International Journal of Mathematical, Engineering and Management Sciences
Vol. 11, No. 4, 1830-1862, 2026
https://doi.org/10.33889/IJMEMS.2026.11.4.074

FedLeaf: A Collaborative Federated Learning based Model for Rice Leaf Disease Detection

Amandeep Kaur
Chitkara University Institute of Engineering and Technology,
Chitkara University, Punjab, India.
E-mail: kaur.amandeep@chitkara.edu.in, 15amankdeep@gmail.com

Kalpna Guleria
Chitkara University Institute of Engineering and Technology,
Chitkara University, Punjab, India.
Corresponding author: kalpna@chitkara.edu.in, guleria.kalpna@gmail.com

(Received on January 8, 2026; Revised on May 7, 2026 & June 7, 2026 & June 11, 2026; Accepted on June 14, 2026)

Abstract
The rice crop is one of the major food crops in India and around the world, which globally necessitates food security. However, the productivity of rice is usually decreased due to various rice leaf diseases. Nowadays, the identification of such diseases has become increasingly feasible through an automated process, which has been made possible due to various improvements and advancements in deep learning technologies. Traditional deep learning methods are heavily reliant on centralized data collection, thus causing significant issues regarding data privacy and ownership for the agricultural stakeholders. Centralized deep learning models require a large memory space and high computation cost for datasets to be stored at a centralized location. Federated learning solves these problems by allowing collaborative model training in a decentralized way, thus no data needs to be shared with a centralized location. This research proposes FedLeaf, a collaborative federated deep learning model for the accurate identification of rice leaf diseases. The FedLeaf design adopts a distributed learning method in which the global model is forwarded to client nodes, and then clients apply local training using their respective datasets. In the proposed research, FedLeaf utilises two deep learning architectures, MobileNetV2 and EfficientNet, in a federated environment, which are named as federated learning based MobileNetV2 and federated learning based EfficientNet. The hyperparameter tuning has been performed to get the best results, at a batch size of 32, learning rate of 0.001, Adam optimizer, seven communication rounds and five epochs per communication round. The performance evaluation of the proposed FedLeaf, has been performed based on its round-wise global and local performance parameters, including accuracy, precision, recall, F1-Score and AUC. In the proposed FedLeaf model, federated learning based EfficientNet has emerged as the optimal model for rice leaf disease detection, exhibiting 96.19% highest accuracy for IID distribution and 95.87% for non-IID data distribution. The results exhibit that the proposed model shows better performance and offers a collaborative and efficient model for automated identification of rice leaf disease.

Keywords- Federated learning, Rice leaf disease, EfficientNet, MobileNetV2, Precision agriculture, FedLeaf.

1. Introduction

Advancements in deep learning have significantly influenced the paradigm shift in data analytics across various industries and domains such as precision agriculture, healthcare, and autonomous systems (Fetzer et al., 1990; Henry and Winston, 2004; Ongsulee, 2017). Rice serves as a dietary staple for more than half of the global population. During rice cultivation, a wide range of pathogens, including bacteria, fungi, and viruses, frequently invade rice plants and thus, make the yield and quality of the grain drop by a large percentage. Various rice leaf diseases, namely bacterial leaf blight, brown spot, and leaf directly impact its productivity worldwide (Jakhar and Kaur, 2020; Erickson et al., 2021). Further, the productivity and quality of rice crops are closely linked to food security and the national economies of Asian countries (Gnanamanickam, 2009; Laha et al., 2017; Fukagawa and Ziska, 2019).

In modern agriculture, timely and accurate detection of leaf diseases remains one of the critical challenges. Traditional manual inspection of leaf disease is quite time-consuming and susceptible to human error. In recent years, deep learning models have shown good accuracy for automatic rice leaf disease detection. However, these models generally depend on large, centralized datasets, which can be difficult to obtain due to data privacy and ownership concerns (Pothen and Pai, 2020; Aggarwal et al., 2024).

On the contrary, federated learning architectures in comparison to deep learning are one of the new techniques that have the potential of further improvements to these methods by facilitating collaborative model training over distributed data sources and also providing data privacy (Aggarwal et al., 2024). This privacy-preserving method provides scalable and efficient solutions for disease detection in agriculture. Therefore, FL based rice leaf disease detection and classification models have become a powerful tool to handle the heterogeneous datasets of rice leaf disease images using collaborative learning (Aggarwal et al., 2023).

In precision agriculture, for automatic rice leaf disease detection the collaborative federated learning is more suitable since it resembles the distributed computing architecture rather than centralized computing. In modern agriculture, data is distributed across multiple farms, locations and sensors which makes centralized learning quite challenging for data sharing due to bandwidth, computation cost and ownership concerns. Whereas a collaborative federated learning based approach enables multiple stakeholders to jointly train robust disease detection models without sharing raw data, thereby supporting better, generalized, privacy-preserving, and real-time decision-making in precision agriculture environments.

In federated learning, the data is distributed across different locations and is often non-Independent and Identically Distributed (Non-IID) in nature. In federated learning (FL) models, it is absolutely necessary to understand different data distribution types. If the data is IID, then all clients will have datasets that are similarly distributed and have the same proportion of disease classes, which makes model training and aggregation less complex. On the other hand, non-IID data causes significant differences between clients in aspects like the number of classes and local contexts. These differences, called statistical heterogeneity can disrupt the process of model convergence and decrease the overall performance (Lu et al., 2024).

Creating models for rice leaf disease detection using federated learning (FL) is complicated by the data variability and class imbalance in non-IID scenarios. In order to overcome these problems and to build efficient models, different aggregation strategies, adaptive weighting mechanisms have been proposed. In the existing studies, authors have shown that with the higher degree of data heterogeneity among the clients, if the federated methods used are properly designed, FL could still attain a high level of classification accuracy (Hsieh et al., 2020). Moreover, during local training, the use of data augmentation and balanced sampling methods helps to neutralize the problem of class imbalance so that the minority disease categories get enough attention in the learning process.

Following are the contributions of this work:

- This research proposes FedLeaf, a collaborative federated learning based model for rice leaf disease detection using two deep learning architectures MobileNetV2, and EfficientNet named as Federated MobileNetV2, Federated EfficientNet.
- In the proposed work, Independent and identically distributed (IID) and non-independent and identically distributed (non-IID) distributions have been considered, and data augmentation has been applied. The dataset used in this work has been taken from the open source repository Kaggle (Saputra, n.d.), wherein 2099 images are taken from this dataset, which are further augmented by applying different augmentation functions and increased to 6297 images.

- The FedLeaf architecture follows a distributed learning approach, where a global model is shared with client nodes for local training on their respective datasets. The federated framework is implemented and tested using base architectures: MobileNetV2 and EfficientNet. Key hyperparameters were adjusted for optimal performance, including a batch size of 32, a learning rate of 0.001, the Adam optimizer, and seven communication rounds and 5 epochs per communication round.
- The federated learning-based EfficientNet model shows better performance by achieving the highest global accuracy, which proves to be an efficient model in a collaborative environment. In all the communication rounds, federated learning based EfficientNet was able to maintain very good global accuracy values, and finally, it exhibits 96.19% highest accuracy for IID distribution and 95.87% for non-IID data distribution.
- The research provides a federated learning based collaborative learning model for automatic rice leaf disease detection, in distributed scenario for early disease detection and multiclass classification contributing to the field of precision agriculture.

The work introduces FedLeaf, a federated learning based framework for rice leaf disease detection. It allows clients to train the model collaboratively at the local locations without sharing the raw data with a centralized server. FedLeaf employs a decentralized architecture that collectively builds a global model while preserving data confidentiality. The proposed framework is efficient and lightweight which is suitable for deployment on resource-constrained edge devices. Further, client training has been performed using MobileNetV2 and EfficientNet in the federated environment. Experiments were conducted on two clients involving multiple communication rounds to assess the model's performance. The results indicate that the federated learning approach is capable of achieving high classification accuracy for rice leaf disease in a distributed environment without sharing the data to the centralized server. Federated learning-based EfficientNet model consistently delivers better performance, achieving the highest global accuracy of 96.19% at round 7 for IID data and 95.87% for non-IID data, which proves to be an efficient model in a collaborative environment. In comparison, federated learning based MobileNetV2 shows an accuracy of 0.7672 under the same scenario. EfficientNet proves to be an effective model within the federated framework in the two-client configuration. The paper is organized in five sections wherein the first section covers introduction part of the paper and second section covers the literature review and existing methodologies for rice leaf disease detection model. Material and proposed research methodology is covered in the third section of this paper, and fourth section describes the results and discussion of the proposed FedLeaf federated learning based model. Finally, the conclusion of this proposed FedLeaf framework is presented in the fifth section.

2. Literature Review on Federated Learning Frameworks and Methodologies for Rice Leaf Disease Detection

Federated learning (FL) framework typically involves a client-server architecture. Initially server shares the global model with clients. In such a configuration, several client devices, like mobile or edge systems, hold local rice leaf image datasets and independently carry out initial model training using their data. Periodically, these clients send the learned model parameters or updates to a central server. The server then combines the received parameters through the usage of federated averaging (FedAvg) or other similar aggregation methods in order to create an improved global model. This updated model is thus returned to the clients for the subsequent training round, enabling collective model refinement without exchanging raw data (Aggarwal et al., 2024).

The federated architecture preserves the privacy and data integrity of edge devices which have limited processing and communication capabilities. Privacy-centric approaches like secure aggregation are also utilized to hinder data leakage or model inversion attacks during the parameter transmission. In the past, authors have tested federated learning frameworks and have achieved considerable performance showing good trade-offs between accuracy, computational efficiency, and privacy preservation (Aggarwal et al., 2023).

In Federated Transfer Learning (FTL), the feature space and sample space differ significantly, with minimum and nothing in common, and it is achieved by utilising the transfer learning paradigm by applying various CNN-based pretrained models in a distributed federated scenario. Authors (Fahim-Ul-Islam et al., 2024) proposed a federated learning model to classify plant diseases using transfer learning model along with attention mechanism and vision transformer in a distributed scenario. Besides, FL models can generalize better across different clients, these are more flexible towards data heterogeneity without requiring centralized data sharing, maintaining privacy principles, which make them feasible for real-world plant leaf disease monitoring systems implementation (Shao et al., 2024).

In the existing works, Spotted Hyena Archimedes Optimizer (SHAO) features various mechanisms for adaptive learning rate adjustment which are very suitable for situations with noisy gradients in federated setups. The use of these optimizers leads to the improvement of different metrics which, include precision, recall, and F-measure and also contributes to the faster convergence process along with the reduction of the loss values (Elaraby et al., 2022). The authors, concluded that the usage of advanced optimization algorithms, together with efficient neural architectures with federated framework, can achieve good performance.

Centralized learning can face scalability challenges, as the central server needs high computational power and storage capacity to handle large datasets and it do not comply with privacy regulations like GDPR due to the centralized data storage. However, several devices or clients work together to cooperatively train a common model using federated learning, a decentralized method that keeps client data local (Mohammed, 2025; Soni et al., 2025). Clients execute local model using their respective datasets and exchange only model updates with a centralized server, which merges these updates to enhance the global model, as opposed to providing raw data to a central server. This method enhances data privacy and security, as sensitive data is not transferred to other locations, which helps to overcome the danger of data breaches, and it complies with privacy regulations. Federated learning distributes the computational load across many clients, enabling the system to handle diverse datasets without overwhelming a central server.

Figure 1 shows a difference between centralized and federated learning approach. Federated learning can be categorized into three main types based on the distribution and organization of data among clients. Horizontal federated learning (HFL) is known as a sample-based federated learning, it is relevant when datasets across various clients share the same feature space but differ in samples. The second one is vertical federated learning (VFL) is known as feature-based Federated Learning. Vertical federated learning is applicable when datasets across different clients share the same samples but differ in feature space. This means that each client has information about the same data points but with different features. For example, a bank and an insurance company have data about the same set of customers. The bank has financial data, while the insurance company has health-related data about these customers. Federated transfer learning (FTL) federated transfer learning is suitable when both the feature spaces and samples are different across clients, but there is a small overlap in the sample space that can be used for model transfer. This type leverages transfer learning techniques to handle different feature spaces. For example, a mobile application and a wearable device have completely different data, but there is a small subset of users who use both the mobile application and the wearable device. This subset can be used to transfer knowledge between the two datasets.

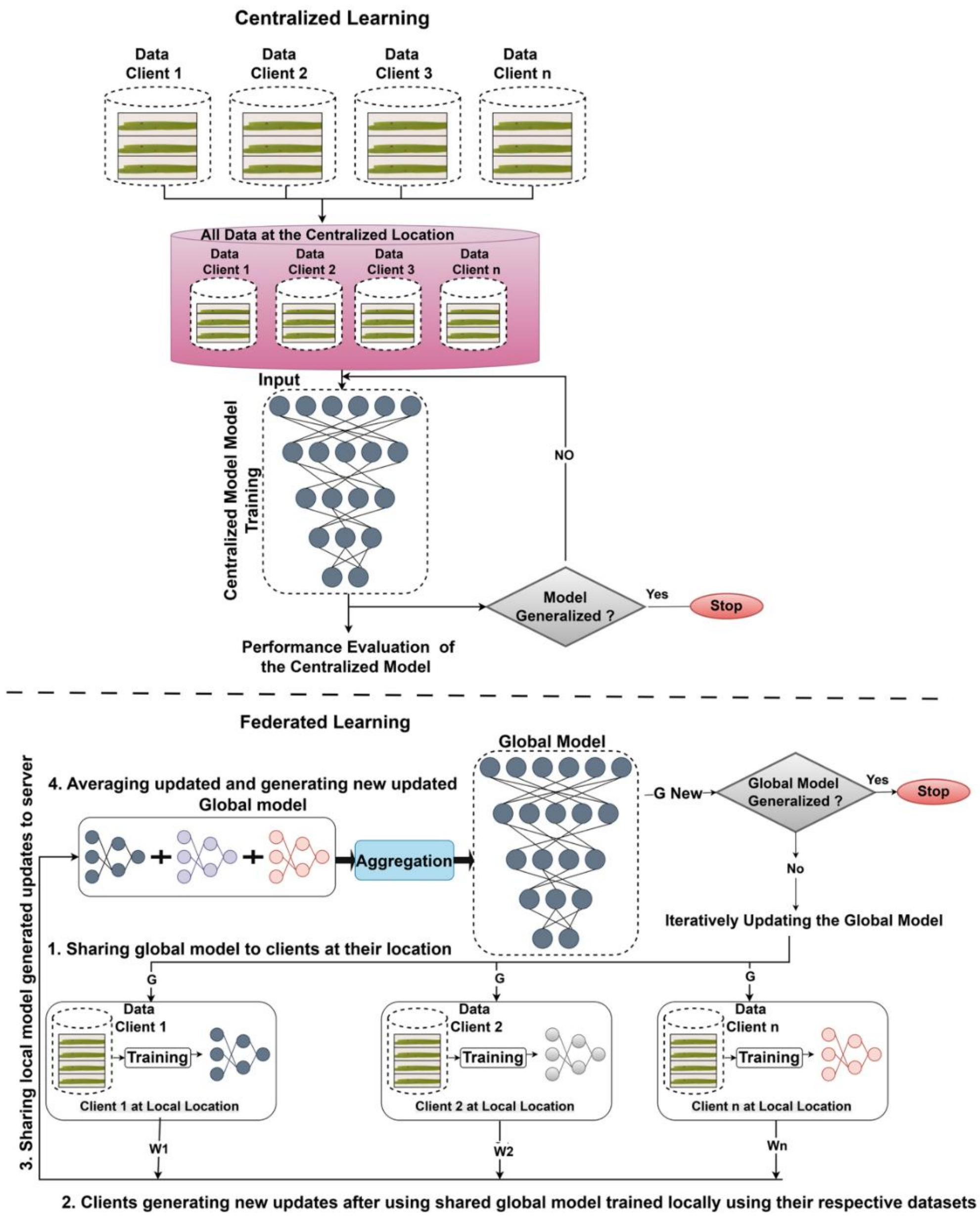

Figure 1. Difference between centralized learning and federated learning.

Figure 2 shows a decentralized federated learning model for disease prediction. It shows in first step the server shares the initial global model with the clients and clients updating the global model using private data and generating new updates. The third step is to send the updates to the server and thereafter server performs averaging on the updates and generates the new global model. This iterative process allows for collaborative model training while ensuring that sensitive data is never shared by clients, thereby preserving privacy. This proposed work introduces FedLeaf, a federated learning model designed specifically for the early and automatic detection of rice leaf diseases. In this work, two different DL architectures evaluated which are known as MobileNet and EfficientNet as base models within the federated framework to identify the effective approach.

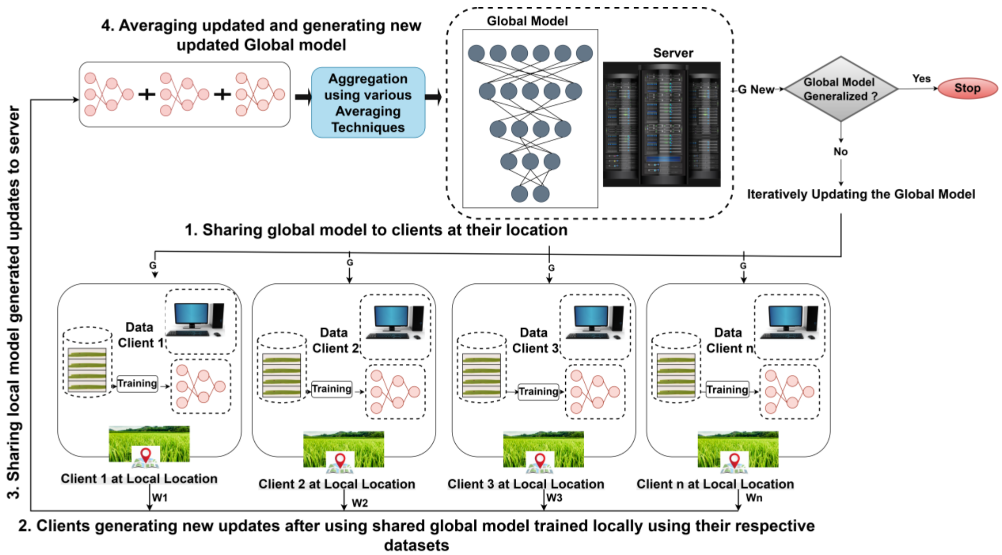

Figure 2. Architecture of federated learning.

Figure 3 shows the standard steps that are typically present in the federated learning process. First step is initializing the global model. Then next step after the client's selection is the distribution of the model in which the server shares the current global model with the chosen clients. Then, each of these clients carries out local training which is a training of the local model with their own local dataset. Every client, therefore, generates its local model update that indicates the changed model parameters relative to the global model they received. These updates, which are generally the changed weights, are then sent back to the server during the update transmission phase. The server performs aggregation where it aggregates the updates received from all the clients that participated. This can be achieved by different methods such as simple averaging, weighted averaging, or the Federated Averaging (FedAvg) approach. The combined result is then utilized to update the global model; hence, it keeps a record of the learning from all the clients. Steps 2 to 8 are repeated in a loop until the global model converges or the predetermined number of rounds is reached. In the last phase, the refined model undergoes evaluation through a separate test dataset. The performance and the ability to generalize the model are the major aspects to be measured in federated learning.

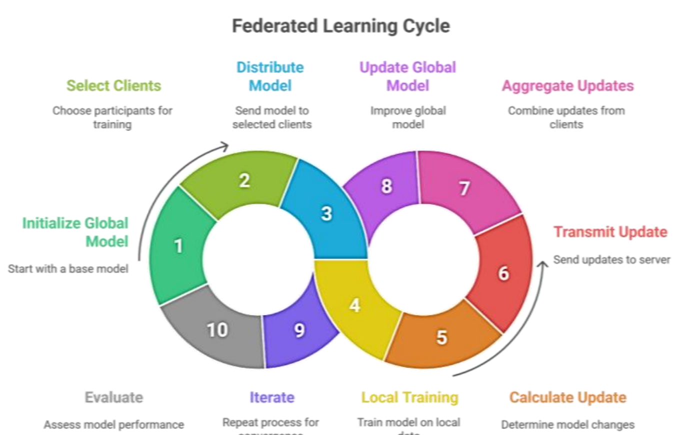

Figure 3. Steps of federated learning process.

There are different techniques like machine learning, deep learning and transfer learning to generate automated system for the rice leaf disease classification and several researches has been done on it. But it has been observed that several researchers used very small dataset for building these models due to the less availability of the datasets. Centralized models face challenges such as high storage space requirements, communication costs and issues with unbalanced and insufficient data. Federated learning offers a solution to these problems. Additionally, the lack of open-access datasets is a significant issue due to data privacy concerns. Therefore, federated learning provides a solution to this problem. Data is not transferred from the actual location as distributed model will be established or trained at the data-points and only updates will be forwarded to the centralized server. Hence, collaborative federated learning is one of the best techniques for the rice leaf disease classification (Chitta et al., 2024; Fu et al., 2024). AI techniques, particularly machine learning, offer solutions for disease detection; however, the distributed nature of computing resources necessitates strategies like FL for training models.

A detailed literature review of various federated learning based plant leaf disease detection models is presented in the **Table 1** below, which summarizes key aspects of the existing studies.

Table 1. Detailed literature review on federated deep learning based for plant leaf disease detection models.

Reference	Used technology	Number of rounds	Number of diseases	Dataset size	Focused area	Summary
Aggarwal et al. (2023)	EfficientNetB3, Dense Neural Network	10	4 (Rice Leaf Diseases)	5932 images	Lightweight federated deep learning for rice leaf disease classification	In this framework FL outperformed conventional pre-trained models. Achieved 99% accuracy for IID data, and 95% accuracy for non-IID data.
Aggarwal et al. (2024)	CNN, 8 transfer learning models (MobileNet V2, EfficientNetB3)	--	4 (Rice Leaf Diseases)	5932 images	Federated transfer learning for rice-leaf disease classification across multiclient cross-silo datasets	The federated learning framework and achieved accuracy of 99.1% on IID data and 99.8% on Non-IID data.
Antico et al. (2024)	5 CNN models	50	Maize Leaf Diseases	3852 images	Evaluation of Federated Learning for maize leaf disease prediction	Federated Learning can potentially enhance data privacy in heterogeneous domains.
Deng et al. (2022)	Improved Faster R-CNN with ResNet-101, soft-NMS	100	8 (Apple Diseases and Pests)	15,522 images (expanded from 445)	Multiple diseases and pests detection in orchards	The Faster R-CNN with federated learning achieved accuracy and a 59% improvement in training speed.
Hari et al. (2024)	Hierarchical Convolutional Neural Network (H-CNN)	4	Plant Leaf Diseases	46672	An improved federated deep learning for plant leaf disease identification	The proposed Federated Deep Learning with a lightweight H-CNN addresses the limitations of Centralized Learning (CL) for plant leaf disease identification.
Idoje et al. (2023)	Federated Averaging with SGD and Adam optimizers	100 epochs	3 (Chickpea, Rice, Maize)	Smart farm dataset	Crop classification in a smart farm decentralized network	Decentralized models achieve faster convergence and higher accuracy than centralized network models. The Adam optimizer performed better than SGD.
Kabala et al. (2023)	CNNs (ResNet50, VGG-16, etc.), Vision Transformers (ViT)	10, 30, 50, 100	Multiple (from 4 plant types)	PlantVillage dataset (over 50,000 images)	Image-based crop disease detection with federated learning	Model performance is influenced by the number of clients, communication rounds, local iterations, and data quality. ResNet50 performed better in federated learning scenario.

Table 1 continued…

Behera et al. (2025)	Fast R-CNN, Mask R-CNN, RetinaNet	--	Wheat Diseases	--	Federated learning for collaborative crop disease monitoring in wheat production	The models demonstrated effectiveness in disease detection and classification, contributing to global food security and sustainable agriculture.
Nilsson et al. (2018)	FedAvg, FSVRG, CO-OP	1000 (FedAvg), 50 (FSVRG), Every 10th update (CO-OP)	10 (Handwritten Digits)	MNIST dataset	Performance evaluation of federated learning algorithms	FedAvg aggregation achieved the highest accuracy among the federated algorithms.
Tong et al. (2022)	Improved Faster R-CNN with ResNet-101 and soft-NMS	100	8 (Orchard Pests and Diseases)	15,522 images (expanded from 445)	Detection of multiple pests and diseases in orchard	The improved Faster R-CNN with federated learning achieved 89.34% performance and the model training speed was increased by 59%.
Zhao et al. (2018)	CNN with data-sharing strategy	500	10 (Image Classes)	MNIST, CIFAR-10, Speech commands datasets	Federated learning with non-IID data	A data-sharing strategy can significantly improve the accuracy of federated learning on non-IID data. Accuracy can be increased by 30% on CIFAR-10 with only 5% globally shared data.
Tripathy et al. (2024)	Optimization (Spotted Hyena +Archimedes)	--	3 (Rice leaf classes)	120 rice leaf disease images	Federated learning with IID data	The authors presented federated learning framework with SHAO_LeNet. T. Accuracy value has not been mentioned. This work shows precision- 0.913, recall- 0.922 and F measure- 0.917 in this experiment, however the dataset is considerably small.
Leema and Balakrishnan (2023)	Vision Transformer	100	3 (Rice leaf disease classes)	120 rice leaf disease images	Federated learning with non-IID data	The authors worked on federated learning based rice leaf disease detection model using vision transformers and achieved accuracy of 99%.
Chorney et al. (2025)	Federated learning with ADSNN	650	8(7 rice leaf disease class and 1 healthy)	10553	Federated Learning with non-IID data	Authors worked on federated learning with FedAvg, enhanced by (q-FedAvg), pre-training strategies (LF-FedAvg, PT-FedAvg). This work achieved accuracy of 83.14% using adaptive deep spiking neural network with bayesian optimization.

This section presents related works on various existing FL based models for disease detection in agricultural. **Table 1** provides a detailed summary of these models, highlighting their architecture, performance, and research gaps or future scope. It has been observed through the literature review that existing studies have explored various FL approaches for plant leaf disease detection in agriculture whereas less works have been done in federated learning architectures for rice leaf disease detection. Further, in federated learning, several limitations persist in terms of handling of non-IID data, comprehensive evaluation, limited dataset size, limited work on rice leaf diseases using federated learning etc (Zhao et al., 2018; Idoje et al., 2023; Kabala et al., 2023; Antico et al., 2024; Behera et al., 2025). In the existing works, utilization of small datasets and a limited number of disease classes, may restrict model generalizability (Tripathy et al., 2024). Furthermore, a significant portion of the existing research focuses on general datasets, including maize, orchard crops, and benchmark datasets, rather than specifically addressing rice leaf disease detection (Zhao et al., 2018; Deng et al., 2022; Idoje et al., 2023; Kabala et al., 2023; Antico et al., 2024). In the existing work, authors (Tripathy et al., 2024) presented a federated learning based model that integrates hybrid meta-heuristic optimizer for rice leaf disease detection. Two optimizer spotted hyena and Archimedes optimizer are combined with LeNet which is a CNN architecture. In this study, the authors used a dataset of 120 images, which consists of three classes of rice leaf diseases known as brown spot, bacterial blight and leaf smut. In this work, the authors work on IID data and achieved a recall of 0.922, however the dataset considered in this research was very small. In Leema and Balakrishnan (2024), authors presented a federated learning based framework using vision transformers CoAtNet and SwinTransformer for the rice leaf disease classification. In this work authors used dataset which consists three rice leaf disease classes named as bacterial blight, leaf smut and brown spot. The authors achieved an accuracy of 99% however, the dataset size was limited to 120 images and only three disease classes were considered. Authors (Chorney et al., 2025) proposed a multi-site rice leaf disease detection model in which heterogeneous non-IID data distribution and 8 classes (7 rice leaf disease classes and 1 healthy) were considered. Federated learning model was implemented using MobileNetV2 and ADSNN-BO (Adaptive Deep Spiking Neural Network with Bayesian Optimization) however the achieved accuracy of the model was limited to 83.14% The proposed work provides a federated learning based collaborative model for automatic rice leaf disease detection, wherein 6 classes (five disease classes and a healthy rice leaf class) have been considered on a large dataset with IID and non-IID data distribution for early disease detection and multiclass classification contributing to the field of precision agriculture.

3. Materials and Methodology for FedLeaf: A Collaborative Federated Learning based Model for Rice Leaf Disease Detection

This section describes materials and methodology used in the proposed FedLeaf, a federated learning based model for rice leaf disease detection.

3.1 Dataset Description

The dataset used in this work is collected from the open source repository Kaggle (Saputra, n.d.), wherein 2099 images are taken from this dataset, which are further augmented by applying different augmentation functions and increased to 6297 images. The dataset has five disease classes (narrow brown spot, leaf scald, leaf blast, brown spot, and bacterial leaf blight and one healthy leaf class (Saputra, n.d.). All images of the dataset are resized to 224x224 from 300x300, since EfficientNet and MobileNetV2 require an image size of 224x224. This helps to reduce the computation cost of the model and maintains the sufficient feature details for classification. To increase dataset diversity and improve model generalization, extensive data augmentation was applied using rotation (image rotated randomly between -20° to +20°), zoom (zoom in or out upto ±25%), shift (image shift left/right up to 15% of width), and horizontal-flip (randomly flips the image left ↔ right) transformations. These transformations are randomly combined, so that every generated image is unique. For every original image, two new augmented images were generated using random transformation with the above-mentioned functions for the 3x expansion of the dataset, resulting in a total of 6297 images after augmentation. Each rice leaf disease class has been expanded from approximately 350 original images to around 1020-1050 images, providing a more balanced and enriched dataset. Further, the augmented dataset is divided into a training and testing set of 80/20, which consists of 5037 training images and 1260 testing images. This augmented dataset was later divided into IID (Independent and Identically Distributed) splits for two clients, ensuring uniform class distribution of the images to form the federated learning-based model. Each client consists of 2518-2519 images with an identical per-class count of around 420 samples per class. In case of non-IID data distribution client 1 has 3393 images and client 2 has 1644 images which are further divided into six classes for each client. **Figures 4** to **9** presents a sample images of rice leaf diseases known as leaf smut, brown spot, leaf blast, bacterial blight and healthy leaf.

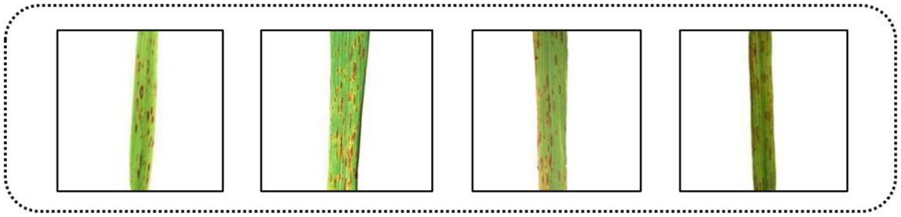

Figure 4. Sample image of narrow brown spot.

Figure 5. Sample image of leaf blast.

Figure 6. Sample image of bacterial blight.

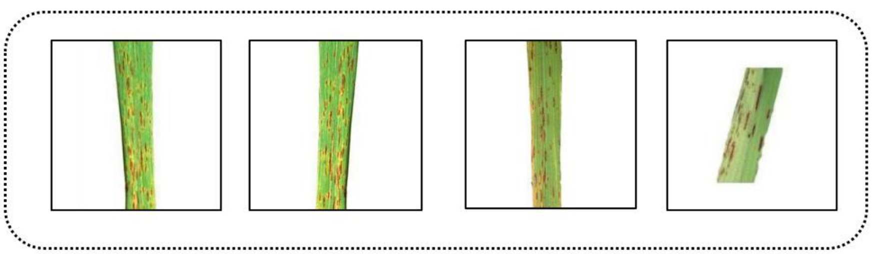

Figure 7. Sample image of brown spot.

Figure 8. Sample image of leaf scald.

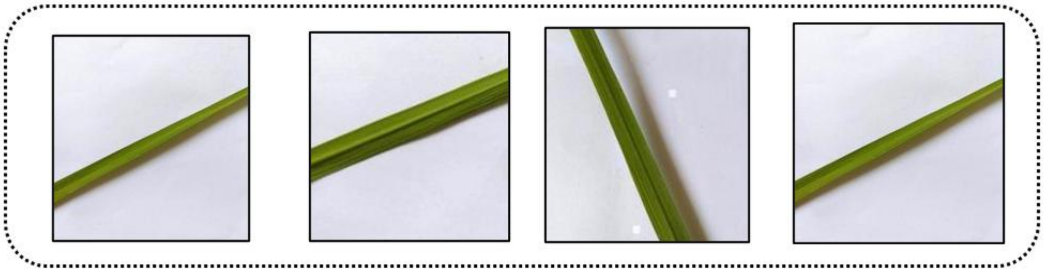

Figure 9. Sample image of healthy leaf.

3.2 Proposed Methodology for FedLeaf Model for Rice Leaf Disease Detection

This section presents the description of the proposed methodology for FedLeaf, a federated learning based model for rice leaf disease detection. This work integrating a deep learning architecture with federated learning to create a system that can detect rice leaf diseases automatically from the images.

Figure 10 shows the general framework of how rice leaf disease classification can be carried out using federated learning. The entire federated learning (FL) process for rice leaf disease classification begins with collecting leaf images for disease classification. During data pre-processing various operations are applied for compatibility with the model. Pre-processing is very important as it updates the raw images into an understandable format that can be fed into the model by adjusting images to have the same size. Implementing preprocessing strategies can improve classification accuracy as well as model performance. Whereas, employing deep learning or convolutional neural networks on non-pre-processed images can result in poor classification performance. Thereafter, a pre-processed portion of the dataset is given to each client, and the data of each client is handled separately. This leads to a federated learning framework in which each client uses its model locally with its own data subset. In the collaborative decentralized approach, the models are trained locally and then aggregated on the central server. Federated learning (FL) provides a way of addressing the problem of data distribution and decentralized computing by establishing a firm foundation for the efficient training of model.

Once the local training is done, each client sends the changes made in its local model to a central server for global aggregation. The server adopts the federated averaging algorithm to combine these changes into a single global model, thus enabling the entire system to learn from the heterogeneous data distributed across all the clients and thereafter update the global model that is shared with the clients so that local training can proceed for another round. This establishes a continuous iterative process of local training, model update exchange and global aggregation. The model is, therefore, subjected to several such rounds and at last, it is tested on a different dataset to measure its accuracy as well as generalization ability. This iterative federated learning approach enhances model robustness and ensures effective performance across varying data distributions. The model makes use of the loss function, like categorical cross-entropy, to direct the training as well as the evaluation metrics such as accuracy and F1-score to measure the performance. In the end all these updates are aggregated by the central server to improve and refine the global model. This method enhances data privacy as data remains on local devices and is not exposed to potential breaches during transmission.

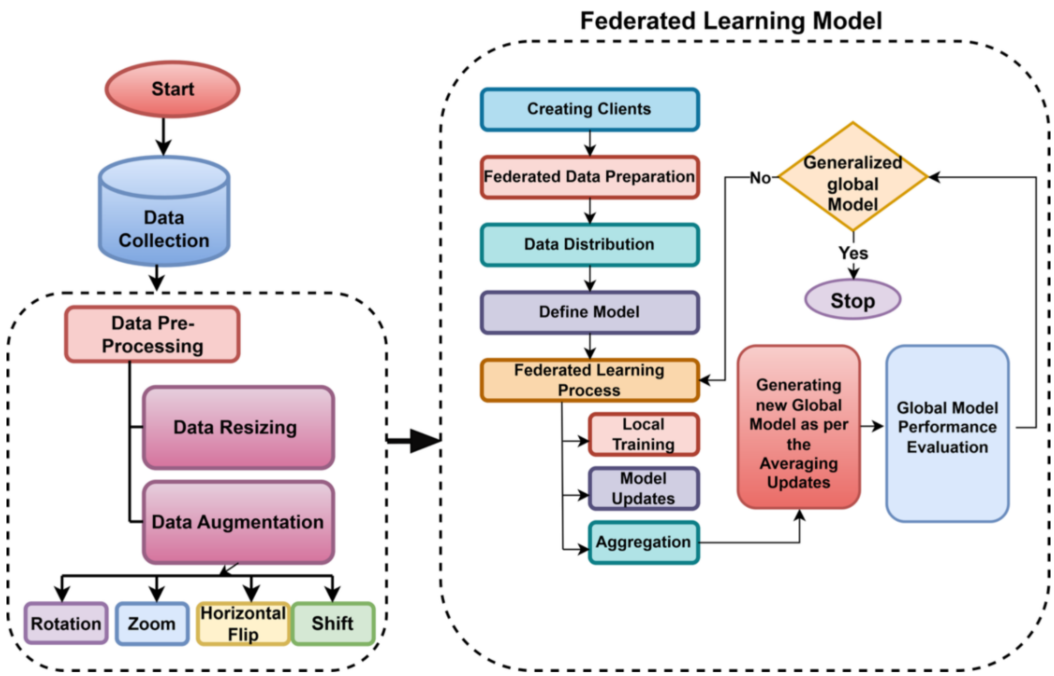

Figure 10. General methodology of federated learning based rice leaf disease detection model.

The steps mentioned below show a numerical presentation of the federated learning framework.

- k: Total number of clients.
- n_k: Number of data points on k clients.
- n: Total number of data points across all clients ($n=\sum_{k=1}^{K} n_k$) (1)
- w_k: model parameter trained locally on client k.
- w_t: global model parameters at round t.

The global model parameters will be updated as: $w_{t+1}=\sum_{k=1}^{K}\frac{n_k}{n}w_k$ (2)

where, $\frac{n_k}{n}$ is the weight proportional to the client's dataset size.

Figure 11 depicts the entire workflow of the FedLeaf a federated learning based rice leaf disease detection model, which is designed to be executed with two deep learning architectures MobileNetV2 and EfficientNet in federated environment considering two local clients in each scenario. The process flow is initiated by a central server distributing a preliminary global model to both clients. A client-side training of the model is performed by each client with its private dataset, and new weights (W1 and W2) are generated. Locally, the updates are forwarded to the server, where a federated averaging operation is performed on the weights returned from the clients to obtain the aggregated global model. The FedAvg aggregation scheme has been applied at the server side. The updated model is then shared with the clients once again, and the procedure is carried out through several communication rounds. Furthermore, the figure indicates that client-side data storage is independent of each other, implying that raw data is not shared with the server or between the local locations. In essence, the workflow is a depiction of how federated learning is used to train models collaboratively over distributed data sources while at the same time, data privacy is maintained.

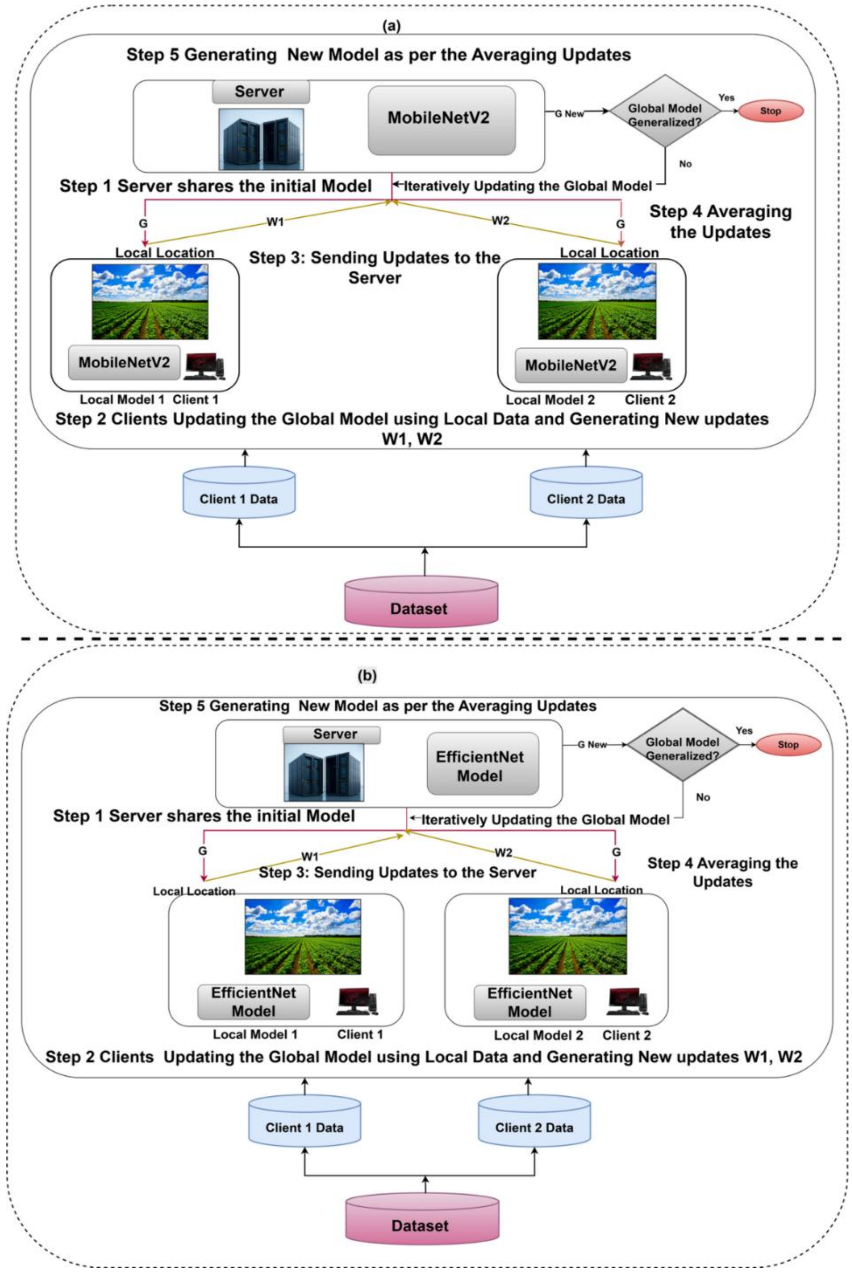

Figure 11. Proposed FedLeaf: a collaborative model for rice leaf disease detection (a) Federated learning based MobileNetV2 (b) Federated learning based EfficientNet.

3.3 FedLeaf: Federated Learning Model Training and Evaluation

This section presents the results of the proposed FedLeaf model for predicting rice leaf diseases. The key experimental parameters and their corresponding values used for implementing federated learning FedLeaf model for rice leaf disease detection are summarized in **Table 2**. These parameters define the dataset configuration, model training settings, optimization strategy, evaluation metrics, and overall experimental setup, of the proposed FedLeaf model using MobileNetV2 and EfficientNet.

Table 2. A detailed description of implementation hyperparameters and performance evaluation parameters for the FedLeaf model.

Parameter	Value	Description
Deep learning architectures for building FedLeaf	MobileNetV2, EfficientNet-B0	FedLeaf model has been constructed using two deep learning architectures namely MobileNetV2 and EfficientNet-B0 as the base architecture in federated environment.
Clients	2	Two federated clients are simulated on a single machine (Google Colab). Each client has its own training and validation dataset.
Batch size	32	Each batch contains 32 images. The model updates weights after processing each batch (mini-batch gradient descent).
Learning rate	0.001	The code uses the Adam optimizer with default learning rate 0.001.
Communication rounds	7	Federated Learning runs for 7 communication rounds. After each round, client weights are aggregated at the server using FedAvg.
Optimizer	Adam	The optimizer used is Adam, a popular adaptive learning rate optimization algorithm.
Performance metric	Global and Local performance evaluation for Accuracy, Precision, Recall, F1- Score, AUC	Metrics are computed using Precision, Recall, F1, Accuracy and AUC for local and global performance evaluation of FedLeaf model.
Epochs per communication round	5	Each client trains locally for 5 epochs per communication round.
No of classes	6	The dataset has 6 classes (5 rice leaf disease + 1 healthy).
Train: test ratio	80:20	When preparing the dataset, each client has train and test folders with 80% training and 20% validation images.

4. Results and Discussion

This section elaborates on performance evaluation of FedLeaf federated learning model for rice leaf disease detection.

4.1 Performance Evaluation of Federated Model using MobileNetV2 for Rice Leaf Disease Detection on Client 1, Client 2 from Round 1-7 on IID Data

Table 3 illustrates the results of FedLeaf Federated MobileNetV2 model for multiclass classification of rice leaf disease detection. Across the seven rounds of communications, it indicates a consistent improvement in accuracy of both client-side models and the global model. Client 1 and client 2 initially have accuracies of 0.7245 and 0.7147 in Round 1, respectively. Model learns gradually and the performance eventually reaches 0.8071 and 0.7990, respectively for client 1 and 2. The global model's performance improved from 0.7196 to 0.7942 from round 1 to round. The model exhibits a slight decrease in global accuracy from 0.7942 in round 6 to 0.7925 in round 7 which shows a drop of 0.0017. In federated learning scenarios, it happens due to the stochastic nature of local client training and global model aggregation wherein global model outcome is not always improvise in a monotonic manner in all the communication rounds. Hence, the model has achieved highest accuracy of 0.7942 in round 6.

Figure 12 is the graphical representation of the client-wise accuracy using federated MobileNetV2 model over the 7 communication rounds. In the initial rounds both clients exhibit a consistent performance in accuracy 0.7245 and 0.7147 respectively for client 1 and client 2. The highest accuracy achieved by the client 1 is 0.8071 in round 6 and client 2 achieved highest accuracy of 0.7990 in round 7.

Figure 13 represents a precision graph which shows a client-wise precision value across the 7 federated rounds. In round 1 both clients achieved 0.7334 and 0.7228 respectively. Further, improvement was observed throughout the rounds and the highest precision of 0.8071 was achieved for client 1 in round 6 and 0.8004 for client 2 in round 7. The increasing precision values indicates that the model is more effective in minimizing false positive and correctly identifying rice leaf disease classes.

Table 3. Performance evaluation of FedLeaf, federated model using MobileNetV2 for rice leaf disease detection on client 1, client 2 from round 1-7 on IID data.

Model	Client	Accuracy	Precision	Recall	F1 Score	AUC
MobileNetV2 Round 1 (5 epochs)	Client 1	0.7245	0.7334	0.7245	0.7266	0.9369
	Client 2	0.7147	0.7228	0.7147	0.7153	0.9418
	Global Metrics After Round 1	0.7196	0.7281	0.7196	0.7209	0.9394
MobileNetV2 Round 2 (10 epochs)	Client 1	0.7504	0.7542	0.7504	0.7473	0.9508
	Client 2	0.7569	0.7643	0.7569	0.7528	0.9567
	Global Metrics After Round 2	0.7536	0.7593	0.7536	0.7500	0.9537
MobileNetV2 Round 3 (15 epochs)	Client 1	0.7618	0.7694	0.7618	0.7612	0.9542
	Client 2	0.7699	0.7772	0.7699	0.7680	0.9620
	Global Metrics After Round 3	0.7658	0.7733	0.7658	0.7646	0.9581
MobileNetV2 Round 4 (20 epochs)	Client 1	0.7666	0.7704	0.7666	0.7644	0.9572
	Client 2	0.7699	0.7732	0.7699	0.7672	0.9656
	Global Metrics After Round 4	0.7682	0.7718	0.7682	0.7658	0.9614
MobileNetV2 Round 5 (25 epochs)	Client 1	0.7804	0.7806	0.7804	0.7771	0.9629
	Client 2	0.7909	0.7928	0.7909	0.7882	0.9662
	Global Metrics After Round 5	0.7804	0.7806	0.7804	0.7771	0.9629
MobileNetV2 Round 6 (30 epochs)	Client 1	0.8071	0.8071	0.8071	0.8045	0.9685
	Client 2	0.7812	0.7776	0.7812	0.7772	0.9606
	Global Metrics After Round 6	**0.7942**	**0.7924**	**0.7942**	**0.7908**	**0.9645**
MobileNetV2 Round 7 (35 epochs)	Client 1	0.7861	0.7863	0.7861	0.7842	0.9619
	Client 2	0.7990	0.8004	0.7990	0.7968	0.9692
	Global Metrics After Round 7	0.7925	0.7933	0.7925	0.7905	0.9656

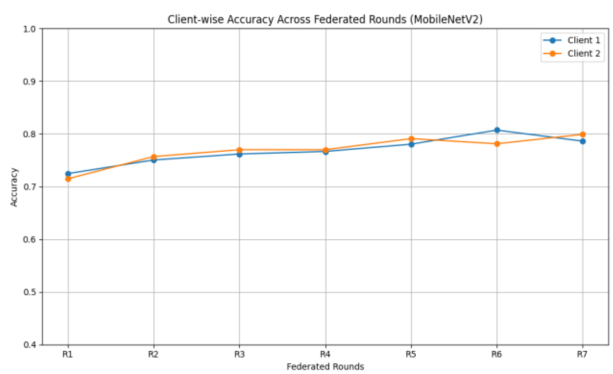

Figure 12. Client-wise accuracy of federated model using MobileNetV2 for rice leaf disease detection on round 1-7 on IID data.

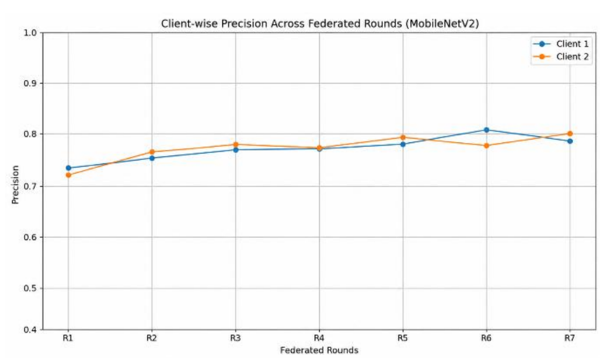

Figure 13. Client-wise precision of federated model using MobileNetV2 for rice leaf disease detection on round 1-7 on IID data.

Figure 14 shows the recall performance of client 1 and client 2 using federated MobileNetV2 across all federated rounds. In round 1 client 1 achieved recall value of 0.7245 and client 2 achieved 0.7147. Subsequent rounds enhance recall performance. The highest recall value of 0.8071 was achieved by client 1 in round 6.

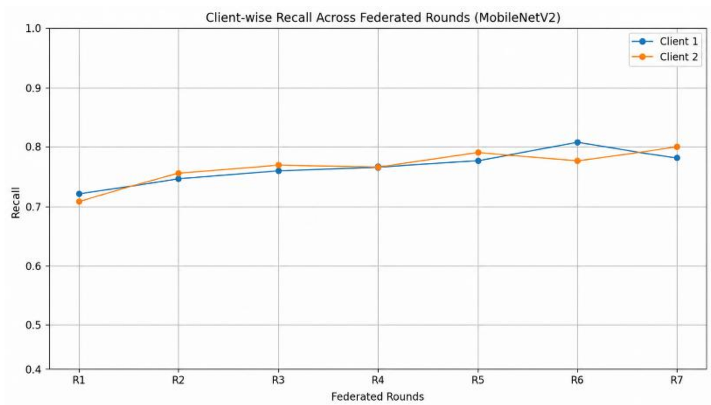

Figure 14. Client-wise recall of federated model using MobileNetV2 for rice leaf disease detection on round 1-7 on IID data.

Figure 15 is graphical representation of the F1 score of federated MobileNetV2 for rice leaf disease detection from round 1-7 on IID data. The F1-score increased from 0.7266 in round 1 to 0.8045 in round 6 for client 1. Client 1 achieved the highest individual F1-score of 0.8045 in round 6, while client 2 reached 0.7968 in the final round. These findings confirm that the federated model maintained a balanced classification performance across all rounds.

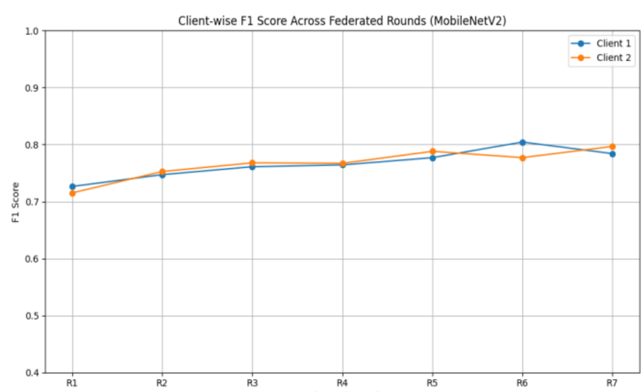

Figure 15. Client-wise F1 score of federated model using MobileNetV2 for rice leaf disease detection on round 1-7 on IID data.

Figure 16 represents a client-wise AUC of federated MobileNetV2 model for rice leaf disease detection on round 1-7 on IID data. The Area Under the Curve (AUC) values remained consistently high throughout the training process, indicating strong discriminative capability. The highest AUC for client 1 is 0.9685 in round 6 while client 2 attained the highest AUC value of 0.9692 in round 7. These high AUC values demonstrate the model's ability to distinguish between healthy and diseased rice leaf classes.

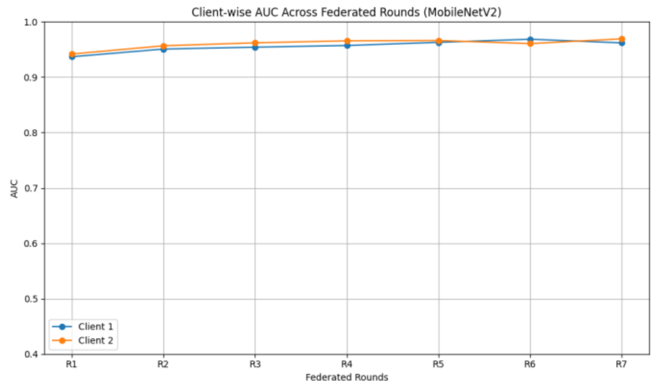

Figure 16. Client-wise AUC of federated model using MobileNetV2 for rice leaf disease detection on round 1-7 on IID data.

4.2 Performance Evaluation of Federated Model using EfficientNet for Rice Leaf Disease detection for Client 1, Client 2 on IID Data Distribution

Table 4 shows the results of the federated learning model using EfficientNet for both clients 1 and 2 and the global model through seven communication rounds. In round 1, client 1 and client 2 started with accuracies of 0.9762 and 0.9617, respectively, and both performed very well throughout the rounds, reaching the highest accuracies of 0.9964 and 0.9968 in round 6. The global model exhibits stable learning, improving from 0.9238 in Round 1 to 0.9619 in round 5, with AUC values staying above 0.99 all the time, which shows that the model is very good at distinguishing between classes. Precision, recall, and F1-score are very close to each other in different rounds, which is a strong indication that classification behavior is balanced for all clients. To sum up, EfficientNet demonstrated strong convergence in a federated setting, keeping a high value of accuracy and AUC metrics on both clients and ensuring stable global performance during all rounds.

Table 4. Performance evaluation of federated model using EfficientNet for rice leaf disease detection on client 1, client 2 from round 1-7 on IID data.

Model	Client	Accuracy	Precision	Recall	F1 Score	AUC
EfficientNet Round 1 (5 epochs)	Client 1	0.9762	0.9763	0.9762	0.9760	0.9992
	Client 2	0.9617	0.9633	0.9617	0.9613	0.9988
	Global Metrics After Round 1	0.9238	0.9237	0.9238	0.9224	0.9948
EfficientNet Round 2 (10 epochs)	Client 1	0.9887	0.9890	0.9887	0.9887	0.9999
	Client 2	0.9855	0.9855	0.9855	0.9854	0.9997
	Global Metrics After Round 2	0.9441	0.9469	0.9441	0.9438	0.9970
EfficientNet Round 3 (15 epochs)	Client 1	0.9919	0.9920	0.9919	0.9919	1.0000
	Client 2	0.9883	0.9885	0.9883	0.9882	0.9999
	Global Metrics After Round 3	0.9457	0.9478	0.9457	0.9453	0.9971
EfficientNet Round 4 (20 epochs)	Client 1	0.9919	0.9920	0.9919	0.9919	0.9999
	Client 2	0.9895	0.9897	0.9895	0.9895	0.9999
	Global Metrics After Round 4	0.9587	0.9598	0.9587	0.9583	0.9971
EfficientNet Round 5 (25 epochs)	Client 1	0.9611	0.9625	0.9611	0.9607	0.9969
	Client 2	0.9972	0.9972	0.9972	0.9972	1.0000
	Global Metrics After Round 5	**0.9619**	**0.9619**	**0.9619**	**0.9617**	**0.9979**
EfficientNet Round 6 (30 epochs)	Client 1	0.9964	0.9964	0.9964	0.9964	1.0000
	Client 2	0.9968	0.9968	0.9968	0.9968	1.0000
	Global Metrics After Round 6	0.9579	0.9579	0.9579	0.9576	0.9976
EfficientNet Round 7 (35 epochs)	Client 1	0.9611	0.9611	0.9611	0.9609	0.9976
	Client 2	0.9546	0.9548	0.9546	0.9542	0.9976
	Global Metrics After Round 7	0.9457	0.9478	0.9457	0.9453	0.9971

Figure 17 illustrates the client-wise accuracy of the federated EfficientNet model across seven communication rounds. In round 1, client 1 achieved an accuracy of 0.9762, while client 2 achieved 0.9617. As training progressed, both clients demonstrated continuous improvement in classification performance. The highest accuracies were observed in round 6, where client 1 and client 2 achieved 0.9964 and 0.9968, respectively.

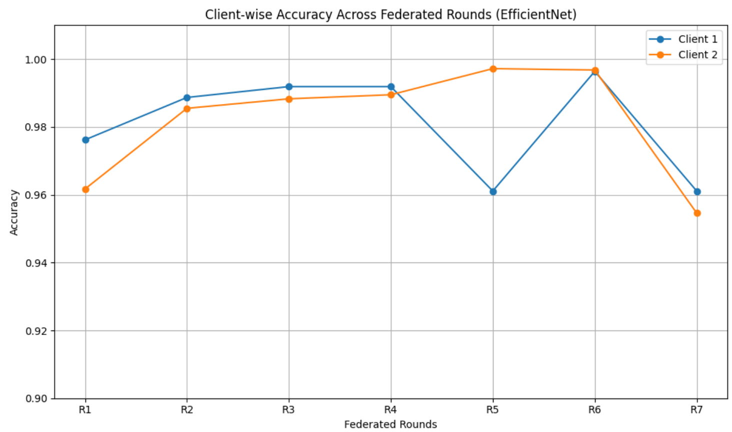

Figure 17. Client-wise accuracy of federated model using EfficientNet for rice leaf disease detection on client 1, 2 on round 1-7 on IID data.

Figure 18 represents a precision graph which shows a consistently high precision values for both clients throughout the training process. In first round client 1 and client 2 achieved precision value of 0.9763 and 0.9633 respectively. Further client 1 and client 2 reached their highest levels in round 6, where both clients achieved 0.9964, 0.9968 precision respectively.

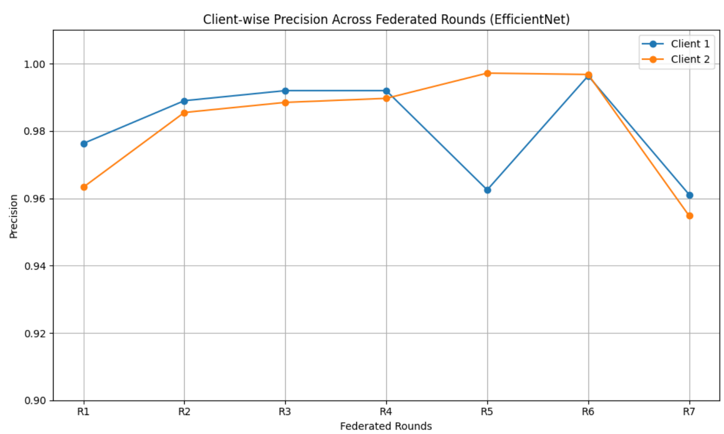

Figure 18. Client-wise precision of federated model using EfficientNet for rice leaf disease detection on client 1, 2 from round 1-7 on IID data.

Figure 19 presents the recall performance achieved by both clients over seven federated communication rounds on IID data. Initially, client 1 and client 2 obtained recall values of 0.9762 and 0.9617, respectively, in round 1. The highest recall values were achieved in round 6 wherein client 1 shows a recall of 0.9964 and client 2 exhibits a 0.9968 recall outcome.

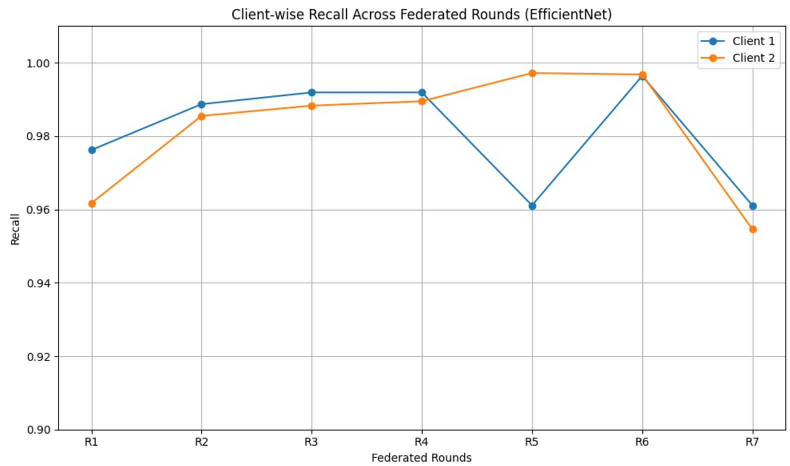

Figure 19. Client-wise recall of federated model using EfficientNet for rice leaf disease detection on client 1,2 from round 1-7 on IID data.

In **Figure 20** results of F1-score demonstrate balanced improvements in precision and recall throughout federated process. In first round, client 1 achieved F1 score value of 0.9760 and client 2 achieved 0.9613. These values increased to 0.9887 for client 1 and 0.9854 for client 2 in round 2 and further improved to 0.9919 and 0.9882 in round 3, while the highest F1-scores of 0.9964 achieved by client 1 and 0.9968 by client 2 in round 6. The consistently high F1-scores indicate that the EfficientNet model maintained an excellent balance between precision and recall throughout the federated learning process.

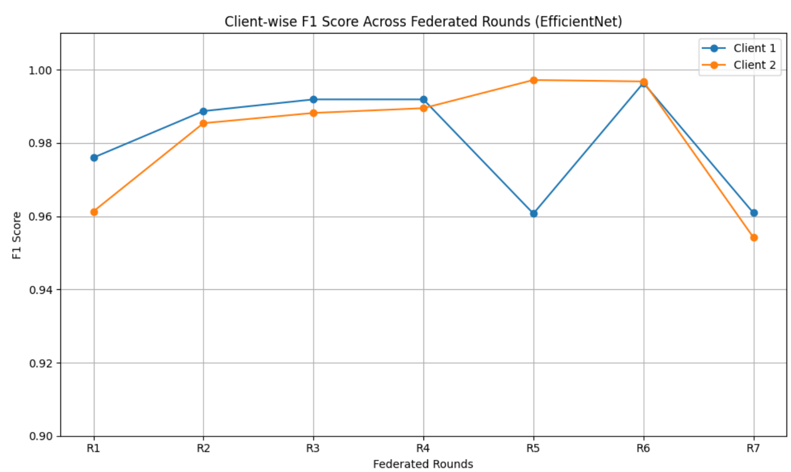

Figure 20. Client-wise F1-score of federated model using EfficientNet for rice leaf disease detection on client 1, 2 from round 1-7 on IID data.

Figure 21 demonstrates that Area Under the Curve (AUC) values are high across all communication rounds using federated EfficientNet for rice leaf disease detection. In round 1 client 1 and client 2 achieved AUC values of 0.9992 and 0.9988 respectively. The AUC further improved in round 2 to 0.9999 and 0.9997 while client 1 and 2 achieved a perfect AUC value of 1.0000 in round 6.

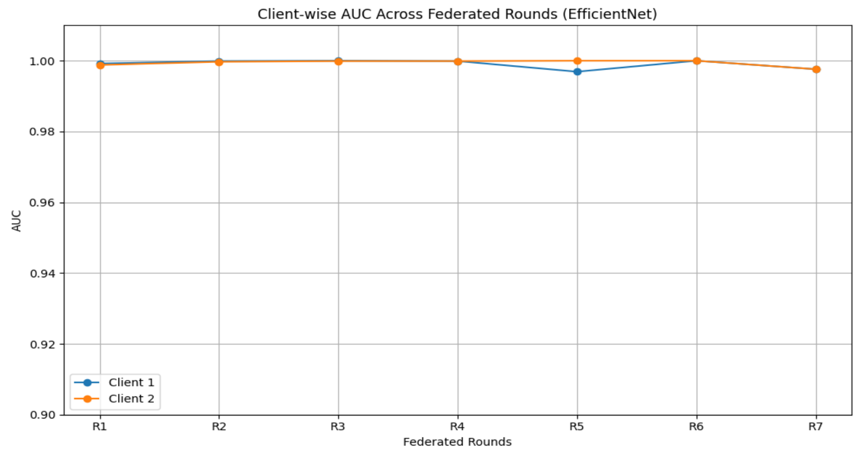

Figure 21. Client-wise AUC of federated model using EfficientNet for rice leaf disease detection on client 1,2 from round 1-7 on IID data.

Figure 22 represents the comparison of the global accuracy of federated MobileNetV2 and federated EfficientNet models across seven communication rounds. MobileNetV2 achieved a global accuracy of 0.7196 in round 1 which increased to 0.7925 by round 7. In comparison, EfficientNet started with a substantially higher global accuracy of 0.9238 in round 1 and reached to 0.9619 in round 5. Throughout, all communication rounds, EfficientNet consistently outperformed MobileNetV2.

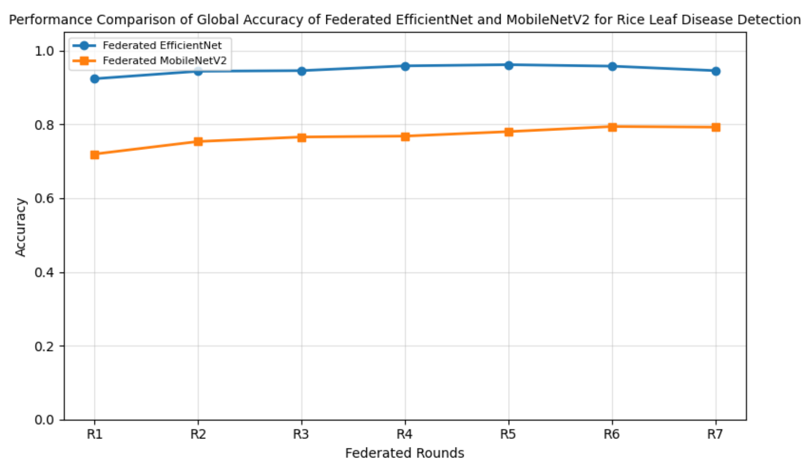

Figure 22. Performance comparison of global accuracy of federated EfficientNet and MobileNetV2 for rice leaf disease detection on IID data.

Figure 23 represents the comparison of the global precision of federated MobileNetV2 and federated EfficientNet models across seven communication rounds. MobileNetV2 achieved a global precision of 0.7281 in round 1 and 0.7933 by round 7. In comparison, EfficientNet achieved a higher global precision of 0.9237 in round 1 and 0.9619 in round 5.

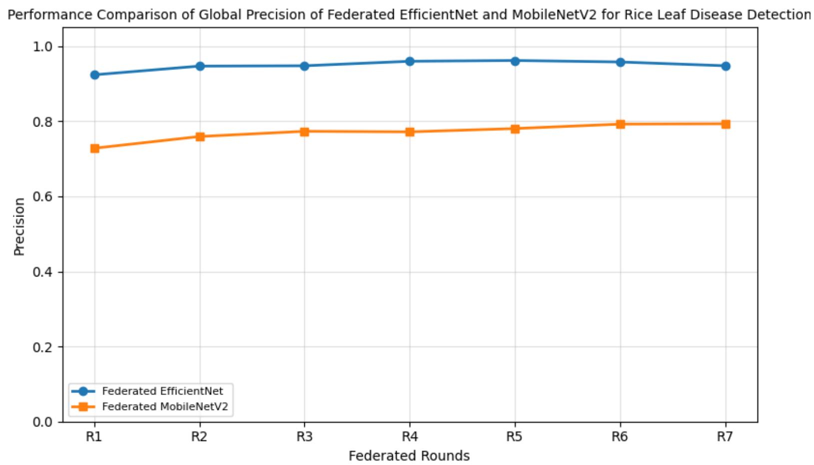

Figure 23. Performance comparison of global precision of federated EfficientNet and ederated MobileNetV2 for rice leaf disease detection on IID data.

Figure 24 illustrates the global recall comparison between federated MobileNetV2 and EfficientNet across seven federated communication rounds. MobileNetV2 showed an improvement with recall increasing from 0.7196 in round 1 to 0.7925 in round 7. In contrast, EfficientNet achieved a recall of 0.9238 in round 1 and attained highest value of 0.9619 in round 5. In all rounds EfficientNet maintained a higher recall values than MobileNetV2, which demonstrate a stronger ability to correctly identify diseased rice leaf samples.

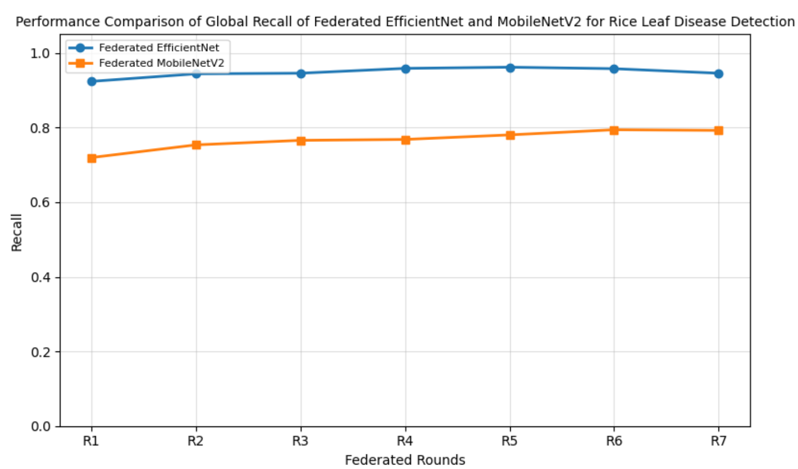

Figure 24. Performance comparison of global recall of federated EfficientNet and federated MobileNetV2 for rice leaf disease detection on IID data.

Figure 25 compares the global F1-scores of federated MobileNetV2 and EfficientNet the federated learning process. MobileNetV2 achieved F1 score value of 0.7209 in round 1 and 0.7905 in round 7, reflecting balanced improvements in precision and recall. EfficientNet, however, achieved a much higher F1-score of 0.9224 in round 1 and reached a maximum of 0.9617 in round 5.

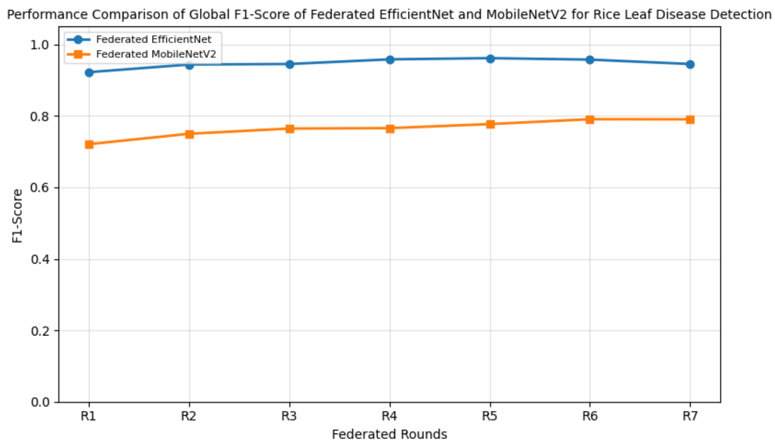

Figure 25. Performance comparison of global F1 score of federated efficientnet and federated MobileNetV2 for rice leaf disease detection on IID data.

Figure 26 presents the global AUC comparison between federated MobileNetV2 and EfficientNet for rice leaf disease detection. Federated MobileNetV2 achieved an AUC of 0.9394 in round 1 and 0.9656 in round 7. Although these values indicate good classification capability, EfficientNet consistently achieved higher AUC values, starting from 0.9948 in round 1 and 0.9979 in round 5. The AUC values obtained by EfficientNet demonstrate its exceptional ability to distinguish between healthy and diseased rice leaf classes.

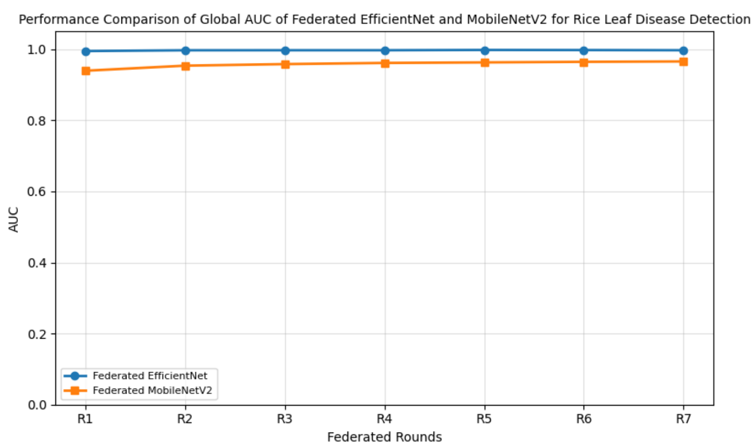

Figure 26. Performance comparison of global AUC of federated EfficientNet and federated MobileNetV2 for rice leaf disease detection on IID data.

The comparison between federated learning based model using MobileNetV2 and EfficientNet reveals that EfficientNet is the effective model in this scenario in terms of evaluation metrics. In all the rounds, EfficientNet was able to maintain high global accuracy values of 0.9441 in round 2, and it was 0.9619 in round 5. The global AUC value of EfficientNet was 0.997 in round 5, depicting that the separation of classes was rightly made by the EfficientNet. MobileNetV2 was not perform well, though it had improvements across rounds but it was able to reach a maximum global accuracy of 0.7942 in round 7. Moreover, its precision, recall, F1-score, and AUC was also quite good but still, there exists significant performance gap between Federated EfficientNet and Federated MobileNetV2. Federated EfficientNet exhibits better learning, good generalization, and faster convergence.

4.3 Global Performance Comparison of FedLeaf: Federated Learning Model using MobileNetV2 and Federated Learning Model using EfficientNet on IID Data Distribution

Table 5 presents a global performance comparison of FedLeaf: federated learning model using MobileNetv2 and federated learning model using EfficientNet on IID data distribution. The performance parameters show that EfficientNet has achieved a very good global accuracy of 0.9619 for rice leaf disease detection. The comparison based on performance metrics compares the best performances of the two federated deep learning models, MobileNetV2 and EfficientNet, visually through five primary evaluation metrics: accuracy, precision, recall, F1-score, and AUC. From the **Table 5**, it is seen that EfficientNet obtains the highest scores in all the metrics, and thus, it proves to have the good ability to learn and generalize the disease patterns from the rice leaf dataset. The results indicate that EfficientNet outperforms MobileNet, delivering the highest accuracy between the two.

Table 5. Global performance comparison of FedLeaf: federated learning model using MobileNetv2 and federated learning model using EfficientNet on IID data.

Model	Accuracy	Precision	Recall	F1 Score	AUC
Federated Learning model using MobileNetV2	0.7942	0.7921	0.7942	0.7908	0.9648
Federated Learning Model using EfficientNet	0.9619	0.9620	0.9619	0.9617	1.0000

Figure 27 presents a comparative evaluation of Federated MobileNetV2 and Federated EfficientNet which exhibits that Federated EfficientNet outperformed and achieved an accuracy of 0.9619 on IID distribution. It also shows high precision, recall, F1-score, and a perfect AUC of 1.0000. Federated EfficientNet utilizes the compound scaling method which harmonizes network depth, width, and resolution, and hence able to extract more robust and discriminative features even from complex augmented datasets. Federated MobileNetV2 achieves an accuracy of 0.7942. The comparison demonstrates that EfficientNet is the best model for rice leaf disease classification in federated learning, thereby exhibiting better accuracy, generalization, and stability across rounds.

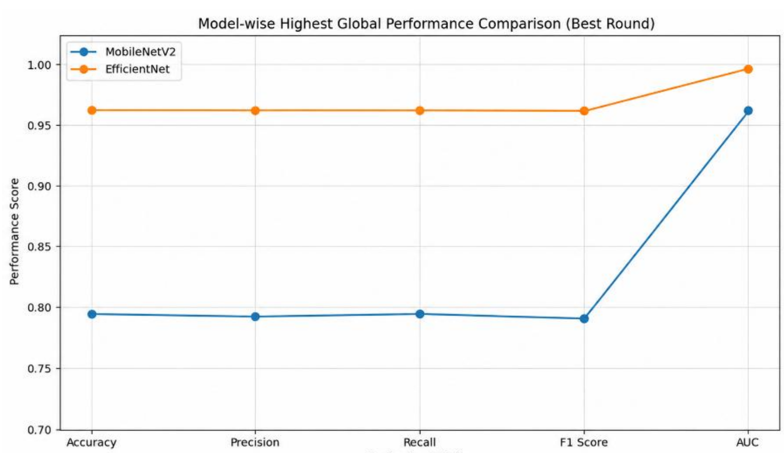

Figure 27. Global performance comparison of federated learning based MobileNetV2 model and federated learning based EfficientNet model.

4.4 Performance Evaluation of Federated Model using EfficientNet for Rice Leaf Disease detection on Client 1, Client 2 for Non-IID data Distribution

Federated learning Model using EfficientNet shows better performance for IID data distribution; hence, this model has been further utilised for the rice leaf disease detection for the non-IID data distribution, which is heterogeneous in nature. Non- IID data distribution refers to the heterogeneity of data distribution between clients, hence the split among the classes is not the same. In the case of non-IID data distribution, client 1 has 3393 images and client 2 has 1644 images, which have further varied class distribution in both clients. **Table 6** shows federated learning results with EfficientNet, which gives a strong and consistent performance for both clients and the global model through seven communication rounds on non-IID distribution. In the first round, the global model begins with an accuracy of 0.8667, indicating a good learning performance. As training progresses through multiple communication rounds, both client models achieve high local accuracies reaching up to 0.99 in later rounds, while the global model steadily improves, reaching a final accuracy of 0.9587, with corresponding precision of 0.9587, recall 0.9587, F1-score 0.9585, and AUC 0.9975. Overall, the results illustrate that Federated model using EfficientNet, achieves high accuracy and robust generalization in decentralized and heterogeneous data environments.

Table 6. Performance evaluation of federated model using EfficientNet for rice leaf disease detection on client 1, client 2 on non IID data distribution.

Model	Client	Accuracy	Precision	Recall	F1 Score	AUC
EfficientNet Round 1 (5 epochs)	Client 1	0.8960	0.8968	0.8960	0.8923	0.9874
	Client 2	0.8822	0.8829	0.8822	0.8789	0.9840
	Global Metrics After Round 1	0.8667	0.8666	0.8667	0.8616	0.9807
EfficientNet Round 2 (10 epochs)	Client 1	0.9527	0.9532	0.9527	0.9521	0.9971
	Client 2	0.9389	0.9394	0.9389	0.9380	0.9961
	Global Metrics After Round 2	0.9071	0.9062	0.9071	0.9054	0.9919
EfficientNet Round 3 (15 epochs)	Client 1	0.9704	0.9706	0.9704	0.9702	0.9989
	Client 2	0.9609	0.9610	0.9609	0.9607	0.9984
	Global Metrics After Round 3	0.9278	0.9271	0.9278	0.9270	0.9947
EfficientNet Round 4 (20 epochs)	Client 1	0.9805	0.9806	0.9805	0.9804	0.9995
	Client 2	0.9729	0.9730	0.9729	0.9728	0.9991
	Global Metrics After Round 4	0.9397	0.9395	0.9397	0.9392	0.9960
EfficientNet Round 5 (25 epochs)	Client 1	0.9842	0.9843	0.9842	0.9842	0.9997
	Client 2	0.9817	0.9818	0.9817	0.9817	0.9995
	Global Metrics After Round 5	0.9524	0.9523	0.9524	0.9520	0.9967
EfficientNet Round 6 (30 epochs)	Client 1	0.9874	0.9874	0.9874	0.9874	0.9998
	Client 2	0.9842	0.9843	0.9842	0.9842	0.9997
	Global Metrics After Round 6	0.9524	0.9524	0.9524	0.9520	0.9970
EfficientNet Round 7 (35 epochs)	Client 1	0.9924	0.9924	0.9924	0.9924	0.9999
	Client 2	0.9905	0.9906	0.9905	0.9905	0.9998
	Global Metrics After Round 7	0.9587	0.9587	0.9587	0.9585	0.9975

Figure 28 illustrates the client-wise accuracy of the federated EfficientNet model across seven communication rounds for rice leaf disease detection on non-IID data. In round 1, client 1 achieved an accuracy of 0.8960 while client 2 achieved 0.8822. As training progressed, both clients showed a consistent improvement in performance. The accuracy of client 1 increased to 0.9527 in round 2, 0.9704 in round 3 and 0.9805 in round 4. Similarly, client 2 improved from 0.9389 in round 2 to 0.9729 in round 4. The highest accuracies were achieved by the federated model using EfficientNet in round 7, where client 1 achieved a value of 0.9924 and client 2 achieved 0.9905.

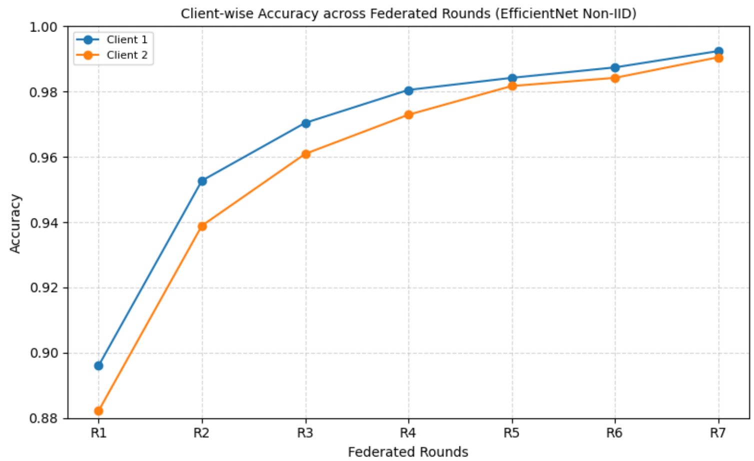

Figure 28. Client-wise accuracy of federated model using EfficientNet for rice leaf disease detection on client 1, client 2 for non-IID data distribution.

Figure 29 presents the precision values achieved by both clients throughout the federated learning process using federated EfficientNet for rice leaf disease detection on non-IID data. In round 1, client 1 and client 2 obtained precision scores of 0.8968 and 0.8829, respectively. Further improvements were observed in subsequent rounds, wherein client 1 achieved 0.9874 and client 2 achieved 0.9843 in round 6. The highest precision values were achieved in round 7 by achieving 0.9924 for client 1 and 0.9906 for client 2.

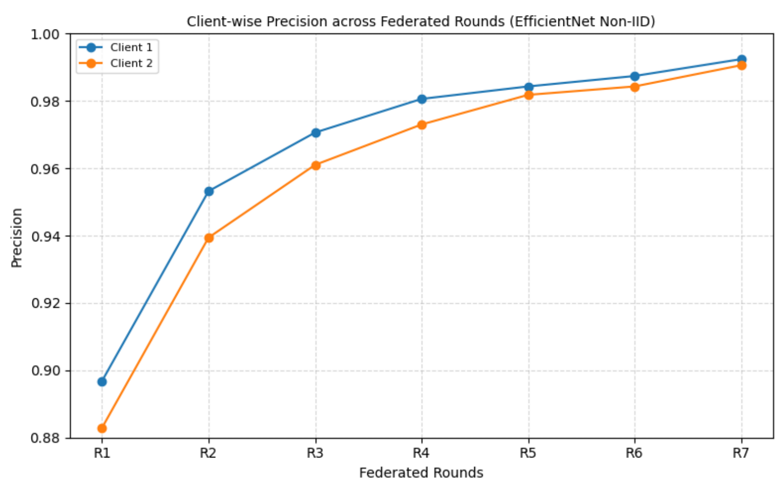

Figure 29. Client-wise precision of federated model using EfficientNet for rice leaf disease detection on client 1, client 2 for non-IID data distribution.

Figure 30 shows the recall performance of both clients across seven communication rounds using federated EfficientNet on non-IID data. Initially, client 1 and client 2 achieved recall values of 0.8960 and 0.8822, respectively. The highest recall values were observed in round 7, where client 1 achieved 0.9924 and client 2 achieved 0.9905. The consistently increasing recall values indicate that the model became more effective at correctly identifying rice leaf disease samples while reducing misclassification values.

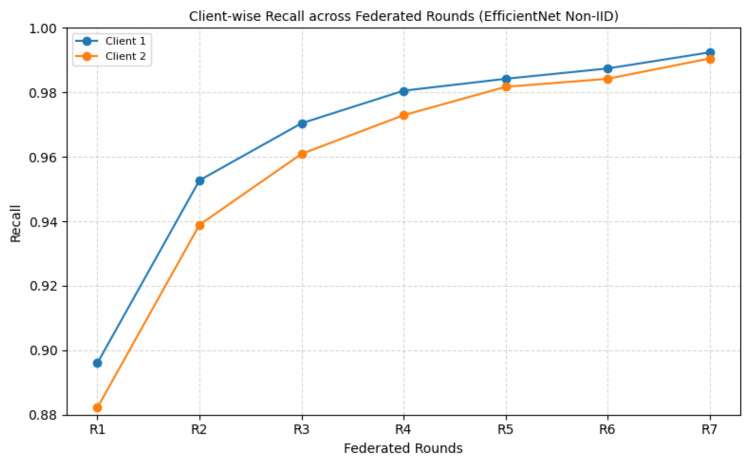

Figure 30. Client-wise recall of federated model using EfficientNet for rice leaf disease detection on client 1, client 2 for non-IID data distribution.

Figure 31 shows F1-score results, which demonstrate balanced improvements in precision and recall throughout federated training using federated EfficientNet on non-IID data distribution. In round 1, client 1 achieved an F1-score of 0.8923 while client 2 achieved 0.8789. These values increased to 0.9521 and 0.9380 in round 2 and further improved to 0.9702 and 0.9607 in round 3 for client 1 and client 2 respectively. By round 5 both clients achieved F1-scores of 0.9842, 0.9817 respectively and the highest values were achieved in round 7 as 0.9924 for client 1 and 0.9905 for client 2.

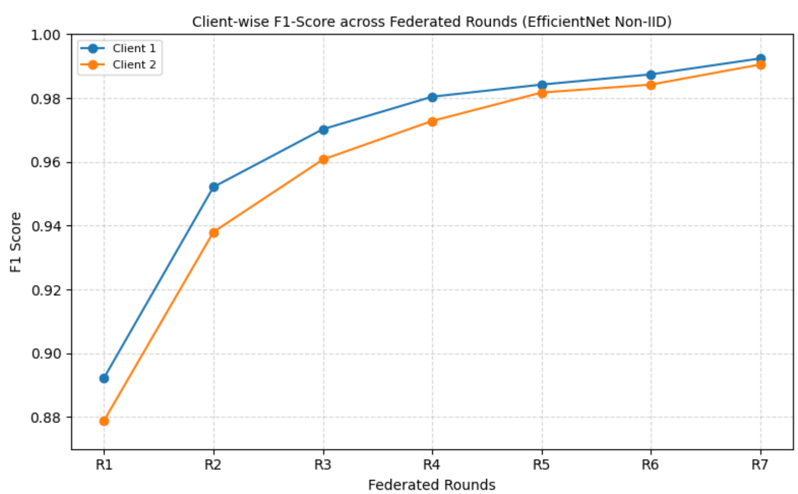

Figure 31. Client-wise F1-score of federated model using EfficientNet for rice leaf disease detection on client 1, client 2 for non-IID data distribution.

Figure 32 illustrates the Area Under the Curve (AUC) values obtained by both clients over seven communication rounds using federated EfficientNet on non-IID data distribution. In round 1, client 1 achieved an AUC of 0.9874 and client 2 achieved 0.9840. Finally, an AUC value of 0.9999 was achieved by client 1 in round 7, and client 2 achieved AUC of 0.9998 in the final round.

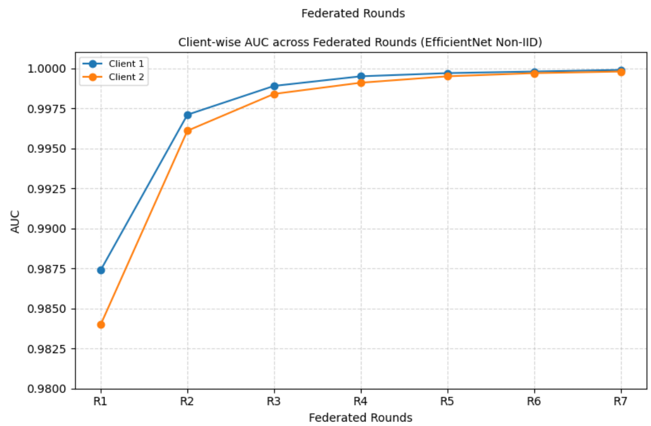

Figure 32. Client-wise AUC of federated model using EfficientNet for rice leaf disease detection on client 1, client 2 for non-IID data distribution.

The **Figures 33** to **37** show a graphical representation of the global performance of the federated learning model based model using EfficientNet on non-IID data distribution. Federated EfficientNet achieves global accuracy of 0.8667 in round 1 due to its strong feature extraction capability and effective transfer learning. Further, it exhibits a steady and consistent improvement in global accuracy eventually, reaching approximately 95.87% in round 7 indicating efficient aggregation of clients through federated averaging. The global performance metrics of the EfficientNet-based federated learning model demonstrated substantial improvement throughout the federated rounds. Global precision increased from 0.8666 in round 1 to its highest value of 0.9587 in round 7, while global recall improved from 0.8667 in round 1 to 0.9587 in round 7, indicating enhanced disease detection capability. Similarly, the global F1-score increased from 0.8616 in round 1 to 0.9585 in round 7, reflecting a balanced improvement in both precision and recall. The global AUC also increased from an already high value of 0.9807 in round 1 to 0.9975 in round 7, demonstrating model's strong ability to distinguish among different rice leaf disease categories.

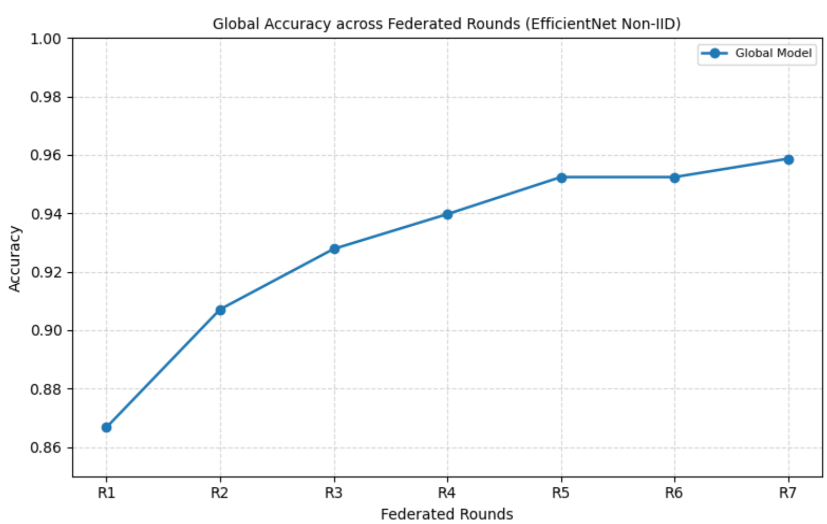

Figure 33. Global accuracy of federated EfficientNet for rice leaf disease detection on non-IID data distribution.

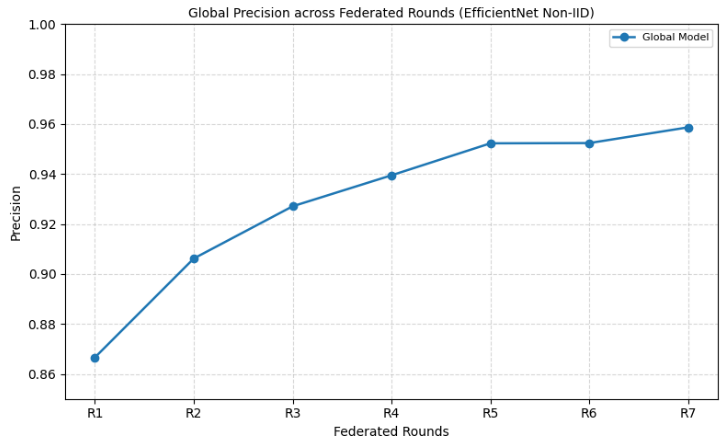

Figure 34. Global precision of federated EfficientNet for rice leaf disease detection on non-IID data.

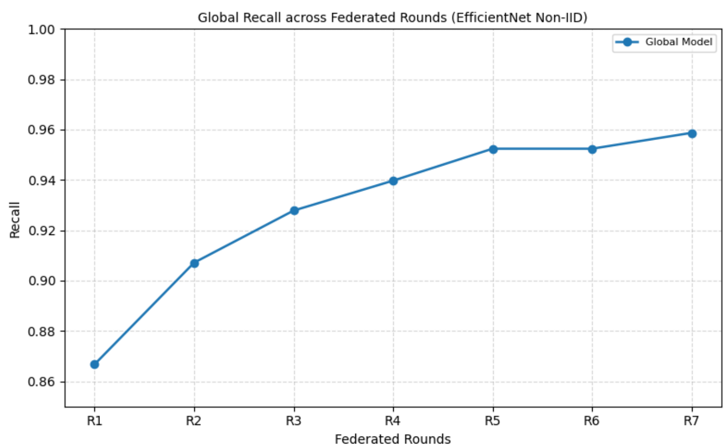

Figure 35. Global recall of federated EfficientNet for rice leaf disease detection on non-IID data.

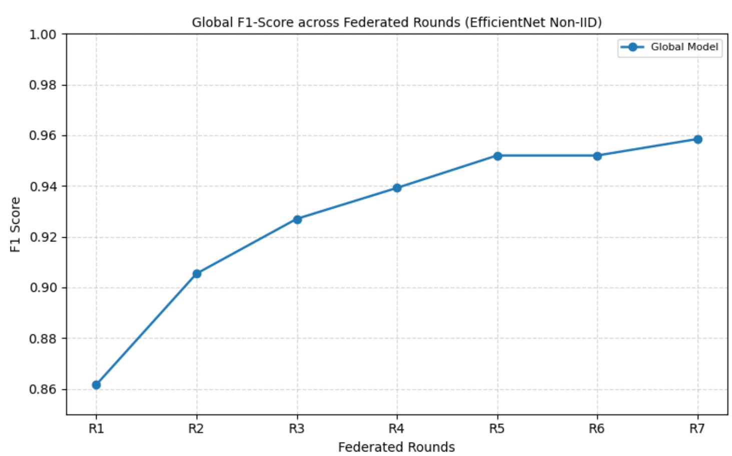

Figure 36. Global F1 score of federated EfficientNet for rice leaf disease detection on non-IID data.

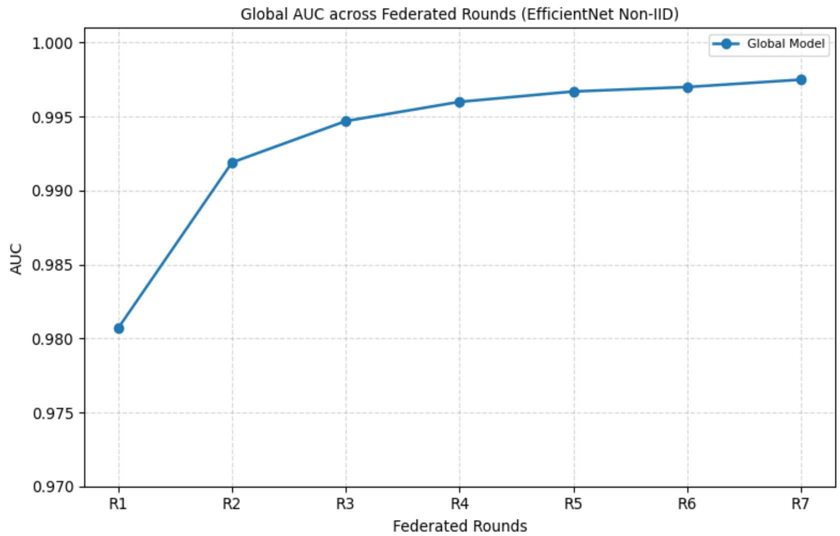

Figure 37. Global AUC of federated EfficientNet for rice leaf disease detection on non-IID data.

4.5 Performance Comparison with State-of-the-art Models

Table 7 summarizes the comparison of the proposed work along with the existing works for the detection of rice leaf diseases using collaborative federated learning based models considering dataset size, number of communication rounds, epochs per communication round, learning rate, achieved accuracy, strengths and future scope.

Authors (Leema et al., 2023; Tripathy et al., 2024) considered only 120 images that have 3 rice leaf diseases, trained for 100 epochs, used 2 clients and a batch size of 32 and yet, in these favorable conditions and small datasets, the accuracy achieved by Leema et al. (2023) was 99% and Tripathy et al. (2024) has achieved recall value of 0.922. It is due to the fact that the models were trained on a small and less complex dataset, which may lead to a high risk of overfitting. In Chorney et al. (2025), authors considered large size dataset of 10553 images with 8 rice leaf classes (7 disease and one healthy) however, the model's accuracy was limited to 83.14%. In Aggarwal et al. (2023), the authors have achieved good competitive accuracy of 95% on Non-IID data distribution, though the dataset size is considerably good but it has comparatively small dataset size and less number of disease classes as compared to the proposed work. The proposed work demonstrates the federated learning performance on a comparatively large and complex dataset of 6297 images, 6 classes, including 5 disease categories and one healthy class. The model was implemented using 7 communication rounds with a batch size of 32 and 5 epochs per communication round across two client scenarios. The proposed model, therefore, achieves a higher accuracy of 96.19% on IID data distribution and 95.87% on non-IID distribution under higher heterogeneity and more disease classes, which exhibits its good robustness, and generalization ability. The comparison shows that the proposed method achieves high accuracy on a significantly large, multi-class dataset in a federated learning environment that is more appropriate for the real-world applications of precision agriculture for rice leaf disease detection.

Table 7. Performance comparison with state-of-the-art models.

References	Dataset size	No. of classes	No. of rounds	No. of epochs per communic-ation round	Learning rate	Accuracy in non-IID data	Strength/Pros	Future scope/Limitations
Tripathy et al. (2024)	120	3	--Not Mentioned--	20-100	--Not Mentioned	Accuracy has not mentioned Non-IID data has not been considered	●Optimization technique applied for efficiency.	●Small Size of dataset. ●Less number of disease classes. ●High risk of overfitting. ●Worked only on IID data, Non-IID data has not been considered.
Leema et al. (2023)	120	3	100	1	0.001	99% on Non-IID Data	●Non-IID data used. ●Self attention mechanism for feature extraction.	●Small Size of dataset. ●Less number of classes. ●High risk of overfitting.
Chorney et al. (2025)	10553	8(7 disease classes, 1 healthy leaf class)	650	1	0.01	83.14% on Non-IID	●Non-IID data used. ●More disease classes. ●Large dataset size.	●Partial data sharing. ●Less Accuracy of the model.
Aggarwal et al. (2023)	5932	4	10	10	0.001	95% on Non-IID	●IID and Non-IID data distribution were considered.	●Less number of disease classes were considered.
Proposed FedLeaf Federated Learning based EfficientNet	6297 (After augmentation)	6 (5 disease classes, 1 healthy leaf class)	7	5	0.001	95.87% on Non-IID	●Large size of dataset. ●More disease classes. ●Experiment done on both IID and Non-IID data.	●Model scalability can be enhanced by increasing more number of clients.

5. Conclusion

The overall experimental results exhibit that the proposed FedLeaf, a collaborative federated learning system has promising classification performance for rice leaf disease detection, across different model architectures for IID and non-IID data distributions. FedLeaf has been constructed using two deep learning architectures, namely federated learning based MobileNetV2 and federated learning based EfficientNet. Between the two proposed architectures of FedLeaf, findings consistently illustrate that EfficientNet, achieves maximum global accuracy of 96.19% together with the highest values of precision, recall, F1-score, and AUC for IID data distribution. MobileNetV2 demonstrates average yet consistent performance. Though, existing works show good accuracy, most of the results are obtained using a comparatively small dataset and a smaller number of classes. On the other hand, the proposed work is conducted on a considerably large dataset of 6297 images for 6 classes (five rice leaf diseases and 1 healthy class). The federated EfficientNet model shows an accuracy of 96.19% for IID distribution and 95.87% for non-IID data distribution, proving that the proposed system exhibits high performance and is better suited for early rice leaf disease detection. The performance comparison on IID and non-IID data distribution, and model-wise comparisons exhibit that FedLeaf, a federated learning approach, can ensure a better collaborative model that deals with heterogeneity of data and at the same time demonstrates a high predictive performance. Specifically, federated learning based EfficientNet exhibits better performance for the multiclass rice leaf disease detection because it is an efficient model with balanced depth-width-resolution scaling and good feature extraction capabilities. The proposed FedLeaf, collaborative federated learning model is more appropriate for real-world precision agricultural scenarios for automatic rice leaf disease detection. In future work, the research can consider multi-client setups with an increased number of clients to make the model more scalable.

Conflicts of Interest
The Authors declare that they do not have any conflicts of interest.

Acknowledgments
The author would like to thank the editors and anonymous reviewers for their valuable comments and suggestions. This research did not receive any specific grant from funding agencies in public, commercial, or-not-for-profit sector.

The contribution of the authors is: Amandeep Kaur: Conceptualization, Methodology, Writing - original draft, Data Curation, Formal analysis, Validation; Kalpna Guleria: Conceptualization, Methodology, Formal analysis, Validation, Writing -review & editing, Supervision.

The dataset was obtained from the open-source Kaggle repository.
https://www.kaggle.com/datasets/dedeikhsandwisaputra/rice-leafs-disease-dataset.

AI Discourse
The author(s) declare that no assistance is taken from generative AI to write this article.

References

Aggarwal, M., Khullar, V., Goyal, N., & Prola, T.A. (2024). Resource-efficient federated learning over IoAT for rice leaf disease classification. *Computers and Electronics in Agriculture, 221*, 109001. https://doi.org/10.1016/j.compag.2024.109001

Aggarwal, M., Khullar, V., Goyal, N., Alammari, A., Albahar, M.A., & Singh, A. (2023). Lightweight federated learning for rice leaf disease classification using non-independent and identically distributed images. *Sustainability, 15*(16), 12149. https://doi.org/10.3390/su151612149

Antico, T.M., Moreira, L.F.R., & Moreira, R. (2024). Evaluating the potential of federated learning for maize leaf disease prediction. *Machine Learning. arXiv preprint arXiv:2412.07872*

Behera, S., Padhy, N., Panigrahi, R., & Kuanar, S.K. (2025). Crop disease prediction using deep learning in a federated learning environment: ensuring data privacy and agricultural sustainability. *Procedia Computer Science, 254*, 137-146. https://doi.org/10.1016/j.procs.2025.02.072

Chitta, S., Yandrapalli, V.K., & Sharma, S. (2024). Deep learning for precision agriculture: evaluating CNNs and vision transformers in rice disease classification. In *Proceedings of the OPJU International Technology Conference on Smart Computing for Innovation and Advancement in Industry 4.0* (pp. 1-6). IEEE. Raigarh, India. https://doi.org/10.1109/OTCON60325.2024.10687983

Chorney, W., Rahman, A., Wang, Y., Wang, H., & Peng, Z. (2025). Federated learning for heterogeneous multi-site crop disease diagnosis. *Mathematics*, *13*(9), 1401. https://doi.org/10.3390/math13091401

Deng, F., Mao, W., Zeng, Z., Zeng, H., & Wei, B. (2022). Multiple diseases and pests detection based on federated learning and improved faster R-CNN. *IEEE Transactions on Instrumentation and Measurement, 71*, 1-11. https://doi.org/10.1109/TIM.2022.3201937

Elaraby, A., Hamdy, W., & Alruwaili, M. (2022). Optimization of deep learning model for plant disease detection using particle swarm optimizer. *Computers, Materials & Continua, 71*(2), 4019-4031. http://dx.doi.org/10.32604/cmc.2022.022161

Erickson, B.J. (2021). Basic artificial intelligence techniques: machine learning and deep learning. *Radiologic Clinics, 59*(6), 933-940. https://doi.org/10.1016/j.rcl.2021.06.004

Fahim-Ul-Islam, M., Chakrabarty, A., Ahmed, S.T., Rahman, R., Kwon, H.H., & Piran, M.J. (2024). A comprehensive approach towards wheat leaf disease identification leveraging transformer models and federated learning. *IEEE Access, 12,* 09128-109156. https://doi.org/10.1109/ACCESS.2024.3438544

Fetzer, J.H. (1990). What is artificial intelligence? In: Fetzer, J.H. (ed) *Artificial Intelligence: its Scope and Limits*. Springer, Netherlands, Dordrecht, pp. 3-27. https://doi.org/10.1007/978-94-009-1900-6_1

Fu, X., Ma, Q., Yang, F., Zhang, C., Zhao, X., Chang, F., & Han, L. (2024). Crop pest image recognition based on the improved ViT method. *Information Processing in Agriculture, 11*(2), 249-259. https://doi.org/10.1016/j.inpa.2023.02.007

Fukagawa, N.K., & Ziska, L.H. (2019). Rice: importance for global nutrition. *Journal of Nutritional Science and Vitaminology, 65*(Supplement), S2-S3. https://doi.org/10.3177/jnsv.65.S2

Gnanamanickam, S.S. (2009). Rice and its importance to human life. In *Biological Control of Rice Diseases* (pp. 1-11). Springer, Netherlands, Dordrecht. https://doi.org/10.1007/978-90-481-2465-7_1

Hari, P., Singh, M.P., & Singh, A.K. (2024). An improved federated deep learning for plant leaf disease detection. *Multimedia Tools and Applications, 83*(35), 83471-83491. https://doi.org/10.1007/s11042-024-18867-9

Henry, W.P., & Winston, P.H. (2004). *Artificial intelligence*. Addison Wesley Longman Publishing Co., Inc. ISBN: 0201855046.

Hsieh, K., Phanishayee, A., Mutlu, O., & Gibbons, P. (2020). The non-IID data quagmire of decentralized machine learning. *Proceedings of 37th International Conference on Machine Learning* (Vol. 119, pp. 4387-4398). PMLR. https://proceedings.mlr.press/v119/hsieh20a.html

Idoje, G., Dagiuklas, T., & Iqbal, M. (2023). Federated learning: crop classification in a smart farm decentralised network. *Smart Agricultural Technology, 5*, 100277. https://doi.org/10.1016/j.atech.2023.100277

Jakhar, D., & Kaur, I. (2020). Artificial intelligence, machine learning and deep learning: definitions and differences. *Clinical and Experimental Dermatology, 45*(1), 131-132. https://doi.org/10.1111/ced.14029

Kabala, D.M., Hafiane, A., Bobelin, L., & Canals, R. (2023). Image-based crop disease detection with federated learning. *Scientific Reports, 13*(1), 19220. https://doi.org/10.1038/s41598-023-46218-5

Laha, G.S., Singh, R., Ladhalakshmi, D., Sunder, S., Prasad, M.S., Dagar, C.S., & Babu, V.R. (2017). Importance and management of rice diseases: a global perspective. In: Chauhan, B.S., Jabran, K., Mahajan, G. (eds) *Rice Production Worldwide* (pp. 303-360). Springer International Publishing, Cham. https://doi.org/10.1007/978-3-319-47516-5_13

Leema, A.A., & Balakrishnan, P. (2024). Detection and categorization of rice leaf diseases through federated learning and improved vision transformer models. *Journal of Wireless Mobile Networks, Ubiquitous Computing, and Dependable Applications, 15*(4), 370-379. https://doi.org/10.58346/JOWUA.2024.I4.025

Lu, Z., Pan, H., Dai, Y., Si, X., & Zhang, Y. (2024). Federated learning with non-IID data: a survey. *IEEE Internet of Things Journal, 11*(11), 19188-19209. https://doi.org/1109/JIOT.2024.3376548

Mohammed, B. (2025). A comprehensive overview of federated learning for next-generation smart agriculture: current trends, challenges, and future directions. *Informatica, 49*(1), 117-136. https://doi.org/10.31449/inf.v49i1.6764

Nilsson, A., Smith, S., Ulm, G., Gustavsson, E., & Jirstrand, M. (2018). A performance evaluation of federated learning algorithms. In *Proceedings of the Workshop on Distributed Infrastructures for Deep Learning* (pp. 1-8). Association for Computing Machinery. New York, USA. https://doi.org/10.1145/3286490.3286559

Ongsulee, P. (2017). Artificial intelligence, machine learning and deep learning. In *2017 Proceedings of the 15th International Conference on ICT and Knowledge Engineering* (pp. 1-6). IEEE. Bangkok, Thailand. https://doi.org/10.1109/ICTKE.2017.8259629

Pothen, M.E., & Pai, M.L. (2020). Detection of rice leaf diseases using image processing. In *2020 Proceedings of the International Conference on Computing Methodologies and Communication* (pp. 424-430). IEEE. Erode, India. https://doi.org/10.1109/ICCMC48092.2020.ICCMC-00080

Saputra, D.I.D. (n.d.). Rice leafs disease dataset. *Kaggle*. https://www.kaggle.com/datasets/dedeikhsandwisaputra/rice-leafs-disease-dataset

Shao, Y., Yang, W., Lu, Z., Geng, H., & Chen, D. (2024). Critical information mining network: identifying crop diseases in noisy environments. *Symmetry, 16*(6), 652. https://doi.org/10.3390/sym16060652

Soni, J.A., Pandya, N.R., Patel, R.B., & Raval, J.S. (2025). Privacy-preserving AI in agriculture: a review of federated learning approaches. *SPU-Journal of Science, Technology and Management Research*, *2*(01), 36-46.

Tong, J., Zou, Y., Zeng, Z., & Deng, F. (2022). Detection technology of several pests and diseases in orchard based on federated learning and improved faster R-CNN. In *2nd International Conference on Artificial Intelligence, Big Data and Algorithms* (pp. 1-4). VDE. Nanjing, China.

Tripathy, R., Mandala, J., Pappu, S.R., & Gopisetty, G.K.D. (2024). Optimization based rice leaf disease classification in federated learning. *Multimedia Tools and Applications, 83*(29), 72491-72517. https://doi.org/10.1007/s11042-023-18085-9

Zhao, Y., Li, M., Lai, L., Suda, N., Civin, D., & Chandra, V. (2018). Federated learning with non-IID data. *Machine Learning.* arXiv preprint arXiv:1806.00582