A novel deep learning-based hierarchical attention feature fusion network for automated detection of rice leaf diseases

Tejashwini C. Gadag, D. R. Kumar Raja
School of Computer Science and Engineering, REVA University, Bengaluru, India

ABSTRACT

Rice, as a staple crop, plays a crucial role in global food security, especially in developing countries. However, rice production is significantly impacted by diseases such as Brown Spot, Hispa, and Leaf Blast, which can reduce crop yield. Traditional methods of disease detection rely on manual inspection, which is time-consuming, labour-intensive, and prone to errors. To address these challenges, this paper presents a novel deep learning-based model for automated rice leaf disease detection. This paper proposes a novel deep learning-based model, the hierarchical attention feature fusion network (HAFFN), designed to enhance rice leaf disease detection accuracy by addressing key limitations in existing methods. The HAFFN model integrates multi-level feature extraction with a hierarchical attention mechanism to improve the detection of both small and large infected areas. The core novelty of the proposed approach lies in the combination of the deep multiscale feature fusion network (DMFN), the adaptive multiscale feature aggregator (AMFA), and the deep hierarchical attention module (DHAM). The model was trained and tested on a publicly available rice leaf disease dataset and demonstrated superior performance compared to benchmark models like LeafNet, Xception, and MobileNetV2.

Corresponding Author:

Tejashwini C. Gadag
School of Computer Science and Engineering, REVA University
Bengaluru, India
Email: tejashwinicg_12@rediffmail.com

1. INTRODUCTION

Particularly in developing countries, the agriculture sector is essential to rural inhabitants' access to food, money, and employment possibilities. This industry contributes around 6.4% of the world's economic production [1]. According to a 2016 survey, 65 percent of those who live in poverty make their main source of income from agriculture. Furthermore, a large percentage of consumer needs for agricultural commodities—particularly in rural areas—are satisfied by agriculture. The establishment of an efficient and well-functioning agricultural system is necessary to provide dependable food security for humans. To maintain this state, the agricultural production system must be managed effectively. Many farmers in rural regions lack the expertise needed to identify and address crop issues early on in the growing season. They are frequently harmed by different illnesses, which makes it more difficult for them to get the necessary output from the crops they grow.

For billions of people worldwide, rice is a staple grain that plays a critical role in nutrition. It has been noted that 61.54% of all countries on Earth cultivate it. The problem of a growing population and a decreasing amount of arable land has made the issue of a gradual food scarcity worse, leading to a greater dependence on rice. In addition, there are ongoing obstacles to rice production, such as the yearly appearance of several illnesses, especially harmful rice leaf diseases. Fungal and viral diseases interfere

with rice leaves' natural growth, structure, and color. The majority of people in India's mostly agricultural economy eat rice as a staple diet [2], [3]. India, which is said to be the world's biggest delta, has rivers that originate in the Himalayas. Even with its rich soil and ideal seasons, the country has not yet reached its full agricultural potential. The deficiency is ascribed to restricted availability of cutting-edge technology and a reluctance to embrace novel technologies, resulting from several elements such as misinterpretation, ambiguity, and perplexity. The amount of rice produced annually—roughly 135.76 million metric tons—is insufficient to feed the world's expanding population. The main reason for this shortfall and the reason for lower crop yields is infections in the rice leaves [4]. To increase productivity, it is important to use appropriate care measures and evaluate crop health consistently. The diagnosis of rice leaf diseases by hand is associated with high expenses, a large time commitment, and a substantial labor demand [5].

Producing wholesome food requires the early detection of plant diseases. A crucial part of conventional research on plant ecology is the detection of plant diseases. Farmers face difficulties in accurately identifying plant illnesses because of the subtle symptoms that these diseases display. If diseases brought on by bacteria, fungus, or viruses are not adequately treated, farmers may suffer significant losses [6]. Brown spot, hispa, and leaf blast are common diseases that harm rice [7]. Effective disease management techniques, such as crop rotation, the use of resistant cultivars, and the timely administration of fungicides, can lessen the impact of these diseases on rice production [8]. Chemical insecticides are one method of controlling rice illnesses [9]. These materials are sprayed on rice fields to either eradicate or stop the growth of harmful fungus, viruses, and bacteria. Depending on the kind and stage of the illness, these compounds can be sprayed on the leaves, in the soil, or as seed treatments [8]. Chemical pesticides have the potential to negatively impact human health as well as the environment, yet they can be useful in controlling rice illnesses [10]. Timely detection and precise diagnosis are essential for the effective management of rice illnesses. By quickly diagnosing the disease and taking timely action to assess its severity, farmers can reduce the disease's spread and productivity losses. Early identification makes it easier to use safer and more targeted control methods, such cultural practices, or less toxic pesticides. A quick and accurate diagnosis is made possible by the combination of visual inspection, laboratory testing, and diagnostic tools, such as molecular tests. The early detection of rice illnesses may be improved by utilizing a variety of contemporary diagnostic technologies, such as remote sensing and machine learning techniques.

Neural network, hybrid, and machine learning algorithm-based methods were used in earlier research on rice leaf disease identification. Even if classic machine learning algorithms worked well in some situations, their performance was frequently compromised by their reliance on time-consuming, human feature engineering. The use of hand-crafted parts resulted in longer development times and more work needed, and also made it more difficult to capture complex patterns. With the advent of neural networks, convolutional neural network (CNN)-based techniques that combine transfer learning with pre-trained models and customized architectures have become more and more common. Because of the large magnitude of their trainable parameters, these methods frequently show significant asymptotic complexity. This feature makes them less useful for real-time applications and makes it difficult to deploy on devices with limited resources [8], [9].

Rice leaf diseases, such as Brown Spot, Hispa, and Leaf Blast, pose significant threats to crop yields, particularly in developing countries like Bangladesh, where rice is a vital staple. Manual detection of these diseases is labor-intensive, error-prone, and inefficient, highlighting the need for automated, reliable detection methods. While various deep learning models like CNNs have been used for disease detection, existing approaches often suffer from limitations such as poor performance on less infected areas and high computational complexity. This research aims to address these challenges by proposing a novel deep learning-based model that integrates multi-level feature retrieval and an attention scheme to enhance disease detection accuracy across varying infection sizes. By improving detection efficiency, this work seeks to promote more sustainable agricultural practices and contribute to global food security, further detailed contribution of research work is given as follows:

− We propose the hierarchical attention feature fusion network (HAFFN), a unique deep learning architecture that integrates multi-level and multi-scale feature extraction with an attention mechanism. This model is specifically designed to address challenges in detecting rice leaf diseases with varying infection sizes, from small localized spots to larger infected areas, ensuring more accurate and robust disease classification.
− The adaptive multiscale feature aggregator (AMFA) enhances the model's ability to retain crucial information from different feature layers, particularly improving the detection of less infected areas, which are often missed by existing models. This method enables the fusion of semantic and texture-based features, improving the precision of the detection process.

- The deep hierarchical attention module (DHAM) is introduced to dynamically focus on the most relevant features across different scales, allowing the model to better differentiate between healthy and diseased leaf areas, even when the infection size varies. This attention mechanism improves detection accuracy by adapting to both large and small infected regions.
- The HAFFN model is thoroughly evaluated against benchmark models like Xception and MobileNetV2, demonstrating superior performance in terms of accuracy, precision, recall, and F1-scores across multiple disease classes.

2. RELATED WORK

Over the past several years, there have been significant improvements in deep neural networks. In the field of crop leaf picture recognition, CNNs integrated with transformer-based architectures have attracted a lot of interest [11]. Using multi-level and multi-scale feature extraction and representation approaches is crucial since crop leaf photographs usually have complex structures, textures, and shapes. CNNs are excellent image processors because they are good at catching local features [12], whereas transformer-based networks are good at gathering global contextual data [13], which improves the accuracy of picture identification. This section provides a thorough summary of the research done on crop disease identification, emphasizing the main conclusions and the approaches used by the investigators. Particularly in relation to the diagnosis of agricultural leaf diseases, recent developments in transformer network design have been noted [14].

Yu *et al.* [15] used an image transformer (ICVT) in conjunction with an Inception neural network to show significant improvements in the diagnosis of plant diseases. The intricacy of the deep network architecture used in this technique may necessitate more computer power and longer training times in scenarios involving larger datasets. The foundation for the creation of the ICVT architecture is the convolutional vision transformer (CvT) [16] design, which integrates ideas from CNN and ViT. The procedure comprised using convolutional algorithms to divide each input picture into patches of the same size. These patches are first flattened using depth-wise separable convolution operations into the query, key, and value matrices that are necessary for the attention mechanism before being merged into the multi-head attention layer.

The CNN and ViT strengths are combined in this architecture, but the parameter-intensive attention and multilayer perceptron (MLP) layers remain unaltered. Many researchers have shown interest in transformer-based systems with lightweight architectures. Touvron et al. presented the data-efficient image transformers (DeiT) model [17]. Information distillation strategies are used by the self-attention mechanism [18] to accomplish effective training and inference. Even though DeiT has improved data efficiency, its performance results may still be influenced by the amount of training data in some tasks and datasets. Zhao *et al.* [19] presented a "cross-feature attention" technique that controls attention via the feature dimension, decreasing the computing cost of transformers by producing linear complexity. Luo *et al.* [20] introduced a group-wise transformer intended for vision-and-language tasks. The method minimizes the parameters and computations related to the transformer while preserving its two key features: effective attention modeling over many MHA subspaces and the expanding-scaling feature transformation of MLP.

Hossen *et al.* [21] provide a two-dimensional CNN model intended for the classification and diagnosis of wheat illnesses. With respect to a dataset of 4,800 photos, the suggested model performs with a 98.84% accuracy rate in differentiating between wheat fields that are infected and those that are not. The study makes use of feature extraction, scaling, and segmentation techniques to improve pre-processing and disease detection precision. A deep neural network classification method is presented by [22] to distinguish between wheat heads infected with Fusarium head blight and healthy ones in hyperspectral pictures. With the use of a hybrid neural network that combines convolutional and bidirectional recurrent layers, the model transforms two-dimensional pixel spectra data. An accuracy of 0.743 and an F1 score of 0.75 were obtained from the testing dataset, indicating increased performance over previous models. With an accuracy of 84.54%, the Hussain *et al.* [23] suggest a CNN-based technique for identifying diseases of wheat crops. Four different types of wheat illnesses are included in the training dataset, highlighting the potential value of the model in helping farmers protect their wheat harvests. According to [24], SimpleNet is a portable CNN model intended for the automated diagnosis of wheat ear disorders in photos taken in the field. Two essential parts of SimpleNet are the convolutional block attention module (CBAM) and the feature fusion module, which improve representation capabilities while maintaining important information.

The experiment produced a 94.1% identification accuracy. A few-shot learning process was used by Alharbi *et al.* [25] to develop a unique approach for diagnosing illnesses in wheat crops. Using EfficientNet as the base model, the study incorporates an attention mechanism to improve feature selection

efficiency. On a manually assembled dataset of 18 wheat diseases taken from the internet, the suggested network shows an accuracy of 93.19%, suggesting remarkable performance. Scenario analysis is used by [26] to investigate how changes in the world's conditions may affect wheat health by 2050. The study offers a scenario analysis approach that considers three fictitious agrosystems' worth of factors, including pesticide use, host plant resistance, and climate change. The findings highlight the use of scenario analysis in assessing the development of crop health, showing a general rise in risk probabilities and magnitudes throughout agrosystems.

In a different study [27], ResNet50 and SVM models were trained using 5932 pictures of four types of rice leaf diseases. The SVM classifier is not suited for large datasets. The performance shows abnormalities since ResNet50 requires a large amount of data to be input concurrently. Recently, Ghosal and Sarkar [28] suggested a CNN architecture using the pretrained VGG16 backbone, leveraging transfer learning. Compared to other models like ResNet, VGG has a more expansive design and requires more time to process images. Other pretrained models are not considered. Shrivastava and Pradhan [29] explores 14 different color spaces and uses color features to derive four colors from each color channel. The lighting affects the leaves' color, which fluctuates even if it is not consistent. It has been observed that the efficacy of SVM classifier implementation is restricted in situations when there are substantial datasets. Considering all local properties of a picture, our suggested approach overcomes the shortcomings noted in the literature by using a model with a significantly smaller parameter size than previous approaches.

3.   PROPOSED METHOD

The proposed methodology of this paper focuses on detecting diseases in the leaves of rice. By implementing a combination of various advanced techniques of deep learning for enhancement of detection. Figure 1 given below shows the framework of the HAFFN model.
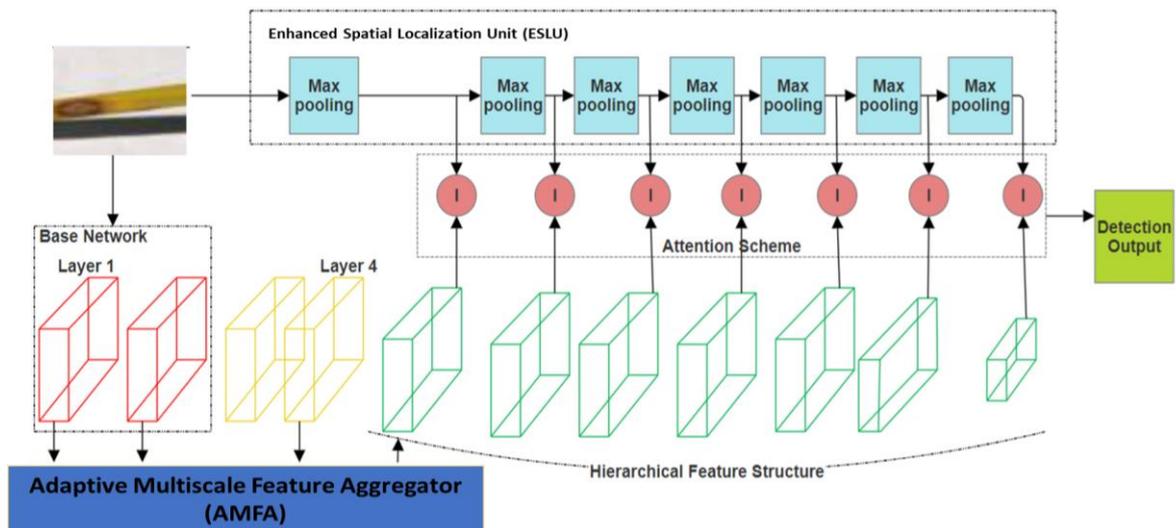

Figure 1. HAFFN model for rice leaf disease detection and classification

In this case, we observe the technique developed for the issue of disease detection that has areas of infection of various dimensions, here the infected areas are predicted using attribute maps of various receptor fields. Completely linked layers combined with convolutional and deconvolution layers are implemented. For this study, various techniques of AMFA, feature space area as well as DHAM is implemented to aid a deep multiscale feature fusion network (DMFN) result in an enhanced performance for the disease detection in leaves. Colored (Red, green, blue) images are used as an input for the network. In the figure, the red and yellow layers of convolution are parts of the initial DMFN. The AMFA is developed as a combination of attribute maps from various levels, which along with the withheld DMFN phase leads to multiple level diseased areas feature extraction, this leads to the generation of a new hierarchical structure of features. Additionally, feature-space is based on the network area that is used to improvise the data for the area. In conclusion, the attention scheme is used for building a multiple level detection model having a detection level that is adaptive.

3.1. AMFA

The dimension of most of the infected areas in the leaf is as small as $32\ by\ 32$, which is assumed as small objects. Because of down sampling, the attribute map generated for DMFN has the dimension of $1\ by\ 1$. After this convolutional layer, the data for small areas disappears, that leads to lesser perception for smaller objects. For more enhancement of performance for the HAFFN model, the representation of the attributes that are retrieved use feature fusion. We observe from figure 1, the AMFA segment uses the attribute maps that are completely linked to the DMFN switched by the atrous convolutional layer. While considering the top-down approach, the AMFA is built having an architecture joining semantic attributes having lesser resolution as well as texture attributes having increased resolution of various stages from feature maps. This is due to the layers having increased location as well as detailed data of infected areas in the leaves, the performance of AMFA is widely improvised.

The dimension of the attribute map that is generated is $64\ by\ 64$. The HAFFN model consists of deconvolutional layer with two stride for the generation of an attribute map of the dimension $64\ by\ 64$. For the convolutional layer 4, two layers of deconvolution are used having kernel dimensions of $2\ by\ 2$ that are used for the generation of an attribute map having dimensions $64\ by\ 64$. Every layer of deconvolution has a convolutional layer as well as ReLU following it for enhancing the semantic data that is used to represent the attributes. The first layer of convolution, having output size of 256 as well as a ReLu layer is used having a constant dimension of the attribute map. Furthermore, the layer implemented are integrated using a concatenation scheme. Gradually, a layer of batch normalization is applied for improvision of generalization capabilities of the network. Once this attribute fusion is completed, the initial DMFN is used as a reference to design a new hierarchical attribute extractor.

3.2. Enhanced spatial localization unit (ESLU)

The disease detecting task in leaves focuses on locating the infected areas precisely, which implies that the texture data is essential. The feature-space based feature retrieval method is implemented in the proposed study. For the hue-saturation-brightness (value), the Saturation in this case is explained as the ratio of the chosen colour's purity and the max purity that the chosen colour can have. The hue here, shows the location of the colour on the spectrum where it is located, this is evaluated by considering the score from 0 to $360\ degrees$. Value or brightness has a range from $0-1$. The images obtained here are from the environment having natural lighting, exposure and shadows implying that they are more exposed to luminance. Therefore, the RGB space is unable to express luminance efficiently, hence HSV space is implemented. It is possible that the Hue of the image could omit the original attributes of the infected areas, the Saturation of the image could decrease the features of the infected areas. Hence, max-pooling is applied to attain improved texture-based data as the retrieval of texture data for an area is performed more efficiently. This operation evaluates the score of the surrounding of brightness channel of rust having dimension $2\ by\ 2$ filter along with 2 step size, this implies that the max pooling withholds data that is more useful in comparison the average pooling for brightness channel in the images of leaves having infected areas. Hence, in the proposed feature-space location, maximum pooling is implemented as down sampling of the image to attain a sequence of attribute map, that are observed to be in the similar scale as the DMFN layers of prediction. The layers that are generated in the feature-space location as well as the attribute maps obtained using the new hierarchical feature is expressed by the equation given below in (1). Considering the (1), the count of layers used for prediction is represented as p. The channel size summation for the combined technique is used to result in fusion of infected area data as well as semantic data which is formulated as given in (2).

$$X_r = \{x_1, x_2, .., x_p\}$$
$$H_r = \{h_1, h_2, .., h_p\} \tag{1}$$

$$R_k = \mu_1 H_k + \mu_1 X_k \tag{2}$$

Here, μ_1 is used to denote the added convolutional layer of dimension $1\ by\ 1$. As well as the operation of batch normalization once the concatenation of both the layer of H_k and X_k is done. Then, seven various levels of prediction layers are introduced.

3.3. DHAM

The dataset of the leaves that are infected with diseases contains small object areas such as spots of infections in areas having dimensions $32\ by\ 32$. Along with larger object areas where the infected areas are spread out over a larger section with dimensions $96\ by\ 96$. Additionally, there are few infections such

as rusts as well as brown areas having medium object area with the dimensions ranging between these two sizes. Therefore, a single level detection mechanism cannot be applied for this research.

As the leaves have infection areas of various dimensions, a DHAM is implemented in this mode for automatic learning of essential attribute channels at various levels for improvised accuracy detection of diseased area at various stages. For improvisation of the performance of detecting diseased areas, the DHAM implemented in the proposed prediction model uses the integration of feature maps of various dimensions. The layer of convolution is used in place of completely linked layers and to further enhance the capacity to represent attributes. The DHAM is developed by integration of the attention scheme as well as the hierarchical features. The model is capable of automatically grasping the significance of every feature channel based on the dimensions of the surrounding box by aiming at the correlation of various channels as well as automatic allocation of weights for various attribute maps. Furthermore, the level of adaptation is recognized based on the level of significance that results in detection that is accurate at various stages of infected areas. This is observed in Figure 2 and Figure 3 given below.
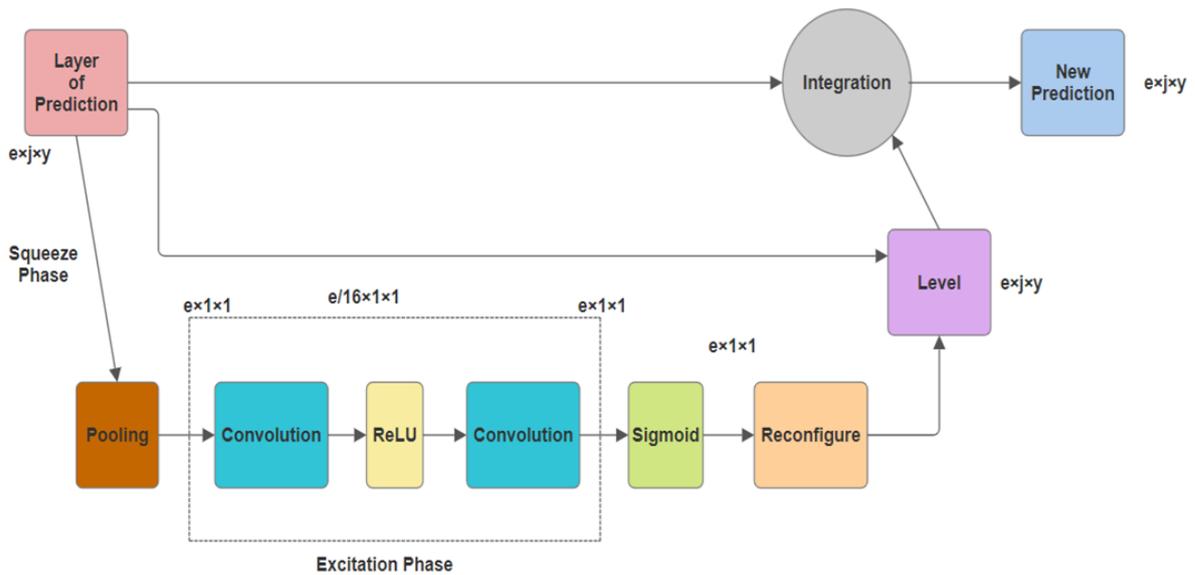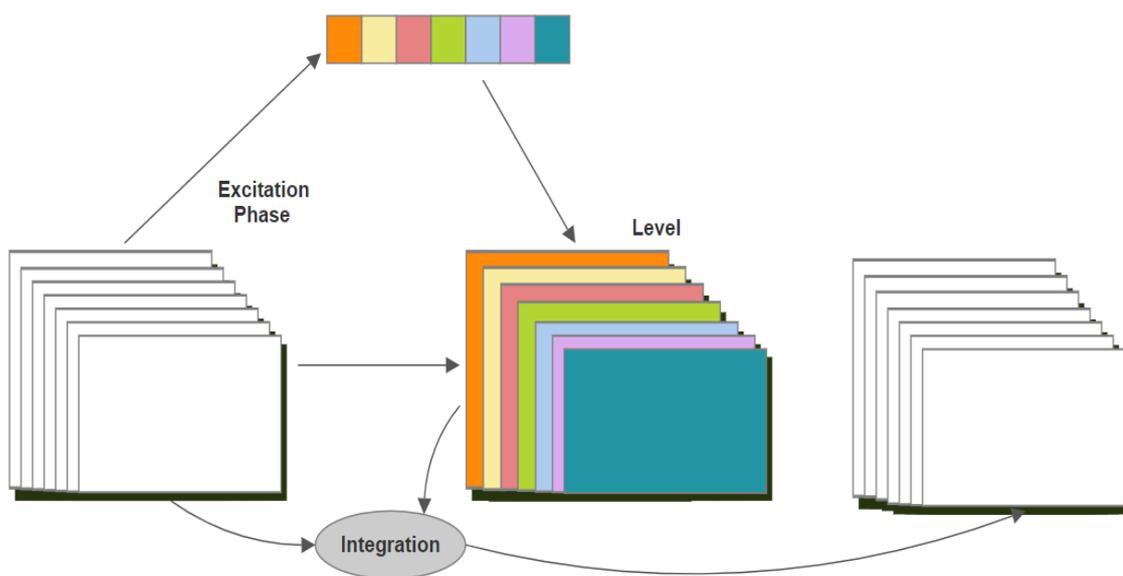

Figure 2. HAFFN model attention function

Figure 3. DHAM implementation of HAFFN model

At the squeeze phase, the average pooling is implemented for encoding the complete spatial attributes into features of one single dimension. This converts every channel attribute map into the output of dimension $1 \ by \ 1 \ by \ C$. Considering the input is expressed as $z \ belongs \ to \ T^{J \times Y \times E}$, here we formulate for b as given in (3). While we consider the excitation phase, the given in (4) is utilized to understand the nonlinear relation of various channels as given in (4).

$$b_e = H_{squeeze} z_e = (Y \times J)^{-1} \left(\sum_{k=1}^{Y} \sum_{l=1}^{J} z_e(k,l) \right) \quad b_e \ belongs \ to \ T^E \tag{3}$$

$$u = H_{excitation}(b,E) = sigmoid(E_2 ReLU(E_1 b)) \quad u \ belongs \ to \ T^{1 \times 1 \times E} \tag{4}$$

Here, $E_1 \ belongs \ to \ T^{\left(\frac{E}{t}\right) \times 1 \times 1}$ and $E_2 \ belongs \ to \ T^{E \times 1 \times 1}$, where $t = 16$. The capacity of generalization is enhanced in the HAFFN model for detection of diseases in leaves by adding two layers of convolution to the attention scheme. Further, these attention layers u_e is utilized to re-level z_e, this is formulated as given below in the (5). In conclusion, the resulting $\tilde{z}_e$ is integrated with the multiple level attribute maps that is generated by the DMFN method for improved performance in detection of multiple level disease detection in leaves.

$$\tilde{z}_e = H_{level}(z_e, u_e) = u_e . z_e \tag{5}$$

3.5. HAFFN model training

For the proposed feature space based multiple level feature integration DMFN, every infected area in the picture used as input, the box is chosen in default having max Intersection of Union being a positive sample. Furthermore, the Intersection of Union higher than 0.5 box from the other boxes is chosen in order to match the area that is infected. Simultaneously, it is essential that we maintain a balance for both the negative as well as positive samples that are utilized during the process of training, the initial m negative data samples having highest score of loss is withheld, a ratio of 1:3 is maintained which represents the positive to negative sample ratio making the optimization of the model enhanced in performance and increasing the stability of training. The rate of learning initially is maintained at $\frac{1}{10^4}$ for the initial 80 thousand iterations, this value is reduced at $\frac{1}{10^5}$ for 80 thousand iteration and then set at $\frac{1}{10^6}$ at the next 100 thousand iterations. Further, momentum $= 0.9$. The objection function used for training is given by the (6) below that has a integrated weight between the loss of localization as well as confidence. For the (6) given above, the loss of localization is based on Smooth $N1$ and the confidence loss is based on softmax loss. Here, P represented the count of boxes that are matched. The bound box is represented as n and the ground truth is denoted as i. The weight is $\gamma = 1$.

$$N(z,e,n,i) = (P)^{-1} \left(N_{confidence}(z,e) + \gamma N_{localization}(z,n,i) \right) \tag{6}$$

4. RESULTS AND DISCUSSION
4.1. Dataset details

In this work, rice leaf diseases were categorized using image analysis of the leaves. Four separate classes have been created from the dataset: brown spot, leaf blast, hispa, and healthy. 2,658 tagged photos in all were taken from Kaggle datasets that were available to the public [30]. Following collection, the dataset was split into training, validation, and testing sets. During training, the model was only given access to the training and validation datasets; it was not given access to the testing dataset. The model was evaluated by means of the testing and validation sets. Four example photos, one for each class in the dataset, are shown in Figure 4.

4.2. Comparison model
4.2.1. Leafnet model

The LeafNet model [31] consists of eleven convolutional layers followed by three fully-connected layers. Figure 4 displays the architecture of the LeafNet model. The model has distinct kernel filters in each convolutional layer, a 256×256 input size, and a 3×3 max pooling filter. The filters in the preceding two layers were 9×9, while the filters in the following two levels were adjusted to 5×5. The next levels kept the settings of the 3×3 kernel filter intact. The architecture of the model consisted of two thick layers, each with 2048 neurons. Its output layer had a SoftMax activation function, and all of its

layers were fully connected. The output layer of the model was set up to represent four classes instead of 185.
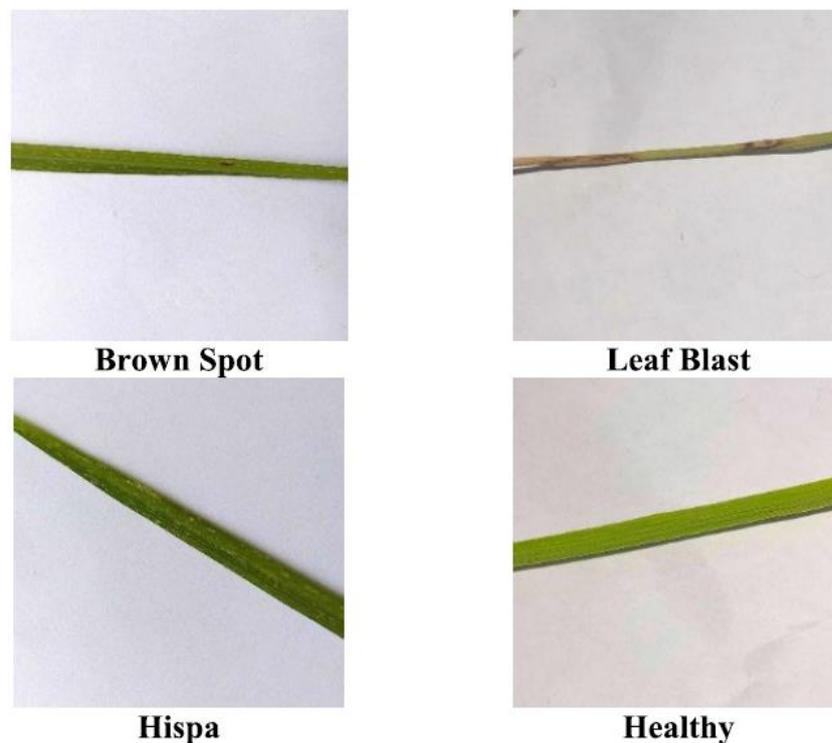

Figure 4. sample dataset

4.2.2. Modified leafnet model

To improve rice leaf disease categorization, a modified version of the LeafNet model [32] was created. Several changes were made to the LeafNet architecture to improve its performance for this application. Convolutional layers with filters measuring 9×9, 5×5, and 3×3 were used in the development of the LeafNet model, allowing for the integration of a variety of kernel sizes into its architecture. The modified version uses a 3×3 kernel size for all convolutional layers. The purpose of this modification was to improve the model's effectiveness and dependability in recognizing items in pictures of rice leaves. In its pooling layer, the original LeafNet architecture used a 2×2 max-pooling filter. The 3×3 max-pooling layer in the new model improves downsampling and feature extraction, particularly in pictures with small variations like those of unhealthy rice leaves.

4.2.3. Pretrained models

This study employed the Xception and MobileNetV2 models as comparison frameworks to highlight the differences between intricate structures and simpler alternatives. Deep CNNs, distinguished by their distinct input dimensions and architectural configurations, are exemplified by the two. The Xception model was set up with input dimensions of $299 \times 299 \times 3$, whereas the MobileNetV2 model used an input size of $224 \times 224 \times 3$. To achieve uniformity, the input dimensions for both models in this investigation were standardized to $224 \times 224 \times 3$. The benchmarks used in this work to demonstrate the trade-offs between complicated and lightweight designs are the Xception and MobileNetV2 models. Optimizing performance is given top priority in the model suggested for the rice leaf dataset. The outcomes of these pretrained models provide as a standard by which to compare the models' appropriateness and comparative effectiveness.

4.3. Results

Figure 5 provides the precision values for five different models—Modified LeafNet, LeafNet, Xception, MobileNetV2, and the HAFFN model—across four plant disease categories: Brown Spot, Healthy, Hispa, and Leaf Blast. The HAFFN model exhibits the highest precision across all categories,

with standout values in Leaf Blast 0.9835 and Brown Spot 0.9545, making it the most effective model for classifying these diseases. In contrast, the Modified LeafNet also performs well, especially in Leaf Blast 0.96 and Healthy 0.93, but falls slightly behind the HAFFN model. For the LeafNet model, the precision drops significantly in the Hispa category 0.57, indicating a challenge in accurately identifying this class. The Xception and MobileNetV2 models display the least overall precision, particularly for Brown Spot both at 0.68, suggesting lower reliability in disease classification. Additionally, both models struggle with Hispa, achieving precision scores of 0.58 and 0.6, respectively. The MobileNetV2 model has consistently lower precision across all categories, with its lowest value in Healthy and Leaf Blast 0.68 and 0.74. In summary, the HAFFN model significantly outperforms the other models, particularly excelling in difficult classes like Hispa 0.8595 and Leaf Blast, which further highlights its robustness and reliability across diverse disease categories. The Modified LeafNet still shows competitive performance but with a slight drop in accuracy for certain categories like Hispa. The other models, particularly Xception and MobileNetV2, show clear limitations and would likely require improvement or fine-tuning for better classification results in more challenging disease classes.

Figure 6 presents values for five models—Modified LeafNet, LeafNet, Xception, MobileNetV2, and the HAFFN model—across four disease classes: Brown Spot, Healthy, Hispa, and Leaf Blast. The HAFFN model stands out with near-perfect precision in multiple categories, achieving perfect scores for Brown Spot 1.0 and the highest precision for Healthy 0.9909, Hispa 0.8995, and Leaf Blast 0.9932. This model demonstrates its ability to consistently and accurately classify all categories, especially challenging ones like Hispa. Modified LeafNet follows closely behind, particularly with strong performance in Brown Spot 1.0 and Leaf Blast 0.96, but it lags in the Hispa class with a precision of only 0.66, indicating some difficulty in classifying this disease compared to the HAFFN model. LeafNet performs well in Brown Spot 0.98 and Healthy 0.93, but it significantly underperforms in Leaf Blast 0.48, showing a substantial drop in precision in this category, which may affect its overall reliability for this particular class. The Xception model demonstrates solid performance in Hispa 0.78, surpassing LeafNet and MobileNetV2 in this category, but it underperforms in the Brown Spot and Healthy classes with scores of 0.87 and 0.73, respectively. MobileNetV2 shows weaker performance overall, with its best result in Brown Spot 0.78 but low precision for the remaining classes, particularly Healthy 0.7 and Leaf Blast 0.6, indicating significant challenges in handling these disease types.
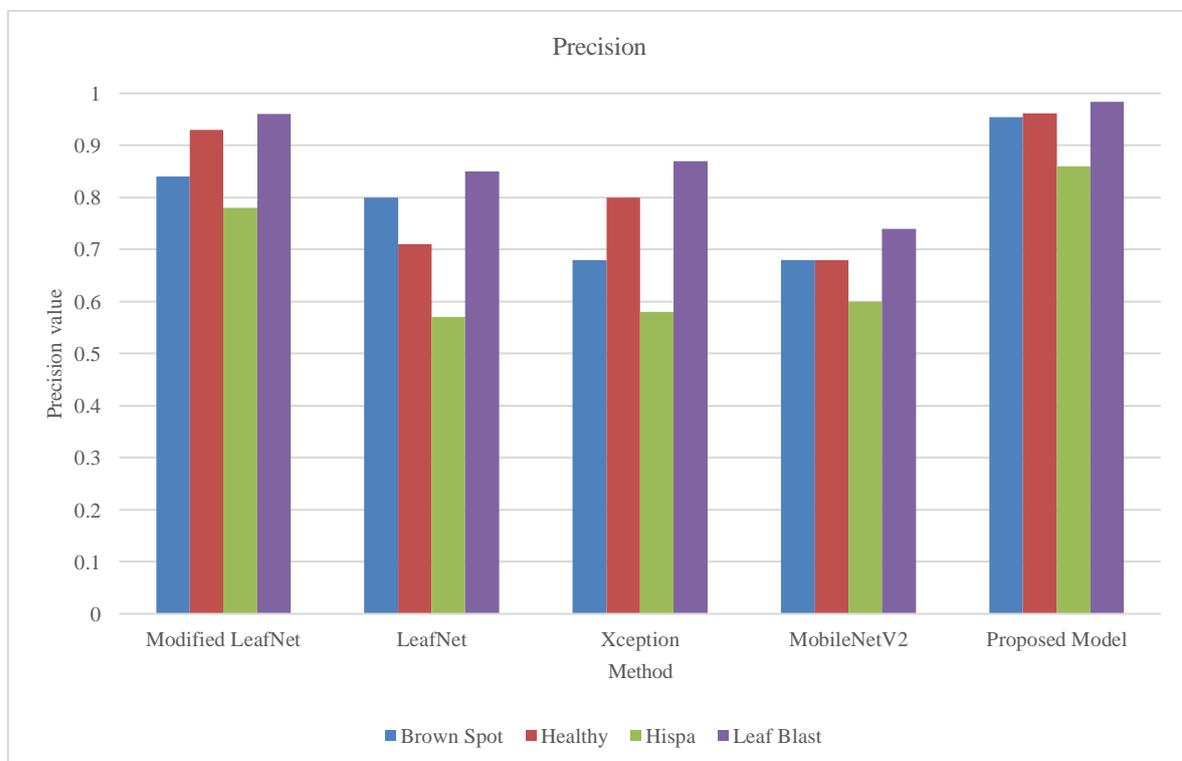

Figure 5. Precision graph

The F1-score analysis shows in Figure 7 that the HAFFN model consistently outperforms other models across all disease categories. For Brown Spot and Healthy, it achieves the highest scores of 0.9692 and 0.9559, respectively, significantly surpassing LeafNet and MobileNetV2. In the challenging Hispa class, the HAFFN model scores 0.8538, outperforming LeafNet (0.58) and Xception (0.66), demonstrating its ability to detect subtle infections. For Leaf Blast, it leads with 0.9061, again outperforming the other models. These results indicate the HAFFN model's superior accuracy and reliability in detecting rice leaf diseases.
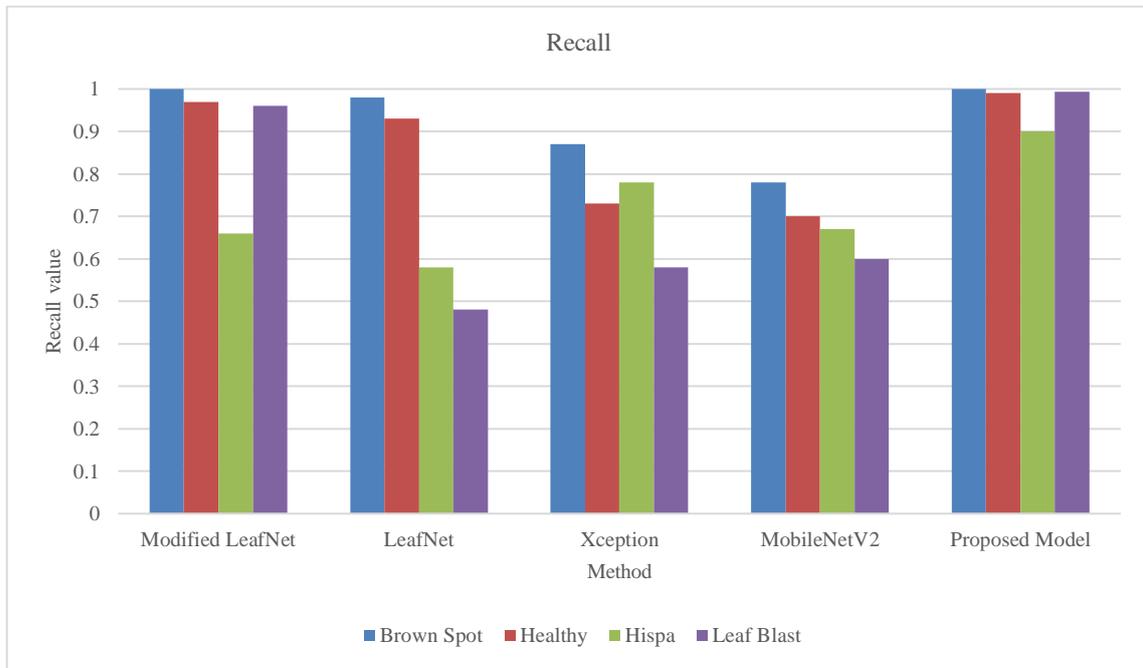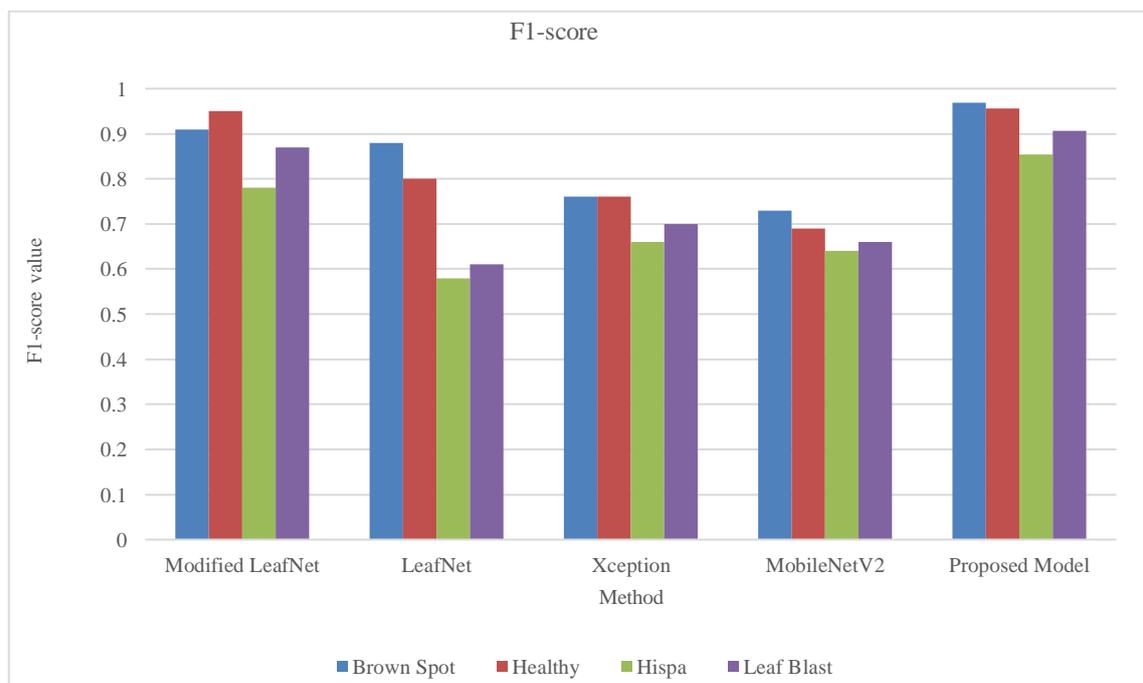

Figure 6. Recall comparison

Figure 7. F1-score comparison

Figure 8 presents accuracy data highlights the performance differences among five models—Modified LeafNet, LeafNet, Xception, MobileNetV2, and the HAFFN model—in terms of overall classification accuracy. The HAFFN model exhibits the highest accuracy at 93.4%, significantly outperforming the other models, making it the most reliable in classification tasks. The modified LeafNet follows with an accuracy of 87.76%, which is a strong performance but still notably lower than the HAFFN model. Both LeafNet (71.84%) and Xception (71.95%) show similar results, with only a slight edge for Xception, but their accuracy lags considerably behind the top two models. MobileNetV2 records the lowest accuracy at 67.68%, indicating substantial difficulties in classifying the target classes compared to the others. This data reinforces the conclusion that the HAFFN model is the most effective and accurate model overall, while modified LeafNet also shows competitive performance. In contrast, LeafNet, Xception, and MobileNetV2 exhibit weaker classification capabilities and would require optimization to improve their accuracy.
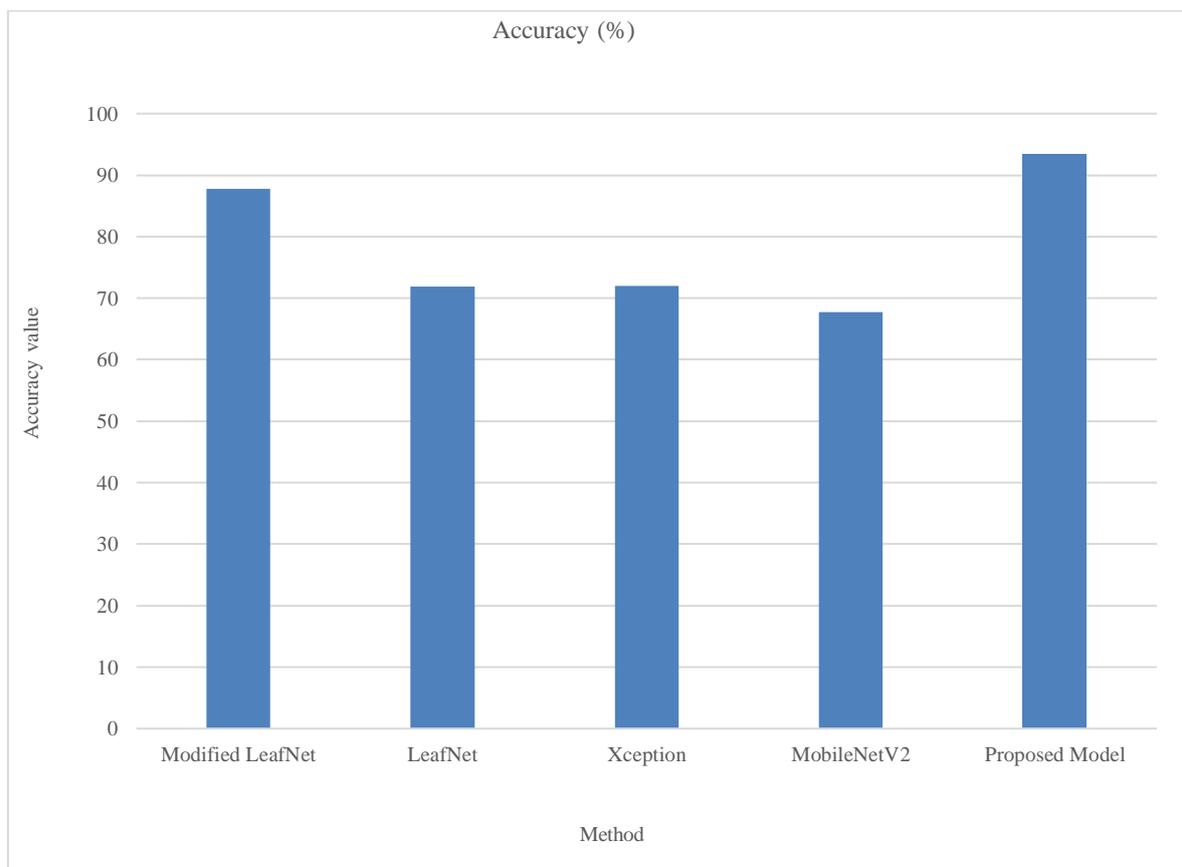

Figure 8. Accuracy comparison

5. CONCLUSION

This research introduces a novel HAFFN aimed at enhancing the detection of rice leaf diseases through advanced deep learning techniques. The HAFFN model addresses key challenges in rice disease detection, particularly the difficulty in identifying less infected areas, which are often overlooked by traditional methods. By integrating the DMFN with an AMFA and a DHAM, the HAFFN model can capture both small-scale and large-scale disease features. This multi-level feature extraction and attention mechanism contributes to more precise and adaptive detection of diseases such as Brown Spot, Hispa, and Leaf Blast. The model was trained and tested on a comprehensive rice leaf disease dataset, and the results demonstrated significant improvements over benchmark models, including LeafNet, Xception, and MobileNetV2. The HAFFN model achieved a notable 93.4% accuracy, with strong precision and recall metrics across all disease classes. Its ability to detect challenging disease features, such as less infected areas, makes it a reliable tool for large-scale, real-time agricultural disease management. This could help

mitigate the labor-intensive, time-consuming, and error-prone nature of manual disease detection in agriculture. Findings of this research have important implications for sustainable farming practices. The automated detection of rice leaf diseases enables early diagnosis and timely intervention, which can reduce the use of chemical pesticides and prevent crop losses, contributing to global food security.

AUTHOR CONTRIBUTION

This journal uses the Contributor Roles Taxonomy (CRediT) to recognize individual author contributions, reduce authorship disputes, and facilitate collaboration.

Name of Author	C	M	So	Va	Fo	I	R	D	O	E	Vi	Su	P	Fu
Tejaswini C. Gadag	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓			✓	
D. R. Kumar Raja								✓	✓	✓	✓	✓		

C : **C**onceptualization	I : **I**nvestigation	Vi : **Vi**sualization
M : **M**ethodology	R : **R**esources	Su : **Su**pervision
So : **So**ftware	D : **D**ata Curation	P : **P**roject administration
Va : **Va**lidation	O : Writing - **O**riginal Draft	Fu : **Fu**nding acquisition
Fo : **Fo**rmal analysis	E : Writing - Review & **E**diting	

CONFLICT OF INTEREST

Author declares no conflict of interest.

DATA AVAILABILTY

Dataset is utilized in this research mentioned in reference [30].

REFERENCES

[1]   K. Bashir, M. Rehman, and M. Bari, "Detection and classification of rice diseases: an automated approach using textural features," *Mehran University Research Journal of Engineering and Technology*, vol. 38, no. 1, pp. 239–250, Jan. 2019, doi: 10.22581/muet1982.1901.20.
[2]   S. M. Shaheen *et al.*, "Sustainable applications of rice feedstock in agro-environmental and construction sectors: a global perspective," *Renewable and Sustainable Energy Reviews*, vol. 153, p. 111791, Jan. 2022, doi: 10.1016/j.rser.2021.111791.
[3]   M. H. Bijoy *et al.*, "Towards sustainable agriculture: a novel approach for rice leaf disease detection using dCNN and enhanced dataset," *IEEE Access*, vol. 12, pp. 34174–34191, 2024, doi: 10.1109/ACCESS.2024.3371511.
[4]   M. M. Rahman *et al.*, "Ganges-brahmaputra-meghna delta, bangladesh and india: a transnational mega-delta," *Deltas in the Anthropocene*, pp. 23–51, 2020, [Online]. Available: https://doi.org/10.1007/978-3-030-23517-8_2.
[5]   M. D. Nirmal, P. P. Jadhav, and S. Pawar, "RETRACTED ARTICLE: pomegranate leaf disease detection using supervised and unsupervised algorithm techniques," *Cybernetics and Systems*, pp. 1–12, Jan. 2023, doi: 10.1080/01969722.2023.2166192.
[6]   "Rice Disease Management," 2023. https://extension.missouri.edu/programs/rice-extension/rice-diseases/ricedisease-management (accessed Jul. 01, 2023).
[7]   W.-C. Chen, T.-Y. Chiou, A. L. Delgado, and C.-S. Liao, "The control of rice blast disease by the novel biofungicide formulations," *Sustainability*, vol. 11, no. 12, p. 3449, Jun. 2019, doi: 10.3390/su11123449.
[8]   J. M. Al-Khayri *et al.*, "The role of nanoparticles in response of plants to abiotic stress at physiological, biochemical, and molecular levels," *Plants*, vol. 12, no. 2, p. 292, Jan. 2023, doi: 10.3390/plants12020292.

[9]   R. R. Patil and S. Kumar, "Rice-fusion: a multimodality data fusion framework for rice disease diagnosis," *IEEE Access*, vol. 10, pp. 5207–5222, 2022, doi: 10.1109/ACCESS.2022.3140815.

[10] L. Wei, Y. Luo, L. Xu, Q. Zhang, Q. Cai, and M. Shen, "Deep convolutional neural network for rice density prescription map at ripening stage using unmanned aerial vehicle-based remotely sensed images," *Remote Sensing*, vol. 14, no. 1, p. 46, Dec. 2021, doi: 10.3390/rs14010046.

[11] Y. Yuan, L. Chen, H. Wu, and L. Li, "Advanced agricultural disease image recognition technologies: a review," *Information Processing in Agriculture*, vol. 9, no. 1, pp. 48–59, Mar. 2022, doi: 10.1016/j.inpa.2021.01.003.

[12] Y. Kurmi, P. Saxena, B. S. Kirar, S. Gangwar, V. Chaurasia, and A. Goel, "Deep CNN model for crops' diseases detection using leaf images," *Multidimensional Systems and Signal Processing*, vol. 33, no. 3, pp. 981–1000, Sep. 2022, doi: 10.1007/s11045-022-00820-4.

[13] P. S. Thakur, P. Khanna, T. Sheorey, and A. Ojha, "Explainable vision transformer enabled convolutional neural network for plant disease identification: PlantXViT," *arXiv:2207.07919*, 2022, [Online]. Available: http://arxiv.org/abs/2207.07919.

[14] S. Wu, Y. Sun, and H. Huang, "Multi-granularity feature extraction based on vision transformer for tomato leaf disease recognition," in *2021 3rd International Academic Exchange Conference on Science and Technology Innovation (IAECST)*, Dec. 2021, pp. 387–390, doi: 10.1109/IAECST54258.2021.9695688.

[15] S. Yu, L. Xie, and Q. Huang, "Inception convolutional vision transformers for plant disease identification," *Internet of Things*, vol. 21, p. 100650, Apr. 2023, doi: 10.1016/j.iot.2022.100650.

[16] H. Wu *et al.*, "CvT: introducing convolutions to vision transformers," in *2021 IEEE/CVF International Conference on Computer Vision (ICCV)*, Oct. 2021, pp. 22–31, doi: 10.1109/ICCV48922.2021.00009.

[17] H. Touvron, M. Cord, M. Douze, F. Massa, A. Sablayrolles, and H. Jégou, "Training data-efficient image transformers and distillation through attention," *Proceedings of Machine Learning Research*, vol. 139, pp. 10347–10357, 2021, [Online]. Available:https://proceedings.mlr.press/v139/touvron21a.html.

[18] J. Gou, B. Yu, S. J. Maybank, and D. Tao, "Knowledge distillation: a survey," *International Journal of Computer Vision*, vol. 129, no. 6, pp. 1789–1819, Jun. 2021, doi: 10.1007/s11263-021-01453-z.

[19] Y. Zhao, H. Tang, Y. Jiang, Y. A, and Q. Wu, "Lightweight vision transformer with cross feature attention," *arXiv:2207.07268*, 2022, [Online]. Available: http://arxiv.org/abs/2207.07268.

[20] G. Luo *et al.*, "Towards lightweight transformer via group-wise transformation for vision-and-language tasks," *IEEE Transactions on Image Processing*, vol. 31, pp. 3386–3398, 2022, doi: 10.1109/TIP.2021.3139234.

[21] M. H. Hossen, M. Mohibullah, C. S. Muzammel, T. Ahmed, S. Acharjee, and M. B. Panna, "Wheat diseases detection and classification using convolutional neural network (CNN)," *International Journal of Advanced Computer Science and Applications*, vol. 13, no. 11, pp. 719–726, 2022, doi: 10.14569/IJACSA.2022.0131183.

[22] X. Jin, L. Jie, S. Wang, H. Qi, and S. Li, "Classifying wheat hyperspectral pixels of healthy heads and fusarium head blight disease using a deep neural network in the wild field," *Remote Sensing*, vol. 10, no. 3, p. 395, Mar. 2018, doi: 10.3390/rs10030395.

[23] A. Hussain, M. Ahmad, I. A. Mughal, and H. Ali, "Automatic disease detection in wheat crop using convolution neural network," in *The 4th International Conference on Next Generation Computing 2018*, Dec. 2018, pp. 7–10, doi: 10.13140/RG.2.2.14191.46244.

[24] W. Bao, X. Yang, D. Liang, G. Hu, and X. Yang, "Lightweight convolutional neural network model for field wheat ear disease identification," *Computers and Electronics in Agriculture*, vol. 189, p. 106367, Oct. 2021, doi: 10.1016/j.compag.2021.106367.

[25] A. Alharbi, M. U. G. Khan, and B. Tayyaba, "Wheat disease classification using continual learning," *IEEE Access*, vol. 11, pp. 90016–90026, 2023, doi: 10.1109/ACCESS.2023.3304358.

[26] S. Savary *et al.*, "A white paper on global wheat health based on scenario development and analysis," *Phytopathology®*, vol. 107, no. 10, pp. 1109–1122, Oct. 2017, doi: 10.1094/PHYTO-01-17-0027-FI.

[27] P. K. Sethy, N. K. Barpanda, A. K. Rath, and S. K. Behera, "Deep feature based rice leaf disease identification using support vector machine," *Computers and Electronics in Agriculture*, vol. 175, p. 105527, Aug. 2020, doi: 10.1016/j.compag.2020.105527.

[28] S. Ghosal and K. Sarkar, "Rice leaf diseases classification using CNN with transfer learning," in *2020 IEEE Calcutta Conference (CALCON)*, Feb. 2020, pp. 230–236, doi: 10.1109/CALCON49167.2020.9106423.

[29] V. K. Shrivastava and M. K. Pradhan, "Rice plant disease classification using color features: a machine learning paradigm," *Journal of Plant Pathology*, vol. 103, pp. 17–26, 2020, doi: 10.1007/s42161-020-00683-3.

[30] Vbookshelf, "Rice leaf diseases," *Kaggle*, 2020. https://www.kaggle.com/datasets/vbookshelf/rice-leaf-diseases (accessed May 01, 2023).

[31] P. Barré, B. C. Stöver, K. F. Müller, and V. Steinhage, "LeafNet: a computer vision system for automatic plant species identification," *Ecological Informatics*, vol. 40, pp. 50–56, Jul. 2017, doi: 10.1016/j.ecoinf.2017.05.005.

[32] W. I. A. E. Altabaji, M. Umair, W.-H. Tan, Y.-L. Foo, and C.-P. Ooi, "Comparative analysis of transfer learning, leafnet, and modified LeafNet models for accurate rice leaf diseases classification," *IEEE Access*, vol. 12, pp. 36622–36635, 2024, doi: 10.1109/ACCESS.2024.3373000.

BIOGRAPHIES OF AUTHORS

Tejashwini C. Gadag 🆔 ⸬ ⒮⒞ ◖ holds a master's computer network engineering degree from Ramaiah Institute of Technology, Bengaluru, India. She is currently pursuing a Ph.D. in computer science and engineering from Reva University, Bengaluru. She is currently working as an assistant professor in the School of Computer Science and Engineering at Reva University. Her research interests include machine learning, deep learning, and cloud computing. She can be contacted at email: tejashwinicg_12@rediffmail.com.

Dr. D. R. Kumar Raja 🆔 ⸬ ⒮⒞ ◖ is currently working as professor in the School of Computer Science and Engineering at REVA University Bangaluru, Karnataka, India. He received his Bachelor of Technology (B.Tech.) from JNTUA College of Engineering and Master of Technology (M.Tech.) from National Institute of Technology Karnataka (NITK) Surathkal, Karnataka, India. He received a Doctorate of Philosophy (Ph.D.) from St Peters University, Chennai, India for an effective context-driven recommender system for e-commerce applications. His research areas include the IoT, data mining, machine learning and AI. He can be contacted at email: Kumarraja.dr@reva.edu.in.