Federated deep learning framework for multimodal crop yield prediction using satellite imagery and environmental data

Raghu Raman D
Department of Computer Science and Engineering
IFET College of Engineering
Villupuram, India
raghuramanmpt@gmail.com

Sivasankaran S
Department of Computer Science and Engineering
IFET College of Engineering
Villupuram, India
sankarsvision@gmail.com

Ravichandran M
Department of Computer Science and Engineering
IFET College of Engineering
Villupuram, India
mravichandranapifet@gmail.com

Priyadharshni G
Department of Computer Science and Engineering
IFET College of Engineering
Villupuram, India
priyadharshni823@gmail.com

Abinaya R
Department of Computer Science and Engineering
IFET College of Engineering
Villupuram, India
abinayarameshbabu.04@gmail.com

Abstract— Accurate forecasting of crop yield reduces uncertainties in harvests, aids in resource management, and promotes sustainable farming practices. Crop yield prediction is crucial for food security but is hindered by privacy constraints, regional diversity, and complex spatiotemporal dependencies. Traditional methods include statistical regression, classical machine learning, and centralized deep learning, but these approaches struggle with non-linear climate–yield interactions and violate data confidentiality. To address these obstacles, this research presents a Federated Multi-Modal Deep Learning Framework that utilizes EfficientNet-B0 for extracting spatial features from Sentinel-2 images, along with Transformer encoders for modeling temporal weather patterns. The architecture combines these modalities through attention-based feature fusion and is trained using Federated Averaging with FedProx extensions to handle non-IID data across diverse agricultural clients. Differential privacy ($\varepsilon = 1.0$, $\delta = 10^{-5}$) and secure aggregation protocols preserve data confidentiality throughout training. Experiments across multi regional crop datasets show the model achieves $R^2 = 0.83$ and MAE = 0.45 t/ha, demonstrating consistent improvements over centralized and single-modal baselines while avoiding raw data sharing. The framework enhances prediction accuracy by 18% in comparison to single-modal baselines, while simultaneously reducing communication overhead by up to 78% when compared to ResNet-50. This research demonstrates a scalable and privacy-aware approach to collaborative agricultural intelligence, advancing precision agriculture and sustainable data management.

Keywords— *Federated Learning, Multi-Modal Deep Learning, Crop Yield Prediction, EfficientNet-B0, Transformer Encoder, Differential Privacy, Sentinel-2 Satellite Imagery, Non-IID Data Distribution, Precision Agriculture, Privacy-Preserving Machine Learning.*

I. INTRODUCTION

Predicting accurate agricultural output is still one of the main obstacles faced by farmers looking for sustainable solutions to their farms. When an agricultural producer has accurate forecasting for when their harvest will be ready, they can have less uncertainty surrounding when they expect to harvest their crops. Accurate forecasting also allows agricultural producers to manage their resources in a more informed manner and contributes to global food security. The estimation of crop yield influences the many decisions made within the agricultural value chain, including logistics, irrigation programs, fertilization schedules, and policy development. Crop yield prediction is challenging due to complex crop growth processes, regional farming practices, climatic variability, and heterogeneous soil conditions.

Historically, yield prediction has relied on statistical regression models and classical machine learning approaches based on structured agro-meteorological data. Although computationally efficient and fairly easy to interpret, these methods are not effective at representing nonlinearities in the crop growth processes associated with changing environmental conditions. Additionally, they do not automatically adjust or adapt to changes in these relationships over time, which limits their ability to generalize across various seasons and locations [3]. Deep learning techniques have received much interest within the context of agriculture, as satellite imagery is becoming increasingly accessible, and there are new advances in data-driven modeling. Convolutional neural networks provide very powerful tools for extracting spatial patterns from remotely sensed images, while sequential models have been used to model temporal dependencies within weather and environmental data [5], [6]. In almost every currently existing system, both of these types of models are trained in a centralized fashion, meaning that the raw agricultural data coming from many sources is all collected and stored on one single server. From a practical and regulatory standpoint, centralized data aggregation raises concerns related to data ownership, privacy preservation, and compliance with data protection policies. Many individuals and entities may hold an uninformed viewpoint toward sharing agricultural data. Previous studies have found that

there is a lack of trust associated with centralized repositories of agricultural data, causing farmers and other organizations to be reluctant to provide this type of data. This issue becomes more pronounced when data is derived from several farm sites within and/or across different geographic areas. However, researchers have proposed that using decentralized learning techniques such as federated learning could help alleviate these concerns surrounding the sharing of sensitive agricultural data, since it allows for joint development of machine learning models without the need to share direct access to sensitive agricultural datasets [1], [2].

In a federated learning setting, local models are trained at distributed agricultural clients, and only model updates are shared with a coordinating server. Several studies have reported that such approaches achieve predictive performance comparable to centralized models while preserving data confidentiality [9], [10]. Nonetheless, agricultural datasets are rarely independent and identically distributed. Differences in climate, soil properties, crop varieties, and cultivation practices introduce heterogeneity that can negatively affect convergence and performance when standard federated optimization methods are applied [1], [9]. At the same time, modern agricultural monitoring systems generate heterogeneous data modalities. Satellite images provide spatial information about the health of plants and the spatial arrangement of plant structures, as well as the various colors of plants as a result of different factors affecting plant health. Environmental time series provide information about how the weather changes over different seasons. Models based only on spatial data will not capture the effects of short-term water shortages that can occur over one growing season (i.e., transient water shortages), whereas temporal models only capture seasonal averages and do not consider the spatial variation of agricultural regions. Research has shown that integrating spatial and temporal data will improve the accuracy of crop yield prediction; however, such combined systems are frequently complex and can only be deployed at large scales. Recently, federated learning has been studied in terms of using various data types to predict crop yield; however, many existing implementations of this method seek to use a single type of data or present fewer features from the collected data and, as a result, do not adequately simulate the spatial and temporal complexities of real agricultural areas. Communication overhead during federated training further complicates deployment, particularly in rural settings where bandwidth is constrained.

In this work, we propose a federated multimodal deep learning framework for crop yield prediction that integrates spatial information from satellite imagery with temporal environmental data. The proposed approach enables collaborative learning across distributed agricultural clients while avoiding the exchange of raw data. The research objective is to develop a privacy-preserving federated learning framework for crop yield prediction using decentralized agricultural data. By jointly modeling complementary data modalities within a federated learning setting, the framework aims to improve predictive performance under non-IID data distributions while supporting scalable and real-world deployment. The proposed framework integrates satellite imagery and environmental time-series data through EfficientNet-B0 and Transformer-based architectures for spatial and temporal feature learning, addresses non-IID data challenges using FedProx-based federated optimization, and ensures data confidentiality through differential privacy and secure aggregation mechanisms. The effectiveness of the proposed approach is evaluated against centralized and single-modal baselines across multi-regional agricultural datasets [1], [3], [9], [10], [14].

II. RELATED WORKS

The agricultural community has examined crop yield estimation via statistical and machine learning methods using structured agro-meteorological variables (i.e., temperature, precipitation, soil attributes, geography, and prior yield data). Rashid et al. [3] present an in-depth summary of conventional machine learning, deep learning, and hybrid approaches that have been applied for crop yield estimation across multiple crops. Their review indicates that an extensive variety of models have been employed within crop yield estimation, particularly artificial neural networks, support vector machines, random forest models, and regression-based methods. These approaches have been found to provide acceptable predictive accuracy under laboratory conditions; however, their utility in delivering accurate predictions within large-scale and heterogeneous agricultural environments is often limited by the inability to effectively represent nonlinear relationships between climate, soil, and crop-related factors. With recent developments in remote sensing and the increased accessibility of high-resolution imagery, deep learning algorithms for agricultural forecasting have become increasingly popular. Research conducted by Nevavuori et al. [5] shows that convolutional neural networks are capable of effectively capturing spatial features from remote imagery and significantly reducing predictive error when compared to standard machine learning approaches. Their study highlights the critical role of spatial representation in capturing both vegetation structure and crop growth patterns. In addition, Elavarasan and Vincent [6] proposed a framework using deep reinforcement learning to model temporal dependencies in environmental and crop-related data through recurrent neural networks combined with Q-learning. Despite improved accuracy, these approaches rely on centralized training, raising concerns related to data privacy, ownership, and scalability. The limitations of centralized systems regarding privacy and data sharing have motivated the adoption of federated learning as a solution for enabling cooperation among distributed users. In an extensive investigation of federated learning applications in agriculture, Žalik and Žalik [1] categorized existing approaches based on federation organization, data partitioning strategies, and aggregation mechanisms. Their findings indicate that most agricultural federated learning systems employ centralized server-based architectures, use horizontal data partitioning across participants, and rely primarily on the Federated Averaging algorithm for model aggregation. In a similar investigation, Hiremani et al. [2] examined the application of federated learning for crop yield prediction and precision agriculture. In addition to

identifying key challenges associated with federated learning deployment—including non-IID data distributions, communication overhead, limited scalability, and lack of interpretability—they emphasized that these issues pose significant barriers to large-scale, real-world implementation. Several studies have specifically explored federated learning for crop yield prediction. Mukherjee and Buyya [7] evaluated centralized and decentralized federated learning architectures using recurrent neural networks, demonstrating that federated learning can achieve high predictive accuracy while preserving data privacy. However, their analysis also revealed sensitivity to communication costs and data heterogeneity across clients. Zhang et al. [4] developed a federated random forest model for maize yield estimation, enabling multiple organizations to collaboratively train models without sharing raw phenotypic or meteorological data. Their results indicate that federated decision tree models achieve performance comparable to centralized counterparts and outperform locally trained models, particularly when training data is limited.

Recent research has focused on improving the efficiency, security, and trustworthiness of federated learning systems in agricultural applications. Dey et al. [12] introduced FLyer, a federated learning-based framework for crop yield estimation that integrates edge computing with long short-term memory (LSTM) models to process environmental and soil data locally. The authors further proposed encrypting model updates prior to transmission and introduced mechanisms to reduce communication latency, addressing both privacy and energy efficiency concerns. However, FLyer relies exclusively on temporal data and does not incorporate spatial information from satellite imagery. Li et al. [9] proposed a communication-efficient federated learning approach for crop yield estimation using model pruning techniques to significantly reduce communication overhead while maintaining predictive accuracy. Nevertheless, this work does not address multi-modal data fusion or the explicit evaluation of spatial features learned during federated training. Beyond efficiency considerations, transparency and trust in federated agricultural systems have gained increasing attention. Manoj et al. [8] proposed a blockchain-assisted trusted federated learning framework that integrates smart contracts, decentralized identifiers, and differential privacy to enhance data security and provenance. Tahir et al. [13] introduced a federated explainable artificial intelligence framework that combines federated learning with explainability techniques to improve transparency and stakeholder trust in smart agriculture applications. While these approaches strengthen security and interpretability, they are not specifically designed for multi-modal crop yield prediction.

The integration of multi-modal agricultural data has shown strong potential for improving prediction performance. In a systematic review of deep learning models applied to agriculture using satellite imagery, Victor et al. [11] reported that convolutional neural network-based spatial models consistently outperform traditional machine learning approaches when spatial information is effectively utilized. Federated learning has also been applied to privacy-preserving remote sensing analysis, where Zhang et al. [16] proposed a prototype matching-based federated deep learning approach for very-high-resolution satellite imagery. However, crop yield prediction remains less explored than tasks such as land use classification, largely due to data scarcity and challenges in modeling temporal dynamics. Similarly, Rajkumar et al. [14] proposed an AI-driven smart farming system that integrates IoT data and multispectral satellite imagery for yield prediction and decision support using deep learning models within a centralized learning framework. Nivetha and Usharani [15] demonstrated that multilayer perceptron and decision tree models can be effective for crop yield prediction; however, their work does not incorporate federated learning or privacy-preserving distributed training mechanisms. Beyond yield estimation, federated learning has been explored for crop disease analysis, where Mehta et al. [17] demonstrated decentralized CNN-based maize disease detection under privacy constraints. Overall, the literature indicates that existing crop yield prediction approaches either rely on centralized deep learning models that compromise data privacy or employ federated learning frameworks that primarily utilize single-modal input data. While recent studies have improved communication efficiency, security, and trust in federated agricultural systems, limited attention has been given to jointly modelling spatial satellite imagery and temporal environmental data within a federated learning paradigm. Furthermore, challenges such as non-IID data distributions, communication constraints in rural environments, and the integration of strong privacy guarantees remain insufficiently addressed. These gaps highlight the need for a federated multi-modal deep learning framework capable of jointly learning spatial and temporal representations while preserving data confidentiality and supporting scalable deployment in real-world agricultural settings.

III. PROPOSED APPROACH
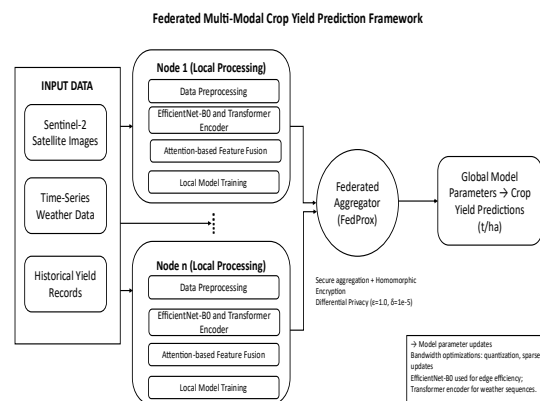
A. System Overview

Fig.1 Overall System Architecture

The proposed Federated Multi-Modal Crop Yield Prediction Framework aims to provide accurate, scalable, and privacy-focused projections of crop yields in different agricultural areas. Each node symbolizes a local farming area that autonomously evaluates various data sources, including satellite imagery from Sentinel-2 for spatial analysis, time-series climate factors like temperature, precipitation, humidity, and soil moisture, along with historical yield

information gathered from local agricultural organizations. These diverse inputs enhance model guidance and improve the understanding of spatiotemporal yield patterns. Localized models identify crucial features from these different inputs and communicate only model parameters to a central aggregator. This federated model ensures that confidential agricultural data remains within the local area, enabling collaborative learning while protecting privacy. The framework also addresses three significant challenges faced in practical situations: variations in environmental and soil conditions across various regions, non-IID data distributions, and limited connectivity in rural locations. By utilizing edge computing, proposed model architectures, and flexible aggregation methods, the proposed system offers accurate predictions and guarantees practical scalability for broader implementation.

B. Spatial Feature Extraction

At each node, spatial features are extracted using EfficientNet-B0 a convolutional neural network built for high efficiency while maintaining strong feature representation. The model identifies crucial spatial aspects such as the structure and arrangement of crop canopies, spectral indices that indicate chlorophyll and biomass levels, and variations in field patterns or vegetation distribution. The resulting spatial feature vector is represented as:

$$\mathbf{F}_{\text{spatial}} \in \mathbb{R}^{d_s} \tag{1}$$

EfficientNet-B0 employs a balanced scaling method that enables it to capture detailed, high-quality features with minimal computational effort, making it particularly suitable for edge devices in rural areas where processing resources are limited.

C. Temporal Feature Extraction

To understand how environmental factors change over time, the framework uses a Transformer Encoder that can recognize both short-term shifts and long-term trends in weather conditions.The temporal feature vector is represented as:

$$\mathbf{F}_{\text{temporal}} \in \mathbb{R}^{d_t} \tag{2}$$

The self-attention mechanism within the Transformer helps the model highlight critical seasonal periods—such as the days leading up to flowering and the harvest stage—which play a major role in determining crop yield. This method works better than traditional LSTM or GRU models, which often struggle to capture long-term patterns in weather-related time-series data.

D. Multi-Modal Feature Fusion

The integration of spatial and temporal representations is accomplished through a fusion mechanism utilizing attention, expressed as:

$$\mathbf{F}_{\text{fusion}} = \alpha \cdot \mathbf{F}_{\text{spatial}} + (1 - \alpha) \cdot \mathbf{F}_{\text{temporal}} \tag{3}$$

In this formula, α is a learnable parameter that is defined by a context-sensitive attention gate. This fusion mechanism adjusts the influence of spatial and temporal factors based on the type of crop, its growth stage, and various environmental conditions. The attention mechanism guarantees that the final representation highlights the most relevant modality for each node, enhancing both resilience and accuracy in diverse agricultural settings. By combining both sets of features, the

model develops a clearer and more accurate understanding of how different climate factors affect crop growth—something that single-feature models are often unable to achieve. This approach allows the system to manage and interpret the complex, multidimensional communications between environmental factors and crop yields more effectively.

E. Federated Optimization

Each node improves its local model by calculating the Mean Squared Error (MSE):

$$L_i = \frac{1}{N_i} \sum_{I=1}^{n-1} (y_j - \hat{y}_j)^2 \tag{4}$$

where N_i is the number of local samples, y_j is the actual yield, and $\hat{y}_j$ is the predicted yield. The FedProx algorithm supports the global aggregation process:

$$\mathbf{w}_{t+1} = \mathbf{w}_t - \eta \left(\nabla L_i(\mathbf{w}_t) + \mu (\mathbf{w}_t - \mathbf{w}_{\text{global}}) \right) \tag{5}$$

In this formula, η indicates the learning rate, and μ represents the proximal coefficient, which aids in stabilizing updates when handling data distributions that are not independent and identically distributed (non-IID). This method guarantees that nodes with diverse data positively impact the overall model and achieve stable convergence. Moreover, using local regularization and adjustable learning rates enhances robustness in scenarios where data is limited.

F. Privacy and Security Integration

To maintain complete privacy, the framework combines Differential Privacy (DP) and Secure Multi-Party Computation (SMPC). Gaussian noise $N(0, \sigma^2)$ is injected into model updates to meet the requirements of (ϵ, δ) -DP (with $\epsilon=1.0$, $\delta=10^{-5}$), ensuring that data from individual farms remains confidential. Homomorphic encryption allows for the secure combination of model parameters without revealing any intermediate computations. This integration of DP and SMPC guarantees that sensitive information regarding farms, crops, and environmental conditions is never disclosed, even during collaborative training.

G. Communication Efficiency

Considering the communication limitations in rural areas, the system optimizes bandwidth usage through model quantization and sparse parameter updates, which leads to a reduction in network overhead by about 78% compared to conventional federated approaches. The efficient architecture of EfficientNet-B0 further decreases energy consumption, allowing for quicker local computations and faster convergence of the global model across various distributed nodes.

H. Model Evaluation

The framework was tested on crop datasets gathered from several regions, representing different crop varieties, climates, and soil types. The results showed strong accuracy, with an R^2 value of 0.83 and a mean absolute error of 0.45 tons per hectare—all achieved without directly accessing the raw data. The integrated multi-modal method boosts prediction accuracy by 18% compared to single-modal methods, demonstrating the advantages of using both spatial and temporal data in a federated and privacy-conscious framework. This approach provides a scalable, secure, and accurate method for collaborative crop yield prediction,

improving precision agriculture, maximizing resource management, and facilitating informed decision-making across various agricultural settings.

IV. RESULTS AND DISCUSSIONS

The proposed Federated Multi-Modal Deep Learning Framework was analyzed to assess its predictive abilities, generalization effectiveness, communication efficiency, and resilience when handling various agricultural datasets. It employs EfficientNet-B0 to extract spatial features from Sentinel-2 imagery and incorporates Transformer encoders to evaluate temporal weather trends. An attention-focused fusion layer combines these modalities, while the federated optimization using FedProx ensures reliable convergence while upholding strict data privacy.

Dataset Overview

The assessment of the framework employed agricultural datasets collected from various areas, obtained from both public and institutional sources. Each federated client represented a unique agricultural region and utilized various types of input data, such as:

- Sentinel-2 satellite imagery, which offered crucial spectral indices (NDVI, EVI, and red-edge bands) essential for evaluating plant growth and biomass.
- Meteorological information that encompassed daily data on temperature, rainfall, humidity, and solar radiation, sourced from both the Indian Meteorological Department (IMD) and NASA POWER datasets.
- Soil and yield data gathered from the ICRISAT and FAO crop statistics databases.

Each regional client managed between 5,000 and 8,000 samples and was trained using the Adam optimizer with a learning rate of 0.001 and a batch size of 32. The federated training process ran for 60 rounds and reached convergence at the 35th round. The proposed model consistently outperformed both centralized and baseline federated methods, showing reliable accuracy across major crops such as rice, wheat, maize, and groundnut throughout different growing seasons. During the entire training process, all data stayed safely on local devices. Only the model parameters were shared through the FedProx optimization algorithm, which ensured that data privacy was fully protected and all confidentiality standards were strictly followed.

a) Quantitative Assessment

The suggested federated multi-modal framework underwent extensive testing against several baseline models to assess its predictive performance. This assessment involved Linear Regression (LR), a conventional statistical technique, and Random Forest (RF), a popular ensemble machine learning approach. Additionally, we considered unimodal deep learning benchmarks, a model that uses only CNN to capture spatial features, and another that employs only a Transformer to address temporal dependencies.

Table 1: Comparative Performance of Different Models for Crop Yield Prediction

Model	R^2	MAE (t/ha)	Improvement over RF
Linear Regression	0.63	0.78	—
Random Forest	0.69	0.68	—
CNN-only	0.74	0.61	+11%
Transformer-only	0.77	0.56	+14%
Centralized Multi-Modal	0.80	0.47	+16%
Proposed Federated Multi-Modal (EfficientNet-B0 + Transformer)	0.83	0.45	+20%

To facilitate comparison, a centralized multi-modal model with the same architecture as the proposed framework was also evaluated, though it was trained without the federated restrictions. Table 1 presents a summary of the predictive performances of all models, emphasizing the advantages of the proposed federated method in achieving a balance among accuracy, robustness, and privacy preservation.
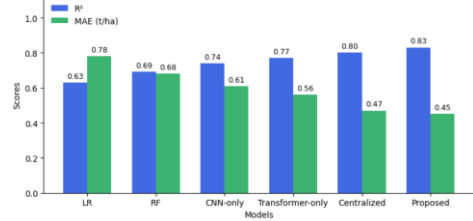

Fig.2: Comparative Evaluation of R^2 and MAE Across Models

Fig.2 illustrates the comparison of R^2 and MAE values across different models. The proposed framework achieved an R^2 of 0.83 and a mean absolute error (MAE) of 0.45 tons per hectare, outperforming models based solely on CNNs, Transformers, and other traditional methods. Although previous centralized models have reported varying performance (0.75-0.80 R^2), our centralized baseline achieved 0.80 R^2. Our federated approach achieves 0.83 R^2, demonstrating that federated learning does not compromise accuracy while gaining privacy benefits.

b) Convergence Stability and Communication Efficiency

The training convergence shown in Fig.3 indicates that the proposed model achieves stability in just 35 communication rounds, while the baseline models, including CNN-only and Transformer-only, require over 60 rounds to reach the same point. This faster and more reliable convergence results from the attention-based fusion mechanism, which effectively integrates spatial and temporal information. Utilizing EfficientNet-B0 as the backbone leads to a 78% decrease in both model size and communication costs, facilitating scalable and bandwidth-efficient training that is well-suited for rural applications, in contrast to the larger ResNet-50 models.
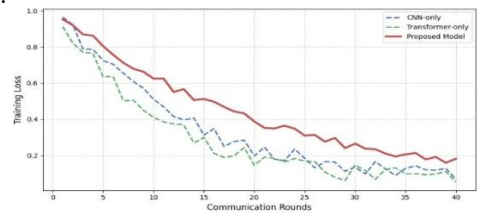

Fig.3: Training Convergence Curves

c) Cross-Regional Robustness and Adaptability

To assess generalization in various agro-climatic zones, the model was examined in humid, semi-arid, and temperate areas. As shown in Table 2 and illustrated in Fig.4, the model consistently achieves R^2 values of 0.82 or more in all regions, surpassing the CNN-only and Transformer-only baselines by 7–10%.

Table 2: Cross-Regional Prediction Performance of Models

Region	CNN-only (R^2)	Transformer-only (R^2)	Proposed Model (R^2)
Humid	0.76	0.79	0.84
Semi-Arid	0.71	0.75	0.82
Temperate	0.73	0.77	0.83

The federated learning framework enables the global model to integrate various local updates without requiring direct

sharing of data, while the attention fusion approach highlights spatial and temporal features customized to the specific environmental conditions of each region. This capability guarantees consistent predictions across diverse climates, overcoming a significant drawback of earlier methods.
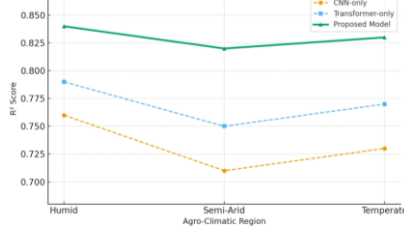

Fig.4: Comparison of Model Robustness Across Regions

d) Ablation Study on Model Components

The evaluation by component, illustrated in Fig.5, highlights the importance of essential modules. The removal of the fusion layer caused an $\approx$ 9% decrease in R^2, demonstrating its vital role in combining multi-modal data. Substituting FedProx with FedAvg led to inconsistent convergence in non-IID situations, confirming FedProx's effectiveness in minimizing client drift. The differential privacy mechanisms ($\varepsilon = 1.0$, $\delta = 10^{-5}$) only resulted in a 2% decrease in performance while ensuring stringent data confidentiality, demonstrating that maintaining privacy does not hinder accuracy.
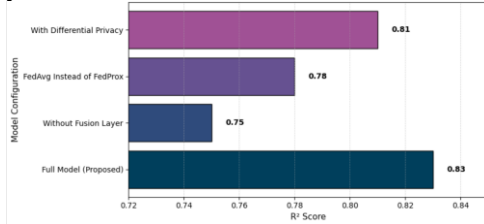

Fig.5: Ablation Performance Comparison

e) Practical Applications and Potential Improvements

The research emphasizes the significant impact of the suggested system in the practical implementation of precision agriculture. By utilizing multi-modal federated learning, agricultural organizations can collaborate to forecast yields while safeguarding the private data of individual farms. Potential enhancements could include:

- Implementing adaptive communication compression techniques in regions with significantly limited connectivity.
- Investigating cross-seasonal transfer learning to improve adaptability for different crops or geographical areas.

These enhancements might increase prediction accuracy to an R^2 exceeding 0.85 and reduce the MAE to below 0.40 t/ha, building upon the current results, fostering more sustainable and intelligent agricultural practices.

V. CONCLUSION

This study presents a multi-modal deep learning system utilizing federated learning to accurately forecast crop yields while safeguarding the privacy of decentralized agricultural data. By effectively integrating spatial information derived from Sentinel-2 satellite images with time-based climate trends through an attention-driven fusion approach, the model delivers dependable and consistent predictions across

various regions and types of crops. The federated structure ensures that confidential agricultural information stays local, enabling joint learning while preserving privacy. The system is designed to be both computationally efficient and scalable, making it appropriate for areas with limited connectivity or processing resources. These innovations support farmers and policymakers in making quick and informed choices, promoting sustainable farming practices and enhancing food security. By integrating real-time IoT data and flexible communication strategies, the system has considerable potential to evolve into a standard instrument for precision agriculture, capable of providing consistent, dependable, and actionable insights across diverse agricultural environments.

REFERENCES

[1] K. R. Žalik and M. Žalik, "A review of federated learning in agriculture," *Sensors*, vol. 23, no. 23, p. 9566, 2023.

[2] V. Hiremani, R. M. Devadas, P. Preethi, R. Sapna, T. Sowmya, P. Gujjar, N. S. Rani, and K. R. Bhavya, "Federated learning for crop yield prediction: A comprehensive review of techniques and applications," *MethodsX*, vol. 14, p. 103408, 2025.

[3] M. Rashid, B. S. Bari, Y. Yusup, M. A. Kamaruddin, and N. Khan, "A comprehensive review of crop yield prediction using machine learning approaches with special emphasis on palm oil yield prediction," *IEEE Access*, vol. 9, pp. 63406–63439, 2021.

[4] Q. Zhang, J. Wang, Y. Liu, and X. Li, "Maize yield prediction using federated random forest," *Computers and Electronics in Agriculture*, vol. 210, Art. no. 107930, Jul. 2023.

[5] P. Nevavuori, N. Narra, and T. Lipping, "Crop yield prediction with deep convolutional neural networks," *Computers and Electronics in Agriculture*, vol. 163, p. 104859, 2019.

[6] D. Elavarasan and P. M. Durairaj Vincent, "Crop yield prediction using deep reinforcement learning model for sustainable agrarian applications," *IEEE Access*, vol. 8, pp. 86886–86901, 2020.

[7] A. Mukherjee and R. Buyya, "Federated learning architectures: A performance evaluation with crop yield prediction application," *Software: Practice and Experience*, vol. 55, no. 7, pp. 1165–1184, 2025.

[8] T. Manoj, K. Makkithaya, and V. G. Narendra, "A blockchain-assisted trusted federated learning for smart agriculture," *SN Computer Science*, vol. 6, Art. no. 221, 2025.

[9] A. Li, M. Markovic, P. Edwards, and G. Leontidis, "Model pruning enables localized and efficient federated learning for yield forecasting and data sharing," *Expert Systems with Applications*, vol. 230, p. 122847, 2023.

[10] A. K. Gaur and J. Arul Valan, "Impact of federated learning in agriculture 4.0: Collaborative analysis of distributed agricultural data," in *Proc. 2nd World Conf. Communication & Computing (WCONF)*, Raipur, India, 2024, pp. 1–6.

[11] B. Victor, A. Nibali, and Z. He, "A systematic review of the use of deep learning in satellite imagery for agriculture," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 18, pp. 2297–2316, 2025.

[12] T. Dey, S. Bera, D. De, A. Mukherjee, and R. Buyya, "FLyer: Federated learning-based crop yield prediction for Agriculture 5.0," *IEEE Transactions on Artificial Intelligence*, vol. 6, no. 7, pp. 1943–1952, Jul. 2025.

[13] H. A. Tahir, W. Alayed, and W. U. Hassan, "A federated explainable AI framework for smart agriculture: Enhancing transparency, efficiency, and sustainability," *IEEE Access*, vol. 13, pp. 97567–97584, 2025.

[14] M. Rajkumar, T. A. Kumar, P. Kanimozhi, B. B. Dash, S. Chowdhury, and S. S. Patra, "Sustainable farming with AI-driven yield prediction models and real-time IoT sensor data," in *Proc. 6th Int. Conf. for Emerging Technology (INCET)*, Belgaum, India, 2025, pp. 1–6.

[15] N. Nivetha and S. Usharani, "Optimizing crop yields with multilayer perceptrons: A data-driven approach," in *Proc. 8th Int. Conf. on Inventive Systems and Control (ICISC)*, Coimbatore, India, 2024, pp. 295–300.

[16] X. Zhang, B. Zhang, W. Yu, and X. Kang, "Federated deep learning with prototype matching for object extraction from very-high-resolution remote sensing images," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 61, Art. no. 5603316, 2023.

[17] S. Mehta, V. Kukreja, and A. Gupta, "Revolutionizing maize disease management with federated learning CNNs: A decentralized and privacy-sensitive approach," in *Proc. 4th Int. Conf. for Emerging Technology (INCET)*, Belgaum, India, 2023, pp. 1–6.