Paddy leaf symptom-based disease classification using deep CNN with ResNet-50

Pushpa Athisaya Sakila Rani [a,1,*], N.Suresh Singh [b]

[a] Research Scholar, Department of Computer Science, Malankara Catholic College, India
[b] Associate Professor and Head, Department of Computer Applications, Malankara Catholic College, India
[1] pushpasakila@gmail.com
* corresponding author

ARTICLE INFO

Article history
Received May 13, 2022
Revised June 29, 2022
Accepted July 12, 2022

Keywords
Rice disease identification
Classification
Logistic regression
Deep CNN
ResNet-50

ABSTRACT

Agriculture is a globally important occupation. Food is a fundamental need for all living things on the earth, hence it plays a significant role. As a result, agricultural product quality has to be improved. Paddy is susceptible to a variety of illnesses, just like all other crops. Diseases vary by region and season. Despite the fact that the number of new technologies being implemented in agriculture is rapidly expanding, farmers in our nation continue to rely on old methods for disease detection. Machine learning relies heavily on features to classify images. The advancement of the deep convolutional neural network paves the path for disease detection in rice based on deep characteristics, with the expectation of excellent yields. Field photos of four forms of rice leaf diseases, including bacterial leaf blight, brown spot, leaf smut, and tungro, were introduced using this proposed method. The model is trained using the deep CNN classification technique. The pre-trained ResNet-50 model is also included to increase the model's prediction accuracy. Existing approaches are outperformed by the combination of Deep CNN and ResNet-50. Accuracy was used as a criterion for evaluating performance.

1. Introduction

Agriculture is one of the world's most important professions. Crop diseases, in particular, are thought to be responsible for 42 percent of the world's most important food crops losing output. As a result, it's critical to recognise disorders early on. Crop diseases can sometimes wipe out the entire agricultural industry. As a result, farmers must learn everything they can about crop diseases as soon as possible so that they can effectively manage them. A plant's life cycle is divided into several stages. This involves soil preparation, sowing, adding manure and fertiliser, irrigation operations, disease detection, pesticide use, and crop harvesting. Food is necessary for our bodies to function properly, yet excessive fertilizer use has harmed not only plants, but also humans.

Paddy diseases are key problems because they can result in considerable detections in both the quality and quantity of crops produced in agriculture. As a result, disease identification and classification are a critical duty. Traditionally, specialists naked eye observation has been the primary method for detecting and identifying rice infections in practice. However, this necessitates expert monitoring on a continual basis, which could be prohibitively expensive in large forms. Furthermore, in some developing nations, farmers may have to travel considerable distances to reach experts, making consulting exerts prohibitively expensive and time consuming, and farmers may be uninformed of non-native diseases.

Automatic detection of plant diseases is a hot study area because it might help monitor large fields of crops and, as a result, automatically detect diseases based on symptoms on the plant leaves. This enables image-based automatic inspection, process control, and robot guiding using machine vision. Plant disease identification by sight is a more time-consuming and inaccurate task that can only be performed in limited locations. Automatic detection, on the other hand, requires fewer efforts, takes less time, and is more accurate. Our goal is to distinguish the faults identified in a paddy using Deep CNN with ResNet-50. It minimizes the loss of detection of diseased paddy crops at the early stage. The existing system supports small data sets and is limited to identifying three diseases. But we came up with Deep CNN with ResNet-50 to identify bacterial leaf blight, brown spot, leaf smut and tungro diseases.

2. Related Works

Nilam *et al.,* [1] Paddy leaf illnesses were identified utilising evolutionary and machine learning methods including AdaBoost and Bagging Classifier, as well as evolutionary algorithms like Genetic algorithm. In this proposed system they can obtained the accuracy by GA classifier is 96% at training feature-set and 91% at testing feature-set. The accuracy attained by AdaBoost is 88% at training feature-set and 84% at testing feature-set whereas the accuracy attained by Bagging algorithm is 86% at training feature-set and 81% at testing feature-set. Vimal K.Shrivastava et al. [2] used deep CNN transfer learning to classify rice leaf diseases. For the first time, deep CNN transfer learning was used to classify rice plant diseases in this work. However, for the 80 percent - 20% training-testing partition, this model was able to classify rice illnesses with a classification accuracy of 91.37 percent. Ruoling et al. [3] used deep learning to construct an automatic diagnosis system for rice illnesses. In this study, ensemble learning was achieved by combining the three best network sub models such as DenseNet-121, ResNeSt-50, and SE-ResNet-50. But the limitation was that the Ensemble Model has many parameters, which may affect the speed of identification of rice diseases. Shivam *et al*., [4] developed a model to recognize rice plant infection using deep neural network systems. In this work they does not concentrate to get better accuracy by applying some image pre-processing techniques before training the model. G.Jayanthi *et al.,* [5] presented the detailed study of different image processing techniques to detect the disease in rice plant. In segmentation, primary colours are RGB images that are used to identify the disease. This paper will aid scholars in their understanding of rice disease detection through computer vision. They were utilised for image enhancement in digital image processing techniques. For feature extraction, GLCM and SURF features were used. For segmentation, edge detection and FCM are utilised. For categorization, ANN is utilised. D.Swathi and A.Bharathi [6] described about disease classification of paddy leaves using HSI feature extraction and SVM technique. This system does not demonstrate that the feature extraction technique should include the extraction of more features, nor does it demonstrate that stem cell samples are also engaged in determining disease accuracy with colour and chlorophyll content at each stage of growth.. Kawcher Ahmed *et al.,* [7] machine learning was offered as a method for detecting rice leaf disease. This study discovered three of the most frequent rice plant diseases: leaf smut, bacterial leaf blight, and brown spot diseases. The input was clear photos of harmed rice leaves on a white backdrop. The dataset was trained using a variety of machine learning methods, including KNN (K-Nearest Neighbor), J48 (Decision Tree), Naive Bayes, and Logistic Regression, after necessary pre-processing. However, when applied to the test dataset, the Decision tree algorithm attained an accuracy of over 97 percent after 10-fold cross validation. Wan-jie Liang *et al.*,[8] proposed a novel rice blast recognition method based on CNN. For training and testing the CNN model, a dataset of 2906 positive samples and 2902 negative samples was created. This research does not focus on other deep neural network architectures and instead makes full use of deep learning methods to increase classification accuracy and the reliability and resilience of rice disease diagnosis systems.

3. Metodology

A system based on Deep CNN with ResNet-50 was developed for the identification of four types of rice leaf diseases. In the pre-processing stage, grey scale conversion method is used to convert the input image into grey image. Median filter algorithm is used to remove the noise from the images and a deep feature is extracted with the help of Logistic regression algorithm. Again, the feature learning approach was applied for the identification of rice diseases in the Deep CNN with resnet 50 model. Figure1 represent the steps involved in classification of rice diseases. The proposed technique involves the following steps:

• Image Acquisition.

• Image Pre-processing using median filter algorithm.

• Image Segmentation using K-means clustering

• Feature Extraction using logistic regression method.

• Image Classification using deep CNN with ResNet-50.

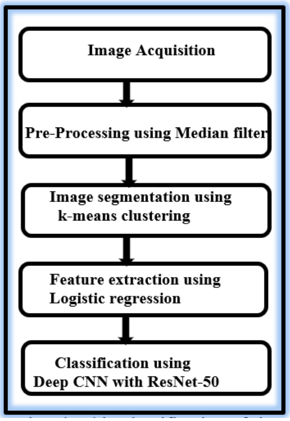

Figure 1. Steps Involved in Classification of Rice Diseases

A. Image Acquisition

Image acquisition means to collect different type of samples for the formation of the input dataset. The data for this research activity is collected from Kaggle repository [9] and Mendeley data [10]. Recently Kaggle is found to the most famous benchmark and real dataset provider for machine learning and artificial intelligence. Dataset images further go through the various steps.

B. Image Pre-processing Using Median Filter

The median filtering algorithm is employed in the second phase to eliminate noise from rice leaf images. The median value of the neighbourhood replaces the noisy value of the digital image or sequence (mask). The mask's pixels are ranked in order of their grey levels, and the group's median value is used to replace the noisy value. The median filtering output is $z(x, y) = \text{med}\{f(x - i, y - j), i, j \in W\}$ where $f(x, y)$ and $z(x, y)$ are the original image and the output image respectively, W is the two-dimensional mask: the mask size is n x n(where n is commonly odd) such as 3 x 3, 5 x 5, and etc.; the mask shape may be linear, square, circular, cross, and etc.

C. Image Segmentation Using K-means Clustering

For segmentation, K-means clustering is used. Clustering is the technique of grouping photographs based on specified criteria. The photos of infected and non-diseased leaves are grouped using the clustering approach. This approach is predicted to cluster both the afflicted and non-diseased parts. The K-means method divides an image (in this case a pixel) into p classes based on a set of features. The classification is accomplished by minimising the sum of squared distances between the objects and mapping to the centroid of associated groups. The K-Means clustering algorithm divides q photos into K clusters, with each image belonging to a cluster based on its centroid, mean intensity, and area. This algorithm creates a total of p clusters. The good number of clusters p that lead to the separation (distance) is not known in advance and must be calculated from the data set. K-Means clustering is useful for reducing total cluster variance or the square function.

D. Feature Extraction Using Logistic Regression Method

The process of converting input data into a set of features that can accurately describe the original data is known as feature extraction. It's a type of dimensionality reduction that's different from the others. The colour and texture of the crop are used to extract features in agricultural applications. Using the feature Extraction approach, we can create new features that are a linear mixture of current features. When compared to the original feature values, the new set of features will have a different value. The fundamental goal is to collect the same information with fewer characteristics. Principal component analysis is a technique for extracting key variables (in the form of components) from a large number of variables in a data set. It has a proclivity towards locating the highest variation (spread) in data. When working with three-dimensional or higher-dimensional data, PCA is more useful. The algorithm for the logistic regression approach is shown in the stages below.

The first step is to standardise the data (Leaf img).

Step 2: Find the Covariance Matrix.

Step 3: For the Covariance-matrix, calculate the Eigenvector and Eigenvalues.

Step 4: Sort all of the Eigenvalues in ascending order.

Step 5: Take the ordered Eigenvalues and normalise them.

Step6: Stack the Normalized_ Eigenvalues =W matrix horizontally.

Step7: X PCA=Leaf img.dot (W matrix) X PCA=Leaf img.dot (W matrix)

E. Rice Leaf Image Classification Using Deep CNN with Resnet-50

ResNet-50 is a sophisticated deep neural network that excelled in the classification challenge. Convolution, pooling, activation, and fully-connected layers are all stacked one on top of the other in a Deep Residual Network. The identity link between the layers is the only construction that turns a simple network into a residual network. ResNet-50 follows a four-stage architecture. Figure 2 illustrates this.

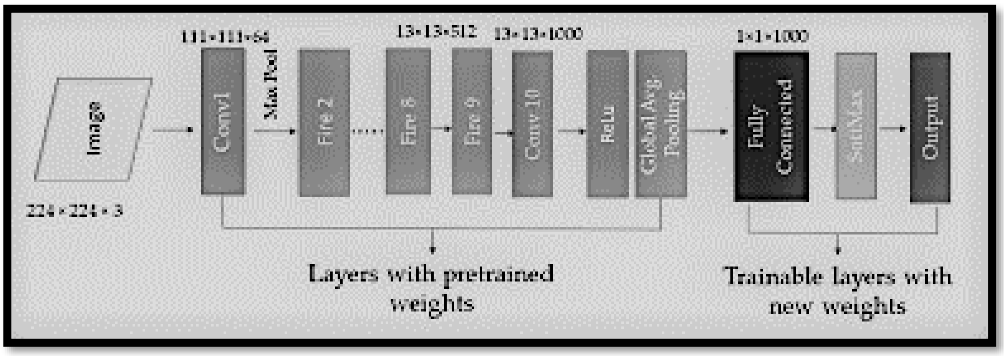

Figure 2. Architecture diagram of ResNet-50

The network can handle images with heights and widths that are multiples of 32, as well as a channel width of three. We'll assume the input size is 224 × 224 x 3 for the sake of simplicity. Initial convolution and max-pooling are performed with 77% and 33% kernel sizes, respectively, in every ResNet architecture. The network then enters Stage 1, which consists of three Residual blocks, each with three layers. The size of the kernels used to perform the convolution operation in each of the three levels of stage 1's block is 64, 64, and 128. The identity relationship is represented by the curved arrows. The dashed connecting arrow indicates that the Residual Block's convolution operation uses stride 2, resulting in a half-size input in terms of height and breadth but a doubling of channel width. The channel width is doubled and the size of the input is reduced to half as we advance from one stage to the next. Bottleneck architecture is utilised for deeper networks like ResNet-50, ResNet152, and so on. Three layers are placed one on top of the other for each residual function F. Convolutions are used in the first, third, and eleventh layers. The 11 convolution layers are in charge of shrinking and then expanding the dimensions.

4. Result and Discussion

The image recognition is done using the deep CNN with ResNet-50 method, and rice leaf illnesses in rice plants are discovered. This study uses a mixture of two separate datasets of 120 and 1308 rice leaf picture samples, respectively. The first set of data includes 120 samples from three disease classes: brown spot, bacterial leaf blight, and leaf smut. This dataset contains a total of 40, 40, and 40 samples, respectively. For Tungro illness, a total of 1308 photos are used in the second dataset. A total of 1428 pictures are used for pre-processing and classification after combining both datasets. The input rice disease leaf images were pre-processed using the median filter approach, and then the segmented image was further segmented using the k-means clustering algorithm. The segmented images were extracted with logistic regression and then categorised with deep CNN and the ResNet-50 algorithm. Biological leaf blight, brown spot, leaf smut, and tungro are the diseases that affect rice leaves. The output of the deep CNN using the ResNet-50 technique is shown in the diagram below. Figure 4 depicts a diseased rice leaf as an input image. Figures 5 and 6 show a rice leaf image that has been median filtered and segmented. Brown spot has been identified on rice leaves. Table 1 compares the proposed Deep CNN with ResNet-50 against several machine learning techniques. Figure 7 compares the performance of various machine learning methods and the suggested technique in terms of accuracy.

Figure 4. Input image

Figure 5. Filtered image

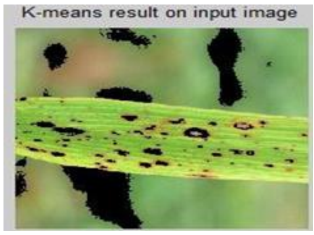

Figure 6. Segmented image

Table1. Comparison between various machine learning algorithm and proposed Deep CNN with ResNet-50

Techniques	**Accuracy**
Genetic Algorithm	96%
Adaboost Algorithm	88%
Bagging Algorithm	86%
Deep CNN with ResNet-50	97.3%

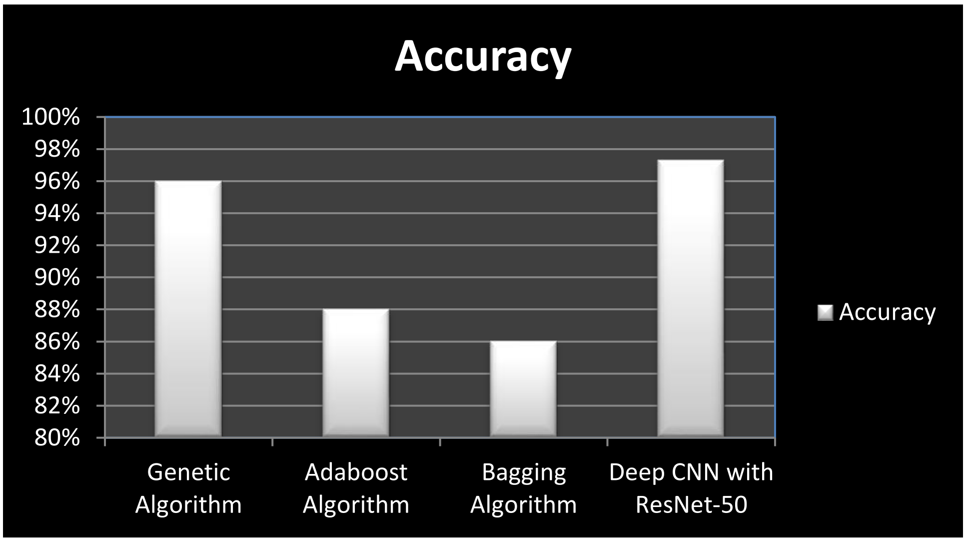

Figure 7. Performance comparison

5. Conclusion

This work presents a reliable investigation into the detection of rice leaf disease. Various approaches are utilized to identify rice illnesses, but the deep CNN with ResNet-50 algorithm is employed in this method to detect the disease that affects the plant. Because rice is used on a regular basis all around the world, it is critical to protect the crops. This strategy can be used in the agricultural field to protect crops from a variety of diseases and harm. Precautions must be taken to protect the crops. To produce the output, the deep CNN method may estimate and produce the result for a huge number of data sets. The condition can be identified with 97.3% accuracy. As a result, this approach may be able to protect the crop from illness. This proposed technology can be used to swiftly identify paddy leaf diseases on the Android and Windows platforms. It will undoubtedly assist farmers in classifying infected paddy leaves at an early stage, allowing them to prevent additional damage to their crops.

References

[1] Nilam Sachin Patil and E. Kannan, “Identification of Paddy Leaf Diseases using Evolutionary and Machine Learning Methods”,Turkish Journal of Computer and Mathematics Education, Vol.12 No.2 (2021), 1672-1686, 2021.

[2] Vimal K. Shrivastava , Monoj K. Pradhan, Sonajharia Minz , Mahesh P. Thakur, “Rice plant disease classification using transfer learning of deep convolution neural network”, The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, Volume XLII-3/W6, 2019.

[3] Ruoling Deng, Ming Tao, Hang Xing , Xiuli Yang , Chuang Liu , Kaifeng Liao and Long Qi, “Automatic Diagnosis of Rice Diseases Using Deep Learning”, Front. Plant Sci., 19 August 2021 | https://doi.org/10.3389/fpls.2021.701038.

[4] Shivam, Surya Pratap Singh and Indrajeet Kumar,”Rice Plant Infection Recognition using Deep Neural Network Systems”, International Semantic Intelligence Conference (ISIC 2021), Feb 25-27, New Delhi, India, 2021.

[5] G. Jayanthi, K.S. Archana and A. Saritha,” Analysis of Automatic Rice Disease Classification Using Image Processing Techniques”, International Journal of Engineering and Advanced Technology (IJEAT) ISSN: 2249 – 8958, Volume-8, Issue-3S, February 2019.

[6] D. Swathi and A. Bharathi, “Disease Classification of Paddy Leaves Using HSI Feature Extraction and SVM Technique”, IJSRD - International Journal for Scientific Research & Development| Vol. 4, Issue 02| ISSN (online): 2321-0613, 2016.

[7] Kawcher Ahmed, Tasmia Rahman Shahidi, Syed Md. Irfanul Alam and Sifat Momen,” Rice Leaf Disease Detection Using Machine Learning Techniques”, International Conference on Sustainable Technologies for Industry 4.0 (STI), 24-25 December, Dhaka,2019.

[8] Wan-jie Liang, Hong Zhang, Gu-feng Zhang and Hong-xinCao, ”Rice Blast Disease Recognition Using a Deep Convolutional Neural Network”, Scientific Reports | 9:2869 | https://doi.org/10.1038/s41598-019-38966-0, 2019.

[9] https://www.kaggle.com/minhhuy2810/rice-diseases-image-dataset

[10] https://data.mendeley.com/datasets/fwcj7stb8r/1