weighting, and combining sensor data based on environmental conditions and system performance. Through continuous interaction with the environment, the RL model improves perception accuracy and decision quality over time.

This approach enables autonomous vehicles to detect hidden obstacles, predict traffic situations more accurately, and respond faster to potential hazards. By leveraging **V2I communication, multi-sensor fusion, and reinforcement learning**, the proposed system significantly enhances the reliability and safety of autonomous vehicle perception.

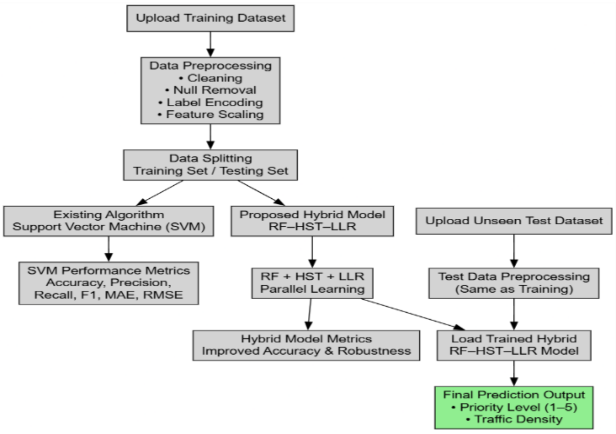
ADVANTAGES OF THE PROPOSED SYSTEM

- Improved perception accuracy using multiple sensor sources
- Extended sensing range beyond the vehicle's onboard sensors
- Better detection of occluded or hidden obstacles
- Adaptive decision-making using reinforcement learning
- Enhanced safety in complex traffic environments
- Efficient real-time data processing through V2I communication networks

4.IMPLEMENTATION

Step 1: Dataset Identification and Collection

The first step involves identifying and selecting an appropriate dataset relevant to V2I communication, traffic behavior, and infrastructure-level parameters. The dataset contains vehicle dynamics, network characteristics, and service-level attributes such as speed, signal strength, latency, bandwidth, congestion indicators, and QoS-related metrics. Target variables include Priority Level (classification) and Traffic Density (regression). This dataset provides a realistic representation of heterogeneous traffic and infrastructure conditions required for intelligent perception and decision-making.



Step 2: Dataset Preprocessing

Raw datasets often contain inconsistencies, missing values, and non-numeric attributes that can degrade model performance. In this step, data preprocessing is performed to clean and standardize the dataset. Null and irrelevant values are removed to ensure data integrity. Categorical attributes such as service type, vehicle category, or node identifiers are transformed into numerical representations using label encoding. Feature normalization and scaling are applied where required to bring all attributes into a common range, enabling fair contribution during model training.

Step 3: Existing Model Building (SVM and Random Forest)

In this research, conventional machine learning models such as Support Vector Machine (SVM) and Random Forest (RF) are implemented as baseline models. These algorithms are trained separately for classification (priority level prediction) and regression (traffic density estimation). The performance of these models serves as a benchmark for evaluating the effectiveness of the proposed ensemble-based approach.

Step 4: Proposed Model Building (RF-HST-LLR with Advanced Ensemble Voting)

The proposed system introduces a stacked ensemble learning model combining Random Forest, HST, and Local Linear Regression (LLR). Each model captures different data characteristics—Random Forest handles nonlinear relationships, HST improves hierarchical decision learning, and LLR refines local predictions. An advanced ensemble voting mechanism aggregates outputs from individual learners to generate robust and accurate predictions. This hybrid design significantly enhances generalization and reduces prediction uncertainty in dynamic traffic environments.

Step 5: Performance Evaluation

Model performance is evaluated using standard metrics. For classification tasks, accuracy, precision, recall, and F1-score are used to assess priority prediction effectiveness. For regression tasks, MAE, RMSE, and R^2 score evaluate traffic density prediction accuracy. Comparative analysis between existing and proposed models highlights the improvements achieved through ensemble learning.

Step 6: Prediction on Unseen Test Data

Finally, the trained models are applied to unseen test data to validate real-world applicability. This step ensures that the system can generalize well beyond training samples. The prediction module outputs priority levels and traffic density estimates, supporting real-time autonomous vehicle perception and decision-making.

4.2 Data Preprocessing in This Research

Data preprocessing plays a crucial role in improving model reliability and accuracy. Initially, missing and null values are identified and removed to prevent biased learning. Duplicate and irrelevant attributes are discarded to reduce noise and dimensionality. Categorical features such as vehicle type, service category, and cloud provider are converted into numerical values using label encoding, enabling compatibility with machine learning algorithms.

Feature engineering is also performed by deriving meaningful attributes such as congestion level indicators, normalized bandwidth usage, and latency ratios. Continuous numerical features are scaled to avoid dominance of high-magnitude values. Textual or symbolic identifiers are standardized to maintain consistency. These preprocessing steps ensure clean, structured, and machine-readable data suitable for advanced modeling.

4.3 Exploratory Data Analysis (EDA)

Exploratory Data Analysis is conducted to understand the underlying patterns, relationships, and distributions within the dataset. EDA examines how features such as speed, bandwidth, latency, and congestion correlate with priority levels and traffic density. Visualization techniques are used to identify skewness, outliers, and feature importance trends.

During EDA, the dataset is logically divided into training and testing subsets to ensure unbiased evaluation. This separation ensures that models learn patterns from historical data while being validated on unseen samples. EDA helps justify feature selection and model choice by revealing nonlinear relationships and dependencies critical to autonomous perception systems.

4.4 Train-Test Split

The dataset is divided into training and testing subsets to evaluate model generalization. The training set is used to learn patterns and relationships, while the testing set evaluates predictive capability. This separation prevents overfitting and ensures fair performance assessment. By maintaining representative samples in both subsets, the system achieves reliable and consistent evaluation outcomes under real-world conditions.

4.5 Model Building

The model-building phase involves training both existing and proposed algorithms on preprocessed data. Feature vectors are mapped to corresponding target variables for classification and regression tasks. Hyperparameters are tuned to optimize performance. The ensemble model integrates multiple learners to achieve higher robustness, stability, and predictive accuracy compared to standalone models.

4.5.1 Existing Algorithm – Support Vector Machine (SVM)

Definition:

Support Vector Machine (SVM) is a supervised machine learning algorithm used for both classification and regression tasks. SVM aims to identify an optimal decision boundary (hyperplane) that maximizes the margin between different classes. For regression, it estimates a function that deviates minimally from actual target values while maintaining model simplicity.

Working Principle:

SVM transforms input data into a high-dimensional feature space using kernel functions such as linear, polynomial, or radial basis functions. It then determines the hyperplane that best separates classes or fits regression targets. Support vectors—critical data points near the boundary—define the model, making SVM effective in handling complex nonlinear relationships.

Algorithm Steps (Architecture):

[1] Input feature extraction and normalization
[2] Kernel selection and transformation
[3] Margin optimization using support vectors
[4] Hyperplane construction
[5] Classification or regression prediction

Disadvantages:

Despite its effectiveness, SVM suffers from high computational complexity for large datasets. Kernel selection and parameter tuning are challenging and time-consuming. SVM also lacks interpretability and performs poorly when noise and overlapping classes are present, limiting its scalability for large-scale V2I environments

Hybrid RF–HST–LLR Model: Definition and Detailed Explanation

The **Hybrid RF–HST–LLR model** is an advanced ensemble learning framework that integrates **Random Forest (RF)**, **Hierarchical Soft Tree (HST)**, and **Local Linear Regression (LLR)** to improve prediction accuracy and robustness in complex V2I-based autonomous vehicle perception systems. This hybrid architecture combines the global learning capability of Random Forest, the hierarchical decision refinement of HST, and the local approximation strength of LLR. By leveraging diverse learning behaviors, the model effectively captures nonlinear patterns, hierarchical relationships, and localized variations in traffic and infrastructure data, making it highly suitable for priority classification and traffic density regression tasks.

Working Principle of Hybrid RF–HST–LLR

The hybrid model operates by processing the same input feature set through three complementary learners in parallel. Random Forest captures complex nonlinear interactions and feature importance at a global scale. HST further refines decisions through a layered, soft-decision hierarchy that improves interpretability and class boundary smoothness. LLR focuses on local neighborhoods of data points, providing fine-grained regression adjustments. The outputs from all three models are then combined using an **advanced ensemble voting mechanism**, which dynamically weighs predictions to generate a final, more accurate output. This cooperative learning strategy reduces bias, variance, and uncertainty.

Algorithm Steps

1. Input data acquisition and preprocessing
2. Parallel feature learning using RF, HST, and LLR
3. Independent prediction generation from each model
4. Confidence-based ensemble voting and aggregation
5. Final classification and regression output generation

Advantages of Hybrid RF–HST–LLR

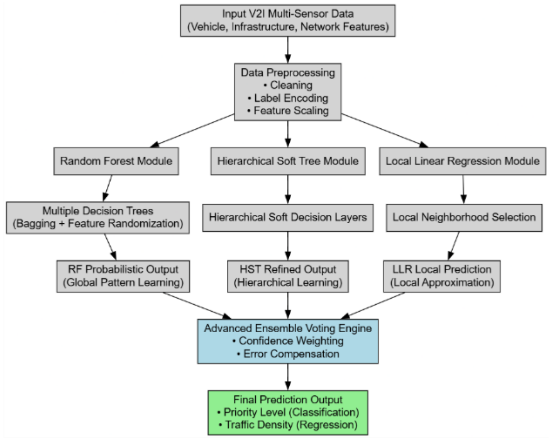
- Improved prediction accuracy through model diversity
- Reduced overfitting and variance
- Better handling of nonlinear and hierarchical data
- Robust performance under dynamic traffic conditions
- Enhanced generalization on unseen data

Internal Operational Steps of RF–HST–LLR with Advanced Ensemble Voting

1. RF generates probabilistic predictions using multiple decision trees
2. HST refines outputs through hierarchical soft decision layers
3. LLR adjusts predictions based on local data patterns
4. Ensemble voting assigns adaptive weights to each model
5. Aggregated output produces final prediction

Mathematical Intuition

The mathematical intuition behind the proposed hybrid model lies in **ensemble diversity and error decomposition**. Random Forest minimizes variance through bagging, HST reduces bias by structured decision refinement, and LLR minimizes local approximation error. The ensemble voting mechanism combines these complementary error profiles, leading to improved overall generalization. Instead of relying on a single hypothesis space, the model aggregates multiple perspectives of the data distribution, resulting in stable and accurate predictions.



CONCLUSION

This work presents a comprehensive intelligent service management and prediction framework by integrating machine learning with cloud–edge and VANET-inspired environments. The proposed system successfully demonstrates how real-time operational metrics such as CPU utilization, memory and storage usage, network bandwidth, latency, response time, throughput, and QoS indicators can be leveraged to make informed decisions about optimal service placement and performance optimization. By employing advanced machine learning models such as Support Vector Machines, Random Forests, and hybrid ensemble approaches, the system is capable of accurately classifying service priority, predicting traffic density or service load, and supporting proactive resource allocation. The modular Flask-based architecture ensures scalability, ease of deployment, and seamless interaction between data preprocessing, model training, prediction, and visualization components. Overall, the solution enhances decision-making efficiency, reduces service latency, improves resource utilization, and supports reliable Quality of Service in dynamic and heterogeneous computing environments.

FUTURE SCOPE

The scope for future enhancement of this system is significant and multifaceted. The framework can be extended by incorporating deep learning models such as LSTM, GRU, or Graph Neural Networks to better capture temporal dependencies and complex interactions among distributed edge nodes. Real-time data streaming from VANETs, IoT sensors, and 5G/6G networks can be integrated to enable online learning and adaptive model updates, making the system more responsive to rapidly changing network conditions. Additionally, reinforcement learning techniques can be applied to dynamically optimize service placement and load balancing decisions based on continuous feedback from the environment. Security and privacy can be strengthened by integrating federated learning and blockchain-based trust mechanisms, ensuring decentralized model training without exposing sensitive data. Finally, the system can be deployed and evaluated at scale using container orchestration platforms such as Kubernetes, with support for multi-cloud providers, thereby transforming the framework into a production-ready intelligent resource management solution for next-generation smart transportation and edge computing ecosystems.

REFERENCES

1. C. Xiang, C. Feng, X. Xie, B. Shi, H. Lu, Y. Lv, M. Yang, and Z. Niu, "Multi-sensor fusion and cooperative perception for autonomous driving: A review," *IEEE Intell. Transp. Syst. Mag.*, vol. 15, no. 5, pp. 36–58, Sep. 2023, doi: 10.1109/MITS.2023.3283864.

2. J. Wang, Z. Wang, B. Yu, J. Tang, S. L. Song, C. Liu, and Y. Hu, "Data fusion in infrastructure-augmented autonomous driving system: Why? where? And how?," *IEEE Internet Things J.*, vol. 10, no. 18, pp. 15857–15871, Sep. 2023, doi: 10.1109/JIOT.2023.3266247.

3. K. Shan, M. Penlington, S. Gunner, K. Koufos, M. Dianati, A. Fairgrieve, and I. Kirwan, "Experimental study of multi-camera infrastructure perception for V2X-assisted automated driving in highway merging," *IEEE Trans. Intell. Transp. Syst.*, vol. 25, no. 11, pp. 16207–16220, Nov. 2024, doi: 10.1109/TITS.2024.3424673.

4. Z. Wang, S. Fan, X. Huo, T. Xu, Y. Wang, J. Liu, Y. Chen, and Y.-Q. Zhang, "VIMI: Vehicle-infrastructure multi-view intermediate fusion for camera-based 3D object detection," 2023, *arXiv:2303.10975*.

Show in Context

5. Z. Yang, Y. Cai, P. Wang, C. Wang, X. Wang, and N. N. Van, "ViT-FuseNet: Multimodal fusion of vision transformer for vehicle-infrastructure cooperative perception," *IEEE Access*, vol. 12, pp. 31640–31651, 2024, doi: 10.1109/ACCESS.2024.3368404.

6.T. D. Borba, O. Vaculín, H. Marzbani, and R. N. Jazar, "Increasing safety of automated driving by infrastructure-based sensors," *IEEE Access*, vol. 11, pp. 94974–94991, 2023, doi: 10.1109/ACCESS.2023.3311136.

7.Y. Li, D. Ma, Z. An, Z. Wang, Y. Zhong, S. Chen, and C. Feng, "V2X-sim: Multi-agent collaborative perception dataset and benchmark for autonomous driving," *IEEE Robot. Autom. Lett.*, vol. 7, no. 4, pp. 10914–10921, Oct. 2022, doi: 10.1109/LRA.2022.3192802.

8.C. Chen, Q. Tang, X. Hu, and Z. Huang, "Infrastructure sensor-based cooperative perception for early stage connected and automated vehicle deployment," *J. Intell. Transp. Syst.*, vol. 28, no. 6, pp. 956–970, Nov. 2024.

Show in ContextCrossRef

9.S. Yi, H. Zhang, and K. Liu, “V2IViewer: Towards efficient collaborative perception via point cloud data fusion and vehicle-to-infrastructure communications,” *IEEE Trans. Netw. Sci. Eng.*, vol. 11, no. 6, pp. 6219–6230, Nov. 2024, doi: 10.1109/TNSE.2024.3479770.

10.M. Shan, K. Narula, S. Worrall, Y. F. Wong, J. Stephany Berrio Perez, P. Gray, and E. Nebot, “A novel probabilistic V2X data fusion framework for cooperative perception,” in *Proc. IEEE 25th Int. Conf. Intell. Transp. Syst. (ITSC)*, Oct. 2022, pp. 2013–2020, doi: 10.1109/ITSC55140.2022.9922251.

11.T. Jiawei, C. J. Pawase, and K. Chang, “Adaptive sidelink open loop power control optimization strategies for Vehicle-to-Vehicle communications in 5G-NR-V2X,” *IEEE Access*, vol. 12, pp. 25079–25089, 2024, doi: 10.1109/ACCESS.2024.3365133.View Article

12.C. Xiang, L. Zhang, X. Xie, L. Zhao, X. Ke, Z. Niu, and F. Wang, “Multi-sensor fusion algorithm in cooperative vehicle-infrastructure system for blind spot warning,” *Int. J. Distrib. Sensor Netw.*, vol. 18, no. 5, May 2022, Art. no. 155013292211004, doi: 10.1177/15501329221100412.Show in ContextCrossRef

13.J. Yao, J. Zhou, Y. Wang, Z. Gao, and W. Hu, “Infrastructure-assisted 3D detection networks based on camera-lidar early fusion strategy,” *Neurocomputing*, vol. 600, Oct. 2024, Art. no. 128180.Show in ContextCrossRef

14.A. Singh, “Transformer-based sensor fusion for autonomous driving: A survey,” in *Proc. IEEE/CVF Int. Conf. Comput. Vis. Workshops (ICCVW)*, Oct. 2023, pp. 3304–3309, doi: 10.1109/iccvw60793.2023.00355.

15.W. Terapaptommakol, D. Phaoharuhansa, P. Koowattanasuchat, and J. Rajruangrabin, “Design of obstacle avoidance for autonomous vehicle using deep q-network and CARLA simulator,” *World Electric Vehicle J.*, vol. 13, no. 12, p. 239, Dec. 2022, doi: 10.3390/wevj13120239. Show in ContextCrossRef