SSRG International Journal of Electronics and Communication Engineering
ISSN: 2348-8549/ https://doi.org/10.14445/23488549/IJECE-V11I7P121

Volume 11 Issue 7, 208-219, July 2024

Original Article

Optimized Strategy for Rice Plant Disease Detection Using Convolutional Neural Networks

R. Rajakumar[1], D. Vaishnavi[1]*, R. Ramesh[1], J. Ganesh[1]

[1]*Department of CSE, Srinivasa Ramanujan Centre, SASTRA University, Tamilnadu, India.*

Corresponding Author : vaishnavi@src.sastra.edu

Received: 14 May 2024 Revised: 12 June 2024 Accepted: 13 July 2024 Published: 27 July 2024

Abstract - *Image classification has evolved greatly in recent years, owing to the development of machine learning algorithms and the availability of large-scale image datasets. Convolutional Neural Networks (CNNs) have profoundly impacted the field of image classification due to their ability to learn hierarchical representations directly from pixel data. Unlike traditional machine learning algorithms, which rely on handcrafted features, CNNs can extract information hierarchically through multiple convolutional and pooling layers. Initially, compared various feature extraction techniques such as Histogram of Oriented Gradients (HOG), Local Binary Pattern (LBP), Scale Invariant-Future Transform (SIFT), Speeded-Up Robust Features (SURF), and wavelet domains like Discrete Cosine Transform (DCT) and Discrete Wavelet Transform (DWT) with standard classifiers. In order to improve the results, design a novel CNN model for image classification tasks, with an emphasis on hyperparameter optimization, data augmentation, dropout regularization, and efficient data loading. Experimental results on benchmark datasets show that our proposed CNN model outperforms baseline models, with an improving accuracy of 76.00%. These findings demonstrate the usefulness of our approach in advancing the state-of-the-art forward in image classification task.*

Keywords - *Image classification, Feature extraction, Convolutional neural network, Data augmentation, Efficient data loading.*

1. Introduction

Rice, as the preferred staple food for a substantial portion of the world's population, is essential to maintaining global food security. However, a number of diseases that pose rice farming must be recognized and controlled quickly in order to avoid yield losses. In recent years, advances in computer vision, machine learning, and remote sensing technologies have transformed the field of rice plant disease identification. These technologies provide non-destructive, quick, and cost-effective ways to diagnose illnesses in rice plants, allowing for prompt treatments to reduce crop losses and boost agricultural production.

This study provides an overview of the most recent advances in rice plant disease detection strategies, including the utilization of image processing techniques, deep learning algorithms, and Unmanned Aerial Vehicles (UAVs). It also addresses the field's challenges, possibilities, and future directions, with the goal of helping to improve sustainable agriculture and food security projects around the world.

Image feature extraction is significant in computer vision because it transforms raw pixel data into structured representations that improve efficiency, robustness, and compatibility with machine learning models. It reduces dimensionality, improves interpretability, and facilitates tasks like classification and detection across a variety of conditions. This study conducts a comparative analysis of feature extraction techniques for image classification tasks using Convolutional Neural Networks (CNNs), Scale-Invariant Feature Transform (SIFT), and Histogram of Oriented Gradients (HOG), among others. Furthermore, we explore the effectiveness of several classifiers when combined with these feature extraction strategies, which include both classic machine learning algorithms and deep learning models.

Data augmentation and hyperparameter tuning are important strategies for optimizing image classification algorithms. Data augmentation enhances the training dataset by introducing transformations such as rotations, flips, and zooms, which diversity the data observed by the model. This method improves model robustness, allowing it to generalize more well to new data and lowering the danger of overfitting. Furthermore, data augmentation is especially useful when labeled data is limited, maximizing the utility of available samples without requiring further annotation efforts. Hyperparameter tuning, on the other hand, is concerned with optimizing parameters that the model does not directly learn, such as learning rates, batch sizes, and network designs. This iterative method entails experimenting with various settings, training the model, and assessing performance metrics using validation data. Practitioners may greatly improve model performance by fine-tuning hyperparameters based on individual tasks and datasets, ensuring optimal accuracy and efficiency in image classification.

By conducting a rigorous comparative analysis, we aim to provide insights into the strengths and weaknesses of different image feature extraction techniques and classifiers. Our findings will be useful for researchers, practitioners, and stakeholders working on the development and deployment of image classification systems across diverse domains.

2. Related Works

Recent advancements in computer vision have significantly influenced various domains, including human detection, texture classification, and large-scale image recognition. Dalal and Triggs [1] proposed HOG for human detection, providing a robust method for capturing local object appearance and shape. Lowe [2] introduced Scale-Invariant Feature Transform (SIFT), enabling distinctive image feature extraction across different scales and orientations.

Ojala, Pietikäinen, and Maenpaa [3] developed Local Binary Patterns (LBP) for texture classification, offering a computationally efficient approach invariant to gray-scale and rotation. These traditional methods laid the foundation for subsequent advancements in deep learning architectures. Simonyan and Zisserman [4] introduced Very Deep Convolutional Networks (VGG) for large-scale image understanding, demonstrating the efficacy of deep learning in learning hierarchical representations.

He et al. [5] offered Deep Residual Learning, addressing the challenges of training deeper networks by introducing residual connections. These works collectively highlight the evolution of computer vision methodologies from traditional feature-based approaches to deep learning paradigms, driving progress in various applications such as object recognition and image classification.

Over the past few years, there has been a paradigm shift in computer vision towards more sophisticated architectures and methodologies. Szegedy et al. [6] introduced the Inception architecture, redefining the design principles for Convolutional Neural Networks (CNNs) by incorporating multiple parallel convolutional pathways. This approach aimed to improve computational efficiency while maintaining high accuracy in image recognition tasks.

Building upon this, Tan and Le [7] proposed EfficientNet, a novel scaling method that balances model depth, width, and resolution to optimize overall efficiency and performance. Dosovitskiy et al. [8] further expanded the horizon with transformers, a class of models originally developed for natural language processing but adapted successfully for image recognition tasks. This pioneering work demonstrated the potential of leveraging self-attention mechanisms for capturing long-range dependencies in images. In parallel, the concept of radiomics emerged, emphasizing the extraction and analysis of quantitative features from medical images as data beyond mere visual representations [9]. This holistic approach has led to significant advancements in medical imaging analysis and diagnosis.

Additionally, attention mechanisms have gained prominence in the deep learning community, with Vaswani et al. [10] proposing the Transformer model, which relies solely on attention mechanisms for feature extraction and classification, showing promising results in various domains beyond natural language processing. Together, these innovations underscore the diverse range of approaches reshaping the landscape of computer vision, from reimagining neural network architectures to harnessing the power of attention mechanisms and quantitative image analysis techniques.

Concurrently, novel architectures like MobileNetV2, introduced by Sandler et al. [11], have pushed the boundaries of convolutional neural networks (CNNs) by incorporating inverted residuals and linear bottlenecks, thereby improving model efficiency without compromising accuracy.

Huang et al. [12] proposed densely connected convolutional networks, which establish direct connections between all layers within a block, enabling feature reuse and augmenting information flow throughout the network. Furthermore, Redmon et al. [13] presented You Only Look Once (YOLO), a unified real-time object detection system that processes images in a single pass, demonstrating remarkable efficiency and accuracy in detecting objects of interest. These pioneering works collectively underscore the importance of comparative evaluations and the continuous innovation in designing architectures that address key challenges in computer vision, such as feature extraction, classification, and real-time object detection.

Recent years have seen amazing progress in the fields of computer vision and machine learning, thanks to ground-breaking research projects. Russakovsky et al. [14] established a seminal milestone with the ImageNet Large Scale Visual Recognition Challenge, catalyzing progress in image classification and object recognition. Lin et al. [15] introduced focal loss as a pivotal technique for enhancing dense object detection, contributing to the refinement of object detection systems.

Shoaib Muhammad et al. [16] provided valuable insights into the application of machine learning for crop disease detection, showcasing the transformative potential of these technologies in agricultural science. Jones and Wang [17] conducted an extensive review of deep learning methodologies for plant disease finding, underscoring the significance of leveraging advanced computational techniques in agricultural informatics.

Meanwhile, Saleem et al. [18] focused on optimizing convolutional neural network architectures specifically for crop disease classification, with a case study centered on rice plants, highlighting the importance of tailored solutions for addressing domain-specific challenges in agriculture. These studies collectively underscore the profound impact of machine learning and deep learning approaches in revolutionizing various aspects of plant disease detection and agricultural practices.

Shamsuzzaman [19] investigated the potential of transfer learning in the domain of crop disease detection, shedding some light on the advantages and difficulties of this strategy in agricultural technology. Their research emphasizes how crucial it is to use pre-trained models and information transfer from relevant fields in order to improve the efficiency of agricultural disease detection systems.

Liu et al. [20] delved into scalable machine learning solutions tailored for precision agriculture, offering insights into the development of scalable and efficient algorithms to address the unique challenges of large-scale crop monitoring and management.

Additionally, Sandler et al. [11] introduced MobileNetV2, a novel architecture featuring inverted residuals and linear bottlenecks, which has occurred as a significant advancement in the arena of computer vision, mainly in the context of mobile and embedded systems. Using the 320-image Rice-leaf-disease dataset from Kaggle, Ahmad et al. developed an Xception model using transfer learning specifically for rice disease identification [21].

The study's notable performance measurements showed that the Xception model maintained a robust validation accuracy of 90% and reached 93% accuracy in training. The Xception model regularly outperformed VGG16, MobileNetV2, and EfficientNetV2 in studies involving disease categorization.

Based on evaluation metrics, the model performed well, with an accuracy of 0.89. Using leaf image analysis, Nandi Sunandar and Joko Sutopo used Artificial Neural Networks (ANN) to classify different forms of illnesses in rice plants [22]. Their study used training data from the UCI Machine Learning Repository, which included 24 records for testing and 56 records for training, to reach an 83% accuracy rate in disease detection.

Furthermore, Huang et al. [23] proposed densely connected convolutional networks, presenting a powerful framework for feature extraction and representation learning, thus contributing to the development of robust and effective deep learning models for various computer vision tasks.

Collectively, these studies underscore the pivotal role of advanced machine learning and deep learning techniques in addressing critical challenges in agricultural technology, ranging from crop disease detection to precision agriculture and model optimization for resource-constrained environments.

3. Materials and Methods

3.1. Major Contributions

Predictive modeling still faces numerous issues in the field of rice plant disease detection, according to the literature currently under publication. These difficulties include the interpretability of models, dynamic system behavior, complex and noisy real-world data, and a lack of labeled data.

Overcoming these challenges will require interdisciplinary collaboration and innovative thinking in order to produce trustworthy and transparent prediction models. Our research focuses on the following plans are as follows:

1. To detect rice plant diseases, various feature extraction techniques were compared using an optimal classifier.
2. Following the comparison, in order to improve the prediction rate, apply cosine and wavelet platforms.
3. After that, applied several deep learning and transfer learning models for accuracy improvisation.
4. To improve the prediction rate, lastly design a novel convolutional neural network model with appropriate optimization algorithms.

3.2. Proposed Model

3.2.1. Dataset Description

Initially, the dataset rice plant disease detection dataset downloaded from Kaggle [5] was used in this work for the experiments. It contains 120 images of multiple classes of rice diseases, such as blast, bacterial leaf blight, and sheath blight, among others. A sample dataset for rice plant disease detection is shown in Figures 1 and 2.

It depicts Brown Spots, bacterial leaf blight, and leaf smut. Each class has 40 images. It is possible that data augmentation techniques were employed to increase the dataset's diversity and robustness. All things considered, the dataset is a helpful resource for developing and testing machine learning models for the identification and categorization of illnesses affecting rice plants.

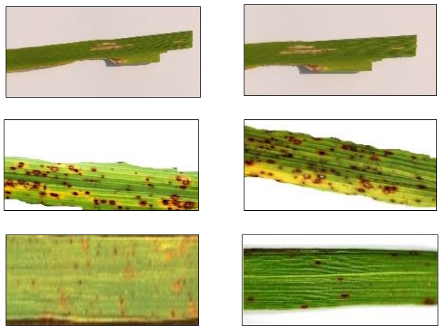

Fig. 1 Rice plant disease detection dataset samples

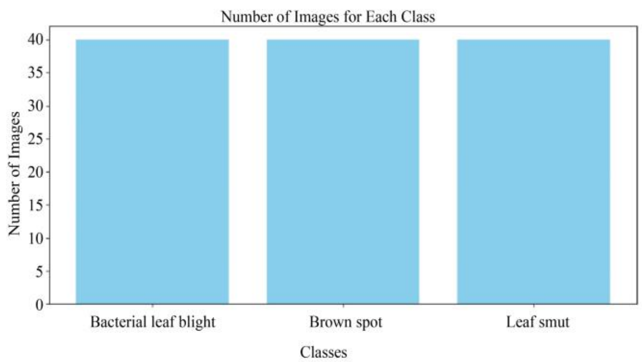

Fig. 2 EDA dataset exploratory data analytics

3.2.2. Preprocessing

Preprocessing assures that machine learning models can accurately distinguish between healthy and diseased plants by enhancing features, reducing noise, and normalizing input images. Here, we used a variety of preprocessing techniques like noise reduction, image segmentation, and contrast enhancement over the rice plant disease detection dataset. Preprocessing techniques improve the detection process by concentrating on pertinent features and not only setting up images for analysis but also raising the overall accuracy of rice plant disease identification.

3.2.3. Feature Extraction Methods with Classification

Concentrate several feature extraction strategies over the rice plant disease detection dataset after completing the preprocessing processes. Feature extraction plays a crucial part in image classification, as it goals to capture discriminative information from images that facilitate effective classification.

Initially explore several popular feature extraction methods, Histogram-based Techniques: Histogram of Oriented Gradients (HOG), Color Histograms, Discrete Wavelet Transform (DWT), Discrete Cosine Transform (DCT), Scale-Invariant Feature Transform (SIFT), Speeded-Up Robust Features (SURF), Local Binary Patterns (LBP). For each feature extraction method, utilize a Random Forests classifier and performance is measured using the conventional measure called accuracy.

Histogram of Oriented Gradients (HOG)

HOG is frequently used for applications involving object recognition and image categorization. It captures the local shape and texture information by computing histograms of gradient orientations in localized regions of the image. HOG descriptors are effective in scenarios where the spatial arrangement of gradients is essential for discrimination. HOG creates tiny cell divisions in the image. It calculates gradient orientations (angles) and magnitudes for every cell. Next, these gradients are assembled into bigger blocks. A feature vector of 36 points is gathered from every block.

Local Binary Patterns (LBP)

LBP focuses on texture information in images and is commonly used for texture classification tasks. It describes the local texture patterns by comparing the intensity of a central pixel with its neighbors, resulting in binary patterns' descriptors that are robust to changes in illumination and contrast, making them suitable for various texture-based classification tasks.

If a local binary pattern has no more than two 0-1 or 1-0 transitions, it is referred to as "uniform." By using uniform patterns, the feature vector length for a single cell is shortened from 256 to 59. The 58 consistent binary patterns map to distinct numbers, enabling a more straightforward and rotation-invariant descriptor.

Scale-Invariant Feature Transform (SIFT)

It detects and describes key points in images that are invariant to rotation, scale and illumination changes. It computes 128 feature vectors of key point descriptors based on gradient information in local image patches around key points. SIFT descriptors are effective for tasks requiring robustness to viewpoint changes, such as object recognition and image stitching.

Speeded-Up Robust Features (SURF)

It is an enhancement of SIFT designed for faster computation while maintaining similar performance. It uses integral images and box filters to accelerate key point detection and descriptor computation. SURF descriptors are suitable for real-time applications and scenarios where computational efficiency is critical.

SURF uses a precomputed integral image to approximate the determinant of the Hessian blob detector in an integer manner, which allows it to identify points of interest. Based on the sum of the wavelet responses surrounding the point of interest, the SURF feature descriptor is a 64-dimensional vector.

The following are the results of our feature extraction technique, classification, and performance calculations made with the aid of the methods mentioned above.

Table 1. Feature extraction performance analysis

Feature Extraction Model	Train $accu_y$ (%)	Test $accu_y$ (%)
HOG	0.66	0.41
LBP	0.75	0.41
SIFT	0.87	0.54
SURF	0.81	0.66

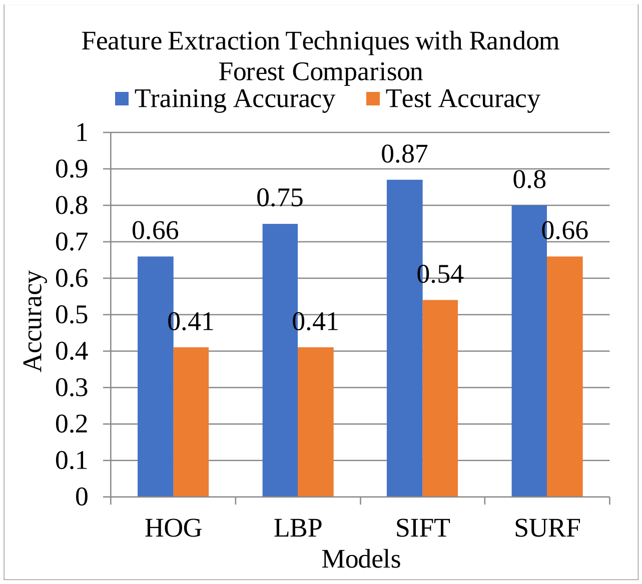

Fig. 3 Performance analysis for different feature extraction techniques

Table 1 and Figures 3 and 4. provide a brief comparative study of different feature extraction techniques with a Random forest classifier for the rice plant disease detection dataset. We have compared various feature extraction techniques like HOG, LBP, SIFT and SURF in terms of u_y. The experimental outcomes inferred that the Histogram of Oriented Gradients (HOG) algorithm had reached the lowest performance at $accu_y$ 41.00 %, whereas the Speeded-Up Robust Features (SURF) algorithm showed a somewhat high $accu_y$ of 66.00%.

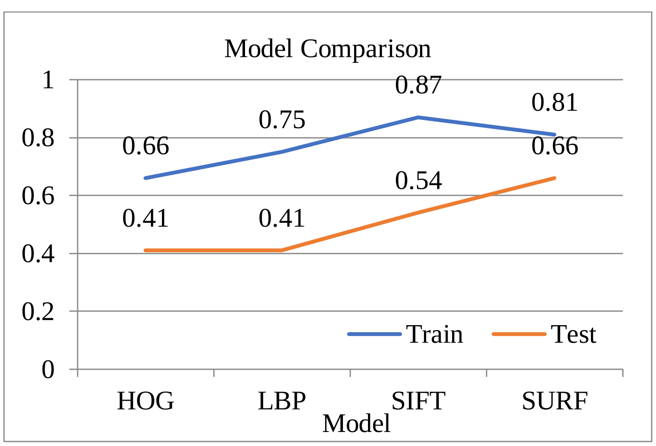

Fig. 4 Model accuracy using different feature extraction techniques

3.3. Transform Domain with Classification

Here, we have focused on the transform domain for rice plant disease identification in an effort to improve accuracy. So, the traditional and efficient transform domain-based algorithms such as DCT and DWT are employed in this work.

3.3.1. Discrete Cosine Transform (DCT)

DCT has garnered significant attention for its ability to represent image data in the frequency domain efficiently. By decomposing rice plant images into frequency components, DCT facilitates the extraction of essential features related to disease symptoms and plant health.

This work provides a comprehensive overview of the application of DCT in rice plant disease detection, highlighting its advantages, challenges, and potential for enhancing disease diagnosis accuracy and efficiency.

3.3.2. Discrete Wavelet Transform (DWT)

It is a prevailing signal processing technique that decomposes an image into different frequency bands using wavelet functions. From the perspective of rice plant disease recognition, DWT can be applied to analyze the texture and structural characteristics of rice plant images, which are essential for identifying disease-related patterns.

The image (n) is successively filtered and down-sampled to obtain approximation coefficients cA and detail coefficients cD at different scales. These coefficients capture image characteristics such as texture, edges, and structural information at different scales.

This process results in a multi-resolution representation of the image, capturing both global and local image features. The classifier learns to distinguish between healthy and diseased rice plants based on the characteristic patterns present in the feature space.

This understanding facilitates the optimization of DWT parameters and the development of efficient classification algorithms tailored to the problem domain. After applying the DCT and DWT separately, we have obtained the results using a random forest classification algorithm and its results are displayed in Table 2.

Table 2. Cosine and wavelet transform accuracy comparison

Models	Classifier	Train $accu_y$ (%)	Test $accu_y$ (%)
Discrete Cosine Transform (DCT)	Random Forest	0.66	0.75
Discrete Wavelet Transform (DWT)	Random Forest	0.66	0.41

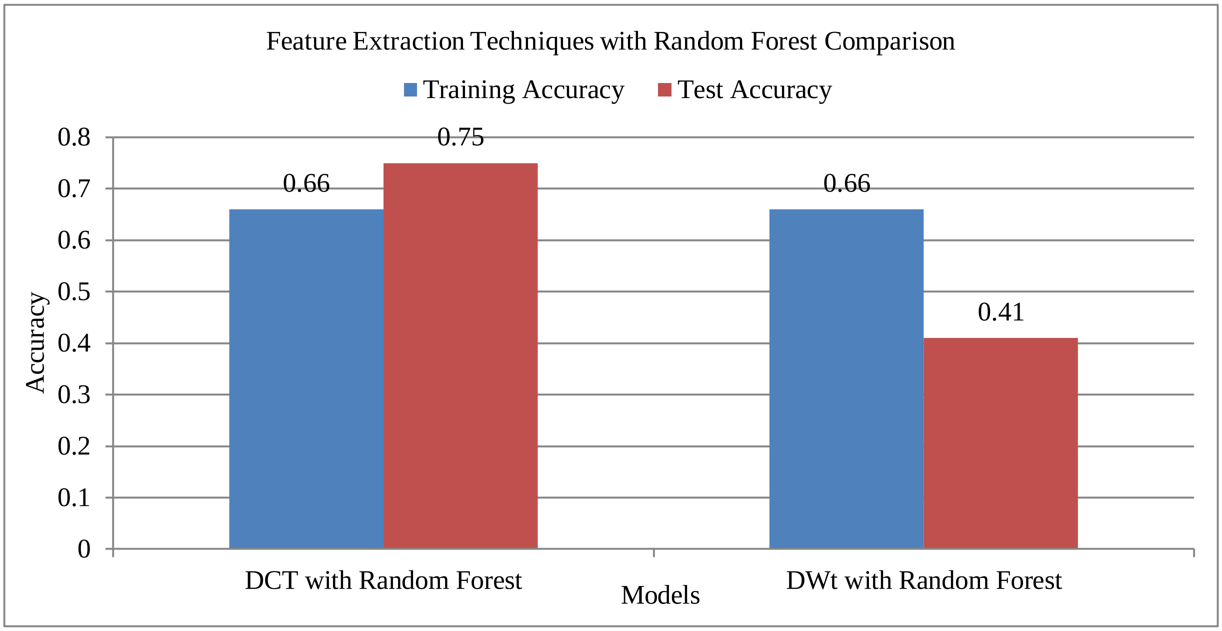

Fig. 5 Cosine and wavelet transform accuracy comparison

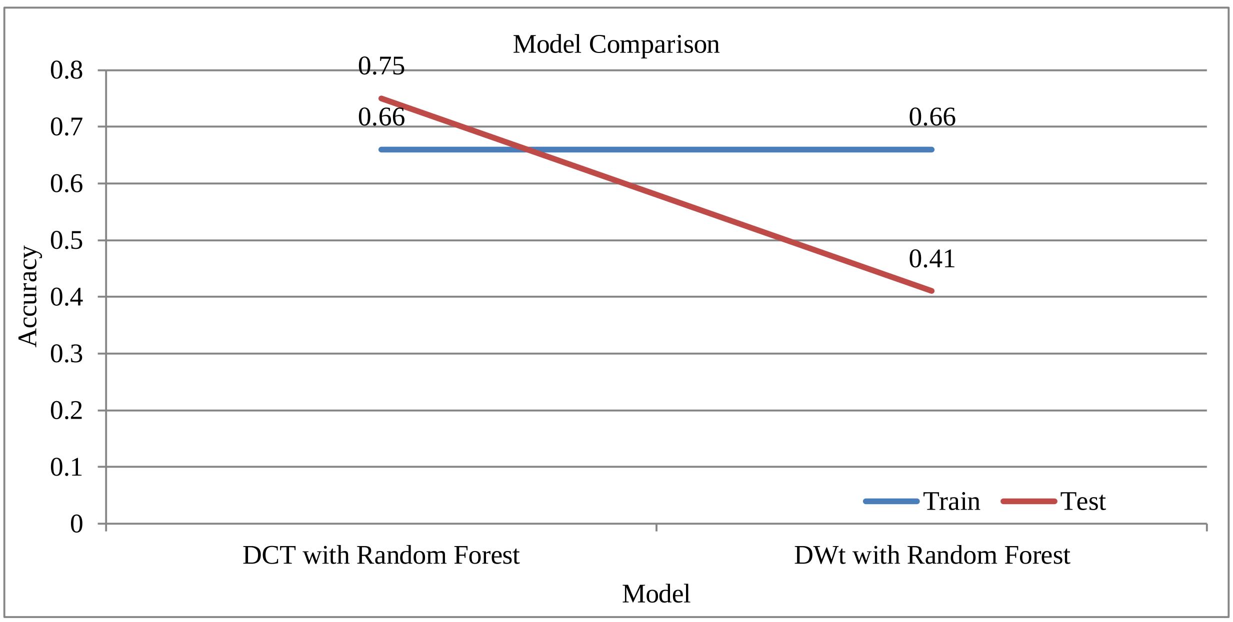

Fig. 6 DCT and DWT with different accuracy comparison

Table 2 and Figures 5 and 6. provide a brief comparative study of different wavelet domain techniques with different classifiers for the rice plant disease detection dataset. We have compared transform domain techniques like DCT and DWT in terms of $u_{y.}$

The experimental outcomes inferred that though both transforms had obtained 66.00% training accuracy, the DWT algorithm reached the lowest performance with $accu_y$ 41.00 %, whereas the DCT technique showed a somewhat high $accu_y$ 75.00% test accuracy.

3.4. CNN and Transfer Learning Model Analysis

A particular kind of deep neural network called CNNs is made especially for handling structured grid data, like picture data. They employ convolutional layers to detect patterns within an image, enabling hierarchical feature learning.

Transfer learning, on the other hand, involves leveraging pre-trained models on vast datasets to tackle new tasks efficiently by fine-tuning them on smaller, domain-specific datasets.

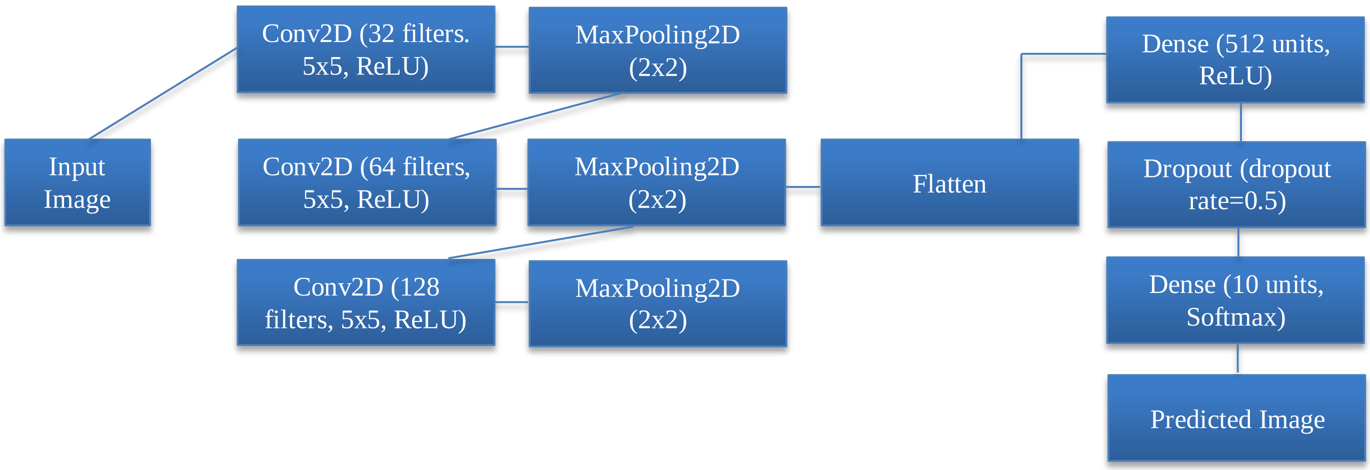

Fig. 7 Architecture for CNN

Mathematically, CNNs utilize convolution operations, activation functions like Rectified Linear Unit (ReLU), pooling layers, and fully connected layers to process image data. Let us denote:

- I as the input image,
- W as the weights of the convolutional filters,
- b as the bias terms,
- f as the activation function (e.g., ReLU),
- P as the pooling operation,
- F as the fully connected layer with weights Wf and biases bf,
- y as the predicted output.

From Equations (1) to (3), the forward pass of a CNN involves convolving the input image with the filters, applying the activation function, performing pooling to downsample features, and finally, passing the flattened output through fully connected layers to obtain predictions:

$$\text{Convolution}: z = f\big((I * W) + b\big) \tag{1}$$

$$Pooling: p = P(z) \tag{2}$$

$$Fully\ Connected: y = F(p) \tag{3}$$

Transfer learning adds another layer of mathematical sophistication by adapting pre-trained models. Let M be a pre-trained CNN model with parameters Θ, and D be a dataset of rice plant images with corresponding disease labels. By fine-tuning M on D, the new model learns to minimize a loss function L defined over the dataset D:

$$\min_{\theta} L(M(I;\theta), D) \tag{4}$$

Equation (4). This optimization process adjusts the parameters Θ of the pre-trained model better to suit the characteristics of rice plant disease images, resulting in a highly accurate and efficient disease detection system. Through this mathematical framework, CNNs and transfer learning models synergize to empower precision agriculture, enabling timely and accurate identification of diseases in rice plants, ultimately enhancing crop management and ensuring food security.

Table 3. Performance analysis CNN with transfer learning models

Model	Training $accu_y$ (%)	Test $accu_y$ (%)
CNN	0.56	0.50
VGG16 with CNN	0.75	0.50
ResNet50 with CNN	0.72	0.58
InceptionV3 with CNN	0.94	0.66
MobileNetV2 with CNN	0.80	0.66
Xception with CNN	0.77	0.75

Table 3 provides a brief comparative study of different transfer learning models with CNN for the rice plant disease detection dataset. We have compared various transfer learning models like VGG16 with CNN. ResNet50 with CNN, InceptionV3 with CNN, MobileNetV2 with CNN, Xception with CNN and CNN in terms of $accu_y$. The experimental outcomes inferred that the basic CNN has reached the lowest performance with $accu_y$ of 50.00 %, whereas the Xception with CNN model has shown a somewhat high $accu_y$ of 75.00 %.

3.5. Optimized Convolutional Neural Networks

The architecture of the Convolutional Neural Network (CNN) is designed specifically for image classification tasks, and it is furnished in Figure 8. It incorporates multiple convolutional layers with fluctuating filter sizes and depths, subsequently max-pooling layers for spatial downsampling and feature extraction. The architecture also includes fully connected layers for high-level feature representation and classification.

3.5.1. Convolutional Layers

Convolutional layers apply filters to input data to extract features. A set of learnable weights represents each filter. The output feature map is obtained by convolving the input with the filters. Following the application of an activation function and filter convalescence, the output feature map is produced. Mathematically, the output feature map YY of a convolutional layer can be expressed as:

$$Y[\,i\,,j\,] = f\left(\sum\nolimits_m \sum\nolimits_n X[i+m, j+n]\,.W[m,n\,] + b\right) \quad (5)$$

Where, XX is the input tensor, WW represents the filter weights, bb is the bias term, ff is the activation function, and ii and jj represent the spatial dimensions of the output feature map.

3.5.2. Max Pooling Layers

It uses the maximum value inside each pooling window to downsample the feature maps. In doing so, the feature maps' spatial dimensions are decreased, yet their most notable traits are maintained. Mathematically, the output of a max pooling operation can be expressed as:

$$Y[\,i\,,j\,] = max_{x,n}\, X\,[\,i\,*\,s+m, j\,*\,s+n\,] \quad (6)$$

where ss is the stride of the pooling operation and mm and nn iterate over the pooling window.

Fully Connected Layers

Dense layers that are fully connected link each neuron in the layer below to every other neuron in the layer above. Mathematically, the output YY of a fully connected layer can be calculated as:

$$Y = f\,(\sum\nolimits_i X\,[i].W\,[i] + b) \quad (7)$$

Equation (7), where XX represents the input vector, WW are the weights, bb is the bias term, and ff is the activation function.

3.5.3. Dropout Regularization

In order to avoid overfitting, dropout regularization randomly changes a portion of input units to zero during training. Mathematically, the dropout process can be expressed as:

$$Y\,[\,i\,] = \left\{ \begin{matrix} 0 & with\ propability\ p \\ \dfrac{x\,[\,i\,]}{1-p} & otherwise \end{matrix} \right\} \quad (8)$$

Equation (8), where X is the input vector, Y is the output vector after dropout, and pp is the dropout probability.

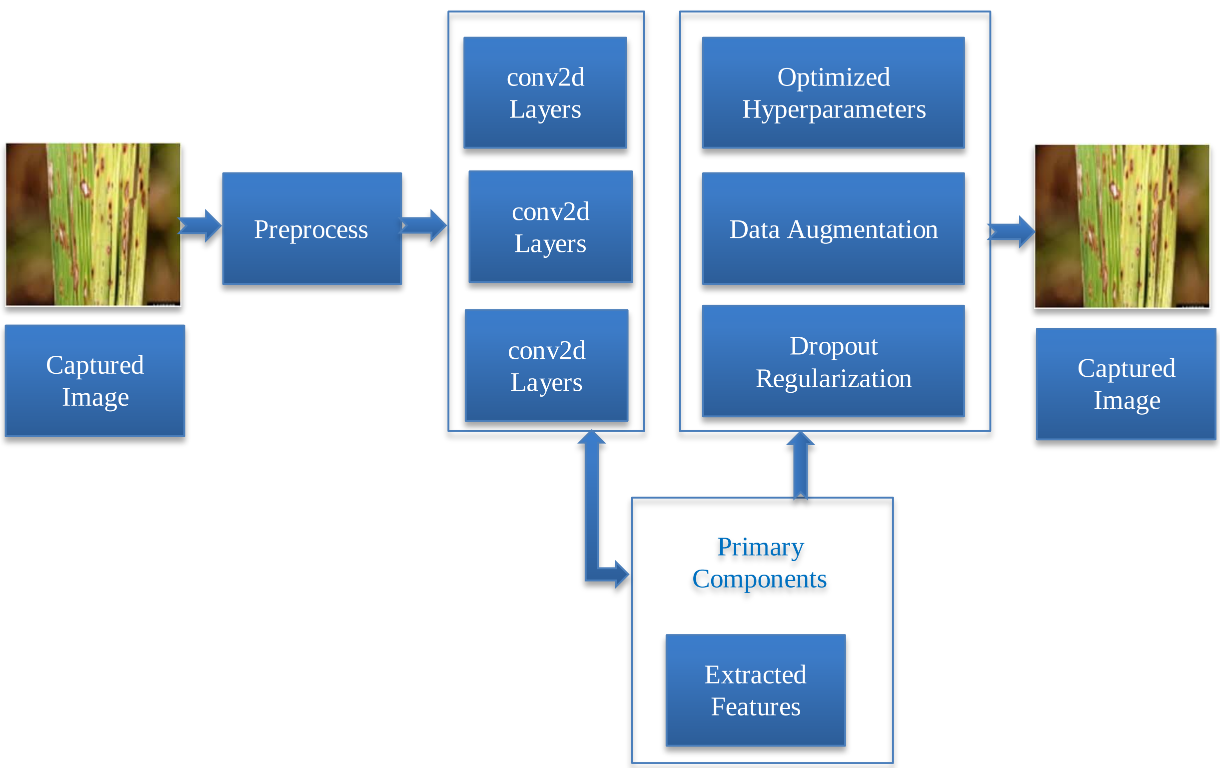

Fig. 8 Novel Convolutional Neural Network (CNN) architecture

The novelty lies in the customized architecture, optimized hyperparameters, utilization of data augmentation and dropout regularization techniques, and efficient data loading methods, all of which contribute to the improved performance and effectiveness of the CNN model for image classification tasks and are attached in Table 4. An implementation of CNN architecture can be summarized as follows:

Table 4. Novel CNN model summary

Layer (type)	Output Shape	Param #
conv2d_27 (Conv2D)	(None, 146, 146, 32)	2432
max_pooling2d_27 (MaxPooling 2D)	(None, 73, 73, 32)	0
conv2d_28 (Conv2D)	(None, 69, 69, 64)	51264
max_pooling2d_28 (MaxPooling 2D)	(None, 34, 34, 64)	0
conv2d_29 (Conv2D)	(None, 30, 30, 128)	204928
max_pooling2d_28 (MaxPooling 2D)	(None, 15, 15, 128)	0
flatten_9 (Flatten)	(None, 28800)	0
dense_18 (Dense)	(None, 512)	14746112
dropout_9 (Dropout)	(None, 512)	0
dense_19 (Dense)	(None, 3)	1539

Optimized Hyperparameters

The choice of hyperparameters, such as filter sizes, depths, and activation functions, is optimized for the task of image classification. These hyperparameters are selected based on empirical studies and experimentation to achieve optimal performance on the given dataset.

Data Augmentation

To diversify the training dataset, the algorithm incorporates data augmentation techniques. The model is strengthened and made less prone to overfitting by randomly transforming the input images through rotation, shifting, and flipping.

Dropout Regularization

To keep the model from overfitting, dropout regularization is used. Improved generalization performance results from the model learning more robust characteristics and being less dependent on individual neurons through the random drop of a portion of the neurons during training.

Efficient Data Loading

The code efficiently loads and preprocesses image data using the flow_from_directory method, which allows for easy integration with large-scale datasets organized into directories by class labels. This approach streamlines the data-loading process and improves the overall efficiency of model training.

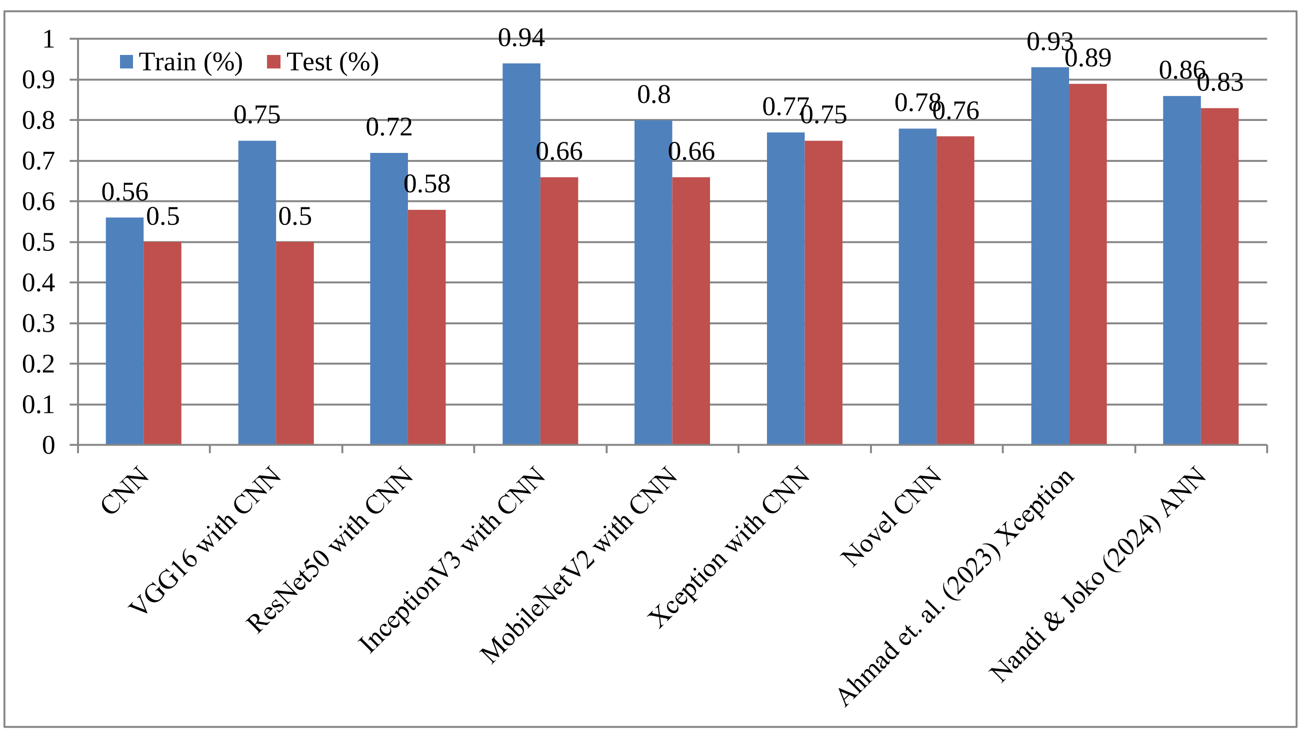

Fig. 9 Comparative analysis with existing models

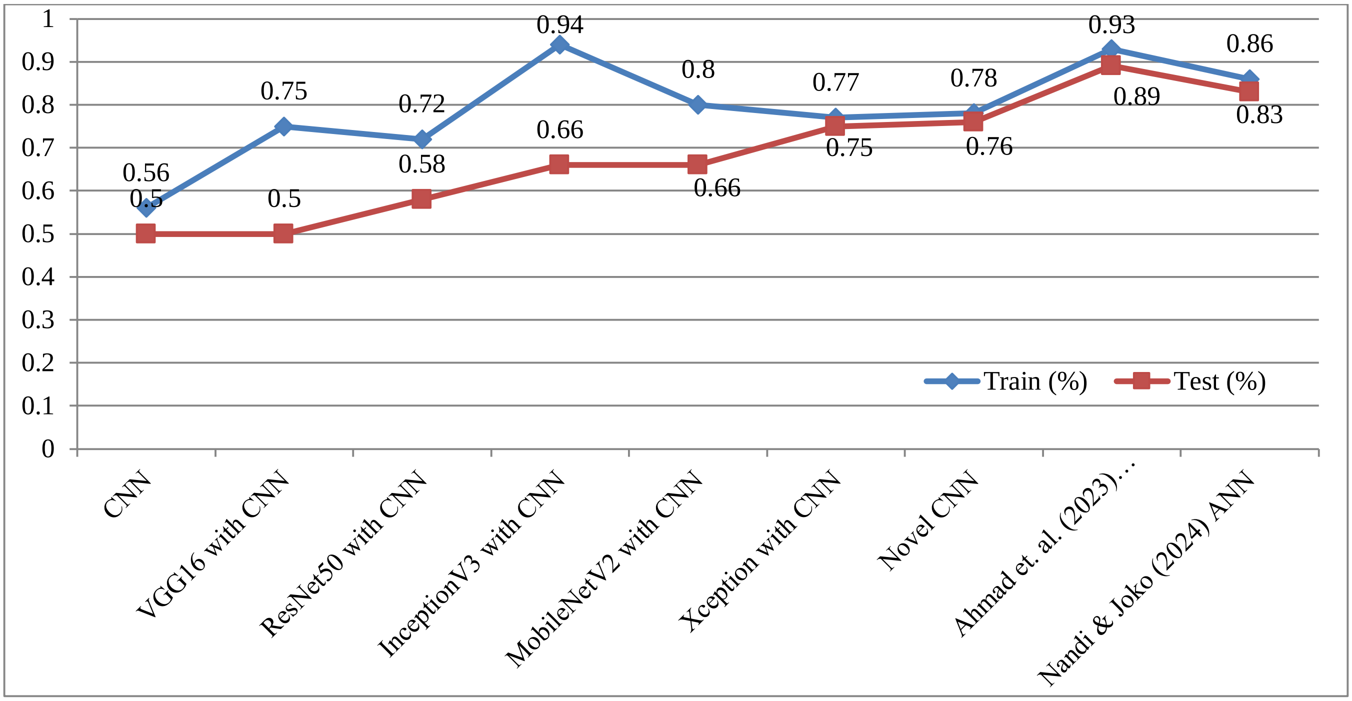

Fig. 10 Comparative analysis for model performance

Table 4 shows that the proposed CNN layers are organized around a series of specialized layers intended for the effective processing of visual input. Convolutional layers are among the essential parts. They apply filters to the input image to extract features like textures and edges. After that, these features are spatially down-sampled using pooling layers, which lowers computational complexity without sacrificing crucial information. Non-linearities are introduced by activation functions such as ReLU, which are essential for capturing intricate correlations in the data. The network's fully connected layers combine the features that have been extracted for tasks like regression or classification. Because CNNs can automatically learn hierarchical representations of visual information, they are highly effective in tasks like object detection, segmentation, and image recognition, which makes them an essential tool in contemporary computer vision applications.

Table 5. Enhanced CNN with other models comparative study

Model	Train (%)	Test (%)
CNN	0.56	0.50
VGG16 with CNN	0.75	0.50
ResNet50 with CNN	0.72	0.58
InceptionV3 with CNN	0.94	0.66
MobileNetV2 with CNN	0.8	0.66
Xception with CNN	0.77	0.75
Novel CNN	**0.78**	**0.76**
Ahmad et al. (2023) Xception	0.93	0.89
Nandi & Joko (2024) ANN	0.86	0.83

4. Results and Discussion

Table 5 provides a brief comparative study of different transfer learning models with CNN for the rice plant disease detection dataset. We have compared various transfer learning models like VGG16 with CNN. ResNet50 with CNN, InceptionV3 with CNN, MobileNetV2 with CNN, Xception with CNN and CNN in terms of $accu_y$. The experimental outcomes inferred that the basic Xception with CNN has reached the lowest performance with $accu_y$ of 75.00 %, whereas the Enhanced CNN model has shown a somewhat high $accu_y$ of 76.00 %. When we compared all the other models our proposed CNN model has given the greater accuracy of this problem. Though it has obtained a lower performance than the methods presented by Ahmad et al. (2023) using Xception and Nandi & Joko (2024) using ANN, the study presented an in-depth exploration of rice plant disease detection using Convolutional Neural Networks (CNNs) and optimization techniques. The same is graphically explained in Figures 9 and 10. However, this method extensively demonstrated the performance of the proposed method by various feature extraction methods and they were compared. It has been recorded that the novel CNNs has obtained a superior performance. Our novel CNN architecture, optimized with hyperparameters, data augmentation, and dropout regularization, achieved promising results in accurately identifying rice plant diseases.

5. Conclusion

Finally, our study offers a thorough examination of methods for detecting rice plant diseases, with an emphasis on Convolutional Neural Networks (CNNs) and optimization techniques for image classification. Conducted a comparative analysis of several feature extraction techniques and classifiers, emphasizing the improved performance that can be obtained using deep learning-based methods, especially CNNs. With the use of dropout regularization, data augmentation, and hyperparameter optimization, our suggested CNN architecture produced a superior result of 76.00%, outperforming the results of the other feature extraction techniques, and showed encouraging results in the accurate identification of rice plant illnesses. Though it has obtained a significant performance, it has given less performance than the existing methods presented by Ahmad et al. (2023) using Xception and Nandi & Joko (2024) using ANN.

Several areas for improvement are noted for future work. First, by exploring attention processes inside the CNN architecture and fine-tuning its hyperparameters more, the model may be better able to extract and focus on pertinent features, which would increase classification accuracy. Furthermore, utilizing transfer learning strategies and adding new data modalities, such as spectrum imaging, may offer insightful information for reliable illness identification, especially in a variety of environmental settings.

References

[1] Navneet Dalal, and Bill Triggs, "Histograms of Oriented Gradients for Human Detection," *2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05)*, San Diego, CA, USA, pp. 886-893, 2005. [CrossRef] [Google Scholar] [Publisher Link]

[2] David G. Lowe, "Distinctive Image Features from Scale-Invariant Keypoints," *International Journal of Computer Vision*, vol. 60, pp. 91-110, 2004. [CrossRef] [Google Scholar] [Publisher Link]

[3] Timo Ojala, Matti Pietikainen, and Topi Maenpaa, "Multiresolution Gray-Scale and Rotation Invariant Texture Classification with Local Binary Patterns," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 24, no. 7, pp. 971-987, 2002. [CrossRef] [Google Scholar] [Publisher Link]

[4] Karen Simonyan, and Andrew Zisserman, "Very Deep Convolutional Networks for Large-Scale Image Recognition," *arXiv*, pp. 1-14, 2014. [CrossRef] [Google Scholar] [Publisher Link]

[5] Kaiming He et al., "Deep Residual Learning for Image Recognition," *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Las Vegas, NV, USA, pp. 770-778, 2016. [CrossRef] [Google Scholar] [Publisher Link]

[6] Christian Szegedy et al., "Rethinking the Inception Architecture for Computer Vision," *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Las Vegas, NV, USA, pp. 2818-2826, 2016. [CrossRef] [Google Scholar] [Publisher Link]

[7] Mingxing Tan, and Quoc Le, "Efficientnet: Rethinking Model Scaling for Convolutional Neural Networks," *Proceedings of the 36th International Conference on Machine Learning*, vol. 97, pp. 6105-6114, 2019. [Google Scholar] [Publisher Link]

[8] Alexey Dosovitskiy et al., "An Image is Worth 16x16 Words: Transformers for Image Recognition at Scale," *arXiv*, pp. 1-22, 2020. [CrossRef] [Google Scholar] [Publisher Link]

[9] Robert J. Gillies, Paul E. Kinahan, and Hedvig Hricak, "Radiomics: Images are More Than Pictures, They are Data," *Radiology*, vol. 278, no. 2, pp. 563-577, 2016. [CrossRef] [Google Scholar] [Publisher Link]

[10] Ashish Vaswani, "Attention is All You Need," *Advances in Neural Information Processing Systems*, vol. 30, pp. 1-11, 2017. [Google Scholar] [Publisher Link]

[11] Mark Sandler et al., "MobileNetV2: Inverted Residuals and Linear Bottlenecks," *2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 4510-4520, 2018. [CrossRef] [Google Scholar] [Publisher Link]

[12] Gao Huang et al., "Densely Connected Convolutional Networks," *2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Honolulu, HI, USA, pp. 4700-4708, 2017. [CrossRef] [Google Scholar] [Publisher Link]

[13] Joseph Redmon et al., "You Only Look Once: Unified, Real-Time Object Detection," *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Las Vegas, NV, USA, pp. 779-788, 2016. [CrossRef] [Google Scholar] [Publisher Link]

[14] Olga Russakovsky et al., "ImageNet Large Scale Visual Recognition Challenge," *International Journal of Computer Vision*, vol. 115, pp. 211-252, 2015. [CrossRef] [Google Scholar] [Publisher Link]

[15] Tsung-Yi Lin et al., "Focal Loss for Dense Object Detection," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 42, no. 2, pp. 318-327, 2017. [CrossRef] [Google Scholar] [Publisher Link]

[16] Muhammad Shoaib et al. "An Advanced Deep Learning Models-Based Plant Disease Detection: A Review of Recent Research," *Frontiers in Plant Science*, vol. 14, 2023. [CrossRef] [Google Scholar] [Publisher Link]

[17] Lukas Wiku Kuswidiyanto, Hyun-Ho Noh, and Xiongzhe Han, "Plant Disease Diagnosis Using Deep Learning Based on Aerial Hyperspectral Images: A Review," *Remote Sensing*, vol. 14, no. 23, pp. 1-24, 2022. [CrossRef] [Google Scholar] [Publisher Link]

[18] Muhammad Asif Saleem et al., "An Optimized Convolution Neural Network Architecture for Paddy Disease Classification," *Computers, Materials & Continua*, vol. 71, no. 3, pp. 6053-6067, 2022. [CrossRef] [Google Scholar] [Publisher Link]

[19] Md. Tarikul Islam, and Md. Shamsuzzaman, "Transfer Learning Based Approach to Crops Leaf Disease Detection: A Diversion Changer in Agriculture," *International Journal of Engineering and Technical Research*, vol. 11, no. 9, pp. 87-94, 2022. [Google Scholar] [Publisher Link]

[20] Liu Huajian, and Xu Zhanyou, "Machine Vision and Machine Learning for Plant Phenotyping and Precision Agriculture," *Frontiers in Plant Science*, vol. 14, 2023. [CrossRef] [Google Scholar] [Publisher Link]

[21] Ahmad Rofiqul Muslikh, De Rosal Ignatius Moses Setiadi, and Arnold Adimabua Ojugo, "Rice Disease Recognition Using Transfer Learning Xception Convolutional Neural Network," *Journal of Information Engineering*, vol. 4, no. 6, pp. 1535-1540, 2023. [CrossRef] [Google Scholar] [Publisher Link]

[22] Nandi Sunandar, and Joko Sutopo, "Utilization of Artificial Neural Network in Rice Plant Disease Classification Using Leaf Image," *International Journal of Research In Science & Engineering (IJRISE)*, vol. 4, no. 2, pp. 1-10, 2024. [CrossRef] [Google Scholar] [Publisher Link]

[23] Gao Huang et al., "Densely Connected Convolutional Networks," *2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Honolulu, HI, USA, pp. 4700-4708, 2017. [CrossRef] [Google Scholar] [Publisher Link]