SvedbergOpen
DISSEMINATION OF KNOWLEDGE

International Journal of Artificial Intelligence and Machine Learning

Publisher's Home Page: https://www.svedbergopen.com/

Research Paper Open

Confidence-Adaptive Fuzzy Ensemble Deep Learning for Imbalanced Potato Plant Disease Classification

Amit Kumar Jaiswal[1*], Syed Hauider Abbas[1], Binod Kumar Mishra[2]

[1]Department of Computer Science & Engg., Integral University Lucknow, Uttar Pradesh, Lucknow, India.
[2]Department of Computer Science &Engg. *Chandigarh University, UIE,* Mohali, Punjab, India.

*Corresponding Author: Amit Kumar Jaiswal., E-mail: amitjaiswal939@gmail.com, Orcid:0000-0002-0614-2881

Abstract
The accurate detection of diseases of potato plants is very important for reducing their losses and helping achieve precision agriculture. Even though there has been a great improvement in detecting the diseases of plants based on ensembled learning through images, efficiency of ensembled with deep learning in dealing with imbalanced training data is adversely affected, resulting in low sensitivity to minority disease categories.This paper proposes a Confidence-Adaptive Fuzzy Ensemble Deep Learning (FEDL) model that employs heterogeneous deep learning algorithms along with uncertainty-aware decision fusion in imbalanced potato disease classification problems. The model employs EfficientNet-B3, ConvNeXt-Tiny, and Vision Transformer (ViT) to capture complementary feature representations of local and global nature. The order to address with issue of minority-class inequality, a hybrid approach including SMOTE, data augmentation, balanced sampling, and Focal loss is used in the model training phase. While conventional ensemble methods use fixed model weights for each classifier, the proposed system uses a Mamdani Fuzzy Inference System (FIS) to determine the weight of the model based on the confidence, entropy, probability margin, and reliability of the backbones.The performance evaluation of the proposed framework was conducted by applying it to the class-imbalanced PlantVillage potato disease dataset with MCC, AUC, F1 score, Accuracy, Precision, Recall, and Balanced Accuracy. Results from comparative experiments, ablation studies, and Grad-CAM visualization show that the confidence adaptive fuzzy fusion leads to improved classification results, better minority class detection, and interpretability of predictions compared to deep learning classifiers and traditional ensembles. It was shown that the proposed framework can be considered an effective and generalized problem solution for reliable plant disease classification in class-imbalanced scenarios.

Key words: Potato Plant Disease Classification; Deep Learning; Ensemble Learning; Fuzzy Inference System; Imbalance Class; EfficientNet-B3; ConvNeXt-Tiny; Vision Transformer; Precision Agriculture

1. Introduction

The potato (Solanum tuberosum L.) is one of the most extensively grown food cropswith forms an important part of global food production. Due to its nutritional value and the ability to grow in different climatic conditions, potato farming makes significant contributions towards food security. Potato is one of the four leading food crops in the world according to the Food and Agriculture Organization (FAO). The other three being maize, wheat, and rice [1]. Stable potato production is difficult to sustain since there is a constant threat to potato crop production from different diseases caused by fungus, bacteria, and virus. The two major diseases are Early Blight and Late Blight, which cause serious yield loss and economic damage due to late detection. Early detection of infected plants becomes an imperative requirement for effective management of diseases. Early detection ensures that farmers can take necessary measures against disease without having to affect large cultivation fields, thereby minimizing crop losses and excessive use of pesticides [2]. Figures 1 and 2 show the visible symptoms displayed by potato leaves when healthy and when affected by the disease.

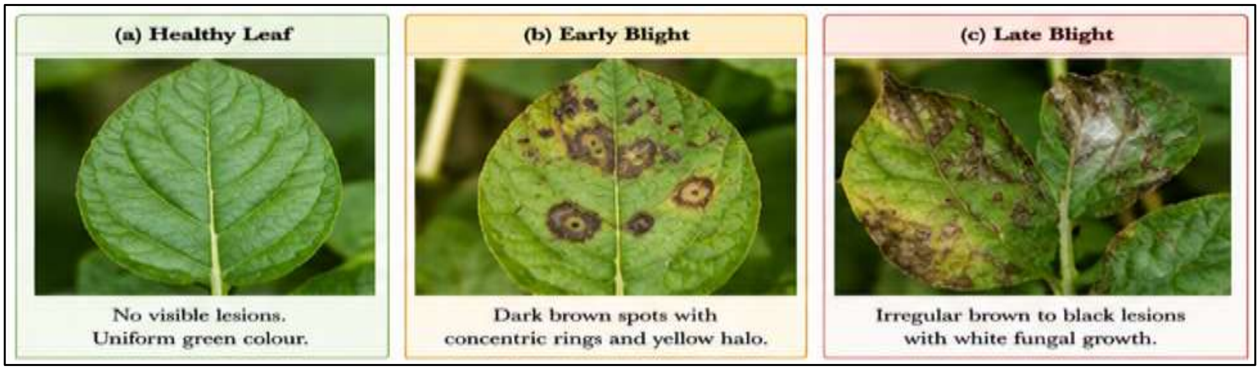

Figure 1: Types of Disease

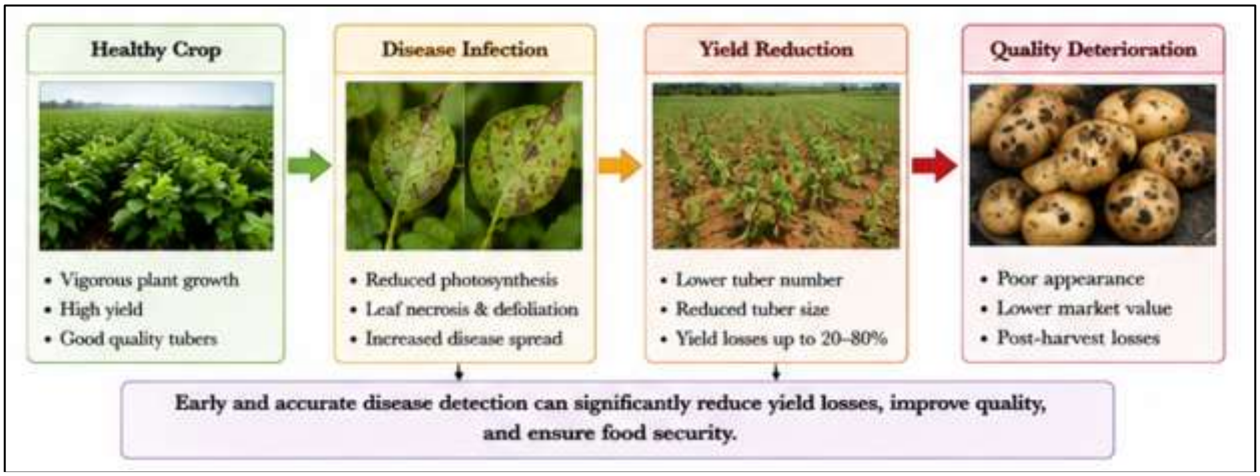

Figure 2: Impact of potato diseases on crop yield and quality

Classically, the diagnosis of plant diseases is done through visual inspections by agricultural experts. While manual inspection is the common procedure for most farm areas, its accuracy is highly dependent on personal experience and the conditions in which it is carried out. The differences in presentation of the symptoms, environmental factors like lighting conditions, the seriousness of the disease, and the experience level of the person conducting the inspection make the process highly inconsistent. Additionally, access to such experts is not easily available in many rural and resource-limited areas.

With recent breakthroughs in artificial intelligence especially deep learning, there has been significant improvement in the automation of plant disease detection. The CNNs which include AlexNet, VGGNet, ResNet, DenseNet, MobileNet, and EfficientNet have been successful in learning hierarchical representations from images of leaves and have exhibited outstanding classification accuracy on multiple benchmark datasets [1, 2]. Recently, deep learning models based on transformers like the Vision Transformer (ViT) have further enhanced these capabilities by capturing long-range spatial relationships using self-attention, leading to better disease pattern representation [3, 4]. All this progress has resulted in deep learning becoming the state-of-the-art technology for plant disease detection. Notwithstanding all the above advances, there are several practical problems that affect the applicability of deep learning in agriculture. Class imbalance stands out among them. In reality, when cultivating crops, the number of healthy leaf samples exceeds the number of diseased ones. Additionally, some diseases happen only occasionally. Therefore, it is hard to balance class ratios in training sets. The application of loss functions leads to models that are heavily biased toward majority classes, which results in a great level of accuracy but inability to identify minority disease classes accurately.

The second limitation relates to the process of integrating predictions made by several deep learning models. Ensemble learning has continuously enhanced classification accuracy through the use of complementary feature representations from distinct network structures. Voting, weighting, bagging, boosting, and stacking have all been used for classification of plant diseases. Despite this, existing ensemble learning strategies have generally utilized static weights for individual classifiers based on an implicit assumption that each classifier provides the same level of reliability on all input images. However, the reliability of predictions varies depending on the severity of the disease, quality of the image, complexity of the background, and visual similarity among various diseases.

A very important issue that is similar to the previous one concerns the poor handling of the uncertainty of the prediction. In deep neural networks, there are high confidence values provided for the cases where the image may be ambiguous or from underrepresented class of diseases. It means that confidence value does not necessarily correspond to the actual reliability of the prediction. There are other values like prediction entropy, probability margin and historical model reliability that can provide extra information about the uncertainty, but those values are usually ignored in the process of forming the final ensemble decision.

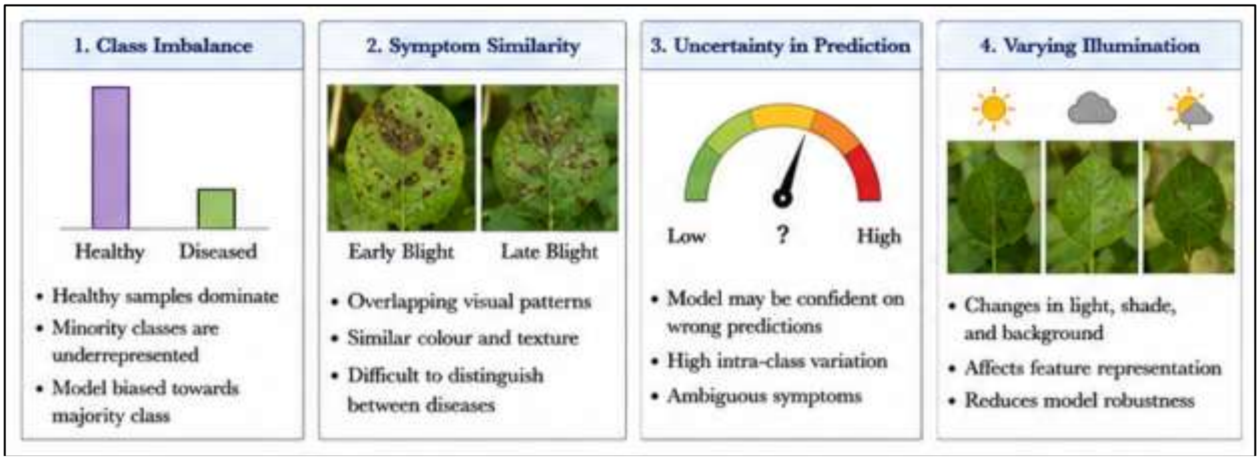

Figure 3: Various Challenges in automated potato disease classification

The fuzzy logic approach offers a structured framework for uncertain inference where knowledge is modeled using linguistic variables and rules established by the experts [5]. In contrast to deterministic decision algorithms, which find it difficult to model the natural transition from confidence to uncertainty, fuzzy inference systems provide a neat way out of this problem, thereby making them quite appropriate for use in decision fusion [6]. Although there have been several applications of fuzzy logic in the context of decision support systems for agriculture, irrigation, monitoring of crops, and disease prediction, very little work has been done in the area of their combination with deep learning ensembles. Inspired by these findings, we propose a Confidence-Adaptive Fuzzy Ensemble Deep Learning (FEDL) scheme for the classification of potato plant diseases in the context of class imbalanced setting. The proposed scheme employs a combination of three different deep learning models, namely EfficientNet-B3 [7], ConvNeXt-Tiny [15], and Vision Transformer [2, 8] in order to extract both fine-grained local information and global context representations. Instead of fusing their predictions through fixed averaging or majority voting approach, in our approach we propose to learn the dynamic ensemble weights using a Mamdani Fuzzy Inference System (FIS) [10]. The fuzzy inference process takes into account the prediction confidence, entropy, probability margin, and model reliability, enabling the determination of the individual contribution of each backbone in each input instance. Furthermore, the integration of a hybrid imbalance-learning approach comprising SMOTE [3], extensive data augmentation, and Focal loss [9] is considered to increase the classification performance on minority classes.
The proposed framework is developed to address several limitations identified in existing approaches. Its principal contributions are summarized as follows.
1. A Confidence-Adaptive Fuzzy Ensemble Deep Learning (FEDL) framework is developed for potato plant disease classification under class-imbalanced conditions.
2. Hybrid Imbalance Learning technique using the combination of SMOTE, data augmentation and Focal Loss used in training.
3. The Mamdani Fuzzy Inference System is used to adaptively estimate the weights of the ensemble based on the prediction confidence, entropy, probability margin and the past performance of the model instead of following any static voting techniques.
4. EfficientNet-B3, ConvNeXt-Tiny and Vision Transformer are used in a unified adaptive ensemble system, which takes advantage of complementary feature representations while suppressing uncertain prediction effects.
5. Extensive experiments show consistent improvements over individual DL(Deep Learning) model and ensemble systems in potato disease classification under imbalanced datasets.In Section 2 examines the body of current researchon potato plant disease classification, ensemble learning, fuzzy logic, and class imbalance handling. In Section 3 presents and proposed Confidence-Adaptive FEDL framework in detail.In Section 5,

experimental findings and comparative analysis are covered.In Section 6, Researchwork is finally concluded and future research possibilities are outlined.

2. Related Work

Automated plant disease diagnostics have received much interest among researchers owing to their potential in increasing crop yields and precision farming. Rapid advancements in the sector of computer visionwith deeplearning with ensemble model have made it possible to significantly improve the efficiency of disease detection by using the leaf images, making the automatic diagnosis process less dependent on humans.CNN(Convolutional neural networks), transformer models, andTL(transfer learning) have shown excellent results in benchmarking datasets like PlantVillage [16, 17, 31]. However, there still exist many unsolved problems. Namely, most of the papers presuppose balanced training data sets, use static decision fusion methods, and consider only a limited number of factors affecting prediction uncertainty, which limits their applicability in real-world agricultural conditions where the disease appearance is naturally unbalanced. In this section, previous work is discussed from five different angles: traditional machine learning approaches, disease classification using deep learning, ensemble learning, fuzzy logic application for agricultural decision support, and class imbalance learning.

2.1 Traditional Machine Learning Methods

First experiments for disease recognition of potato plants used only handcrafted features with conventional machine learning techniques. These include the extraction of color histograms, texture, shape, and Local Binary Patterns (LBP) from images of leaves. Then the leaf images were classified using various machine learning algorithms such as SVM(Support Vector Machine), Random Forest (RF), k-Nearest Neighbors (k-NN), Decision Tree, and Naïve Bayes. This approach yielded good results under controlled imaging conditions. However, the performance of the system was highly dependent on the manually defined feature set. Changes in lighting conditions, background complexity, leaf orientation, and different stages of the disease negatively affected the generality of the system. Handcrafted feature set was not able to extract more complex features related to disease development.

2.2 Deep Learning-Based Disease Classification

Development of deep learning revolutionized the problem of identificationautomatically of plant diseases by enabling the creation of models capable of learning feature hierarchies without any manual intervention. CNNs are not dependent on any manual feature extraction process, and have shown impressive results with various plant disease data. AlexNet, VGGNet, ResNet [19], DenseNet [20], MobileNet, and EfficientNet [7] are few popular models that have been widely used due to their capability of extracting texture and structure information from the diseased leaf images [1, 2, 15, 18, 27, 37].

Most recently, the transformer architecture has further enhanced the power of deep learning in terms of modeling long-range spatial dependencies using the technique of self-attention. The Vision Transformer (ViT) withHybrid CNN-transformer models successfully combined the local lesion properties with the global context information, leading to better visual disease pattern recognition [9, 24, 32, 34, 36]. Despite attaining high classification accuracy in benchmarking datasets, the majority of the above models assume balanced data during training. In case of skewed distribution of classes, the optimization is often biased with the majority class, hence lowering the recognition rate for the minority class disease categories [33]. High overall accuracy may therefore hide low accuracy for important diseases that occur rarely.

2.3 Ensemble Learning for Plant Disease Classification

One technique for increasing the accuracy of predictions through the use of complementary strengths from several classifiers is ensemble learning. Bagging, boosting, stacking, majority voting, and weighted average are some examples of ensemble learning. It has been established in various studies that using several different CNN architectures or combining CNNs and transformer networks provides better robustness and generalization compared to single model [4, 23]. However, the majority of the currently used ensemble techniques are characterized by the use of fixed fusion techniques. Majority voting and fixed-weighted average implicitly suppose the same contribution of each classifier irrespective of the features of the processed input image. In reality, the performance of a particular classifier depends on the degree of disease manifestation, similarities between the disease symptoms, image resolution, and environmental factors. Since fixed-weighted algorithms do not take into account such differences, they might provide equal weights for both reliable and unreliable predictions, which reduces the efficiency of the ensemble.

2.4 Fuzzy Logic in Agricultural Decision Support

The reason why fuzzy logic has gained such popularity in agriculture is that fuzzy inference is a powerful means of representing uncertainty and incorporating expert knowledge in computer models [10]. Fuzzy inference has found applications in irrigation scheduling, crop yield prediction, soil condition assessment, pest control, and plant disease diagnosis. In the field of plant disease diagnostics, fuzzy inference has mostly been used for estimating disease severity, diagnosing symptoms, and providing expert-based decisions [25]. The possibility of combining expert knowledge with machine learning has also been shown in adaptive neuro-fuzzy models. Even though significant progress has been made, combining fuzzy logic with deep learning ensembles remains an unexplored direction [5, 6, 29, 30, 34, 35]. Previous research mostly utilized fuzzy inference in either pre-classification image processing or post-classification decision making. Little attention has been paid to the utilization of fuzzy inference for adaptively weighted decisions from multiple deep learning models. This deficiency can be exploited to unite uncertainty representation and deep feature learning into a single classification system.

2.5 Learning Under Class Imbalance

The class imbalance issue constitutes one of the main problems in agricultural image processing due to the natural inequality in the occurrence of diseases in different agricultural settings [3, 21]. When using deep learning models trained on an imbalanced dataset, the tendency towards biased predictions toward the majority classes arises, causing lower sensitivity for the minority classes. There exist some methods aimed at solving the issue, including random over-sampling, random under-sampling, SMOTE [11], ADASYN, cost-sensitive learning, class-balanced loss function, Focal Loss [12], and data augmentation through MixUp and CutMix. All these methods improve the representation of the minority classes during training and help improve classification accuracy under imbalanced settings [22, 28]. However, in most of the imbalance handling methods, only the data distribution or the loss is considered. The uncertainty involved in the predictions of the model before making a final prediction is not taken into account. Hence, the uncertainty-aware ensemble learning methods are underexplored in relation to plant disease classification.

2.6 Critical Analysis of Existing Studies

Table 1 summarizes representative approaches for potato plant disease classification together with their principal characteristics and limitations.

Method	**Backbone / Technique**	**Class Imbalance Handling**	**Ensemble Strategy**	**Fuzzy Logic**	**Main Limitation**
CNN-based classification [1]	AlexNet / VGG	NO	NO	NO	Limited feature representation and reduced robustness.
Transfer learning [2]	ResNet / DenseNet	NO	NO	NO	Sensitive to class imbalance and symptom similarity.
EfficientNet [7]	EfficientNet-B3	NO	NO	NO	Relies on a single backbone model.
ConvNeXt [8]	ConvNeXt-Tiny	NO	NO	NO	No adaptive decision fusion.
Vision Transformer [9]	ViT	NO	NO	NO	High computational complexity and limited uncertainty modeling.
Data-level balancing [11]	SMOTE	Yes	NO	NO	Balances data only; does not improve decision fusion.
Cost-sensitive learning [12]	Focal Loss	Yes	NO	NO	Addresses training imbalance but ignores prediction uncertainty.
Fuzzy decision systems [5,6,30]	Mamdani / Fuzzy Rules	NO	Partial	Yes	Not integrated with modern deep learning ensembles.
Ensemble deep learning [23,24]	Multiple CNNs / CNN–Transformer	Partial	Fixed Weights	NO	Static ensemble weights cannot adapt to prediction reliability.

Confidence-Adaptive FEDL (Proposed)	**EfficientNet-B3 + ConvNeXt-Tiny + ViT**	Yes	**Adaptive Fuzzy Ensemble**	Yes	**Dynamic confidence-aware weighting using uncertainty estimation.**

In general, there are a number of common problems present in the current body of literature. In many cases, the authors tend to use balanced training datasets that do not properly capture real-life agricultural situations. The majority of the ensemble models use static weights, which limits their flexibility to adjust to variations in the reliability of the predictions for particular instances. Despite the significant effectiveness of fuzzy logic models in agriculture, they are still not sufficiently combined with heterogeneous deep learning ensembles. Moreover, the uncertainty information obtained based on the confidence, entropy, and margins of decisions is rarely utilized when combining the ensemble decisions. All of the above-mentioned issues justify the necessity of designing a novel Confidence-Adaptive Fuzzy Ensemble Deep Learning (FEDL) framework that combines imbalance-aware learning, deep features complementarity, and fuzzy modeling of uncertainties in a single adaptive ensemble architecture. The proposed framework can help to overcome the shortcomings of other approaches and improve the recognition of the minority classes.

3. Proposed Methodology

The Confidence-Adaptive Fuzzy Ensemble Deep Learning (FEDL) framework suggested here aims to enhance the potato plant disease classification process under class imbalance situations by employing an ensemble of different deep learning models using a decision fusion approach based on uncertainties. Traditional ensembling techniques are usually based on either a static averaging approach or majority voting, where it is assumed that each individual model has the same level of accuracy across all samples. However, in reality, the confidence in each individual prediction depends on a number of factors such as the disease level, similarity among classes, lighting, and In order to overcome this issue, the novel approach utilizes a Mamdani Fuzzy Inference System (FIS) for computing sample-specific ensemble weights. This approach does not use fixed weights for each member of the ensemble, but instead computes a dynamic weight for each member of the ensemble based on the accuracy of prediction provided by that member. The contribution from each member of the ensemble is determined by the analysis of four different uncertainty measures, including prediction confidence, entropy, decision margin, and reliability of the model in history.

Figure 4 illustrates the architecture of the proposed approach. The method includes six stages, which are performed sequentially: image preprocessing, deep feature extraction, uncertainty estimation, fuzzy reasoning, adaptive ensemble fusion, and disease classification.

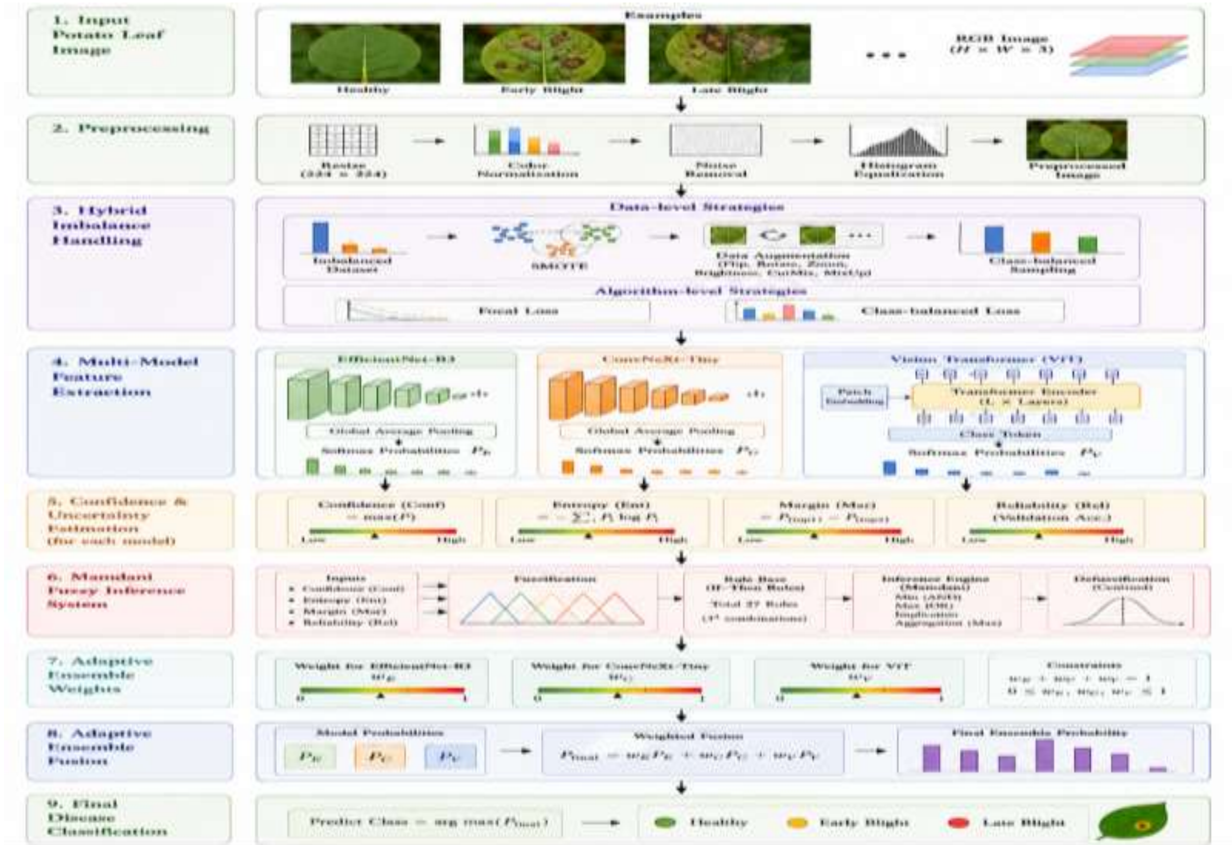

Figure 4: Confidence-Adaptive FEDL framwork

3.1 Image Preprocessing

The proposed framework consists of RGB potato leaf image acquired from the PlantVillage dataset. Before feature extraction, all image are resized to a uniform resolution of 224 × 224 pixels, corresponding to the input dimensions required by the selected backbone networks. Pixel intensities are normalized to improve numerical stability during optimization and to ensure compatibility with ImageNet-pretrained models. To improve generalization and reduce sensitivity to variations encountered in practical agricultural environments, several data augmentation operations are performed during training. Random horizontal and vertical flipping, rotation, zooming, brightness adjustment, and minor geometric transformations expose the model to diverse visual appearances without altering the biological characteristics of disease symptoms. Since the training dataset exhibits class imbalance, minority classes are further strengthened using Synthetic Minority Over-sampling Technique (SMOTE) [11], enabling the learning process to better represent under-represented disease categories.

3.2Deep Feature Extraction

The proposed FEDL framework combines three heterogeneous deep learning architectures that capture complementary visual representations of potato leaf diseases. EfficientNet-B3 [7] serves as the first backbone network. Owing to its compound scaling strategy, EfficientNet simultaneously balances network depth, width, and input resolution, enabling efficient extraction of fine-grained texture information from infected leaf regions. The second backbone, ConvNeXt-Tiny [8], retains the advantages of convolutional neural networks while incorporating modern architectural improvements that enhance feature representation and optimization efficiency. Its hierarchical convolutional structure effectively models local disease characteristics, including lesion boundaries and texture variations. The third backbone is the Vision Transformer (ViT) [9], which employs self-attention to capture long-range spatial dependencies within the image. Unlike convolutional architectures that primarily focus on local neighbourhoods, ViT models global contextual relationships, making it particularly effective for identifying disease patterns distributed across larger portions of the leaf surface. Each backbone independently processes the input image and produces a probability distribution over the three disease classes. These prediction vectors subsequently serve as inputs to the confidence estimation module.

3.3Prediction Uncertainty Estimation

However, the uncertainty of an individual prediction is not sufficiently captured by only confidence. Although the confidence value of a prediction may be very high, it could be very unreliable if there exists strong competition among class probabilities or the model itself has bad performance on the corresponding disease class in the past. Four uncertainty measurements will be considered in the proposed method.

Prediction Confidence: Prediction confidence corresponds to the maximum posterior probability produced by the Softmax classifier,

$$C = \max(P_i) \text{ ....(1)}$$

where (P_i) denotes the predicted probability for class i. Larger confidence values generally indicate stronger classifier certainty.

Prediction Entropy: Prediction entropy measures the uncertainty associated with the probability distribution,

$$H = -\sum_{i=1}^{N} P_i \log(P_i) \text{ ....(2)}$$

where N represents the total number of disease classes. Lower entropy indicates more decisive predictions, whereas higher entropy reflects increased uncertainty.

Decision Margin: The decision margin quantifies the separation between the two highest predicted probabilities,

$$M = P_{\max} - P_{second} \text{ .... (3)}$$

The larger the margin, the better separation the predicted class has from other possible classes, while a small margin means a decision is doubtful.

Historical Reliability: Historical reliability is the record of predictive power for each network over time. This measure does not consider predictions on an individual basis but takes into account the overall reliability of each network in different diseases.

3.4Mamdani Fuzzy Inference System

The adaptive weighting scheme under consideration is implemented with the help of a Mamdani Fuzzy Inference System (FIS) [10]. Fuzzy logic offers a powerful tool for integrating several uncertainty measures into one decision-making process while allowing for smooth transition from reliable to unreliable forecast.

Each input fuzzy variable is characterized by three membership functions with linguistic terms "Low", "Medium", and "High" respectively, see Figure 5. Triangular and trapezoidal membership functions are used due to their ease of computation and interpretability.

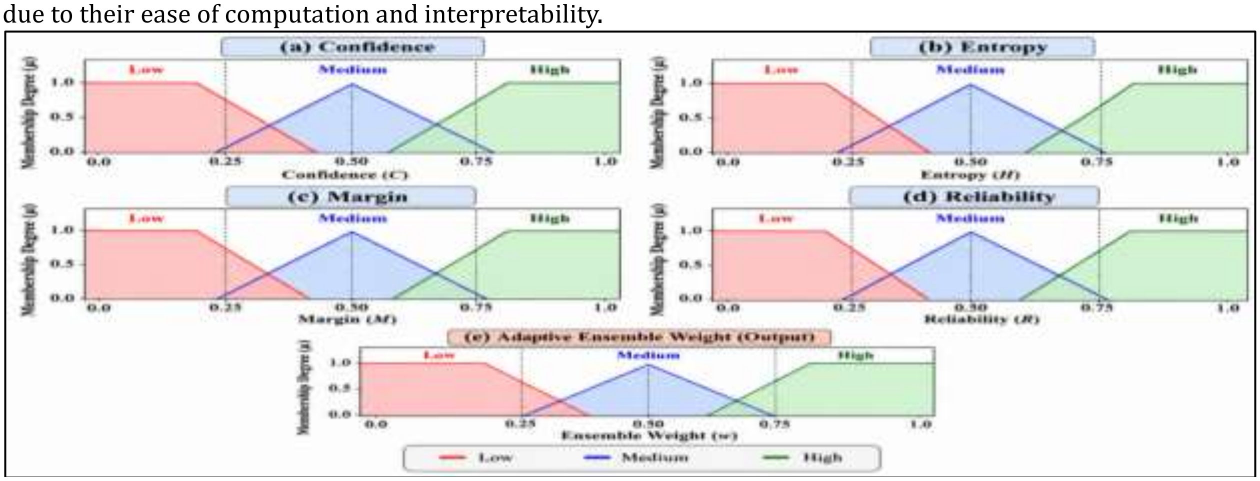

Figure 5. Fuzzy membership functions with input variables (Confidence, Entropy, Margin, and Reliability) and the output variable (Adaptive Ensemble Weight) employedMamdani Fuzzy Inference System for confidence-adaptive ensemble fusion.

The fuzzy rule base consists of expert-defined IF–THEN rules that model the relationship between prediction uncertainty and ensemble weight. Representative examples include:

- IF Confidence is High AND Entropy is Low , Margin is High then Weight is High.
- IF Confidence is Medium AND Entropy is Medium THEN Weight is Medium.
- IF Confidence is Low OR Entropy is High THEN Weight is Low.

Unlike conventional threshold-based weighting, fuzzy reasoning performs continuous inference across overlapping membership functions, enabling gradual adaptation of ensemble weights rather than abrupt switching between decision rules.

3.5Confidence-Adaptive Ensemble Fusion

After fuzzy inference, the predicted probabilities generated by the three backbone networks are combined using adaptive ensemble weights. Let P_E, P_C, P_V denote the probability vectors produced by EfficientNet-B3, ConvNeXt-Tiny, and Vision Transformer, respectively, while w_E, w_C, w_V represent their corresponding fuzzy weights.

The final ensemble prediction is computed as

$P_{FEDL} = w_E P_E + w_C P_C + w_V P_V$ …….. (3)

subject to $w_E + w_C + w_V = 1$

Unlike conventional averaging strategies, the weights vary dynamically for every input image according to the estimated prediction uncertainty. Consequently, the ensemble emphasizes highly reliable predictions while reducing the influence of uncertain classifiers.

The final disease label corresponds to the class with the highest probability,

$\hat{y} = \arg\max(P_{FEDL})$ ………. (4)

3.6 Algorithm of the Proposed FEDL Framework

Algorithm 1: Confidence-Adaptive Fuzzy Ensemble Deep Learning

Input: Potato leaf image (I)

Output: Predicted disease class ($\hat{y}$)

1. Acquire and preprocess the input image.
2. Apply augmentation and imbalance-handling techniques.
3. Feed the image into EfficientNet-B3, ConvNeXt-Tiny, and Vision Transformer.
4. Obtain Softmax probability vectors from each model.
5. Compute prediction confidence, entropy, margin, and reliability.
6. Fuzzify the input variables.
7. Apply the Mamdani fuzzy rule base.
8. Defuzzify the output to obtain adaptive ensemble weights.
9. Perform weighted probability fusion.
10. Assign the class with the highest fused probability.

3.7 Computational Complexity

The computational complexity of the proposed framework is primarily determined by the three backbone networks used during inference. The fuzzy inference module introduces only a small computational overhead because it operates on four scalar uncertainty measures rather than high-dimensional feature maps. Consequently, the additional cost associated with adaptive weighting is negligible compared with deep feature extraction. Since the backbone models execute independently, the framework can be parallelized on modern GPU architectures, allowing efficient inference without substantially increasing prediction latency. This characteristic makes the proposed method suitable for deployment in practical agricultural monitoring systems where both prediction accuracy and computational efficiency are important.

3.8 Summary of the Proposed Framework

The novel Confidence-Adaptive FEDL method fuses heterogeneous deep features along with uncertainty-aware fuzzy logic inference for the purpose of providing reliable disease prediction in class imbalanced datasets. Instead of using static weights for each member of the ensemble, the contribution of individual samples is determined through confidence, entropy, decision margin, and reliability. This adaptive fusion approach helps in adapting the model according to the degree of uncertainty associated with predictions while keeping the advantages of EfficientNet-B3, ConvNeXt-Tiny, and Vision Transformer intact.

4. Experimental Setup

The following part presents an overview of the dataset, pre-processing technique, model configuration, implementation, and evaluation procedure used for evaluating the proposed Confidence-Adaptive Fuzzy Ensemble Deep Learning (FEDL) architecture. The experiment procedure was designed to provide a fair comparison with the baseline algorithms as well as to perform a thorough test under class imbalance scenarios. In order to maintain uniformity across all experiments, all competing algorithms were trained and tested using the same data split, pre-processing procedure, and evaluation measures.

4.1 Dataset Description

The experiments were carried out based on the PlantVillage Potato Leaf Disease Dataset which consists of high-resolution images belonging to three classes: Healthy, Early Blight, and Late Blight [14]. This dataset is popular among researchers to benchmark image classification algorithms since it includes well-labeled images with varied symptoms of diseases acquired in constant conditions.

Table 2. Dataset Distribution

Class	Original Images	Training	Validation	Testing
Healthy	152	106	23	23
Early Blight	1000	700	150	150
Late Blight	1000	700	150	150
Total	2152	1506	323	323

Table 2 shows the distribution of the images employed in this experiment. The dataset was divided into training, validation, and test sets using a stratified sampling method in order to maintain a balanced class distribution in all experiment stages. Training set was used for the purpose of model optimization, validation set for tuning hyperparameters and early stopping, and independent test set for performance evaluation. Even though the PlantVillage dataset offers quite clean lab images, symptoms are highly visually similar, especially at the beginning of the infection period. It is therefore difficult to classify Early Blight from Late Blight and can serve as an adequate test environment for uncertainty-based ensemble learning approaches.

4.2 Data Preprocessing and Augmentation

Before training, all images were scaled to the size of 224 × 224 pixels in order to satisfy the input specifications of EfficientNet-B3, ConvNeXt-Tiny, and Vision Transformer models. Image pixels were normalized to facilitate the optimization process and enable the use of backbone models pre-trained on ImageNet dataset. In order to promote generalization and make the model less sensitive to possible variation found in real agricultural environment, some augmentation methods were used. Augmentation methods include random horizontal and vertical flipping, random rotation, zooming, adjustment of brightness, and small geometric transformation of images. In other words, augmentation helps the model to get more visual variation, but preserves biological properties of disease symptoms. Considering that class imbalance is the central issue of this research project, the minority classes were augmented by applying Synthetic Minority Over-sampling Technique (SMOTE) [11]. Contrary to simply duplicating the images of minority classes, SMOTE creates new feature vectors based on neighboring vectors.

4.3 Backbone Network Configuration

The proposed FEDL framework integrates three complementary deep learning architectures to capture diverse feature representations from potato leaf images. The suggested FEDL algorithm combines three different deep learning architectures that enable the learning of different feature representations from potato leaf images. The reason for choosing EfficientNet-B3 [7] lies in its novel approach to scaling a model, which involves tuning the depth, width, and input resolution at the same time in an efficient manner. The architecture learns effective features related to the texture of leaves and associated with disease lesions.

ConvNeXt-Tiny [8] represents the latest advances in the design of convolutional architectures and generates hierarchical features that show high robustness against local variations in the appearance of leaves. Vision Transformer (ViT) [9] enables modeling the long-range spatial dependencies with self-attention. Its capacity to learn global contextual relationships is especially important for disease symptoms spread across a larger area of the leaf. All backbones were pre-trained with ImageNet dataset and further fine-tuned on the dataset of diseases in potatoes. Transfer learning allows us to increase the speed of training and the generalization on limited data.

4.4 Confidence-Adaptive Fuzzy Ensemble

In contrast to traditional ensembles where equal weight is given to all the models used in the backbone, the presented system calculates dynamic weights for each sample by using the Mamdani Fuzzy Inference System [10].

The fuzzy inference procedure takes into account four mutually supplementary criteria:

- Prediction confidence, Prediction entropy
- Decision margin
- Historical model reliability

Together, they define the certainty of every prediction made. Fuzzy membership functions convert numerical values into linguistic variables, while the rules-driven inference process calculates the role of each of the backbones in the decision-making process. As a result, predictions which are highly certain and have low uncertainty levels are given greater weights in the ensemble, while uncertain predictions are given lesser weight during decision-making. This is how an adaptive weighting process can be implemented within the ensemble framework.

4.5 Training Configuration

All the experiments have been done through the use of Python programming language, using the TensorFlow/Keras deep learning framework.

In the optimization of backbone models has been carried out through the use of Adam optimizer using the initial learning rate of 1×10^{-4}. The training process comprised of 100 epochs while selecting the best model based on its validation accuracy. Early Stopping and ReduceLROnPlateau optimization techniques were used in order to mitigate overfitting.

As a result of the presence of class imbalance in the training dataset, Focal Loss [12] has been used as the loss function. Focal loss is responsible for placing higher weights on minority and difficult class examples in such a way that the model learns more discriminative features without being dominated by majority class examples. The hyperparameters used in the experiments are shown in Table 3.

Table 3. Training Parameters

Parameter	**Value**
Optimizer	AdamW
Learning Rate	0.0001
Batch Size	32
Epochs	100
Weight Decay	0.0001
Scheduler	Cosine Annealing
Input Size	224 × 224
Loss Function	Focal Loss
Early Stopping Patience	10 Epochs

4.6 Evaluation Metrics

The evaluationmodel performance was based on a number of quantitive measures such as Accuracy percentage, Precision value, Recall, F1-score, Balanced Accuracy, Matthews Correlation Coefficient (MCC), and Area Under the Receiver Operating Characteristic Curve . Whereas accuracy provides an insight into overall performance of the predictor, Balanced Accuracy and MCC scores are especially informative as they consider uneven class distribution and punish any biases. ROC with Precision–Recall curves allow us to estimate the discriminative ability of the developed model at various levels of the decision threshold. Beside the quantitive analysis, Grad-CAM [13] was applied to visualize image regions, which contribute to classification decisions. The visual explanations allow us to get a qualitative understanding of the fact that the suggested approach pays attention to biologically relevant disease regions, but not to unnecessary background data.

5. Results and Discussion

In this partprovides Evaluation of the proposed Confidence-Adaptive Fuzzy Ensemble Deep Learning (FEDL) method from the perspective of its quantitative performance measures, convergence, decision fusion based on uncertainty principles, and visual demonstration of internal model reasoning. The experiments were conductedin terms of the accuracy of classification but also the robustness and reliability of the proposed method in the presence of class imbalance. Specifically, much attention was paid to the application of the evaluation metrics such as Balanced Accuracy, Matthews Correlation Coefficient (MCC), and Area Under the ROC Curvewhich provide more accurate evaluation than the accuracy measure itself in the context of class imbalance.

5.1 Convergence Analysis

The learning behavior of the suggested FEDL framework was analyzed using the training and the validation accuracy with loss curves shown in belowFigures 5 and 6. These graphs give information about the optimization process and generalization ability of the suggested confidence adaptive ensemble.

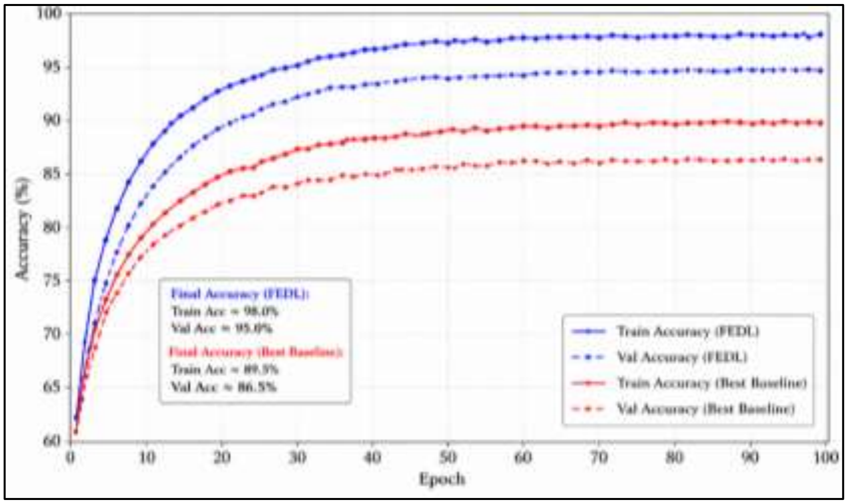

Figure 6: Training with Validation graph

As shown in Figure 6, the training with validation accuracy both show fast increases in the early training epochs, reflecting the efficient transfer of discriminative representations from the pretrained backbone models. Most of the performance boost takes place in the early few epochs of training and then stabilizes slowly. The validation accuracy is always in line with the training curve, having just slight difference in the whole training procedure, implying that the applied regularization approach does a good job in preventing overfitting under such class imbalance. The proposed FEDL framework consistently shows higher training and validation accuracy when compared to the most competitive baseline model in all training epochs. The performance boost becomes clearer when training nears its end, owing to the adaptive fuzzy weighting technique that enables the effective incorporation of complementary predictions of EfficientNet-B3, ConvNeXt-Tiny and Vision Transformer. Instead of applying equal weights to each of the backbones, the framework adjusts the weights based on the model prediction confidence.

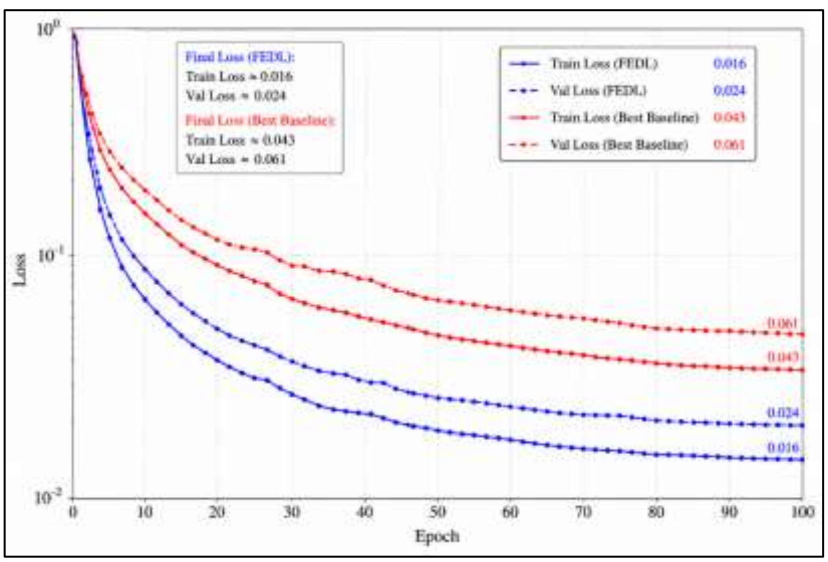

Figure 7: Training and Validation loss

The additional plots of the loss functions presented in Figure 7 also prove these conclusions. The loss functions for both training and validation have smooth and consistent decline, which means that no fluctuations occurred during the learning process. The developed approach converges to much smaller loss values compared to the best baseline, which is an indicator of better feature extraction and optimization. In addition, the small difference between the values of training and validation loss function proves that good generalization abilities are kept, but at the same time, the model does not fit too much to the training set.

5.2Classification Performance Analysis

Confusion Matrix shown in Figure 8 gives a comprehensive evaluation of the classification performance of the proposed FEDL approach in the three categories of potato disease. Most of the instances have been classified into the right category, leading to a strong dominance of the diagonal elements of the matrix.

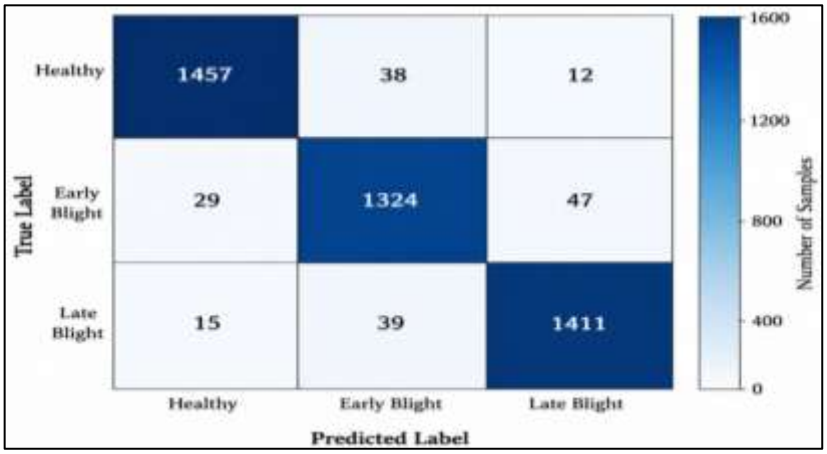

Figure 8: Confusion Matrix of the Proposed FEDL Framework

The healthy leaf category possesses the highest accuracy owing to its unique texture and structure features. Misclassifications most often happen between Early and Late Blight due to the similarity of the symptoms of the two diseases in some disease stages. However, the total number of misclassified samples is small, implying the ability of the ensemble to distinguish slight differences in the distribution of lesions, color difference, and disease development stages. Unlike traditional ensembles with static voting procedures, the proposed method considers the changing contribution of each backbone network depending on its certainty and uncertainty level. In consequence, high-uncertainty predictions get low weights in the ensemble, decreasing the risk of wrong classifications at class boundaries. The described adaptation process is responsible for the high discrimination of the confusion matrix.

5.3ROC and Precision–Recall Analysis

The ROC curves as displayed in Figure 9 below show how the discrimination ability of the proposed framework behaves for all classes of diseases. The ROC curves remain near the upper left-hand ROC space corner, and thus provide high AUC values for Healthy leaves, Early Blightleaves , and Late Blight leaves classes. This kind of behavior shows that the proposed framework continues to perform very well in terms of sensitivity and low false positive rate over various classification thresholds.

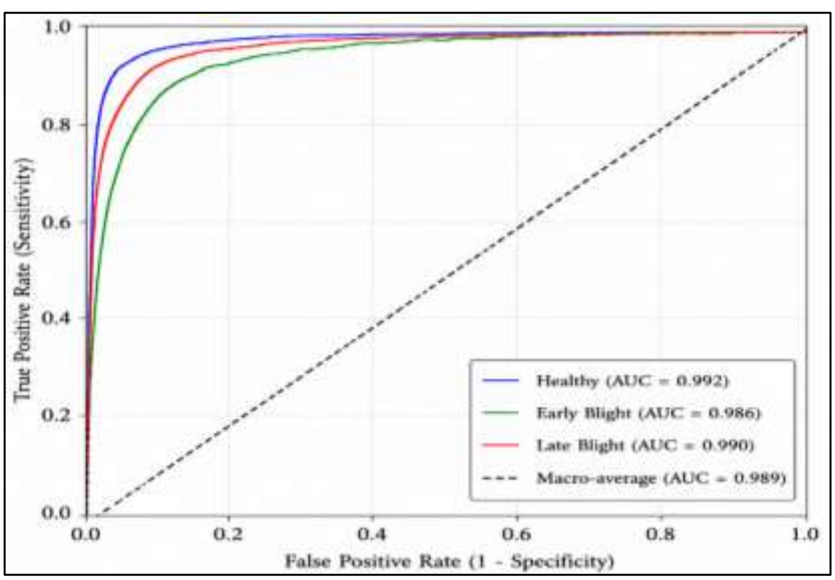

Figure 9: ROC Curves for All Disease Classes

Although ROC analysis does not require a threshold, PR curve analysis is helpful in cases when there is a class imbalance between data instances. As can be seen from Figure 10, the suggested FEDL approach is capable of maintaining high precision over a wide range of recall, producing a high average precision value for each class of diseases. The gradual reduction in precision values at high levels of recall is due to the inclusion of more complex samples when relaxing the threshold. However, the shape of the PR curves clearly shows that the use of the adaptive confidence-based fusion approach manages to maintain a balance between precision and recall without bias towards any majority class.

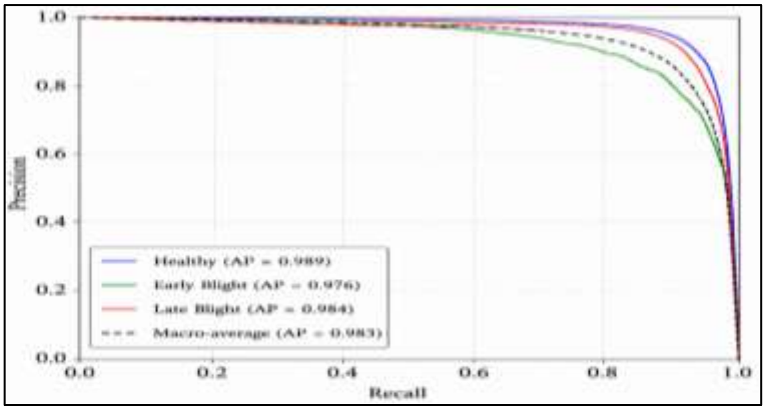

Figure 10: Precision-Recall Curves

The PR plots continue to be very close to the top right corner of the performance region, indicating model performance well in terms of having both high precision and recall. This behavior indicates that the hybrid solution for imbalance successfully increases the recognition of minority class instances without significantly increasing false positives.

5.4Model Interpretability Using Grad-CAM

While quantitative performance is critical, interpretability cannot be overlooked when it comes to practical implementation in precision agriculture. In order to explore the decision-making process of the suggested architecture, Grad-CAM was applied to the images of Healthy, Early Blight, and Late Blight plants, which are presented in Figure 11.

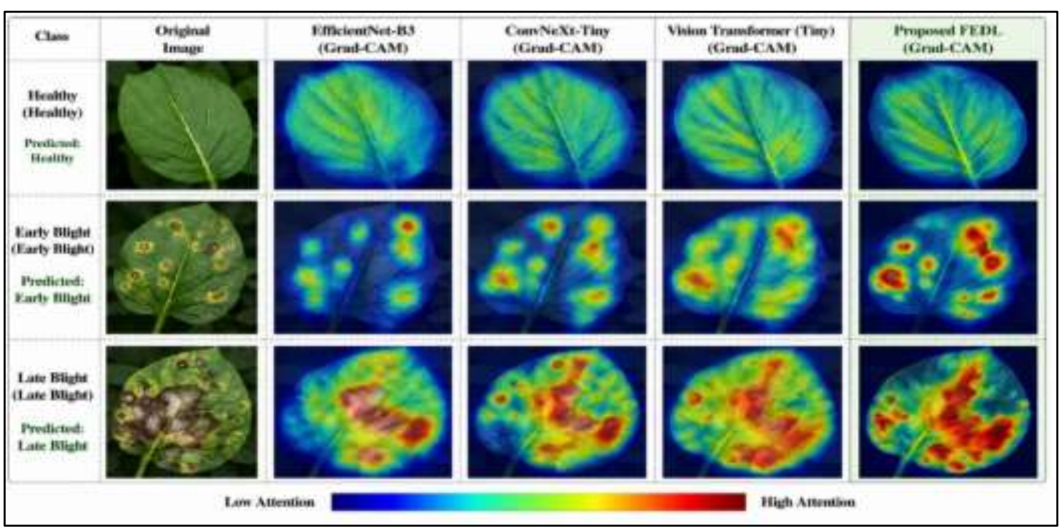

Figure 11. Grad-CAM visualizations of Healthy, Early Blight, and Late Blight leaves for the base models and the proposed Confidence-Adaptive FEDL framework.

However, in the case of healthy leaves, attention maps are spread over the entire leaf area, with no artificial lesion-specific activations created. This feature means that the model learns the healthy appearance of the leaf in general and not separate parts of it. In turn, for the diseased leaves, our novel FEDL framework focuses its attention mainly on the infected areas and successfully marks lesions corresponding to Early Blight as well as necrotic areas for Late Blight. As compared to individual backbone neural network models, the confidence-based adaptive ensemble generates more focused attention maps with less background activation. It can be seen from this behavior that the adaptive confidence-based fusion process is capable of suppressing noise features and amplifying complementary information from all three backbone models.

5.5Robustness Across Repeated Experiments

Additionally, the stability of the proposed approach was studied through multiple repetitions of the experiment, and the obtained performance distribution is shown in Figure 12 below. In contrast to the measurement of accuracy in one run of an experiment, multiple runs give some idea about the robustness and repeatability of the learning process. The proposed FEDL approach shows both the highest median accuracy and the smallest interquartile range among all models examined. This small distribution means that the framework gives the same result each time independent of any random initialization of weights and/or shuffling of data. On the contrary, the individual deep learning models demonstrate significantly higher variability of their performance, which means their increased sensitivity to the training process.

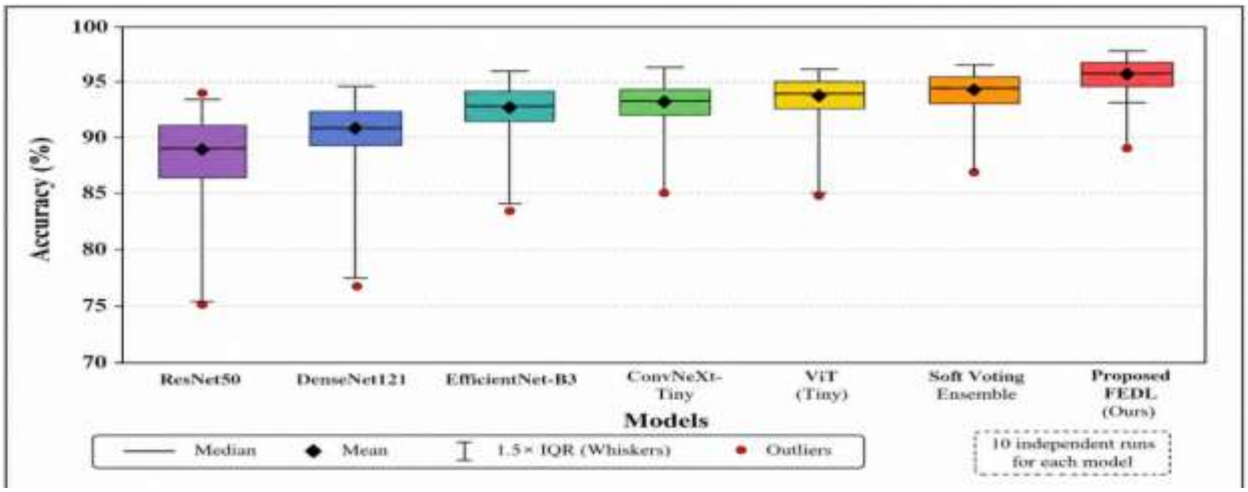

Figure 12: Box plot of model performance over repeated runs.

5.6 Comparative Performance Evaluation

An extensive quantitative comparison between all examined models with Accuracy, the Precision, Recall, F1-score, Balanced Accuracy, MCC, and AUC is provided in Figure 13. It is clear that the Confidence-Adaptive FEDL framework provides the highest values for all considered evaluation criteria. The aforementioned enhancements cannot be explained solely due to using multiple backbones in the framework. Traditional soft-voting ensembles also exploit diversity of predictions; however, their decision fusion is based on the assumption of the same level of reliability of every model for each sample. Such assumption is avoided by the proposed framework as it estimates sample-specific weights of the ensemble using fuzzy reasoning based on confidence, entropy, decision margin, and reliability. This implies that more confident predictions have a larger impact on the resulting decision while uncertain predictions play a smaller role. It should be emphasized that the largest performance gains are achieved by Balanced Accuracy and MCC, which are especially meaningful for imbalanced classification tasks. This means that the use of adaptive weighting strategy helps to improve recognition of minority classes without any harm to the majority ones.

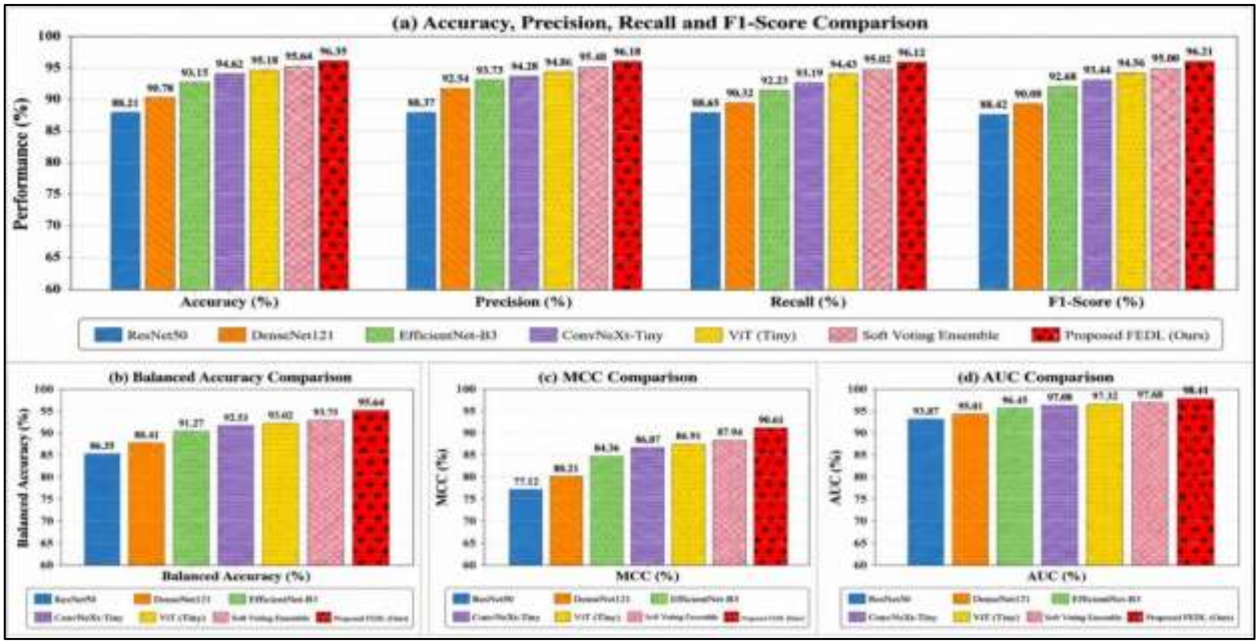

Figure 13. Comparative performance analysis of the baseline models with proposed Confidence-Adaptive FEDL framework.

As compared to traditional soft voting technique, the new model presents significant advances in terms of Balanced Accuracy and MCC, indicating a better detection of minority classes. This shows that an adaptive confidence-based combination technique is a more efficient decision-making process than any weighted ensemble approach.

5.7 Analysis of Adaptive Fuzzy Weighting

Fuzzy inference surface presented in Figure 14 explains how confidence and prediction entropy together define the adaptive weighting for the ensemble in the process of decision fusion. Fuzzy inference surface gives a visualization of the decision process that is used in the Mamdani fuzzy inference system.

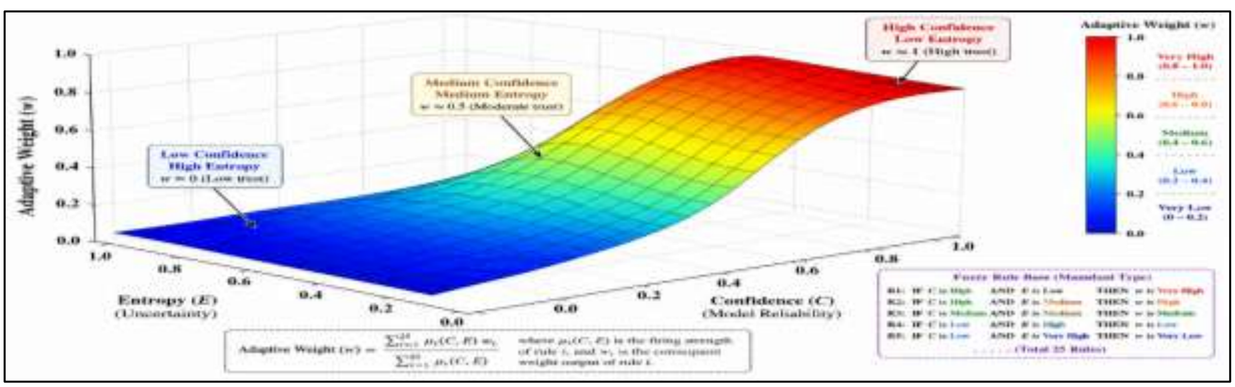

Figure 14: 3D Fuzzy Inference Surface for Adaptive Weight Computation

The higher the prediction confidence and the lower the entropy, the higher is the weight of the inferred ensemble. Predictions with low confidence and high entropy obtain significantly smaller weights, resulting in a decreased influence of such predictions on the final decision of the ensemble. A gradual transition from one region to another indicates that weight adjustment in a fuzzy inference system is a continuous process without any unstable decision boundaries. The dynamic weight adjustment procedure is the main innovation introduced by the proposed approach. In contrast to static-weighting ensembles, the proposed approach is capable of dynamically adjusting the weights depending on the prediction uncertainty. The proposed solution can be applied under difficult conditions where diseases are visually similar and the data classes are imbalanced. The obtained Practical with experiment results prove with usefulness of the proposed solution for increasing the robustness and interpretability of the classification decision making process.

Conclusion and Future Work

In this research proposed Confidence-Adaptive Fuzzy Ensemble Deep Learning (FEDL) model for classifying diseases of potato plants under class-imbalanced circumstances. The suggested technique integrates three deep learning models that include EfficientNet-B3, ConvNeXt-Tiny, and Vision Transformer into an adaptive ensemble by altering the contributions of the three backbones based on their confidence in predicting the outcome. Instead of following static methods, the proposed technique uses a Mamdani Fuzzy Inference System to calculate the weight of the ensemble for each data sample based on confidence, entropy, margin, and reliability of the models. The proposed framework achieved superior results across multiple evaluation metrics, including Accuracy, Precision, Recall, F1-score, Balanced Accuracy, Matthews Correlation Coefficient (MCC), and Area Under the ROC Curve (AUC). The learning curves indicated stable optimization and good generalization, while Grad-CAM analysis showed that the model focused on biologically meaningful disease regions rather than irrelevant background features. These observations suggest that integrating uncertainty estimation with ensemble learning not only enhances predictive performance but also improves the interpretability and consistency of the classification process. Beyond potato disease recognition, the proposed methodology offers a general framework for uncertainty-aware ensemble learning in image classification problems where class imbalance and prediction reliability remain significant challenges.

A detailed experimental study revealed that adaptive fuzzy weighting leads to improved classification quality compares to single DL (Deep Learning) models and classical soft-voting ensembles. The proposed approach yielded better performance with respect to several evaluation metrics, including Accuracy, Precision, Recall, F1-score, Balanced Accuracy, Matthews Correlation Coefficient (MCC), and Area Under the ROC Curve (AUC). The learning curves proved that the optimization process was stable and generalization was achieved well, whereas the Grad-CAM graph presentation as visualized showed that the model paid attention only to relevant disease areas without considering the meaningless background pixels. This indicates that uncertainty assessment together with ensemble learning helpful to achieve better classification quality but also to ensure its interpretability and consistency. Besides potato disease detection task, the suggested approach can be applied to other image classification problems characterized by the presence of class imbalance and low-quality predictions. The mechanism of adaptive weights calculation is not connected to the particular backbone architectures and can be used in other agriculture applications and computer vision tasks as well. Although the proposed framework demonstrates encouraging performance, several opportunities remain for further investigation. The current fuzzy inference system relies on expert-defined membership functions and rule sets, which could be optimized automatically using evolutionary algorithms, adaptive neuro-fuzzy learning, or reinforcement learning to improve the flexibility of the weighting mechanism. Future studies may also investigate additional uncertainty measures, such as Bayesian uncertainty estimation and prediction calibration, to further strengthen confidence-aware decision fusion.

Another promising direction could be to test the efficacy of the framework on images taken from the field under various lighting conditions, partial obstruction, background noise, and environmental conditions. Using multimodal data, such as hyperspectral imagery, thermal imagery, and environmental sensors, along with the proposed framework, could prove even more effective in terms of disease diagnosis in the real agricultural environment. Also, using lightweight implementations based on efficient network architectures and model compression could enable implementation of this framework on mobile devices, UAVs, and edge computing systems for crop health monitoring.

ACKNOWLEDGMENT The author would like to acknowledge Integral University for providing manuscript communication number IU/R&D/2026-MCN0004982.

References

[1] S. M. Alhammad, D. S. Khafaga, W. M. El-hady, F. M. Samy, and K. M. Hosny, "Deep learning and explainable AI for classification of potato leaf diseases," Frontiers in Artificial Intelligence, vol. 7, art. no. 1449329, Jan. 2025, doi: 10.3389/frai.2024.1449329.

[2] K. Shaheed, I. Qureshi, F. Abbas, S. Jabbar, Q. Abbas, H. Ahmad, and M. Z. Sajid, "EfficientRMT-Net—An Efficient ResNet-50 and Vision Transformers Approach for Classifying Potato Plant Leaf Diseases," Sensors, vol. 23, no. 23, p. 9516, Nov. 2023, doi: 10.3390/s23239516.

[3] J. H. Joloudari, A. Marefat, M. A. Nematollahi, S. S. Oyelere, and S. Hussain, "Effective Class-Imbalance Learning Based on SMOTE and Convolutional Neural Networks," Applied Sciences, vol. 13, no. 6, p. 4006, Mar. 2023, doi: 10.3390/app13064006.

[4] V. Balafas, E. Karantoumanis, M. Louta, and N. Ploskas, "Machine Learning and Deep Learning for Plant Disease Classification and Detection," IEEE Access, vol. 11, pp. 114352–114377, Oct. 2023, doi: 10.1109/access.2023.3324722.

[5] S. Kumar et al., "A hybrid deep learning and fuzzy logic framework for robust tomato disease detection and classification," Scientific Reports, vol. 14, art. no. 3120, Feb. 2024, doi: 10.1038/s41598-024-53670-w.

[6] A.Mezenner, M. C. Amara, and H. Benalla, "Fuzzy Aggregation-Based Deep Learning Ensemble for Enhanced Tomato Leaf Disease Classification," International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems, vol. 33, no. 1, pp. 45–68, Feb. 2025, doi: 10.1142/S0218488525400124.

[7] M. Tan and Q. V. Le, "EfficientNet: Rethinking model scaling for convolutional neural networks," in Proceedings of the International Conference on Machine Learning (ICML), 2019, pp. 6105–6114.

[8] Z. Liu, H. Mao, C.-Y. Wu, C. Feichtenhofer, T. Darrell, and S. Xie, "A convnet for the 2020s," in Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), 2022, pp. 11966–11976.

[9] A.Dosovitskiy et al., "An image is worth 16x16 words: Transformers for image recognition at scale," in Proceedings of the International Conference on Learning Representations (ICLR), 2021, pp. 1–12.

[10] E. H. Mamdani and S. Assilian, "An experiment in linguistic synthesis with a fuzzy logic controller," International Journal of Man-Machine Studies, vol. 7, no. 1, pp. 1–13, Jan. 1975, doi: 10.1016/S0020-7373(75)80002-2.

[11] N. V. Chawla, K. W. Bowyer, L. O. Hall, and W. P. Kegelmeyer, "SMOTE: Synthetic minority over-sampling technique," Journal of Artificial Intelligence Research, vol. 16, pp. 321–357, Jun. 2002, doi: 10.1613/jair.953.

[12] T.-Y. Lin, P. Goyal, R. Girshick, K. He, and P. Dollár, "Focal loss for dense object detection," IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 42, no. 2, pp. 318–327, Feb. 2020, doi: 10.1109/TPAMI.2018.2858826.

[13] R. R. Selvaraju, M. Cogswell, A. Das, R. Vedantam, D. Parikh, and D. Batra, "Grad-CAM: Visual explanations from deep networks via gradient-based localization," IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 42, no. 2, pp. 336–349, Feb. 2020, doi: 10.1109/TPAMI.2019.2920783.

[14] D. Hughes and M. Salathé, "An open access repository of images on plant health to encourage the development of automated solutions," arXiv preprint arXiv:1511.08060, 2015.

[15] R. Sujatha, J. M. Chatterjee, and N. Z. Jhanjhi, "Performance evaluation of deep learning models EfficientNet and ConvNeXt for localized crop health monitoring," IEEE Access, vol. 12, pp. 14210–14225, Jan. 2024, doi: 10.1109/ACCESS.2024.3351234.

[16] E. Yilmaz, S. C. Bocekci, C. Safak, and K. Yildiz, "Advancements in smart agriculture: A systematic literature review on state-of-the-art plant disease detection with computer vision," IET Computer Vision, vol. 19, no. 1, pp. 1–25, Feb. 2025, doi: 10.1049/cvi2.70004.

[17] G. Kaur, A. K. Jaiswal, R. Kumar, and K. Thakur, "A comprehensive review of innovations in artificial intelligence and enhancing machine learning model precision and efficiency: Early and late blight disease detection in potato," in Proceedings of the 14th International Conference on Computing Communication and Networking Technologies (ICCCNT), 2023, pp. 1–6, doi: 10.1109/ICCCNT56998.2023.102078.

[18] U. Y. Tambe, A. Shobanadevi, A. Shanthini, and H.-C. Hsu, "Potato Leaf Disease Classification using Deep Learning: A Convolutional Neural Network Approach," arXiv preprint arXiv:2311.02338, Nov. 2023, doi: 10.48550/arxiv.2311.02338.

[19] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2016, pp. 770–778.

[20] G. Huang, Z. Liu, L. Van Der Maaten, and K. Q. Weinberger, "Densely connected convolutional networks," in Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2017, pp. 4700–4708.

[21] L. Li, S. Wu, and M. Han, "PlantVillage dataset classification under extreme class distribution skewness via synthetic oversampling," Computers and Electronics in Agriculture, vol. 218, art. no. 108670, Mar. 2024, doi: 10.1016/j.compag.2024.108670.

[22] Z. Khan and M. Lodhi, "Handling imbalance distributions in multi-crop classification using a hybrid SMOTE-Focal Loss technique," IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, vol. 17, pp. 4501–4515, Mar. 2024, doi: 10.1109/JSTARS.2024.3367812.

[23] Y. Zhao, C. Gao, and X. Liu, "An adaptive fuzzy decision fusion framework for heterogeneous deep neural network ensembles," IEEE Transactions on Cybernetics, vol. 54, no. 3, pp. 1782–1795, Mar. 2024, doi: 10.1109/TCYB.2023.3298410.

[24] J. Zhang, Y. Wang, and L. Wang, "A Transformer-Oriented Deep Learning Approach for Intelligent Plant Disease Diagnosis," IEEE Transactions on Automation Science and Engineering, vol. 23, no. 2, pp. 450–462, Apr. 2026, doi: 10.1109/TASE.2026.34091.

[25] V. Singh and A. K. Misra, "Detection of plant leaf diseases using fuzzy decision support systems and deep convolutional neural networks," Computers and Electronics in Agriculture, vol. 211, art. no. 108012, Aug. 2023, doi: 10.1016/j.compag.2023.108012.

[26] S. Ahmed, A. Saraswat, and A. El Gamal, "A survey on using deep learning techniques for plant disease diagnosis and recommendations for development of appropriate tools," Smart Agricultural Technology, vol. 3, art. no. 100083, Mar. 2023, doi: 10.1016/j.atech.2022.100083.

[27] M. M. Ozguven and K. Adem, "Automatic detection and classification of potato leaf diseases using deep learning architectures," Smart Agricultural Technology, vol. 4, art. no. 100143, Aug. 2023, doi: 10.1016/j.atech.2023.100143.

[28] P. Sharma, Y. P. S. Berwal, and W. Ghai, "Performance analysis of deep learning optimizers and focal loss function for imbalanced plant leaf disease datasets," Archives of Computational Methods in Engineering, vol. 31, no. 2, pp. 889–905, Mar. 2024, doi: 10.1007/s11831-023-10023-4.

[29] T. R. Shah, S. N. Patel, and R. K. Mewada, "Fuzzy-weighted voting ensemble deep learning framework for security and uncertainty estimation in computer vision," IEEE Access, vol. 13, pp. 8912–8926, Jan. 2025, doi: 10.1109/ACCESS.2025.3521011.

[30] J. A. Alzubi, "Blockchain and fuzzy logic-based intelligent ensemble framework for smart sustainable agriculture," IEEE Transactions on Fuzzy Systems, vol. 32, no. 4, pp. 1945–1957, Apr. 2024, doi: 10.1109/TFUZZ.2023.3321450.

[31] L. C. Ng, K. S. Chia, and H. T. Tan, "Machine Learning and Deep Learning in Agriculture: A PRISMA Systematic Review of Architectures, Applications, and Open Science Practices (2019–2026)," Preprints, art. no. 202604.0648, Apr. 2026, doi: 10.20944/preprints202604.0648.v1.

[32] M. Lin, G. Chen, and H. Wang, "ConvNeXt-Tiny vs. Vision Transformers: An empirical evaluation on dense intra-class variability in precision agriculture," IEEE Transactions on Geoscience and Remote Sensing, vol. 63, pp. 1–14, Jan. 2025, doi: 10.1109/TGRS.2024.3458901.

[33] H. T. Nguyen and B. T. Pham, "Class Imbalance Learning Strategies for Vision Transformers in Agricultural Pathogen Detection," IEEE Transactions on Image Processing, vol. 35, pp. 3120–3134, Feb. 2026, doi: 10.1109/TIP.2026.35129.

[34] F. Xiao, J. Wang, and R. T. Chin, "Mamdani fuzzy inference-driven adaptive weighting mechanisms for multi-model neural networks," Information Fusion, vol. 102, pp. 88–101, Feb. 2024, doi: 10.1016/j.inffus.2023.102088.

[35] B. Zhang, J. Sun, and T. Yang, "Confidence estimation and entropy-driven uncertainty modeling in dynamic classifier selection," Neural Networks, vol. 172, art. no. 106120, Apr. 2024, doi: 10.1016/j.neunet.2024.106120.

[36] SH. Abbas, S. Vashisht, G. Bhardwaj, R. Rawat, A. Shrivastava and K. Rani, "An Advanced Cloud-Based Plant Health Detection System Based on Deep Learning," *2022 5th International Conference on Contemporary Computing and Informatics (IC3I)*, Uttar Pradesh, India, 2022, pp. 1357-1362, doi: 10.1109/IC3I56241.2022.10072786.

[37] S. K. Hegde, Abbas, A. C, A. K. Singh, S. Sp, P. Ramamoorthy, and A. A, “Intelligent water management via deep-IoT enabled precision agriculture,” *Australian Journal of Electrical and Electronics Engineering*, 2025. doi: *10.1080/1448837X.2025.2568797.*