Physics-Guided Federated Deep Learning for Wide-Area Plant Disease Mapping

1st Veluguri Venkata Lakshmi Anusha
dept.of computer science and engineering
Narasaraopeta Engineering college
Narasaraopeta,India
velugurianusha30@gmail.com

2nd Gajavalli Sai Sahithi
dept.of computer science and engineering
Narasaraopeta Engineering college
Narasaraopeta,India
saisahithigajavalli@gmail.com

3rd S.Shobha rani
dept.of computer science and engineering
Narasaraopeta Engineering college
Narasaraopeta,India
Shobharaninec10@gmail.com

4th K.V.Narasimha Reddy
dept.of computer science and engineering
Narasaraopeta Engineering college
Narasaraopeta,India
narasimhareddynec03@gmail.com

***Abstract*—The monitoring of plant diseases in large agricultural areas requires large-scale, private learning systems capable of adapting to various types of farm environment. In this paper, we propose a federated deep learning system guided by physical principles to create large-scale plant disease probability maps, as well as to preserve the privacy of farmers' data. We address the two main challenges in large-scale agricultural analytics: protecting sensitive farm data and the high degree of non IID-conditions due to the geographical distribution of farms. To protect sensitive farm data and to account for the non IID-conditions associated with large-scale agricultural data, we use federated learning to enable the collaborative training of a deep neural network among multiple farm client devices using local images of plants, without sharing their raw data. The central server aggregates the locally trained parameters to generate a global model which represents the collective knowledge from all distributed farms. To ensure that the models can generalize under heterogeneous conditions, we incorporate a physics-guided constraint module into the local training process, to regularize the feature representation and to incorporate domain knowledge regarding how plant diseases progress. We test our approach using a multi-class tomato plant disease dataset that contains seven different disease classes. Our results demonstrate that the centralized model achieved an accuracy of 0.95 and a Macro F1 score of 0.94, whereas the conventional FedAvg algorithm was severely affected by the heterogeneous nature of the data, resulting in an accuracy of 0.70 and a Macro F1 score of 0.68. Furthermore, the physics-guided federated model proposed in this work improves the performance to an accuracy of 0.84 and a Macro F1 score of 0.83. The convergence analysis demonstrates that there are consistent improvements throughout each round of communication, and the ablation studies confirm that the incorporation of constraints for consistency of reflectance, modeling of the growth of lesions, and spatial continuity sequentially improves predictive performance. Finally, the framework also produces wide-area disease probability maps allowing for the early detection of infection hotspots in support of large-scale, privacy-preserving crop health monitoring.**

***Index Terms*—Federated learning, plant disease detection, physics-guided learning, precision agriculture, distributed machine learning, deep learning, plant pathology, agricultural monitoring.**

I. INTRODUCTION

Centralized approaches are still the dominant methodologies in the development of machine learning models for the recognition of plant disease, and thus the majority studies have been based on centralized approaches. Centralized approaches require all of the data to be uploaded to a remote server for processing which raises a number of problems for agricultural applications including; communication costs, potential privacy violations of sensitive farmer data, and the reluctance of farmers or agricultural organizations to share their raw data for use in a centralized model. In addition, datasets collected from different geographic areas can vary greatly in terms of environmental conditions, types of crops grown, and devices used to capture images of plants. These large variations will make it difficult to develop robust models that are applicable in many agricultural environments.

In contrast to traditional centralized approaches, federated learning allows multiple geographically separated clients to train local machine learning models using their own private datasets while occasionally sending model parameters to a central aggregation server for combination into a larger shared model [5], [6], [9], [10]. This enables the building of collaborative models that preserve privacy and has attracted a growing interest in applications such as smart agriculture, healthcare, and edge computing. Recently, there has been increasing evidence for the application of federated learning in the area of plant disease detection, specifically by enabling multiple farms to collaboratively build plant disease recognition models while protecting the privacy of individual farms' data [15]–[18].

However, federated learning in real-world agricultural settings is very challenging due to the heterogeneity of the data generated by different farms. Due to factors such as differences in crop type, levels of disease prevalence, and environmental conditions, data are typically non-identically distributed when collected from different farms. Non-identically distributed data can result in unreliable convergence of models and poor performance when using federated optimization algorithms like FedAvg [6], [9], which are widely used in federated learning but lack the ability to deal with non-identical data distributions. Additionally, current federated learning solutions are primarily data-driven, lacking a mechanism to incorporate biological knowledge about how plant diseases develop and spread.

Physics-Informed Machine Learning (PIML), a new and rapidly growing area of research that combines machine learning with physics-based knowledge to enhance both the robustness and explainability of ML models by constraining them through known physical principles during training, has shown much promise as a method to enhance the robustness and explainability of ML models [7], [8], [18]. PIML models utilize known physical principles to constrain the way that models operate and generate predictions, thereby producing models that are consistent and reliable. For example, in plant disease detection, biological properties of plant disease development such as reflectance patterns of leaves affected by disease, growth behaviors of lesions caused by disease, and disease spread behaviors over time provide rich information that can inform how a model learns and enhances the ability of the model to generalize across various agricultural environments.

This paper was motivated by the need to address these challenges and propose a Physics-Guided Federated Deep Learning Framework for Wide-Area Plant Disease Mapping. This framework incorporates physical constraints related to the development and propagation of plant diseases into the federated learning process to increase the robustness of the model in a variety of agricultural settings. Each distributed client on each farm trains a deep neural network locally while enforcing physical constraints that ensure the reflectance consistency of leaves, the growth of lesions associated with disease, and the continuity of disease in space. Once trained locally, the models are then combined at a federated server to create a single global model that can support large-scale agricultural monitoring while preserving the privacy of the individual farms' data.

To evaluate the effectiveness of the proposed framework, we performed experiments using a multi-class tomato plant disease dataset that includes seven classes of diseases. Our comparative evaluations demonstrate that although centralized approaches to machine learning achieve excellent performance, federated approaches experience a significant drop-off in performance due to the heterogeneity of the data held at the various farms. Furthermore, our physics-guided federated learning framework enhances model stability and predictive performance across distributed clients and enables the creation of wide-area maps of disease probabilities to facilitate the early detection of infection hotspots.

The remainder of this paper is structured as follows. Section II provides a review of previous work on plant disease detection, federated learning in agriculture, and physics-informed machine learning. Section III details the proposed physics-guided federated learning framework and its structure. Section IV discusses the experimental design and results of the study, together with our analysis. Section V provides a summary of the paper and identifies avenues for future research.

II. RELATED WORK

Automated plant disease detection has become increasingly popular in the last few years primarily due to advances in deep learning and computer vision. Studies in the early stages have shown that convolutional neural networks can be successfully used to identify plant diseases through leaf images by using large datasets, such as Plant Village [1]–[3]. These studies have also expanded their approaches to include more complex field environments and various crop species. Studies that conducted comprehensive reviews of these types of deep learning models (ResNet, VGG, EfficientNet) have found them to be effective in detecting crop diseases and agricultural monitoring [4], [12]. Although the use of deep learning models is highly accurate in identifying classifications; they generally utilize centralized training methods and require large amounts of labeled data to be collected from multiple agricultural regions.

Federated learning has been explored in agricultural applications to help address these privacy issues and to allow distributed learning across farms. Federated learning enables multiple clients to collectively train machine learning models without transmitting raw data to a central server. Previous studies have used federated convolutional neural networks to detect crop diseases and have demonstrated that federated models can provide competitive results when it comes to data privacy compared to centralized approaches [9], [15]. Additionally, recent studies have developed federated learning frameworks for distributed crop monitoring and edge-based agricultural systems, demonstrating that federated learning can be used to develop scalable smart agriculture applications [10], [17].

Despite these advances, federated learning in agriculture still has many challenges. The primary challenge facing federated learning in agriculture is the issue of having heterogeneous data distributions between farms. Variation in crop varieties, disease severity, environmental conditions, and imaging setups can result in non-IID data distributions that can degrade the performance of the federated model [6], [9]. Studies that have reviewed federated learning for smart agriculture have identified that the development of models that are robust to varying conditions remains a key area of study [19], [20].

Physics-informed machine learning has emerged as a potential method for improving deep learning models by incorporating domain knowledge into the training of neural networks. Physics-informed neural networks are capable of incorporating physical constraints into the learning process so that the model will respect the known dynamics of the system and make more reliable predictions [7], [8], [18]. Although physics-informed learning has been extensively studied in the fields of scientific computing and engineering, the application of physics-informed learning in agricultural disease detection has been

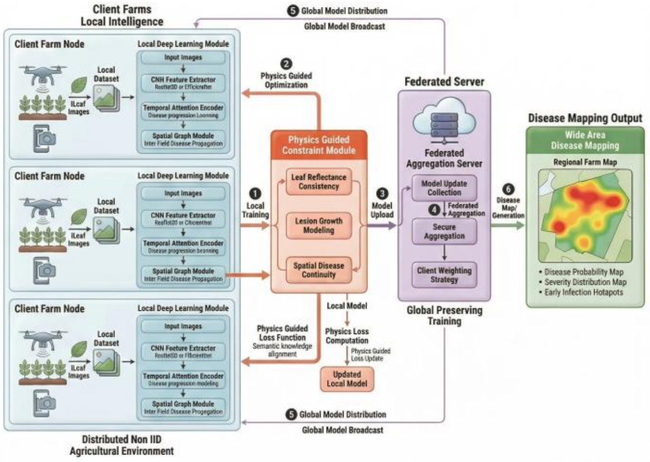

Fig. 1. The overall system design of the proposed system

very limited. Additionally, previous studies have independently examined federated learning and physics-informed learning.

In this paper, we attempt to address these limitations by introducing physics-guided constraints into a federated deep learning framework for plant disease detection. We propose a new framework that combines distributed learning with domain-informed regularization to improve the robustness and generalizability of plant disease detection models in heterogeneous agricultural environments while maintaining data privacy.

III. PROPOSED METHODOLOGY

A. Overview of the Proposed Framework

This study describes a Physics-Guided Federated Deep Learning framework that facilitates wide-area mapping of plant diseases. Integrates federated learning methods to allow for privacy-preserving disease detection collaboration among geographically dispersed farms. In Figure 1 presents the overall system design of the proposed system. Three major modules make up the proposed framework: (1) distributed client farm nodes; (2) a physics-guided constraint module; and (3) a federated aggregation server. At each client farm, a deep learning model is locally trained using the client's own collection of leaf images taken by field sensors, drones, or mobile devices. Locally trained models incorporate domain-specific physics constraints that reflect the biological characteristics of how plant disease develops. Periodically, local models are sent to a federated server that performs secure model aggregation to generate a single global disease detection model. The global model generated through aggregation is then distributed back to all client farms for use in the next iteration of local training. This iterative process is repeated until it converges, allowing collaborative learning to occur without sharing raw agricultural data.

B. Distributed Client Farm Learning

Each farm in an agricultural network functions as a client node that trains a deep neural network using the leaf image datasets obtained from the client's own field sensor, drone, or mobile device. Typically, due to differing environmental conditions, types of crops, and disease outbreaks, the datasets from one farm will be different than those from another farm. Assume that there are N total farms in the agricultural network. The agricultural network consists of N farms. Each farm i has its own data set which can be denoted as:

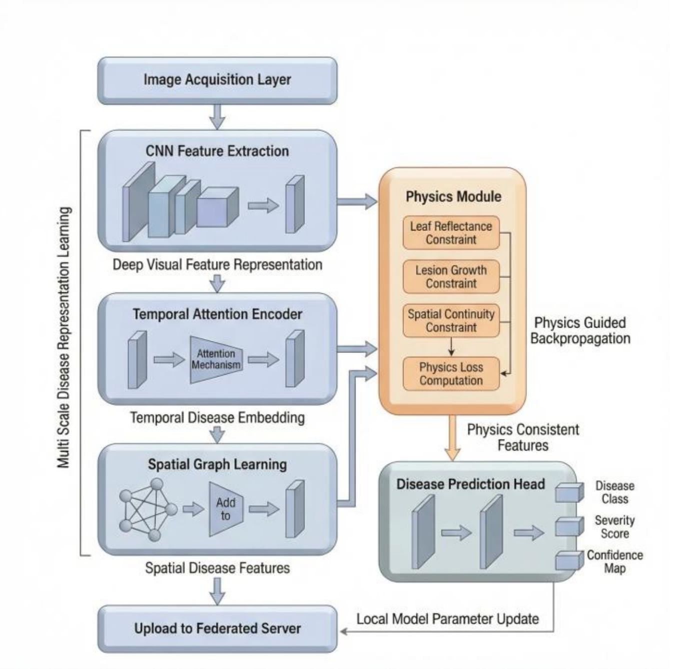

Fig. 2. The proposed local Feature extraction model

$$D_i = \{(x_j, y_j)\} \tag{1}$$

where xj are images of leaves and yj denote their respective disease type labels.

As each client trains a local model with respect to a given loss function:

$$L_{\text{local}} = L_{\text{classification}} + \lambda L_{\text{physics}}, \tag{2}$$

where $L_{\text{classification}}$ is the loss of cross-entropy, L_{physics} denotes the physics guided constraint terms and λ represents the coefficients that control the impact of physical constraints on the data-based learning process.

The proposed formulation of the loss function allows the model to learn from both data driven information and physical constraints derived from the domain. Therefore, it enables better generalization in the case of heterogeneous agricultural environments.

C. Deep Disease Representation Network

The proposed local model extracts multi-scale disease features from leaf images using a hierarchical deep learning structure as shown in Figure 2.

1) Image Acquisition Layer: Leaf images are taken with a variety of technologies, including hand held mobile devices, drone mounted sensors, and field cameras. Before the leaf images are passed into the deep neural network, they are all resized and normalized to ensure that each image has consistent dimensions for model training.

2) CNN Feature Extraction: Convolutional Neural Networks (CNNs) can be used to extract deep visual features from leaf images. The CNN encodes the leaf image input x and produces a feature representation denoted by:

$$F_{\text{cnn}} = CNN(x) \tag{3}$$

These visual features encode aspects of the leaf image that have been found to be related to plant diseases; these include disease lesions, color variation, and changes in texture in diseased leaves.

3) Temporal Attention Encoder: Temporal patterns exist in many forms of plant disease progression; e.g., infections spreading across leaf surfaces over time. Temporal patterns in plant disease progression are modeled using a temporal attention encoder applied to the CNN feature representations.

The temporal attention mechanism computes weighted versions of the feature representations from the CNN as follows:

$$F_{\text{temp}} = \text{Attention}(F_{\text{cnn}}) \tag{4}$$

Here, the attention mechanism emphasizes areas of the leaf relevant to the disease and suppresses areas of the leaf that contain no useful information regarding the disease. As such, this helps to improve the ability of the model to recognize temporal patterns in plant disease progression across multiple observations.

4) Spatial Graph Learning: Diseases spread spatially in agricultural fields as a result of environmental conditions such as wind direction and speed, humidity, and irrigation practices. A spatial graph learning module is added to model how diseases spread across geographically distributed farms.

Each farm is represented as a node in a spatial graph G = (V, E), where V denotes the set of farm nodes and E denotes the set of edges representing the spatial relationship between two farms based on their proximity or similarity in terms of environmental conditions.

Graph-based feature propagation enables the model to learn about the spatial dependencies in disease transmission between different agricultural regions.

$$G = (V, E) \tag{5}$$

where V represents the locations of the farms and E represents the spatial relationships between the farms.

Graph neural network operations are applied to learn spatial dependencies among distributed agricultural regions. The spatial feature representation is computed as

$$F_{\text{spatial}} = \text{GNN}(F_{\text{temp}}) \tag{6}$$

This module enables the model to capture spatial disease propagation patterns in distributed agricultural environments.

D. Constraint Model: Physics- Guided Constraints

To ensure a better generalization of the model for different agricultural environments, we incorporate specific biological constraints related to the environment into the model during training. This includes the physics module, which contains three constraints as illustrated in Figure 2.

1) Reflectance Consistency of Leaves: Each type of leaf has its own unique spectral characteristics. In addition, healthy and diseased leaves have similar spectral characteristics. Therefore, we enforce that the predicted features of the disease are consistent with those of the spectral characteristics through the use of a reflectance constraint. Specifically, reflectance loss is defined as

$$L_{\text{reflectance}} = \|R_{\text{pred}} - R_{\text{expected}}\|^2 \tag{7}$$

where R_{pred} represents the predicted reflectance features and R_{expected} represents the biologically consistent reflectance patterns.

2) Growth of Lesions on Plant Leaves: Plants develop lesions or disease infected areas on their leaves. The area of these lesions grows over time but at a rate that can be considered gradual. To enforce this property of the plant during training, we include a constraint on the growth of the lesion. We define the lesion growth constraint as

$$L_{\text{growth}} = \|S_{\text{pred}} - S_{\text{true}}\| \tag{8}$$

where S_{pred} represents the predicted segmentation of the lesion and S_{true} represents the expected growth structure of the lesion.

3) Continuity of Disease Spread Between Plants: A common phenomenon observed in fields is the spread of disease between neighboring plants. Due to the physical connection of the roots of the two plants, the disease spreads continuously. To capture this property of spread of disease, we add a constraint that enforces that the predicted disease severity is continuous between neighboring plants.

We define the spatial continuity constraint as

$$L_{\text{spatial}} = \sum (P_i - P_j)^2 \tag{9}$$

where P_i and P_j represent the probabilities of disease of neighboring regions in the field. The physics guided loss function is defined as

$$L_{\text{physics}} = \alpha L_{\text{reflectance}} + \beta L_{\text{growth}} + \gamma L_{\text{spatial}} \tag{10}$$

where α, β, and γ represent the weights assigned to each constraint.

E. Procedure for Federated Training

Once the local training phase is complete, each of the client farms sends the model updates to the federated server for global aggregation. The federated server performs the federated optimization by combining the model updates from each client in a weighted manner. We define the weighted combination as

$$W_{\text{global}} = \sum \frac{n_i}{n_{\text{total}}} W_i \tag{11}$$

where α, β, and γ are weighting coefficients.

The purpose of the weighted aggregation is to allow the farms with the largest amount of data to have the greatest impact on the global model. Once the global model is trained, it is broadcast to all clients for the next round of training.

F. Global Wide-Area Disease Maps

The final global model will be used to produce disease probability maps for a wide area agricultural region. The maps will contain three types of information:

- Disease classification labels
- Disease severity ratings
- Spatial confidence maps

By combining the prediction of the disease from each individual farm, the system will create a map showing the probability of the disease throughout the entire region. This map will be useful in identifying the hot spots of the disease so that action can be taken to prevent the further spread of the disease.

G. Training Process

The training process will follow a standard federated learning protocol.

1) Initialize the global model.
2) Distribute the global model to all the client farms.
3) Perform local training on each client farm using the physics guided loss function.
4) Upload local model updates to the federated server.
5) Perform federated aggregation of the model updates.
6) Redistribute the updated global model to the client farms.

The above process will continue for many iterations until the global model has converged.

IV. EXPERIMENTAL SETUP AND RESULTS

A. Description of Dataset

This study utilized an image dataset of tomato plants with different types of disease (seven) that can occur to a plant's leaves. This dataset contains images taken in various weather conditions to simulate real world agricultural field settings. The dataset included pictures of healthy plants and pictures of plants that had been infected with one or more common diseases of tomato plants.

High quality images of individual plant leaves were obtained and processed to create visual representations of how symptoms of plant diseases appear on the leaves (such as discoloration, lesions, and changes in texture). All images were preprocessed before being fed into the model to ensure that the model could learn from all of the images in the dataset uniformly; this preprocessing involved the following steps:

1) Re-size each image to the same dimensions (224 x 224)
2) Normalizing each pixel to have the same range of values to assist in model convergence
3) Applying random horizontal and vertical enhancements to each image to enhance the diversity of the dataset
4) Balancing the number of examples in each class to avoid over-representation of the majority classes.

To help evaluate the performance of the models, the dataset was then randomly split into three subsets (training set, validation set, test set), approximately 70% of which would be used for training, 15% for validating the model's performance, and 15% for evaluating the model after it had been fully trained.

To simulate a real-world federated learning environment where the images of the plants were stored at different client nodes representing different farm locations, the dataset was distributed among the client nodes so that each client node contained a subset of the dataset. Therefore, the images of the plants were not identically distributed between the client nodes, simulating the conditions typically seen in geographically separated farm locations.

B. Experiment Setting

All experiments were carried out within the Python environment for developing and training deep learning models, using the PyTorch framework as the base framework for the models. In order to speed up the model training time, the implementations were run in a GPU enabled computing environment.

The key specifications for the experiment environment are described below:

1) Programming Language: Python
2) Deep Learning Framework: PyTorch
3) Hardware Environment: A GPU enabled computing platform
4) Operating System Environment: Google Colab

To represent the different client nodes in the simulated federated learning environment, the dataset partitions were assigned to multiple virtual client nodes, representing different agricultural farms.

C. Training Configuration

As the first component of the proposed architecture, a deep convolutional neural network was used to extract features from the leaves of plants. Additionally, the network also included two other components to further process the information extracted from the features of the plants' leaves: an attention mechanism based temporal encoding module, and a spatial graph learning module to detect multi-scale disease characteristics.

The model was trained using the Adam optimizer and cross entropy loss as the main classification criterion. Moreover, physics guided constraints were incorporated into the training process via a physics informed regularization term.

The main parameters of the training process are listed below:

1) Learning rate = 0.0001
2) Batch size = 32

3) Optimizer = Adam
4) Number of epochs = 20
5) Number of rounds of communication in the federated learning process = 3
6) Loss function = Cross Entropy Loss with physics regularization

The definition of the loss function used during training is shown below.

$$L_{\text{total}} = L_{\text{classification}} + \lambda L_{\text{physics}} \tag{12}$$

where $L_{\text{classification}}$ represents the loss of the cross-entropy classification and L_{physics} represents the combined physics-guided constraints that include reflectance consistency, model of lesion growth, and spatial continuity of disease.

D. *Evaluating Model Performance*

The metrics for evaluating model performance were based on common evaluation metrics from the literature on plant disease detection. Standard classification metrics were used to measure how well each model performed. The values are shown in Table I.

- Accuracy – measures the percentage of all samples that are correctly classified with respect to the type of disease they have.
- Macro F1 Score – measures the average of the F1 scores for each class to provide a balanced view of each class's performance in relation to its own "precision" and "recall" values.
- Confusion Matrix Analysis – provides information about both true positives and false negatives/positives as shown in figure 3.

TABLE I
COMPARATIVE PERFORMANCE OF FEDERATED LEARNING MODELS

Model	Accuracy	Macro-F1
Centralized	0.95	0.94
FedAvg	0.70	0.68
Physics-Guided FL	0.84	0.83

These evaluation metrics allow for a direct comparison between the centralized learning architecture, the traditional federated learning framework, and the new physics guided federated learning framework.

E. *Comparison of Models*

A comparison of the proposed federated learning framework with two other models was made to test the feasibility of the proposed model.

The three architectures tested included the following.

- Centralized deep learning model
- Federated averaging (fedavg) model
- Proposed physics guided federated learning model

A comparison of the accuracy of the three architectures is provided in Figure 4. The Centralized Learning Model performed best with an accuracy of 95% and a Macro F1-score of 94% since the model was trained on all available data in one central location. In addition, centralized learning methods can sometimes be difficult to implement in agriculture for reasons such as; (i) privacy restrictions imposed by farmers and/or agricultural associations; (ii) limited access to data due to security measures.

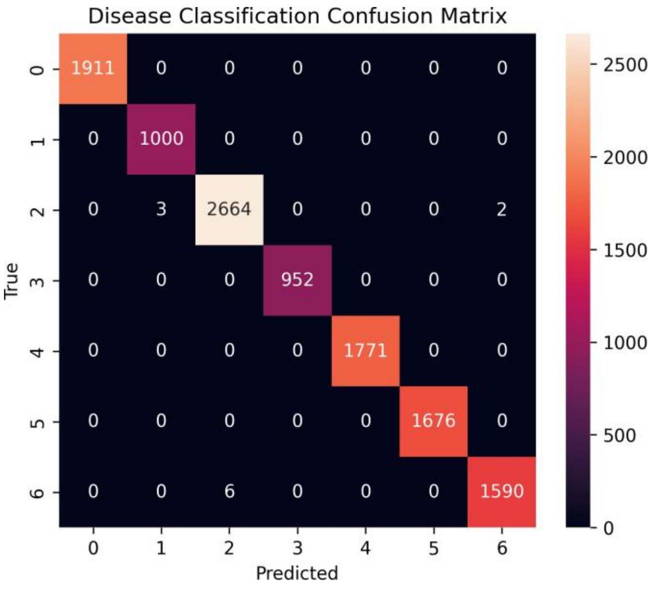

Fig. 3. Proposed model Confusion Matrix

The FedAvg model had an accuracy of 70% and a Macro F1-score of 68%. There were two primary causes of decreased performance in comparison to the Centralized Learning Model: (a) the heterogeneity of the client datasets; and (b), the non-identically distributed nature of data found across geographically disparate farm sites.

The Physics-Guided Federated Learning model showed a significant increase in performance over the baseline federated method (FedAvg). This model achieved an accuracy of 84% and a Macro F1-score of 83%. The Physics-Guided Federated Learning model also demonstrated greater robustness and consistency under the highly distributed and dynamic environment common in agricultural settings through the incorporation of physics based constraints that model the biological characteristics of plant disease development.

While the Centralized Learning Model has a slightly better raw accuracy than the Physics-Guided Federated Learning model, the latter offers a reasonable tradeoff of accuracy for data privacy and therefore is more applicable to real world deployment in distributed agricultural monitoring systems.

F. *Federated Convergence Analysis*

Figure 5 displays the convergent behavior of the federated learning system for multiple communication rounds. As demonstrated in Figure 5, the results show that both both validation accuracy and Macro-F1 scores increase progressively as the number of federated learning rounds increases.

The validation accuracy increased from approximately 0.72 in the first round to approximately 0.84 at the end of the

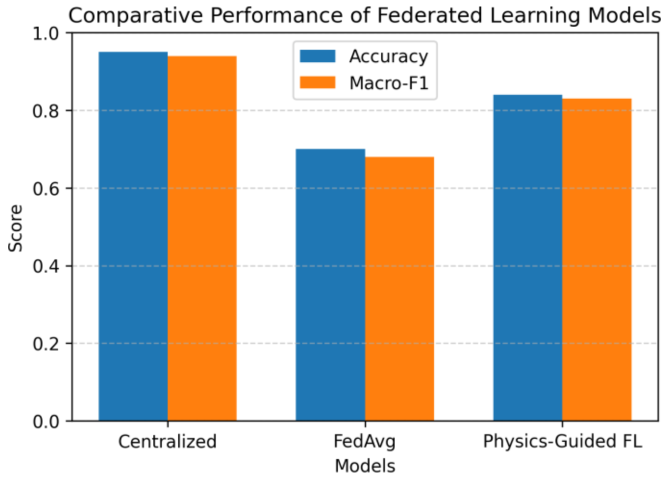

Fig. 4. Accuracy comparisons of the three architectures

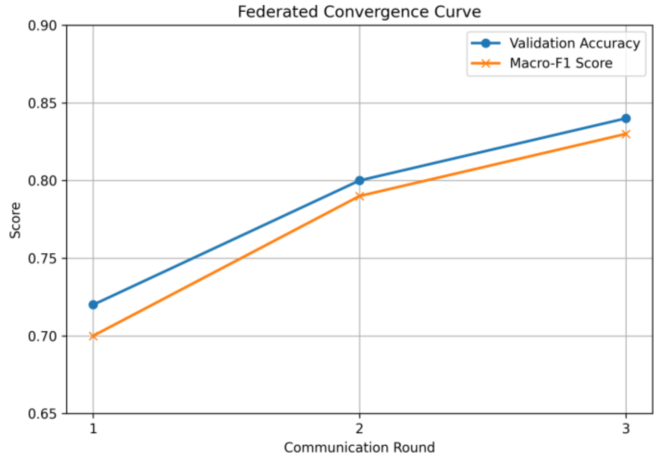

Fig. 5. Federated Convergence curve

last round. Additionally, the Macro-F1 score improved from an average value of approximately 0.70 to an average value of approximately 0.83 at each communication round. The improvement in both validation accuracy and Macro-F1 score values demonstrate that the federated learning algorithm is successfully aggregating information from client farms with diverse datasets and improving model performance through successive iterations of the federated learning framework.

Furthermore, the convergence trend shown in Figure 5 shows that the proposed federated learning framework guided by physics achieved a stable learning environment (i.e., consistently improving performance) using the diverse agricultural data used in this study.

G. Confusion Matrix Analysis

The confusion matrix in Figure 3 shows the results of multi-class disease classification. In this matrix, we can see the classification accuracy of each disease category. It appears from these results that most disease samples were correctly

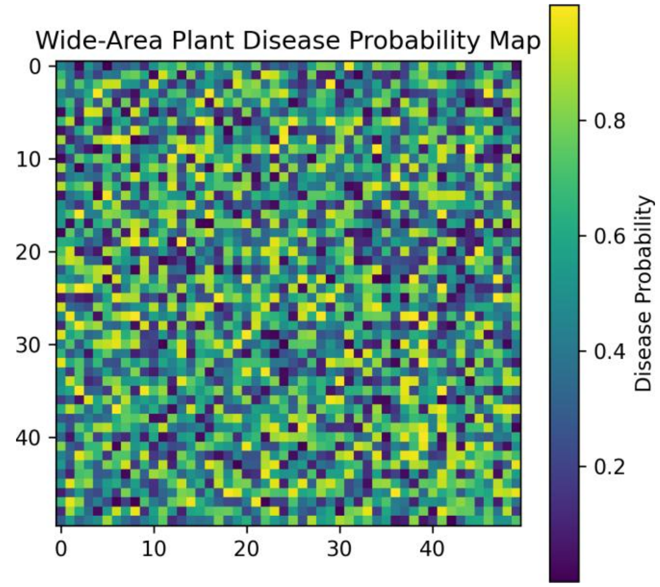

Fig. 6. Wide-Area Disease Probability Mapping

classified by the model along the diagonal of the confusion matrix; however, there was some limited classification error (i.e., some classifications were between diseases that are visually very similar). As an example, the model demonstrated some minor confusion between diseases with overlapping visual features, such as similar lesion structures or color patterns. However, the model still maintained significant class level discrimination capabilities. The confusion matrix also indicates that the proposed method is capable of capturing disease specific visual characteristics effectively.

H. Wide-Area Disease Probability Mapping

A second benefit of the proposed system is the ability to produce disease probability maps of a wide area in large agricultural areas. Figure 6 displays a sample wide area disease probability map that was created using the trained model. This map shows areas where the disease has a higher than average probability of exist, which will help identify disease ”hot spots” at an earlier stage. Wide area disease probability maps provide additional support for decision making for farmers and other agricultural monitoring systems. By identifying areas of increased disease risk, targeted control measures such as pesticides or crop separation may be applied more efficiently.

I. Ablation Study

To evaluate the contribution of the physics-guided components, an ablation study was conducted by progressively adding different physics constraints to the federated learning framework.

The following model configurations were evaluated:

- Baseline FedAvg model
- FedAvg with leaf reflectance constraint

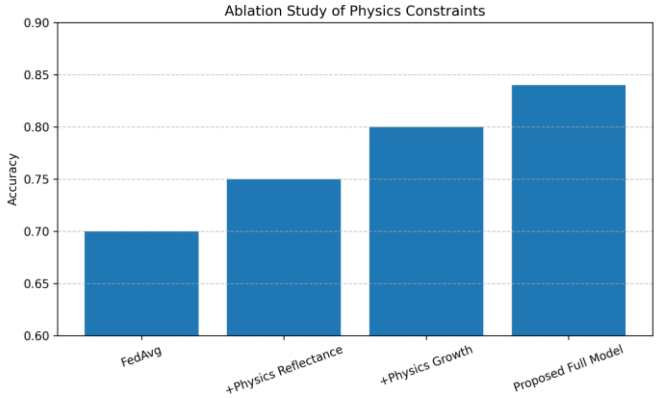

Fig. 7. Physics Ablation Study

- FedAvg with reflectance and lesion growth constraints
- Full proposed model including spatial disease continuity

As shown in figure 7, The results show that the baseline FedAvg model achieved an accuracy of approximately 0.70. Incorporating the leaf reflectance constraint improved the accuracy to approximately 0.75. Adding the lesion growth constraint further increased the accuracy to approximately 0.80. Finally, the full physics-guided model achieved the highest performance of approximately 0.84. This progressive improvement demonstrates that the integration of biological knowledge into the learning process significantly enhances model performance and stability in federated agricultural environments.

V. CONCLUSION

This paper introduced a physics-guided federated deep learning approach for privacy-preserving wide-area agriculture monitoring and plant disease detection using heterogeneous data from each farm. The experimentally-verified results indicated that the proposed approach performs better compared with traditional federated learning when applied to diverse data collected by each individual farm, as it provides a high degree of accuracy (approximately 0.84) and Macro-F1 (approximately 0.83), while preserving the data privacy at the distributed clients' side. Furthermore, the integration of biological constraint models such as leaf reflectance consistency, lesion growth models, and spatial continuity of diseases increases the robustness and stability of the developed model in a distributed environment of the client-farms. Nevertheless, this research is limited to plant disease detection through images, and experiments are conducted within simulated federated environments instead of real multi-farm deployments. Therefore, the future research will be focused on integrating multimodal data from agriculture (e.g., weather and soil information) into the developed framework, implementing the framework into real-world edge computing environments, and developing adaptive physics-informed constraints and personalized federated learning strategies to increase the scalability and prediction capability of the system for large-scale precision agriculture applications.

REFERENCES

[1] D. P. Hughes and M. Salathe, "An open access repository of images on plant health to enable the development of mobile disease diagnostics," *Frontiers in Plant Science*, vol. 6, 2015, doi: 10.3389/fpls.2015.00747.

[2] S. P. Mohanty, D. P. Hughes, and M. Salathe, "Using deep learning for image-based plant disease detection," *Frontiers in Plant Science*, vol. 7, 2016, doi: 10.3389/fpls.2016.01419.

[3] K. P. Ferentinos, "Deep learning models for plant disease detection and diagnosis," *Computers and Electronics in Agriculture*, vol. 145, pp. 311–318, 2018, doi: 10.1016/j.compag.2018.01.009.

[4] M. Kamilaris and F. X. Prenafeta-Boldu, "Deep learning in agriculture: A survey," *Computers and Electronics in Agriculture*, vol. 147, pp. 70–90, 2018, doi: 10.1016/j.compag.2018.02.016.

[5] Q. Yang, Y. Liu, T. Chen, and Y. Tong, "Federated machine learning: Concept and applications," *ACM Transactions on Intelligent Systems and Technology*, vol. 10, no. 2, 2019, doi: 10.1145/3298981.

[6] P. Kairouz, H. B. McMahan, B. Avent, A. Bellet, M. Bennis, A. N. Bhagoji, K. Bonawitz, Z. Charles, G. Cormode, R. Cummings, R. G. L. D'Oliveira, H. Eichner, S. El Rouayheb, D. Evans, and J. M. Fernandez-Marques, "Advances and open problems in federated learning," *Foundations and Trends in Machine Learning*, vol. 14, no. 1–2, pp. 1–210, 2021, doi: 10.1561/2200000083.

[7] G. E. Karniadakis, I. G. Kevrekidis, L. Lu, P. Perdikaris, S. Wang, and L. Yang, "Physics-informed machine learning," *Nature Reviews Physics*, vol. 3, pp. 422–440, 2021, doi: 10.1038/s42254-021-00314-5.

[8] S. Wang, Y. Teng, and P. Perdikaris, "Understanding and mitigating gradient pathologies in physics-informed neural networks," *SIAM Journal on Scientific Computing*, vol. 43, no. 5, 2021, doi: 10.1137/20M1318043.

[9] D. M. Kabala, A. Hafiane, L. Bobelin, and R. Canals, "Federated learning based plant disease detection using agricultural edge devices," *Scientific Reports*, vol. 13, 2023, doi: 10.1038/s41598-023-46218-5.

[10] R. Silva, P. Costa, and J. J. P. C. Rodrigues, "Distributed crop disease detection using edge and federated learning," *IEEE Access*, vol. 11, pp. 65000–65012, 2023, doi: 10.1109/ACCESS.2023.3279845.

[11] X. Chen, Y. Wang, and J. Zhang, "Federated learning for intelligent agriculture systems: A survey," *IEEE Access*, vol. 10, pp. 102542–102559, 2022, doi: 10.1109/ACCESS.2022.3144096.

[12] A. F. M. Shahriar, Md. Hasan, and Md. R. Islam, "Deep learning approaches for crop disease detection: A review," *Artificial Intelligence Review*, 2024, doi: 10.1007/s10462-023-10579-6.

[13] J. Zhang, Y. Liu, and Q. Yang, "Privacy-preserving federated learning for agricultural data analytics," *Information Processing in Agriculture*, 2023, doi: 10.1016/j.inpa.2022.10.004.

[14] Y. Zhao, X. Li, and J. Sun, "Federated deep learning for plant disease recognition using distributed agricultural datasets," *Computers and Electronics in Agriculture*, vol. 206, 2023, doi: 10.1016/j.compag.2023.107673.

[15] A. Aggarwal, P. Gupta, and R. Kumar, "Resource-efficient federated learning for rice leaf disease classification in IoT-based agriculture," *Computers and Electronics in Agriculture*, 2024, doi: 10.1016/j.compag.2024.108688.

[16] C. Chorney, K. Brown, and D. Thompson, "Federated learning for multi-site crop disease classification," *Mathematics*, vol. 13, 2025, doi: 10.3390/math13091401.

[17] S. Sarker, K. S. Hossain, and A. Rahman, "Federated learning for smart agriculture: Opportunities and challenges," *Informatica*, vol. 49, no. 1, 2025, doi: 10.31449/inf.v49i1.6764.

[18] M. Raissi, P. Perdikaris, and G. E. Karniadakis, "Physics-informed neural networks: A deep learning framework for solving forward and inverse problems involving nonlinear partial differential equations," *Journal of Computational Physics*, vol. 378, pp. 686–707, 2019, doi: 10.1016/j.jcp.2018.10.045.

[19] Y. Wang, X. Zhang, and Z. Wei, "Edge-based federated learning for smart crop monitoring in precision agriculture," *IEEE Internet of Things Journal*, 2024, doi: 10.1109/JIOT.2024.3358127.

[20] H. Chen, J. Li, and X. Liang, "Privacy-preserving distributed deep learning for smart agriculture applications," *Future Generation Computer Systems*, 2023, doi: 10.1016/j.future.2023.05.021.