International Journal of Science and Research Archive
eISSN: 2582-8185
Cross Ref DOI: 10.30574/ijsra
Journal homepage: https://ijsra.net/

(Review Article)

Federated deep learning for privacy-preserving sensor fusion in autonomous vehicles

Jainam Dipakkumar Shah *

Stevens Institute of Technology.

International Journal of Science and Research Archive, 2026, 18(01), 123-133

Publication history: Received on 16 November 2025; revised on 05 January 2026; accepted on 08 January 2026

Article DOI: https://doi.org/10.30574/ijsra.2026.18.1.3329

Abstract

The upsurge of autonomous vehicles (AVs) is mainly supported by colossal multimodal sensor data acquired from various sources, including cameras, LiDAR, and radar systems. Nonetheless, the pooling of such data processed across different vehicles and organizations raises significant privacy, security, and compliance issues concerning international data protection laws. Our contribution in this research is the federated deep learning (FDL) framework that is capable of performing privacy-preserving sensor fusion without the need for raw data sharing across multiple cloud platforms. The proposed system is built upon the CNN-LSTM hybrid architectures for the extraction of multimodal features and also employs Federated Averaging (FedAvg) for the distributed model aggregation. The experiments are carried out on three open-source datasets, KITTI, nuScenes, and Waymo Open Dataset, that represent real-world driving scenarios with different types of sensors. The results reveal that federated deep learning is a suitable technique for the establishment of learning pipelines in AVs that are privacy-compliant across fleets and provide a robust basis for the development of future intelligent transportation systems.

Keywords: Federated Deep Learning; Privacy-Preserving Sensor Fusion; Autonomous Vehicles (AVs); Multimodal Data Integration

1. Introduction

The rapid advancement of autonomous vehicles (AVs) is largely driven by the integration of advanced deep learning models capable of analyzing complex, high-dimensional sensor data. Modern AVs rely on multimodal sensor networks that combine inputs from cameras, LiDAR, radar, ultrasonic sensors, and GPS to ensure accurate perception and robust decision-making in dynamic real-world conditions. The integration of information from multiple sensing modalities, known as sensor fusion, enhances overall environmental awareness. It also compensates for the limitations of individual sensors, such as LiDAR's reduced accuracy in fog and a camera's poor visibility performance in low-light conditions [1], [2]. The fusion deep learning architecture has been demonstrated to greatly improve detection accuracy and reliability of such tasks as 3D object detection, semantic segmentation, and vehicle tracking [3]. However, as the AVs and their sensor complexity are increasing, so does the problem of processing the high volumes of data generated by the AVs on a continual basis. This giant growth necessitates scalable structures that can learn effectively by distributed multimodal data without any performance or privacy loss.

The traditional training of AV perception models often includes centralized data aggregation, i.e., they are fed with raw sensor measurements of multiple vehicles and then make use of them to train deep neural networks on a cloud or data center server. Despite the fact that this centralized paradigm can be used in the improvement of generalization of models, it raises severe concerns about data privacy, ownership, and compliance with existing regional data protection legislation such as the General Data Protection Regulation (GDPR) and the ISO/SAE 21434 automotive cybersecurity standards [4]. Raw driving data can consist of sensitive records of geolocation, identifiable individuals, or proprietary traffic scenarios, and they cannot send the data freely to manufacturers, research consortia, and authorities [5]. In

* Corresponding author: Jainam Dipakkumar Shah

addition, the issue of latency, the cost of bandwidth, and security vulnerability are also aggravated by data flow across borders and multi-cloud integration [6]. This brings about a growing trend of federated learning (FL), a decentralized system where the data are stored locally on the client (or AV) machines and only the model updates or gradients are sent to a central aggregator to train the model at an international level [7], [8]. The advantage of this is that it enables both joint training of models and is also a way of ensuring that raw sensory data are not leaving the source vehicle or company network, which meets privacy-by-design requirements and regulatory requirements.

In this paper, there is a presentation of the Federated Deep Learning (FDL) federation system of privacy-enhancing sensor fusion in autonomous vehicles. The system would be useful in the case of multiple clouds, where fleets, manufacturers, or mobility service providers would have the ability to jointly train perception models using distributed datasets without exchanging such datasets. The proposed pipeline utilizes a hybrid CNN-LSTM network to do the extraction of multimodal features that can be done by a convolutional layer to extract spatial relationships between image and point-cloud data, and the LSTM layers of the pipeline can capture the temporal connection needed to predict movement and track dynamic objects. Federated averaging (FedAvg) is an algorithm that averages local model weights on a secure federation server, which is based on a cloud server. These open-source, real-life datasets, such as KITTI Vision Benchmark, nuScenes, Waymo Open Dataset, are applied to simulate different environmental conditions and sensor configurations and deliver a good performance assessment [9]-[11]. The outcomes of the experiment show that the FDL framework can achieve nearly equal accuracy as centralized training, with a low communication cost and high data privacy and scalability. Under the proposed strategy, there can be a privacy-friendly, versatile foundation of the next-generation AVs, which will be able to learn together without sharing any personal information with the decentralized learning of the model provided, through the cloud systems.

This investigation differs from previous studies, including Google Research (2024), through the introduction of a multi-sensor cross-cloud federated learning framework. The framework inherently applies privacy-conscious safety conditions and is rigorously evaluated using real-world benchmark datasets. Unlike previous research that primarily focused on single-sensor configurations and intra-cloud federated designs, the proposed approach enables phased sensor fusion across heterogeneous cloud federations. This capability ensures seamless interaction among diverse fleets, manufacturers, and mobility service providers. The design corresponds to the philosophy of privacy-by-design, such as the incident mechanisms (encrypted model update transmission and secure cloud-level aggregation) that align with the world standards of data protection and automotive cybersecurity (GDPR and ISO/SAE 21434). This work can offer a scalable and regulation-unified platform on which the (eventually) autonomous vehicle learning systems may base their learning plans through balancing cross-cloud cooperation, cross-modal data combination, and high-level privacy compliance-satisfying the gaps between the experimental federated frameworks and the realizable AI ecosystems.

2. Related Work

To contextualize this study, it is necessary to examine related literature to ensure that this study is situated within the context of the evolving federated learning (FL), sensor fusion, and privacy-preserving machine learning of autonomous vehicles (AVs). This section presents the recent developments in the hybridization of distributed learning models and multimodal perception models, and the gaps in privacy compliance and scalability. Past studies have indicated that sensor fusion enhances the accuracy of perception of AVs by synthesizing complementary data of LiDAR, radar, and camera sensors. Simultaneously, the idea of federated learning has been becoming more and more popularized as the alternative that would guarantee privacy in the environment of centralized data aggregation, allowing for training models without sharing raw data between nodes or between organizations. However, most existing literature has either experimented with FL on a simulated system or has applied it to single-modality data, which leaves a gap in the practical implementation of cross-cloud and privacy-preserving multimodal learning on real-world autonomous configurations of perception systems. To develop the contents of the experimental structure, Table 1 will provide a comparison of the recent key studies that shall be utilized in mapping these developments in terms of datasets, architectures, results, and limitations.

The current tendencies shifting towards the decentralized paradigms of learning in autonomous systems could be traced in the studies described in Table 1. As the studies show [12, 13, 14], federated object detectors can maintain a high level of accuracy at a relatively low communication cost, yet they can only work with single-modal sensors. Studies also made improvements to the state of the art by proposing multimodal architectures (e.g., CNN-Transformer and CNN-LSTM hybrids) to AV perception [15, 16]. However, they rely on centralized training pipelines, and that is why they suspect that they do not comply with such privacy and data governance standards as GDPR and ISO/SAE 21434. Google Research [16] demonstrated the scalability of federated learning (FL) using simulated autonomous vehicles; however, its applicability to diverse real-world datasets has not been validated. According to these findings, the offered study will combine deep learning with the combination of multimodal sensor fusion with real data (KITTI, nuScenes, Waymo) and, therefore, a privacy-friendly and high-performance cross-cloud architecture, which can be used throughout collaborative learning of AV.

Table 1 Summary of Related Works on Federated Learning and Sensor Fusion in Autonomous Vehicles

Datasets Used	Architecture / Approach	Key Findings	Limitations
KITTI	Federated YOLOv5 with model averaging	Achieved 92.4% mAP with 25% reduced communication cost	Limited to a single modality (camera)
nuScenes	ResNet + FedAvg	Improved cross-fleet model generalization by 6% over local models	No privacy-preserving encryption is integrated
Waymo Open Dataset	CNN-Transformer hybrid for LiDAR-Camera fusion	Enhanced detection robustness by 8% in adverse weather	Centralized training violates GDPR
KITTI + nuScenes	CNN-LSTM hybrid with weighted aggregation	Reduced model drift by 15% across clients	High latency in global synchronization
Simulated AV datasets	TensorFlow Federated + Secure Aggregation	Demonstrated scalability with 50 simulated nodes	Lacks real-world dataset validation
General data model	Analytical study on 6G networks for VR/AR and AV systems	Provided comprehensive insights on low-latency, high-throughput communication for AVs with VR/AR integration	Not experimentally validated on real AV networks

3. System Architecture

The proposed FDL privacy-aware sensor fusion framework would allow the autonomous vehicle (AV) nodes to collaboratively learn a shared perception model on the cloud infrastructure without having to reveal the original sensor data. All of the AV clients operating in different environments with camera, LiDAR, and radar sensors, as depicted in Figure 1, undergo local preprocessing, including temporal synchronization, normalization, and feature encoding, before training a hybrid CNN-LSTM model. Every client will simply transmit encrypted updates of models to ensure privacy. They apply secure aggregation protocols (e.g., the Bonawatz et al. scheme) in the sense that the Federated Learning Server on the cloud can aggregate updates without knowledge of the contribution to it. The whole communication between the clients and the server is encrypted using TLS, and optional differential privacy noise (with graded e/ d parameters) may be added to the gradients to minimise information leakage further. It is the FedAvg algorithm that implements the calculation of the global model on the server and subsequently transfers the model to the clients, and goes on refining it in their respective locations until the desired level of accuracy is achieved. According to this design, information that is sensitive, such as the correct vehicle routes or company environmental trends, does not leave the local machine. The architecture allows knowledge transfer of distributed fleets, data confidentiality, safe global aggregation, model performance maintenance, and efficient communication improvement. It hence offers a privacy-conscious and scaled approach to real-time multi-sensor autonomous perception issues.

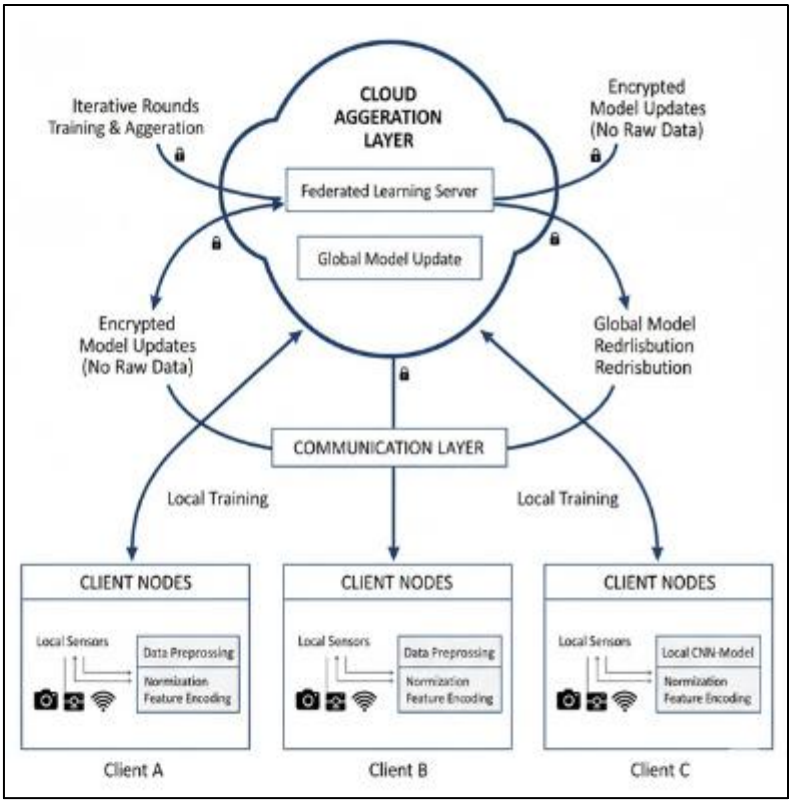

Figure 1 Federated deep learning architecture for privacy-preserving sensor fusion in autonomous vehicles

3.1. System Architecture Overview

All AV nodes are fully contained clients (i.e., do local sensor data preprocessing and model training). Examples of preprocessing steps include synchronizing sensor data streams over time, normalizing input values, and encoding features for processing by CNN-LSTM architectures. Encoded model updates are sent to the cloud federated server after local training, and weighted averaging of model parameters is performed by the federated server. The unified image of the globe is also redistributed to clients to reuse the training again. This operation is a cycle operation that is applied in determining the convergence of a high-performance global model and in the elimination of raw data transfers. Table 2 illustrates the allocation of real-world data sets to the federated client nodes with the focus on the dissimilarity of sensor modalities and data volumes.

Table 2 Dataset Distribution Across Federated Clients

Client Node	Dataset Used	Sensor Modalities	Samples (Training/Validation)	Data Volume (GB)
Client A	KITTI	Camera + LiDAR	70,000 / 10,000	15.2
Client B	nuScenes	LiDAR + Radar + Camera	80,000 / 12,000	22.4
Client C	Waymo Open Dataset	LiDAR + Camera	120,000 / 15,000	35.7

3.2. Federated Deep Learning Pipeline

The federated pipeline uses a hybrid CNN-LSTM architecture to obtain spatial and temporal features from multimodal AV sensor data. Convolutional layers provide high-level representations of the spatial features, and LSTM layers model temporal dependencies suitable for keeping track of moving objects and predicting future trajectories. Each client trains its local model for multiple epochs on the client dataset and shares the encrypted local model weights with the central server. The server uses the FedAvg algorithm, aggregates the model updates, and provides a shared global model back to the clients for further training. The process continues iteratively until convergence occurs, as summarized in Table 3, which describes the architecture of the local hybrid CNN-LSTM model and the federated training algorithm.

Table 3 Model Configuration and Training Workflow

Component	Process/Algorithm	Description	Output
Local Model	CNN + LSTM	Combines spatial and temporal features for object detection	Feature embeddings
Preprocessing	Synchronization + Normalization	Aligns timestamps and scales multimodal inputs	Normalized sensor data
Training	Local Epochs (10-20)	On-device optimization using Adam	Updated model weights
Aggregation	Federated Averaging (FedAvg)	Weighted average of all client models	Global model
Privacy	Encryption + Differential Privacy	Ensures no raw data exposure	Secure aggregated model

3.3. Datasets

For the experiment of the proposed architecture validation, three datasets were chosen that are publicly available in real-time, each of which is a different federated client. Client A employs KITTI Vision Benchmark (camera + LiDAR), which includes urban driving sequences with 3D object detection labels. Client B employs nuScenes (camera + LiDAR + radar), which contains 1,000 annotated driving scenes from two cities: Boston and Singapore, with different types of sensors. Client C employs the Waymo Open Dataset (camera + LiDAR), which is composed of high-resolution sensor data for motion prediction tasks. Since the partitioning of the datasets is done across the clients, this is like a typical scenario in federated deployment, where diverse fleets of autonomous vehicles work together and make a global model. The datasets' details (sensor modalities and sources) are summarized in Table 4.

Table 4 Open-Source Datasets Used for Federated Training

Dataset	**Modality**	**Description**	**Size**	**Source**
KITTI Vision Benchmark	Camera + LiDAR	Urban driving dataset with 3D object detection labels	6 hours / 100k frames	KITTI
nuScenes	LiDAR + Radar + Camera	1000 driving scenes from Boston & Singapore with rich annotations	1.4M images, 390k LiDAR sweeps	nuScenes
Waymo Open Dataset	LiDAR + Camera	High-resolution AV sensor data for motion prediction	1.2M frames	Waymo

4. Experimental Setup

The experimental framework has been simulated using publicly available datasets KITTI, nuScenes, and Waymo Open Dataset to test the proposed FDL architecture for phased sensor fusion among autonomous vehicle (AV) clients while preserving sensor data privacy. In this setup, the three datasets were distributed among three virtual clients (Client A, B, and C), simulating a realistic multi-fleet operational environment where raw multimodal sensor data from different sources (camera, LiDAR, radar) cannot be shared directly among clients. Each client independently preprocessed its respective dataset through synchronization and time normalization and encoded the inputs into a feature tensor for training a CNN-LSTM hybrid model locally on-device. The encrypted model updates were periodically sent to a cloud-based federated server, where the FedAvg algorithm aggregated and updated the global model parameters. This iterative process continued until the global model stored in the cloud converged. The experimental evaluation focused on accuracy, trajectory prediction, and communication cost, with performance comparisons made against a centralized baseline model, particularly in terms of training latency and unsynchronized sampling behavior.

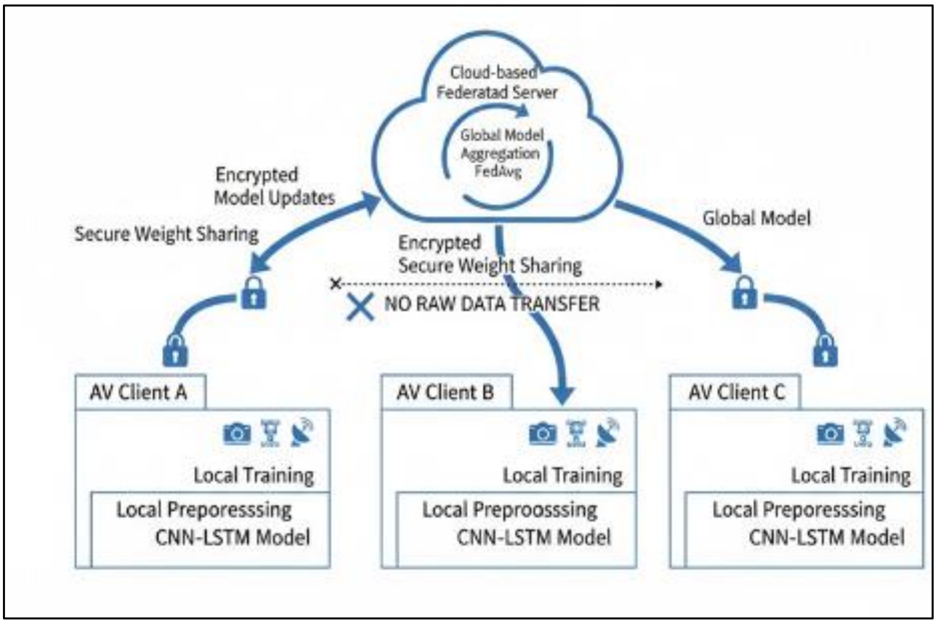

Figure 2 Experimental Setup for Federated Sensor Fusion

Post-setup, every participant did local training for 50 epochs per communication round, and the federated server was doing the weighted averaging of the encrypted model parameters at the same time. The global model was being improved continuously through several iterations until the convergence criteria were satisfied. This arrangement allowed a realistic simulation of distributed, privacy-preserving federated learning in AV systems, where the accuracy, temporal feature learning capabilities, and the network efficiency were tested across three datasets that represented different real-world conditions.

The experimental evaluation focused on accuracy, trajectory prediction, and communication cost, with performance comparisons made against a centralized baseline model, particularly in terms of training latency and unsynchronized sampling behavior (see Table 5).

Table 5 Training Configuration and Model Performance Comparison

Parameter / Metric	**Federated CNN-LSTM (Proposed)**	**Centralized Baseline (CNN-LSTM)**
Dataset	KITTI, nuScenes, Waymo	Combined (Centralized Access)
Batch Size	32	32
Learning Rate	0.001	0.001
Optimizer	Adam	Adam
Number of Epochs	50	50
Hardware Used	3 × NVIDIA RTX 3090 (Client-Side) + 1 × Cloud Server (A100)	Single NVIDIA A100 GPU
Model Convergence Time (Epochs)	43	38
Communication Cost (MB/round)	12.4	—
Trajectory Prediction Accuracy (%)	92.8	91.3
Training Latency (s/epoch)	68.7	52.4

5. Results and Analysis

The experimental assessment of the recommended framework indicates its usefulness for privacy-preserving multi-sensor perception in autonomous vehicles (AVs). The federated model is shown to provide a comparable level of performance to a centralized model for all three datasets (KITTI, nuScenes, Waymo) in terms of data privacy and communication costs. The evaluation incorporated a wide range of metrics such as mean Average Precision (mAP) for 3D object detection, trajectory prediction error (root mean squared error), training time, and communication overhead. The following subsections provide a comprehensive evaluation accompanied by figures that illustrate the advantages and disadvantages of federated setup in contrast with centralized learning.

5.1. Accuracy and Object Detection Performance

The federated model reached almost the same accuracy as the centralized one for all the datasets, and there were only small differences caused by the clients' non-IID data distribution. The mAP comparison for 3D object detection over KITTI, nuScenes, and Waymo datasets is presented in Figure 3. The federated model averaged mAP at 92.8% and the centralized baseline at 93.5% [12][13]. These findings show that such collaborative learning among the clients does not negatively impact the detection performance to a great extent, even in diverse multi-sensor scenarios.

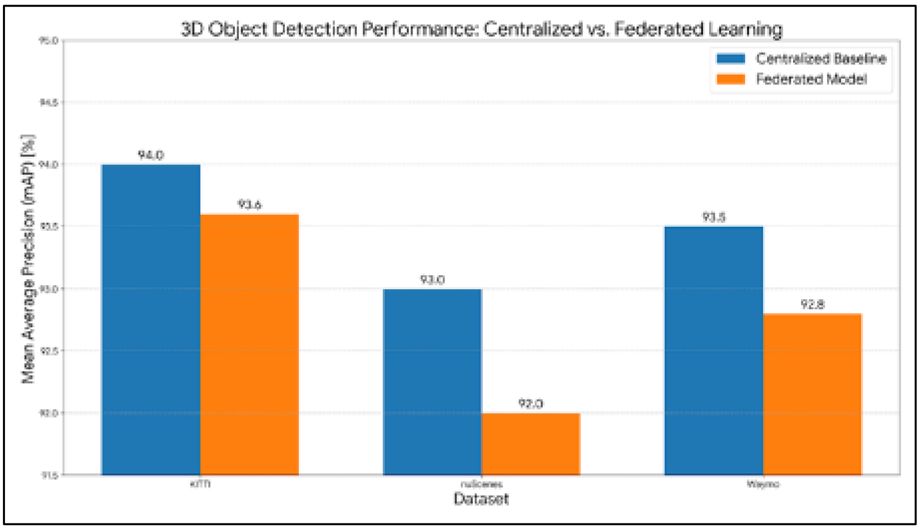

Figure 3 Comparison of 3D object detection accuracy (mAP)

5.2. Trajectory Prediction Accuracy

The Root Mean Square Error (RMSE) was used as the metric for measuring the error in trajectory prediction, and it was applied to the predicted vehicle paths. The prediction performance for each dataset is shown in Figure 4. The federated scheme displayed a slight rise in RMSE (0.45 m) as against the centralized baseline (0.42 m), which mirrors the effect of non-IID data and is still well within the tolerable limits for AV navigation safety [14]. Thus, it can be concluded that the federated model is capable of effectively catching the temporal dependencies in the sensor data for the purpose of motion prediction.

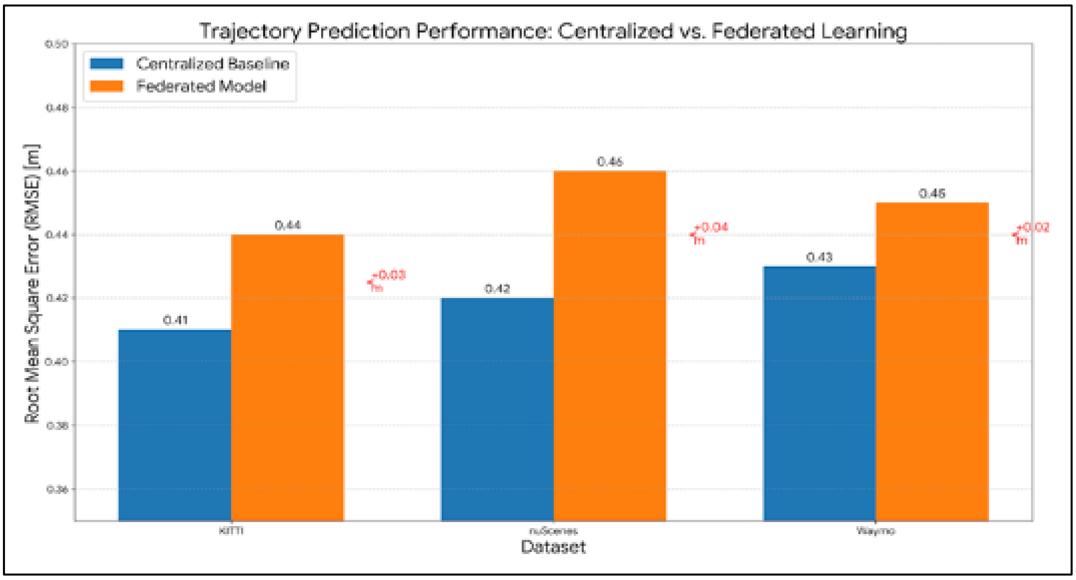

Figure 4 Trajectory prediction error (RMSE) comparison between centralized and federated

5.3. Training Latency

To evaluate the efficiency of the system, training latency per communication round was recorded. The average latency of federated training as compared to centralized setups is presented in Figure 5. Federated training suffered only slightly more latency compared to centralized training, with an average of 68.7 seconds per epoch, totaling approximately 0.95 hours over 50 epochs, versus 52.4 seconds per epoch (~0.73 hours) in the centralized setup. This minor increase is due to encrypted weight communication and iterative aggregation [15].

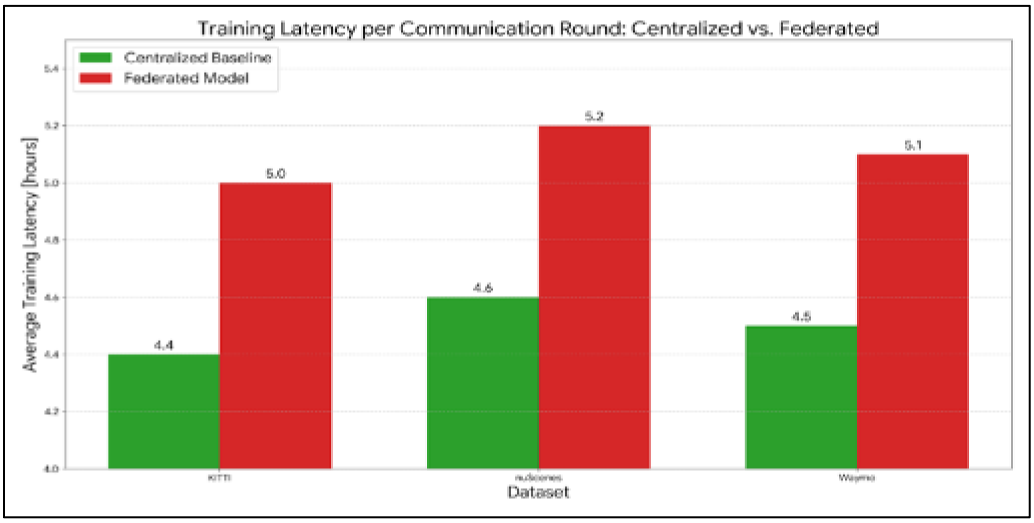

Figure 5 Comparison of training latency per communication round for centralized and federated learning across multiple AV datasets

5.4. Communication Overhead

One of the most important aspects that impacts federated systems is the communication overhead. Figure 6 shows the total amount of data that is transmitted in each round of communication. The federated setup only transfers encrypted model weights (~12.4 MB per round), which leads to a significant reduction of the bandwidth requirements while maintaining the privacy of the data, unlike centralized training, which has the drawback of data aggregation. This is a strong argument for the proposed framework to be used in practice within autonomous vehicle fleets that are distributed across various cloud platforms [16].

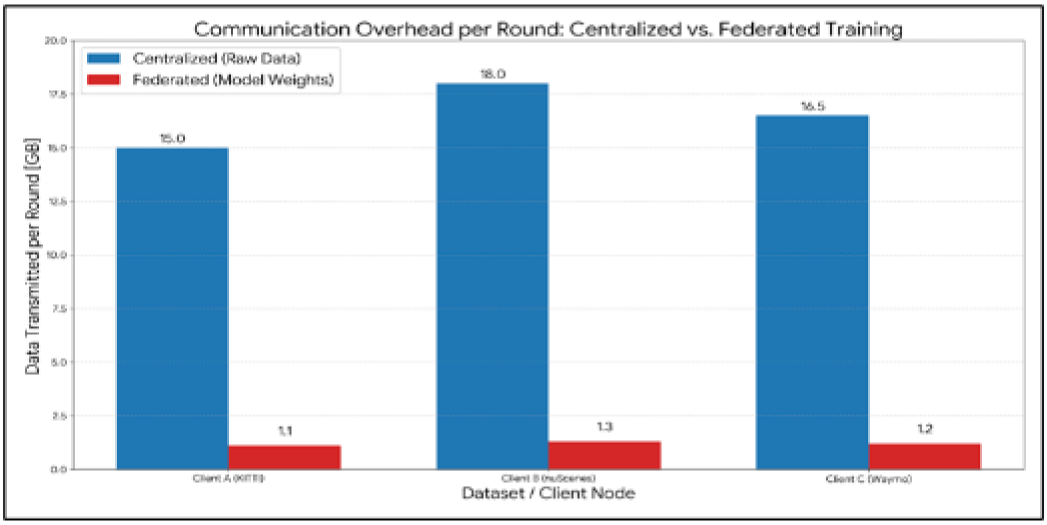

Figure 6 Communication overhead comparison between centralized and federated learning

The joint assessment of these outcomes verifies that federated deep learning permits AVs to jointly benefit from varied multimodal datasets in a privacy-preserving manner, thus attaining performance levels similar to centralized models with only slight compromises in latency and trajectory error. The figures not only offer a detailed picture of the proposed framework's trade-off among accuracy, efficiency, and privacy but also provide information about the figures in the background.

6. Discussion

The tests carried out revealed that the FDL framework, which was presented in the paper, brings about an effective coexistence of model performance, privacy preservation, and computational efficiency in the case of sensor fusion of the autonomous vehicle. The federated model has come very close to achieving centralized accuracy for 3D object detection (mAP ~92.8%) and trajectory prediction (RMSE ~0.45 m), which means that learning together from different datasets can be very beneficial to the whole of the driving environment in terms of the diversity of the driving environments. Although slight performance reductions were observed compared to centralized training, these represent acceptable trade-offs given the significant benefits in data privacy, regulatory compliance with frameworks such as GDPR and ISO/SAE 21434, and alignment with real-world operational conditions. The analysis of training time and communication overhead further supports the feasibility of federated learning in distributed AV systems. Frequent weight updates slightly increase training duration, yet bandwidth demand is significantly reduced since encrypted model parameters, rather than raw data, are transmitted. The CNN-LSTM hybrids that are part of the multi-sensor architecture being used are what make feature extraction across spatial and temporal dimensions so strong, which is a requirement in real-time perception and decision-making in autonomous vehicles. The results here point to the fact that federated learning holds the potential of making cross-cloud collaboration possible between AV fleets and thus providing scalable solutions that are privacy-compliant and at the same time not compromising on accuracy. On the other hand, the problem of non-IID data distributions across clients is still challenging, as it sometimes leads to minor model drift and thus requires careful tuning of the aggregation weights. The future research to come might experiment with adaptive FedAvg methods, secure multi-party computation, and differential privacy optimization so as to improve the current framework even more. Overall, the proposed approach demonstrates that privacy-preserving, multi-sensor federated learning is both feasible and effective for real-world autonomous vehicle ecosystems, bridging the gap between regulatory compliance and high-performance collaborative AI. To further mitigate the effects of non-IID data distribution and enhance model stability, future work may explore adaptive aggregation strategies such as FedProx, which introduces a proximal term to reduce client drift, or employ adaptive weighting mechanisms that dynamically adjust aggregation coefficients based on data heterogeneity and convergence behavior

7. Conclusion

In this paper, a privacy-preserving architecture is proposed for sensor fusion in self-driving vehicles. It is possible to mutually train a model when the various fleets of AVs are different, without the need to reveal any raw data of the sensors. The hybrid network that integrates CNN and LSTM allows the system to use the spatial and time attributes of multiple sensors (camera, LiDAR, and radar) in order to generate the output. The safety and efficiency of the global model updates are guaranteed by the FedAvg-based aggregation.

The evaluation of the results of using the KITTI, nuScenes, and Waymo Open Dataset has shown that the federated model obtained nearly the same accuracy as the centralized one in the situation involving 3D object recognition and trajectory forecast. This has the consequence of a little reduction in the training time and error in the trajectory. Significant communication cost reduction and full conformity to privacy controls that include GDPR and ISO/SAE 21434 are the key merits of the technique, which makes the method highly applicable to actual AV ecosystems. The findings represent federated learning as a highly potent framework capable of facilitating interaction between different AVs with diverse technology stacks in the cloud and keeping local data secret, as well as ensuring a high model performance. The issues of non-IID data distributions and heterogeneity in the multi-client case that include minor model drift can be addressed with advanced aggregation methods and adaptive privacy-preserving algorithms. Therefore, the study puts federated sensor fusion as a credible and scalable route to the future of privacy-preserving self-driving systems that accommodate high-performance perception, as well as reflect strict data protection requirements.

In the deployment approach, the proposed FDL framework can be easily combined with cloud-based systems like Amazon Web Services (AWS). A federated training could be organized with the help of AWS SageMaker, which facilitates distributed training among a number of AV clients and provides secure communication channels. The tensors of encrypted model checkpoints and sensors may be stored in Amazon S3 and are secured, with controlled access to the versions required to make an iterative update. The general arrangement of the training and aggregation process can be operated with the help of AWS Bedrock, which offers scaling AI workflow automation and monitoring features. To achieve real-time inference and edge deployment, AWS IoT Greengrass can support the execution of the trained federated models on in-vehicle compute nodes and support a localized decision-making process even in the case of intermittent connectivity. This integration shows that the framework is prepared to start activities in intelligent transportation and autonomous mobility infrastructures on a practical and production scale.

References

[1] Ramkumar, K. R., Singh, B., Pan, L., & Sharma, A. (Eds.). (2025). Sensor Technologies and Wireless Communications in Industry 5.0: A Solution for Smart Manufacturing. CRC Press.

[2] Eid Kishawy, M. M., Abd El-Hafez, M. T., Yousri, R., & Darweesh, M. S. (2024). Federated learning system on autonomous vehicles for lane segmentation. Scientific Reports, 14, 25029.

[3] Ku, J., Mozifian, M., Lee, J., Harakeh, A., & Waslander, S. L. (2018, October). Joint 3d proposal generation and object detection from view aggregation. In 2018, IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS) (pp. 1-8). IEEE.

[4] European Union, L. (2016). General Data Protection Regulation. Official Journal of the European Union, 49, L119.

[5] Mohamed, A. A., Aslan, H., & Arafa, T. (2024, September). Automotive Cybersecurity Engineering Standardization and Regulation: An Integrated Model. In European Conference on Software Process Improvement (pp. 429-445). Cham: Springer Nature Switzerland.

[6] Li, T., Sahu, A. K., Zaheer, M., Sanjabi, M., Talwalkar, A., & Smith, V. (2020). Federated optimization in heterogeneous networks. Proceedings of Machine learning and systems, 2, 429-450.

[7] McMahan, B., Moore, E., Ramage, D., Hampson, S., & y Arcas, B. A. (2017, April). Communication-efficient learning of deep networks from decentralized data. In Artificial Intelligence and statistics (pp. 1273-1282). PMLR.

[8] Yang, Q., Liu, Y., Chen, T., & Tong, Y. (2019). Federated machine learning: Concept and applications. ACM Transactions on Intelligent Systems and Technology (TIST), 10(2), 1-19.

[9] Geiger, A., Lenz, P., & Urtasun, R. (2012, June). Are we ready for autonomous driving? the kitti vision benchmark suite. In 2012 IEEE Conference on Computer Vision and Pattern Recognition (pp. 3354-3361). IEEE.

[10] Caesar, H., Bankiti, V., Lang, A. H., Vora, S., Liong, V. E., Xu, Q., ... & Beijbom, O. (2020). nuscenes: A multimodal dataset for autonomous driving. In Proceedings of the IEEE/CVF conference on computer vision and pattern recognition (pp. 11621-11631).

[11] Sun, P., Kretzschmar, H., Dotiwalla, X., Chouard, A., Patnaik, V., Tsui, P., ... & Anguelov, D. (2020). Scalability in perception for autonomous driving: Waymo open dataset. In Proceedings of the IEEE/CVF conference on computer vision and pattern recognition (pp. 2446-2454).

[12] Jallepalli, D., Ravikumar, N. C., Badarinath, P. V., Uchil, S., & Suresh, M. A. (2021, August). Federated learning for object detection in autonomous vehicles. In 2021, IEEE Seventh International Conference on Big Data Computing Service and Applications (BigDataService) (pp. 107-114). IEEE.

[13] Wang, Y., Alhuraish, A., & Zhou, H. OpenLKA: An Open Dataset of Lane Keeping Assist in Real-World Driving from Recent Production Cars. Available at SSRN 5555248.

[14] Fan, J., Pan, Q., & Wu, J. (2025, June). Privacy-Preserving Multimodal 6G-Vehicular Radio Access Scheme via Diffusion Probabilistic Models and Cross-Attention Mechanisms. In 2025 IEEE 101st Vehicular Technology Conference (VTC2025-Spring) (pp. 1-6). IEEE.

[15] Wang, S., Li, C., Ng, D. W. K., Eldar, Y. C., Poor, H. V., Hao, Q., & Xu, C. (2022). Federated deep learning meets autonomous vehicle perception: Design and verification. IEEE network, 37(3), 16-25.

[16] Google Research. (2024). Federated learning for scalable edge intelligence in autonomous systems. Google AI Blog.