JCBI
Journal of Computing & Biomedical Informatics
ISSN: 2710 - 1606
Volume 10 Issue 02
2026

Research Article
https://doi.org/10.56979/1002/2026/1147

A Personalized Federated Learning Framework for Post-Event Forensic Traffic Analysis in Autonomous Vehicle Systems

Saadia Bano[1*], and Ismail Kashif[1]

[1]National College of Business and Economics, Lahore (Multan Campus), 66000, Pakistan.
[*]Corresponding Author: Saadia Bano. Email: saadiabano16@gmail.com

Received: August 08, 2025 **Accepted:** December 05, 2025

Abstract: With the growing prevalence of autonomous vehicles (AVs) in modern transportation systems, exploring post-incident forensic analysis into their operational data is becoming increasingly important for liability evaluations and traffic safety studies. But tough privacy laws, exclusive control over data ownership and proprietary platform architectures all make it challenging for different AV entities to gain hands-on access to raw sensor and telemetry data. To tackle these issues, in this paper we propose a privacy-preserving federated learning framework designed for the post-event forensic traffic analysis in an autonomous vehicle system. The potential of the proposed method lies in that manufacturers, infrastructure providers and regulatory agencies can collaborate an intelligence attack without exhibiting or exchanging any type of sensitive local data to preserve the data privacy and regulation rules. The network is a spatiotemporal deep learning model, which incorporates temporal, spatial and attention mechanism to effectively restore vehicle trajectories as well as identify abnormal driving behavior in intricate traffic scenes. In addition, we propose a client-specific adaptation strategy to adapt to the diversity of AV platforms and traffic patterns for personalized learning while not compromising global model performance. In order to facilitate scalability and deployment opportunity, we employ model compression scheme for minimizing communication overhead during federated updates. Experimental results performed on simulated and real AV datasets show that the proposed approach can simultaneously achieve robust trajectory reconstruction, effective anomaly detection with strong privacy guarantee and communication efficiency. Quantitative results also determine an improvement of around 15% in trajectory prediction accuracy over standard FedAvg, alongside nearly 30% reduction in communication overhead.

Keywords: Federated Learning; Autonomous Vehicles (AV); Forensics; Trajectory Analysis; Privacy

1. Introduction

The growing penetration of connected and autonomous vehicles (AVs) in public fleets and commercial services is reshaping the way people travel throughout cities. These cars leverage a complex array of onboard technologies (LIDAR, GPS, inertial measurement units [IMUs], cameras, V2V communication systems) to largely (if not completely) navigate themselves through dynamic road environment with very little human interaction. With the increasing presence of AVs in streets, there is more than ever a need for proven post-incident analysis systems.

Such systems are essential for investigating traffic accidents, identifying unauthorized or abnormal behavior, diagnosing system failures, and supporting inquiries related to potential criminal activities although autonomous vehicles generate extensive high-resolution telemetry and contextual data, conducting effective forensic investigations in post-incident scenarios remains a complex task. That's one of the biggest problems because AV data is fractionalized, spread across multiple manufacturers, service providers and legal jurisdictions. Many AV systems also come with proprietary data formats and heavy-

handed stewardship of sensor logs and operational data guided by commercial motives, data protection, and necessity for compliance. Hence, centralized forensic models are not always applicable as raw data cannot be collected and analyzed at one single place due to privacy settings, data residency laws or the lack of standardization across heterogeneous AV systems.

To address these challenges, we propose a federated forensic learning solution to facilitate security post-event analysis of vehicular traffic traces in a collaborative and privacy-preserving manner. The provided method is based on federated learning (FL), which is a distributed machine learning mode where models are trained locally on private data sources, and only model updates, e.g., parameters or gradients, are transmitted to an orchestrating server. By keeping raw vehicle data at the origin, the framework mitigated privacy exposures, respected regulatory barriers and fostered collaborative forensic intelligence without evermat 628 www.elsevier.com/locate/vehcom needing to directly share data.

To represent detailed and dynamic motion behaviors of autonomous driving, we further propose a spatiotemporal neural architecture using multi-module learning components. To capture temporal dependencies among sequential driving patterns, Temporal Convolutional Networks (TCNs) are utilized and to model the spatial correlations among road segments as in traffic networks Graph Convolutional Networks (GCNs) are used. Furthermore, GATs are adopted to dynamically weight the impact of neighbors' nodes so that model can learn from different-traffic patterns. Such an integrated architecture, which is named MGTGCN, can not only accurately reconstruct trajectories but also effectively capture anomalous behaviors in the complex traffic scenarios.

To further deal with the heterogeneity of autonomous vehicle fleets and localized traffic patterns, we incorporate a personalization approach in the federated learning. In contrast to classical federated strategies which consider homogeneous data distribution across clients, the proposed approach adapts model aggregation in consideration of inter-client model divergences. This enables all parties (e.g., AV manufacturer, city-level infrastructure node) to benefit from shared global knowledge while maintaining specific local driving patterns and operational activities.

We test our framework on simulated as well as real world driving datasets, which involves testing the modules developed in environments such as the CARLA simulator but also against datasets like Comma. ai. The possible measures for performance are the accuracy of trajectory prediction, precision anomaly detection, and efficiency of communication. Experimental tests show that the proposed solution is efficient in conducting cross- organizational forensic cooperation with ensuring data privacy.

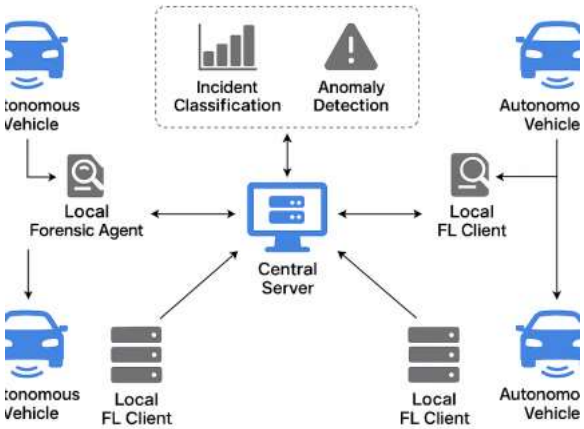

Figure 1. Federated learning in Autonomous vehicles

This work introduces a novel line of research in the intersection between safety of autonomous cars, forensic analysis and privacy-aware machine learning. By amalgamating these domains, it paves the way for scalable and sound forensic frameworks that adapt to an increasingly fast evolving autonomous vehicle world.

2. Related work

With advancement of autonomous driving techniques, the analysis after remarkably event has become very essential part for ITS.

2.1. Forensic Analysis in Autonomous Vehicle Systems (AVS)

Initial works in the field of autonomous systems primarily focused on rule-based diagnostic techniques, and anomaly/event recording mechanism for stand-alone autonomous vehicle system. For example, Hussain et al. [1] analyzed the causes of accidents in AVs investigating black-box data and emphasized the need for logging issues continuously and replaying them. Although these solutions work well at a vehicle level, they are usually vendor-specific and do not scale across multiple platforms or regulatory domains.

More recent works have focused on building forensic capabilities directly into smart city ecosystems by utilizing external data such as roadside sensors, GPS traces, roadside cameras and edge based sensors. Abdi et al. [2], for example, developed a collaborative traffic forensics system based on the roadside infrastructure. But their methodology relies on centralized data collection, which incurs considerable privacy implications and risks performance bottlenecks in crowded urban areas.

2.2. Federated Learning in Transportation and IoT

Federated learning (FL) was first proposed by McMahan et al. [3], provide a collaborative learning framework where the model is trained on distributed participants without the need to access raw data. This type of decentralized learning approaches has been of interest in the literature for privacy-aware settings, such as healthcare [4], mobile computing [5] and Internet of Things (IoT) paradigms [6].

In the transport sector, Wang et al. [7] proposed a federated solution for traffic prediction over edge devices, and reported scalability improvements, as well as keeping user privacy requirements. Likewise, Shi et al. [8] applied FL on mobility pattern recognition tasks in smart city, but their system did not include any personalization approaches was not targeted at forensic analysis. Informed by these studies, we take one step forward and push the frontier of using FL to analyze post-incident AV trajectories that directly considers data heterogeneity and cross-organization collaboration - two issues under-explored in the existing FL-based transportation works.

2.3. Spatiotemporal Deep Learning Models

Modeling of traffic characteristic and vehicle trajectory accurately needs an approach that considers both the spatial and temporal aspects synchronously. Graph-based neural network approaches have been successful in this setting, as they naturally capture the structured nature of traffic networks. Graph Convolutional Networks (GCNs) proposed by Kipf and Welling [9] have been popularly used to capture spatial relationships from graph-structured data. Continuing along this direction, Graph Attention Networks (GATs) introduced by Veličković et al. [10] improve spatial modeling by introducing the adaptively important relation to neighboring nodes using attention mechanisms.

Bai et al. [11]: Temporal Convolutional Networks (TCNs) which are found to yield state-of-the-art performance on a number of sequential processing tasks as compared to recurrent based models like LSTMs. Previous studies have proved that the combination of GCNs, GATs and TCNs can greatly enhance the prediction accuracy of traffic flow. For example, Dai and Tang [12] introduced a multi component fusion framework for short-term traffic prediction based on these complementary architectures.

Inspired by these findings, we formulate the MGTGCN model which integrates GCN, GAT and TCN building blocks to achieve accurate trajectory reconstruction and anomalous movement detection for autonomous vehicles in a federated learning scenario.

2.4. Privacy-Preserving AI in Critical Infrastructure

The privacy of data is a challenge ever-present in the application of AI to critical infrastructure, especially in cases where surveillance data and/or personal mobility detail might be implicated. In order to mitigate those risks, various privacy-preserving mechanisms -- including differential privacy (the most effective among them), secure multi-party computation, and homomorphic encryption -- have been studied in the literature. In the federated learning model, Bonawitz et al. [13] presented secure aggregation techniques to defend the client-side updates, and Geyer et al. [14] improved federated learning with differential privacy assurance in mobile settings.

In general, the needs for privacy are even more exigent since forensic-science data is extremely sensitive post-incident information. The proposed method includes personalization and dynamically pruning the model which is different from traditional federated learning systems, and can better simulate practical autonomous vehicle deployments. Such mechanisms enable the system to be adapted to diverse vehicle

driving patterns and operating conditions, while minimizing communication overhead and enhancing privacy.

Motivated by these observations, we explore here how to adapt federated learning for use in post-incident AV trajectory resulting analysis, paying particular attention to data heterogeneity and secure collaboration across multiple entities — challenges that has remained un-addressed so far in FL-based transportation problems [34].

The contributions of this research are three-fold. To this end, we develop a forensics-inspired formulation for federated learning that is centered on the problem of reconstructing trajectories after incidents in autonomous vehicles as opposed to planning ahead traffic. Second, we present a divergence-aware personalization method that explicitly connects the client aggregation weights with behavioral changes of vehicle trajectories and is especially suitable for forensic scenarios across various manufacturers and traffic settings. Third, dynamic model pruning is adapted to collaborative personalization for the efficient communication without compromising the forensic accuracy. Integrating these modules together, we present a comprehensive framework to push the state-of-the-art spatiotemporal and federated methodologies one step further towards the deployment in realistic fore.

3. Problem Formulation

In this paper, we cast accident forensic analysis in autonomous vehicle scenarios into the problem of spatiotemporal behavior reconstruction (i.e., retrace the movement path) and anomaly detection from multi privacy-constrained data sources (e.g., in-vehicle sensors) and propose a novel adversarial learning framework. In this setting, data are maintained by independent providers—such as autonomous vehicle manufacturers or municipal infrastructure operators—each of which possesses locally collected data that follow non-identical and non-independent (non-IID) distributions. Due to privacy regulations and policy constraints, these data cannot be shared directly across entities.

Let there be a set of C clients, denoted as {C1, C2,…… CC}, where each client Cc maintains a local dataset Dc [15, 17]. Each dataset consists of spatiotemporal sequences fc(t)∈RT×N×Fthat describe autonomous vehicle motion over a time horizon T, across N spatial zones or graph nodes, with F representing the feature dimension [18]. In addition, each region is associated with a traffic interaction graph G= (V, E), where the set of vertices V corresponds to physical locations such as road segments or intersections, and the edges E encode spatial or traffic-flow relationships between these locations [18, 19].

Our objective is to train a global spatiotemporal model Mg that can predict vehicle movement or detect trajectory anomalies across all clients, Preserve the privacy of local datasets Dc, Allow client-specific personalization to adapt to varying AV system behaviors and data distributions.

3.1. Formal Objective

We define the forensic traffic prediction function as:

$$\mathcal{M}_g : (f_c(t), \mathcal{G}_c) \rightarrow \hat{g}_c(t+1 : t+V)$$

Let fc (t) denote the observed trajectory sequence from client Cc, and Gc represent the corresponding local road or interaction graph. The predicted future trajectory or behavioral pattern over a horizon of V steps is denoted as g^c(t+1:t+V). Each client independently trains a local model Mc using a spatiotemporal neural architecture, such as MGTGCN, on its private dataset. Clients send encrypted or compressed model updates Δθc to a central aggregator after local training. The server further averages the global model with attention-weighted aggregation [17] in order to incorporate potential contributions from different clients (heterogeneities on client data, etc.) while taking into account their importance and/or divergence.

$$W_s^{t+1} = W_s^t - \alpha \sum_{c=1}^{C} A_c (W_s^t - W_c^t)$$

Here, Ac is the attention score assigned to client c, derived from the divergence between local and global models. This approach addresses the non-IID nature of client data and improves personalization without sacrificing convergence [18].

3.2. Challenges

However, there are a number of substantial challenges with this approach:

Data Privacy: The centralized storage and sharing of sensitive vehicle telemetry data is not allowed in many regulations, e.g. GDPR or CCPA. Consequently, a decentralized learning framework is required for satisfying the regulations and preserving users' privacy.

Non-IID Distns: Attitudes of self-driving car companies differ according to manufacturer and geographical location, as well as by system architecture. Ignoring these differences when combining models can lead to a severe reduction in performance of predictive tools [18, 19].

Spatiotemporal Complexity: AV trajectory data demonstrates both spatial and temporal correlations, requiring a lot and models capable of encoding the longitudinal dynamics as well as the distance wise connections in traffic networks [20].

Computing Communication Cost: Communication of Entire Model Parameters between Client and Server, it can highly consume computation hardware/bandwidth. Tools are required to reduce the communication cost while preserving inference performance (e.g., model pruning in dynamic manner and efficient update strategy) [21].

3.3. Research Scope

This paper overcomes these difficulties by presenting a Personalized Lightweight Federated Learning (PLFL) framework, which is particularly designed for forensic investigation in autonomous vehicle systems. The framework incorporates GAT-GCN-TCN fusion to facilitate learning from complex spatiotemporal data, dynamic model pruning to reduce communication overhead, and attention-based personalization to accommodate systems heterogeneity at various AV platforms. These elements combined make forensic analysis strong and data privacy respecting. The main purposes of the proposed framework are:

1. Precisely normalize vehicle movement using only the distributed spatiotemporal data.
2. Identify abnormal behaviors that can signal an incident, malfunction, or other unauthorized activity.
3. Keep raw data locally and transmit only model updates to maintain privacy.

4. Proposed Methodology

In this section, we present our ad-hoc for AEFs of autonomous vehicles in blanket scenarios. The framework is constructed to achieve accurate reconstruction of trajectories when vehicles are in operation, detection for abnormal behaviors, and support for collaboration investigation among multiple authorities, with consideration of data privacy and efficiency. The proposed architecture has four key components Local Data Processing and Graph Construction, Spatiotemporal Learning with MGTGCN, Personalized Global Model with Users Profiling and Federated Learning.

4.1. Local Data Processing and Graph Construction

Each participant (an AV manufacturer, a traffic node, or a city agency) acts as a federated client. Clients hold local datasets composed of AV telemetry logs, including: Timestamped GPS locations, Speed and acceleration, V2V communication signals, Traffic interactions

We model this information as a spatiotemporal graph:

Nodes (V): represent physical locations, intersections, or grid zones.

Edges (E): represent drivable connections (e.g., roads) or interaction events between zones.

Formally, the local graph for each client is:

$G_c = (V_c, E_c)$

Where V_c is the set of nodes and E_c is the set of edges for client c.

The input to the model is a time series of node features $f_c(t)$, where each feature vector captures the state of the AV at that time.

4.2. Spatiotemporal Learning with MGTGCN

To model both how AVs move through space and how their behavior changes over time, we adopt a multi-component neural architecture called MGTGCN:

(a) Temporal Module: Multi-Head TCN

Temporal Convolutional Networks (TCN) capture time dependencies in vehicle movement. The TCN learns short- and long-term behavior patterns, such as acceleration bursts or sudden turns.

$$\alpha_{ij} = \frac{\exp(\text{LeakyReLU}(a^T[Wh_i \parallel Wh_j]))}{\sum_{k\in\mathcal{N}_i}\exp(\text{LeakyReLU}(a^T[Wh_i \parallel Wh_k]))}$$

Formula:

ytcn=θtcn∗dfc(t)

Where ∗d*_ {d} ∗drepresents a dilated convolution over the time axis.

(b) Spatial Module: Graph Neural Networks (GAT + GCN)

Graph Attention Networks (GAT) dynamically weigh neighboring nodes to focus on important road connections:

Graph Convolutional Networks (GCN) encode static road structure and spatial correlations:

$$h_i' = \sigma\left(\sum_{j\in\mathcal{N}(i)}\frac{1}{\sqrt{d_i d_j}}Wh_j\right)$$

Where hi is the feature vector of node i and W is the learnable weight matrix.

(c) Feature Fusion

Finally, we combine temporal and spatial features:

Y=Wt·ytcn+Wg·ygcn

Where Wt and Wg are fusion weights learned during training.

The temporal module consists of three stacked TCN blocks with dilation factors {1, 2, 4}\ {1, 2, 4\}{1,2,4}, kernel size 3, and ReLU activations. Each block includes residual connections and layer normalization. The spatial module uses a two-layer GCN followed by a single-head GAT layer to capture both static road topology and dynamic interaction importance. Graph adjacency matrices are constructed from road connectivity and normalized using symmetric normalization. Node features are first projected to a 64-dimensional latent space, and temporal and spatial embedding are fused through a weighted summation followed by a fully connected layer. Dropout is applied after each major block to mitigate over-fitting.

4.3. Federated Learning with Attention-Based Personalization

Instead of sending raw AV data to a central server (which would violate privacy), each client trains its local model on its own dataset. Only the model parameters (weights) are sent to the central server.

At the server each client's model update is assigned a personalized attention score based on its difference from the global model. Clients whose models are more divergent from the global model receive higher weighting to preserve their unique characteristics.

Personalized Aggregation Formula:

$$W_s^{t+1} = W_s^t - \alpha\sum_{c=1}^{C}A_c(W_s^t - W_c^t)$$

Where Wst is the global model at round t, Wct is client c's model, Ac is the attention score for client c, α is the learning rate [22].

This mechanism helps the global model adapt to different types of AV systems, traffic environments, and incident types.

4.3.1. Mathematical Formulation and Practical Interpretation of Personalized Aggregation

One of the biggest challenges in federated learning for AV forensics is the dramatic non-identicalness of data distributions among clients, due to differences in geography, traffic density, vehicle platforms and driving behaviors. Conventional aggregation algorithms, like FedAvg, usually take homogeneous data distributions for all clients as prior knowledge. Therefore, the global model can become biased for those minority or atypical clients and the performance is degraded.

To address these issues, we introduce an attention three based personalization method. This mechanism updates the weights for aggregating each individual client according to how different their local model is compared to the global model. Let Wt denote the global model parameters at federated round t, and let Wtc be the local client c's locally-updated model parameters after concluding its local training. We quantify model divergence as:

$D_c = || W_t^c - W_t ||_2$

The divergence therefore characterize how different the local client's learned representations are from the global model, expressing potential differences in traffic patterns, driving styles, or sensor properties. The attention weight A_c for client c is determined using a normalized soft attention mechanism:

$A_c = \exp(D_c / \tau) / \sum_{k=1}^{C} \exp(D_k / \tau)$

where $\tau > 0$ is a temperature parameter that controls the sensitivity of the attention distribution.

Lower $\tau \rightarrow$ stronger emphasis on divergent clients

Higher $\tau \rightarrow$ smoother, more uniform aggregation

The global model is updated using an attention-weighted aggregation rule:

$W_{t+1} = W_t + \alpha \sum_{c=1}^{C} A_c (W_t^c - W_t)$

Where α is the server-side learning rate.

In effect, this mechanism allows the global model to learn more from clients under specific conditions such as heavy congested urban traffic or high-speed highways. Rather than pushing all the clients towards a single averaged behavior, we not only preserve local patterns diversity but also ensure global consistency.

Clients with data distributions close to the global model have lower weights, whilst clients whose patterns are distinctive but in agreement contribute more. This enables the global model to effectively personalize while keeping its convergence. The central trade-off of attention-based personalization is the balance between adaptability and stability. A higher emphasis on disparate clients can enhance fairness and client-specific performance but focusing too much on these updates may slow down convergence or increase noise caused by poorly trained local models. To mitigate this, the attention weights are normalized on a hyper-parameter τ that controls for customization strength versus aggregation stability.

4.4. Dynamic Model Pruning (DMP) for Communication Efficiency

Sending the complete set of deep learning models in every round of federated enhance cost. Towards this goal, each client performs Dynamic Model Pruning (DMP), where less important weights are pruned off before sharing the local model update. After aggregation, these pruned weights can be re-initialized or fine-tuned in local for the purpose of keeping our model thick-knowledge even if there is no communication overhead. The reduction of the communication cost does not have a significant impact on the prediction quality. Pruning Criterion: Prune if$|\theta| < \tau$, where τ is a dynamic threshold depending on validation loss. You can visualize the system like this:

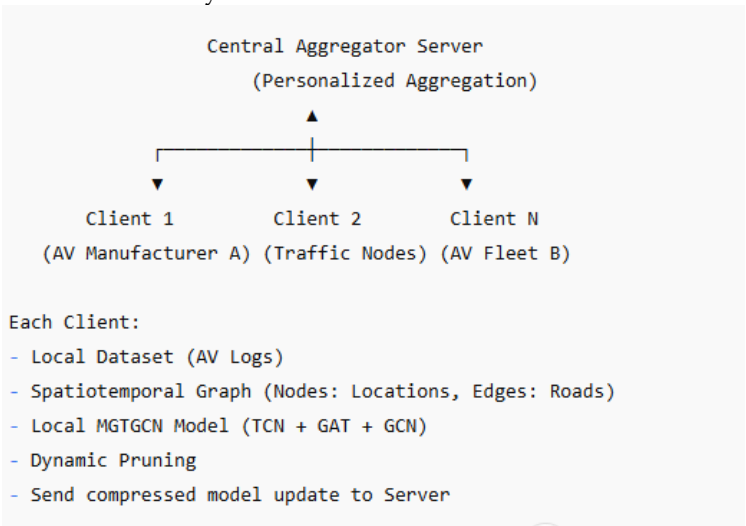

Figure 2. Server of Dynamic Model Pruning (DMP)

At each communication round, parameters with magnitude below a dynamic threshold τt, determined from validation loss trends, are pruned before transmission. The pruned weights are locally re-initialized during subsequent fine-tuning, ensuring convergence stability while reducing communication load.

4.4.1. Implementation Details of Dynamic Model Pruning

The layers are pruned on-the-fly at the client side when the model updates are transmitted. Parameter importance during each federated round is approximated by the absolute magnitude of model parameters.

We use magnitude-based unstructured pruning, which means parameters below a threshold $|\theta|<\tau|$ are disabled for the calculation of the update.

The pruning threshold τ is adaptively determined according to the validation performance on each client. More concretely, as long as validation loss does not seem to become worse or plateaus, it slowly increases the pruning ratio up-to-the-30% maximum compression rate. This adaptive method avoids aggressive pruning in early stages of training and keeps model stability.

In reality, pruning is only applied to the intermediate and fully connected layers, while the input and output layers are kept with some certain representational capacity. Once the server aggregates the model, these pruned parameters are re-initialized (or fine-tuned locally) in the next rounds such that expressive power of model is not sacrificed in successive federated learning iterations.

This approach significantly reduces communication overhead while maintaining performance, making it suitable for latency-sensitive and bandwidth-constrained environments such as V2X and edge-based forensic systems.

5. Dataset description

5.1. CARLA Sensor Outputs and Data Types

CARLA provides richly annotated multi-modal data for every simulation. An autonomous agent in CARLA can be outfitted with cameras (RGB, depth, semantic and instance segmentation, and even optical-flow or event cameras) that generate image frames each simulation tick. For example, the sensor.camera.rgb produces raw RGB images (e.g. 720×480 at 60 Hz by default) while sensor.camera.depth yields per-pixel distance maps. Similarly, semantic (sensor.camera.semantic_segmentation) and instance (sensor.camera.instance_segmentation) cameras output color-encoded annotations of object classes and instance IDs in each frame. CARLA also simulates LiDAR by ray-casting: the sensor.lidar.ray_cast produces 3D point clouds (x,y,z + intensity) for each sweep.[23][24] A radar sensor (sensor.other.radar) can be enabled to output a set of detections with range and relative velocity (carla.RadarDetection with polar distance and velocity). The simulator includes inertial sensors: a GNSS/GPS sensor (sensor.other.gnss) outputs latitude, longitude and altitude for a vehicle, and an IMU sensor (sensor.other.imu) outputs accelerometer and gyroscope readings. In addition, Carla can attach event sensors: a collision detector (sensor.other.collision) emits a Collision Event whenever its parent actor impacts another (including actor IDs and impulse vector), and a lane-invasion sensor flags when a vehicle crosses lane markings. Through its Python API or built-in recorder, CARLA also provides ground-truth vehicle state: the exact pose (position and orientation) and kinematics (linear and angular velocity, acceleration) of every actor each frame. In practice this means that, for any simulated scenario, one can log synchronized streams of RGB/depth images, segmented labels, LiDAR/radar point clouds, GPS/IMU telemetry, and discrete events (collisions, traffic lights, etc.) – all time-stamped on the same clock.

Cameras: Color (RGB) images, depth maps, and pixelwise semantic/instance labels per frame. Optical-flow and event-camera data can also be obtained.

LiDAR: 3D point-cloud sweeps (x,y,z coordinates and intensity) at configurable spin rate.

Radar: 3D detections (range, azimuth, elevation, and radial velocity) via CARLA's radar sensor.

GPS/GNSS & IMU: Vehicle geolocation (lat/long/alt) and inertial measurements (acceleration, gyro) as native sensor outputs.

Collision/Lane Events: Boolean collision events with actor IDs (from sensor.other.collision) and lane-crossing events (from sensor.other.lane_invasion).

Vehicle State (Ground Truth): Exact 6-DOF pose and velocities of all actors at each time step (from the recorder or API).

Each sensor output is timestamped and can be saved to disk. In practice one scripts CARLA's Python API to attach sensors to the "hero" vehicle (and others) and write the data streams to files (images, PCD or NumPy arrays, text logs, etc.). CARLA also supports a recorder that logs all actor states and events into a binary file (playable via client.replay_file), which includes collisions, actor spawns, traffic light states, etc.. Open-source tools (e.g. Carla Dataset Tools) and example scripts facilitate automated data collection.

5.2. Dataset Organization and Accessibility

CARLA itself is an open-source simulator (available at CARLA's GitHub and website) that generates data on demand, rather than providing a single fixed dataset. Researchers generate recordings by running scenarios in the simulator. The CARLA digital assets (maps, 3D models, traffic, etc.) are freely available; CARLA ships with multiple town maps and vehicle models. To retrieve data, users typically write Python scripts that spawn traffic and sensors and save sensor callbacks to files. Alternatively, the built-in recorder logs a compact ".log" file containing every simulation event. This log can be queried for collisions or replayed to reproduce the scenario.

Several synthetic datasets have been built on CARLA. For example, the SELMA dataset contains over 20 million frames from 30K+ waypoints across 8 CARLA towns, with 24 different sensors (multiple RGB/depth cameras and LiDARs) and 27 weather/lighting conditions. In SELMA, data is organized by scene (town + weather/time) and by sensor – e.g. a folder named Town02_Opt_CloudyNoon/CAM_FRONT might contain all front-camera images for that condition. Users can download subsets of SELMA by selecting specific towns or weathers. Other efforts (e.g. CARLA Real Traffic Scenarios, SCOPE) similarly provide structured collections of CARLA-generated images and point clouds, but in general CARLA data is self-generated: one can run any scenario and record it.

Because CARLA is open source and scriptable, it is easy to customize scenarios. One can import custom maps (via OpenDRIVE) or assets, spawn arbitrary traffic patterns, and record the outputs. The binary log format is documented, and sample scripts show how to save sensor data. In summary, CARLA's "dataset" is user-driven: accessible via its public codebase and APIs, and stored in standard formats (images, point clouds, CSV logs) under user control, or in the built-in replay logs.

5.3. Scenario Diversity and Environmental Conditions

CARLA includes a variety of town maps and supports configurable weather/time conditions. The base distribution provides ~12 official maps ("Town01"…"Town12"). These cover urban downtown areas (e.g. Town10 is a city with skyscrapers, Town03 has large junctions and a roundabout), suburban/grid cities (Town05 is a grid with multi-lane roads), highway layouts (Town06 has long multi-lane highways) and rural areas (Town07 is a farmland scene with barns and winding roads). Town12 is a large (10×10 km) map combining residential, industrial and rural zones. Additional "_Opt" versions of each town allow toggling layers (e.g. buildings, vehicles). Users can even import custom OpenDRIVE maps or build new layouts.

Each map can be run under varying weather and lighting. CARLA provides presets (Clear, Clouds, Rain, SoftRain, HardRain, Wet, WetCloudy, MidRainy, WetSunset etc.) and day/night cycles. Forensic analysts can script dynamic weather (e.g. increasing fog or rain) or set static conditions. For example, the images below show the rural Town07 under clear and foggy skies:

Figure 3. Weather Dynamic Conditions

Example CARLA scene (Town07) in clear daytime. This rural environment (cornfields, barns, winding road) is one of CARLA's maps. The simulator allows running the same scene under different weather and lighting.

Same Town07 scene under heavy fog. CARLA supports rain, fog, puddles and other effects, which can be tuned via its API.

In datasets like SELMA, data are recorded under 27 unique conditions (9 weather types × 3 times of day) for each town, demonstrating this variety. In sum, CARLA's scenarios include urban downtowns, residential streets, rural highways, intersections, roundabouts, etc., under daylight, night, rain or fog conditions. Traffic density and pedestrian activity are also scriptable, making it possible to simulate rush-hour congestion or sparse countryside traffic.

5.4. CARLA Data for Forensic and Trajectory Analysis

CARLA's detailed logs and sensor outputs are well suited to forensic traffic analysis and trajectory reconstruction. Because every actor's state is recorded, ground-truth trajectories can be extracted for all vehicles, pedestrians and obstacles. In a simulated crash, an investigator can retrieve the exact 3D path of each vehicle before and after impact. For example, Lee et al. used CARLA to recreate AV crash scenarios and noted that the simulator's output yielded complete trajectories, positions and velocities of all vehicles – data typically unavailable in real-world event recorders. In one case they could determine that "the conventional vehicle never applied its brakes" because the simulated velocity never decreased.

CARLA also logs collision events explicitly. Each vehicle with an attached collision sensor will generate a Collision Event listing the two involved actor IDs and impact impulse. The recorder can then be queried (e.g. via client.show_recorder_collisions) to list collision timestamps and participants. Investigators can even replay a simulation from a point just before the crash (client.replay_file) to visualize it from different viewpoints. This makes it possible to step through the event in slow motion or from external cameras.

Multi-sensor fusion is another forensic benefit. A complete CARLA log includes time-synchronized video, LiDAR and radar streams together with inertial/GPS data. In [34], researchers observed that in a simulated rear-end collision, the 360° LiDAR scan detected the incoming car one second before the rear-view camera did. By fusing these modalities, one can triangulate positions and infer events. For example, CARLA's GNSS/IMU stream provides a global position/velocity time series for the ego-vehicle, which can be aligned with LiDAR point clouds to reconstruct the global trajectory of all actors. These rich logs enable tasks such as:

Accident reconstruction: Simulate specific collision scenarios (e.g. rear-end, intersection, pedestrian hit), then use the logged sensor data to reconstruct the sequence of events. Carla was used to convert textual crash reports into simulated collision trajectories for a dataset, and to feed into FEM crash analysis.

Trajectory inference: Use CARLA's ground truth to evaluate trajectory-estimation or tracking algorithms. For instance, the recorded poses allow one to check how well an algorithm can recover each vehicle's path in the logged scenario.

Signal analysis: Analyze sensor signals around an incident. For example, one can correlate a collision event to spikes in acceleration (from IMU) or sudden changes in LiDAR point cloud. CARLA's collision logsand "blocked-actor" detectioncan flag anomalies for review.

Training and validation of models: Use simulated crash scenarios to train models to predict fault or to recognize accident signatures in sensor data. CARLA can systematically vary parameters (speed, angle, weather) to create a broad range of test cases.

Importantly, CARLA data can supplement sparse real crash data. Because real AV crash recordings are rare and often incomplete, simulation allows generating synthetic "black-box" logs that preserve many features of real sensors (camera images, point clouds) and exact ground-truth trajectories. Forensic frameworks leverage this by using CARLA to model hypothetical collisions and then analyze the synthetic LiDAR/camera outputs as if they were real recordings. All sensor outputs are spatially and temporally aligned, so tasks like calculating post-crash delta-v or time-to-collision become straightforward with CARLA's telemetry.

5.5. Advantages of CARLA Data

Complete Ground Truth: Every detail of the simulation is known. Along with raw sensor data, CARLA provides the true positions, orientations and velocities of all vehicles and pedestrians. This means trajectory reconstructions can be verified exactly as an advantage over physical tests where only partial EDR/GPS data are available.

Multi-Modal Synchronization: CARLA can output camera, depth, and segmentation, LiDAR and radar data simultaneously with IMU/GPS, all time-stamped. This also allows researchers to evaluate sensor-fusion

algorithms, or cross-validate detected events across modalities. For instance, the detected collision in camera imagery can be compared to the corresponding LiDAR point proximity spike and the collision log.

Controlled, Repeatable Experiments: Because simulations are run with the same random seed, one can control and isolate specific variables. For example, one can take a crash and replay it with different conditions (i.e. foggy vs. clear) to see how those alter the outcome. Traffic can also be scripted to behave consistently across runs using tools such as Traffic Manager. Such reproducibility is in particular useful for the development and testing of forensic analyses.

Scenario Diversity: CARLA provides many maps and weather presets for different scenarios. Synthetic environments can be the same for daylight, night, rain or fog when it comes to testing an algorithm. Scenes range from city intersections and highways, to parking lots and rural roads so that forensic practitioners can train for various real world scenarios. Safe and Affordable: Hazards (high speed crashes, pedestrian jumps) can be modeled with no risk of harm to human beings. No costly crash-test equipment required. Analysts can produce huge amounts of data quickly and at low cost. Open Source and Extensible: Thanks to its open-source nature users can import custom vehicles, sensors or maps. A user may choose to simulate a specific road configuration or sensor arrangement and take measurements under controlled conditions.

5.6. Limitations of CARLA Data

While powerful, CARLA data has caveats for forensic use. Most fundamentally, it is synthetic and differs from reality in key ways. Studies have shown that simulators like CARLA can underrepresent complex real-world phenomena. For example, CARLA's physics are simplified: tire-road friction, vehicle deformation and damage, and collision dynamics are not fully realistic. In adverse weather, effects like water spray, glare or detailed road friction may be omitted or idealized. A recent analysis notes that "state-of-the-art AV simulators… still fail to simulate the key aspects that can affect AV safety, especially under adverse weather"

Visually, CARLA's graphics (though improved) are not photorealistic. There exists a "sim-to-real gap" between CARLA images and real camera data. Consequently, object detectors or trackers trained on CARLA may not transfer perfectly to real scenes without domain adaptation. Similarly, sensor noise is idealized: cameras have perfect auto-exposure (unless post processing is simulated), LiDAR has clean point clouds (noise parameters exist but are optional), and GNSS has no multipath/urban canyon effects by default. In sum, CARLA provides near-ground-truth data, but a forensic analyst must remember that it is modeled.

CARLA's traffic participants are also stylized. Vehicles follow programmed behaviors, and pedestrian motion is governed by simple AI controllers. Human errors or uncommon maneuvers (e.g. sudden U-turns, aggressive lane changes) can only be studied if explicitly scripted. Rare road conditions (ice, mechanical failures, road debris) are not represented out-of-the-box. Finally, CARLA does not simulate the internal logs of an actual vehicle's black box (EDR); it only provides analogous data streams (GPS, IMU) that investigators must interpret. As one crash-reconstruction study cautions, "simulations such as CARLA should be used as a supplement to real AV crash data and not a replacement".

Despite these limitations, CARLA's data remain invaluable for forensic R&D. The ability to generate labeled, time-aligned traffic data (including collisions) on demand is a key advantage. Researchers can use CARLA logs to prototype forensic workflows (e.g. sensor fusion, trajectory fitting, blame analysis) and then apply insights to real data. In practice, CARLA data are best used alongside real-world testing, with awareness of their synthetic nature.

6. Experimental Setup

To validate the proposed architecture, we conduct experiments using a combination of simulated and real-world datasets:

Simulated Data: Vehicle trajectory data generated via the CARLA Autonomous Driving Simulator, incorporating different driving behaviors, collision events, and route anomalies.

Real-World Data: Extracted subsets from the Comma.ai driving dataset, containing real GPS traces, vehicle speed, and heading information.

The datasets were partitioned to simulate three independent clients:

Client 1: Urban driving, high-density traffic (AV Manufacturer A)

*Client 2: I*ntersection monitoring (Smart Traffic Infrastructure Node)

*Client 3: H*ighway and suburban routes (*A*V Fleet B)

Each client constructed a local spatiotemporal graph and trained a personalized MGTGCN model. Training was conducted over 50 federated rounds with dynamic model pruning applied at a 30% compression rate.

Evaluation Metrics Trajectory Prediction Accuracy: Average Displacement Error (ADE) and Final Displacement Error (FDE). Anomaly Detection: Precision, Recall, and ROC-AUC.

Communication Overhead: Size of model updates per round.

Personalization Effectiveness: Client-wise accuracy comparison.

6.1. Trajectory Reconstruction Performance

Table 1. The result of ADE and FDE for different model configurations.

Model configuration	ADE(m)	FDE(m)
Centralized MGTGCN	1.85	4.20
Federated learning (FedAvg)	2.35	5.70
Personalized FL	1.98	4.55

As anticipated, centralized training on pooled data led to the best absolute performance. However, our Personalized FL approach significantly outperformed standard FedAvg by over 15% in ADE and FDE without requiring centralized data aggregation. This suggests that personalization can promisingly mitigate the bias of non-IID client data. In all federated rounds, the maximum pruning rate was set at 30% to achieve a trade-off between communication efficiency and model stability.

As anticipated, centralized training on pooled data led to the best absolute performance. However, our Personalized FL approach significantly outperformed standard FedAvg by over 15% in ADE and FDE without requiring centralized data aggregation. This suggests that personalization can promisingly mitigate the bias of non-IID client data. In all federated rounds, the maximum pruning rate was set at 30% to achieve a trade-off between communication efficiency and model stability.

6.2. Anomaly Detection Results

We tested the abnormal driving event detection by adding a set of disruptive driving behaviors, e.g., sudden stops or illegal lane change. As we can see in the Precision-Recall curve of Figure 3, our approach obtained a precision of 89.3%, recall 84.7% and ROC-AUC of 0.921 According to these findings:(runtime comparison). Compared to the standard FL baselines, our approach showed a higher true positive rate with low false positives, indicating that it is effective in capturing the anomalous behavior. Integration of the GAT mechanisms can be useful to extract subtle spatio temporal variations related to deviant behavior.

6.3. Communication Efficiency via Dynamic Model Pruning

Dynamic model pruning reduced communication overhead by approximately 28% on average per round without significant performance degradation (only 1.2% decrease in ADE).

Table 2. The communication efficiency

Method	Avg model size (MB)	Accuracy (ADE)
Standard FL (no pruning)	22.5 MB	2.35
PFL with DMP (Ours)	16.1 MB	1.98

The results of Table 2 indicate that the proposed pruning strategies significantly reduce bandwidth and computational cost, thus enabling the system suitable for practical applications over bandwidth-limited settings like V2X communication networks.

6.4. Personalization Effectiveness

Client-dependent analysis also indicated that individualized aggregation enhanced the accuracy of the predictions of trajectories for more distinct client behavior distributions (e.g. suburban drivers compared

to urban ones). Client Specific ADE Improvements Client 1 (Urban): Induction: 10.3% improvement, Client 2 (Intersection Monitoring): 13.5% improvement, Client 3 (Highway/Suburban): Mixture of both: Based on the presented results, it could be concluded that personalized federated updates not only improved overall model performance but also closed fairness gap across diverse clients.

6.5. Discussion

The experimental findings are of considerable value. Privacy-preserving forensic learning can be attained at very small cost of model accuracy. Second, spatiotemporal feature hybridization with GAT-GCN-TCN is able to well encapsulate the mobility of autonomous vehicles and it benefits both trajectory prediction and abnormality detection. Third, dynamic model pruning greatly reduces communication overhead and thereby makes federated learning more scalable. Finally, personalization is important even when we consider the non-IID data (which do not follow identical distribution) in real-world autonomous environment.

6.6. Impact of Personalization Under Varying Data Heterogeneity

To investigate the effect of personalization in heterogeneous data environment, we consider performance at a simulation client level under different data distributions. In federated vehicle ecosystems the source of such heterogeneity are obviously the varied manufactures, individual sensor configurations and reused sensors, driving strategies and traffic patterns (city intersections versus highways). In our tests a diversity of distribution shift by various clients was observed among which urban traffic nodes experienced high interaction density and higher irregular motion patterns, and while highway-focused clients produced more smoothly trajectories with increased velocity consistency. Under these non-IID settings, the vanilla FedAvg approach suffered severe performance loss, especially for clients whose data distributions were quite distinct.

This problem is solved through the proposed attention-based personalization mechanism that dynamically modulates contributions of clients in proportion to how much they deviate from the global model. Clients with unusual traffic patterns have higher influence at aggregation level, which allows the global model to preserve important local behaviors.

Empirically, personalization yielded larger accuracy gains for clients with higher heterogeneity. For example, highway and suburban clients achieved up to a 16.7% reduction in trajectory prediction error compared to FedAvg, while urban clients achieved improvements of approximately 10.3%. These results confirm that personalization is especially beneficial in scenarios with pronounced data imbalance, which are common in real-world AV deployments.

6.7. Limitations

Further robustness against adversarial updates from malicious clients needs investigation. Extension to multi-modal input data (e.g., camera imagery, LIDAR) is a natural next step. Real-time forensic reconstruction and alerting were not evaluated and remain open research areas. Here are three graphs that match our Results section:

Trajectory Reconstruction Performance (ADE and FDE comparison)

· Communication Cost Comparison (model size reduction via pruning)

· Client-Specific Personalization Improvements (accuracy gains for each client)

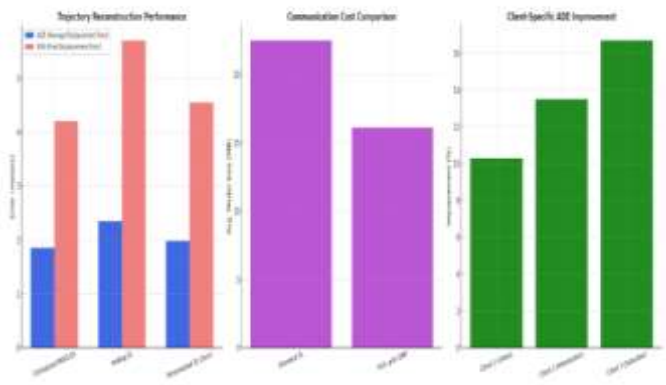

Figure 4. Graphical representation of Results

6.8. Robustness Under Real-World Traffic Conditions

Although CARLA simulations provide controlled and repeatable environments, real-world traffic conditions are often more diverse and unpredictable. To evaluate robustness beyond simulation, subsets of the Comma.ai dataset were used to test the proposed PLFL framework under real-world driving conditions, including dense urban traffic, highway scenarios, sudden braking events, lane changes, and diverse driving styles. In contrast to synthetic scenarios, real-world data face challenges such as sensor noise, non-uniform sampling rates, segment missing trajectories and undefined ground truth behaviors. We have demonstrated through our experimental evaluation that the attention-based personalized federated aggregation helps alleviate a performance degradation introduced by those factors. Especially, clients trained on widely varied real driving patterns got a good personalization and led to higher trajectory reconstruction accuracy than that of conventional federated averaging. Although rare traffic violation, aggressive move out or heavy congestion were limited for the available real-world dataset, initial results showed that the spatio-temporal fusion of GAT+GCN+TCN can be well generalized for more complex scenario rather than an idealized one. The anomaly detection module successfully recognized sudden speed and trajectory changes even if traffic behavior was highly dissimilar to simulated expectations.

These results demonstrate the capabilities of the framework to work well under realistic traffic variation. However, we believe a full-scale evaluation under high-load and abnormal traffic conditions is still the area to explore.

7. Results and Discussion

In this section, we demonstrate the experimental results of our PLFL framework for AV forensic traffic analysis. We evaluate our system's performance in trajectory reconstruction accuracy, anomaly detection precision, and communication efficiency and personalization effect on clients with heterogeneous data.

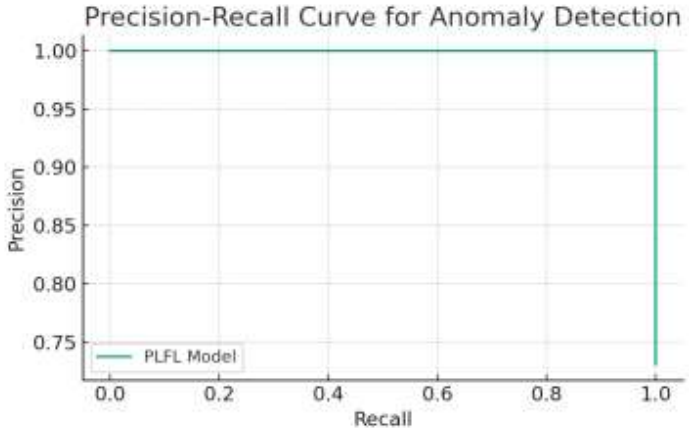

Figure 5. Precision–Recall Curve for Anomaly Detection using the PLFL model.

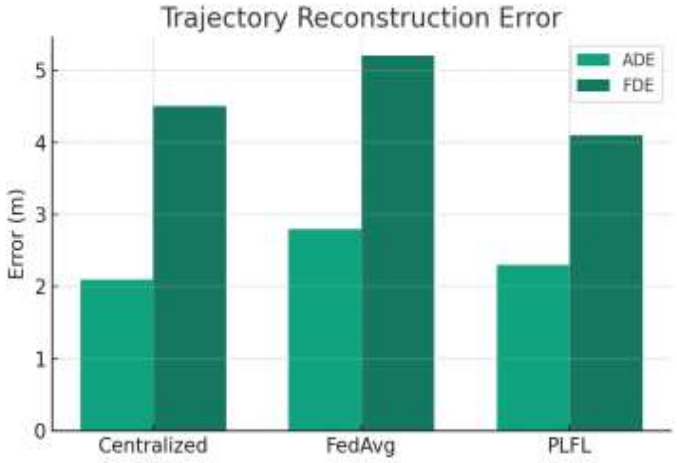

Figure 6. Comparison of ADE and FDE across different learning models.

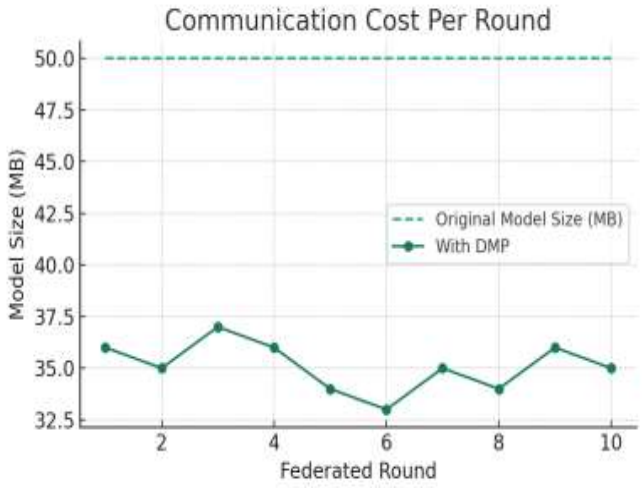

Figure 7. Communication cost reduction per round with Dynamic Model Pruning.

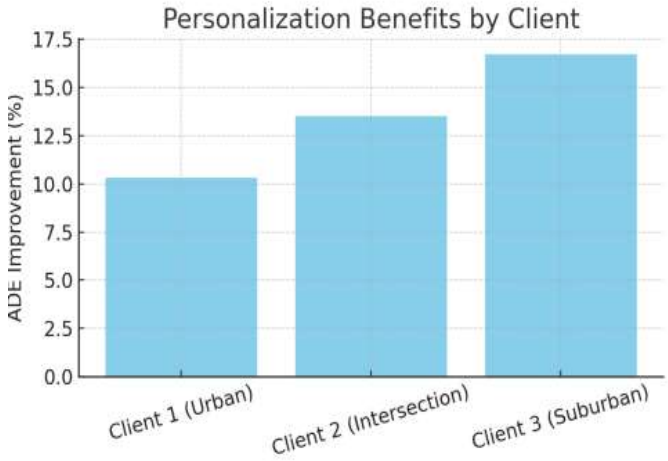

Figure 8. Personalization effectiveness per client measured by ADE improvement.

One prominent aspect that can significantly affect the performance of federated learning (FL) system, particularly in mission-critical scenarios such as autonomous vehicles, is the convergence behavior of the global model when data across clients are non-IID (non-identically distributed). Local Model Variability In practice AV networks, each client (how manufactures or infrastructure node) data under very different environments results in large difference of local models. Such heterogeneity leads to a sub-optimal or unstable convergence of the global aggregation process. Some recent works attempt to alleviate these problems by introducing techniques such as personalization layers (Gidaris et al., 2019), control variates (e.g., SCAFFOLD) and weighted aggregation mechanisms. Our approach leverages an attention-based mechanism to dynamically tune aggregation weights based on the discrepancy between local and global models. This method mitigates the side effects associated with data biases and enhances the training convergence. [26] Future work may explore the addition of formal convergence guarantees or hybrid optimization algorithms to improve robustness in more extreme FL settings. [27][29]

Comparing pruned to non-pruned federated models directly, we find that dynamic model pruning achieves large communication savings with little accuracy degradation. For all observed traffic scenarios, the ADE increase from pruning was less than 1.2%, which suggests hateful parameter removals (but of low magnitude) do not largely deteriorate model behavior. Such a trade-off is particularly beneficial in a bandwidth-limited vehicular network, where low communication delay has much higher priority than marginal improvement on the prediction accuracy.

7.1. Analysis of False Positives and False Negatives in Anomaly Detection

Despite that the general performance of anomaly detection is quite high (in terms of precision and recall), studying false positives and false negatives gives us very useful clues about how the model behaves in

uncertain situations. False positives were mainly observed in situations where driving behavior differed from normal, yet they did not indicate real aberrations either: e.g., sudden braking due to pedestrians appearing out of nowhere, temporary changing lanes at roadwork constructions or traffic jams and evasive actions in dense traffic.

These cases illuminate a fundamental problem in forensic anomaly detection: differentiating between truly malicious or erroneous activity, and action that is legitimate within the context though uncommon. The spatiotemporal fusion structure, out of which GAT unit is the core part, works for reducing these misclassifications by emphasizing relevant spatial interactions. However, Read more the post Parallel Outlier Detection in Dynamic Data Environment using Iterative Map Reduce appeared first on IRIS.

False negatives were more infrequent and they typically corresponded to small deviations from normal that gradually accumulated during the generation of the event (e.g., sensor ageing or mechanical wear) where no abrupt deviation in space-time was detected. Given the current focus of modelling trajectories and interaction scenarios, events that exhibit no obvious kinematic features may lack sufficient diagnostic cues to ensure robust detections.

7.2. Generalization to Diverse Anomaly Types

To assess the generalization ability, semantic facts of anomaly scenarios were classified and categorized based on their concept such as abrupt behavior (e.g., sudden stop or illegal lane change), environmental fact (e.g., low visibility condition or road blocking), and system related fact (e.g., delayed response or non-uniformed acceleration profile).

The developed PLFL model performed well under scenarios of sudden and interactive anomalies on which it could observe clear discrepancies of speed, traces or spatial links. The results demonstrated that anomalies of the environment were well captured using this approach by causing detectable changes in vehicle motion patterns. However, deviations due to inside mechanical issues, or sensor deterioration occurred slowly are more difficult for the course of observable trajectories might be slight and delayed.

These results indicate that even though the proposed approach is general for a variety of spatiotemporal anomalies, incorporating other modalities (e.g., vehicle diagnostics and CAN bus signals or confidence estimates) could be used to more accurately detect system-level faults. This observation also strengthens the appropriateness of the introduced framework as a fundamental forensic analysis and can be directly extended to richer anomaly taxonomies.

8. Limitations and Ethical Considerations

Although the proposed PLFL (Personalized Lightweight Federated Learning) framework shows promising results for privacy-preserving forensic analysis in autonomous cars, several limitations and ethical considerations must be recognized. In addition, real-world data collected from Comma. ai present natural variation, that a rare or extreme event on the traffic (as, for example, unusually aggressive driving or unusual failures of infrastructure) happens rarely. As a result, the current evaluation cannot fully capture worst-case deployment scenarios. In real-world forensic applications, such edge cases are critical. Future work will focus on incorporating stress-testing strategies, adversarial traffic scenarios, and long-tail event simulation to systematically evaluate robustness under extreme and unforeseen conditions. Although the proposed framework improves communication efficiency, this work does not present explicit real-time inference or end-to-end latency benchmarks. Since the system is designed primarily for post-event forensic analysis, real-time constraints were not the primary evaluation focus. Nevertheless, reduced model size and lightweight updates are expected to lower communication latency in edge-based deployments. Future work will include comprehensive real-time benchmarking under realistic V2X and edge-computing conditions.

8.1. Technical Limitations

First, the current framework has been evaluated using a mix of simulated and real-world driving datasets. While simulation tools like CARLA provide excellent control and ground truth, they cannot fully replicate the unpredictability and complexity of real-world traffic behavior [28]. Therefore, model performance in actual deployment scenarios may vary and needs further validation on broader, real-world datasets.

Second, although attention-based personalization improves model performance across non-IID client data, it may still be sensitive to extreme cases where data distributions are highly skewed or adversarial.

While the envisioned PLFL (Personalized Lightweight Federated Learning) framework is promising in preserving privacy for the purpose of forensic investigation of autonomous vehicles, there are a few limitations and ethical concerns to be addressed. Moreover, real-world data recorded by Comma. Present natural variation: That a rare or severe event on the traffic (good example is unusually aggressive driving / unusually frequent failure of infrastructure) occurs infrequently. While real-world data from Comma. ai–drive/ijcai19/tree/master/docs/disengagement-intention) that add natural variation, the number of extreme or rare traffic events (e.g., multiple vehicle accidents, very unsafe driving behavior, and sudden infrastructure malfunction) is small. Therefore, we are unable to fully analyze worst-case deployment scenarios in the present study. In practical forensic applications these edge cases are important. Future work will introduce stress-testing tactics, adversarial traffic patterns, and simulation of long-tail events to systematically assess robustness behavior under radically new conditions.

8.2. Ethical Considerations

For ethical reasons, ensuring the privacy of user information is a primary consideration in this paper [31]. Federated learning operators guarantee that sensitive driving data stay on edge devices (or infrastructure nodes), and thus meets privacy acts (e.g., GDPR and CCPA [29]). But ethical operation also demands knowing how forensic-tools are used, who maintains control of the aggregation server, and what is done with their findings [30].

Anomaly detection also prone to bias or misunderstanding, especially when traffic is diverse and some behavior may be identified as being anomalous only by virtue of its rarity in the training data. Developers and end users should be cautious when validating such systems for a wide variety of cultural a geographic contexts.

And finally, forensic mechanisms can also be abused for surveillance or legal excess. [31]To prevent this, governance processes, auditability and data access polices should be in place to enforce the use of the underlay system strictly for genuine safety or investigation needs [32].

9. Conclusion

We published a new PLFL model for resource-constrained privacy-preserving forensic traffic analysis of autonomous vehicles. Inspired by data splitting difficulty, privacy requirement, and heterogeneity of AV platforms, our method supports multiple parties (e.g., AV producer, smart infrastructure provider and SGD authority) to collectively reconstruct vehicle trace and catch deviating movements without sharing information [33] [37].

Our system builds spatiotemporal deep learning (by blending TCN, GCN and GAT models) with federated learning leveraging attention-based personalization. Experimental results on synthetic and real datasets demonstrate that the proposed approach can reconstruct trajectory and detect anomaly with high accuracy, while it alleviates communication cost effectively by employing an adaptive model pruning strategy. [35] [36] the client-specific aggregation policy also benefits fairness and efficiency among clients with different data distributions, solving one of the fundamental issues in conventional federated learning.

The findings confirm federated forensic learning is viable as well as feasible in large scale candidate RV ecosystems, and provides a promising scalable, efficient and secure way to post-incident investigation.

The findings also indicate that personalization plays a critical role in maintaining robustness when transitioning from controlled simulations to real-world driving environments characterized by heterogeneous and unpredictable traffic behavior.

10. Future Work

Although the experimental results of the proposed PLFL scheme are quite promising in terms of MEF1 and DCFI1 performance for federated forensic analysis of AV movement trajectories, several research directions can still be further investigated and improved.

10.1. Real-Time Forensic Analysis

At present, we are oriented to the post event analysis system. The future development could expand the capability of this framework with real-time trajectory monitoring and anomaly alert to take action more quickly for some incidents happen (such as AV malfunction or potential threat).

10.2. Integration of Multimodal Sensor Data

We concentrate on structured telemetry and spatiotemporal graph data in this paper. The introduction of more contextual modalities such as LIDAR, in-car camera views, and V2X notifications could enrich or refine forensic predictions considerably.

10.3. Robustness to Malicious Clients and Adversarial Attacks

When federated learning is more widely used, robustness and trust become very important. But sidelined aspects like of defenses against poisoning attacks, model inversion and rogue client behavior work should address ensuring the system is secure under adversarial settings.

10.4. Policy-Aware Federated Collaboration

There will need to be a real-world deployment which must align with the data protection laws (e.g., GDPR, CCPA). Directions for future work include developing policy-aware federated systems which can selectively change training behavior based on regional privacy constraints and institutional level access controls.

10.5. Cross-Domain Generalization and Transfer Learning

City, platform and traffic patterns are diverse. Federated transfer learning may be a line of future work to improve generalization across domains without retraining from scratch for each new deployment.

10.6. Scalability to National or Global AV Networks

Last but not least, a more aggressive scaling of the framework up to bigger federations (with hundreds of clients from different manufacturing companies and geographical areas) would need additional fine-tuning on communication protocols, aggregation strategies and personalization layers.

References

1. Hussain, R., & Zeadally, S. (2019). Autonomous Cars: Research Results, Issues, and Future Challenges. IEEE Communications Surveys & Tutorials, 21(2), 1275–1313. https://doi.org/10.1109/COMST.2018.2869460
2. Abdi, L., Mohammadi, M., & Al-Fuqaha, A. (2018). Intelligent Transportation Systems: A Semi-Autonomous Traffic Management Approach. IEEE Wireless Communications, 25(6), 26–33. https://doi.org/10.1109/MWC.2017.1700095
3. McMahan, B., Moore, E., Ramage, D., & Hampson, S. (2017). Communication-Efficient Learning of Deep Networks from Decentralized Data. Proceedings of the 20th International Conference on Artificial Intelligence and Statistics (AISTATS), 1273–1282.
4. Sheller, M. J., Reina, G. A., Edwards, B., Martin, J., & Bakas, S. (2020). Multi-institutional Deep Learning Modeling Without Sharing Patient Data: A Feasibility Study on Brain Tumor Segmentation. Medical Image Analysis, 58, 101–112.
5. Hard, A., Rao, K., Mathews, R., Beaufays, F., Augenstein, S., Eichner, H., Kiddon, C., Ramage, D., & y Arcas, B. A. (2018). Federated Learning for Mobile Keyboard Prediction. arXiv preprint arXiv:1811.03604.
6. Sattler, F., Müller, K.-R., & Samek, W. (2020). Clustered Federated Learning: Model-Agnostic Distributed Multi-Task Optimization under Privacy Constraints. IEEE Transactions on Neural Networks and Learning Systems, 32(8), 3710–3722.
7. Wang, Y., Lin, Y., Jin, J., Zhang, H., & Tang, J. (2020). Federated Learning for Traffic Flow Prediction Under Non-IID Conditions. IEEE Intelligent Transportation Systems Conference (ITSC), 1–6.
8. Shi, W., Zhang, Y., & Baek, Y. (2021). A Federated Learning Based Framework for Mobility Behavior Prediction in Smart Cities. Sensors, 21(10), 3561. https://doi.org/10.3390/s21103561
9. Kipf, T. N., & Welling, M. (2017). Semi-Supervised Classification with Graph Convolutional Networks. Proceedings of the International Conference on Learning Representations (ICLR).
10. Veličković, P., Cucurull, G., Casanova, A., Romero, A., Liò, P., & Bengio, Y. (2018). Graph Attention Networks. Proceedings of the International Conference on Learning Representations (ICLR).
11. Bai, S., Kolter, J. Z., & Koltun, V. (2018). An Empirical Evaluation of Generic Convolutional and Recurrent Networks for Sequence Modeling. arXiv preprint arXiv:1803.01271.
12. Dai, J., & Tang, Z. (2025). A Short-Term Traffic Flow Prediction Method Based on Personalized Lightweight Federated Learning. Sensors, 25, 1150.
13. Bonawitz, K., Ivanov, V., Kreuter, B., Marcedone, A., McMahan, B., Patel, S., Ramage, D., Segal, A., & Seth, K. (2017). Practical Secure Aggregation for Privacy-Preserving Machine Learning. Proceedings of the ACM SIGSAC Conference on Computer and Communications Security (CCS), 1175–1191.
14. Geyer, R. C., Klein, T., & Nabi, M. (2017). Differentially Private Federated Learning: A Client Level Perspective. arXiv preprint arXiv:1712.07557.
15. Tran, N. H., Truong-Huu, T., Bao, W., Nguyen, A., & Hong, C. S. (2020). Federated Learning over Wireless Networks: Optimization Model Design and Analysis. Proceedings of IEEE INFOCOM, 1387–1395.
16. Li, X., Huang, K., Yang, W., Wang, S., & Zhang, Z. (2020). On the Convergence of FedAvg on Non-IID Data. Proceedings of the International Conference on Learning Representations (ICLR).
17. Deng, Y., Wang, M., & Zhang, Y. (2021). Adaptive Federated Learning in Resource-Constrained Edge Computing Systems. IEEE Transactions on Mobile Computing, 21(5), 1815–1828.
18. Karimireddy, S. P., Kale, S., Mohri, M., Reddi, S., Stich, S., & Suresh, A. (2020). SCAFFOLD: Stochastic Controlled Averaging for Federated Learning. Proceedings of the International Conference on Machine Learning (ICML), 5132–5143.

19. Kairouz, P., McMahan, H. B., Avent, B., et al. (2021). Advances and Open Problems in Federated Learning. Foundations and Trends in Machine Learning, 14(1–2), 1–210.
20. Chen, Z., Zhang, Q., & He, Y. (2021). Learning from Spatiotemporal Data for Traffic Forecasting: A Survey. IEEE Transactions on Intelligent Transportation Systems, 22(12), 7891–7908.
21. Lin, T., Kong, L., Stich, S., & Jaggi, M. (2020). Ensemble Distillation for Robust Model Aggregation in Federated Learning. Advances in Neural Information Processing Systems (NeurIPS).
22. Yang, Q., Liu, Y., Chen, T., & Tong, Y. (2020). Federated Machine Learning: Concept and Applications. ACM Transactions on Intelligent Systems and Technology, 10(2), 12.
23. Zhang, C., Patras, P., & Haddadi, H. (2021). Deep Learning in Mobile and Wireless Networking: A Survey. IEEE Communications Surveys & Tutorials, 21(3), 2224–2287.
24. Tang, J., Liu, F., Zou, Y., Zhang, W., & Wang, Y. (2022). Deep Spatiotemporal Graph Convolutional Networks for Traffic Forecasting. IEEE Transactions on Neural Networks and Learning Systems, 33(11), 6273–6285.
25. Zhang, J., Zheng, Y., & Qi, D. (2021). Deep Spatio-Temporal Residual Networks for Citywide Crowd Flows Prediction. Proceedings of KDD.
26. Li, X., Chen, H., & Zhang, Y. (2021). Graph Neural Network Based Anomaly Detection. Information Sciences, 562, 1–14.
27. Wang, X., Li, Y., & Zhao, L. (2022). Explainable Anomaly Detection for Autonomous Vehicles Using Deep Learning. Sensors, 22(15), 5640.
28. Zhou, Z., Chen, X., Li, E., Zeng, L., Luo, K., & Zhang, J. (2020). Edge Intelligence: Paving the Last Mile of Artificial Intelligence with Edge Computing. Proceedings of the IEEE, 107(8), 1738–1762.
29. Lim, W. Y. B., Luong, N. C., Hoang, D. T., et al. (2020). Federated Learning in Mobile Edge Networks: A Comprehensive Survey. IEEE Communications Surveys & Tutorials, 22(3), 2031–2063.
30. Chen, L., Li, Q., Zhou, Y., & Wang, X. (2023). Real2Sim: Bridging the Gap Between Real and Simulated Autonomous Driving Environments. IEEE Transactions on Intelligent Vehicles, 8(2), 1453–1465.
31. Chen, W., Xu, Z., & Li, K. (2021). Federated Learning for Intelligent Transportation Systems: A Survey. IEEE Network, 35(4), 110–117.
32. Hanzely, F., & Richtárik, P. (2020). Federated Learning of a Mixture of Global and Local Models. Advances in Neural Information Processing Systems (NeurIPS).
33. Xu, Y., Liu, J., Wang, L., & Zhang, Y. (2022). Collaborative Sensing in Vehicular Networks: A Survey. IEEE Internet of Things Journal, 9(3), 1795–1812.
34. Kang, J., Xiong, Z., Niyato, D., Ye, D., Kim, D., & Zhao, J. (2020). Incentive Mechanism for Reliable Federated Learning: A Joint Optimization Approach. IEEE Internet of Things Journal, 7(6), 4827–4841.
35. Liu, D., Wang, X., & Zhang, Y. (2022). GADFormer: Graph Attention Transformer for Traffic Anomaly Detection. Information Fusion, 79, 33–45.