Computers and Electronics in Agriculture 229 (2025) 109720

Contents lists available at ScienceDirect

Computers and Electronics in Agriculture

journal homepage: www.elsevier.com/locate/compag

Adaptive knowledge transfer using federated deep learning for plant disease detection

Pragya Hari *, Maheshwari Prasad Singh

Department of Computer Science and Engineering, National Institute of Technology Patna, India

ARTICLE INFO

Keywords:
Federated deep learning
Adaptive collaborative learning
Plant disease detection
Non-IID

ABSTRACT

Federated learning has revolutionized machine learning in decentralized settings, enabling collaborative model training across distributed devices. In the domain of plant leaf disease detection, where timely intervention is vital to protect crop yields, federated learning holds immense promise. However, the challenge of an imbalanced, non-independent, and identically distributed (Non-IID) plant disease dataset from diverse agricultural regions poses obstacles in achieving accurate and efficient models. Thus, this paper introduces Intelligent Weight Transferring method, an approach tailored for plant leaf disease detection in federated deep learning (FDL-IWT). The method adapts weightage coefficients based on child (local) model knowledge analysis, integrating more representative knowledge for improved model convergence and accuracy. Moreover, an efficient and lightweight Parallel MultiScale CNN architecture based on attention mechanism (PMACNN) with 0.13 million parameters is proposed for parent–child entity of federated learning. In order to assess the robustness of the proposed FDL-IWT method, it is evaluated over two distinct publicly available plant leaf disease dataset, simulating the configuration of federated learning. The obtained simulation results show that the proposed FDL-IWT method attains the highest testing accuracy rate of 97.5%, outperforming seven distinct state-of-the-art methods: FedAvg, FedAdam, FedAdagrad, CWT, FedAdp, Median and Trimmed-mean. The findings enforce the potential of Intelligent Weight Transferring method to significantly elevate the effectiveness of plant leaf disease detection in federated deep learning, empowering sustainable and efficient agricultural practices.

1. Introduction

Leaf diseases have a profound impact on human life by affecting agriculture, economy, and well-being. As a result, crop loss leads to food scarcity and higher prices, especially in less developed countries where annual losses of 30%–50% are common for major crops.[1] Detection of such diseases is a critical task to ensure early identification and timely intervention to prevent crop losses. Deep learning has proven highly effective in this due to its capacity to learn intricate patterns from images. Techniques like convolutional neural networks (CNNs) automatically extract relevant features from leaf images, supported by transfer learning and data augmentation for improved performance. Nonetheless, deep learning encounters challenges such as need for large amount of datasets and huge computational resources. In an analysis, it is shown that out of total stored data by an organization, at least 30% of them are redundant, obsolete, or trivial (ROT).[2] Therefore, storing extensive amount of data possess high operational costs, data management complexity, compliance challenges and performance degradation, emphasizing the necessity for robust and efficient federated framework with sustainable solutions.

Federated learning offers feasible solutions to such challenges of deep learning. It allows training of multiple local models collaboratively, without centralizing the dataset. Considering, training of robust leaf disease detection models across nation without accumulating the agricultural dataset at a single place. However, in real-world scenarios, the dataset collected from various regions may exhibit non-independent and identically distributed (Non-IID) characteristics due to variations in climate, soil conditions and farming practices. This presents considerable challenges for federated learning, as it may hinder model convergence and accuracy. Aggarwal et al. (2023a,b) present basic implementation of federated learning over IID and non-IID Rice diseased dataset, in order to show superiority of federated learning over

* Corresponding author.
E-mail addresses: pragyah.phd20.cs@nitp.ac.in (P. Hari), mps@nitp.ac.in (M.P. Singh).

1 www.britannica.com/science/plant-disease

2 blogs.manageengine.com/active-directory/datasecurity-plus-active-directory/2020/10/16/dealing-with-data-glut-why-rot-data-is-an-issue-and-how-to-manage-it.html

https://doi.org/10.1016/j.compag.2024.109720
Received 12 October 2023; Received in revised form 19 September 2024; Accepted 24 November 2024
Available online 6 December 2024

traditional deep learning techniques. The challenges which may arise because of non-IID dataset such as statistical heterogeneity, and concept drift are still unexplored.

This paper proposes the "Intelligent Weight Transferring" method coupled with "Federated Deep Learning" tailored specifically for plant leaf disease detection. This approach employs an analysis of the performance of child (local) models using a decision system to extract knowledge and subsequently determines optimized weightage coefficients for these models. By integrating this adaptive method, the federated framework achieves enhanced accuracy and convergence across non-IID datasets.

The main contributions of this paper are as follows:

1. This work proposes a novel adaptive federated deep learning method in the presence of Non-IID dataset. It intelligently assigns optimal weightage coefficients to each local model in order to determine their contributions in aggregation. These optimal coefficients are determined by following two steps: (i) Developing a decision-making system based on online-trained parameters and (ii) Determining optimal coefficients using the Dirichlet distribution through a developed decision-making system.
2. In addition to the proposed FDL-IWT (Federated Deep Learning with Intelligent Weight transfer) method, this work also presents formulation of two distinct Non-IID plant diseased datasets, simulating the federated learning configurations. The datasets include: (i) Potato, Strawberry, Tomato; (ii) Banana, Coconut, Grape.
3. Additionally, a Parallel MultiScale Attention Convolutional Neural Network (PMACNN) architecture is designed for both parent and child models of FDL-IWT. The architecture is characterized by its efficiency and lightweight nature, containing 0.13 million parameters and a model size of 0.57 MB, which may potentially contribute in reducing communication costs.

The rest of this paper is organized as follows. Section 2 presents the related work of deep learning based models and adaptive federated learning. Section 3 elaborates the proposed intelligently transferring knowledge method in federated framework(FDL-IWT) for detecting plant leaf diseases and proposed Parallel MultiScale Attention CNN(PMACNN) architecture for parent and child models. Then, Section 4 shows the obtained simulation results on two distinct formulated plant diseased federated datasets and compares with several state-of-the-art methods including FedAvg, FedAdam, FedAdagrad, CWT, FedAdp, Median, and Trimmed-mean. Section 5 discusses the effectiveness of proposed FDL-IWT method and PMACNN architecture. Section 6 concludes this paper.

2. Related work

This section discusses related work of deep learning techniques for plant disease detection followed by adaptive federated learning methods.

2.1. Deep learning techniques

Deep learning (DL) has gained significant importance in the agriculture domain (Zheng et al., 2023) due to its ability to process and analyze large amounts of data, leading to several benefits and advancements. DL based architectures has the potential to significantly improve productivity, reduce costs and promote sustainable farming practices (Huang et al., 2023). Arun and Umamaheswari (2023) propose Complete Concatenated Deep Learning (CCDL) architecture for detecting crop diseases. Complete Concatenated Block (CCB) is the core of this architecture consisting pointwise convolution followed by standard convolution. The study introduces three different variations of proposed architecture namely (i) CCDL with Standard Convolutional Technique (CCDL-SCT), (ii) CCDL with Depthwise Separable Convolutional Technique (CCDL-DSCT) and (iii) CCDL with Partial Standard Convolutional Technique (CCDL-PSCT). Among these, CCDL-PSCT attains highest efficiency rate of 98.14%.

Thakur et al. (2023) introduce "VGG-ICNN", a lightweight CNN model comprising four initial layers of VGG16 and three blocks of GoogleNet InceptionV7. The proposed model evaluated over five publicly available datasets of varying sizes viz., PlantVillage(38), Embrapa[3] (93), Apple (4), Maize (4) and Rice (5). The results show that the proposed model attains 99.16% performance rate over PlantVillage dataset and consistently performs well on all five datasets.

Kaya and Gürsoy (2023) develop Fused DenseNet121 for plant disease detection by fusing RGB and segmented images. Authors obtained the accuracy of 98.17% over the 38 classes of PlantVillage datasets. Amin et al. (2022) utilize concatenation of EfficientNetB0 and DenseNet121 for classifying diseases in corn leaf. The authors obtain an accuracy rate of 98.56%, showing superiority over ResNet152 (98.37%) and InceptionV3 (96.26%). Chen et al. (2021) utilize MobileNetV2 to enhance the learning capability for minute lesion features, and incorporates the attention module (Zhang et al., 2021b) to learn the importance of inter-channel relationship and spatial points for input features. The proposed model obtains 99.67% over Rice dataset.

Yang et al. (2023) propose RE-GoogLeNet based on residual network and attention mechanism. The proposed network shows an improvement of 1.72% over the original GoogLeNet model. This network replaces the 7×7 conv kernel in the first layer of GoogLeNet with three 3×3 conv kernel and adds an Efficient Channel Attention (ECA) mechanism to the inception module. Sharma et al. (2023) suggest deep lightweight multi-class (DLMC-Net) model, consisting sequence of collective blocks along with the passage layer. The proposed model attains promising results over citrus(93.56%), cucumber(92.34%), grapes(99.50%) and tomato(96.56%). Table 1 summarizes the related works of plant disease detection.

2.2. Adaptive federated learning

The review on adaptive federated learning (FL) methods with non-IID datasets reveals a growing interest in addressing data heterogeneity challenges in federated learning settings. Wu and Wang (2021) propose FedAdp (Federated Adaptive) algorithm that accelerates model convergence rate in the presence of non-IID dataset. This algorithm assigns adaptive weights based on node contributions and dataset size, reducing communication rounds up to 54.1% for MNIST and 45.4% for FashionMNIST, compared to FedAvg (McMahan et al., 2017). Zhang et al. (2021a) integrates DRL (Deep Reinforcement Learning) algorithm with FL to address non-IID data training under resource constraints. This approach adapts to dataset heterogeneity by dynamically adjusting the learning rate and aggregating model updates based on their respective local dataset.

Ilhan et al. (2023) propose ScaleFL framework, addressing client resource heterogeneity. This approach involves scaling down deep neural networks in width and depth, followed by optimizing aggregated client knowledge using self-distillation. Experimental results on CIFAR-10/100 and ILSVRC2012 datasets showcase remarkable reductions in inference latency (up to 2×) and model size (up to 4×) with minimal performance drop (below 2%). Ahmad et al. (2023) introduce Hessian-Weighted Aggregation (HWA) method for achieving robust Federated Learning under statistical heterogeneity. This method attains 78.02% accuracy, outperforming FedAvg (McMahan et al., 2017) (74.87%), FedProx (74.78%), and FedCurv (75.08%) on FEMNIST dataset.

Zhang et al. (2023a) propose a novel personalized federated learning technique called Federated Learning with Adaptive Local Aggregation (FedALA). This technique adaptively aggregates global and local

[3] www.digipathos-rep.cnptia.embrapa.br

Table 1
Summary of related works for plant disease classification.

Authors	Methods	Dataset	Accuracy	Parameters
Arun and Umamaheswari (2023)	CCDL	PlantVillage	98.14%	2.87 millions
Amin et al. (2022)	Fused EfficientNetB0-DenseNet121	Corn	98.56%	16.20 millions
Thakur et al. (2023)	VGG-ICNN	PlantVillage	99.16%	6 millions
Kaya and Gürsoy (2023)	Fused DenseNet121	PlantVillage	98.17%	–
Chen et al. (2021)	Mobile-Atten	Rice	99.67%	–
Yang et al. (2023)	RE-GoogLeNet	Rice	99.58%	9,183,189
Sharma et al. (2023)	DLMC-Net	Citrus Cucumber Grapes Tomato	93.56% 92.34% 99.50% 96.56%	24 millions

Table 2
Summary of related works for adaptive federated learning.

Authors	Proposed objective	Techniques used	Performance measure
FedAdp (Wu and Wang, 2021)	Accelerate Global Model Convergence with Non-IID dataset	Measurement of Node Contribution and then assigning adaptive weights to participating nodes	Communication rounds reduced up to 54.1% - MNIST and 45.4% - FashionMNIST
Adaptive-B (Zhang et al., 2021a)	Optimize local training and global aggregation with Non-IID dataset	Selecting subset of participants on the basis of hyperparameters(batch size, and number of local updates) and applying Experience driven DRL to control training of local models adaptively	Model accuracy improves by up to 30%
ScaleFL (Ilhan et al., 2023)	Client resource heterogeneity	Adaptively scales down DNN model by using early exits to find best fits models for resource-aware local training and utilizes self-distillation technique to enhance the aggregation process	latency reduces up to 2×and model size reduces up to 4×with negligible performance drop (below 2%)
HWA (Ahmad et al., 2023)	Statistical heterogeneity	Using Hessian as a scaling matrix to Quasi-Newton methods, assigns adaptive weights to local models during aggregation	92.18% - MNIST, 76.88% - Fashion-MNIST
FedALA (Zhang et al., 2023a)	Statistical Heterogeneity	Adaptive Local Aggregation module used to initialize local model before training	Improves test accuracy by up to 3.27%

models at an element-wise level, outperforming eleven standard baselines by up to 3.27% in testing accuracy across diverse benchmark datasets (MNIST, CIFAR10/100, Tiny-ImageNet, AG News and fastText). Hou et al. (2023) present the Adaptive Training and Aggregation Federated Learning (ATAFL) framework, incorporating Twin-Delayed Deep Deterministic (TD3) algorithm tailored for multi-tier networks. They address the joint optimization problem involving training, aggregation node selection, and resource allocation (TASRS). The related works of adaptive federated learning are presented in Table 2. In order to simulate the federated configuration for plant disease detection, this work presents an adaptive approach to handle challenges of non-IID dataset. However, this work can also be explored for IID dataset.

3. Proposed intelligently transferring knowledge method in federated framework

Federated learning (FL) assumes that each child possesses its own local dataset, and these datasets may have different statistical characteristics of different crops, which makes it difficult to assign appropriate weights to child models due to Non-IID data distribution. Considering number of samples per child may lead to biased aggregation, as some child may have more representative or important data for the overall learning task. Thus, this paper proposes a method for transferring knowledge among child models in Federated Framework under the presence of Non-IID dataset.

The proposed method includes two steps: (i) Developing Decisive System using Online Trained Parameters to analyze knowledge performance and (ii) Finding Optimized Dirichlet based Weightage Coefficient for Enhanced Model Training. These steps are discussed as follows

1. **Developing Decisive System using Online Trained Parameters to analyze knowledge performance**

This is first step of the proposed method, which develops a decision making system driven by multi batch training and knowledge performance analysis. A key component of this stage involves the training of a machine learning regressor model, which plays a crucial role in selecting the optimal weightage coefficients for child models. The workflow for developing Decisive System using Online Trained Parameters are shown in Fig. 1 and illustrated below in detail:

(a) Initial Chunking: The available dataset is divided into different smaller subsets, known as batches (In this case batch size is 100). This step is performed to process the data in smaller portions efficiently.
(b) Iterative Training: The pre-trained InceptionV3 (Szegedy et al., 2016) model is trained using each individual batch of data. This involves passing the batched data through the model and adjusting its parameters to improve its performance on the specific batch.
(c) Storing Knowledge and Performance: As the model is trained on each batch, the corresponding weights (generated by the model) and the achieved performance are stored in a CSV file. This allows for tracking the model's progress and performance across different batches.
(d) Assessing Knowledge and Performance Change: Following training with all dataset batches, the method computes changes in both the model's knowledge and performance, gauging variations in learned parameters and accuracy improvements over the training process.
(e) Decisive System and Metric Evaluation: A machine learning regressor model is trained on a newly formed CSV dataset containing knowledge change and performance change values. This model predicts performance change from knowledge change, termed as the decisive system

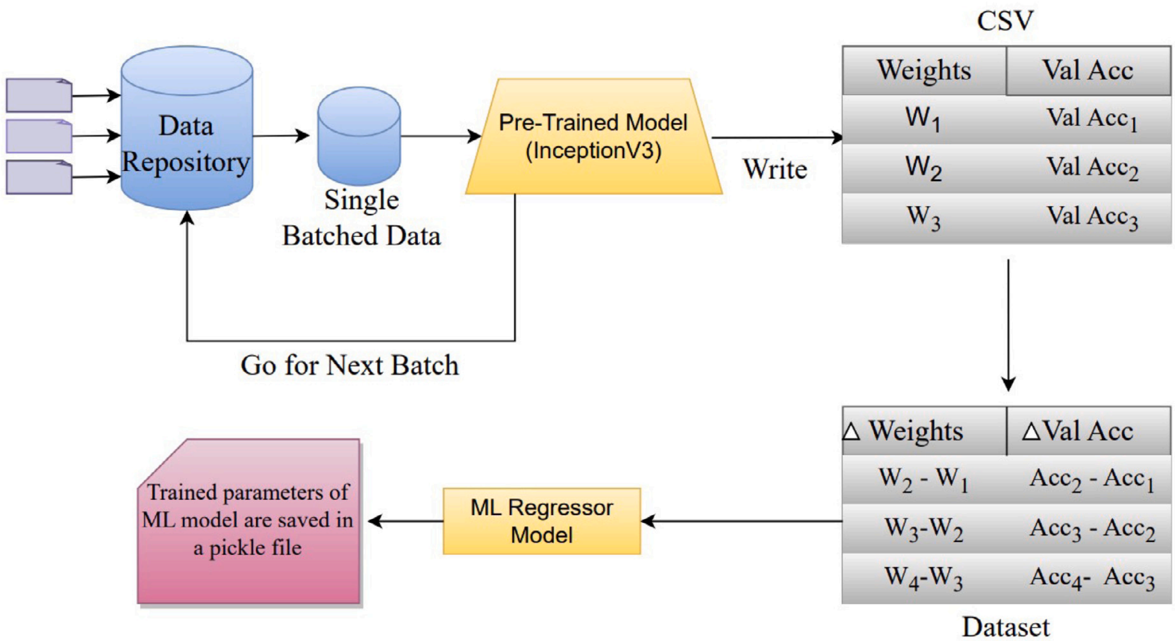

Fig. 1. Workflow of Decisive System using Online Trained Parameters to analyze knowledge performance.

analyzing knowledge performance. Evaluation is done using mean squared error (MSE), root mean square error (RMSE), mean absolute error (MAE) and coefficient of determination (R^2).

(f) Saving the Model: The ML regressor model with the optimal performance or the lowest RMSE value is saved in a pickle file, encapsulating its parameters and trained behavior. This pickle file can be readily loaded into memory, enabling the estimation of performance changes based on varying knowledge updates.

2. **Finding Optimized Dirichlet based Weightage Coefficient for Enhanced Model Training**
The workflow for Finding Optimized Dirichlet based Weightage Coefficient for Enhanced Model Training are shown in Fig. 2 and illustrated below in detail:

(a) Dirichlet mathematical model (Khorchef et al., 2023): The Dirichlet mathematical model is employed to derive a list of potential combinations of weightage coefficients. The model calculates these coefficients based on concentration parameters and point coordinates, using Eq. (1).

$$f(x_1, x_2, \ldots, x_d; \alpha_1, \alpha_2, \ldots, \alpha_d) = (1/\beta(\alpha)) * \pi(x_i \exp(\alpha_i - 1)) \tag{1}$$

$\alpha = (\alpha_1, \alpha_2, \ldots, \alpha_d)$ are concentration parameters x_1, x_2, $\ldots$, x_d are the coordinates of a point $\beta(\alpha) = (\pi\gamma(\alpha_i)) / \gamma(\sum \alpha_i)$ where γ is the gamma function

(b) Determining knowledge over combinations: The parent entity assesses its knowledge over all the generated combinations of weightage coefficients. This involves evaluating the model's performance and understanding the impact of each combination on the overall knowledge.

(c) Calculating knowledge change: The change in knowledge is calculated by comparing the current knowledge with the previously acquired knowledge. This step quantifies the difference in the model's understanding and representation of the data.

(d) Predicting performance change: With the help of developed decisive system (saved ML model) from the previous stage, the change in performance rate is predicted for each knowledge change. This system is utilized to estimate how the performance of the child (local) model is likely to be affected by variations in knowledge.

(e) Optimal Weightage Selection and Parent's Knowledge Updates: The process involves identifying the optimal weightage coefficient by comparing predicted performance changes and selecting the combination that maximizes performance. This coefficient is then utilized to update the parent's knowledge, enhancing the model's data understanding. In the next communication phase, the best coefficient is used as a concentration parameter in the Dirichlet model, influencing knowledge update distribution and guiding future learning iterations.

3.1. Proposed parallel multiscale attention CNN architecture for parent and child

In addition to the proposed federated method, this work also presents Parallel Multiscale CNN architecture based on attention mechanism (PMACNN) for both parent and child entities. In order to perform efficiently with limited computational resources, this newly proposed model is designed to be lightweight, featuring 0.13 million parameters and 0.57 MB model size. Fig. 3 depicts the proposed PMACNN architecture.

The input images undergo processing in two parallel modules, extracting features at different scales. Each module consists of three convolutional layers with varying kernel sizes, two maxpool layers and a concatenate layer. The initial stage involves individual processing within each module, passing through two consecutive convolutional(conv) layers.

The convolutions within the first module employ kernel sizes of 4 × 4 and 3 × 3, while the second module utilizes kernel sizes of 5 × 5 and 2 × 2. The features extracted from the second conv layer are simultaneously fed to the third conv layer with a kernel size of 1 × 1 and a maxpool layer. Furthermore, the extracted features from this last conv layer are parallelly processed with other maxpool layer and simultaneously concatenated with the features obtained from the second conv layer.

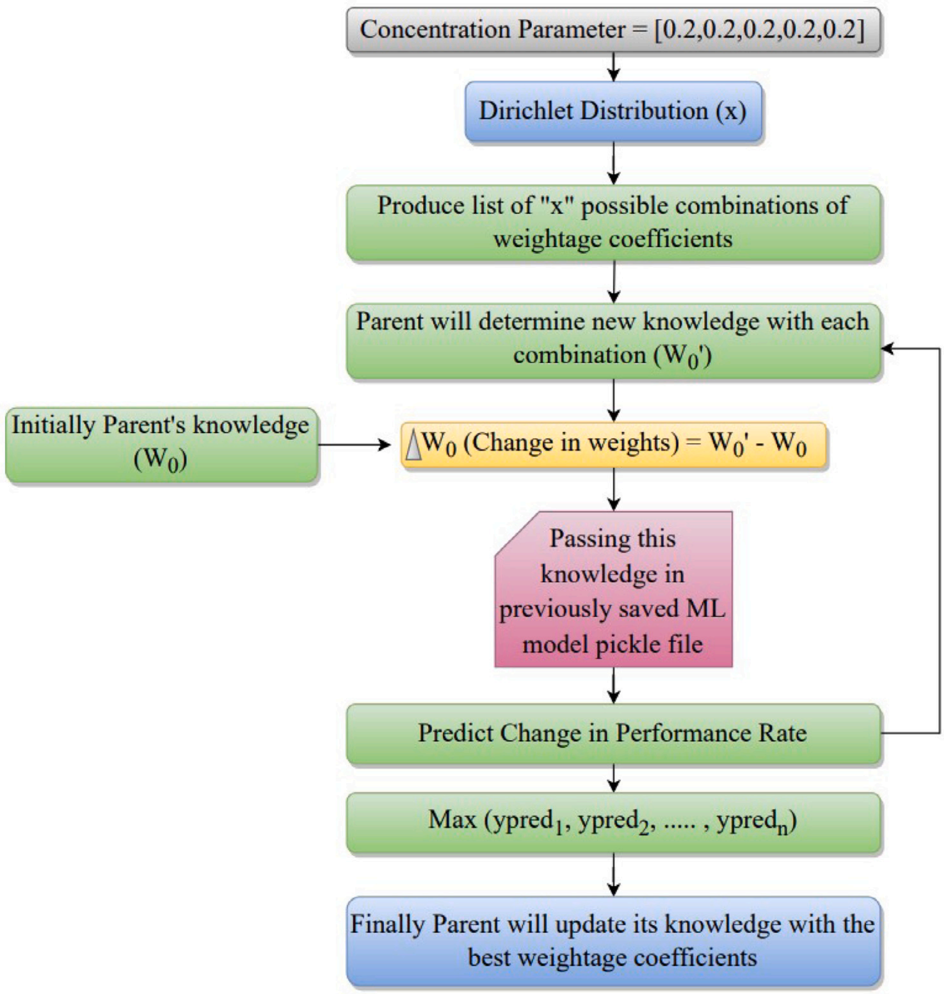

Fig. 2. Workflow for finding optimized weightage coefficient for enhanced model training.

Table 3
Dataset repository formulated by using following crops in multiple batches.

Group	Plant leaf	Category
Group1	Apple, Maize, Peach	Mixed
Group2	Potato, Strawberry, Tomato	Mixed
Group3	Banana, Cherry, Grape	Fruit
Group4	Jackfruit, Okra, Tomato	Vegetable
Group5	Apple, Betel, Peach, Strawberry	Fruit
Group6	Guava, Jamun, Lemon, Mango, Pomegranate	Fruit
Group7	Soybean, Maize, Olive	Vegetable
Group8	Coffee, Tea	Cash
Group9	Potato, Pepper, Cassava	Vegetable
Group10	Cotton, Sugarcane	Cash

Subsequently, the concatenated convolutional features from both the modules are subject to processing within an attention module. This module encompasses convolutional layers, maxpool layer and Averagepool layer. The resulting features obtained from addition of concatenate layers of both modules, are passed to the conv layer. These convolved features are then extracted and concurrently fed into the AveragePool layer and MaxPool layer. The resultant features from these pooling layers are concatenated to emphasize the most relevant features. These concatenated features are multiplied with the outcomes obtained from the concatenation of two added maxpool layers in each module. The resulting features are further fed into a conv layer followed by a global maxpool layer, which flattens the multidimensional features. Subsequently, they are passed into a dense layer with softmax activation function to classify the input images into N distinct classes.

4. Simulation results

This section provides an overview of the formulated dataset. Subsequently, the findings obtained by decisive system and the proposed FDL-IWT method are shown. Further, it shows the comparisons between the proposed method and several state-of-the-art baselines, namely FedAvg, FedAdam, FedAdagrad, CWT, FedAdp, Median and Trimmed-Mean on Non-IID datasets.

4.1. Dataset formulation

Dataset for Decisive System

There are total of 85 classes containing diseased and healthy leaves of 25 distinct plants are used for this study. Diseased and healthy leaves of different plants are collected from Plant Village (Samir, 2018), Banana (Bhuiyan et al., 2023), Yellow Vein (Gadde, 2021), Betel (Kumar S, 2021), Leaf images (Siddharth et al., 2019), Soybean (Mignoni, 2021), Olive (Uğuz and Uysal, 2021), Coffee (Parraga-Alava et al., 2019), Tea (Kimutai and Förster, 2022), Cassava (Iranga, 2021), Cotton (Sarosh, 2021) and Sugarcane (Ali, 2022) repositories. The dataset formulation process utilizes ten distinct groups of crops, which are categorized under Cash crops, Fruit crops, Vegetable crops and Mixed crops (fruit and vegetable), as shown in Table 3.

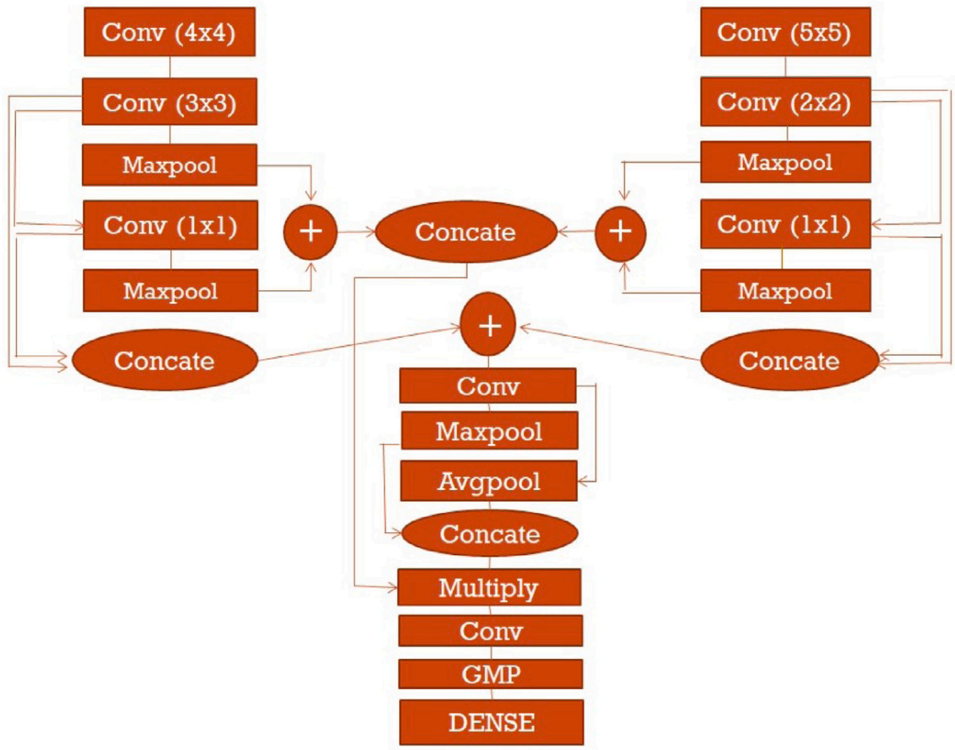

Fig. 3. Proposed Parallel MultiScale Attention CNN (PMACNN) architecture for parent and child.

Table 4
Categorization of Dataset to simulate Federated configuration.

Categorization	Maharashtra	West Bengal	Uttar Pradesh	Gujarat
High Prevalence	Tomato Bacterialspot (Pawar, 2014)	Strawberry Leafspot (Sun et al., 2023), Tomato Target Leafspot (Kamei et al., 2019)	Tomato SpiderMites (Singh and Raghuraman, 2011)	Potato Earlyblight (Sharma et al., 2020) Tomato Leafcurl
Limited Prevalence	Strawberry Leafspot, Tomato Target Leafspot, Tomato SpiderMites (Bhatt, 2018), Potato Earlyblight, Tomato Leafcurl	Tomato Bacterialspot, Tomato SpiderMites (Bhatt, 2018), Potato Earlyblight, Tomato Leafcurl	Tomato Bacterialspot, Strawberry Leafspot, Tomato Target Leafspot, Potato Earlyblight, Tomato Leafcurl	Tomato Bacterialspot, Strawberry Leafspot, Tomato Target Leafspot, Tomato SpiderMites (Bhatt, 2018)
Moderate Prevalence	Potato Lateblight, Potato Healthy, Strawberry Healthy, Tomato Healthy			

Further, each individual group creates ten additional subgroups of the same category, each consisting of unique samples. This resulting in a total of 100 groups ordered in a randomized way, in order to obtain unique learning patterns.

Dataset for Proposed FDL with IWT

This paper presents the formulation of Non-IID (Non-Independent and Identically Distributed) datasets, simulating the configuration of federated deep learning in agricultural scenario. The occurrence and prevalence of plant diseases exhibit significant regional variations. These variations are categorized under three distinct forms namely, (i) Highly Prevalence (ii) Limited Prevalence and (iii) Moderate Prevalence. Algorithm 1 outlines the procedure for formulating this Non-IID dataset.

Consequently, a particular plant disease that is highly prevalent in one location due to favorable factors such as environmental conditions and farming practices may be less prevalent in other areas. To replicate this comprehensive scenario, 50,238 samples of diseased and healthy leaves from 10 distinct classes of Potato, Strawberry and Tomato (PST) plants are collected from PlantVillage (Samir, 2018). These samples were gathered across four distinct regions: Maharashtra, West Bengal, Uttar Pradesh and Gujarat. Table 4 shows the disease categorization in respective regions.

As part of dataset simulation, another set of 40,278 samples was formulated. This dataset includes 13 classes, encompassing diseased and healthy leaves of Banana (Bhuiyan et al., 2023), Coconut (Shravana, 2023) and Grape (Samir, 2018) (BCG) crops.

4.2. Results

Simulation Setup: The experimentation took place on a Windows 10 computer equipped with an Intel(R) Xeon(R) CPU E5-1660 v4 processor, 32 GB of RAM, and a 4 GB NVIDIA Quadro M2000 GPU. The implementation utilizes the Python programming language along with standard deep learning libraries, including Keras, TensorFlow and Numpy, to train and evaluate the deep learning models.

In this study, a comprehensive performance evaluation is conducted to assess the performance and effectiveness of the proposed FDL-IWT. The proposed method leverages the knowledge performance analysis by using decisive system. To establish this decision making process, various ML regressor models including Adaboost (Solomatine and Shrestha, 2004), Decision Tree (Loh, 2011), Linear Regression (James et al., 2023), RandomForest (Breiman, 2001) and Support Vector (Zhang and O'Donnell, 2020) are assessed. Therefore, in order to evaluate the performance of these models more comprehensively, mean squared

Algorithm 1 Non-IID Dataset Formulation

Require: d : list of sample count per class; ch : number of child models; $comm$: list of common classes; $pr1$, $pr2$,.. : prevalence class list for child1, child2,..; $phase$: number of phases
Ensure: Distributed list of lists $c1$, $c2$, .. : Non-IID samples in each child
Set Parent percentage (G) randomly within (25,35); $ph \leftarrow 100 - G$
List $ratio$ with normalized percentage of phases randomly selected within (10,25), such that $\sum_{j=1}^{phase} Pj = ph$; Set list $Li \leftarrow ratio \times d$
for a in Li **do**
  $c1ph, c2ph, .. \leftarrow$ PhaseDist($a, pr1, .., comm$); Append $c1ph$ to $c1$,..
end for
procedure PhaseDist(a, $pr1$, .., $comm$)
  for $i11$ in a **do**
    for k in $pr1$ **do**
      if $i = k$ **then** ▷ Initially $i \leftarrow 0$
        $cr2 \leftarrow$ random.uniform(0.0, 0.1)
        $crp2 \leftarrow$ Multiply($i11, cr2$)▷ Repeat similarly, for other child
        $crp1 \leftarrow$ Substract($i11, (crp2 + ..)$)
        Append $crp1$ to $c1ph$, $crp2$ to $c2ph$, and so on..
      end if
    end for ▷ Repeat similar blocks for $pr2$, pr,.. and comm..
    $i \leftarrow i + 1$;
  end for
  return $c1ph$, $c2ph$, ..
end procedure

Table 5
Results obtained by different machine learning regression models.

Machine learning models	MSE	RMSE	MAE	R^2
Adaboost Regressor	0.03	0.17	0.14	−0.03
Decision Tree Regressor	0.06	0.25	0.21	−1.25
Linear Regressor	**0.009**	**0.09**	**0.07**	**0.67**
Random Forest Regressor	0.03	0.18	0.14	−0.15
Support Vector Regressor	0.03	0.17	0.13	−0.03

error (MSE), root mean square error (RMSE), mean absolute error (MAE) and coefficient of determination (R^2) (Zhang et al., 2023b) were used. This evaluation is performed over CSV dataset having knowledge updates and change in performance rate. Table 5 presents the obtained MSE, RMSE, MAE and R^2 outcomes.

It is obtained that the Linear Regression model attains the lowest mse, rmse, mae value of 0.009, 0.09 and 0.07 respectively. Whereas Adaboost, Decision Tree, RandomForest, and Support Vector regression models attain more error rate with mse value greater than 0.025, rmse value greater than 0.15 and mae greater than 0.10. Therefore, the Linear Regression model is employed as the Decisive System to determine optimized weightage coefficients for child models based on their performance in the proposed FDL-IWT method.

In addition to this, the effectiveness of proposed FDL-IWT is evaluated on two distinct formulated federated dataset under Non-IID settings (i) Potato, Strawberry, and Tomato (PST) (ii) Banana, Coconut and Grapes (BCG) crops. For this evaluation, two different performance metrics are employed: (i) Testing Accuracy using parent (global) knowledge and (ii) Validation Accuracy using child (local) knowledge. Each child model after aggregation get evaluated with their local dataset or testing dataset. This local dataset further split into training dataset and validation dataset with 75:25 splitting ratio. The child model gets trained on the training set and the trained model is employed to predict the responses for the observations in the validation set.

Table 6 shows the respective findings obtained over PST dataset under four different phases. Phase1 is termed as initialization phase, where each child model gets initialized by the parent's knowledge. The testing accuracy of each individual child model obtained over their own local dataset ranges from 94 to 98%. This performance is either improving or consistent with the subsequent phases (Phase2, Phase3, and Phase4). The performance range improves by up to 2% at last phase, depicting more accurate detection of plant leaf diseases.

In addition to this, the evaluation of child models over validation dataset provides an unbiased estimate of model's performance on unseen dataset. This value initially ranges from 97 to 99% and improves by up to maximum of 99% at last phase.

The learning curves of child models are shown in Fig. 4, to observe the generalizing ability of the models. Fig. 4(a), 4(b), and 4(c) show high fluctuation rate up to 25th epochs, 35th epochs, 33rd epochs respectively, after which they starts converging. Whereas Fig. 4(d) shows that child model 4 performs consistently better than the other child models (child1, child2, and child3). There is a significant spike observed in learning curves of each child model, this may be because of the use of Adam optimizer which accidentally pushes the weights away from the minimum. However, it starts converging and stabilizing after a few epochs. This shows that the models are neither underfitted nor overfitted.

4.3. Comparison with standard baselines

This work presents a comparative analysis of various federated learning methods including FedAvg (Li et al., 2019), FedAdam (Reddi et al., 2020), FedAdp (Wu and Wang, 2021), FedAdagrad (Reddi et al., 2020), CWT (Chang et al., 2018), Median (Yin et al., 2018), Trimmed-mean (Yin et al., 2018) and proposed Federated Deep Learning (FDL) with Intelligently Weight Transferring (IWT) over PST dataset and BCG dataset. The comparative findings are illustrated in Tables 7 and 8 respectively.

The findings shown in Table 7, indicate that FedAvg exhibits inconsistent performance during the overall phases, likely due to the heterogeneous distribution of datasets. This shows the lack of knowledge retention in local models, leading to misclassification. Conversely, FedAdam and FedAdp demonstrate improved stability. Both the methods, use Adam optimizer which facilitates faster convergence, achieving 96% and 91% test accuracy, respectively. Additionally, FedAdagrad maintains consistent performance throughout the phases, albeit obtaining the second lowest test accuracy of 90%. This consistency shows that the Adagrad optimizer's adaptive learning rate adjustment based on prior gradients effectively handles varying data distributions. CWT achieves the second-highest performance rate of 97%, showcasing its efficacy.

Furthermore, Median and Trimmed-mean are two statistical methods that follows co-ordinate wise aggregation rule are also compared. The finding shows that the Median method first deteriorates the performance in Phase2 and then it starts improving gradually with each increasing phase. However, the result shows that Median method slower the convergence rate and manages to acquire 76% testing accuracy. While, the Trimmed-mean method performs comparatively similar to the proposed FDL-IWT method and achieves 97.5% recognition rate. Therefore, the proposed FDL-IWT method outperforms FedAvg (93%), Fed-Adam (96%), FedAdagrad (90%), FedAdp (91%), CWT (97%), Median (76%) and comparable to Trimmed Mean (97.5%) by attaining an impressive testing accuracy of 97.5%. As a result, the proposed method enhances knowledge exchange, model synchronization, and convergence speed, ultimately leading to more accurate and robust local models.

The outcomes obtained by state-of-the-art methods over BCG dataset are illustrated in Table 8. The results depict that, FedAvg method deteriorating after phase2, this may be due to the presence of Non-IID dataset which leads to conflicting and noisy updates, resulting in a degradation of model's performance. While FedAdam is showing inconsistent performance across overall phases, this may be because of the Adam optimizer which makes it sensitive to changes in data distribution.

Table 6
Simulation results obtained by Proposed FDL with IWT over Potato, Strawberry, and Tomato.

Phases	Accuracy	Child1	Child2	Child3	Child4
Phase1	Test accuracy using Global Knowledge	98%	94%	96%	97%
	Validation accuracy using Local Knowledge	99%	99%	97%	99%
Phase2	Test accuracy using Global Knowledge	98%	96%	97%	97%
	Validation accuracy using Local Knowledge	98%	98%	98%	99%
Phase3	Test accuracy using Global Knowledge	98%	98%	96%	98%
	Validation accuracy using Local Knowledge	99%	99%	99%	99%
Phase4	Test accuracy using Global Knowledge	98%	99%	96%	99%
	Validation accuracy using Local Knowledge	99%	99%	99%	99%

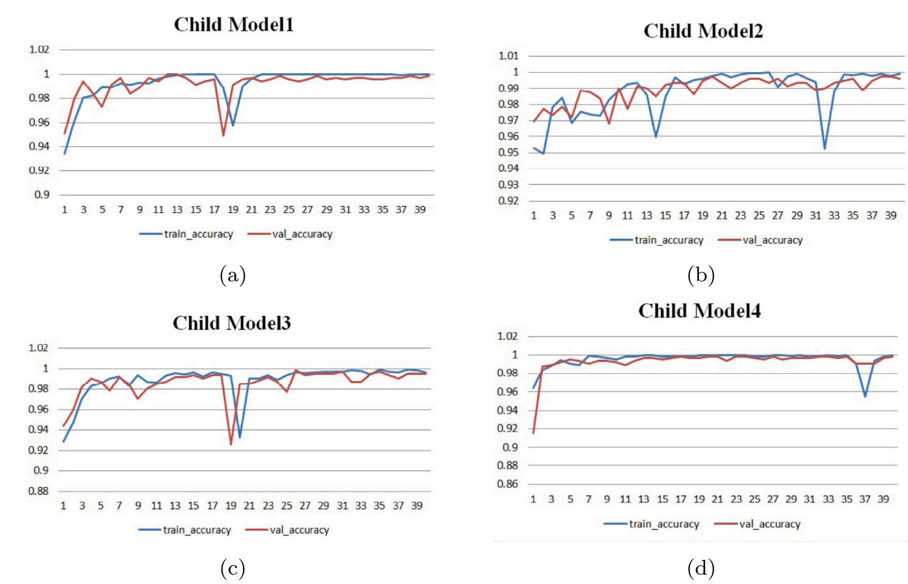

Fig. 4. Learning curves (Accuracy vs Epochs) with respect to number of iterations for (a) Child Model1 (b) Child Model2 (c) Child Model3 (d) Child Model4.

Table 7
Comparison with standard baselines over Potato, Strawberry and Tomato Crops using PMACNN.

Methods	Phase1	Phase2	Phase3	Phase4
FedAvg	(92.5 ± 4.5)%	(97.5 ± 0.5)%	(94 ± 4)%	(93 ± 3)%
FedAdam	(94 ± 4)%	(96 ± 2)%	(98 ± 1)%	(96 ± 2)%
FedAdagrad	(82 ± 2)%	(87.5 ± 4.5)%	(88.5 ± 3.5)%	(90 ± 2)%
CWT	(93 ± 5)%	(96.5 ± 0.5)%	(95.5 ± 3.5)%	(97 ± 2)%
FedAdp	(91 ± 6)%	(95.5 ± 2.5)%	(96.5 ± 1.5)%	(91 ± 7)%
Median	(96.5 ± 1.5)%	(52.5 ± 41.5)%	(69.5 ± 23.5)%	(76 ± 20)%
Trimmed mean	(96.5 ± 1.5)%	(96.5 ± 1.5)%	(97 ± 1)%	(97.5 ± 1.5)%
Proposed FDL with IWT	(96 ± 2)%	(97 ± 1)%	(97 ± 1)%	(97.5 ± 1.5)%

Table 8
Comparison with state-of-the-art over Banana, Coconut and Grape crops using PMACNN.

Methods	Phase1	Phase2	Phase3	Phase4
FedAvg	(92.5 ± 4.5)%	(97.5 ± 1.5)%	(94 ± 4)%	(91 ± 6)%
FedAdam	(92.5 ± 4.5)%	(95 ± 3)%	(91.5 ± 5.5)%	(94.5 ± 3.5)%
FedAdagrad	(51.5 ± 14.5)%	(67.5 ± 8.5)%	(67.15 ± 12.5)%	(67 ± 14)%
CWT	(88 ± 5)%	(95 ± 2)%	(91 ± 8)%	(91.5 ± 6.5)%
FedAdp	(89.5 ± 7.5)%	(90.5 ± 7.5)%	(93.5 ± 4.5)%	(84.5 ± 11.5)%
Median	(93 ± 3)%	(83.5 ± 11.5)%	(87 ± 8)%	(88.5 ± 5.5)%
Trimmed mean	(93 ± 3)%	(59.5 ± 29.5)%	(76.5 ± 16.5)%	(92.5 ± 4.5)%
Proposed FDL with IWT	(93.5 ± 1.5)%	(94 ± 4)%	(95 ± 3)%	(95.5 ± 3.5)%

Inspite of its inconsistency, FedAdam manages to attain second highest accuracy rate with 94.5%. FedAdagrad shows consistent performance but unable to perform efficiently resulting in the lowest test accuracy rate of 67%. Apart from optimizer based methods, FedAdp is a similarity based method which performs well up to phase3 but it degrades in last phase and attains 84.5% accuracy. This may be because of accumulated noise and discrepancies from different child models. These accumulated noise becomes more prominent in later rounds, adversely affecting convergence. CWT is performing efficiently and attaining 91.5% testing accuracy. Additionally, statistical methods namely Median and Trimmed-mean acquires testing accuracy of 88.5% and 92.5% respectively. After Phase2, both the methods are gradually converging with each increasing communication phase. However, the proposed FDL with IWT outperforms the state-of-the-art methods, FedAvg

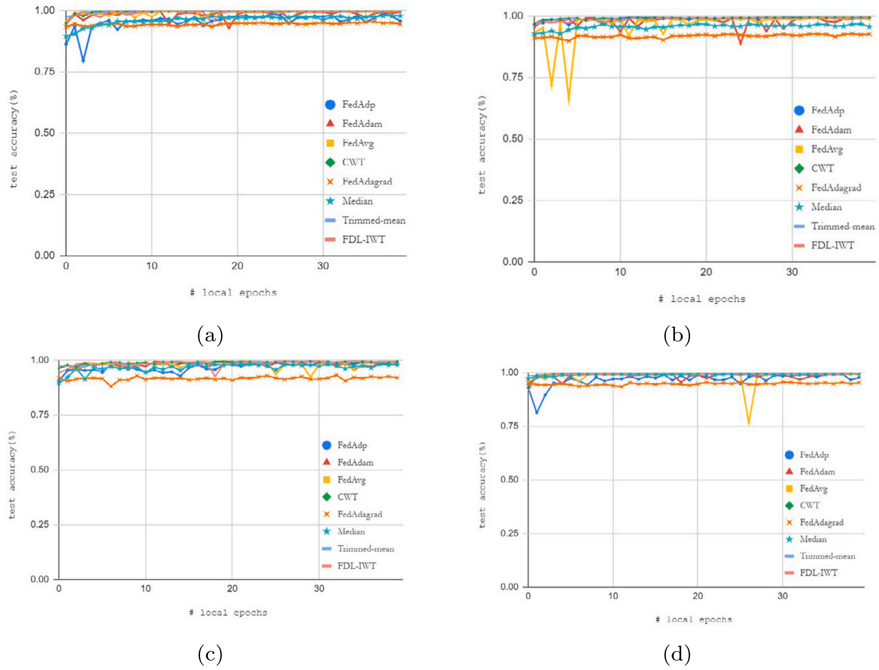

Fig. 5. Test accuracy of child models with different number of local epochs over Potato, Strawberry and Tomato (a) child1 (b) child2 (c) child3 (d) child4.

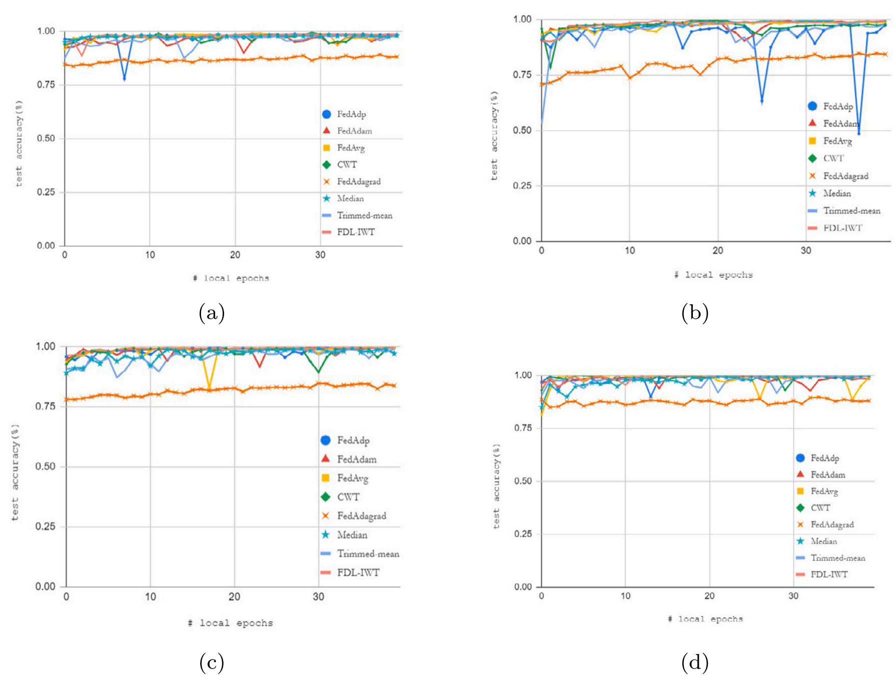

Fig. 6. Test accuracy of child models with different number of local epochs over Banana, Coconut and Grape (a) Child1 (b) Child2 (c) Child3 (d) Child4.

(91%), FedAdam (94.5%), FedAdagrad (67%), CWT (91.5%), FedAdp (85.5%), Median(88.5%), and Trimmed-mean (92.5%) by obtaining testing accuracy rate of 95.5%.

Figs. 5 and 6 illustrate the test accuracy over PST dataset and BCG dataset respectively, obtained by four different child models of proposed FDL-IWT and seven different state-of-the-art methods FedAvg, FedAdam, FedAdagrad, CWT, FedAdp, Median and Trimmed-mean. This shows that the proposed FDL-IWT method obtains a balanced performance rate across distinct child models.

Fig. 7 depicts the top-1 testing accuracy obtained by each child model on both PST and BCG datasets in the presence of Non-IID settings. The most significant difference observed among all childs is 3% in PST and 7% in BCG dataset.

In addition to this, the proposed Parallel MultiScale Attention CNN Architecture along with FDL-IWT is assessed and compared with two existing CNN architectures; VGG (Wang et al., 2023) and CNN (Xu et al., 2023). The respective outcomes are shown in Table 9. The simulated experiments are performed over diseased and healthy leaves of two different datasets: (i) Banana, Coconut and Grapes (BCG) and (ii) Potato, Strawberry and Tomato (PST), simulating the configuration of Non-IID federated learning. It is obtained that VGG, a pre-trained architecture attains 97% testing accuracy over BCG. Although, this architecture obtains higher performance rate, but looses consistency

Table 9
Comparison of PMACNN with existing CNN architectures over PST and BCG datasets.

Dataset	Architectures	Phase1	Phase2	Phase3	Phase4
BCG	VGG	(93.5 ± 3.5)%	(97 ± 2)%	(95.5 ± 2.5)%	(97 ± 1)%
	CNN	(84.5 ± 4.5)%	(86.5 ± 6.5)%	(88.5 ± 6.5)%	(92 ± 4)%
	PMACNN	(93.5 ± 1.5)%	(94 ± 4)%	(95 ± 3)%	(95.5 ± 3.5)%
PST	VGG	(93.5 ± 1.5)%	(95.5 ± 0.5)%	(95.5 ± 1.5)%	(97 ± 1)%
	CNN	(88.5 ± 4.5)%	(91.5 ± 2.5)%	(92.5 ± 1.5)%	(88 ± 6)%
	PMACNN	(96 ± 2)%	(97 ± 1)%	(97 ± 1)%	(97.5 ± 1.5)%

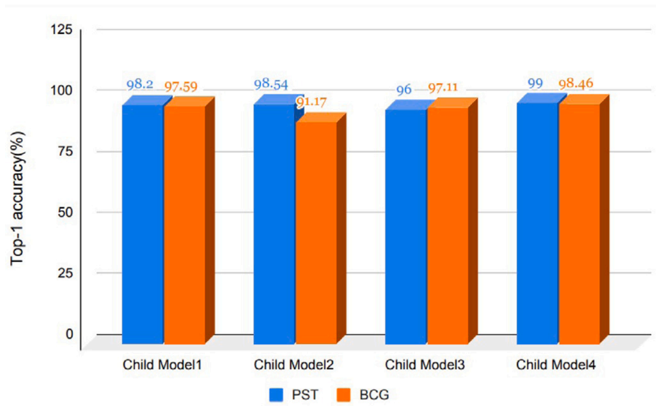

Fig. 7. Top-1 accuracy of each child model on Potato, Strawberry, Tomato (PST) and Banana, Coconut, Grape (BCG) dataset under Non-IID configurations.

across phases. Whereas, CNN and PMACNN obtains a consistent increase in their performances throughout the phases and attains 92% and 95.5% respectively. Whereas, the results obtained by VGG, CNN and PMACNN over PST dataset, show that the performance of all three architectures consistently improving throughout the phases, except CNN which shows a performance drop at last phase and unable to cross 90% mark.

Additionally, PMACNN uses only 0.13 million parameters while VGG uses 14 million parameters and CNN uses 11 million parameters. This shows the lightweight nature of the proposed PMACNN architecture, which can contribute to a reduction in computational costs. Incorporating the proposed PMACNN architecture with FDL-IWT has the potential to further enhance its overall efficiency. Consideration of this lightweight architecture aligns with the goal of achieving improved computational performance in federated learning setups.

5. Discussion

Federated learning is a decentralized machine learning paradigm that enables training of multiple local models without sharing their data. Each individual local (child) model independently gets trained by its own data while periodically exchanging model updates with a global (parent) model. However, local models may have non-independent and identically distributed (non-IID) data, leading to significant variations in their data distributions. Dealing with such data heterogeneity poses challenges for aggregating model updates and achieving convergence to high-quality models. Thus, this work presents an Intelligent Weight Transferring (IWT) method in FDL. It combines an analysis of knowledge performance using a decisive system with the determination of optimized weightage coefficients. This method enables a more adaptable and effective method for handling Non-IID datasets.

In this work, the proposed FDL-IWT method is evaluated over diseased and healthy leaves of Potato, Strawberry and Tomato (PST) simulating their occurrences in Maharashtra, West Bengal, Uttar Pradesh and Gujarat. In order to assess the robustness and generalization of proposed method, it is also evaluated over Banana, Coconut and Grape (BCG) datasets simulated under Non-IID settings. The findings illustrate that Proposed FDL-IWT method attains testing accuracy rate of 97.5% over Potato, Strawberry, Tomato, and outperforms several state-of-the-art methods including FedAvg (93%), FedAdam (96%), FedAdagrad (90%), CWT (97%), FedAdp (91%), Median (76%) and comparable to Trimmed-mean (97.5%) over PST dataset.

The proposed FDL-IWT method attains the highest performance rate by using proposed Parallel MultiScale Attention CNN (PMACNN) architecture as parent and child entities. Therefore, to evaluate the effectiveness of PMACNN with FDL-IWT, it is compared with two different existing CNN architectures VGG (Wang et al., 2023) and CNN (Xu et al., 2023) coupling with FDL-IWT. The findings show that the proposed FDL-IWT method obtains consistent and better performance rate of 95.5% over BCG and 97.5% over PST by using PMACNN as compared to VGG and CNN. Owing to the lightweight nature of PMACNN, it uses 0.13 million parameters whereas VGG employs 14 million parameters and CNN employs 11 million parameters. This depicts that PMACNN may be helpful in mitigating the computational complexity of FDL-IWT.

6. Conclusion

This paper presents Federated Deep Learning with Intelligent Weight Transferring (FDL-IWT) method to adaptively handle the Non-IID plant diseased dataset. It introduces knowledge performance analysis by using decisive system and finding optimized weightage coefficients adaptively to each child models. Additionally, this work also presents parallel multiscale attention CNN (PMACNN) architecture for parent and child models. This proposed PMACNN architecture being lightweight employs 0.13 million parameters. It is obtained from the findings that the proposed FDL-IWT method performs consistent or better with every phase of federated framework attaining 97.5% testing accuracy over Potato, Strawberry and Tomato. In comparison, the proposed method outperforms state-of-the-art methods including FedAvg, FedAdam, FedAdagrad, CWT, FedAdp, Median and Trimmed-mean over diseased and healthy leaves of two Non-IID plant diseased datasets (i) Potato, Strawberry and Tomato (ii) Banana, Coconut and Grape.

In the future, this work can be expanded to different domains, where data privacy is a major concern. In addition to this, the work may be explored for scaling with more local devices and cross-domain adaptation. Furthermore, this approach may extend to "n" clients, as per availability of other regions. The limitations of this work include an asynchronous knowledge update strategy, device heterogeneity, and noisy (or blurred) images.

CRediT authorship contribution statement

Pragya Hari: Writing – review & editing, Writing – original draft, Visualization, Validation, Methodology, Conceptualization. **Maheshwari Prasad Singh:** Writing – review & editing, Validation, Supervision, Methodology, Conceptualization.

Funding

First author, Pragya Hari receives the Ph.D. scholarship from National Institute of Technology, Patna, India.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

Data will be made available on request.

References

Aggarwal, M., Khullar, V., Goyal, N., Alammari, A., Albahar, M.A., Singh, A., 2023a. Lightweight federated learning for rice leaf disease classification using non-independent and identically distributed images. Sustainability 15 (16), 12149.
Aggarwal, M., Khullar, V., Goyal, N., Gautam, R., Alblehai, F., Elghatwary, M., Singh, A., 2023b. Federated transfer learning for rice-leaf disease classification across multiclient cross-silo datasets. Agronomy 13 (10), 2483.
Ahmad, A., Luo, W., Robles-Kelly, A., 2023. Robust federated learning under statistical heterogeneity via hessian-weighted aggregation. Mach. Learn. 112 (2), 633–654.
Ali, K., 2022. Red Rot Sugarcane Disease Leaf Dataset. Kaggle, Available online at www.kaggle.com/datasets/alihussainkhan24/red-rot-sugarcane-disease-leaf-dataset.
Amin, H., Darwish, A., Hassanien, A.E., Soliman, M., 2022. End-to-end deep learning model for corn leaf disease classification. IEEE Access 10, 31103–31115.
Arun, R.A., Umamaheswari, S., 2023. Effective multi-crop disease detection using pruned complete concatenated deep learning model. Expert Syst. Appl. 213, 118905.
Bhatt, N., 2018. Tomato bug, nesidiocoris tenuis (Reuter): A zoophytophagous insect. Life 4486, 985B–9C86FC4F1C40.
Bhuiyan, M.A.B., Abdullah, H.M., Arman, S.E., Rahman, S.S., Al Mahmud, K., 2023. BananaSqueezeNet: A very fast, lightweight convolutional neural network for the diagnosis of three prominent banana leaf diseases. Smart Agric. Technol. 4, 100214.
Breiman, L., 2001. Random forests. Mach. Learn. 45, 5–32.
Chang, K., Balachandar, N., Lam, C., Yi, D., Brown, J., Beers, A., Rosen, B., Rubin, D.L., Kalpathy-Cramer, J., 2018. Distributed deep learning networks among institutions for medical imaging. J. Am. Med. Inform. Assoc. 25 (8), 945–954.
Chen, J., Zhang, D., Zeb, A., Nanehkaran, Y.A., 2021. Identification of rice plant diseases using lightweight attention networks. Expert Syst. Appl. 169, 114514.
Gadde, M., 2021. Yellow Vein Mosaic Disease. Kaggle, Available online at www.kaggle.com/datasets/manojgadde/yellow-vein-mosaic-disease.
Hou, W., Wen, H., Zhang, N., Lei, W., Lin, H., Han, Z., Liu, Q., 2023. Adaptive training and aggregation for federated learning in multi-tier computing networks. IEEE Trans. Mob. Comput..
Huang, W., Wang, X., Xia, J., Li, Y., Zhang, L., Feng, H., Zhang, X., 2023. Flexible sensing enabled agri-food cold chain quality control: A review of mechanism analysis, emerging applications, and system integration. Trends Food Sci. Technol..
Ilhan, F., Su, G., Liu, L., 2023. ScaleFL: Resource-adaptive federated learning with heterogeneous clients. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. pp. 24532–24541.
Iranga, H.A., 2021. Leaf Disease Dataset (Combination). Kaggle, Available online at www.kaggle.com/datasets/asheniranga/leaf-disease-dataset-combination.
James, G., Witten, D., Hastie, T., Tibshirani, R., Taylor, J., 2023. Linear regression. In: An Introduction to Statistical Learning: With Applications in Python. Springer, pp. 69–134.
Kamei, A., Dutta, S., Sarker, K., Das, S., Datta, G., Goldar, S., 2019. Target leaf spot of tomato incited by Corynespora cassiicola , an emerging disease in tomato production under gangetic alluvial region of West Bengal, India. Arch. Phytopathol. Plant Prot. 51, 1–10.
Kaya, Y., Gürsoy, E., 2023. A novel multi-head CNN design to identify plant diseases using the fusion of RGB images. Ecol. Inform. 75, 101998.
Khorchef, M., Ramou, N., Boutiche, Y., Guessoum, A., Chetih, N., 2023. Image forest fire segmentation using Dirichlet process mixture model. In: 2023 International Conference on Advances in Electronics, Control and Communication Systems. ICAECCS, IEEE, pp. 1–5.
Kimutai, G., Förster, A., 2022. tea sickness dataset. Mendeley Data, Available online at data.mendeley.com/datasets/j32xdt2ff5/2.
Kumar S, C., 2021. Betel Leaf. Kaggle, Available online at www.kaggle.com/datasets/chethandekohli/betel-leaf.
Li, X., Huang, K., Yang, W., Wang, S., Zhang, Z., 2019. On the convergence of fedavg on non-iid data. arXiv preprint arXiv:1907.02189.
Loh, W.Y., 2011. Classification and regression trees. Wiley Interdiscip. Rev.: Data Min. Knowl. Discov. 1 (1), 14–23.
McMahan, B., Moore, E., Ramage, D., Hampson, S., y Arcas, B.A., 2017. Communication-efficient learning of deep networks from decentralized data. In: Artificial Intelligence and Statistics. PMLR, pp. 1273–1282.
Mignoni, M.E., 2021. Images of soybean leaves. Mendeley Data, Available online at data.mendeley.com/datasets/bycbh73438/1.
Parraga-Alava, J., Cusme, K., Loor, A., Santander, E., 2019. RoCoLe: A robusta coffee leaf images dataset for evaluation of machine learning based methods in plant diseases recognition. Data Brief 25, 104414.
Pawar, B., 2014. Report of bacterial diseases of tomato from Marathwada region of Maharashtra, India. Res. J. Recent Sci. ISSN 2277, 2502.
Reddi, S., Charles, Z., Zaheer, M., Garrett, Z., Rush, K., Konečnỳ, J., Kumar, S., McMahan, H.B., 2020. Adaptive federated optimization. arXiv preprint arXiv:2003.00295.
Samir, B., 2018. New plant diseases dataset. Available online at www.kaggle.com/datasets/vipoooool/new-plant-diseases-dataset.
Sarosh, K., 2021. cotton Leaf Disease Dataset. Kaggle, Available online at www.kaggle.com/datasets/seroshkarim/cotton-leaf-disease-dataset.
Sharma, R., Patel, J., Patel, R., 2020. Management of early blight of potato (Solanum tuberosum L.) caused by Alternaria solani [(Ellis & Martin) Jones & Grout] through fungicides and its impact on yield. Int. J. Curr. Microbiol. Appl. Sci. 9, 1683–1693.
Sharma, V., Tripathi, A.K., Mittal, H., 2023. DLMC-net: Deeper lightweight multi-class classification model for plant leaf disease detection. Ecol. Inform. 75, 102025.
Shravana, T., 2023. Coconut Leaf Dataset for Pest Identification. Kaggle, Available online at www.kaggle.com/datasets/shravanatirtha/coconut-leaf-dataset-for-pest-identification.
Siddharth, S.C., Ajay, K., Uday, P.S., Sanjeev, J., 2019. A database of leaf images: practice towards plant conservation with plant pathology. Mendeley Data.
Singh, J., Raghuraman, M., 2011. Emerging scenario of important mite pests in north India. Zoosymposia 6, 170–179.
Solomatine, D.P., Shrestha, D.L., 2004. AdaBoost. RT: a boosting algorithm for regression problems. In: 2004 IEEE International Joint Conference on Neural Networks (IEEE Cat. No. 04CH37541), vol. 2, IEEE, pp. 163–1168.
Sun, Y., Jayawardena, R., Sun, J., Wang, Y., 2023. Pestalotioid species associated with medicinal plants in southwest China and Thailand. Microbiol. Spectrum 11 (1), e03987–22.
Szegedy, C., Vanhoucke, V., Ioffe, S., Shlens, J., Wojna, Z., 2016. Rethinking the inception architecture for computer vision. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. pp. 2818–2826.
Thakur, P.S., Sheorey, T., Ojha, A., 2023. VGG-ICNN: A lightweight CNN model for crop disease identification. Multimedia Tools Appl. 82 (1), 497–520.
Uğuz, S., Uysal, N., 2021. Classification of olive leaf diseases using deep convolutional neural networks. Neural Comput. Appl. 33 (9), 4133–4149.
Wang, K., He, Q., Chen, F., Chen, C., Huang, F., Jin, H., Yang, Y., 2023. FlexiFed: Personalized federated learning for edge clients with heterogeneous model architectures. In: Proceedings of the ACM Web Conference 2023. pp. 2979–2990.
Wu, H., Wang, P., 2021. Fast-convergent federated learning with adaptive weighting. IEEE Trans. Cogn. Commun. Netw. 7 (4), 1078–1088.
Xu, J., Tong, X., Huang, S.L., 2023. Personalized federated learning with feature alignment and classifier collaboration. arXiv preprint arXiv:2306.11867.
Yang, L., Yu, X., Zhang, S., Long, H., Zhang, H., Xu, S., Liao, Y., 2023. GoogLeNet based on residual network and attention mechanism identification of rice leaf diseases. Comput. Electron. Agric. 204, 107543.
Yin, D., Chen, Y., Kannan, R., Bartlett, P., 2018. Byzantine-robust distributed learning: Towards optimal statistical rates. In: International Conference on Machine Learning. Pmlr, pp. 5650–5659.
Zhang, J., Guo, S., Qu, Z., Zeng, D., Zhan, Y., Liu, Q., Akerkar, R., 2021a. Adaptive federated learning on non-iid data with resource constraint. IEEE Trans. Comput. 71 (7), 1655–1667.
Zhang, J., Hua, Y., Wang, H., Song, T., Xue, Z., Ma, R., Guan, H., 2023a. Fedala: Adaptive local aggregation for personalized federated learning. In: Proceedings of the AAAI Conference on Artificial Intelligence, vol. 37, pp. 11237–11244.
Zhang, H., Luo, G., Li, J., Wang, F.Y., 2021b. C2FDA: Coarse-to-fine domain adaptation for traffic object detection. IEEE Trans. Intell. Transp. Syst. 23 (8), 12633–12647.
Zhang, F., O'Donnell, L.J., 2020. Support vector regression. In: Machine Learning. Elsevier, pp. 123–140.
Zhang, Q., Zhao, X., Han, Y., Yang, F., Pan, S., Liu, Z., Wang, K., Zhao, C., 2023b. Maize yield prediction using federated random forest. Comput. Electron. Agric. 210, 107930.
Zheng, H., Fan, X., Bo, W., Yang, X., Tjahjadi, T., Jin, S., 2023. A multiscale point-supervised network for counting maize tassels in the wild. Plant Phenomics 5, 0100.