



Article Title: Artificial Intelligence based Classification of Diseases for Rice Leaf Using CNN model

Artificial Intelligence based Classification of Diseases for Rice Leaf Using CNN model

Dr. Misba M/ AP¹, Vivek A², Ratheesh R³, Aslin C⁴

Bethlahem Institute of Engineering, Ulaganvillai, Kanyakumari, Tamil Nadu

ABSTRACT

Rice cultivation is a crucial industry in India, but it is plagued by various diseases that can damage crops at different stages. These diseases are challenging for farmers to identify accurately due to their limited knowledge and expertise. As a result, the farmers often struggle to take appropriate measures to prevent or manage these diseases, which can result in significant losses in crop yield and quality. Therefore, there is a need for advanced technologies and tools to help farmers accurately identify and manage these diseases, ensuring a sustainable and profitable rice cultivation industry in India. Recent advances in Deep Learning have demonstrated that Convolutional Neural Network models can be highly effective in automatic image recognition tasks. These models have shown great potential in addressing the challenges faced by farmers in identifying diseases in crops such as rice. However, in order to train such models, a large and diverse dataset is required, which may not always be readily available. To address this issue, researchers have created their own dataset of rice leaf disease images, which may be smaller in size but sufficient for the task at hand. To develop their CNN model, they have used a technique called Transfer Learning, which used as a starting point to fine-tune already trained models for a new task. The proposed CNN architecture is based on the VGG-16, a widely used pre-trained model used in computer vision tasks. The researchers trained and tested their model with a dataset obtained from rice fields and the Internet. The results show that the proposed model achieves 92.46% accuracy, demonstrating its potential in accurately detecting rice leaf diseases.

Keywords: Convolutional Neural Network, Deep Learning, Fine-tuning, Rice leaf diseases, Transfer learning.

1 Introduction

Rice is a staple food not only in India but also worldwide. However, rice plants are susceptible to various diseases at different stages of cultivation, making it difficult for farmers to detect and effectively control them. Early detection and control of these diseases is critical to ensuring a high-quality, high-yielding crop. However, the vast acreage of each grower combined with the diversity and complexity of these diseases make this a difficult task. Furthermore, rice plants can also be affected by more than one disease simultaneously, which further complicates the process of disease identification and management. As a result, farmers face significant challenges in maintaining the health and productivity of their rice crops, which can have a



Article Title: Artificial Intelligence based Classification of Diseases for Rice Leaf Using CNN model

significant impact on their livelihoods and the food supply for the population. Therefore, the development of effective technologies and tools for disease detection and management is critical for the sustainable cultivation of rice and food security in India and beyond. Agricultural expertise and support are not readily available in remote areas, and even when they are available, it can be a time-consuming process.

To overcome these challenges, automated systems are needed to help farmers accurately and quickly detect and treat plant diseases. To address this need, researchers have explored the use of various machine learning algorithms such as Support Vector Machine (SVM) and Artificial Neural Networks (ANN) to develop automated systems for plant disease detection. However, the accuracy of such systems is highly dependent on the selection of appropriate features. Recent research on Convolutional Neural Networks (CNNs) has shown great promise in image-based detection tasks, including plant disease detection. CNNs eliminate the need for image preprocessing and feature selection by providing inbuilt feature selection capabilities, improving the accuracy of disease detection systems.

However, one major challenge in developing such systems is the difficulty in obtaining a large dataset for training the models. Nonetheless, the application of CNNs for plant disease detection represents a significant advancement in the development of automated systems for the benefit of farmers and sustainable agriculture. For relatively small data sets, it is more advantageous to use a pre-trained model that has been trained on a larger data set. This process is known as transfer learning, where an existing model is used as a fixed feature extractor by removing the last fully linked layer or fine-tuning the last few layers to make the model more specific to the dataset. To make disease detection more accessible to farmers, we have developed an automated system that uses cell phones. Farmers can upload images of diseased leaves to our server, where a neural network is used to detect the disease. The system will then send back disease classification and remedy information to the farmer, allowing them to take action promptly.

This approach has significant advantages, as it enables farmers to obtain expert assistance for disease management quickly and easily, without requiring extensive knowledge or resources. Furthermore, this system has the potential to significantly improve the efficiency and productivity of agricultural systems, leading to better crop yields and food security. In this study, we focused on developing the disease classification component of an automatic rice disease identification system. Following previous work on convolutional neural networks, including the studies cited in references [1]- [2] and [3], we developed a deep learning approach using our own rice disease dataset that we have collected over the past few months. To improve the accuracy of our model, we used transfer learning and fine-tuned the fully linked layers of a pre-trained VGG16 model trained on the extensive ImageNet dataset. By doing so, we were able to adjust the model to better accommodate our own dataset and achieve high classification accuracy. After training and testing the model, we conducted error analysis to understand the



Article Title: Artificial Intelligence based Classification of Diseases for Rice Leaf Using CNN model

reasons for misclassifications. Overall, our study demonstrates the potential of deep learning approaches for automated plant disease detection and highlights the importance of using large data sets and transfer learning to improve model accuracy.

1.1 Type of Rice Diseases and Description of Dataset

The dataset of rice leaf images used in this work was collected over a period of several months from various crop fields in the Indian state of West Bengal and from the internet. The images were captured using mobile cameras and the symptoms and disease knowledge were collected from IRRI Rice Knowledge Bank website. The dataset consists of 1649 images of diseased rice leaves covering three common diseases, namely Rice Leaf Blast, Rice Leaf Blight and Brown Spot, and 507 images of healthy leaves. Due to the limited number of images available for training the CNN, data augmentation techniques were used to increase the size of the dataset. Despite difficulties such as poor lighting and multiple diseases in the same plant, the images were preprocessed using resizing and zooming techniques to improve data quality. The following sections describe the different types of rice leaf diseases on which the study was conducted.



Figure 1: (a)-(c) From Left to Right. (a) Leaf Blight (b) Late Blight and (c) Brown Spot Disease

1.1.1 Leaf Blast

Rice leaf blast is a fungal disease caused by *Magnaporthe oryzae*. The disease is characterized by the appearance of elliptical or spindle-shaped spots on leaves that are white to gray-green with dark red to brownish margins. Some spots are diamond-shaped with broad centers and pointed ends. Figure 1 (a) shows an example of these spindle-shaped lesions with white spots and dark brown margins.

1.1.2 Leaf spot disease

Rice leaf blight is a bacterial disease caused by *Xanthomonas oryzae*. The disease causes infected leaves to turn gray-green and curl, followed by yellowing and eventual wilting and



Article Title: Artificial Intelligence based Classification of Diseases for Rice Leaf Using CNN model

death. Lesions have wavy edges and spread toward the base of the plant. Young lesions may have a bacterial slime resembling morning dew.

1.1.3 Brown Spot

Brown Spot is a fungal disease that causes big spots on rice leaves and can eventually kill the entire leaf. The disease starts with small, circular, dark brown to purple-brown lesions on the leaves, which later develop into circular to oval lesions with a light brown to gray center, surrounded by a reddish brown margin caused by the toxin produced by the fungi. In Figure 1(c), we can see the small dark brown lesions on the leaves affected by Brown Spot.

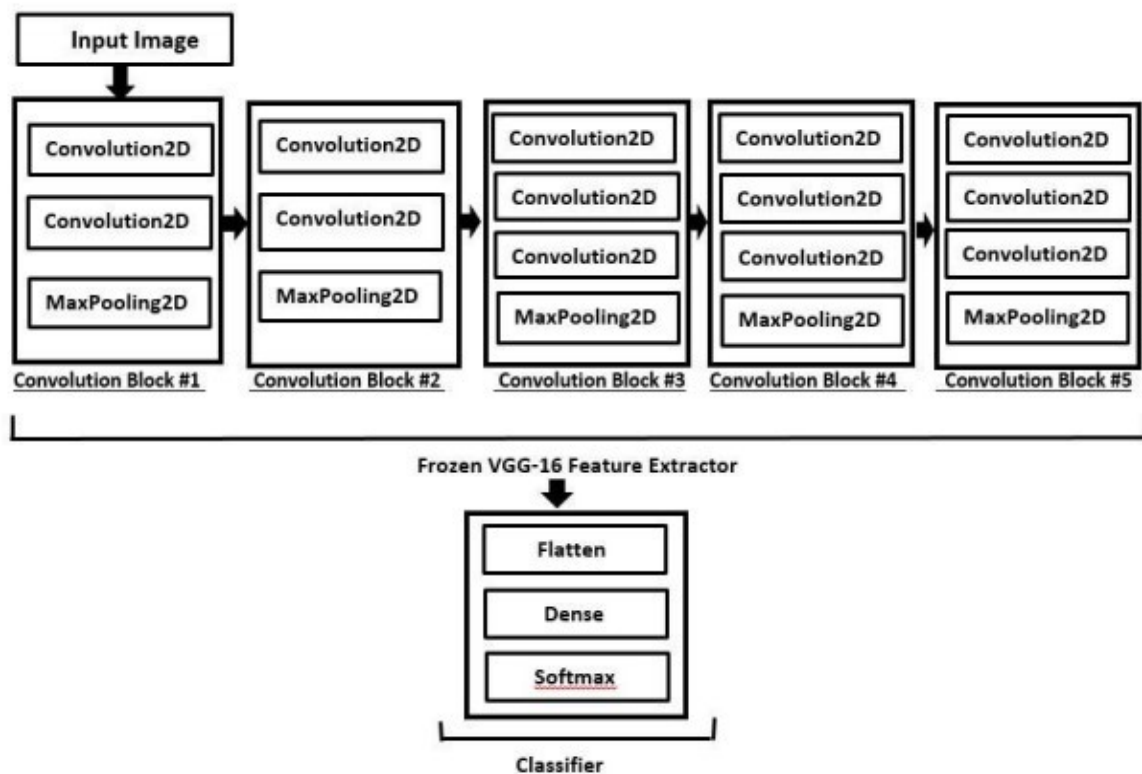


Figure 2: VGG-16 Architecture find-tuned with the last two layer with 128 Dense FC Layer and 4 Dense Softmax Layer as the output

2 Methodology

Convolutional neural networks (CNNs) are deep learning models used for image classification tasks. They consist of convolutional layers, pooling layers, and fully connected layers that together form the feature extractor and classifier. The convolution layer extracts features from the input image using a set of adaptive filters, and the pooling layer reduces the dimensionality of the extracted features. The fully linked layer connects to each of the generated feature maps,



Article Title: Artificial Intelligence based Classification of Diseases for Rice Leaf Using CNN model

and the softmax layer assigns probabilities to each class in a multi-class problem. The VGG16 architecture is a deep CNN with weights pre-trained on the ImageNet database, which contains millions of images. Using transfer learning, the pre-trained weights, architecture, and features learned in one domain can be transferred to another domain. In our model, we froze the first 5 blocks of convolutional layers to use as a feature extractor, and trained the last dense layer of size 128 followed by a softmax layer of size 4 for classification. There are four approaches for transfer learning, depending on the size and similarity of the new dataset with the original dataset.

- The new record is small and familiar with the original record.
- The new record is large and familiar with the original record.
- The new record is small but dissimilar to the original record.
- The new dataset is large and dissimilar to the original dataset.

Due to the limited size and dissimilarity of our dataset with the ImageNet dataset, we have adopted the third approach of transfer learning. Here, the convolutional layers of VGGNet are retained to extract features from our dataset, and the final fully linked layers are fine-tuned to match the number of classes in our dataset. This approach allows us to leverage the pretrained weights and architecture of the VGGNet for feature extraction, while still adapting the model to our specific classification task.

The proposed model architecture is shown in Figure 2. In addition, a CNN model without transfer learning was also developed. This model had 4 convolutional layers, each followed by ReLU activation, maxpooling, and dropout layers, followed by 2 fully linked layers and a softmax layer for classification. However, the performance of this model proved to be worse than that of the model with transfer learning, as shown in the Results section.

2.1 Implementation

2.1.1 Experimental Setup

The experiment was conducted using a Windows 10 PC with a P4000 GPU card and a 64-bit operating system. The CNN-based model was implemented using Keras 2.2.4 deep learning framework with TensorFlow 1.13.1 backend and Python 3.7.2 programming language.

2.1.2 Image Acquisition

The dataset used in the study consists of images of rice plant leaves with four classes: Leaf Blast, Leaf Blight, Brown Spot, and Healthy Plants. The images were collected from real cultivated fields as well as from the internet to increase the diversity of the dataset. The use of real images of crop fields helps to ensure the practicality of the proposed model in real scenarios.



Article Title: Artificial Intelligence based Classification of Diseases for Rice Leaf Using CNN model

2.1.3 Image Preprocessing and Augmentation

To improve the performance of the model and prevent overfitting, data augmentation techniques were used to increase the size and diversity of the dataset. The collected images were first resized to a standard size of 224 x 224 pixels and then augmented using the Image Data Generator module in Keras. This module provides a range of image transformations such as zooming, rotation, horizontal and vertical shifting, flipping, and more, which were applied to the images to generate new variations of the original images. The augmented dataset was then used for training and testing the model.

2.1.4 CNN Model Training

In summary, the image dataset is preprocessed by reducing the images to 224*224 pixels and applying various magnification techniques. The dataset is split into training, test and validation sets using the train_test_split function. The class labels are encoded as integers and then converted to one-hot encoding. The VGG-16 model is loaded from keras and the last fully connected layers are removed. The remaining layers are made untrainable, and the output of the feature extractor is flattened, followed by a fully connected layer and an output layer with softmax. The model is built using the Adam optimizer with categorical_crossentropy as the loss function for classification. The training is stopped at 25 epochs and the steps are shown in Figure 3.

2.1.5 Justification for the Chosen Model

Transfer learning is a technique in which knowledge gained in solving one problem is used to solve a different but related problem. It is often used in Deep Learning to improve model accuracy and reduce training time. Transfer learning uses a pre-trained model as a starting point and then fine-tunes it using a smaller, more specific data set. This approach is particularly useful when there is a lack of labeled data to train a model from scratch. By using a pre-trained model, the neural network can use the knowledge it has already acquired and then adapt to new data sets to solve different tasks.

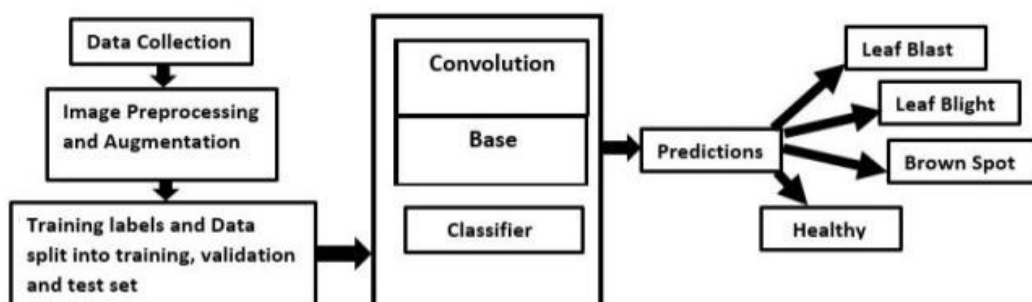


Figure 3: Overview of the steps of the proposed model



Article Title: Artificial Intelligence based Classification of Diseases for Rice Leaf Using CNN model

3 Results

3.1 Calculations

The proposed CNN model with transfer learning achieved 97% accuracy on the training set and 92.4% on the test set after being trained for 25 epochs on the dataset consisting of 1509 training images and 647 test images. In comparison, a CNN model without transfer learning was also trained on the same dataset, but only achieved a maximum accuracy of 74%, even after fine-tuning the stack size, number of epochs, optimizer, and dropout rate. The CNN model without transfer learning consisted of 4 convolutional layers, each followed by ReLU, maxpooling, and dropout layers, followed by 2 fully linked layers and SoftMax. Table 1 summarizes the comparison of accuracy between the proposed CNN model with transfer learning and the CNN model without transfer learning. In addition, Figure 4 shows the training and validation accuracy as a function of the number of epochs for the CNN model with transfer learning.

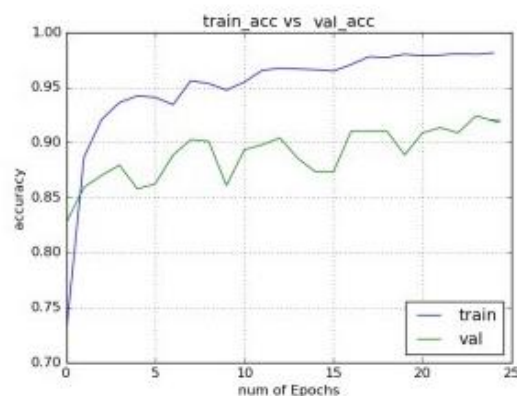


Figure 4: plot of accuracy versus epochs

Table 1: Performance of Comparison of CNN with and Without Transfer Learning

Model	Test Accuracy
CNN With Transfer Learning	92.46%
CNN Without Transfer Learning	74%



Article Title: Artificial Intelligence based Classification of Diseases for Rice Leaf Using CNN model



Figure 5: From Left to Right (a)-(f) Rice Disease images that are misclassified by the model
(a) Rice Blast Disease (b) and (c) Brown Spot (d) and (e) Leaf Blight (f) Healthy

3.1.1 Rice Blast

In the results of our proposed model, we found that image (a), which belongs to the class of rice tracks, was incorrectly classified as a brown stain. The possible reason for this misclassification could be the blurring of the image. In addition, there could be small brown spots on the same rice leaf that led to confusion in the classification.

3.1.2 Leaf spot disease

Images (d) and (e), which are classified as healthy when they actually belong to the blight category, may be due to poor lighting or blurring of the images, which may have made it difficult for the model to accurately distinguish disease symptoms. The model may not have been trained on images with such poor quality, hence, it may not have learned to recognize the disease symptoms under such conditions. It's important to ensure that the dataset used for training is diverse and contains images with a range of lighting conditions and blurriness to improve the model's ability to generalize to different scenarios.

3.1.3 Healthy

Blurring and poor contrast in the image can affect the ability of the model to accurately identify the features and patterns in the image, which can lead to misclassification.

3.1.4 Brown Spot

Images (b) and (c) were originally classified as brown spots, but were incorrectly classified as blasts by the model. One reason for this could be the presence of small blast lesions on the leaf that could be missed during image preprocessing and magnification. Also in image (d), the brown spot lesions may resemble blast lesions, leading to misclassification.



Article Title: Artificial Intelligence based Classification of Diseases for Rice Leaf Using CNN model

4 Conclusion

In this research, a deep-learning model was proposed and evaluated using a dataset of 1509 rice leaf images with four classes - leaf blight, leaf rot, brown spot and healthy. The model achieved 92.46% accuracy on the test data, which is a significant improvement over the traditional CNN model without transfer learning. The use of transfer learning and fine-tuning of VGGNet helped to improve the performance of the model, allowing it to achieve good results with a relatively small data set. The number of epochs used for training was limited to 25 to avoid overfitting. Future work in this area includes expanding the dataset by collecting more images from different sources, cross-validation, and using advanced deep learning models. The proposed model can also be extended to the detection of other leaf diseases, which are an important crop in India. The potential for using deep learning in agriculture is significant, and this study highlights the potential of such techniques to improve crop yields and food security.

References

1. Rishabh Sharma; Vinay Kukreja, Year: 2021, "Rice diseases detection using Convolutional Neural Networks: A Survey", In 2021 International Conference on Advance Computing and Innovative Technologies in Engineering (ICACITE), pp. 995 – 1001.
2. Paidi Haritha, Dr. R. Maruthamuthu, "Rice Leaf Diseases Classification Using Cnn with Transfer Learning".
3. Shreya Ghosal; Kamal Sarkar, year: 2020, "Rice Leaf Diseases Classification Using CNN with Transfer Learning", In 2020 IEEE Calcutta Conference (CALCON), pp. 230 – 236.