International Journal of
INTELLIGENT SYSTEMS AND APPLICATIONS IN ENGINEERING

ISSN:2147-6799 www.ijisae.org Original Research Paper

# Hybrid Approach for Biotic Stress Severity Estimation using Random Forest Regressor

**Leelavathy B[1]*, Ram Mohan Rao Kovvur [2]**

**Submitted**:14/03/2024 **Revised**: 29/04/2024 **Accepted**: 06/05/2024

**Abstract:**Biotic stresses significantly impact paddy cultivation, necessitating accurate and timely disease diagnosis. Manual supervision might not capture the early symptoms of the biotic stress associated with certain diseases. Hence, it requires the best disease management strategies that help to perform good practices under paddy crop cultivation. The proposed work presents a novel approach integrating ensemble convolutional neural networks (ES-DeepNet) predictions and disease severity features for enhanced paddy disease classification and severity prediction for the paddy doctor dataset. ES-DeepNet, comprised of VGG16, Xception, MobileNet, and baseline models, effectively extracts robust features thus improving classification accuracy. Disease severity is assessed quantitatively, and extracted features are integrated into the model with ES-DeepNet model predictions using a random forest regressor. The method demonstrated its effectiveness in both disease classification and severity estimation. This study combines deep learning and machine learning techniques, that contribute to precision agriculture by providing a comprehensive solution for paddy disease management.

***Keywords:****ES-DeepNet, Ensembling, Disease Diagnosis, Severity Assessment, Thresholding, Random Forest Regressor.*

## 1. Introduction

Paddy cultivation faces significant challenges from biotic stresses that affect crop health and yield. Accurate identification and classification of these stresses are crucial for effective disease management. Accurate and timely plant disease detection is crucial for agricultural productivity. Traditional methods have limitations in disease identification and classification, resulting in substantial crop losses. Deep learning techniques have emerged as promising solutions, outperforming traditional approaches in terms of accuracy and efficiency. By addressing challenges like image quality and segmentation, these models offer the potential to significantly improve disease management strategies and ensure food security [2]. Image segmentation isolates leaves and disease regions using Grab Cut and thresholding for feature extraction which were classified using SVM, achieving 93% accuracy in differentiating healthy leaves from rot, esca, and leaf blight.[1] Image processing techniques for paddy disease detection are color features, texture analysis, and shape-based descriptors, these methods have shown some success, but their accuracy is often limited by complex disease symptoms and varying environmental conditions. Recent advancements have leveraged deep learning architectures, particularly Convolutional Neural Networks (CNNs), to extract intricate image features automatically. These models have demonstrated superior performance in classifying paddy diseases and estimating disease severity [3]. Deep learning employs multi-layered neural networks to extract intricate patterns from complex data. By iteratively refining internal parameters through backpropagation, it excels in tasks like image, speech, and text processing. Convolutional networks excel at image analysis, while recurrent networks handle sequential data. These advancements have driven groundbreaking results in various fields [9]. Hyperparameter optimization was conducted to enhance model performance. The proposed model achieved a high accuracy of 93.3% in classifying rice leaf diseases, demonstrating its potential for precision agriculture [4]. Clarifying the intricate relationships between layers and their parameters empowers practitioners to design and manipulate CNNs effectively. Through visual explanations, the guide demystifies complex concepts, enabling users to confidently construct and optimize neural networks for various applications [5]. A public dataset consisting of multispectral and RGB images for rice plant disease detection using multi-modal data which includes multispectral images with Red, Green, and Near-Infrared channels. The results demonstrated that incorporating both multispectral and RGB channels as input improves accuracy when compared to using RGB images alone.[6] The AgriDiet framework for detecting plant diseases and classifying disease severity levels. This framework integrates the conventional INC-VGGN for detection and classification, Kohonen-based deep learning networks by employing a multi-variate grab cut algorithm to resolve occlusion issues and ensure effective segmentation. The pre-trained weights and features are transferred to the new network for specific plant disease detection tasks [20-22]. After computing percentage metrics, the enhanced network classifies severity classes in the training sets [7]. The research article [8] focuses on

*[1,2] Vasavi College of Engineering, Hyderabad, India*
*[1]Research Scholar, Osmania University,, Hyderabad, India*
*[1]ORCID ID :https://orcid.org/0000-0002-8729-0332*
*[2] ORCID ID :https://orcid.org/0000-0003-3741-519X*
**Corresponding author's Email: leelapallava@staff.vce.ac.in*

various CNN architectures used for evaluating plant stress, assessing plant development, and determining postharvest quality. It also discusses advancements in imaging classification, object detection, and image segmentation, highlighting state-of-the-art solutions for specific phenotyping tasks. Pre-trained models (VGG16, SqueezeNet, InceptionV3) were employed as feature extractors, followed by machine learning/deep learning classifiers [4].

Deep convolutional neural networks are used in transfer learning to diagnose plant leaf diseases by applying pre-trained models from extensive datasets to our goal. The Inception module of VGGNet [10], Inception are pre-trained models on the ImageNet dataset on paddy crop biotic stresses brown spot, hispa, and leaf blast, ResNetV2, MobileNetV2, EfficientNetB0 [11], Xception, DenseNet169 [12,13], InceptionV3 and ResNet-50 [14] help reduce the number of parameters and computational cost. Experimental results show that the ResNet-50 model achieved the highest average classification accuracy of 92.61% for paddy crop stress. Using digital images, the hybrid CNN ResLeNet model was used to perform image preprocessing, segmentation, augmentation, feature extraction, and classification to identify fruit and leaf diseases in pumpkins [15].

Existing research has explored various feature extraction and classification techniques for plant disease severity assessment. Traditional methods, such as Hu moments, Haralick features, and color histograms, have been combined with classifiers like SVM, Random Forest, and Decision Trees [16, 17]. More recently, deep learning approaches, including CNN-based feature extraction followed by SVM classification, have been employed to predict severity levels based on the leaf area affected [18]. Hyperspectral imaging has also been investigated for assessing rice leaf blast severity, utilizing spectral reflectance metrics and classifiers like SVM and PNN [19]. These studies demonstrate the potential of image analysis and machine learning in developing accurate and efficient tools for disease severity estimation. GoogLeNet, ResNet, ShuffleNet, ResNeXt, and Wide ResNet [24] are used for paddy leaf diseases. three lightweight CNNs—SE-MobileNet, Mobile-DANet, and MobileNet V2—into a new network called Es-MbNet to recognize plant disease types. Convolutional ensemble network to enhance the model's ability to identify minute plant lesion features [25]. Severity classification is achieved by integrating Convolutional Neural Network (CNN) for feature extraction and these features are then inputted into the LSTM layer which potentially increases the classification accuracy [26]. An Ensemble ResNet-EfficientNet Model that effectively classifies crop diseases by balancing network depth, width, and resolution. Ensembling was performed utilizing ResNet, EfficientNet B4, and EfficientNet B7 models [27]. A unique adaptive minimal ensembling technique, which utilizes only two EfficientNet-b0 models, performs ensembling on feature vectors through a trainable layer rather than the traditional aggregation of outputs [29]. However, there is a paucity of studies comprehensively addressing both disease identification and severity quantification within a unified framework. To address the issues in the context of biotic stress in paddy and its severity assessment using an ensemble model where the state-of-art methods stated importance towards its implementation for precision agricultural practices.

In this study, we propose an enhanced approach for multi-class classification of biotic stress in paddy leaves using ES-DeepNet an ensemble model comprising VGG16, Xception, MobileNet, and baseline and improved baseline CNN architectures. ES-DeepNet integrates these diverse CNN models to extract robust features from leaf images, enhancing the model's ability to distinguish between different types of biotic stress.

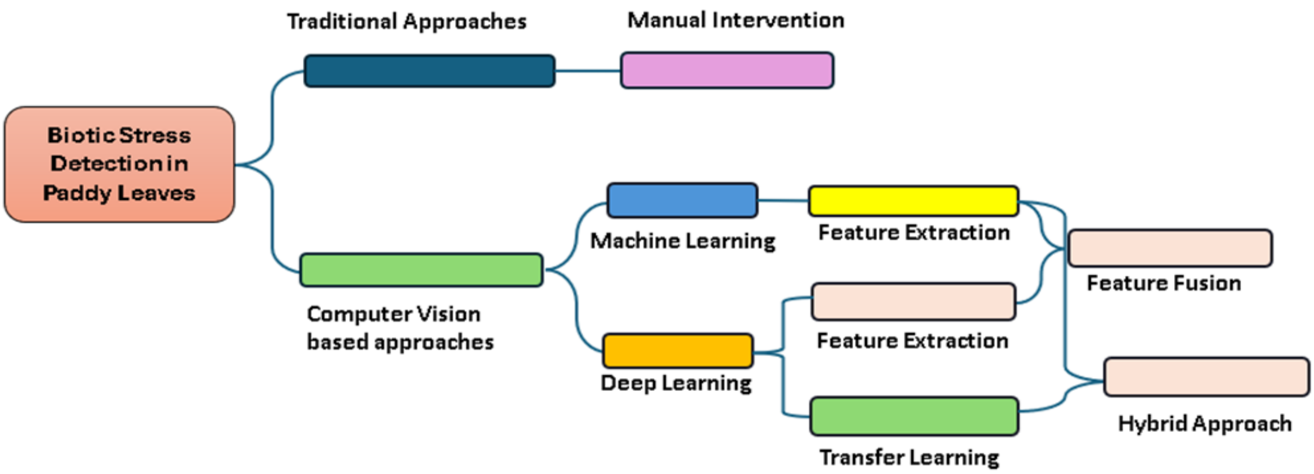


**Fig 1 .** Biotic stress paddy disease detection state-of-the-art methods overview

Additionally, disease severity assessment features, quantitatively measured from the leaf images, are incorporated to improve classification accuracy and provide insights into disease severity levels. The integration of these features is evaluated using a Random Forest Regressor, assessing performance metrics such as Mean Squared Error, Root Mean Squared Error, and R-squared. Performance evaluation of ES-DeepNet focuses on accuracy and F1-score metrics to demonstrate the model's effectiveness in classifying biotic stress types. Meanwhile, the evaluation of the fusion of features with a Random Forest Regressor assesses its capability to predict disease severity levels. Results indicate significant improvements over traditional methods, highlighting the potential of machine learning in advancing precision agriculture practices. This research contributes to the field of agricultural technology by offering a robust framework for automated biotic stress detection and severity assessment in paddy cultivation, supporting sustainable and efficient disease management strategies. Fig. 1 depicts an overview of state-of-the-art methods to address the detection of biotic stress.

## 2. Materials and Methods:

We have a pool of CNN architectures available for image classification and have decided to use pre-trained models such as VGG16, XceptionNet, MobileNet custom models Baseline, and improved Baseline models. Choosing these models for the proposed work on ES-DeepNet was due to several factors that motivated us to use them for the ensemble model. Fig. 2. Corresponds to the deep learning pipeline, emphasizing data preprocessing, model training, ensemble methods, and evaluation.

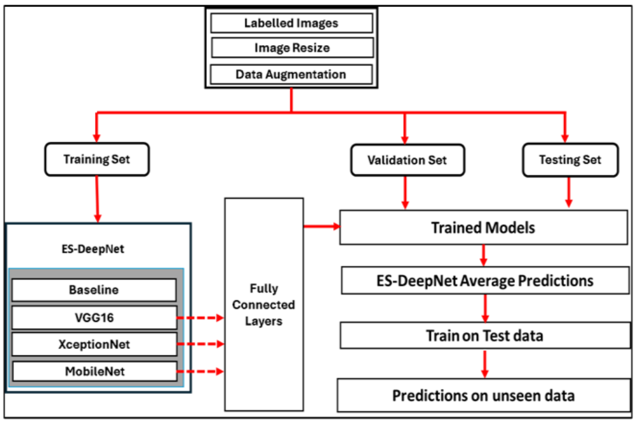


**Fig 2**Architecture of Proposed ES-DeepNet Model**.**

The proposed ES-DeepNet model incorporates the following CNN architectures:

- **VGG16:** Known for its simplicity and depth, VGG16 is effective in feature extraction.
- **Xception:** An extension of the Inception architecture, Xception utilizes depth-wise separable convolutions, making it both efficient and powerful.
- **MobileNet:** Designed for mobile and embedded vision applications, MobileNet is lightweight and efficient, suitable for resource-constrained environments.
- **Baseline Model:** A custom-built CNN tailored for the specific characteristics of the Paddy Doctor dataset.
- **Advanced Baseline Model:** An improved version of the baseline model, incorporating additional layers and regularization techniques for enhanced performance.

### 2.1 Data Acquisition

The Paddy Doctor a benchmark dataset from IEEE data port comprises images of paddy leaves collected from various rice fields, capturing different varieties of leaf diseases. The publicly available dataset includes multiple classes of common paddy leaf diseases, such as Bacterial leaf blight, leaf streak, sheath blight, blast, brown spot, downy mildew, dead heart, hispa, tungro, and healthy. A total of 10,407 images, wherein each image is labeled with the corresponding disease type to facilitate supervised learning.

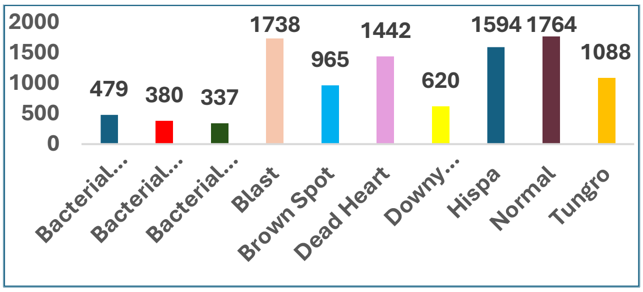


**Fig 3.** Data distribution of disease classes**.**

Fig. 3. shows the data distribution of disease classes.Images were collected from rice fields using high-resolution cameras. These surveys were conducted during different growth stages of the rice plants to capture a variety of disease symptoms. Agricultural experts labeled the images with the correct disease categories, ensuring high-quality and accurate annotations.

### 2.2 Data Preprocessing

Image Augmentation: Techniques such as rotation where this parameter randomly rotates images by up to 20 images, vertical flipping which flips images vertically, and shearing transformations with maximum intensity of 0.2 are applied to increase the diversity of the training set.

Normalization: Pixel values+ are normalized to a range of 0 to 1 to facilitate faster convergence during training.

### 2.3 Transfer Learning

Transfer learning is a technique that is used for attaining knowledge which is solved for one problem to another problem which is of related problems. We use the pre-trained models for image classification on ImageNet data [30], as it has pre-trained weights of more than 14 million images that belong to 1000 generic classes including different species of plants.

### 2.4 Handling Overfitting

Models with too many parameters can easily be overfit as a result, the model performs exceptionally well on the training data but poorly on new, unseen data. To prevent overfitting, we employed augmentation techniques like rotating, flipping, and transformations to increase the dataset size and to stop training when a certain difference between training and validation losses is reached to the defined patience. We employed the training of the model for about 100 epochs and an optimizer Adam was used to train all the models in this proposed work.

### 2.5 Training the models

Instead of learning everything from scratch, the model benefits from the knowledge gained from training on a large dataset. This significantly reduces training time and often improves performance, especially when dealing with limited data. In the diagram, the pre-trained models (VGG16, Xception, MobileNet) represent the transfer learning component. We create a function to generate a sequential model using pre-trained base models to adapt the transfer learning to our biotic stress classification of paddy disease. Fully connected layers are typically added at the end of the deep learning models to map the extracted features to the output classes. To achieve this a Sequential model was implemented by freezing the pre-trained base models flattening the respective model's outputs and then customizing the models by adding dense layers using ReLU and SoftMax activation functions. The purpose of using SoftMax specifically at the output layer as it is specifically designed for multi-class problems, unlike sigmoid which is used for binary classification. it also supports loss function compatibility where it works well with categorical cross-entropy loss, a common loss function for classification tasks. The model is trained with a training set of 80% of the data where it learns patterns and relationships associated with different diseases. For validation, 10% of the dataset is used, wherein it evaluates the model performance to fine-tune hyperparameters as shown in the table and prevent overfitting. For testing 10% of data is reserved for final evaluation of the trained models.

**Table 1. Hyper Parameters used for Training**

| Hyper Parameters | VGG-16 | XCEPTION | MOBILE NET | BASELINE |
|---|---|---|---|---|
| Number of Layers | 8 | 8 | 8 | 14 |
| Pooling | - | - | - | Avg |
| Activation Function | ReLU, SoftMax | ReLU, SoftMax | ReLU, SoftMax | ReLU,SoftMax |
| Batch Size | 32 | 32 | 32 | 32 |
| Number of Epochs | 100 | 100 | 100 | 100 |
| Learning Rate | 0.001 | 0.001 | 0.001 | 0.001 |
| Dropout | - | - | - | 0.3 |
| Optimizer | Adam | Adam | Adam | Adam |

### 2.6 Ensembling and Predictions

Ensembling is a technique that combines multiple models to improve predictive performance. Vgg16, XceptionNet, MobileNet, Baseline, and Improved baseline models are trained, tested, and validated individually on the same dataset [25]. These models are variants of CNNs that are used for ensembling. Each model predicts the given dataset. The predictions from all modules are combined using a specific technique to produce the final predictions.

Common Ensembling Techniques for Combining Predictions obtained by individual models are discussed below:

- **Voting:** For classification, the most common class predicted by the base models is chosen.
- **Averaging:** For regression, the average of the predictions from all models is used.
- **Weighted Averaging:** Similar to averaging, weights are assigned to different models based on their performance.

- **Stacking:** A meta-model is trained to learn the best way to combine predictions.

The predictions from multiple models are combined to improve overall accuracy and robustness. The final predictions are generated by averaging the outputs of the ensemble of models. For the proposed work we have chosen averaging as these predictions are used for regression tasks as they are combined with severity assessment features of disease classes. The average ensemble technique of the predictions often leads to better predictive performance, prevents overfitting, less sensitive to noise in the data when compared to that of individual models. The ES-DeepNet predictions generated by the ES-DeepNet contain the image_id, class_name, class_label. These features are used in combing with disease severity features to perform predictions on unseen data.

Feature Extraction using ES-DeepNet:

1. Ensemble of CNNs: The pre-processed images are fed into an ensemble of CNN models, including VGG16, Xception, MobileNet, Baseline, and Improved Baseline [23, 24].
2. Feature Extraction: Each CNN model extracts features from the input images.
3. Averaging: The extracted features from each CNN are averaged to create a combined feature vector.

**3. Biotic stress severity assessment:**

The disease severity assessment method is implemented based on a comparison of two methods and their features, out of which the best method is selected for the one the results are quite exemplary. Image processing is the promising domain for many image-related real-time applications when it comes to machine learning models or deep learning models. Feature extraction and segmentation techniques play a pivotal role in getting insights from the image. Although there are many techniques available for the proposed work, we implemented using thresholding and image segmentation techniques. It is a fundamental image segmentation technique which is a straightforward and effective approach that partitions an image into two distinct regions based on a predefined intensity level. Essentially, it converts a grayscale image into a binary image, where pixels with intensity values above the threshold are assigned one value (typically white or 1), and those below are assigned another (typically black or 0). The success of thresholding heavily relies on the choice of the threshold value. An optimal threshold effectively separates the objects of interest from the background.

Several methods can be employed to determine the threshold:

- Manual Thresholding: This involves manually selecting a threshold value based on visual inspection of the image histogram which gives the pixel level intensities.
- Global Thresholding: A single threshold value is applied to the entire image. Otsu's method automatically calculates the optimal threshold value by maximizing the inter-class variance between the foreground and background pixels.
- Local Thresholding: Different threshold values are applied to different regions of the image. The threshold value is calculated for each pixel based on its neighbourhood which is called adaptive thresholding.

The severity of biotic stress is implemented using two methods where the first method performs color-based thresholding, and the second method performs disease-specific thresholding as shown in Table 2.

**Table 2**Biotic Stress Severity Assessment Methods**.**

| Method 1: Colour-Based Thresholding (CBT) | Method 2: Disease-Specific Thresholding (DST) |
|---|---|
| Determines disease severity based on the proportion of specific colour pixels (brown and ash) within an image. | Uses predefined colour thresholds for each disease to identify affected areas. |
| Steps:<br>1. Converts image to HSV color space for better color representation.<br>2. Defines color ranges for brown and ash.<br>3. Creates masks for brown and ash pixels using color thresholds.<br>4. Calculate the percentage of brown and ash pixels in the image and combine the percentages to determine overall severity. | Steps:<br>1. Converts image to HSV color space.<br>2. Applies disease-specific color thresholds to create masks.<br>3. Calculate the affected area for each disease.<br>4. Determines severity based on the affected area. |

**3.1 Severity Assessment**

Disease area is calculated by using both thresholding methods, the severity is assessed by following the predefined ranges, and the rating is assigned accordingly as shown in Table 3. Disease severity is rated on an ordinal scale of 0-5 by experts Standard Evaluation System for Rice by IRRI as this ordinal scale is suited for our dataset that has multiple diseases. The scale of ranges are used to assign the feature rating based on the percentage of severity calculated using the CBT, and DST methods. In both the Color-Based Thresholding (CBT) and Disease-Specific Thresholding (DST) methods, severity is calculated as the percentage of the diseased area relative to the total leaf area. The percentage is then mapped to a severity scale, typically ranging from 0 to 5, where:

**Table 3.** Severity scale and descriptions

| Scale | Severity Description |
|---|---|
| 0 | No disease or negligible presence |
| 1 | Early-stage disease, minimal symptoms |
| 2 | Moderate disease progression, visible symptoms |
| 3 | Severe disease has a significant impact on plant health |
| 4 | Very severe disease, extensive damage to plant. |
| 5 | Critical stage, the plant may not recover |

Scale 0: No disease detected.

Scale 1: Very low severity (1-10% of the leaf area affected).
Scale 2: Low severity (10-25% of the leaf area affected).
Scale 3: Moderate severity (25-50% of the leaf area affected).

Scale 4: High severity (50-75% of the leaf area affected).

Scale 5: Very high severity (75% and above of the leaf area affected).

The effectiveness of the severity assessment algorithms was evaluated based on their accuracy in calculating the diseased area and the consistency of the severity ratings across multiple samples. The results showed that the Disease-Specific Thresholding (DST) algorithm, which incorporates unique color thresholds for different diseases, provided more accurate severity assessments

compared to the generic Color-Based Thresholding (CBT) algorithm.

The below Fig 4. displays the severity assessment report generated for the brown spot class where there is an accurate assessment is performed for the DST method when compared to CBT where both the methods reported a scale '0'.

The Fig. 5, infers that the features generated for the test data when the CBT method is used are severity_brown, severity_ash, combined_severity, and scale. When the DBT method is used on the same test data it generates the feature's severity and scale. Thus, the features which are generated by these two methods are further used for an ensemble model where predictions are generated using a random forest regressor.

### 3.2 Random Forest Regressor

The aim is to predict disease severity in paddy crops using a Hybrid Disease Prediction and Severity Estimation using a Random Forest Regressor. A Random Forest Regressor [17] is an ensemble learning method that combines multiple decision trees to make predictions. It can effectively combine numerical features (from ES-DeepNet predictions) with ordinal features (from disease-specific thresholding). By creating numerous trees and averaging their predictions, Random Forest helps to prevent overfitting, a common issue in machine learning. It provides insights into which features contribute most to the prediction. The proposed work is discussed below. The features from both ES-DeepNet predictions and disease-specific thresholding are merged into a single feature vector for each image. The Random Forest model is trained on this combined feature set with corresponding disease severity values as the target variable. The model builds multiple decision trees, each using a random subset of the data and features. For a new image, each decision tree in the forest predicts disease severity. The final prediction is the average of these individual predictions by doing so the quality of the combined features significantly impacts the model's performance.

## 4. Experimental Setup

The proposed work of ES-DeepNet and biotic stress severity assessment method and predictions using combined features is being executed in the verified environment with all dependencies and pre-installed packages Google ColabPro with GPU Memory utilization of 18 GB and disk Space of 38.5 GB.

## 5. Evaluation Metrics

Metrics used for ES-DeepNet are macro F1- score, loss function, and accuracy were calculated for all the training, validation, and test sets. As we got the Combined Features including features extracted from ES-DeepNet predictions and disease-specific thresholds. We used a Random Forest Regressor [31] to predict disease severity based on the combined features. Metrics used to evaluate the regressor model performance are MSE, RMSE, and R-squared. This section gives a detailed description of metrics.

### 5.1 Macro F1-Score

It is the arithmetic mean of the F1-score for each class. we employed the F1-score for all classes that equally contribute to the final score, regardless of several samples in each class. If the dataset is imbalanced, the macro average F1-score shown in Eq. (1), can provide a more informative evaluation than micro average or weighted average.

$$\text{Macro F1-score} = (F1\_BLB + F1\_BLS + F1\_BPB + \ldots\ldots + F1\_TG) / N \quad (1)$$

Where F1_BLB + F1_BLS +F1_BPB+…….. + F1_TG are the F1-scores for each class, N is the total number of classes; in our case, it is N=10.

### 5.2 Loss Function

The loss function is used to measure the model's performance during training. The loss function we used is categorical cross entropy as shown in Eq. (2), suitable for multi-class classification problems like paddy disease detection.

$$\boldsymbol{Loss} = -\sum_{i=0}^{9}\left[y_{true} * \mathbf{log}\left(y_{pred}\right)\right] \quad (2)$$

where y_true is the true label (one-hot encoded), y_pred is the predicted probability distribution, Σ is the summation over all classes where i value ranges from 0-9. In the proposed work all, the models were individually trained where the loss function is calculated to the total number of training samples, where N is the total number of training samples and loss_i is the loss calculated for the i-th sample. Validation loss and testing loss was also calculated in the same manner as done for training as shown in Eq. (3).

$$\textbf{Loss} = \left(\frac{1}{N}\right) * \Sigma(loss_i) \quad \textbf{(3)}$$

### 5.3 Accuracy

We calculated accuracy on the training set, validation, and test set giving the proportion of correctly classified samples.

$$\text{Accuracy} = \frac{TP+TN}{TP+FP+TN+FN} \quad (4)$$

### 5.4 Mean Squared Error

Measures the average squared difference between the predicted and actual values shown in Eq. (5).

$$\text{MSE} = (1/n) * \sum_{i=0}^{9}\left(y_{pred} - y_{true}\right)^2 \quad \textbf{(5)}$$

Where n is the total number of data points, y_pred is the predicted disease severity, y_true is the actual disease severity.

### 5.5 Root Mean Squared Error

The square root of the MSE provides a measure in the same units as the target variable as shown in Eq. (6).

$$\text{RMSE} = \sqrt{MSE} \quad (6)$$

## 6. Experimental Results

### 6.1 Model-wise Performance Analysis

The experimental results insights that VGG-16 Achieves decent accuracy and F1-score, indicating reasonable performance. While XceptionNet Shows improvement in the F1-score compared to VGG-16. The MobileNet has given High training accuracy and F1-score, but performance drops significantly on validation and test sets, where in improved baseline model shows improvement over the baseline model in terms of accuracy and loss, but lags behind more complex models like VGG-16 and XceptionNet in terms of F1-score. The Proposed ES-DeepNet Consistently outperforms all other models in terms of accuracy, loss, and F1-score. Demonstrates strong generalization ability. The Proposed ES-DeepNet model excels in all evaluated metrics. It exhibits high accuracy, low loss, and exceptional F1-score across training, validation, and testing phases. This indicates that the model has effectively learned the underlying patterns in the data and can generalize well to unseen data as shown in Table 4 & Fig. 6.

### 6.2 Combine feature severity predictions

ES-DeepNet Predictions with method 2 use 4 features and show significantly better performance with extremely low MSE and RMSE, suggesting a strong correlation between these features and the target variable. Results are shown in Table 4. Shows the second feature combination outperforms the first one by a significant margin. Reducing the number of features in this case has led to a substantial improvement in model performance. The extremely low MSE and RMSE for the second model indicate a very good fit to the data.

The effectiveness of the proposed fusion of features obtained by ES-DeepNet predictions and the DST algorithm is evaluated using a random forest regressor and compared with that of another algorithm CST algorithm The performance of both DST and CST algorithms is evaluated on the Paddy doctor dataset. Table 4 represents the performance evaluation of metrics on ensemble-based random forest regressor on CST and DST methods.

**Table 5**. Performance Evaluation of Metrics on Hybrid Ensembling with RFR.

| Combined Features | No of Features | MSE | RMSE |
|---|---|---|---|
| Ensemble with RFR - CST | 6 | 0.021 | 0.14 |
| Ensemble with RFR - DST | 4 | 0.014 | 0.12 |

The Ensemble with RFR - DST method has lower MSE (0.014) and RMSE (0.12) values compared to the Ensemble with RFR - CST method which has MSE (0.021) and RMSE (0.14). This indicates that the DST method has a better predictive performance, with smaller errors in the test data.

In Fig. 7 the report consists of the predicted label: tungro, with a severity of 8.67 on, a scale of 1. Most of the predictions were accurate with the DST method compared with that of the CST method.

ES-DeepNet, an ensemble deep learning model, demonstrates superior performance in classification and severity prediction compared to existing methods. Ensemble learning combines predictions from various models to enhance performance.

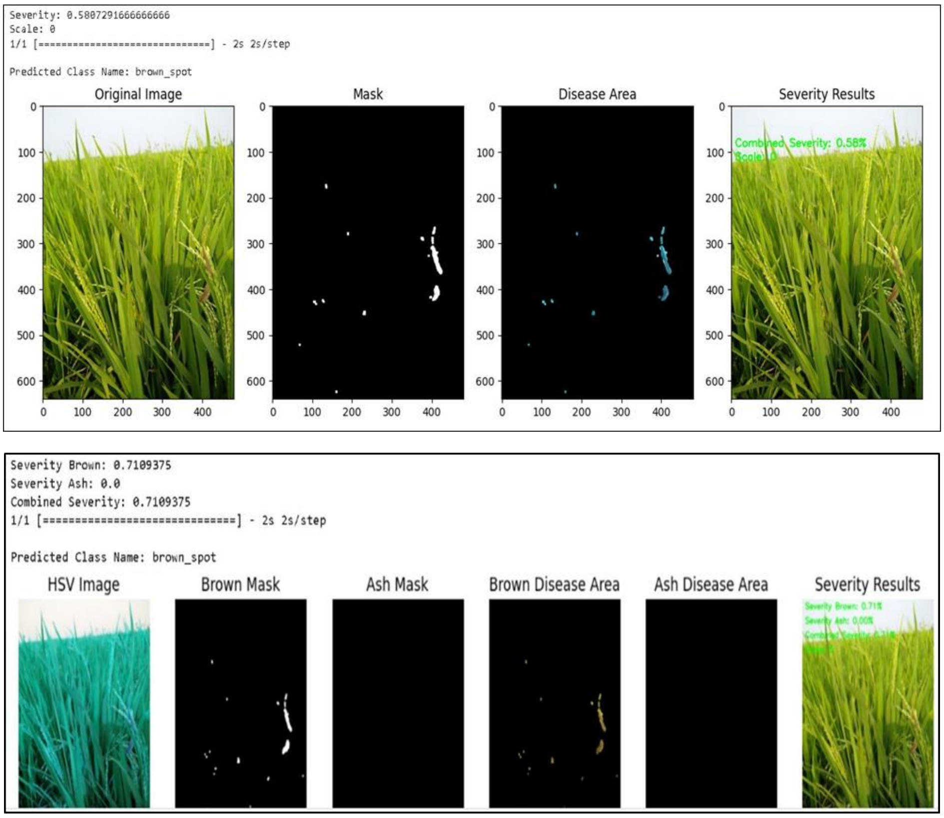


**Fig 4.** Features generated from biotic stress severity assessment methods

**CBT Method**

| | image_id | Severity Brown | Severity Ash | Combined Severity | Scale |
|---|---|---|---|---|---|
| 0 | 203071.jpg | 0.540690 | 0.009766 | 0.550456 | 0 |
| 1 | 202205.jpg | 0.487956 | 0.000000 | 0.487956 | 0 |
| 2 | 202807.jpg | 1.240885 | 0.000000 | 1.240885 | 1 |
| 3 | 203296.jpg | 1.812500 | 1.172852 | 2.985352 | 1 |
| 4 | 202450.jpg | 1.077799 | 0.000000 | 1.077799 | 1 |

**DST Method**

| | image_id | label_predictions | Severity | Scale |
|---|---|---|---|---|
| 0 | 203071.jpg | 2 | 0.540690 | 0 |
| 1 | 202205.jpg | 9 | 0.616536 | 0 |
| 2 | 202807.jpg | 5 | 0.047201 | 0 |
| 3 | 203296.jpg | 1 | 0.622070 | 0 |
| 4 | 202450.jpg | 7 | 0.868164 | 0 |

**Fig 5.** Sample images that are obtained for predicted images of proposed thresholding methods.

**Table 4 . Results obtained for the proposed work.**

| Model Name | Train | Val | Test | Train | Val | Test | Train | Val | Test |
|---|---|---|---|---|---|---|---|---|---|
| | Accuracy | | | Loss | | | F1-Score | | |
| VGG-16 | 0.93 | 0.8 | 0.8 | 0.18 | 0.4 | 0.4 | 0.93 | 0.8 | 0.8 |
| XceptionNet | 0.9 | 0.8 | 0.8 | 0.3 | 0.5 | 0.5 | 0.88 | 0.8 | 0.8 |
| MobileNet | 0.97 | 0.8 | 0.9 | 0.08 | 0.3 | 0.3 | 0.97 | 0.8 | 0.8 |
| Baseline Model | 0.96 | 0.9 | 0.9 | 0.12 | 0.3 | 0.3 | 0.95 | 0.8 | 0.8 |
| ES-DeepNet | 0.98 | 0.98 | 0.98 | 0.12 | 0.12 | 0.14 | 0.98 | 0.9 | 0.9 |

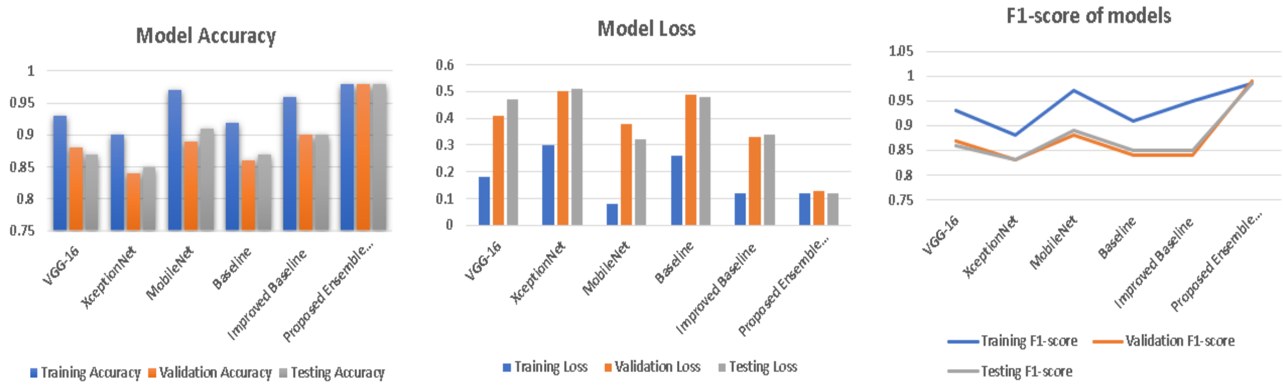


**Figure 4. Visualization of results obtained for the proposed ensembling technique.**

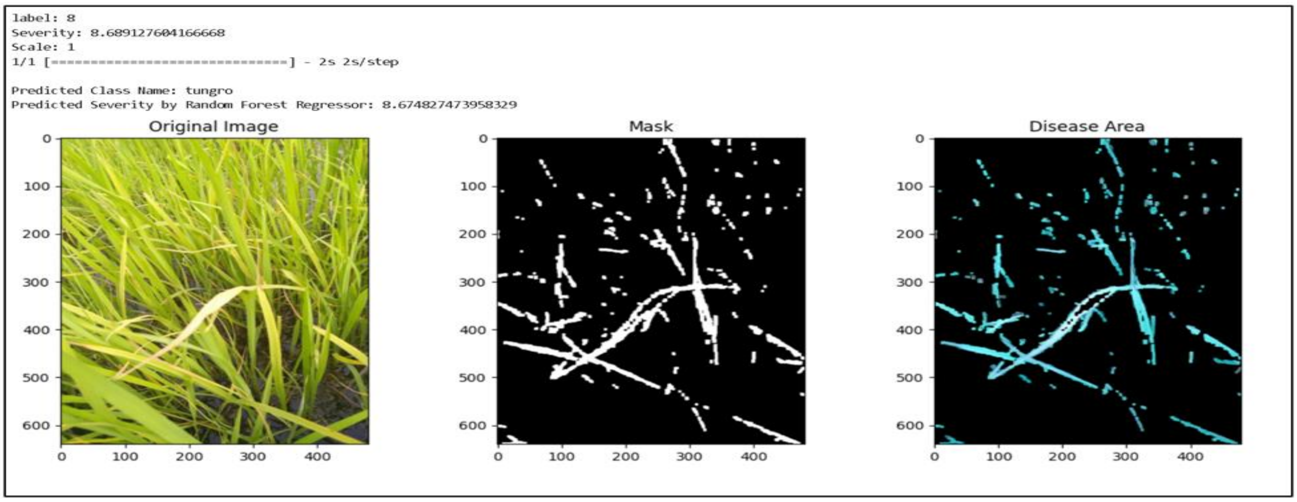


**Figure 5. Severity Assessment Report Using combined ESDeepNet with RFR For Biotic Stress- Tungro**

## 7. Conclusion

This research introduces a novel approach to enhance the biotic stress of paddy for disease diagnosis and severity assessment by synergistically combining ensemble learning and image processing techniques. The proposed ES-DeepNet model, an ensemble of VGG16, Xception, MobileNet, and baseline models, demonstrates superior performance in accurately classifying paddy diseases compared to individual models. The integration of disease severity assessment features, extracted through image processing techniques and analyzed using a Random Forest

Regressor, significantly improves the model's ability to predict disease severity levels. Evaluation metrics, including accuracy, F1-score, mean squared error (MSE), root mean squared error (RMSE) were employed to assess the model's performance. The experimental results consistently demonstrate the superiority of the proposed framework in terms of classification accuracy and severity prediction. The ES-DeepNet model effectively captures intricate patterns within paddy leaf images, leading to precise disease identification. Furthermore, the integration of disease severity assessmentseverity. This multi-model approach helps in extracting complex features from paddy images, enhancing the model’s accuracy and generalization. The ES-DeepNet model was evaluated using the Paddy Doctor dataset, comparing its performance with other ensemble models like AME, CLSTM, and WDM. ES-DeepNet achieved a notable accuracy of 98%. The model also demonstrated lower loss and higher macro F1 scores compared to baseline models. The Ensemble with Random Forest Regressor (RFR) using Disease-Specific Thresholding (DST) showed better performance compared to the Color Specific Thresholding (CST) method, with lower MSE and RMSE. In summary, the chapter presents a robust method for predicting and assessing stress severity in paddy crops by integrating ensemble learning with advanced severity assessment techniques, resulting in improved accuracy and predictive performance. features enhances the model's ability to quantify the extent of disease impact. This research contributes to the advancement of precision agriculture by providing a robust and efficient tool for biotic stress management in paddy. The proposed framework holds the potential to aid farmers in early disease detection, enabling timely interventions and reducing crop losses. However, there are several avenues for further exploration to enhance the system's capabilities like incorporating a more extensive and diverse dataset, including images captured under varying environmental conditions and with different rice varieties, to improve model robustness. The potential of vision transformers can be used for feature extraction and classification. Also, we can develop real-time applications for on-field disease detection and severity assessment using mobile devices or drones by utilizing data related to hyperspectral imaging or weather data, to improve prediction accuracy.

**Conflicts of interest**

The authors declare no conflicts of interest.

**References**

[1] S. M. Jaisakthi, P. Mirunalini, D. Thenmozhi and Vatsala, "Grape Leaf Disease Identification using Machine Learning Techniques," 2019 International Conference on Computational Intelligence in Data Science (ICCIDS), Chennai, India, 2019, pp. 1-6, doi: 10.1109/ICCIDS.2019.8862084.

[2] Jayswal, Hardikkumar& Chaudhari, Jitendra. (2020). Plant Leaf Disease Detection and Classification using Conventional Machine Learning and Deep Learning. 1094-1102.

[3] Ghosal, S., Blystone, D., Singh, A. K., Ganapathysubramanian, B., Singh, A., and Sarkar, S., “An explainable deep machine vision framework for plant stress phenotyping”, Proceedings of the National Academy of Science, vol. 115, no. 18, pp. 4613–4618, 2018. doi:10.1073/pnas.1716999115

[4] Kaur, A., Guleria, K. & Trivedi, N.K. A deep learning-based model for biotic rice leaf disease detection. Multimed Tools Appl (2024). https://doi.org/10.1007/s11042-024-18730-x

[5] Dumoulin, Vincent, and Francesco Visin. "A guide to convolution arithmetic for deep learning." arXiv preprint arXiv:1603.07285 (2016).

[6] Alnaggar, Y.A., Sebaq, A., Amer, K., Naeem, E., Elhelw, M. (2023). Rice Plant Disease Detection and Diagnosis Using Deep Convolutional Neural Networks and Multispectral Imaging. In: Fournier-Viger, P., Hassan, A., Bellatreche, L. (eds) Model and Data Engineering. MEDI 2022. Lecture Notes in Computer Science, vol 13761. Springer, Cham. https://doi.org/10.1007/978-3-031-21595-7_2

[7] Arunangshu Pal, Vinay Kumar, AgriDet: Plant Leaf Disease severity classification using agriculture detection framework, Engineering Applications of Artificial Intelligence, Volume 119, 2023, 105754, ISSN 0952-1976, https://doi.org/10.1016/j.engappai.2022.105754.

[8] Yu Jiang, Changying Li. Convolutional Neural Networks for Image-Based High-Throughput Plant Phenotyping: A Review. Plant Phenomics. 2020;2020:DOI:10.34133/2020/4152816

[9] LeCun, Y., Bengio, Y. & Hinton, G. Deep learning. Nature 521, 436–444 (2015). https://doi.org/10.1038/nature14539

[10] "Junde Chen, Jinxiu Chen, Defu Zhang, Yuandong Sun, Y.A. Nanehkaran, Using deep transfer learning for image-based plant disease identification, Computers and Electronics in Agriculture, Volume 173,2020,105393,ISSN 0168-1699, https://doi.org/10.1016/j.compag.2020.105393"

[11] Hassan, S.M.; Maji, A.K.; Jasiński, M.; Leonowicz, Z.; Jasińska, E. Identification of Plant-Leaf Diseases Using CNN and Transfer-Learning Approach.

Electronics 2021, 10, 1388. https://doi.org/10.3390/electronics10121388

[12] Simhadri, Chinna Gopi &Kondaveeti, Hari. (2023). Transfer Learning for Rice Leaf Disease Detection. 509-515. 10.1109/ICAIS56108.2023.10073711.

[13] Chen J, Zhang D, Nanehkaran YA, Li D. Detection of rice plant diseases based on deep transfer learning. J Sci Food Agric. 2020 May;100(7):3246-3256. doi: 10.1002/jsfa.10365. Epub 2020 Mar 14. PMID: 32124447.

[14] "Naveen N. Malvade, Rajesh Yakkundimath, Girish Saunshi, Mahantesh C. Elemmi, ParashuramBaraki, A comparative analysis of paddy crop biotic stress classification using pre-trained deep neural networks, Artificial Intelligence in Agriculture, Volume 6,2022, Pages 167-175, ISSN 2589-7217, https://doi.org/10.1016/j.aiia.2022.09.001."

[15] YohannesAgegnehuBezabh, BiniyamMulugetaAbuhayi, AlekaMeleseAyalew, HabtamuAyenewAsegie, Classification of pumpkin disease by using a hybrid approach, Smart Agricultural Technology, Volume 7,2024,100398, ISSN 2772-3755,https://doi.org/10.1016/j.atech.2024.100398.

[16] Lamba, S., Kukreja, V., Baliyan, A., Rani, S., & Ahmed, S.H. (2023). A Novel Hybrid Severity Prediction Model for Blast Paddy Disease Using Machine Learning. Sustainability.

[17] Saminathan, K., B. Sowmiya and Devi M Chithra. “Multiclass Classification of Paddy Leaf Diseases Using Random Forest Classifier.” Journal of Image and Graphics (2023):

[18] Naik, H.S., Zhang, J., Lofquist, A. et al. A real-time phenotyping framework using machine learning for plant stress severity rating in soybean. Plant Methods 13, 23 (2017). https://doi.org/10.1186/s13007-017-0173-7

[19] Zhang G, Xu T, Tian Y, Feng S, Zhao D, Guo Z. Classification of rice leaf blast severity using hyperspectral imaging. Sci Rep. 2022 Nov 17;12(1):19757. doi: 10.1038/s41598-022-22074-7. PMID: 36396749; PMCID: PMC9672119.

[20] Jindal, V., Kukreja, V., Mehta, S., Bordoloi, D., & Singh, V. (2023). Severity-Level Assessment of Groundnut Leaf Diseases: A Federated Learning and CNN Approach. 2023 4th IEEE Global Conference for Advancement in Technology (GCAT), 1-6.

[21] Yang, B., Li, M., Li, F. et al. A novel plant type, leaf disease, and severity identification framework using CNN and transformer with multi-label method. Sci Rep 14, 11664 (2024). https://doi.org/10.1038/s41598-024-62452-x

[22] Bedi P, Gole P, Marwaha S. PDSE-Lite: lightweight framework for plant disease severity estimation based on Convolutional Autoencoder and Few-Shot Learning. Front Plant Sci. 2024 Jan 8;14:1319894. doi: 10.3389/fpls.2023.1319894. PMID: 38259916; PMCID: PMC10800669.

[23] M.A. Ganaie, Minghui Hu, A.K. Malik, M. Tanveer, P.N. Suganthan, Ensemble deep learning: A review, Engineering Applications of Artificial Intelligence, Volume 115,2022, 105151, ISSN 0952-1976,https://doi.org/10.1016/j.engappai.2022.105151.

[24] A. Acharya, A. Muvvala, S. Gawali, R. Dhopavkar, R. Kadam and A. Harsola, "Plant Disease detection for paddy crop using Ensemble of CNNs," 2020 IEEE International Conference for Innovation in Technology (INOCON), Bangluru, India, 2020, pp. 1-6, doi: 10.1109/INOCON50539.2020.9298295.

[25] Chen, J., Zeb, A., Nanehkaran, Y.A. et al. Stacking ensemble model of deep learning for plant disease recognition. J Ambient Intell Human Comput 14, 12359–12372 (2023). https://doi.org/10.1007/s12652-022-04334-6

[26] Citation: Lamba, S., Baliyan, A., Kukreja, V., Tripathy, R. (2023). An Ensemble (CNN-LSTM) Model for Severity Detection of Bacterial Blight Rice Disease. In: Marriwala, N., Tripathi, C., Jain, S., Kumar, D. (eds) Mobile Radio Communications and 5G Networks. Lecture Notes in Networks and Systems, vol 588. Springer, Singapore. https://doi.org/10.1007/978-981-19-7982-8_14

[27] Kundur, N. C. ., &Mallikarjuna, P. B. . (2022). Ensemble Efficient Net and ResNet model for Crop Disease Identification. International Journal of Intelligent Systems and Applications in Engineering, 10(4), 378–390. Retrieved from https://ijisae.org/index.php/IJISAE/article/view/2273

[28] M.A. Ganaie, Minghui Hu, A.K. Malik, M. Tanveer, P.N. Suganthan, Ensemble deep learning: A review, Engineering Applications of Artificial Intelligence, Volume 115,2022, 105151, ISSN 0952-1976, https://doi.org/10.1016/j.engappai.2022.105151.

[29] Bruno A, Moroni D, Dainelli R, Rocchi L, Morelli S, Ferrari E, Toscano P, Martinelli M. Improving plant disease classification by adaptive minimal ensembling. Front ArtifIntell. 2022 Sep 8;5:868926. doi: 10.3389/frai.2022.868926. PMID: 36160929; PMCID: PMC9499023.

[30] J. Deng, W. Dong, R. Socher, L. -J. Li, Kai Li, and Li Fei-Fei, "ImageNet: A large-scale hierarchical image

database," 2009 IEEE Conference on Computer Vision and Pattern Recognition, Miami, FL, USA, 2009, pp. 248-255, doi: 10.1109/CVPR.2009.5206848.

[31] Tin Kam Ho, "Random decision forests," Proceedings of 3rd International Conference on Document Analysis and Recognition, Montreal, QC, Canada, 1995, pp. 278-282 vol.1, doi: 10.1109/ICDAR.1995.598994.