

Federated Edge Learning for Distributed Weed Detection in Precision Agriculture Using Multimodal Sensor Fusion [†]

Dasaradha Arangii and Neelamadhab Padhy *

Department of Computer Science and Engineering, School of Engineering and Technology,
GIET University, Gunupur 765022, Odisha, India; dasaradha.arangii@giuet.edu

* Correspondence: dr.neelamadhab@giuet.edu

[†] Presented at the 12th International Electronic Conference on Sensors and Applications (ECSA-12), 12–14 November 2025; Available online: <https://sciforum.net/event/ECSA-12>.

Abstract

Background: Detection is an important component of precision agriculture because accurate weed identification and treatment have a direct impact on crop yield and resource efficiency. Recent breakthroughs in artificial intelligence (AI) have enabled automatic weed recognition systems; nevertheless, standard centralised machine learning models present substantial obstacles such as high communication overhead, privacy issues, and limited scalability in remote farming contexts. To address these restrictions, federated edge learning (FEL) combined with deep learning and multimodal sensor fusion provides a viable solution by allowing for distributed model training while maintaining data privacy. **Objective:** In this work, our goal is to develop a privacy-preserving distributed weed detection and management system. The proposed work is integrated with FEL (Federated Learning), Deep learning with multi-modal sensor fusion to enhance the performance of the model and simultaneously minimize the data transfer, latency, and energy consumption. **Materials and Methods:** In this study, we used Multimodal sensors, such as LiDAR (Light Detection and Ranging), RGB (Red-Green-Blue) cameras, multispectral imaging devices, and soil moisture sensors placed in controlled agricultural plots. Each modality gave complementary information for weed identification: RGB provided texture and colour cues, multispectral collected spectral reflectance patterns, LiDAR delivered structural depth information, and soil sensors supported contextual environmental conditions. For robustness, three sensor fusion techniques were used: Early Fusion (feature-level concatenation), Mid Fusion (intermediate feature aggregation), and Late Fusion (decision-level integration). Deep learning models, such as Convolutional Neural Networks (CNNs), LSTM-CNN hybrids, and Vision Transformers, were trained using standardised parameters. A proposed Federated CNN (FedCNN) was deployed across multiple edge devices, each locally trained on sensor data without exchanging raw data, using **FedAvg** and **FedProx** algorithms. Validation was performed using a stratified 80/20 train-test split combined with 5-fold cross-validation to ensure model generalization. Model performance was assessed using accuracy, precision, recall, F1-score, AUC, latency, and energy consumption, enabling a holistic evaluation of both predictive quality and computational efficiency. The DL models, including CNNs, LSTM-CNN hybrids, and Vision Transformers, are used. A FedCNN model is distributed across many edge nodes, allowing for decentralized training without exchanging raw data. For model performance measures, we used different metrics like accuracy, precision, recall, F1-score, AUC, latency, and energy. **Result:** The experimental work reveals that the model FedCNN performs well in comparison to other models and achieved the highest accuracy of 94.1%, precision is 94.3%, recall

Academic Editor(s): Name

Published: date

Citation: Arangii, D.; Padhy, N.
Federated Edge Learning for
Distributed Weed Detection in
Precision Agriculture Using
Multimodal Sensor Fusion. *Eng.
Proc.* **2025**, volume number, x.
<https://doi.org/10.3390/xxxxx>

Copyright: © 2025 by the authors.
Submitted for possible open access
publication under the terms and
conditions of the Creative Commons
Attribution (CC BY) license
(<https://creativecommons.org/licenses/by/4.0/>).

is 93.9% and F1-score is 94.1%, AUC is 94.1% during hybrid fusion strategies. We compared the centralized and federated learning performance. The FEL (Edge) accuracy is 94.1%, the Latency is 120 ms, the energy consumption is 300 (mWh), and the privacy risk level is low. Conclusion: The combination of FEL and multi-modal sensor fusion provides a reliable and scalable approach for weed detection in precision agriculture. By processing data locally and collaboratively at the edge, the system achieves high accuracy, decreases response time, lowers energy consumption, and preserves data privacy.

Keywords: deep learning; FEL; sensor fusion; weed detection; precision agriculture

1. Introduction

In the 21st century, agriculture plays an important role. For this research work, environmental monitoring plays the key role. The AI (artificial intelligence) is used to make accurate decisions. Machine learning and deep learning techniques are rapidly being used in agriculture to predict crop diseases, monitor plant health, detect weeds, and classify pests. These applications showcase new and high-impact research fields where deep learning has demonstrated great promise. Simultaneously, sensor technologies are fast improving, allowing for more precise soil fertility assessments and supporting data-driven crop recommendation systems. In this scenario, DL (Deep Learning) emerged as one of the strongest techniques to solve the classification task. This technique allows us to handle the hidden patterns automatically from the raw data. However, the dataset's characteristics, feature distribution, and task complexity can all affect how well a deep learning model performs on its own. To address the limitations of typical machine learning models, in this paper, we used advanced fusion techniques to combine outputs from different models, improving robustness and accuracy in weed detection. We use a variety of deep learning (DL) architectures, including Convolutional Neural Networks (CNNs), Long Short-Term Memory (LSTM)-CNN hybrid models, and Vision Transformers (ViTs), to capture spatial and temporal patterns in multi-modal agricultural data. Furthermore, as a Federated Convolutional Neural Network (FedCNN) allows for decentralised training among distributed edge nodes, we suggest using one. By keeping raw data locally at each node and sharing only model parameters like weights and biases with the central server, this federated learning (FL) paradigm improves data security and privacy. Additionally, in order to successfully combine multi-modal inputs including LiDAR, RGB photography, and soil sensor data, three sensor fusion strategies—Early Fusion, Late Fusion, and Hybrid Fusion—are investigated. In practical agricultural settings, this integrated strategy improves energy economy, lowers latency, and protects data privacy while fortifying the system's predictive performance.

This study contains 5 sections, where Section 1 provides the information and background details of weed detection, and Section 2 presents related work of weed detection using deep learning, along with federated learning. Section 3 discusses the entire roadmap of the proposed model for distributed weed detection in precision agriculture using multimodal sensor fusion. Section 4 presents the results and discussion, followed by the conclusion and future scope.

2. Related Work

The author Xia et al. [1] used a multi-modal data fusion technique-based framework for weed resistance assessment. The author has demonstrated how combining various sensor modalities may efficiently capture geographical variability and create a strong link between multi-modal features and resistant weed concentrations. The experimental work

reveals that late fusion gives the best result of R^2 of 0.777 and RMSE of 0.547. Salve, P., et al. [2] used several classification tasks to identify the plant types. The author used the multi-modal fusion techniques to enhance the accuracy of plant type detection. Multiple feature types, such as leaf spectral signatures, leaf colour, and morphological form features, were used to process plant image datasets. The Genetic Algorithm-based Regression (GAR) model achieved an accuracy of 98.03%. Li, L., et al. [4] developed a framework that allows for smart farming. The author's objective was to design a lightweight FL framework, which is called VLLFL. The mentioned framework allows for identifying crop health and maintaining the pesticides. To optimize the crop, the author used FL along with the deep learning models. Their experimental observation found that VLLFL obtained 14.53% improvement and reduced 99.3% communication rounds. Anagnostopoulos, C., et al. [5] discussed how FL is helpful in an AIoT environment. The author discussed the challenges, present problems, etc. The authors' finding was to develop a novel classification strategy. The classification approach name is MMFL, which is used in the four application areas. Hussaini, M., et al. [6] developed a deep learning model for crop detection. The author explored how FL is used for crop identification as well as weed detection. The DL model YOLOv8 was used to train the dataset and apply the FL model FedAvg. The model was trained in homogeneous and heterogeneous environments. The Corn-Weed dataset was employed, which consists of 3575 annotated images. The experimental observation revealed that YOLOv8 performs well in comparison to the other two versions of the YOLO algorithm in terms of the performance metrics accuracy, memory space, and CC (Computational cost) when the training is considered. Mamba Kabala, D., et al. [7] developed a model that helps with crop disease. The author had used the FL models to protect the data. The PlantVillage dataset was taken and used to develop deep learning models, such as ViT, using FL. Rehman, M. U., et al. [8] developed a model that allows to detect weed detection using an advanced deep learning algorithm. The author considered the family of YOLO algorithms to detect weed identification. Their model achieved a precision of 72.5%, a recall of 68.0%

3. Proposed Model for Federated Edge Learning for Distributed Weed Detection in

Data Division: In this section, we split the dataset into three parts, and these are training (70%, testing 15% and the remaining 15% was used for validation purposes.

Model Architectures and Parameters: Here there are different deep learning models are employed for weed detection, and these are mentioned below:

CNN (Convolutional Neural): We used CNN with 5 layers, and each layer with its kernel size of 3×3 , ReLU activation, and batch normalization. Next model is **LSTM-CNN**: This is the hybrid model which combines an LSTM layer with 128 units to capture the patterns. **Vision Transformers (ViTs)** are used with 12 attention heads and patch sizes of 16×16 , along with feedforward networks of 2048 units. The hyperparameters are used for different models, and these are: LR: 0.01, batchsize: 32 for local training and 64 during the centralized evaluation. Apart from this, we also used the Adam optimizer where $\beta_1 = 0.9$ and $\beta_2 = 0.999$. The dropout we considered was 0.5 to avoid overfitting. The binary cross-entropy loss has been considered for weed classification.

The above-mentioned Figure 1 presents the privacy preservation weed detection system, and it consists of seven phases. And these descriptions are as described below: **Phase 1:** In this phase, we used the multi-modal sensors to collect the data from several sources. In order to get the plant structure, we have used a LiDAR sensor that captures 3D spatial types of data. For high-resolution images RGB camera is used to take a snapshot of the weeds. Similarly soil moisture sensor is used for collecting the real-time moisture levels.

Our objective was to collect the data from multi-modal sensors to monitor the weed growth patterns. Apart from other aspects, also. In this research paper, we used a publicly available dataset named “PlantVillage” for pertaining purposes. Apart from this dataset, we also collected the data in-field images using different sensors (including smartphone-based cameras). So we employed the hybrid dataset for identifying the weed. **Phase 2:** During the preprocessing stage, we remove the noise, perform the normalization techniques. To enhance the image quality, the RGB/multispectral images have been used. Before the model selection, we have done extensive way of data preparation and analysis. The outcome of this phase is cleaned data, and it is processed at the edge level. To extract a lightweight CNN-based. To minimize computational load, we have used leaf shape, colour histogram, spectral signatures, and soil metadata that are recovered using lightweight CNN-based models after preprocessing. This procedure minimizes communication overhead and protects privacy by ensuring that only pertinent, condensed features are kept. **Phase 3:** Feature Extraction and Multimodal Sensor Fusion. In this phase, we have identified the relevant features which was extracted from the weed images. These features are called multi-modal features. These features are further utilized for data fusion purposes. The collected sensor data are further fused into the 3 following three categories: Early Fusion: In this case, our collected sensors’ raw data is converted into a unified format. Mid Fusion: Preprocess the sensor’s data independently and fuse the intermediate features and Late fusion is used to combine the final sensor data and provide the average output. This technique is used to classify the weeds and estimate the performance metrics. **Phase 4:** In this phase, we used deep learning models that properly classified the weeds. The fused data is passed into the deep learning model to deep dive to process and weed out the identified. For extracting the spatial features, we used DL models; the LSTM-CNN model is used for spatial-temporal analysis purposes. It is also called a hybrid model. Similarly, a vision transformer (ViT) is used for the analysis of high-dimensional data.

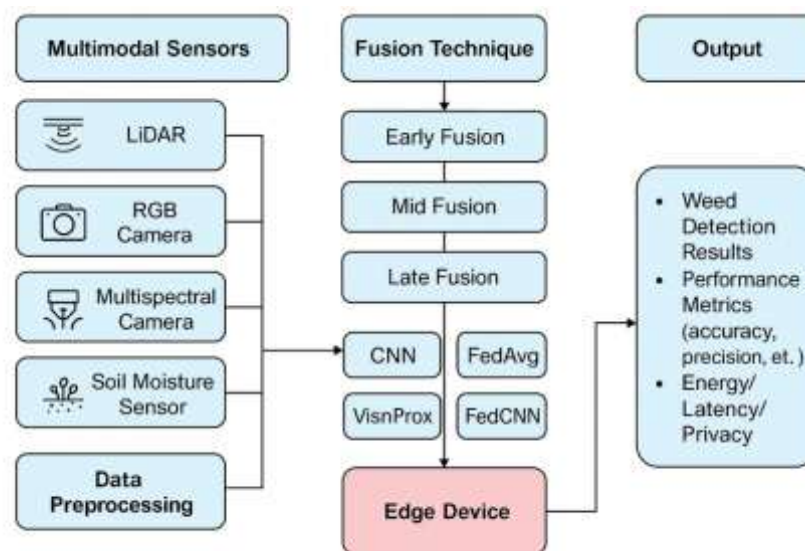


Figure 1. Proposed model for privacy-preserving weed detection system.

The above Figure 2 represents the sequence diagram for Federated Weed Detection, which discusses the time-ordered interactions between the different components associated. Here, the components include are FL server, weed management, sensors, and edge) etc. The data is usually collected from the different sensors and passed into the edge device. The sequence diagram discusses how the components interact with each other, and it follows the hybrid fusion and local FedCNN training, then the model is encrypted and

sent to the federated learning server. In the FL server, the FedProx algorithm is used for aggregation. FL-server sent the updated global model to the edge devices. At the end, the edge device is responsible for deploying the model for real-time weed classification. In this paper, the weed detection accuracy obtained is about 94.1%.

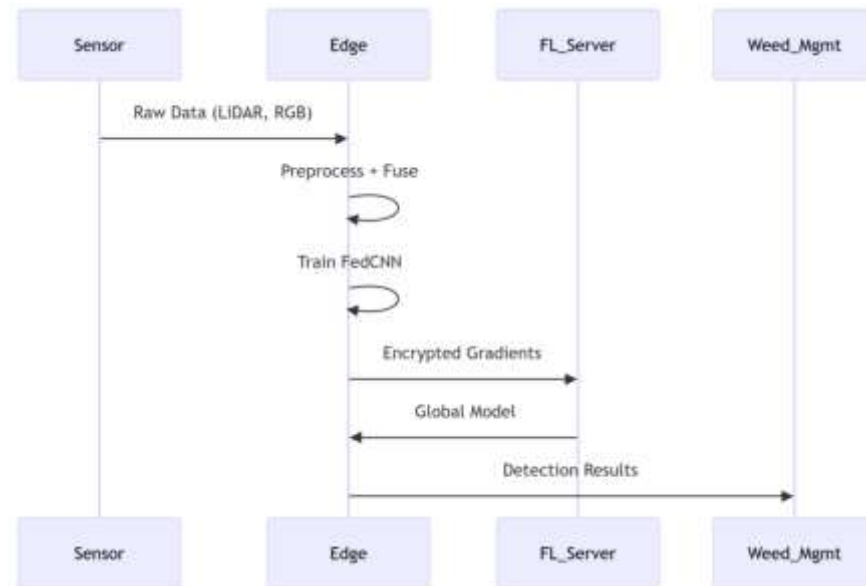


Figure 2. Sequence diagram for Federated Weed Detection.

4. Results and Discussion

Deep learning model creation and processing:

The acquired RGB and LiDAR datasets, as well as other sensor data, were preprocessed through a series of procedures to ensure data quality and compatibility for deep learning models. Images and point cloud data were first aligned using GPS coordinates, then noise and outliers were removed. The data was scaled to 224×224 pixels for deep learning models and normalised with min-max scaling to improve convergence during training. Data augmentation techniques like as rotation, flipping, and brightness modifications were used to boost dataset diversity while reducing overfitting.

Table 1 discusses the performance metrics of the different feature fusion techniques. There are three fusion techniques implemented, and these are Early Fusion, Mid Fusion, Late Fusion, and Hybrid Fusion. All these fusion techniques are used against the FedCNN model, which is used to detect weed identification. It has been observed that the hybrid fusion technique obtained the highest accuracy of 94.1%, precision of 94.3% and recall of 93.9%, F1-score is 94.1%, and AUC-ROC is 94%. It also observed that the 92.8% obtained accuracy in early fusion.

Table 1. Performance Comparison of Different Fusion Strategies with FedCNN.

Fusion Strategy	Accuracy (%)	Precision (%)	Recall (%)	F1-Score (%)	AUC (%)
Early Fusion	92.8	92.9	92.5	92.7	92.8
Mid Fusion	93.5	93.6	93.2	93.4	93.5
Late Fusion	93.7	93.9	93.4	93.6	93.7
Hybrid Fusion	94.1	94.3	93.9	94.1	94.1

From the above-mentioned Table 2 discusses the comparison of different deep learning models for weed identification. It has been observed that the best model is FedCNN, and their accuracy is 94.1% as comparison to the other models. The AUC-ROC curve

obtained 94.1%, F1-score is 94.1%. The FL uses the distributed edge devices for training instead of centralizing data, which improves generalization while keeping privacy. The traditional model, such as CNN, obtained the lowest accuracy of 90.5%. CNN can't capture complex spatial patterns as well as hybrid or transformer-based architectures.

Table 2. Comparison of Deep Learning Models for Weed Detection.

Model	Architecture Type	Accuracy (%)	Precision (%)	Recall (%)	F1-Score (%)	AUC (%)
CNN	Convolutional Neural Net	90.5	91	89.9	90.4	90.6
LSTM-CNN Hybrid	Spatial + Temporal Fusion	92.3	92.5	91.8	92.1	92.4
Vision Transformer	Attention-based Transformer	93.5	93.2	93.1	93.1	93.4
FedCNN (Proposed)	Federated CNN (Edge Device)	94.1	94.3	93.9	94.1	94.1

Table 3 presents the comparison between centralized and federated edge learning. It has been observed that the centralized model performs a little bit better than the federated learning model. The FEL reduced the latency by 120 ms as well, and energy consumption dropped from 500 mWh to 300 mWh. It helps to protect privacy more in comparison to the centralized model.

Table 3. Centralized vs. Federated Edge Learning (FEL) Performance.

Learning Approach	Accuracy (%)	Latency (ms)	Energy Consumption (mWh)	Privacy Risk Level
Centralized Learning	94.5	250	510	High
Federated Edge Learning (FEL)	94.1	120	300	Low

From the above Table 4 presents the comparison between FL algorithms. We have compared the two FL algorithms like FedAvg and Fed Prox. It has been observed that the FedProx FL algorithm's accuracy is 94.1%, precision of 94.3% and F1-score is 94.1%. The algorithm FedAvg is still competitive, but its performance is slightly lower because it is sensitive to non-IID (not independent and identically distributed) data, which makes convergence less stable.

Table 4. Comparison of FL Algorithms (FedAvg vs. FedProx).

FL Algorithm	Accuracy (%)	Precision (%)	Recall (%)	F1-Score (%)	AUC (%)	Latency (ms)	Energy (mWh)
FedAvg	93.7	93.6	93.5	93.5	93.6	150	350
FedProx	94.1	94.3	93.9	94.1	94.1	120	300

In Figure 3, we compared the three different fusion techniques for FL weed detection along with the 5 performance metrics. It has been observed that the Hybrid fusion technique performs well with 94.1% accuracy and 94.3% precision. As well as it is also observed that the late fusion (+1.4%) and early fusion (+3.9%). These findings help us understand that Hybrid Fusions' performance is good at balancing between precision and privacy preservation. But Late Fusion may be preferable for resource-constrained edge deployments.

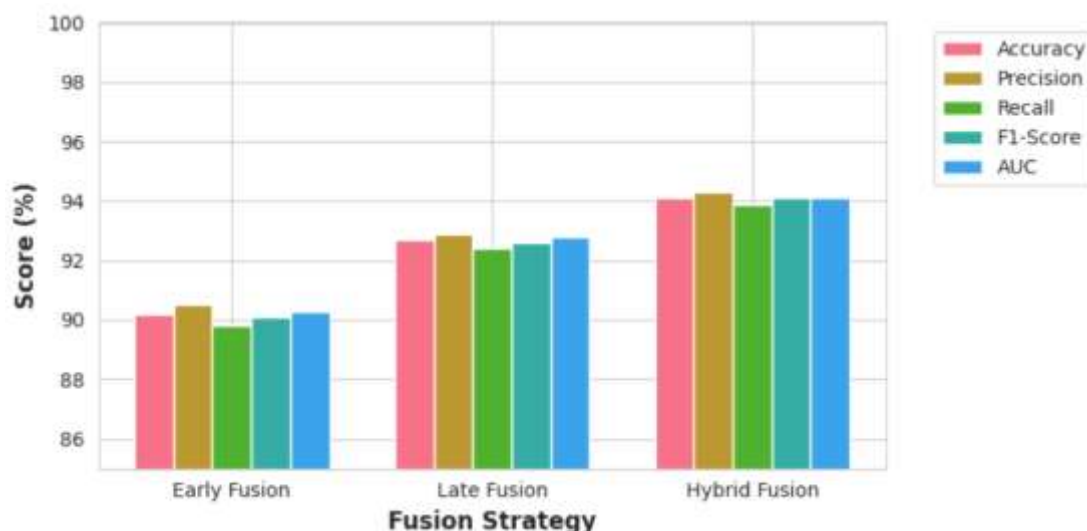


Figure 3. Performance metrics by Fusion strategy.

In Figure 4, we have estimated the critical trade-offs between the centralized and Federated learning environment, where our objective was weed detection. The FedCNN obtained the highest accuracy (**94.1% vs. 93.5%**). The AUC-ROC curve is most important to measure the cost of misclassifying (both TPR and TNR) for weed detection. When the dataset is imbalanced, the traditional metrics like accuracy may not be ideal. In contrast, the ROC curve helps to visualize the trade-off between TPR and TNR. The AUC gives a solid measure of the model's ability to tell the difference between things by putting this performance into a single number.

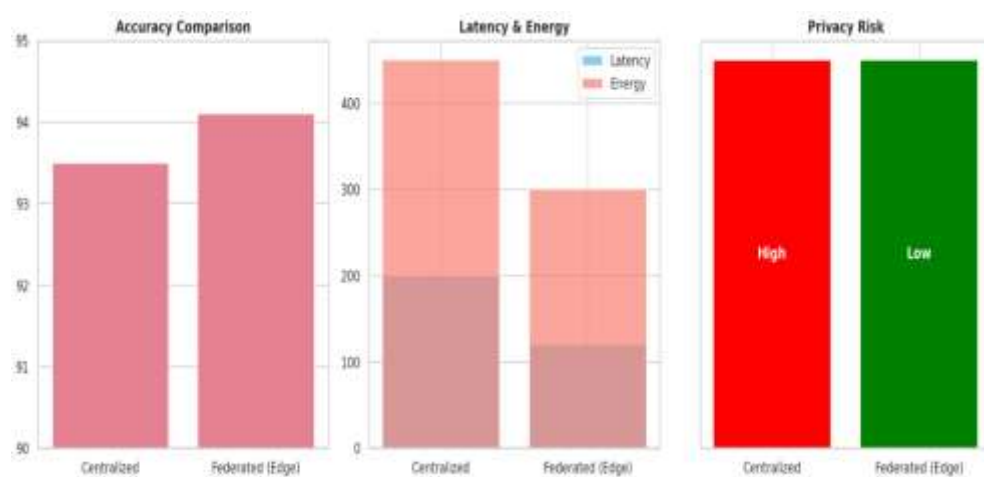


Figure 4. Critical trade-off between centralized and federated learning.

Figure 5 presents the AUC-ROC curve for fusion strategies that are used to determine the contribution to the classification accuracy. During our experimental observation, we found that our hybrid fusion approach obtained the highest AUC of 94.1%. It means that this demonstrates how early weed detection is possible. The other two fusion techniques were also utilized and obtained the performance with an AUC of around 93.7% and 93.5%.

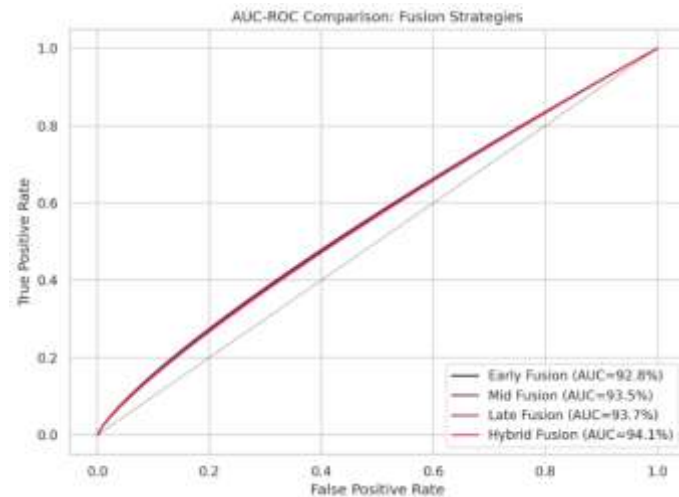


Figure 5. AUC-ROC comparison of fusion strategies.

Figure 6 presents the AUC-ROC curve, which is used to detect the weed using deep learning models. It has been observed that FedCNN obtained the highest AUC Score of 94.1% as compared to the other DL models. Similarly, CNN obtained an AUC score of 90.6%, Vision transformer AUC score is 93.4%. In this graph, the top left is the better model. The FedCNN AUC is one of the general models with having low TFR. It's clear from these results that federated learning and hybrid designs can help models work better on devices that are spread out. This metric is suitable for handling binary class classification. If the AUC is greater, it means that the model or fusion strategy can make more accurate predictions when the decision boundaries change. This is very important in edge-device federated learning settings like FedCNN, where the conditions for a threshold may change on the fly based on how the data is distributed and how it is deployed.

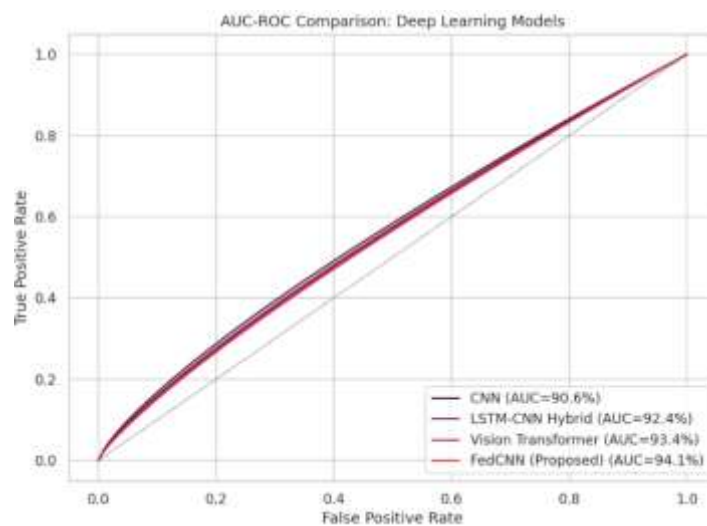


Figure 6. AUC-ROC comparison of different deep learning models.

In the above-mentioned Figure 7, we plot the precision and recall curve for weed classification using the precision and recall scores of the different models. We evaluated the different machine learning models that are used for a real-time insect detection and monitoring system. We used the PR curve, especially when the dataset is imbalanced. It allows researchers to visually assess how well each method balances precision and recall, which helps them choose the highest-performing and most reliable approach for deployment.

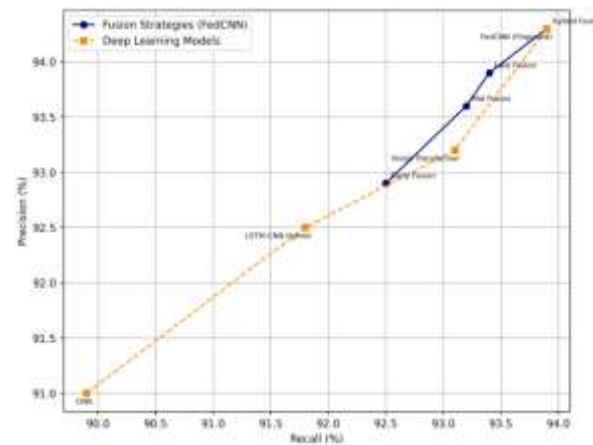


Figure 7. Precision and Recall Curve between fusion strategies and deep learning models.

In Figure 8, we compared the prediction accuracy of the federated learning CNN model and the centralized model concerning the number of epochs. The solid line indicates how the model behaves on the training data, and the dashed line visualizes validation accuracy. This visualization describes how the model behaves the unseen data, whether it is a generalized model or not. The red line shows the highest validation accuracy as compared to the centralized training accuracy. The FL model exhibits little bit slower convergence because of decentralized updates.

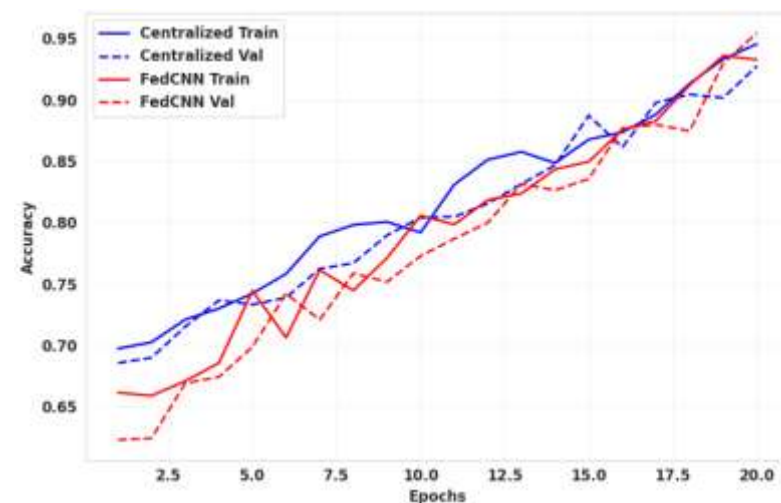


Figure 8. Training time vs. validation accuracy.

The above-mentioned Figure 9 discusses the relationship between training and validation losses. During training, it estimates the cross-entropy loss. If the loss is less than identified, as better optimized. The red line denoted as losses is the lowest one, which is compared to the centralized loss, and it is represented as blue lines.

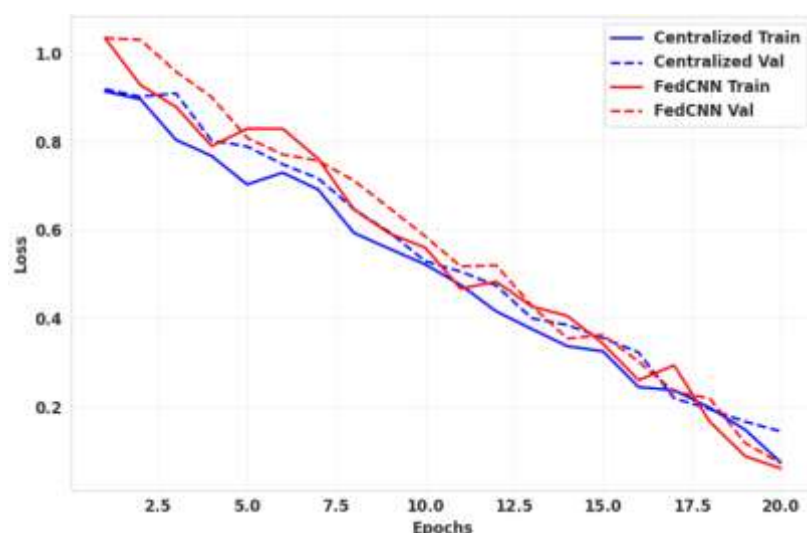


Figure 9. Training vs. validation loss.

Figure 10 demonstrates the performance of different deep learning models and distributed FedCNN. This graph compares the performance of several deep learning models—CNN, LSTM-CNN hybrid, Vision Transformer (ViT), and the distributed FedCNN—across three fusion strategies: Early Fusion, Late Fusion, and Hybrid Fusion. The above figure depicts that the Hybrid Fusion model performs well in comparison to the other models. The model FedCNN shows the best result, especially FedCNN. Apart from this, Vision Transformer also performs well with strong generalization capabilities.

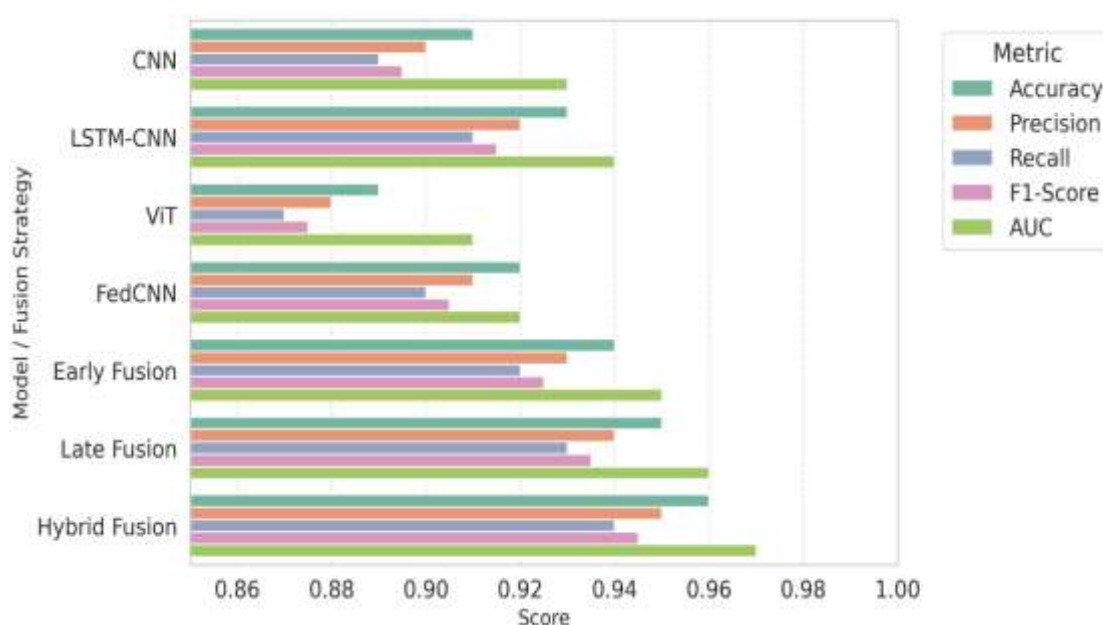


Figure 10. Performance Comparison of Models and Fusion Strategies.

The above-mentioned Figure 11 presents a heatmap representation of the individual machine learning models and fusion strategies. We have compared the fusion strategies side by side.

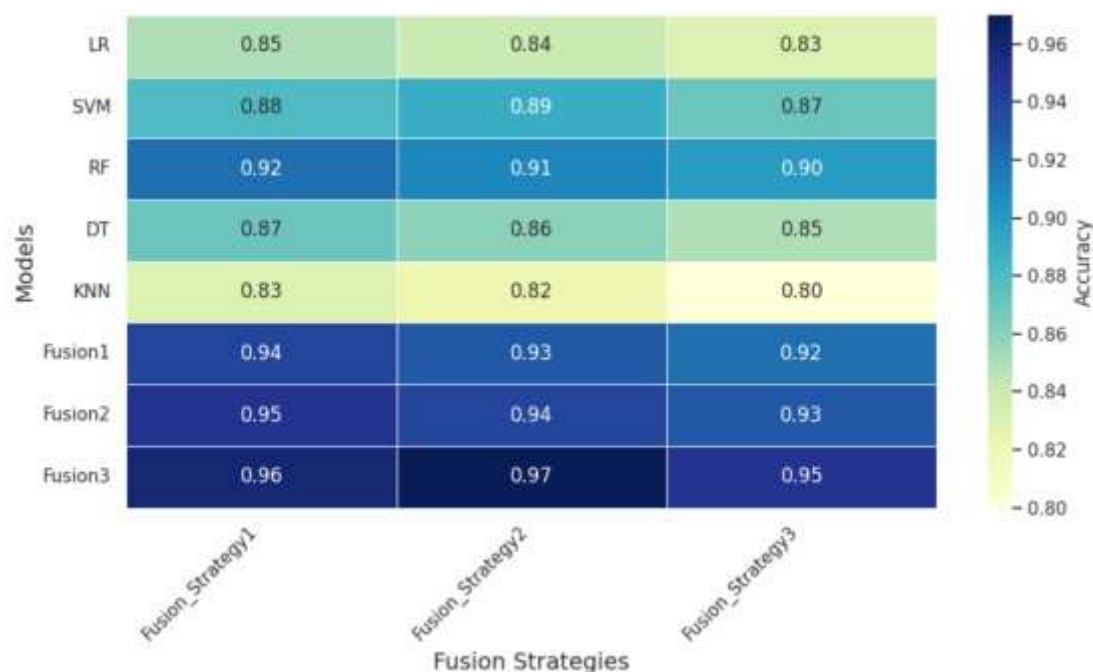


Figure 11. Performance Comparison of Models and Fusion Strategies.

5. Conclusions

This paper discussed the combined approach for weed detection in precision agriculture. We used the fusion techniques along with federated edge learning to detect the weed from the PlantVillage datasets. From our experimental observation, it has been observed that the hybrid fusion techniques provide good results. The performance obtained of 94.1% accuracy because of the three combined fusion strategies. The proposed Federated CNN (FedCNN) achieved 94.1% accuracy, a recall of 94.3% and an F1-score is 94.1%. Overall, the proposed FedCNN with Hybrid Fusion and FedProx optimization provides a robust, efficient, and privacy-preserving method for weed detection, allowing real-time, sustainable, and scalable smart farming applications.

Author Contributions: Conceptualization, D.A. and N.P.; methodology, N.P.; software, D.A.; validation, N.P. and D.A.; formal analysis, D.A.; investigation, N.P.; resources, D.A.; data curation, D.A.; writing—original draft preparation, D.A.; writing—review and editing, D.A.; visualization, D.A.; supervision, N.P.; project administration, D.A.; funding acquisition, N.P. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The dataset generated and analyzed during the current study is available from the corresponding author on reasonable request.

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Xia, F.; Lou, Z.; Sun, D.; Li, H.; Quan, L. Weed resistance assessment through airborne multimodal data fusion and deep learning: A novel approach towards sustainable agriculture. *Int. J. Appl. Earth Obs. Geoinf.* **2023**, *120*, 103352.
2. Salve, P.; Yannawar, P.; Sardesai, M. Multimodal plant recognition through a hybrid feature fusion technique using imaging and non-imaging hyperspectral data. *J. King Saud Univ.-Comput. Inf. Sci.* **2022**, *34*, 1361–1369.

3. Lou, Z.; Quan, L.; Sun, D.; Xia, F.; Li, H.; Guo, Z. Multimodal deep fusion model based on Transformer and multi-layer residuals for assessing the competitiveness of weeds in farmland ecosystems. *Int. J. Appl. Earth Obs. Geo Inf.* **2024**, *127*, 103681.
4. Li, L.; Li, J.; Chen, D.; Pu, L.; Yao, H.; Huang, Y. VLLFL: A Vision-Language Model-Based Lightweight Federated Learning Framework for Smart Agriculture. *arXiv* **2025**, arXiv:2504.13365.
5. Anagnostopoulos, C.; Gkillas, A.; Mavrokefalidis, C.; Pikoulis, E.V.; Piperigkos, N.; Lalos, A.S. Multimodal federated learning in AIoT systems: Existing solutions, applications, and challenges. *IEEE Access* **2024**, *12*, 180864–180902.
6. Hussaini, M.; Voigt, M.; Theiss, J.; Stein, A. Federated Learning for data-sovereign training of weed and crop detection machine learning models. In *45. GIL-Jahrestagung, Digitale Infrastrukturen für eine nachhaltige Land-, Forst-und Ernährungswirtschaft*; Gesellschaft für Informatik e.V.: Bonn, Germany, 2025; pp. 59–70.
7. Mamba Kabala, D.; Hafiane, A.; Bobelin, L.; Canals, R. Image-based crop disease detection with federated learning. *Sci. Rep.* **2023**, *13*, 19220.
8. Rehman, M.U.; Eesaar, H.; Abbas, Z.; Seneviratne, L.; Hussain, I.; Chong, K.T. Advanced drone-based weed detection using a feature-enriched deep learning approach. *Knowl.-Based Syst.* **2024**, *305*, 112655.

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.