International Journal of
INTELLIGENT SYSTEMS AND APPLICATIONS IN ENGINEERING

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

Deep-CNN Based Multi-Class Classification and Accurate Severity Assessment of Biotic Stress in Paddy Leaves

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

Submitted:11/03/2024 **Revised**: 26/04/2024 **Accepted**: 03/05/2024

Abstract: Rice is a staple food for over half of the world's population, particularly in Asia. It's a major source of carbohydrates, providing essential energy for daily activities. Rice cultivation plays a crucial role in the livelihoods of millions of farmers worldwide. It contributes significantly to the GDP of many countries. If left unchecked, biotic stress can cause substantial yield losses, leading to food insecurity and economic hardship for farmers. Early detection and management are crucial for preventing these losses. CNNs are a class of deep learning models well-suited for image classification tasks and can be easily scaled to large datasets and complex classification problems. The Automatic stress severity assessment can save time and resources compared to manual assessments, allowing for more efficiency in terms of decision-making. We proposed a Deep-CNN model, that utilizes the Paddy Doctor dataset with nine stress classes and one healthy class. we also addressed the imbalance in the dataset to avoid overfitting and performed a targeted augmentation technique. Multiple classes were classified and predicted on the proposed model. Extensive experimentation was carried out for tuning the parameters of the model. The proposed model achieved high accuracy of 94.4% while EfficientNetB0 achieved 93.5%. Our findings demonstrate that the model outperformed classification for all the 10 classes of the dataset. Using the predicted image, for every stress class the stress characteristics vary with color, we define color thresholds and apply a mask on the image to evaluate the stressed area and generate a severity report. This research demonstrates a promising solution to combat biotic stress in rice. It offers the potential to revolutionize disease management and empower rice-growing communities worldwide to safeguard their livelihoods and contribute to global food security.

Keywords: *Biotic-Stress, Deep-CNN, Augmentation, Multi-Class Classification, Severity Assessment.*

1. Introduction

Rice is one of the cereal foods called Oryza sativa which is a member of the grass family that includes Oryza glaberrima and Oryza Sativa [1]. It is the primary staple food for most of the South Indian population, providing over 50% of their dietary calorie intake. Rice cultivation is the backbone of many Asian economies, employing millions of people, particularly in rural areas. It contributes significantly to agricultural GDP and export revenues for countries like India, Vietnam, and Thailand. Ensuring a stable and adequate supply of rice is crucial for food security. Rice shortages can lead to price hikes, and widespread hunger, particularly among vulnerable populations. Rice cultivation has significant environmental implications, including water consumption, greenhouse gas emissions (methane from paddy fields), and land usage changes. Sustainable rice production practices are essential to mitigate these impacts. During the regular practices of paddy cultivation, the crops might be prone to biotic stress and abiotic stress. Biotic stress refers to the diminishing impact caused by living organisms on plant health [2]. This could be caused by pathogens like bacteria, fungi, and viruses, pests like insects, and nematodes like worms that can infest plant roots and cause damage [14] diseases that fall under biotic stress in paddy are listed below:

a) Bacterial diseases: Bacterial Leaf Blight, Bacterial Leaf Streak, and Bacterial Panicle Blight.
b) Fungal diseases: Blast, Brown Spot, and Downy Mildew.
c) Insect pests: Dead Heart (caused by stem borers) and Hispa.
d) Virus: Tungro.

Bacterial Leaf Blight(BLB), Bacterial Leaf Streak(BLS), Bacterial Panicle Blight(BPB), Blast(BL), Brown Spot(BS), Dead Heart(DH)(Stem Borer Damage), Downy Mildew(DM), Hispa(HS), Tungro(TG) occurred in paddy, cause a huge loss that in the production of rice. The appearance of stress symptoms might be ambiguous as few stresses might differ with minute variation. This leads to ambiguity when diagnosed manually. The misconception may lead to the wrong treatment of pesticides which rigorously decreases the crop yield, wastage of manpower, and cost spent on those pesticides. Therefore, biotic stress has been quite an important aspect when taking food security and food production. Thus, it has been an immediate requirement to detect these stresses at an early stage. Performing stress detection manually by a pathologist based on visual symptoms may not effectively identify early-stage symptoms. Subsequently, there is a requirement for laboratory tests for analysis and calculation of stress

[1]*Research Scholar, Department of CSE, UCE, Osmania University
ORCID ID : https://orcid.org/0000-0002-8729-0332
[1,2]Department of Information Technology, Vasavi College of Engineering, Hyderabad, India
ORCIDID: https://orcid.org/0000-0003-3741-519X
* Corresponding Author Email: leelapallava@vce.ac.in*

severity. Manual laboratory tests and stress detection are quite expensive and time-consuming. However, these methods are less effective in early-onset stress prediction and stress severity assessment. Hence, preventive measures are to be adopted promptly which can save farmers and result in substantial production of paddy. Several types of stresses can be caused and infected in paddy crops. They can potentially harm the crop at different stages of growth. Several pests/viruses/stresses occur in paddy crops. We have used nine stress classes Bacterial Leaf Blight(BLB), Bacterial Leaf Streak(BLS), Bacterial Panicle Blight

Table 1: Indicative Parameters of Rice Leaf Diseases/ Stresses.

Class	Stress Name	Causal Agent	Color of the Leaf	Stress Region Shape	Stage of Stress	Affected Part	Environmental Conditions
C1	Bacterial Leaf Blight	Xanthomonas oryzae pv. oryzae	Yellow to brown	Initially water-soaked lesions, later turning grayish white with a yellow halo	Tillering to ripening	Leaves	High temperature (25-34°C), high humidity (>90%), wind-driven rain
C2	Bacterial Leaf Streak	Xanthomonas oryzae pv. oryzicola	Brown to black	Narrow, linear to irregular water-soaked streaks, later turning brown with wavy edges	Tillering to booting	Leaves	High temperature (25-30°C), high humidity (>90%), prolonged leaf wetness
C3	Bacterial Panicle Blight	Burkholderia glumae	Pale green to straw-colored	Grains become discolored and chaffy	Panicle emergence to grain filling	Panicles	High temperature (28-36°C), high humidity (>80%), insect damage
C4	Blast	Pyricularia oryzae	Grayish-white with brown margins	Spindle-shaped or diamond-shaped lesions	All growth stages	Leaves, nodes, neck	High humidity (>90%), cool temperature (24-28°C), prolonged leaf wetness

C5	Brown Spot	Bipolaris oryzae	Dark brown	Circular to oval spots with dark brown margins	Seedling to ripening	Leaves	High humidity (>85%), warm temperature (25-30°C), prolonged leaf wetness
C6	Dead Heart (Stem Borer Damage)	Yellow stem borer or striped stem borer (insects)	Yellowing and wilting	Dead heart in young plants, whiteheads in older plants	Tillering to booting	Stem	Suitable temperature and humidity for insect development
C7	Downy Mildew	Sclerospora graminicola	Pale green to yellow	Downward curling leaves with white downy growth on the underside	Seedling to tillering	Leaves	High humidity (>90%), moderate temperature (20-25°C), cloudy weather
C8	Hispa (Rice Hispa)	Dicladispa armiger (insect)	Whitish streaks due to feeding	Irregular feeding patterns on leaf surface	Tillering to booting	Leaves	Warm and dry weather
C9	Tungro	Rice Tungro Bacilliform Virus (RTBV) and Rice Tungro Spherical Virus (RTSV) transmitted by Green Leafhoppers	Yellow to orange yellow	Stunting, discoloration starting from leaf tips, rusty spots	Seedling to ripening	Whole plant	Presence of virus-carrying leafhoppers, favorable weather conditions for insect populations

Blight(BPB), Blast(BL), Brown Spot(BS), Dead Heart(DH) (Stem Borer Damage), Downy Mildew(DM), Hispa(HS), and Tungro(TG) are much more harmful to the crop. These stresses when affected at an early stage of the crop are asymptomatic and not visible to the naked eye. As the severity prolongs from days to weeks or even months. The loss cannot be predictable where the spikelets will be with unfilled

grains at the time of harvest. There is an urge and need to automate the stress detection and stress severity assessment for the early onset of the crop. The responsiveness and popularity of computer vision using Deep Learning techniques for stress identification in rice [7], maize[12], soyabean[9], multi-class classification along with stress severity classification using CNN[14][15][16] and CNN with SVM[13], random forest classifier[16] and prediction with quantification has motivated us to employ these techniques for classification with stress prediction and stress severity assessment. In this research, we implemented a Deep-Convolutional Neural Network for the prediction of these stresses BLB, BLS, BPB, Blast, BS, DH, DM, Hispa, and Tungro for the different varieties of rice like ADT45, Onthanel, Karnataka Ponni, Surya. The work focuses mostly on the disease, type of causal agent, the color of the leaf, the shape of the stressed region, the stage of stress, the part at which the stress is affected, and what could be the environmental aspects of the occurrence of stress as shown in Table 1. The proposed framework can accurately classify and predict normal and stressed images. It also has the potential to detect the area of spread on the predicted image.

Based on the assessment it can predict the stress severity and generate a rating indicating how much percentage of the stress is affected on the specified predicted image. We used an ordinal rating method called Standard Evaluation System for Rice (SES) defined by IRRI[17]. A rating of 0 indicates no severity, likewise, rating-5 indicates highly severe, which means the stress spread on the leaf is more than 75%.

The overview of the proposed work is shown in Fig.1, and the corresponding contributions for the proposed work are listed below:

- Using optimal parameters for training the model.
- Calculate the stress spread on the leaf (disease spread).
- Predicting the disease severity based on the SES ordinal scale.
- Dataset preparation, while we used benchmark dataset from IEEE Data port.
- Developing a novel deep learning framework for stress predictions.

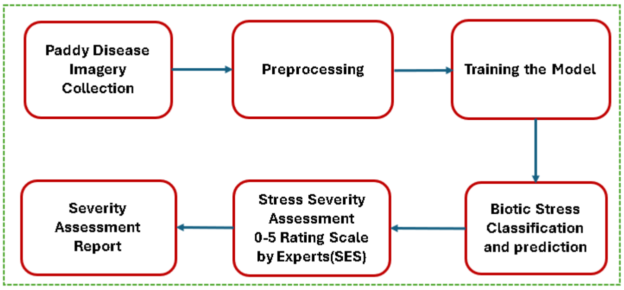


Fig 1. Overview of Proposed Approach

2. Related works

Recognition and Classification of paddy leaf diseases using optimized deep neural network [3] with Jaya Algorithm. A comparative study was performed with ANN, DAE, and DNN and achieved good accuracy for four stress classes. Automatic recognition and classification of biotic and abiotic stress using a deep convolutional neural network framework i.e., VGG-16 using field images of a paddy crop with an accuracy of 92.89% [4]. A two-stage small Convolutional Neural Network architecture [5] was demonstrated which resulted in an accuracy of 93.3%. The state-of-the-art methods in deep learning-based techniques [6] for plant leaf stresses were reviewed with 33 different crops which include vegetables, and fruits using 14 CNN architectures that are used in most of the work. Classification and recognition of rice disease using a hybrid network with a particle swarm optimization algorithm [7] had an accuracy of 94.03%. Disease identification, classification, and recognition from the current state-of-art methods are employed to perform the classification of diseases which alone won't be sufficient to perform severity estimation. The research has to be extended in terms of stress severity estimation based on which we can try to estimate crop loss also in the near future. The error estimates of disease severity in plants can occur which can be solved using Standard Area Diagrams (SADs) which improve accuracy and reliability [8]. Automatic assessment of abiotic stress factor called Iron Deficiency Chlorosis severity on field plots [9] in soybean was performed using classification techniques that help in decision making. One of the biotic stress severities on rice leaf blast using hyperspectral imaging [10] was demonstrated by calculating the standard deviation of the respective spectral reflectance of whole rice leaves with Support Vector Machines and Probabilistic Neural Networks to classify the degree of severity level at different stages of growth. A disease severity classification was implemented as an AgriDiet framework [11] that quantifies the disease severity into mild, moderate, and severe, while the same severity levels were quantified to mild, moderate, severe, and profound using CNN and SVM [12]. Many Studies have addressed the latent of deep learning and machine learning methods for biotic stress classification and stress severity assessment in paddy crops. However, deep learning-based multi-class classification for real-time applications still has a lot of space for improvement, extensive research can be carried out in this context to overcome the limitations and challenges. One of the major challenges to be addressed in the huge data set requirement with annotations, while training this huge data leads to time-consuming. The other challenge is the selection and tuning of model hyperparameters to achieve optimal performance by the proposed model. These challenges prevent the usage of these models for real-life decision-making. To address the above issue, in this research a new framework called Deep-Convolutional Neural Network has been proposed. This model uses a simple CNN architecture with few parameters. Experimental studies infer that the proposed model outperforms state-of-art models like EffiecientNetB0, and VGG-16 for Multi-class Classification and Prediction. A method that employs stress severity automatic assessment and rating is proposed. Further, the visualization of infected regions can be used for decision-making for future analysis of crop estimates.

3. Materials and Methods

In this section, the detailed architecture and working of the proposed framework of Deep-Convolutional Neural Network will be discussed. The high-level overview of the workflow is shown in Fig. The framework is associated with image acquisition, pre-processing, classification, and prediction. Employed a severity assessment technique that rates the stress severity based on a specified scaling method. The description of the activities is elaborated below.

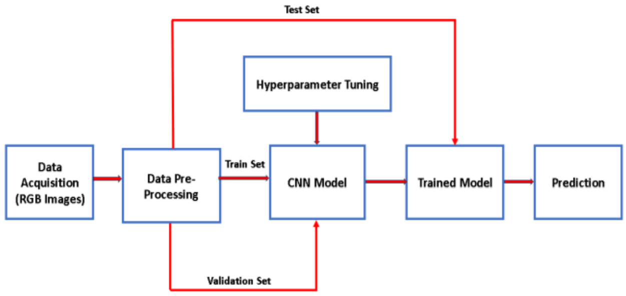


Fig 2. Overview of the workflow.

3.1 Dataset Acquisition

The dataset used for the proposed model is a benchmark dataset that is available at the IEEE Data Port named Paddy Doctor dataset[13]. This dataset consists of 16,225 paddy leaf images with annotations by experts of about 13 classes (12 stress classes and healthy class. The image acquisition was from the paddy fields of Pallamadai, Tamil Nadu. These images were captured from February 2021 to April 2021, with a resolution of 1080x1440 pixels using a smartphone device of CAT S62 Pro. The collected RGB images were cleaned and annotated under the supervision of Agronomist. The dataset is publicly available in the Kaggle competition with 10 classes (9 stress classes and healthy classes). Metadata for every image in the dataset consists of paddy age and variety is being provided. Further, the dataset comprising 10407 images was chosen for the current work. The sample images of the dataset are shown in Fig 3.

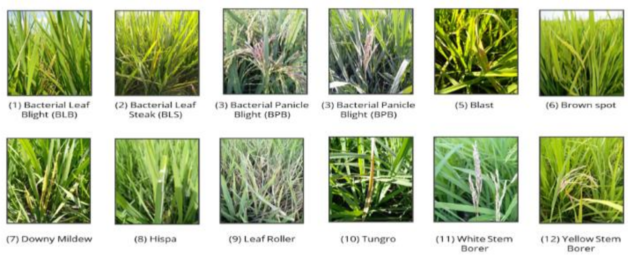


Fig 3. Paddy Doctor Dataset Sample Images

The Collected dataset was accessed through Kaggle-ap, and it was divided into train and validation splits of the ratios 65:35, 70:30, 75:35, and 80:20. The observations noted that the model performs better classification when the ratio is 80:20. We reserved 20% of the images for validation. This is crucial for evaluating the proposed model's performance during training and preventing overfitting. Out of 10,407 images, 8326 images were used for training and 2081 for validation initially without augmentation. The test set consists of 3469 files belonging to 1 class which were used for prediction. After the augmentation out of 12594 images 10,092 were chosen for training, 2522 images were chosen for validation and the test set remained the same.

3.2 Dataset Preparation

The dataset consists of about 10,407 images with imbalance problems among the classes in the dataset. Some diseased classes like bacterial_leaf_blight, bacterial_leaf_streak, bacterial_panicle_blight, downy_mildew, brown_spot, tungro have fewer images than normal, blast, hispa, dead_heart. We first chose the class that had a higher number of images and augmented all the other classes to represent the same number, performing this was not desirable as the model was overfitting. To address this issue targeted augmentation was applied by only focusing on underrepresented classes. Which in turn creates more diverse training data, such that it can improve model performance.

These classes were divided into three classes:

1. Priority classes where the disease classes that need stronger augmentation as their representation is very low in the dataset. Techniques like rotation, shifting, shear transformation, zooming, Flip, and brightness were applied for these classes.
2. Moderate classes are the disease classes that need less intense augmentations like shifting and rotations.
3. Minimal classes are the classes that already have plenty of samples and do not need any augmentation. Class class-wise size of the dataset before augmentation and after augmentation is shown in Table 2.

As the images are stored with pixel values ranging from 0 to 255 (8-bit format), representing the intensity of each color channel of an RGB image. The model typically performs better when input data is normalized to a smaller scale. Rescaling by dividing each pixel value by 255 to normalize the pixel values into the range of 0 to 1, which can lead the model to faster convergence during training.

Table 2: Class-wise dataset size

Class	Original images	Augmented Images	
		Max Num_Class	Targeted Augmentation
normal	1764	1764	1764
blast	1738	1764	1738
hispa	1594	1764	1594
dead_heart	1422	1764	1422

tungro	1088	1764	1333(MC)
brown_spot	965	1764	1212(MC)
downy_mildew	620	1764	869(PC)
bacterial_panicle_blight	479	1764	830(PC)
bacterial_leaf_streak	380	1764	864(PC)
bacterial_leaf_blight	337	1764	968(PC)
Total Images	10407	17640	12594

3.3 Proposed Deep-CNN for classification and prediction:

After pre-processing and augmentation, the images from the Paddy Doctor dataset[13] are subjected for training purposes. The Deep-CNN model is given with diseased samples for the classification task. The proposed sequential model framework consists of the following layers i.e., input, convolution, Average Pooling, dropout, FC Layer, SoftMax, and Output. The architecture of the Deep-CNN Model as shown in Fig 4, is developed for multi-class classification of the dataset comprising of paddy leaves with ten class labels. There is a unique functionality that is performed at every layer of the defined Deep-CNN Model. Relu Activation is used after each convolutional layer to introduce non-linearity and help the network to learn complex patterns. Batch Normalization is used at certain convolutional layers to normalize activations, leading to faster and more stable training. Average Pooling retains the spatial information in the feature maps, which can be beneficial for tasks where precise localization is important. The average pooling layers progressively reduce the spatial dimensions of the feature maps while increasing the number of channels (filters). This allows the network to capture increasingly complex features at different scales. Fully Connected Layers, after flattening, two dense layers are used to combine and interpret the high-level features extracted by the convolutional layers. Dropout layers are added to prevent overfitting.

```
            OPERATION           DATA DIMENSIONS   WEIGHTS(N)   WEIGHTS(%)

                Input   #####    224  224    3
               Conv2D    \|/  -------------------       896     0.0%
                 relu   #####    222  222   32
     AveragePooling2D   Y avg -------------------         0     0.0%
                        #####    111  111   32
               Conv2D    \|/  -------------------     18496     0.4%
                 relu   #####    109  109   64
   BatchNormalization    μ|σ  -------------------       256     0.0%
                        #####    109  109   64
     AveragePooling2D   Y avg -------------------         0     0.0%
                        #####     54   54   64
               Conv2D    \|/  -------------------     73856     1.7%
                 relu   #####     52   52  128
   BatchNormalization    μ|σ  -------------------       512     0.0%
                        #####     52   52  128
     AveragePooling2D   Y avg -------------------         0     0.0%
                        #####     26   26  128
               Conv2D    \|/  -------------------    295168     6.8%
                 relu   #####     24   24  256
   BatchNormalization    μ|σ  -------------------      1024     0.0%
                        #####     24   24  256
     AveragePooling2D   Y avg -------------------         0     0.0%
                        #####     12   12  256
               Conv2D    \|/  -------------------   1180160    27.3%
                 relu   #####     10   10  512
     AveragePooling2D   Y avg -------------------         0     0.0%
                        #####      5    5  512
               Conv2D    \|/  -------------------   2359808    54.5%
                 relu   #####      3    3  512
     AveragePooling2D   Y avg -------------------         0     0.0%
                        #####      1    1  512
              Flatten   ||||| -------------------         0     0.0%
                        #####        512
                Dense   XXXXX -------------------    262656     6.1%
                 relu   #####        512
              Dropout    | || -------------------         0     0.0%
                        #####        512
                Dense   XXXXX -------------------    131328     3.0%
                 relu   #####        256
              Dropout    | || -------------------         0     0.0%
                        #####        256
                Dense   XXXXX -------------------      2570     0.1%
              softmax   #####         10
```

Fig 4: Deep-CNN Model Architecture

3.4 Training the model

The proposed model Deep-CNN is being executed in the verified environment with all dependencies and pre-installed packages Google ColabPro with GPU Memory utilization of 18.26 GB and disk Space of 28 GB. The model is trained using with and without augmentation of the chosen dataset for multi-class classification with nine disease classes and one healthy. The hyperparameters used in training the model are fine-tuned and the details are discussed in the next section. We used the same dataset for training on the Deep-CNN model both with and without augmentation, also on the EfficientNetB0. The model's performance is compared with EfficientNetB0, and the impact of transfer learning will be addressed during the results and discussion.

3.5 Parameter tuning during training

The SoftMax activation, Adam Optimizer, and Sparse Categorical cross-entropy loss were used for training the Deep-CNN model. The learning rate was chosen only after checking the range of learning rates[20] from 0.01,0.001,0.0001 the model was trained using the learning rate of 0.0001 * decay of 0.95 after every 10epochs. It is noted that the learning rate of 0.001 was converging smoothly for the specified loss function. We also employed a model checkpoint with early stopping for the validation loss, with a predefined patience value. If the model doesn't show improvement for patience number of epochs, training is stopped. Doing this prevents the model from overfitting to the training data and saves time by not continuing unnecessary training.

3.6 Evaluation Metrics

The performance of the Deep-CNN and Stress Severity Assessment method was evaluated using confusion matrix, precision, recall, F1-Score, accuracy eq.(4), and stress severity rating. A detailed description is provided below with the necessary formulae. Confusion Matrix is a table that summarizes the performance of a classification model. We have 10 paddy classes; it will be a 10x10 matrix. Each row represents the actual class, and each column represents the predicted class. The elements within the matrix are True Positive (TP), True Negative (TN), False Positive (FP), and False Negative (FN). The confusion matrix for misclassification per each class is shown in Fig-5 below. Once the confusion matrix is obtained these metrics Precision, Recall, and F1 Score can be derived for evaluating the model's performance using the eq.(1) to eq.(3). The formulae are chosen in context with only one single class called blast, the same procedure is considered for all the other classes in the dataset.

Precision (for class BL): The ability of the model to correctly identify samples belonging to class BL (i.e., Blast.)

$$\text{Precision(BL)} = \frac{\text{TP (BL)}}{\text{TP(BL)+FP(BL)}} \quad (1)$$

Recall (for class BL): The ability of the model to find all samples belonging to class BL(i.e., Blast).

$$\text{Recall(BL)} = \frac{\text{TP (BL)}}{\text{TP(BL)+FN(BL)}} \quad (2)$$

F1 Score (for class BL): The harmonic mean of precision and recall, providing a single metric to balance both.

$$\text{F1-Score (BL)} = \frac{2 * \text{Precision(BL)} * \text{Recall(BL)}}{(\text{Precision(BL)} + \text{Recall(BL)})} \quad (3)$$

Accuracy is the ratio of correct samples to all the predicted samples.

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

3.7 Degree of Stress Intensity Spread

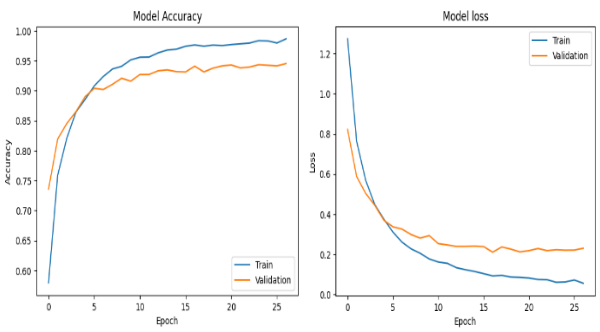
The percentage of infected areas is calculated by segmentation with a proposed method on the chosen dataset. Based on the predicted class name it identifies the stress type and its characteristic. Every biotic stress class image in the dataset consists of three RGB (Red, Green, Blue) color space that represents color information by combining the intensity of red, green, and blue channels, which is useful for displaying images but not for biotic stress analysis. A few diseases might cause subtle color changes that might not be easily distinguished in RGB due to the overlap of color information across channels. Thus, we converted the image to HSV color spaces where HSV color space separates color information (hue) from intensity (value) and saturation. Hue represents the actual color itself (e.g., red, green, yellow). Saturation represents the purity or intensity of the color i.e., 0 for grayscale and 1 for fully saturated color. The value represents the overall brightness of the pixel. This conversion is very crucial for biotic stress analysis. The method retrieves the target-specific color ranges associated with different stress classes, even if the overall intensity (value) changes. Different diseases might cause subtle changes in hue or saturation that are easier to isolate in HSV compared to RGB. Further taking advantage of this, it creates a mask by filtering the HSV image within the defined color range and isolates pixels likely belonging to the diseased area. Morphological operations (opening) are applied to refine the mask and remove noise. Images can often contain noise due to camera sensor imperfections, lighting variations, or compression artifacts. This noise can manifest as isolated pixels or small regions that don't truly represent the diseased area. Morphological opening helps remove such noise and refines the mask on the image. If a pixel's color matches a disease color, the code specially marks that pixel (like highlighting it with a marker). This creates a "mask" that shows the areas in the picture that might have the disease. The diseased area is obtained by applying the mask to the original image. The severity assessment involves converting the stressed area to grayscale. eq-(5) determines the calculation of severity. It helps in Counting non-zero pixels in the grayscale image representing the stressed area, counting the zero pixels in the grayscale image representing the non-stressed area, or simply calculating the total area as counting the number of pixels of a 2D image. Calculating the percentage of stressed area relative to the total image area to get the severity of the disease spread to the corresponding stress class.

$$\left.\begin{array}{l}\text{Non-Stressed Area} = \text{Stressed Area - Total Image Area} \\ \text{Severity} = (\text{Stressed Area / Total Image Area}) * 100\end{array}\right\} \quad (5)$$

4. Results

The experimental results obtained on implementing the Deep-CNN, EfficientNetB0 on the chosen dataset. We trained these models on the images performing with and without augmentation techniques. The models were trained on both the set of images with a batch size of 32 for about 50 and 100 epochs with a preset learning rate of 0.01. The loss function sparse_categorical_crossentropy deals with many classes. It also uses less memory usage and is faster in computation compared to categorical_crossentropy, which uses an additional memory for one-hot encoded vectors. The proposed Deep-CNN, EfficientNETB0 has attained the highest training and validation accuracies of 94.4% and 93.5% respectively with early stopping enabled for 100 epochs. The training and validation accuracy and loss graphs of the proposed Deep-CNN are shown in Fig 5 & 6.

Fig 5&6: Training and Validation Accuracy and Loss



The proposed model predicts the true classes of multi-class classification, which was observed from the confusion matrix generated for the test images. Using this matrix, we calculated the Precision, Recall, and F1-Score for further performance analysis of the model with nine biotic stressed classes and one healthy class. The results in Table 3 clearly state that the target augmented Deep-CNN model outperformed better with precision ranging from 0.89-0.97, recall ranging from 0.88-0.94, and F1-Score ranging from 0.90-0.97 among all the stress classes including healthy. Among all the stressed classes BLS, BPB, DH, DM, and

NM classes performance was extraordinary whereas BLB, BL, DM, HS, and TG classes performance was quite good.

Table 3: Performance Analysis of Proposed Deep-CNN Model

Biotic Stress Class	Precision	Recall	F1-Score	Precision	Recall	F1-Score
	Without Augmentation			With Augmentation		
BLB	0.78	0.82	0.8	0.89	0.92	0.9
BLS	0.84	0.9	0.87	0.97	0.91	0.94
BPB	0.8	0.85	0.82	0.97	0.94	0.96
BL	0.91	0.84	0.87	0.94	0.93	0.93
BS	0.85	0.83	0.84	0.94	0.88	0.91
DH	0.94	0.86	0.9	0.97	0.98	0.97
DM	0.85	0.82	0.83	0.92	0.88	0.9
HS	0.84	0.9	0.87	0.89	0.93	0.91
NM	0.91	0.89	0.9	0.92	0.96	0.94
TG	0.78	0.94	0.85	0.92	0.93	0.93

4.1 Class Labels Prediction

The proposed Deep-CNN model has been saved as the best_model

based on the experimental results with its outstanding performance. The best_model is used as a classifier for which samples of the test dataset are fed with one sample for prediction.

Fig 7: Sample Class Predictions of Biotic Stress of Test Images.

The model prediction method performs prediction and generates an output by mapping the prediction label to the given input image.

The sample predicted images for the classes: BLB, BS, DH, BL, and DM are shown below in Fig 7.

4.2 Comparison with other CNN Models

The overall performance of the proposed Deep-CNN has outperformed with equal to EfficientNetB0 when measured with different performance evaluation metrics including accuracy. The observations led to the understanding that Deep-CNN performed better when compared with the state-of-the-art works described in terms of accuracy in Table 4.

Table 4: Comparison of Deep-CNN with other CNN Models.

Model	No of Samples	Precision (%)	Recall (%)	F1-score (%)	Test Accuracy (%)
VGG16[13]	16,225	93.49	93.19	93.2	93.19
MobileNet	16,225	92.63	92.42	92.39	92.42
CNN[13]	16,225	89.22	88.84	88.81	88.84
EfficientNetB0	12,594	92.5	93	93	93.50
Proposed Deep-CNN	12,594	94	93	93	94.4

4.3 Automatic Stress Severity Assessment:

The proposed Automatic stress severity assessment method performs a crucial part of biotic stress analysis, which helps the end user i.e., the farmers to address the stress mitigations before the symptoms evolve to rise. This method deals with the handling of the intensity of infection and generates an annotation for the predicted set of samples. There are a series of steps involved to perform stress analysis and assess the severity of the stress with a rating and the percentage of spread as shown in Fig 8.

The method accepts input as an image which is the predicted stress class by Deep-CNN and converts the image to HSV color spaces.

Thresholds are defined using specific stress characteristics, to identify potential stress of the corresponding biotic stress class i.e., for all the classes of the dataset, followed by mask creation and noise removal. The masks are applied to the original image to extract regions corresponding to stressed areas.

The extracted stressed area regions are converted to grayscale to simplify severity calculations. The total number of pixels in the original image is the total area which is calculated by multiplying the width and height of the image (in pixels). The stressed area represents the number of pixels in the image that have been identified as belonging to the diseased region which is calculated by counting the white (255) pixels in the mask we create during image processing. Now we can subtract the total area from the stressed area to calculate the non-stressed area. Severity is calculated using the proportion of the entire leaf area of the image that is affected area by the stress which returns the percentage. The severity value will typically range from 0% (no disease detected) to 100% (the entire image is covered by the disease). The severity rating is determined by a rating that is assigned with an ordinal scale[17] of (0-9) based on the severity percentage. Ordinal scales provide more information than nominal scales, allowing for a general assessment of disease progression. The Standard Evaluation System for Rice (SES) uses a 0-9 scale for various diseases, where '0' indicates No disease, a scale of '1-3' indicates Low, '4-6' as Moderate and 7-9 as High severity.

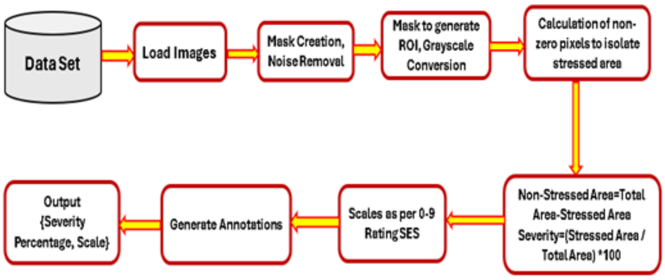


Fig-8: Pipeline of Automatic Stress Severity Assessment.

The proposed method returns the severity percentage and scale/rating as a severity assessment report shown in Fig 9. Different scales that can be used based on the type of disease are described in Table 5. and as well as the type of classification chosen. Assessing accurately and evaluating the severity of disease/stress spread is a crucial step for identifying the correlation between stress severity and yield loss prevention. During the early symptoms due to ambiguity, if it is ignored and not monitored properly, it might lead to a huge loss to the farming community. The assessment of severity can be done using different rating scales provided by the agronomist or experts in the domain. An overview of the rating scales is discussed below:

Qualitative scale: Both nominal and ordinal come under this scale. However, it is a descriptive disease scale with a variety of severity levels.

Table 5: Different rating scales.

Paddy Variety	Stress/ Disease Type	Rating Scale	Rating Method	Scaling Method	Tasks
All Varieties	Healthy	0 (No Disease)	Nominal	-	Image Classification

					(Healthy vs. Diseased)
All Varieties	Multiple Diseases	0-9 (SES)	Ordinal	Standard Evaluation System for Rice (SES)	Multi-Class Classification (Identify disease type)
		0-100% (H-B)	Quantitative	Horsfall-Barratt Scale (H-B)	Severity Assessment (Ordinal/ Quantitative Regression) Yield Loss Estimation (Regression)
Specific Varieties	Blast	0-9 (SES)	Ordinal	Standard Evaluation System for Rice (SES)	Multi-Class Classification (Blast severity levels)
	Brown Spot	0-5 (SES)	Ordinal	Standard Evaluation System for Rice (SES)	Multi-Class Classification (Brown Spot severity levels)
	Bacterial Blight	0-5 (SES)	Ordinal	Standard Evaluation System for Rice (SES)	Multi-Class Classification (Bacterial Blight severity levels)

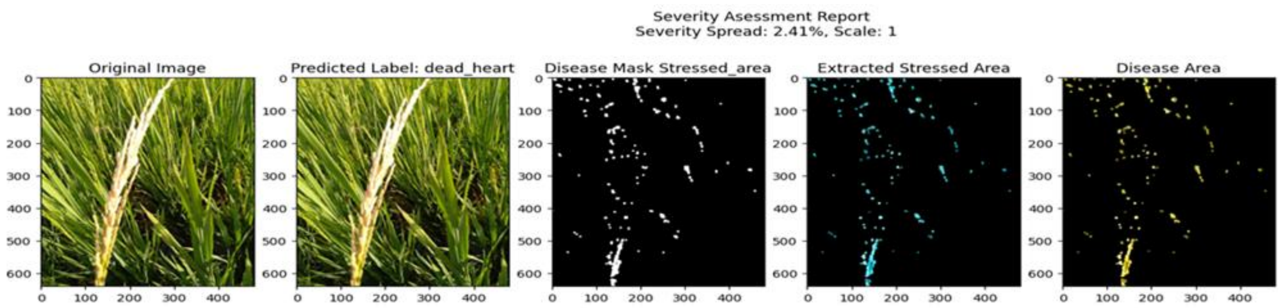
Nominal scales: This type of scaling is suitable for the classification of class labels to identify stress and categories to classify presence or absence. There's no inherent order or ranking among the categories. It can be applied for basic screening or initial detection decision-making.

Ordinal Scales: It categorizes disease severity into ordered classes. The classes have a relative ranking, but the intervals between them may not be equal. It allows a general assessment of stress progression[19].

Quantitative Scales: This method measures disease severity using numerical values using the Horsfall-Barratt Scale[18] which supports unequal intervals. These values can represent a percentage of the affected area, the number of lesions, or other continuous measures. Quantitative scales offer the most precise

measurement of disease severity, making them valuable for research and detailed analysis. We can also have a quantitative

ordinal scale where we use equal intervals to identify the symptomatic area.



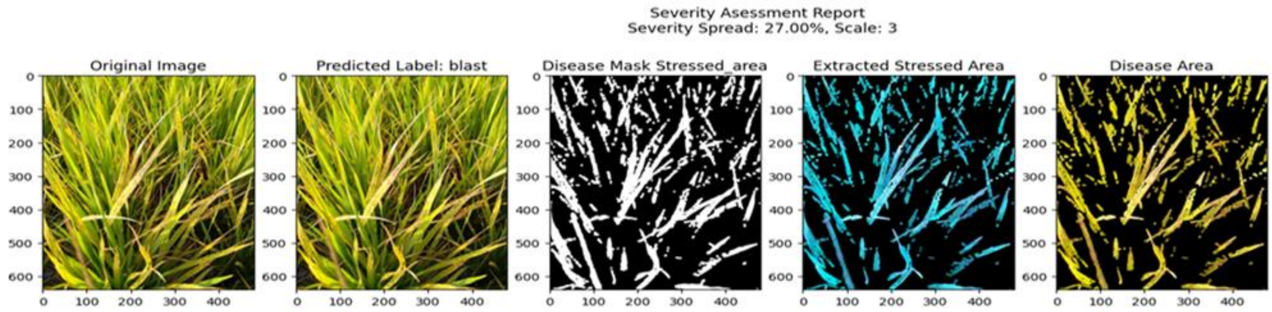


Fig 9 Severity Assessment Report

5. Conclusion & Future Scope

Rice, a cornerstone of global diets, faces escalating demand as population grows. However, biotic stresses pose a significant threat, causing substantial yield losses. Early and accurate identification of these stressors is paramount to prevent widespread damage and ensure food security. To address this critical challenge, we've developed a cutting-edge deep CNN model capable of classifying nine distinct stress classes and healthy rice. Our approach surpasses the performance of other CNN models, even on imbalanced datasets, thanks to targeted augmentation techniques prioritizing crucial stress categories. Our innovation lies in a novel severity assessment method. By applying various masks to extract the stressed areas within predicted images, we generate a comprehensive stress severity report, detailing both the severity level and the specific scale of the identified stress. This method excels in detecting and assessing critical stresses like BL, BS, HS, TG, BLB, BLS, and DH. Looking ahead, we envision extending this work to tackle the complex issue of multiple co-occurring stresses and exploring hybrid models that seamlessly integrate severity assessment and classification. Ultimately, deploying this technology on mobile applications will

empower farmers with real-time, actionable insights, revolutionizing rice disease management.

References

[1] Gnanamanickam, S.S. (2009). Rice and Its Importance to Human Life. In: Biological Control of Rice Diseases. Progress in Biological Control, vol 8. Springer, Dordrecht. https://doi.org/10.1007/978-90-481-2465-7_1

[2] Chapter 20 - Molecular techniques used in plant disease diagnosis, Food Security and Plant Disease Management, Woodhead Publishing, Science Direct, Swapnil Sapre, Iti Gontia-Mishra, Vishwa Vijay Thakur, Sumana Sikdar, Sharad Tiwari.

[3] Recognition and classification of paddy leaf diseases using Optimized Deep Neural network with Jaya algorithm, Information Processing in Agriculture, ScienceDirect, June- 2020, S. Ramesh, D. Vydeki.

[4] Basavaraj S. Anami, Naveen N. Malvade, Surendra Palaiah, Deep learning approach for recognition and classification of yield affecting paddy crop stresses using field images, Artificial Intelligence in Agriculture, Volume 4, 2020, Pages 12-20, ISSN 2589-7217, https://doi.org/10.1016/j.aiia.2020.03.001.

[5] Chowdhury R. Rahman, Preetom S. Arko, Mohammed E. Ali, Mohammad A. Iqbal Khan, Sajid H. Apon, Farzana Nowrin, Abu Wasif, Identification and recognition of rice diseases and pests using convolutional neural networks, Biosystems Engineering, Volume 194,2020, Pages 112-120, ISSN 1537-5110, https://doi.org/10.1016/j.biosystemseng.2020.03.020.

[6] Search Karim Noon, Muhammad Amjad, Muhammad Ali Qureshi, Abdul Mannan, Use of deep learning techniques for identification of plant leaf stresses: A review, Sustainable Computing: Informatics and Systems, Volume 28, 2020, 100443, ISSN 2210-5379, https://doi.org/10.1016/j.suscom.2020.100443.

[7] Lu Y, Du J, Liu P, Zhang Y, Hao Z. Image Classification and Recognition of Rice Diseases: A Hybrid DBN and Particle Swarm Optimization Algorithm. Front Bioeng Biotechnol. 2022 Apr 27;10:855667. doi: 10.3389/fbioe.2022.855667. PMID: 35573246; PMCID: PMC9091375

[8] Bock C, Hotchkiss M, Wood B (2016) Assessing disease severity: accuracy and reliability of rater estimates in relation to number of diagrams in a standard area diagram set. Plant Pathol 65:261–272. https://doi.org/10.1111/ppa.12403

[9] 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

[10] 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.

[11] 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,

[12] Lamba S, Kukreja V, Rashid J, Gadekallu TR, Kim J, Baliyan A, Gupta D, Saini S. A novel fine-tuned deep-learning-based multi-class classifier for severity of paddy leaf diseases. Front Plant Sci. 2023 Sep 5;14:1234067. doi: 10.3389/fpls.2023.1234067. PMID: 37731988; PMCID: PMC10508843.

[13] Petchiammal A, Briskline Kiruba S, Murugan D, Pandarasamy Arjunan, November 18, 2022, "Paddy Doctor: A Visual Image Dataset for Automated Paddy Disease Classification and Benchmarking", IEEE Dataport, doi: https://dx.doi.org/10.21227/hz4v-af08.

[14] P. Singla, Niharika, R. Jain, R. Sharma, V. Kukreja and A. Bansal, "Deep Learning Based Multi-Classification Model for Rice Disease Detection," 2022 10th International Conference on Reliability, Infocom Technologies and Optimization (Trends and Future Directions) (ICRITO), Noida, India, 2022, pp. 1-5, doi: 10.1109/ICRITO56286.2022.9965117.

[15] Saini, Archana et al. “Multiclass Classification of Rice Leaf Disease Using Deep Learning Based Model.” 2023 3rd Asian Conference on Innovation in Technology (ASIANCON) (2023): 1-6.

[16] Saminathan, K., Sowmiya, B., & Chithra, D.M. (2023). Multiclass Classification of Paddy Leaf Diseases Using Random Forest Classifier. Journal of Image and Graphics.

[17] IRRI. (2013). Standard Evaluation System for Rice (SES). International Rice Research Institute.

[18] Horsfall, J.G., and A.E. Dimond. 1957. Plant pathology. An advanced treatise. Vol. 1. Academic Press, New York.

[19] Shi, T., Liu, Y., Zheng, X. et al. “Recent advances in plant disease severity assessment using convolutional neural networks” . Sci Rep 13, 2336 (2023). https://doi.org/10.1038/s41598-023-29230-7.

[20] Nidhi Kundu, Geeta Rani, Vijaypal Singh Dhaka, Kalpit Gupta, Siddaiah Chandra Nayaka, Eugenio Vocaturo, Ester Zumpano, “Disease detection, severity prediction, and crop loss estimation in MaizeCrop using deep learning, Artificial Intelligence in Agriculture”, Volume 6, 2022, Pages 276-291, ISSN 2589-7217, https://doi.org/10.1016/j.aiia.2022.11.002.