Adaptive Local Aggregation for Personalized Federated Learning in Agricultural Terrain Mapping

Vishal Krishna Singh, Nitin Shukla, Deepak Kumar, Neeraj Jain, Chhaya Singh, Rajkumar Singh Rathore and Weiwei Jiang

***Abstract*—Statistical heterogeneity and non-independent and identically distributed (non-IID) data remain major challenges in federated learning, particularly for agricultural terrain mapping where geographically distributed sensing devices collect highly diverse remote sensing data. To address these challenges, this paper proposes an Adaptive Local Aggregation (ALA)-based personalized federated learning framework that improves client-specific model initialization while preserving the standard federated learning communication and global aggregation protocol. Rather than directly replacing local models with the received global model, the proposed framework adaptively initializes selected higher-layer parameters using learnable client-specific aggregation weights, thereby balancing globally shared knowledge with local personalization prior to client-side training. The proposed framework is evaluated on four agricultural remote sensing datasets, namely AgriFieldNet, BigEarthNet $v1$, BigEarthNet $v2$, and CropHarvest, together with three benchmark image classification datasets, MNIST, CIFAR-10, and CIFAR-100, under heterogeneous federated learning settings. Experimental results demonstrate consistent improvements in global classification accuracy and convergence behaviour compared with state-of-the-art federated learning methods, highlighting the effectiveness of the proposed framework for privacy-preserving agricultural terrain mapping under heterogeneous client data distributions.**

***Index Terms*—Precision Agriculture, Federated Learning, Machine Learning, Internet of Things, Unmanned Aerial Vehicles, Image Processing**

I. Introduction

Agricultural terrain mapping and land use classification (ATMLUC) is popularly used for resource management, disaster management and yield management among others. ATMLUC plays a vital role in ensuring global food security and the economy. Successful crop production is heavily reliant on environmental factors such as weather conditions, soil quality, and water availability. Monitoring these factors with accuracy and efficiency is critical to optimize crop yields and minimize resource wastage. Internet of Things (IoT), similar to other applications such as forest fire detection [1], [2], multimedia data processing [3], anomaly detection [4] and traffic networks [5], plays a crucial role in gathering task relevant-real time data [6]. The introduction of unmanned aerial vehicles (UAVs), has significantly enhanced data acquisition capabilities, enabling higher resolution monitoring of agricultural fields and the collection of more detailed information than ever before. Exploiting the advantages of UAV imagery, ATMLUC is being widely used for crop yield detection [7], flood detection [8], animal tracking [9], etc. For example, object tracking through aerial video classification is crucial for enabling UAV intelligence. Greater mobility and bigger coverage are two benefits of aerial surveillance over applications using stationary cameras, but it also has drawbacks including low resolutions and changeable backgrounds. As a result, a lot of emphasis has been paid to creating techniques for spotting moving objects in aerial footage. The three main categories of moving object identification methods are temporal-based, spatial-based, and combination approach. Since motion cues provide the most reliable information for detecting moving objects in video, most proposed methods rely on temporal information, including frame difference [10], [11], optical flow [12], [13] and background subtraction [14], [15]. Some researchers have also proposed using multi-motion layer analysis for detecting moving objects in aerial video from an airborne platform, or exploiting long-term motion patterns to detect moving vehicles [16].

Accuracy in such applications is highly desired and is an important aspect for achieving the results. Existing methods are dominantly led by Convolutional Neural Networks (CNN) and have shown promising results. The general idea is to transmit all the UAV samples to a central processing unit for training and use learning models for accurate classification and mapping of the land. Despite the significant advantages, the FL architecture for ATMLUC faces several practical challenges. For instance, UAVs are prone to technical failures and are susceptible to adverse weather conditions, harsh terrain, and frequently changing environmental conditions. This often results in low-quality images and missing or corrupted data, thereby degrading the overall model training and prediction performance. Furthermore, low-quality images may lead to

V. K. Singh is with the School of Computer Science and Electronics Engineering, University of Essex, Colchester, CO4 3SQ, U.K. (e-mail: v.k.singh@essex.ac.uk)

Nitin Shukla is with School of Computer Science Engineering and Technology, Bennett University, Greater Noida, Uttar Pradesh 201310, India. (e-mail: nitin.shukla@bennett.edu.in).

Deepak Kumar is with Dell Technologies, Bengaluru, Karnataka, 560048, India. (e-mail: mcs22020@iiitl.ac.in)

Neeraj Jain is with School of Computer Science Engineering and Technology, Bennett University, Greater Noida, Uttar Pradesh 201310, India. (e-mail: neeraj.jain@bennett.edu.in).

Chhaya Singh is with School of Computer Application and Technology, Galgotias University, Greater Noida, Uttar Pradesh 203201, India. (e-mail: chhaya.singh@galgotiasuniversity.edu.in).

Rajkumar Singh Rathore is with Department of Computer Science, School of Technologies, Cardiff Metropolitan University, CF5 2YB, United Kingdom. (e-mail: rsrathore@cardiffmet.ac.uk)

Weiwei Jiang is with School of Information and Communication Engineering, Beijing University of Posts and Telecommunications, Beijing, 100088, China.

Corresponding Author: V. K. Singh (e-mail: v.k.singh@essex.ac.uk) and N. Shukla (e-mail: nitin.shukla@bennett.edu.in).

Manuscript received June , 2025.

class imbalance, biased model updates, and overfitting at both the local and global levels [17]. It is important to note that statistical heterogeneity in Federated Learning (FL) architectures significantly limits the ability of the global model to generalize across diverse clients [18]. In agricultural terrain mapping, this challenge becomes even more pronounced because UAVs operating in different geographical locations naturally observe different crop types, soil characteristics, illumination conditions, and environmental variations. Consequently, the data available at individual clients are highly non-independent and non-identically distributed (non-IID), resulting in inconsistent local updates, client drift, slower convergence, and degraded global model performance.

Although conventional FL approaches effectively preserve data privacy by avoiding centralized data collection, they often rely on uniform model initialization and global aggregation strategies that do not adequately account for statistical heterogeneity across clients [19]. Consequently, directly replacing local models with the received global model may weaken client-specific representations and limit personalization under highly heterogeneous data distributions. There is therefore a need for a personalized federated learning framework that adaptively balances global knowledge with client-specific representations while preserving the standard federated learning communication protocol. This work is therefore aimed at developing a personalized FL framework for agricultural terrain mapping to address statistical heterogeneity arising from geographically distributed non-IID data. The proposed Adaptive Local Aggregation (ALA)-based client initialization strategy adaptively initializes selected higher-layer parameters by combining the received global model with the previously trained local model using learnable client-specific aggregation weights. Consequently, the proposed framework improves convergence behaviour and client personalization while preserving the privacy-preserving nature and communication efficiency of conventional federated learning. The major contributions of this work are summarized as follows:

1) A personalized FL framework is developed for agricultural terrain mapping to address statistical heterogeneity and client drift arising from geographically distributed non-IID data while preserving the standard FL communication protocol.
2) An ALA-based client initialization strategy is incorporated into the FL framework, whereby selected higher-layer parameters are adaptively initialized through learnable client-specific aggregation weights prior to local training. Unlike approaches that modify the server-side aggregation process, the proposed framework performs personalization entirely on the client side while remaining compatible with conventional FedAvg-style global aggregation.
3) The proposed framework is extensively evaluated on benchmark image classification datasets (MNIST, CIFAR-10, and CIFAR-100) as well as agricultural remote sensing datasets (AgriFieldNet, BigEarthNet v1, BigEarthNet v2, and CropHarvest) under heterogeneous FL settings, demonstrating improved global classification performance together with stable convergence and scalability compared with representative FL baselines.

The remainder of the paper is organized as follows. Section II presents the problem description, followed by the system model and the proposed method in Section III. Section IV describes the datasets, experimental setup, and comparative performance evaluation under different scenarios. Finally, Section V concludes the paper.

II. Problem Description

Federated architectures for mapping and classification of agricultural lands, face some major challenges such as technical failures of UAVs, adverse weather conditions, harsh terrain and frequently changing environmental conditions. In such a scenario, the captured images are often of low quality and it leads to potential class imbalance and introduction of bias and overfitting at the local and the global level. Additionally, statistical heterogeneity in federated architectures is constrained by the global model's capacity to generalize across diverse clients. In applications such as agricultural terrain mapping, the models performance is further depreciated because of unique variability in data distribution among clients and biased updates. Real-world data distributions are often plagued by non-independent and non-IID characters across different types of devices. In real-world scenarios, data distributions vary significantly among clients, making a one-size-fits-all global model less effective. The FL architecture allows global models to be trained in collaboration with numerous clients while maintaining local data. However, the conventional federated optimization techniques like *FedAvg*, aggregate the model discriminately over all clients without considering the different data distributions and capacities between clients. This causes bad convergence and performance due to variations in local updates.

III. Proposed Methodology

This section presents the proposed ALA-based FL framework. The proposed framework is designed to improve model personalization under heterogeneous (non-IID) client data distributions while preserving effective global model convergence. Each client trains a local four-layer CNN using its private dataset, whereas the ALA module adaptively initializes selected layers of the local model by combining the received global model with the previously learned local model before local optimization.

A. System Architecture

Consider a FL system consisting of N distributed clients, where client i possesses a private local dataset D_i. During the FL process, the local datasets remain entirely on their respective devices and are never shared with the central server. Instead, only model parameters are exchanged during communication rounds to preserve data privacy. Let K_g^t denote the global model maintained by the server at communication round t, and let K_i^t denote the local model maintained by client i. Throughout all experiments, every client employs the

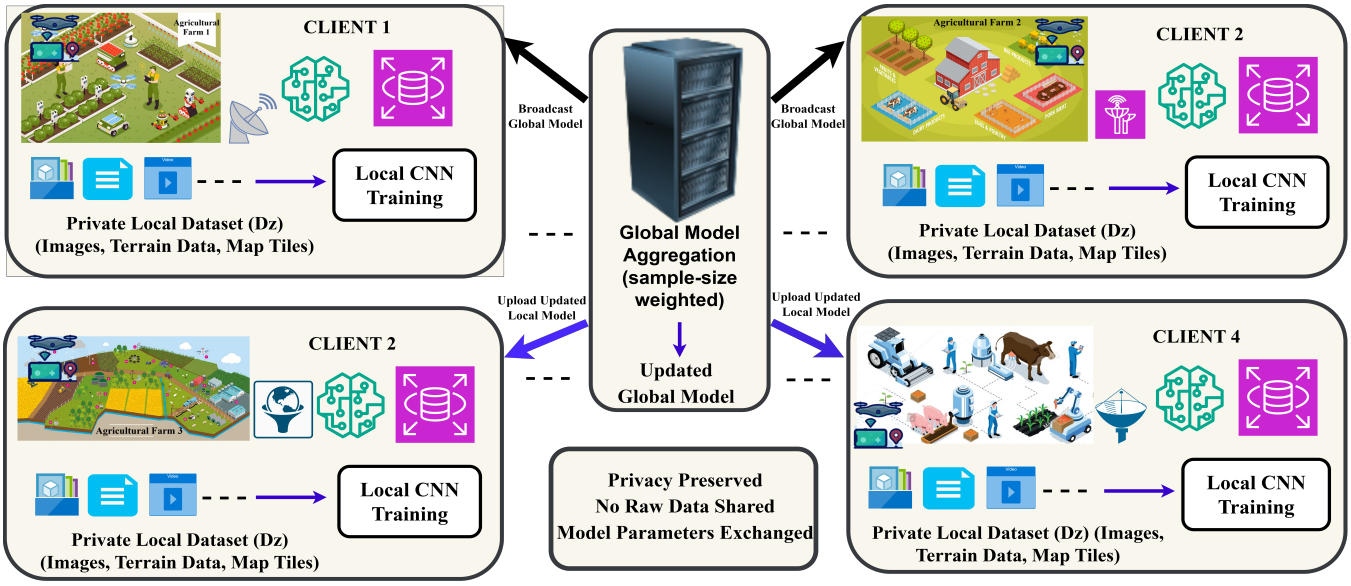


Fig. 1. Federated Agricultural Terrain Mapping Architecture

same four-layer CNN architecture, while the proposed ALA mechanism operates directly on the trainable parameters of these local models. The Fig. 1 shows the proposed federated architecture for terrain mapping. Since client datasets follow heterogeneous and non-IID distributions, directly replacing a locally trained model with the received global model may overwrite client-specific knowledge accumulated during previous communication rounds. Consequently, the proposed framework performs adaptive parameter initialization before local training so that globally shared representations are retained while preserving locally discriminative features.

B. Proposed Framework

The overall workflow of the proposed framework is illustrated in Fig. 2. At the beginning of each communication round, the central server selects a subset of participating clients and broadcasts the current global model K_g^t. Instead of directly replacing the local model with the received global model, each participating client performs ALA to initialize its model before local optimization. The lower layers of the local CNN are directly initialized using the corresponding parameters of the received global model, enabling all clients to share common low-level feature representations. In contrast, the higher layers are adaptively initialized through learnable aggregation weights that balance information from both the previously learned local model and the received global model. This adaptive initialization preserves client-specific knowledge while simultaneously benefiting from globally learned representations.

After initialization, each client independently performs local training using its private dataset. The updated local models are then transmitted to the server, where standard sample-size-weighted federated aggregation is performed to generate the global model for the next communication round. Since ALA is executed entirely at the client side before local optimization, the proposed framework preserves the conventional federated communication protocol without requiring any exchange of raw training data or additional server-side optimization procedures. Directly initializing local models using the received global parameters may degrade personalized representations under heterogeneous data distributions. The proposed ALA mechanism addresses this limitation by adaptively combining higher-layer parameters from the global and local models while preserving globally shared representations within the lower layers. Consequently, the proposed framework achieves an improved balance between client personalization and global model convergence without increasing communication complexity.

C. Adaptive Local Aggregation

The proposed ALA mechanism performs client-specific model initialization before each round of local training. Instead of directly adopting the received global model, ALA adaptively combines the global and previously learned local parameters, thereby preserving personalized representations while benefiting from globally learned knowledge.

Let $K_{i,l}^t$ and $K_{g,l}^t$ denote the parameters of layer l in the local and global models at communication round t, respectively. The CNN layers are divided into two subsets, namely the lower layers $\mathcal{L}_{low}$ and the higher layers $\mathcal{L}_{high}$. Since the lower layers primarily learn generic feature representations, they are initialized directly from the received global model according to

$$K_{i,l}^t \leftarrow K_{g,l}^t, \qquad l \in \mathcal{L}_{low}. \tag{1}$$

The higher layers contain more task- and client-specific representations. Therefore, they are adaptively initialized through learnable element-wise aggregation weights ω_l, yielding

$$\widetilde{K}_{i,l}^t = K_{i,l}^t + \omega_l \odot \left(K_{g,l}^t - K_{i,l}^t\right), \qquad l \in \mathcal{L}_{high}, \tag{2}$$

where $\odot$ denotes element-wise multiplication. During the first communication round, the aggregation weights are initialized to one so that the higher-layer parameters initially coincide with the received global model. In subsequent communication rounds, the aggregation weights are reused and refined using the client's locally available training data.

Before local optimization begins, a randomly sampled subset of the client's training data is used to learn the aggregation weights. During this initialization stage, only the aggregation weights associated with the selected higher layers are updated, whereas the lower layers remain fixed. Given the local loss function $L_i(\cdot)$, the aggregation weights are updated according to

$$\omega_l \leftarrow \omega_l - \eta \left(\nabla_{\widetilde{K}_{i,l}^t} L_i \odot \left(K_{g,l}^t - K_{i,l}^t \right) \right), \tag{3}$$

where η denotes the aggregation-weight learning rate. To maintain numerical stability and ensure a valid interpolation between local and global parameters, the aggregation weights are constrained to the interval $[0, 1]$ according to

$$\omega_l \leftarrow \min\left(1, \max\left(0, \omega_l\right)\right). \tag{4}$$

After the aggregation weights converge, the initialized higher-layer parameters obtained from Eq. (2) are combined with the globally initialized lower layers to construct the personalized client model used for local training. Since the entire ALA procedure is executed locally, no additional communication with the server is required. Furthermore, only model parameters are exchanged during the FL process, thereby preserving client data privacy.

D. Local Training and Global Aggregation

Following ALA, each selected client trains its personalized four-layer CNN using its private local dataset. Local optimization is performed independently by each client using mini-batch optimization without exchanging raw training samples. The local model update is formulated as:

$$K_i^{t+1} = \arg\min_{K_i} L_i(K_i; D_i), \tag{5}$$

where $L_i(\cdot)$ denotes the local loss function computed over the private dataset D_i of client i. After local optimization, only the updated model parameters are transmitted to the central server. Upon receiving the updated client models, the server constructs the global model through sample-size-weighted federated aggregation according to

$$K_g^{t+1} = \sum_{i \in \mathcal{C}_t} \frac{|D_i|}{\sum_{j \in \mathcal{C}_t} |D_j|} K_i^{t+1}, \tag{6}$$

where $\mathcal{C}_t$ denotes the set of participating clients during communication round t, and $|D_i|$ represents the number of training samples available at client i. The aggregated global model is subsequently broadcast to the participating clients at the beginning of the next communication round.

The proposed ALA framework is specifically designed to improve personalization under heterogeneous (non-IID) client data distributions while preserving strong global model convergence. Since the adaptive aggregation process is performed entirely on each client using locally available data, no additional communication overhead is introduced beyond the standard FL protocol. It should be noted that the proposed framework is not intended to address adversarial client behavior. The present work focuses on mitigating statistical heterogeneity through adaptive layer-wise initialization rather than detecting malicious model updates or performing robust aggregation. The proposed framework preserves data privacy through the standard FL paradigm, in which raw client data remain on local devices and only model parameters are exchanged with the central server. The present work does not incorporate a formal Differential Privacy (DP) mechanism, such as gradient perturbation, privacy accounting, or explicit (ϵ, δ) guarantees. The proposed framework employs standard sample-size-weighted global aggregation following the conventional FL protocol. FedAvg and FedProx are included in this work as representative baseline methods for experimental comparison and are not integrated into the proposed optimization framework.

Algorithm 1 Adaptive Local Aggregation-Based Federated Learning

Require: Number of clients N, local datasets $\{D_i\}_{i=1}^N$, communication rounds T, aggregation-weight learning rate η, selected higher-layer index set $\mathcal{L}_{high}$.
Ensure: Final global model K_g.

```
Initialize the global CNN model K_g^0.
for t = 0 to T do
    Select participating clients C_t.
    Broadcast the global model K_g^t to all selected clients.
    for all i ∈ C_t in parallel do
        Initialize lower-layer parameters using the received global model.
        Sample a subset of local training data for ALA initialization.
        Initialize or reuse aggregation weights ω_l.
        Update aggregation weights using Eq. (3).
        Clip aggregation weights using Eq. (4).
        Construct the personalized client model using Eq. (2).
        Train the initialized local model according to Eq. (5).
        Upload the updated model parameters K_i^{t+1} to the server.
    end for
    Aggregate all received client models using Eq. (6).
end for
return final global model K_g^T.
```

IV. Experiments and Results

A. Dataset Description

1) AgriFieldNet [20]: The AgriFieldNet India dataset comprises Sentinel-2 multispectral images with a spatial resolution of 10 meters. The images were collected in the year 2021 – 2022, from the regions in Uttar Pradesh, Bihar, Odisha, and

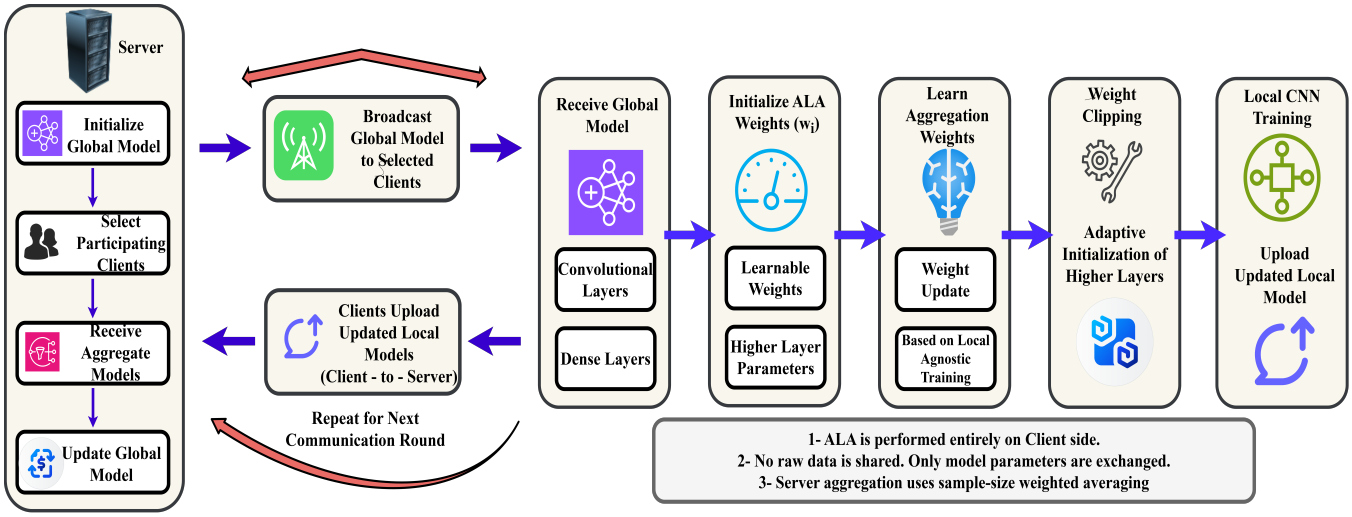


Fig. 2. Proposed ALA based Federated Learning Framework

Rajasthan. The images are of 256×256 pixels and correspond to 13 crop types (e.g., wheat, rice, mustard, sugarcane).

2) BigEarthNet $v1$ [21]: The BigEarthNet $v1$ is a large-scale remote sensing dataset. The dataset has 590326 $Sentinel-2$ image patches, each of 120×120 pixels with geographically diverse regions spread across 10 European countries. Each image is annotated with multi-label land cover classes derived from the CORINE Land Cover (CLC) database. The dataset contains 12 spectral bands (excluding the 10^{th} band).

3) BigEarthNet $v2$ [22]: BigEarthNet $v2$ is a large-scale benchmark designed for remote sensing image analysis using paired $Sentinel$-1 and $Sentinel$-2 observations. The dataset contains $549,488$ co-registered image patch pairs collected from 115 $Sentinel$-2 tiles acquired between June 2017 and May 2018 across ten European countries, namely Austria, Belgium, Finland, Ireland, Kosovo, Lithuania, Luxembourg, Portugal, Serbia, and Switzerland. Each image patch is accompanied by a pixel-level reference map together with multiple land-cover annotations derived from the 2018 CORINE Land Cover database, making the dataset well suited for multi-label land-cover classification tasks.

4) CropHarvest [23]: CropHarvest is an open-access benchmark developed for agricultural remote sensing applications by integrating multiple agricultural land-use datasets with satellite observations. The dataset comprises $95,186$ samples, including $33,205$ (35%) instances with multi-class crop labels, while the remaining samples are annotated using binary crop/non-crop labels. Approximately $70,213$ (74%) samples are linked with complementary remote sensing and environmental information obtained from $Sentinel$-1, $Sentinel$-2, the $SRTM$ Digital Elevation Model, and $ERA5$ climatological data, providing a comprehensive resource for crop monitoring and classification studies.

5) MNIST [24]: The MNIST dataset is one of the most widely adopted benchmarks for handwritten digit recognition. It consists of $60,000$ training images and $10,000$ testing images representing the digits 0–9. The dataset was derived from the NIST handwritten digit collections and contains grayscale images that have been size-normalized and centred within a fixed 20×20 pixel region while preserving their original aspect ratios. Owing to its simplicity and standardized evaluation protocol, MNIST is frequently employed for benchmarking image classification algorithms.

6) CIFAR-10 [25] and CIFAR-100 [26]: The CIFAR-10 and CIFAR-100 datasets are standard benchmarks for image classification and visual recognition tasks. Both datasets contain $60,000$ colour images of size 32×32 pixels, divided into $50,000$ training images and $10,000$ testing images. CIFAR-10 categorizes the images into ten mutually exclusive object classes, whereas CIFAR-100 extends the problem to 100 fine-grained classes organised into 20 superclasses, with 600 images available for each class. The substantially larger number of categories and finer semantic distinctions make CIFAR-100 considerably more challenging than CIFAR-10 and therefore useful for evaluating federated learning algorithms under increased classification complexity.

7) Federated Client Partitioning: To evaluate the proposed framework under realistic FL conditions, all datasets were partitioned among 50 simulated clients. For the benchmark datasets (MNIST, CIFAR-10, and CIFAR-100), heterogeneous client data distributions were generated using both the pathological partitioning strategy and the Dirichlet-based partitioning strategy described in Section IV-B. In the pathological setting, each client was assigned samples from a limited subset of classes, whereas in the practical heterogeneous setting, client data distributions were generated using a Dirichlet distribution with concentration parameter $\beta = 0.1$.

For the agricultural datasets (AgriFieldNet, BigEarthNet $v1$, BigEarthNet $v2$, and CropHarvest), the available training samples were partitioned among the same 50 simulated clients using the Dirichlet-based heterogeneous partitioning strategy

with $\beta = 0.1$. Each training sample was assigned exclusively to a single client without overlap, while the corresponding test sets were retained for global model evaluation. Employing a common partitioning strategy across all datasets ensures a consistent and fair evaluation of the proposed ALA framework and the competing FL methods under heterogeneous client data distributions.

The selected agricultural datasets are representative of modern Earth observation applications based on Sentinel imagery and provide a diverse range of crop monitoring and land-cover classification tasks. In practical deployments, Sentinel-derived remote sensing data are typically generated across geographically distributed locations where centralized data collection may be restricted by data ownership, communication constraints, or privacy considerations. Consequently, FL provides a suitable distributed learning paradigm for these applications by enabling collaborative model training without requiring the transfer of raw satellite imagery. The proposed ALA framework is therefore evaluated on representative Sentinel-based datasets to demonstrate its effectiveness under realistic heterogeneous FL environments.

B. Simulation Setup

Simulations were performed on PyTorch version 1.10 as the primary framework. The experiments were conducted on a machine with NVIDIA GTX 1080 Ti GPU, an Intel $i7$-9700K CPU, and 32 GB of RAM, ensuring efficient computation and memory management. The software stack comprised Ubuntu 20.04 as the operating system, CUDA 11.1 for GPU acceleration, and cuDNN 8.0 for optimized deep learning operations.

All experiments, including the benchmark datasets (MNIST, CIFAR-10, and CIFAR-100) and the agricultural datasets (AgriFieldNet, BigEarthNet v1, BigEarthNet v2, and CropHarvest), were conducted using the same 4-layer CNN backbone. The proposed ALA framework operates on the trainable parameters of this network and is independent of the specific dataset being considered. Simulations were performed with a local learning rate of 0.005 for the 4-layer CNN (0.1 for MNIST following FedAvg). In line with FedAvg, the batch size was set to 32, the hybrid regularization strength was fixed at 0.5, and each selected client performed 5 local training epochs during every communication round. To achieve empirical convergence, all experiments were conducted over 500 communication rounds with 50 participating clients. Table IV summarizes the simulation parameters. Individual analyses of the proposed method on the selected datasets are first presented, followed by a comprehensive comparison with state-of-the-art FL methods, including FedAvg [27], FedProx [28], Per-FedAvg [29], FedAvg-C, FedProx-C, FedALA [30], FedFomo [31], and PartialFed [32].

The selected baseline methods represent widely adopted FL approaches encompassing standard federated optimization (FedAvg), optimization under statistical heterogeneity (FedProx), PFL (Per-FedAvg, FedALA, and FedFomo), and communication-efficient model updating (PartialFed). These methods were selected because they are well-established benchmarks for evaluating personalization and heterogeneous client learning. Although several recent aggregation strategies specifically designed for highly heterogeneous or similarity-based client aggregation have subsequently been proposed, incorporating a broader comparison with these methods is beyond the scope of the present study and will be considered in future work.

TABLE I
PARAMETERS FOR SIMULATION SETUP

Parameter	Value
Learning Rate	0.01
Local Training Epochs	5
Batch Size	32
Number of Clients	50
Hybrid Regularization Strength (λ)	0.5
Communication Rounds	500

C. Model Accuracy and Loss

The performance of the proposed federated architecture is evaluated using the overall test accuracy and training loss on AgriFieldNet (Fig. 3a), BigEarthNet $v1$ (Fig. 3b), BigEarthNet $v2$ (Fig. 3c), and CropHarvest (Fig. 3d) datasets. A comparative analysis with representative state-of-the-art methods is presented in Table II. Although BigEarthNet and CropHarvest are commonly evaluated using additional multi-label performance measures such as precision, recall, F1-score, and mean Average Precision (mAP), the present work primarily investigates the effectiveness of the proposed FL framework under heterogeneous client distributions. Therefore, a unified evaluation based on overall test accuracy is adopted across all datasets to ensure a consistent comparison among the competing FL methods.

The proposed framework demonstrates rapid convergence and consistently achieves higher test accuracy than the compared FL methods across all agricultural datasets. These observations indicate that the adaptive layer-wise initialization effectively improves the learning process under heterogeneous client distributions while maintaining stable convergence behaviour. The results further show that the proposed framework consistently improves performance over the training process across datasets with varying levels of image diversity and complexity, demonstrating its robustness and scalability in federated agricultural learning scenarios. While the benchmark datasets (MNIST, CIFAR-10, and CIFAR-100) are included to facilitate comparison with established FL baselines under controlled heterogeneous settings, the proposed framework is additionally validated on four representative agricultural remote sensing datasets, namely AgriFieldNet, BigEarthNet $v1$, BigEarthNet $v2$, and CropHarvest. These datasets encompass diverse crop monitoring and land-cover classification tasks derived from Sentinel imagery, thereby demonstrating the applicability of the proposed framework to practical agricultural terrain mapping scenarios.

D. Performance Comparison and Analysis

1) Heterogeneous Setting: To evaluate the proposed framework under varying degrees of statistical heterogeneity,

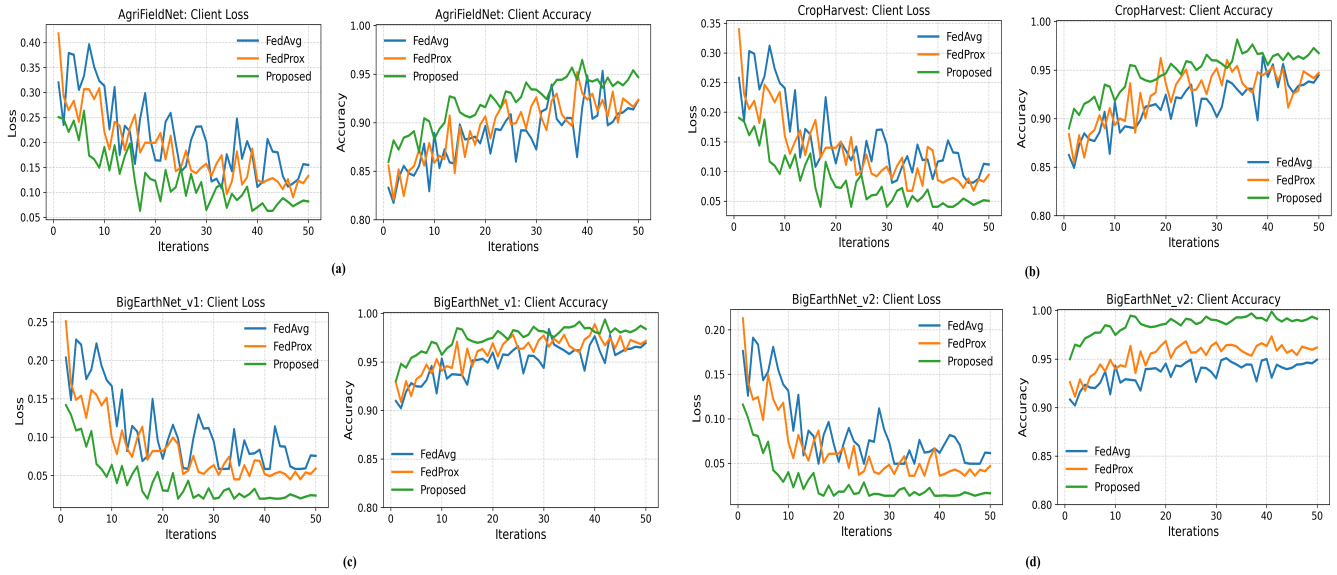


Fig. 3. Agricultural Image Datasets: Client Accuracy/Loss with Iterations: (a) AgriFieldNet, (b) CropHarvest, (c) BigEarthNet (V1) and (d) BigEarthNet (V2)

TABLE II
AGRICULTURAL DATASETS: COMPARATIVE GLOBAL CLASSIFICATION ACCURACY (%)

Method	AgriFieldNet	CropHarvest	BigEarthNet (V1)	BigEarthNet (V2)
FedAvg	92.00±0.35	94.00±0.30	96.50±0.20	94.50±0.18
FedProx	93.00±0.30	95.00±0.24	97.20±0.16	96.20±0.14
Proposed Algorithm	**94.50±0.24**	**96.50±0.18**	**98.20±0.12**	**99.00±0.10**

two widely adopted heterogeneous settings were considered. The first setting, referred to as the *Pathological Heterogeneous Setting*, assigns $2/2/10$ classes to each client for the MNIST/CIFAR-10/CIFAR-100 datasets, respectively, from a total of $10/10/100$ classes, resulting in highly non-IID local data distributions. The second setting is the *Practical Heterogeneous Setting*, which follows a Dirichlet distribution, denoted by $\mathrm{Dir}(\beta)$, where the degree of heterogeneity is inversely proportional to the concentration parameter β. Specifically, smaller values of β correspond to increasingly heterogeneous client data distributions. Unless otherwise specified, β is set to 0.1. The Dirichlet distribution is defined as

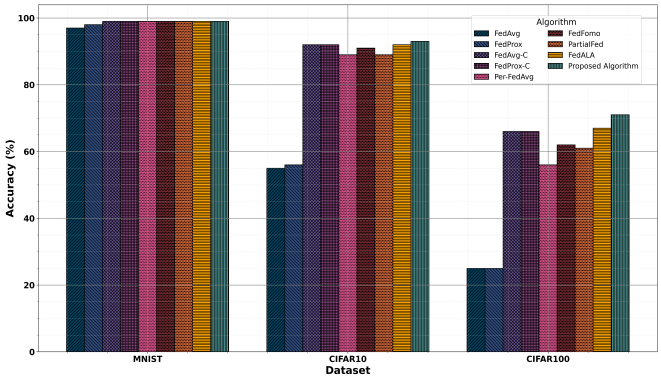
$$P(X_1, X_2, \ldots, X_K; \beta) = \frac{\Gamma\left(\sum_{i=1}^{K} \beta_i\right)}{\prod_{i=1}^{K} \Gamma(\beta_i)} \prod_{i=1}^{K} x_i^{\beta_i - 1}. \quad (7)$$

Table III presents the test accuracies achieved by the competing FL methods under both the pathological and practical heterogeneous settings. Under the pathological setting, FedALA achieves the highest accuracy on MNIST, whereas the proposed ALA framework consistently outperforms all competing methods on the more challenging CIFAR-10 and CIFAR-100 datasets. Similarly, under the practical heterogeneous setting, the proposed framework achieves the highest test accuracy across both benchmark datasets. These results demonstrate that the proposed ALA-based adaptive layer-wise initialization effectively preserves client-specific representations while maintaining strong global model convergence under heterogeneous data distributions. Consequently, the proposed framework achieves an effective balance between personalization and global model performance without modifying the standard federated optimization protocol. Figure 4a further illustrates the comparative performance of the proposed method against the considered state-of-the-art FL approaches.

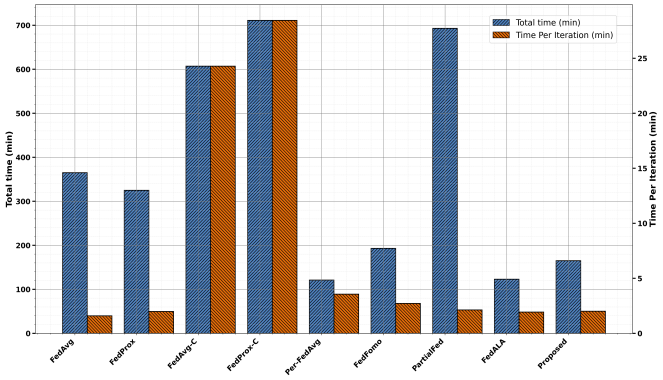
TABLE III
NON-AGRICULTURAL DATASETS: TEST ACCURACY (%) IN DIFFERENT SETTINGS

	Pathological Heterogeneous			Practical Heterogeneous	
Method	MNIST	CIFAR-10	CIFAR-100	CIFAR-10	CIFAR-100
FedAvg	97.93±0.05	55.09±0.83	25.98±0.16	59.16±0.47	31.89±0.47
FedProx	98.01±0.55	56.04±0.75	25.94±0.40	59.21±0.40	31.99±0.41
FedAvg-C	99.79±0.00	92.13±0.03	66.17±0.03	90.34±0.01	51.80±0.02
FedProx-C	99.80±0.04	92.12±0.03	66.07±0.08	90.33±0.01	51.84±0.07
Per-FedAvg	99.63±0.02	89.63±0.23	56.80±0.26	87.74±0.19	44.28±0.33
FedFomo	99.83±0.00	91.85±0.02	62.49±0.22	88.06±0.02	45.39±0.45
PartialFed	99.86±0.01	89.60±0.13	61.39±0.12	87.38±0.08	48.81±0.20
FedALA	99.88±0.01	92.44±0.02	67.83±0.06	90.67±0.03	55.92±0.03
Proposed Algorithm	**99.10±0.01**	**93.93±0.02**	**71.08±0.03**	**95.25±0.01**	**56.37±0.00**

2) Heterogeneity: To study the effectiveness of the proposed algorithm with varying degree of heterogeneity, we vary the β in $\mathrm{Dir}(\beta)$. The Table IV, shows the performance of existing FL methods, under varying degrees of heterogeneity and scalability with different numbers of clients. The performance is validated to verify the applicability of the proposed method, across MNIST and CIFAR-100 datasets. The provided results highlight the performance of the FL methods, specifically FedALA in comparison to the proposed method. The performance of the proposed algorithm is reported to be optimal as it is observed that the accuracy decreases with increasing heterogeneity. The accuracy is reported at 74.24% and 71.20% for $Dir(0.01)$ and $Dir(0.5)$, respectively, on the MNIST dataset. By incorporating ALA, the proposed model ensures a more balanced and representative global model update. This approach addresses the challenges posed by data heterogeneity and enhances the personalization of client models, leading to improved overall performance and stability



(a) Test Accuracy (%) in Heterogeneous Setting



(b) Computation Overhead Analysis

Fig. 4. Non-Agricultural Datasets: Performance Comparison

TABLE IV
NON-AGRICULTURAL DATASETS: TEST ACCURACY (%) WITH VARYING HETEROGENEITY AND SCALABILITY

	Heterogeneity		Scalability	
Dataset	MNIST		CIFAR-100	
Method	Dir(0.01)	Dir(0.5)	Client 50	Client 100
FedAvg	15.70±0.46	21.14±0.47	31.90±0.27	31.95±0.37
FedProx	15.66±0.36	21.22±0.47	31.94±0.30	31.97±0.24
FedAvg-C	49.88±0.11	16.21±0.05	49.82±0.11	47.90±0.12
FedProx-C	49.84±0.02	16.36±0.19	49.79±0.14	48.02±0.02
Per-FedAvg	39.39±0.30	16.36±0.13	44.31±0.20	36.07±0.24
FedFomo	46.36±0.54	11.59±0.11	42.56±0.33	38.91±0.08
PartialFed	49.38±0.02	24.20±0.10	48.95±0.07	39.31±0.01
FedALA	55.75±0.02	27.85±0.06	55.61±0.02	54.68±0.57
Proposed Algorithm	**74.24±0.01**	**71.20±0.13**	**68.25±0.03**	**66.07±0.00**

TABLE V
OVERHEAD ANALYSIS

	Computation Overhead		Communication Overhead
Methods ↓	**Total Time**	**Time per Comm. Round**	**Parameter per iteration**
FedAvg [27]	365 min	1.59 min	$2 * \Theta$
FedProx [28]	325 min	1.99 min	$2 * \Theta$
FedAvg-C [30]	607 min	24.28 min	$2 * \Theta$
FedProx-C [30]	711 min	28.44 min	$2 * \Theta$
Per-FedAvg [29]	121 min	3.56 min	$2 * \Theta$
FedFomo [31]	193 min	2.72 min	$(1 + M) * \Theta$
PartialFed [32]	693 min	2.13 min	$2 * \Theta$
FedALA [30]	7+116 min	1.93 min	$2 * \Theta$
Proposed Method	**165 min**	**2.01 min**	$\mathbf{2 * \Theta}$

in the FL environment.

3) Scalability: In terms of scalability, only a modest reduction in accuracy is observed as the number of participating clients increases from 50 to 100, demonstrating that the proposed framework maintains stable performance under larger-scale FL settings. As reported in Table IV, the proposed method consistently outperforms the competing approaches while exhibiting only a limited degradation in accuracy as the number of participating clients increases. The ALA-based adaptive initialization enables each client to effectively preserve personalized representations while remaining aligned with the global model, thereby improving robustness under increasing system scale.

The proposed ALA framework is intended to improve personalization under heterogeneous client data distributions while preserving strong global model performance. Accordingly, the experimental evaluation in this work primarily focuses on the global test accuracy, which provides a consistent basis for comparison with existing FL approaches. The observed improvements across multiple agricultural and benchmark datasets indicate that the adaptive layer-wise initialization enhances personalization without adversely affecting the overall global model convergence.

The scalability evaluation presented in this work is limited to FL environments comprising up to 100 participating clients. Although the proposed ALA framework performs all adaptive initialization locally and preserves the standard client-server communication protocol without introducing additional communication rounds, further experimental validation on substantially larger federated deployments involving hundreds or thousands of clients would provide a more comprehensive assessment of scalability.

4) Computation and Communication Overhead: The computational overhead of the proposed framework arises primarily from the client-side ALA initialization performed before local model training. Since the aggregation weights are learned only for the selected higher layers using a subset of the local training data, the additional computational cost remains modest. Table V reports the total computation time and the average computation time per communication round for all compared methods. The proposed framework requires only a small increase in computation time compared with FedAvg while providing improved performance under heterogeneous client data distributions. Figure 4b further illustrates the computation overhead of the considered FL methods.

Communication overhead is defined as the amount of model parameters exchanged between each participating client and the central server during a communication round. Since the proposed ALA mechanism performs adaptive layer-wise initialization entirely on the client side before local training, it does not introduce any additional communication beyond the standard FL protocol. Consequently, each client uploads one local model and downloads one global model during every communication round. Table V summarizes the communica-

tion overhead of the compared methods, where Θ denotes the number of model parameters and M represents the number of downloaded client models.

V. Conclusion

FL applications of agricultural terrain mapping are constrained by the unique variability in data distribution among clients and biased updates. Data distributions vary significantly among clients, making a one-size-fits-all global model less effective. This work presents a FL framework for agricultural terrain mapping. The idea is to distribute the preprocessed client data in the form of private datasets to different clients. Personalized variants are created by the clients and are trained locally through a combination of *FedAvg* and *FedProx* aggregation. By assessing these deviations, clients can dynamically adjust the influence of their updates, ensuring that the global model remains on a convergent path while personalized models continue to adapt to local data characteristics. The proposed method is able to reconcile discrepancies between local and global models, applied for image mapping of agricultural lands. Results are validated on 7 publicly available datasets (4 agricultural image datasets and 3 general image datasets), where the performance is validated with existing methods. The proposed method is able to outperform the existing models on diverse parameters and settings, with the proposed method reaching the accuracy of 99.05% on agricultural datasets and 97.25% on non-agricultural datasets. From a practical deployment perspective, agricultural IoT and UAV-assisted monitoring systems often operate under constrained communication bandwidth, intermittent network connectivity, limited onboard computational resources, and highly heterogeneous data collected across different geographical regions and crop types. Since the proposed ALA framework performs adaptive initialization locally while preserving the standard FL communication protocol, it is well suited to such distributed environments without requiring the exchange of raw sensing data or additional communication rounds. Nevertheless, practical deployment in large-scale agricultural systems would require further investigation of issues such as device heterogeneity, asynchronous client participation, communication reliability, and energy-efficient operation. These aspects constitute important directions for future research. Furthermore, Integrating formal DP mechanisms with the proposed ALA framework and incorporating Byzantine-resilient aggregation strategies and adversarial client detection may also be considered in future.

References

[1] V. K. Singh, D. Agarwal, V. K. Gediya, R. S. Rathore, and W. Jiang, "Mitigating class imbalance in forest fire prediction with gan-augmented data fusion," *Information Fusion*, p. 104005, 2025.

[2] V. K. Singh, S. Chakraborty, R. Singh, R. S. Rathore, and W. Jiang, "Bias-aware data quality enhancement for forest fire detection in ai-based remote sensing," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 2026.

[3] V. K. Singh, J. Singh, R. S. Rathore, and D. Nema, "Gradient-based compression and approximate computing for deep network optimization in multimedia data processing," *IEEE Transactions on Consumer Electronics*, 2026.

[4] V. K. Singh, N. Anand, A. Pal, A. Chowdhury, and A. Srivastava, "Anomaly detection in dynamic graphs using multiple encoding strategies via transformers," *IEEE Transactions on Consumer Electronics*, 2025.

[5] V. K. Singh, N. Jain, G. Tripathi, and S. Sahani, "Correlated channeled spatio temporal graph attention network model for traffic prediction," *IEEE Transactions on Intelligent Transportation Systems*, 2025.

[6] V. K. Singh, C. Singh, and H. Raza, "Event classification and intensity discrimination for forest fire inference with iot," *IEEE sensors journal*, vol. 22, no. 9, pp. 8869–8880, 2022.

[7] A. Tariq, J. Yan, A. S. Gagnon, M. Riaz Khan, and F. Mumtaz, "Mapping of cropland, cropping patterns and crop types by combining optical remote sensing images with decision tree classifier and random forest," *Geo-Spatial Information Science*, vol. 26, no. 3, pp. 302–320, 2023.

[8] L. Hashemi-Beni and A. A. Gebrehiwot, "Flood extent mapping: An integrated method using deep learning and region growing using uav optical data," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 14, pp. 2127–2135, 2021.

[9] S. D. Meena and L. Agilandeeswari, "Smart animal detection and counting framework for monitoring livestock in an autonomous unmanned ground vehicle using restricted supervised learning and image fusion," *Neural Processing Letters*, vol. 53, no. 2, pp. 1253–1285, 2021.

[10] Z. Yin and R. Collins, "Moving object localization in thermal imagery by forward-backward motion history images," *Augmented Vision Perception in Infrared: Algorithms and Applied Systems*, pp. 271–291, 2009.

[11] C. Benedek, T. Szirányi, Z. Kato, and J. Zerubia, "Detection of object motion regions in aerial image pairs with a multilayer markovian model," *IEEE Transactions on Image Processing*, vol. 18, no. 10, pp. 2303–2315, 2009.

[12] X. Shi, Z. Huang, W. Bian, D. Li, M. Zhang, K. C. Cheung, S. See, H. Qin, J. Dai, and H. Li, "Videoflow: Exploiting temporal cues for multi-frame optical flow estimation," *arXiv preprint arXiv:2303.08340*, 2023.

[13] S. Anthwal and D. Ganotra, "An overview of optical flow-based approaches for motion segmentation," *The Imaging Science Journal*, vol. 67, no. 5, pp. 284–294, 2019.

[14] Y. Yang, T. Zhang, J. Hu, D. Xu, and G. Xie, "End-to-end background subtraction via a multi-scale spatio-temporal model," *IEEE Access*, vol. 7, pp. 97 949–97 958, 2019.

[15] M. Vijayan, P. Raguraman, and R. Mohan, "A fully residual convolutional neural network for background subtraction," *Pattern Recognition Letters*, vol. 146, pp. 63–69, 2021.

[16] X. Cui, L. Sun, K. Zhang, and W. Zhang, "Multi-motion target recognition and perception method based on information fusion," in *2020 Chinese Automation Congress (CAC)*. IEEE, 2020, pp. 6426–6431.

[17] V. K. Singh, N. Anand, S. K. Sharma, M. K. Shukla, R. S. Rathore, W. Jiang *et al.*, "Low-light image enhancement for edge-based security surveillance in 6g-iot visual systems," *IEEE Internet of Things Journal*, 2025.

[18] N. Jain, C. Singh, V. K. Singh, R. S. Rathore, N. S. Alghamdi, and C. Hewage, "Federated genetic optimization for secure and privacy-preserving sensor localization in consumer iot applications," *IEEE Transactions on Consumer Electronics*, 2025.

[19] S. Srivastava, D. Agarwal, B. K. Chaurasia, V. K. Singh, R. S. Rathore, and W. Jiang, "Federated blockchain-based resilient identity management for securing consumer vehicular networks," *IEEE Transactions on Consumer Electronics*, 2025.

[20] R. E. Foundation, "Agrifieldnet india dataset," https://source.coop/radiantearth/agrifieldnet-competition, 2022, accessed: 2025-05-25.

[21] G. Sumbul, M. Charfuelan, B. Demir, and V. Markl, "Bigearthnet: A large-scale benchmark archive for remote sensing image understanding," *arXiv preprint arXiv:1902.06148*, 2019. [Online]. Available: https://arxiv.org/abs/1902.06148

[22] K. N. Clasen, L. Hackel, T. Burgert, G. Sumbul, B. Demir, and V. Markl, "reben: Refined bigearthnet dataset for remote sensing image analysis," *arXiv preprint arXiv:2407.03653*, 2024.

[23] G. Tseng, I. Zvonkov, C. L. Nakalembe, and H. Kerner, "Cropharvest: A global dataset for crop-type classification," in *Thirty-fifth Conference on Neural Information Processing Systems Datasets and Benchmarks Track (Round 2)*, 2021. [Online]. Available: https://openreview.net/forum?id=JtjzUXPEaCu

[24] Y. LeCun, C. Cortes, and C. J. C. Burges, "Mnist handwritten digit database," *ATT Labs [Online]*, 2010. [Online]. Available: http://yann.lecun.com/exdb/mnist

[25] A. Krizhevsky, "Learning multiple layers of features from tiny images," University of Toronto, Tech. Rep., 2009. [Online]. Available: https://www.cs.toronto.edu/~kriz/cifar.html

[26] ——, "Learning multiple layers of features from tiny images," University of Toronto, Tech. Rep., 2009. [Online]. Available: https://www.cs.toronto.edu/~kriz/cifar.html
[27] B. McMahan, E. Moore, D. Ramage, S. Hampson, and B. A. y Arcas, "Communication-efficient learning of deep networks from decentralized data," in *Artificial intelligence and statistics*. PMLR, 2017, pp. 1273–1282.
[28] T. Li, A. K. Sahu, M. Zaheer, M. Sanjabi, A. Talwalkar, and V. Smith, "Federated optimization in heterogeneous networks," *Proceedings of Machine learning and systems*, vol. 2, pp. 429–450, 2020.
[29] A. Fallah, A. Mokhtari, and A. Ozdaglar, "Personalized federated learning: A meta-learning approach," *arXiv preprint arXiv:2002.07948*, 2020.
[30] J. Zhang, Y. Hua, H. Wang, T. Song, Z. Xue, R. Ma, and H. Guan, "Fedala: Adaptive local aggregation for personalized federated learning," in *Proceedings of the AAAI conference on artificial intelligence*, vol. 37, no. 9, 2023, pp. 11 237–11 244.
[31] M. Zhang, K. Sapra, S. Fidler, S. Yeung, and J. M. Alvarez, "Personalized federated learning with first order model optimization," 2021. [Online]. Available: https://arxiv.org/abs/2012.08565
[32] B. Sun, H. Huo, Y. Yang, and B. Bai, "Partialfed: Cross-domain personalized federated learning via partial initialization," *Advances in Neural Information Processing Systems*, vol. 34, pp. 23 309–23 320, 2021.



Vishal Krishna Singh received his bachelor's degree in Information Technology, in 2010, the master's degree in Computer Technology and Application, in 2013, and PhD degree in Information Technology from Indian Institute of Information Technology, Allahabad, India in 2018. He is currently working as a Lecturer and is associated with the Networks and Communications Research Group at School of Computer Science and Electronics Engineering, University of Essex, Colchester, U.K. His research interests include Internet of Things, Wireless Sensor Networks, In-Network Inference, Machine Learning and Data Analytics.



Nitin Shkula received M.Tech. and PhD degree from the Motilal Nehru National Institute of Technology Allahabad, India. He is currently an assistant professor with the School of Computer Science Engineering and Technology, Bennett University, Greater Noida, India. His research interests include software-defined networks, fog and edge computing, peer-to-peer networks, internet of things, and localization protocols.



Deepak Kumar received the bachelor's degree in Information Technology in 2016, the master's degree in Computer Science from Indian Institute of Information Technology, Lucknow, India in 2021. Currently he is working as Software Developer 2, Dell Technologies, Bengaluru, Karnataka, India. His research interests include Machine Learning, Internet of Things, and Data Analytics.



Neeraj Jain received his bachelor's degree in Information Technology, in 2007, the master's degree in Information Technology, in 2010, and PhD degree in Information Technology from Indian Institute of Information Technology, Allahabad, India in 2018. He is currently working as an Associate Professor at School of Computer Science Engineering and Technology, Bennet University, Greater Noida, India. His research interests include Federated Learning, Machine Learning, Internet of Things, Wireless Sensor Networks, and Localization Protocols.



Chhaya Singh is an Associate Professor in the School of Computer Science and Engineering at Galgotias University, India. She received her Ph.D. from MANIT Bhopal and M.Tech. from IIIT Allahabad. Her research interests include Machine Learning, Data Science, and Fuzzy Logic. She has published several SCOPUS and SCI-indexed research articles and holds a design patent. Dr. Singh has also worked as a Research Assistant on a GCRF-funded project in collaboration with the University of Essex, UK. She has received multiple research fellowships and the Young Scientist Award for her contributions to computational biology.



Dr. Rajkumar Singh Rathore (Senior Member IEEE) is working as Head of Cyber Security of Connected and Autonomous Systems, CINC, Head of Cyber Physical and Networks Systems, CeRISS & Programme Director for MSc Computing and IT in Department of Computer Science at Cardiff Metropolitan University's School of Technologies, United Kingdom. He has gained doctorate degree, dual master's degrees, and bachelor's degree all in Computer Science and Engineering discipline. Dr Rathore is a scholar throughout his career. He is the Fellow of HEA UK. He has several years of rich experience in quality of teaching, learning and research excellence. His research works were fully supported by Nottingham Trent University, United Kingdom and Manchester Metropolitan University, United Kingdom. His research expertise are Wireless Communications, Internet of Things/Cyber Physical Systems, Cyber Security and Privacy, Connected and Autonomous Vehicles, EV Charging Infrastructure Management, Intelligent Networking of Drones and also use cases of AI/ML.



Prof. Weiwei Jiang (Senior Member IEEE) received the B.Sc. Degree of Electronic Engineering and Ph.D. Degree of Information and Communication Engineering from the Department of Electronic Engineering, Tsinghua University, Beijing, China, in 2013 and 2018, respectively. He is currently an assistant professor with the School of Information and Communication Engineering, Beijing University of Posts and Telecommunications, and Key Laboratory of Universal Wireless Communications, Ministry of Education. His current research interests include artificial intelligence, machine learning, big data, wireless communication and edge computing. He has published more than 60 academic papers in IEEE Trans and other journals, with more than 3900 citations in Google Scholar. He is one of 2023 and 2024 Stanford's List of World's Top 2% Scientists.