Classification and Detection of Rice Diseases Using a 3-Stage CNN Architecture with Transfer Learning Approach

Munmi Gogoi 1*, Vikash Kumar 2*, Shahin Ara Begum 3*, Neelesh Sharma 4,5 and Surya Kant 4,5,*

Abstract: Rice is a vital crop for global food security, but its production is vulnerable to various diseases. Early detection and treatment of rice diseases are crucial to minimise yield losses. Convolutional neural networks (CNNs) have shown great potential for disease detection in plant leaves, but training CNNs requires large datasets of labelled images, which can be expensive and time-consuming. Here, we have experimented a 3-Stage CNN architecture with a transfer learning approach that utilises a pre-trained CNN model fine-tuned on a small dataset of rice disease images. The proposed approach significantly reduces the required training data while achieving high accuracy. We also incorporated deep learning techniques such as progressive re-sizing and parametric rectified linear unit (PReLU) to enhance rice disease detection. Progressive re-sizing improves feature learning by gradually increasing image size during training, while PReLU reduces overfitting and enhances model performance. The proposed approach was evaluated on a dataset of 8883 and 1200 images of disease and healthy rice leaves, respectively, achieving an accuracy of 94% when subjected to the 10-fold cross-validation process, significantly higher than other methods. These simulation results for disease detection in rice prove the feasibility and efficiency and offer a cost-effective, accessible solution for the early detection of rice diseases, particularly useful in developing countries with limited resources that can significantly contribute toward sustainable food production.

Keywords: deep learning; image classification; PReLU; progressive re-sizing; transfer learning

1. Introduction

The global population is projected to reach 9.7 billion by 2050, creating a pressing need to increase food production rapidly. Rice is a crucial crop for global food security, being the second-most produced crop worldwide and a staple food for over half of the global population. India is the second-largest producer of rice in the world, accounting for 24% of global production [1,2]. However, numerous factors, such as pests, diseases, bacteria, viruses, temperature, rainfall, and soil fertility, can significantly affect the yield and quality of rice crops. Rice diseases are a significant concern in rice production, as they can cause significant yield losses and impact food security [3]. Rice diseases can be broadly classified into fungal, bacterial, and viral diseases. Fungal diseases, such as brown spot and leaf blast, are caused by fungus. Bacterial diseases, such as bacterial leaf blight, are caused by Xanthomonas. Viral diseases, like tungro, are caused by viruses, such as rice tungro spherical virus (RTSV) and rice tungro bacilliform virus (RTBV) [4]. Each rice disease has specific symptoms that helps in their identification and classification. Understanding the pathogenesis of rice diseases is essential for effective management strategies [5]. Despite dedicated time and resources to managing diseases [6], timely disease detection is a major challenge, and by the time disease symptoms are detected and actions are taken to control disease, significant damage has already been done, resulting in reduced production.
is tedious for farmers to timely inspect the entire crop [7]. Therefore, early detection and identification of crop plant diseases are important in agriculture for preventing yield loss and maintaining product quality [8].

Fortunately, the convergence of computer science and technology has made it possible to detect crop diseases quickly and with minimal training by using artificial intelligence (AI) analytics [9]. Deep learning (DL), particularly convolutional neural networks (CNNs), has recently surpassed machine learning (ML) techniques for image categorisation [9–11]. Several articles have been published on the topic of disease classification using AI-based methods. The CNN model establishes a link between image layers and spatial information, making it useful for image categorization [12]. There have been few studies on rice disease categorisation using CNN. For example, Lu et al. [13] used 500 photos from 10 different categories and analysed by CNN model comprising three layers of convolution, three stochastic pooling layers, and a softmax layer. Rajmohan et al. [14] investigated deep CNN for image de-noising and support vector machine (SVM) for rice plant disease categorisation using colour and shape information on 200 images and reported 87.5% categorisation accuracy.

However, CNN requires a large amount of training data to be properly trained and collecting accurate images of a given type of diseased leaf in rice is a difficult operation. This constraint can be circumvented using the transfer learning technique which modifies the weights of previously trained networks on big data sets for the present small dataset [15]. This study experimented a 3-Stage CNN architecture that incorporates transfer learning and progressive resizing to consider various factors affecting model performance, such as effective regularisation approaches, active data augmentation, and large-scale data [16]. The transfer learning technique shortens training time, mitigates overfitting in networks, and solves the problem of insufficient training data in deep models [17]. In addition, progressive resizing helps fine-tune the final model and boosts accuracy by gradually increasing the scale of images used in the training process from small to large. The parametric rectified linear unit (PReLU) activation function has been used in the dense layers of the proposed model, which is a generalisation of the conventional rectified unit [18]. PReLU improves model fitting with little computational overhead and almost no overfitting risk.

This research implemented the 3-stage deep CNN model to identify five major rice diseases in the midland area of India. The dataset used for training and testing contained images from various regions of India. This study aimed to identify the most effective approach to using AI, neural networks, and ML in rice disease detection. By improving disease identification and classification, farmers can timely control disease and prevent yield loss and improve crop quality, ultimately contributing to global food security.

2. Materials and Methods
2.1. Model Selection for Early Detection of Rice Diseases

With the progress of DL in recent years, various deep architectures have been introduced for image classification, such as CNNs, R-CNNs, Caps Nets, and ResNets, which allow for the extraction of deep-level attributes [19,20]. Both ML and DL algorithms are used to tackle problems in agriculture research [21]. This research aims to automate tasks that are usually performed by humans, with the added value of machines performing these tasks. Deep learning technology has shown great promise in plant disease classification from images, which is a technique based on feature learning from labelled training data sets. For instance, this technique is used for diseases of tea, apple, tomato, grapes, peaches, and pears, mostly using leaves to identify diseases from images [22,23]. Table 1 provides a comparison of various machine-learning techniques used for target disease detection.
Table 1. Published research work examples of disease detection in plants using deep learning models.

<table>
<thead>
<tr>
<th>Year</th>
<th>Techniques</th>
<th>Target Data</th>
<th>Data Set</th>
<th>Number of Classes</th>
<th>Accuracy Rate (%)</th>
<th>Error Rate (%)</th>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td>2012</td>
<td>SVM/Baye’s Classifier</td>
<td>Rice Diseases</td>
<td>450</td>
<td>2</td>
<td>SVM = 68.10, Bayes = 79.50</td>
<td>20.5</td>
<td>[24]</td>
</tr>
<tr>
<td>2017</td>
<td>CNN, HOG, SVM</td>
<td>Rice Diseases</td>
<td>500</td>
<td>10</td>
<td>CNN = 95.5</td>
<td>4.5</td>
<td>[13]</td>
</tr>
<tr>
<td>2015</td>
<td>SVM</td>
<td>Rice leaves</td>
<td>-</td>
<td>1</td>
<td>SVM = 82.00</td>
<td>18</td>
<td>[25]</td>
</tr>
<tr>
<td>2015</td>
<td>SVM, C-DCGAN</td>
<td>Tea leaves</td>
<td>1500</td>
<td>3</td>
<td>SVM = 90</td>
<td>10</td>
<td>[26]</td>
</tr>
<tr>
<td>2016</td>
<td>SVM</td>
<td>Rice seedlings</td>
<td>700</td>
<td>1</td>
<td>SVM = 95.01</td>
<td>4.99</td>
<td>[27]</td>
</tr>
<tr>
<td>2015</td>
<td>CNN</td>
<td>Insect pests</td>
<td>1033</td>
<td>12</td>
<td>SVM = 75.0</td>
<td>25.0</td>
<td>[28]</td>
</tr>
<tr>
<td>2019</td>
<td>CNN</td>
<td>Different crop leaves</td>
<td>1575</td>
<td>10</td>
<td>SVM = 92.89</td>
<td>7.11</td>
<td>[29]</td>
</tr>
<tr>
<td>2020</td>
<td>Deep-CNN</td>
<td>Multiple crops diseases</td>
<td>30,000</td>
<td>12</td>
<td>SVM = 92.0</td>
<td>8.0</td>
<td>[30]</td>
</tr>
<tr>
<td>2017</td>
<td>KNN</td>
<td>Multiple crop plant disease</td>
<td>121,955</td>
<td>17</td>
<td>SVM = 92.07</td>
<td>8.07</td>
<td>[31]</td>
</tr>
<tr>
<td>2017</td>
<td>K-means Clustering</td>
<td>Leaves</td>
<td>-</td>
<td>2</td>
<td>SVM = 92.89</td>
<td>7.11</td>
<td>[32]</td>
</tr>
<tr>
<td>2018</td>
<td>AlexNet</td>
<td>Leaves</td>
<td>600</td>
<td>3</td>
<td>SVM = 91.23</td>
<td>8.77</td>
<td>[33]</td>
</tr>
<tr>
<td>2017</td>
<td>SVM</td>
<td>Rice leaves</td>
<td>50</td>
<td>7</td>
<td>SVM = 92.0</td>
<td>8.0</td>
<td>[34]</td>
</tr>
<tr>
<td>2020</td>
<td>ANN, DAE and DNN-JOA</td>
<td>Rice disease</td>
<td>400</td>
<td>5</td>
<td>SVM = 90.57</td>
<td>9.47</td>
<td>[35]</td>
</tr>
<tr>
<td>2018</td>
<td>SVM/CNN</td>
<td>Rice disease leaves</td>
<td>5808</td>
<td>2</td>
<td>SVM = 95.83</td>
<td>4.17</td>
<td>[36]</td>
</tr>
<tr>
<td>2019</td>
<td>SVM/CNN</td>
<td>Leaves</td>
<td>700</td>
<td>4</td>
<td>SVM = 91.37</td>
<td>8.63</td>
<td>[37]</td>
</tr>
<tr>
<td>2019</td>
<td>SVM</td>
<td>Rice disease leaves</td>
<td>970</td>
<td>2</td>
<td>SVM = 96.7</td>
<td>3.3</td>
<td>[38]</td>
</tr>
<tr>
<td>2019</td>
<td>DNN-CSA and 10-fold SVM</td>
<td>Rice disease leaves</td>
<td>120</td>
<td>3</td>
<td>SVM = 87.46</td>
<td>12.54</td>
<td>[39]</td>
</tr>
<tr>
<td>2021</td>
<td>CNN</td>
<td>Grape leaves</td>
<td>500</td>
<td>4</td>
<td>SVM = 93.75</td>
<td>6.25</td>
<td>[40]</td>
</tr>
<tr>
<td>2022</td>
<td>ELM</td>
<td>Lemon leaves</td>
<td>70</td>
<td>4</td>
<td>SVM = 94.0</td>
<td>6.0</td>
<td>[41]</td>
</tr>
<tr>
<td>2022</td>
<td>NSVMBPN Network</td>
<td>Rice leaves</td>
<td>970</td>
<td>3</td>
<td>SVM = 95.2</td>
<td>4.8</td>
<td>[42]</td>
</tr>
<tr>
<td>2022</td>
<td>DenseNet169-MLP</td>
<td>Rice leaves</td>
<td>1500</td>
<td>3</td>
<td>SVM = 97.68</td>
<td>2.32</td>
<td>[43]</td>
</tr>
</tbody>
</table>

2.2. Data Collection and Training

Images for the dataset were collected from various regions of India, including the states of Uttar Pradesh, Haryana, and Assam (Figure 1). The climate of Uttar Pradesh and Haryana falls under the sub-tropical region. In contrast, the climate of Assam is characterised by alternate cool and warm periods with high humidity. Samples of healthy and diseased rice plants were collected from various locations to identify diseases affecting rice plants in these states.

Figure 1. Study area of rice disease for analysis.
We developed a Deep CNN model based on images to classify rice leaf diseases. The model comprises three phases: data processing, learning, and result analysis. Data processing includes data collection, noise elimination, labelling, and visualisation. During the learning phase, tasks and learning strategies are determined to train the deep neural network. The model’s hidden representations are analysed through performance evaluation and validation. Figure 2 illustrates the classification workflow, with arrows representing the iterative design process. The backward arrows in each phase indicate iterations and updates based on the model’s performance. For example, the Iterate (1) arrow in the data-processing phase indicates room for improvement in the raw instances, while Iterate (2) represents adjusting the data-processing stage based on the learning phase. If the model overfits the training data, Iterate (3) may be used to shorten the training phase. The Iterate (4) arrow indicates that the validation and analysis phase results inform the next data processing round. The Iterate (5) arrow indicates the normal completion of the data processing.

![Figure 2. Image classification workflow.](image)

### 2.3. System Model

The DL approach is a powerful tool that utilises multiple representations to learn feature hierarchies. Its primary goal is to generate features at the top of the hierarchy from those at the bottom. Deep CNN-based image classification has proven to be a highly effective approach in computer vision, outperforming more traditional ML methods (Figure 3). One of the main advantages of CNN networks is their ability to perform automatic feature design [44], thus eliminating the need for more traditional feature extraction techniques like SIFT, HOG, and GIST.

In this implementation, we utilised a split-and-train technique, transfer learning, and progressive scaling to construct deep CNN models. Additionally, the model’s dense layers have been activated with the generalised traditional rectified unit, known as the Parametric Rectified Linear Unit (PreLU), to enhance the accuracy of the representation. This approach has a low risk of overfitting and improves model fitting with almost no discernible increase in computational cost. To enhance comprehension, we divided the concepts and procedures into several steps.
2.3. System Model

The DL approach is a powerful tool that can improve model performance and generalise the model to new datasets. Transfer learning can be implemented in different ways. One way is to use a pre-trained CNN as a feature extractor by freezing its weights and only training the last few layers. This is suitable for small datasets. Another way is to fine-tune a pre-trained CNN by unfreezing its weights and training the entire model on a new dataset, which is more suitable for large datasets.

To implement transfer learning in the proposed 3-Stage CNN model, we used EfficientNet B7, developed by Mingxing Tan and Quoc Le of the Google Research Brain team. In terms of accuracy and computing efficiency for image classification, EfficientNet surpasses the current state of the art. Based on the EfficientNet architecture, the network’s classification accuracy improves with increasing depth, width, and resolution. Efficient net adheres to the compound scaling formula, making it more efficient than other previously developed pre-trained networks. The EfficientNet family is scaled up in various block layers (from B0 to B7 using a compound scaling method). By using EfficientNet B7 as a transfer learning network, we were able to leverage the network’s pre-trained weights and significantly speed up the training process while achieving high classification accuracy.

2.3.2. Progressive Re-Sizing

DL networks incorporated a progressive image resizing technique to enhance their performance, as documented by [46]. This method involves training CNN models with a gradually increasing dataset size, starting from smaller images such as 128 × 128 pixels. The weights of the first model are then utilised to train a second model on larger images of size 256 × 256 pixels, and so on, to fine-tune the models and improve their accuracy scores. The architecture of larger models incorporates the layers and weights of earlier smaller models to further refine their performance. Despite the resolution of the images being changed from 64 × 64 to 128 × 128 pixels, the human eye may not discern the difference, but it provides CNN models with a fresh dataset to practice and improve their learning as shown in Figure 4.
2.3.3. PreLU Activation

Additionally, we used the PreLU activation function on the dense layers of the 3-Stage CNN model, which generalises the standard rectified unit [47] by introducing a learnable slope parameter for negative values. This allows the PreLU function to better capture the non-linearities in the data, which can lead to improved performance on classification tasks. PreLU also has the benefit of being computationally efficient and less prone to overfitting than other activation functions. The activation function of a neural network can be defined as in Equation (1):

\[
L(x_i) = \begin{cases} 
 X_i & \text{if } X_i > 0 \\
 a_i X_i & \text{if } X_i < 0 
\end{cases}
\]

(1)

The slope of the negative component of the PreLU function is determined by the coefficient \(X_i\), which is the input to the \(i^{th}\) nonlinear activation function \(L\). This allows the PreLU function to be differentially expressed in different channels, as shown by the \(i^{th}\) in \(a_i\). ReLU is the name given to the activation function when the coefficient \((a_i = 0)\), while PreLU is the name given to Equation (1) when \((a_i)\) is a learnable parameter [47]. Figure 5 presents the results of the experiment using the Deep CNN model.

Figure 5. System architecture of Deep 3-Stage CNN architecture with transfer learning [48].

The proposed Deep 3-Stage CNN model was built in three stages using the pretrained architecture of EfficientNet7 as a transfer learning model (Figure 5) [48]. Layers were added...
using the compound scaling concept. The PreLU activation was used in the dense layers of the proposed architecture, which allows the parameter to be adaptively learned from the data. The controlling coefficient \((a_i)\) regulates the negative slope by adaptively learning the parameters. The proposed Deep 3-Stage CNN model incorporates a progressive rescaling technique on images to boost accuracy while taking both image size and the regularization parameter into account. The split and train strategy, which is an enhancement of progressive learning, was implemented to achieve a faster training time and better re-sizing dataset.

Adding more layers to a network increases its complexity, making it more difficult to train. This is because the network has more parameters to learn, and the loss function can become more jagged, making it difficult to find the optimal solution. To address this challenge, hyperparameter optimization was used to tune the parameters of the network, such as the learning rate, batch size, and number of epochs. This can help the network to learn more effectively and achieve better accuracy. In addition, strategies like SGD, Adam, AdaGrad, and AdaDelta were implemented for optimisation [49]. This is because different optimisers work better for different networks. For example, the SGD optimiser is a good choice for small networks, while the Adam optimiser is a better choice for large networks. By fine-tuning the optimiser, the proposed Deep CNN model can achieve better accuracy on the rice leaf datasets with fewer parameters. The model’s Hyperparameters with values are shown in Table 2.

**Table 2.** Hyperparameter values of the CNN model.

<table>
<thead>
<tr>
<th>Hyper-Parameters</th>
<th>Values</th>
</tr>
</thead>
<tbody>
<tr>
<td>Optimizer</td>
<td>Adam</td>
</tr>
<tr>
<td>Batch size</td>
<td>32</td>
</tr>
<tr>
<td>Momentum</td>
<td>0.8</td>
</tr>
<tr>
<td>Learning Rate</td>
<td>Decay learning rate</td>
</tr>
</tbody>
</table>

2.4. Dataset and Implementation Setup

The rice image dataset used in this study were captured using a Samsung 52S mobile camera (Samsung Electronics, Noida, India) and a Canon EOS 1500D DSLR (Canon Inc., Tokyo, Japan). The dataset comprises a total of 8883 images, including those of rice leaves affected by four diseases: brown spot, leaf blast, bacterial leaf blight (BLB), and tungro, and 1200 images of healthy leaves, presented in Table 3. Brown spot infected leaves have large spots, which initially appear as small, circular, dark brown to purple-brown lesions. As the disease progresses, lesions become circular to oval with a light brown to grey centre and a reddish-brown margin (Figure 6a) [50]. In leaf blast disease, small necrotic regions grow into chlorotic margins. The disease affects rice plant leaves, collars, nodes, internodes, leaf blades, and panicles (Figure 6b) [51]. BLB infected leaves turn greyish-green and roll up, then turn yellow, discolour straw, and die (Figure 6d). Wavy margins spread toward the base. Bacterial ooze resembles early-morning dew in younger leaves [52]. The most serious rice disease; tungro, in South and Southeast Asia is caused by rice tungro spherical virus (RTSV), transmitted by the green leafhopper (Nephotettix virescens), and rice tungro bacilliform virus (RTBV). RTSV infection causes yellowing and redness in leaves and stunt plant growth (Figure 6e). RTBV causes severe symptoms without leafhopper transmission [53].

**Table 3.** Dataset description of training and validation images.

<table>
<thead>
<tr>
<th>Name of Class</th>
<th>Training Images</th>
<th>Validation Images</th>
</tr>
</thead>
<tbody>
<tr>
<td>Brown Spot</td>
<td>1770</td>
<td>443</td>
</tr>
<tr>
<td>Leaf Blast</td>
<td>1860</td>
<td>465</td>
</tr>
<tr>
<td>Tungro</td>
<td>1201</td>
<td>300</td>
</tr>
<tr>
<td>BLB</td>
<td>1300</td>
<td>324</td>
</tr>
<tr>
<td>Healthy</td>
<td>980</td>
<td>220</td>
</tr>
</tbody>
</table>
2.4. Dataset and Implementation Setup

The rice image dataset used in this study were captured using a Samsung 52S mobile camera and a Canon EOS 1500D DSLR. The data set comprises a total of 8883 images, including those of rice leaves affected by four diseases: brown spot, leaf blast, bacterial leaf blight (BLB), and tungro, and 1200 images of healthy leaves, presented in Table 3. Brown spot infected leaves have large spots, which initially appear as small, circular, dark brown to purple-brown lesions. As the disease progresses, lesions become circular to oval with a light brown to grey centre and a reddish-brown margin (Figure 6a) [50]. In leaf blast disease, small necrotic regions grow into chlorotic margins. The disease affects rice plant leaves, collars, nodes, internodes, leaf blades, and panicles (Figure 6b) [51]. BLB infected leaves turn greyish-green and roll up, then turn yellow, discolour straw, and die (Figure 6d). Wavy margins spread toward the base. Bacterial ooze resembles early-morning dew in younger leaves [52]. The most serious rice disease; tungro, in South and Southeast Asia is caused by rice tungro spherical virus (RTSV), transmitted by the green leafhopper (Neptovirella virescens), and rice tungro bacilliform virus (RTBV). RTSV infection causes yellowing and redness in leaves and stunted plant growth (Figure 6e). RTBV causes severe symptoms without leafhopper transmission [53].

(a) Brown spot (b) Leaf blast (c) Healthy (d) BLB (e) Tungro

Figure 6. Sample images of the disease dataset.

Implementation setup: Python version 3.7 was used for implementing the network architecture, with Keras 2.3.1 and TensorFlow 1.15 serving as the frontend and backend, respectively. We employed model and data-parallelism techniques for conducting image classification experiments on a Linux-based GPU environment with over 2000 CPU cores, 1.5 TB RAM, and NVIDIA Tesla V100 32 GB GPU accelerators (Hewlett Packard Enterprise, Nvidia, Santa Clara, CA, USA) running on Google Collab Pro.

To train network on the rice leaf dataset, which includes five distinct classes (BLB, leaf blast, tungro, healthy, and brown spot), we used the GPU configuration platform to create a pretrained network using ImageNet, which was then fine-tuned. Furthermore, we used a decreasing learning rate schedule of 5% per 10 epochs for an adaptive learning schedule, with each iteration comprising 150 epochs.

Performance matrices: Accuracy and error rate calculations were used to make model comparisons [54]. In most cases, the best models are those with the fewest errors and the highest accuracy. The precision and the margin of error are defined as follows:

\[
\text{Accuracy of model} = \frac{TN + TP}{FN + FP + TP + TN}.
\]

\[
\text{Error in model} = \frac{FN + FP}{FP + FN + TP + TN}.
\]

The numbers of true negatives, false negatives, true positives, and false positives are denoted as \( TN, FN, TP, \) and \( FP \).
3. Results and Discussion

Here we describe the accuracy of the deep 3-Stage CNN model that classifies a dataset of rice leaf images into five classes. The proposed 3-Stage CNN model trained under suitable hyperparameter values for the best results and generalizing the model. The hyperparameter values are described in Table 2. For this experimentation, the proposed 3-Stage CNN model mentioned in Figure 5 was trained using two configurations: (i) progressive resizing and (ii) ReLU versus PReLU activation with transfer learning. Previous research shows that the CNN model outperformed with precision variations ranging from 3% to 28.9% [55,56]. Mohanty et al. [57] trained models to identify 14 crop species and 26 diseases using the Plant Village database, achieving a 99.4% accuracy on a held-out test set. Picon, Alvarez-Gila et al. [58] focused on three wheat diseases and achieved a balanced accuracy of 98% for Septoria and 96% for Rust using a mobile app.

The efficiency of deep neural networks (DNN) training and accuracy improvement is significantly influenced by image size. In this experiment, three parameters were considered while training the network with images of various aspect ratios. The experiment revealed that regularisation parameters employed by the model could affect its accuracy. To achieve higher accuracy when incorporating progressive rescaling on image data, it is recommended that regularisation parameters be adjusted instead of using fixed regularisations, which can lead to inaccurate results. More stringent regularisation is essential to prevent overfitting in large models. For example, EfficientNet-B7 employs higher dropout rates and more stringent data augmentations than EfficientNet-B0. In this experiment, regularisation methods such as Dropout [59] and Rand Augment [60] were used in conjunction with the progressive rescaling of images during three training phases for various image dimensions.

The present study conducted experiments on rice leaf disease datasets to evaluate the performance of a proposed 3-Stage CNN model for image classification. The experimental setup is presented in Table 4, with 100 epochs set for each stage. As illustrated in Figure 7, an accuracy of 89.5% was achieved when the image size was $64 \times 64$ pixels, and the regularisation was less stringent (RandAugment = 5, Dropout 0.1). However, accuracy was increased to 93.99% after the third training stage when the image size was increased and applied more stringent regularisation.

Table 4. The model’s progressive rescaling setup parameters.

<table>
<thead>
<tr>
<th>Training Stages</th>
<th>Image Size</th>
<th>Rand Augment</th>
<th>Dropout</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Min</td>
<td>Max</td>
<td>Min</td>
</tr>
<tr>
<td>First Training Stage</td>
<td>64</td>
<td>224</td>
<td>5</td>
</tr>
<tr>
<td>Second Training Stage</td>
<td>64</td>
<td>224</td>
<td>5</td>
</tr>
<tr>
<td>Third Training Stage</td>
<td>64</td>
<td>224</td>
<td>5</td>
</tr>
</tbody>
</table>

The accuracy of model improved from 90.79% to 93.99% when we moved from stage I to stage III with a larger image size and stronger regularisation parameters. Furthermore, the accuracy of model and training time were improved when the dataset was rescaled in progressive steps. The loss and accuracy plot of the deep model is shown in Figure 8, where the accuracy on the rice leaf dataset was 93.99%.

Our proposed model, which was trained for a total of 26 h using 36 million parameters, achieved an accuracy of 93.99% on categorising the five distinct types of images, as shown in Figure 9. We used fewer FLOPs and fewer parameters than other benchmark models in the literature, which suggests that increasing the image size is more effective for this data set than increasing the model size by adding layers.
The present study conducted experiments on rice leaf disease datasets to evaluate the performance of a proposed 3-Stage CNN model for image classification. The loss and accuracy plot of the deep model is shown in Figure 8, where the accuracy on the rice leaf dataset was 93.99%.

We observed an improved accuracy of PReLU over ReLU in the activation function setup of our proposed deep 3-Stage CNN architecture for rice disease classification over five distinct types of images, as shown in Figure 9. We used fewer FLOPs and fewer parameters than other benchmark models in the literature, which suggests that increasing the image size is more effective than increasing the model size by adding layers. We observed an improved accuracy of PReLU over ReLU in the activation function from stage I to stage III, showing that PReLU is 3.2% more accurate than ReLU.

Table 4. Representation of accuracy, parameters, and training hours of the Deep model.

<table>
<thead>
<tr>
<th>Training Stages</th>
<th>Image Size</th>
<th>Rand Augment</th>
<th>Dropout</th>
<th>Parameter</th>
<th>Training hrs</th>
</tr>
</thead>
<tbody>
<tr>
<td>First Training Stage</td>
<td>64 224</td>
<td>5</td>
<td>10</td>
<td>36 million</td>
<td>26</td>
</tr>
<tr>
<td>Second Training Stage</td>
<td>64 224</td>
<td>5</td>
<td>15</td>
<td>36 million</td>
<td>26</td>
</tr>
<tr>
<td>Third Training Stage</td>
<td>64 224</td>
<td>5</td>
<td>0.2</td>
<td>36 million</td>
<td>26</td>
</tr>
</tbody>
</table>

Figure 7. Accuracy comparison of the CNN model in progressive resizing and activation function setup.

Figure 8. Loss and accuracy plot of the model on rice leaf dataset.

Figure 9. Representation of accuracy, parameters, and training hours of the Deep model.

We observed an improved accuracy of PReLU over ReLU in the activation function setup of our proposed deep 3-Stage CNN architecture for rice disease classification over
the datasets. On the dataset, the training implementation achieved 93.99% accuracy after being tested with 10 different views. Using progressive training on different blocks without loading the weight of the previous block, ReLUs were used to train the model across all layers. As a result, accuracy in all three sections improved by around 3%. Later, the model was trained by switching out all of the ReLUs for PreLUs, and progressive training was also performed by loading weights from the prior block. Table 5 provides the results’ specifics, showing that PreLU is 3.2% more accurate than ReLU from stage I to stage III in the rice leaf dataset. The more favourable value is presented in boldface for each of the stages.

Table 5. Classification accuracy of rice dataset deep model.

<table>
<thead>
<tr>
<th>Training Stages</th>
<th>Activation Function</th>
<th>Rand Augment</th>
<th>Dropout Rate</th>
<th>Image Size in Pixels</th>
<th>Accuracy</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Min</td>
<td>Max</td>
<td>Min</td>
<td>Max</td>
</tr>
<tr>
<td>Stage I</td>
<td>ReLU</td>
<td>5</td>
<td>10</td>
<td>0.1</td>
<td>0.3</td>
</tr>
<tr>
<td></td>
<td>PreLU</td>
<td>5</td>
<td>10</td>
<td>0.1</td>
<td>0.3</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Stage II</td>
<td>ReLU</td>
<td>5</td>
<td>15</td>
<td>0.2</td>
<td>0.5</td>
</tr>
<tr>
<td></td>
<td>PreLU</td>
<td>5</td>
<td>15</td>
<td>0.2</td>
<td>0.5</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Stage III</td>
<td>ReLU</td>
<td>5</td>
<td>20</td>
<td>0.1</td>
<td>0.5</td>
</tr>
<tr>
<td></td>
<td>PreLU</td>
<td>5</td>
<td>20</td>
<td>0.1</td>
<td>0.5</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

The results indicate that (i) using the PreLU activation function on model layers generalises the standard rectified unit and offers various benefits, including improved model fitting, reduced risk of overfitting, superior efficiency on datasets that incorporate rescaled images, and essentially no additional computing cost. (ii) Fine-tuning the Deep network’s hyperparameters and optimisers and increasing the training dataset’s size can lead to competitive results in the high-accuracy regime. (iii) The improved performance on the dataset with rescaled images can be attributed to the use of PreLU activation function on the model layers, which generalises the classic rectified unit and improves model fitting with almost no additional computing cost and no risk of overfitting.

4. Conclusions and Future Work

Computer vision and AI frameworks have spread to many areas of the food industry, including farming and processing. The application of these techniques is effective to computerize laborious tasks in a risk-free manner, thereby producing sufficient data for subsequent investigation. These techniques are applicable to monitor the disease severity thereby timely controlling diseases in crops with reduced losses and increased yield. Here, we propose a deep 3-Stage CNN model and showed how to use transfer learning, progressive scaling, and a split-and-train strategy to improve upon a pre-trained deep CNN of EfficientNetB7. In addition, the thick layers of the model were activated using the PreLU, a generalisation of the classic rectified unit. This was done to improve the accuracy of the simulation and ensure that the representation was as accurate as possible.
PReLU reduces the risk of overfitting while increasing model fitting quality with negligible computational overhead.

This study considered the four most common diseases affecting rice plants: leaf blast, bacterial blight, brown spot, and tungro. We conducted experiments by splitting the entire dataset into training and testing sets at varying ratios. Our proposed model achieved a classification accuracy of 93.99% for rice diseases. However, the lack of standard labelled rice disease images made it inappropriate to benchmark the proposed model against the literature. Nevertheless, our suggested 3-Stage CNN architecture has better performance, shorter execution time, and fewer parameters than the other architectures of the literature on rice leaf benchmark datasets. The model took shorter execution time as we have implemented transfer learning and fine-tuned the hyperparameters, and it has fewer parameters as compared to other models. A larger dataset of rice disease images would further improve the model’s performance. Framework’s application can be effectively extended to other crops and diseases. The proposed framework can be effectively used for more accurate detection of diseases affecting rice as well as other crops, which cause a significant economic loss in the agricultural sector.

Author Contributions: Investigation, methodology, data curation, writing-original manuscript draft, M.G. and V.K.; formal analysis, methodology, writing-review and editing, S.A.B., N.S. and S.K.; investigation, supervision, resources, S.K. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Data Availability Statement: Data is available on request from authors.

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

References


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