Published in: IEEE Transactions on Vehicular Technology

Two-layer Federated Learning with Heterogeneous Model Aggregation for 6G Supported Internet of Vehicles

Authors:
Xiaokang Zhou, Wei Liang
Jinhua She, Zheng Yan, and
Kevin I-Kai Wang

Presented By:
Ittehad Saleh Chowdhury
Reyadath Ullah Akib

Introduction

1. **Innovative Framework:** The paper presents a new two-layer federated learning approach for 6G vehicular networks.

2. **Enhanced Data Handling:** It introduces a unique model that effectively combines local and global data for improved learning efficiency.

3. **Superior Performance:** The model shows better accuracy in intelligent object detection, outperforming existing methods in dense vehicular environments.

Challenges

1. **Data Privacy and Security:** It tackles concerns related to the privacy and security of data in distributed vehicular environments.

2. **Communication Efficiency:** The model is designed to reduce the communication overhead typically associated with federated learning systems.

3. **Handling Heterogeneous Data:** Introduces a solution for efficiently aggregating and processing heterogeneous data from various local and global sources.

4. **Need for Distributed Machine Learning:** Distributed machine learning methods that can efficiently utilize the vast and diverse data generated at the edge of massive interconnected networks.

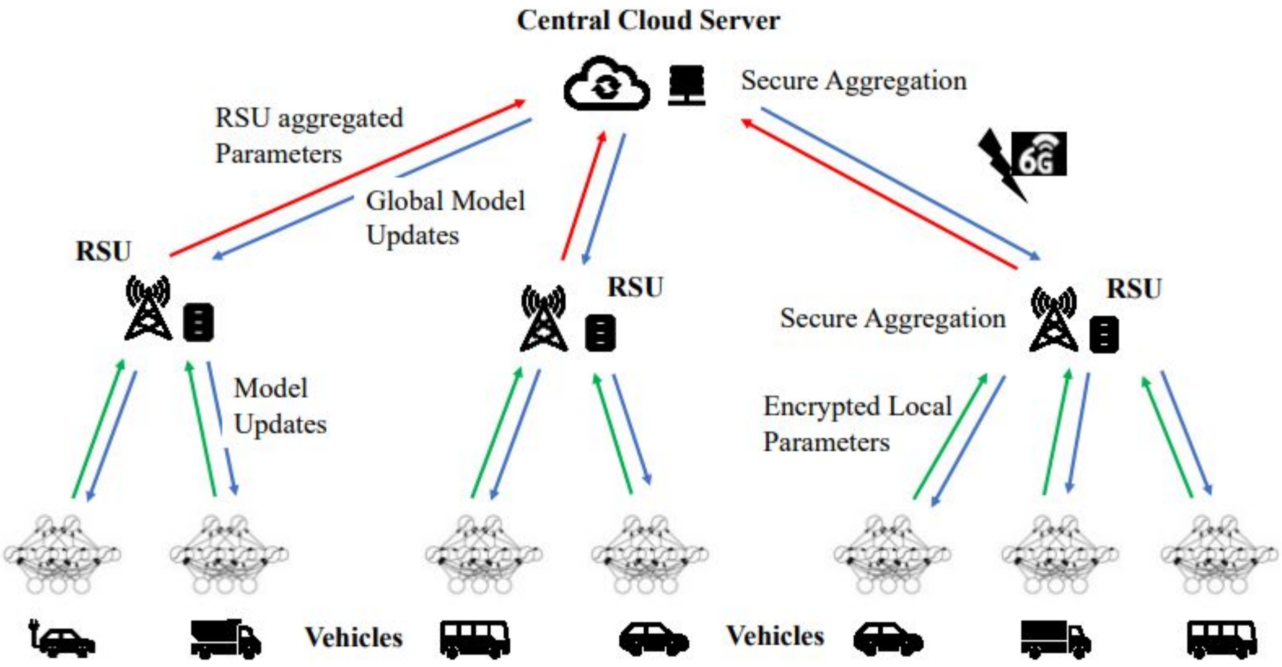
Applications

1. **Intelligent Transportation Systems:** Enhances object detection for safer and more efficient traffic management.

2. **6G Vehicular Networks:** Offers a practical model for data processing in next-generation vehicular communication systems.

3. **Driving and Traffic Enhancements:** The technology aids in recognizing and sharing roadside information like traffic signs and parking details, leading to improvements in autonomous driving, traffic management, and route planning.

Basic Model Architecture



Continued.

1. **Top Layer (Central Cloud Server):** Offers high-performance computing and higher-level knowledge sharing for global caching and aggregation tasks.

2. **Middle Layer (Roadside Units - RSUs):** RSUs with limited caching and computing capabilities supervise vehicles in their coverage, aggregating learning parameters and contextual information like vehicle locations.

3. **Bottom Layer (Vehicles):** Vehicles generate raw data and perform light computing tasks like training models for object detection. They operate under a unified deep learning model and contribute to the federated learning system while ensuring privacy and efficient knowledge sharing.

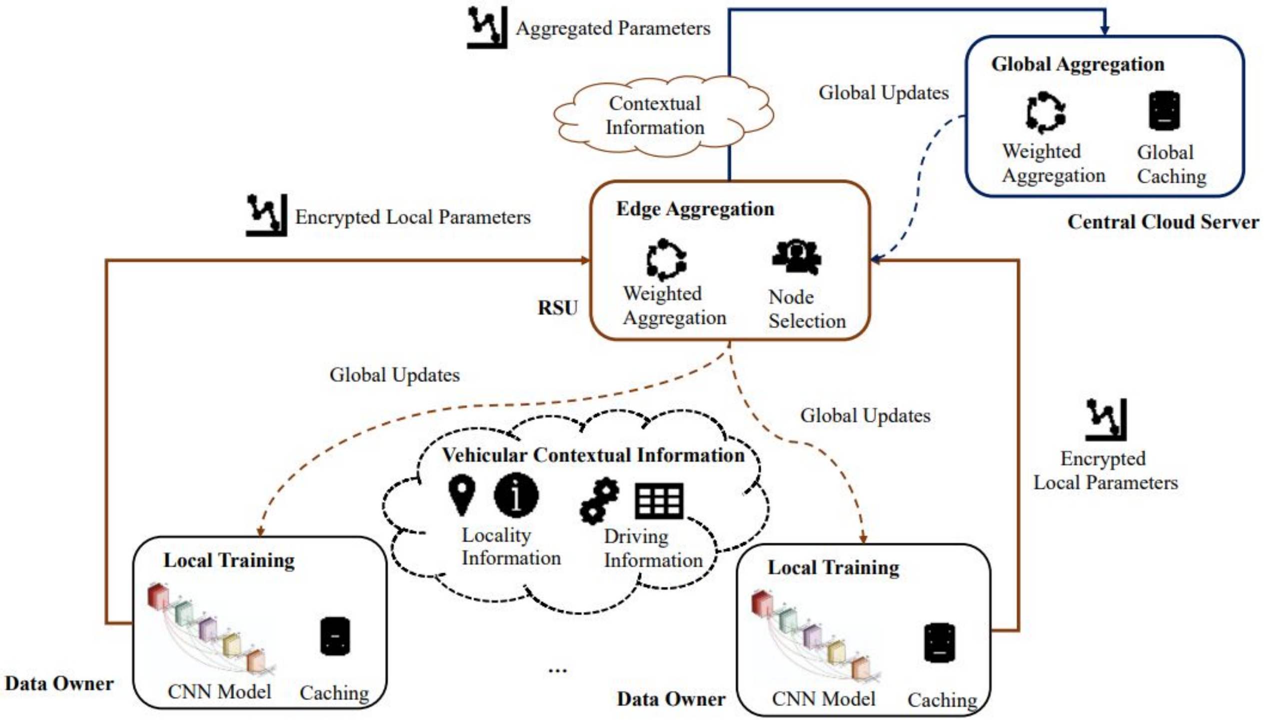
Problem Definition

- A global deep learning model $\boldsymbol{M = \hbar(x, \omega)}$ is trained on the distributed dataset $\boldsymbol{D_{total}}$ across the 6G supported vehicular network.

- The goal of the proposed framework is to recognize the objects or road signs during the real-time autonomous driving scenario based on samples in $\boldsymbol{D_{total}}$.

- Parameters computed via local training by $\boldsymbol{v_i}$ need to be encrypted and then sent to the corresponding RSU.

CNN Based Two-Layer Federated Learning



Continued

CNN model, which is introduced to perform an object detection task can be represented as the hypothesis $\boldsymbol{h(x_i, \omega)}$ and trained locally by the data owner $\boldsymbol{v_i}$

.

$$h(x_i, \omega) = \mathrm{FC}(\mathrm{Pool}(\mathrm{Conv}(x_i, \omega_{Conv}), \omega_{Pool}), \omega_{FC})$$

The result of $\boldsymbol{h(x_i, \omega)}$ is predicted by a SoftMax classifier as follows.

$$y^{predict} = \mathrm{SoftMax}(h(x_i, \omega))$$

Cost Function

We define the cost function as follows.

$$J_i(\omega) = -\frac{1}{k}\left[\sum_{q=1}^{k}\sum_{p=1}^{l}\{y_i = p\}\log\left(\frac{e^{\omega_p^T x_i}}{\sum_{q=1}^{k} e^{\omega_q^T x_i}}\right)\right]$$

where we assume there are ***k*** samples for data owner $\boldsymbol{v_i}$.

Multi-Layer heterogeneous Selection and Aggregation Scheme

Both the **vehicular** and **RSU** contextual information are taken into account in the aggregation process.

The total parameter $w_{\mathbf{j}}^{\mathbf{RSU}}(t)$ by RSU r_j is aggregated by the following equation:

$$\omega_j^{RSU}(t) = \frac{1}{J}\sum_{i=1}^{J} \delta_i * \omega_i(t) \qquad (4)$$

$w_i(t)$ = parameters gained by data owner $\boldsymbol{i}$ at iteration $t \in T$

J = total number of the vehicles supervised by an RSU r_j

δ_i = weighted coefficient of vehicle v_i

Multi-Layer heterogeneous Selection and Aggregation Scheme (local aggregation)

δ_i is applied to measure how the local training result can be influenced by its locality and driving contextual information. It is being calculated based on two factors:

Distance: A distance-based measurement between an RSU r_j and vehicle v_i is defined to represent the locality attribute of v_i

$$dis(v_i, r_j) = \sqrt{(lat_i - lat_j)^2 + (lot_i - lot_j)^2} \qquad (5)$$

Image quality: The quality of captured raw data by vehicles may be influenced by the direction and velocity of the moving vehicles

$$drv(v_i) = \frac{\theta_i}{90} * \log(vlc_i) \qquad (6)$$

θ_i = Direction of the velocity of the vehicle

vlc_i = The instantaneous velocity of the vehicle

Multi-Layer heterogeneous Selection and Aggregation Scheme (local aggregation)

The weighted coefficient $\boldsymbol{\delta_i}$ of vehicle $\mathbf{v_i}$ can be calculated using the vehicular contextual information using equation (5) and (6) in the following way:

$$\delta_i = \frac{dis(v_i, r_j)}{\sum_{j=1}^{J} dis(v_i, r_j)} + drv(v_i) \qquad (7)$$

* $\sum_{i=1}^{J} dis(v_i, r_j)$ (??)

Multi-Layer heterogeneous Selection and Aggregation Scheme (global aggregation)

Once the local models are aggregated by individual RSUs, multiple RSUs will upload their aggregated parameters to the central cloud server CS where the global model selection and aggregation will happen.
The weighted coefficient ξ_j for an RSU r_j can be calculated based on two factors:

- Computational Power of the RSU : $P(r_j)$
- Number of vehicles being supervised by r_j : J

$$\xi_j = \frac{J}{n} * \log\left(\frac{P(r_j)}{P(CS)}\right) \quad (8)$$

n = whole vehicular network capacity

P(CS) = Computational power of Cloud Server

Intelligent Object Detection Algorithm

Based on the applied horizontal federated learning framework, we can share an identical global model and update the training parameters within the whole 6G supported vehicular network.

We define a gradient descent calculation for parameters in each iteration as follows:

$$\omega(t+1) = \omega(t) + \nabla(\omega) \quad (9)$$

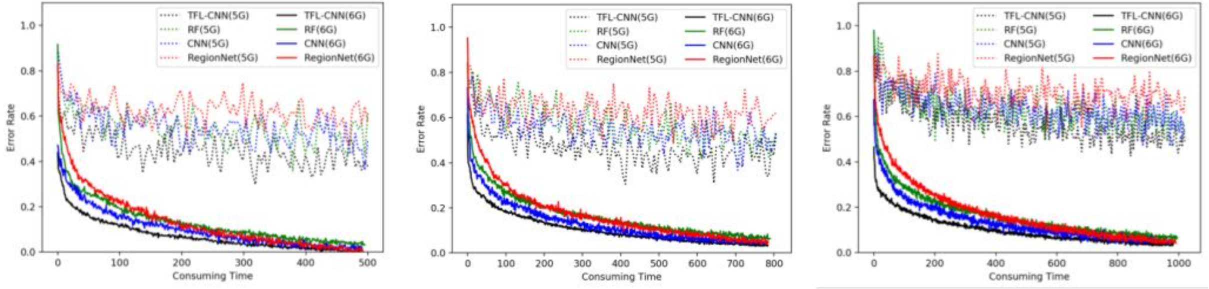
ω indicates the aggregated parameter in the global model

Intelligent Object Detection Algorithm

Input: A set of raw data provided by data owners $D = \{d_i\}_{i=1}^{n}$
Participated data owners $V = \{v_i\}_{i=1}^{n}$
Participated RSUs $R = \{r_j\}_{j=1}^{m}$
Output: A trained global object detection model M

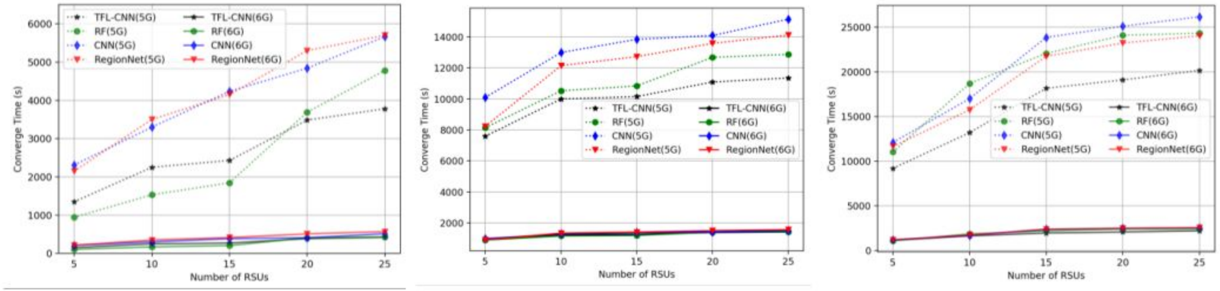
Initialize model parameter ω, maximum iterations T
for $t = 1$ to T **do**
  for each RSU $r_j \in R$ **do**
    for each data owner $v_i \in V$ **do**
      if v_i supervised by r_j **do**
        Conduct local training on dataset d_i
        Calculate the contextual information for v_i by Eq. (5), Eq. (6)
        Submit the local training parameter $\omega_i(t)$ to r_j
    end for
    Aggregate the parameters in r_j by Eq. (4)-(7)
    Upload the parameters to server CS
  end for
  Calculate the global parameters $\omega(t+1)$ for model M by Eq. (9)
  if $\nabla(\omega)$ reach the convergency threshold **goto** 17
  Broadcast $\omega(t+1)$ to the network
end for
return M

Result Analysis



(a) 0.5ms E2E Latency (b) 1ms E2E Latency (c) 5ms E2E Latency

Fig. 4. Error Rate Observation Under 5G/6G Configuration in Federated Learning Process



(a) 0.5ms E2E Latency (b) 1ms E2E Latency (c) 5ms E2E Latency

Fig. 5. Total Convergence Time(s) Under 5G/6G Configuration to Meet the Threshold of Error Rate

Result Analysis

TABLE 2. COMPARISONS ON TRAFFIC SIGN RECOGNITION PERFORMANCE

Methods	Precision	Recall	F1 score
CNN	0.889	0.821	0.854
RegionNet	0.899	0.910	0.904
RF	0.746	0.852	0.795
TFL-CNN	**0.906**	**0.968**	**0.936**

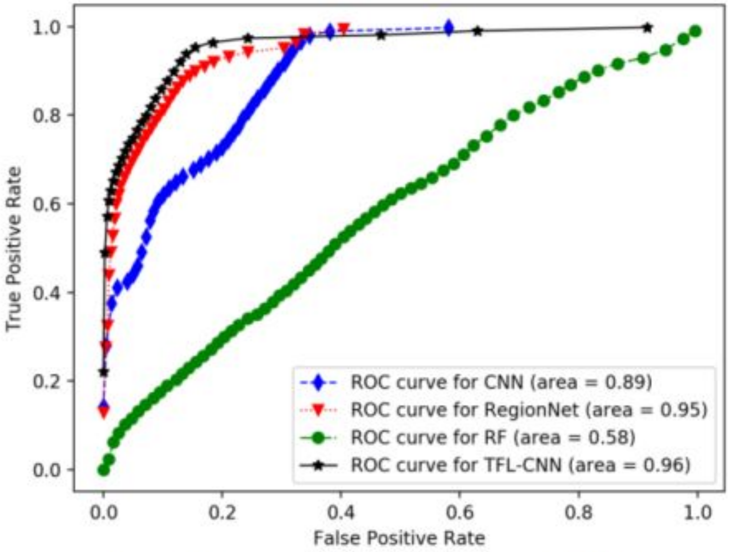


Fig. 6. Performance Evaluation Based on ROC

Observations!

Thank You!