Published by :
http://www.ijert.org

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 13 Issue 11, November 2024

Agriculture and Machine Learning: Detection of Disease in Rice Leaves

Sreedeep Krishnan
Dept of ICE
Kalasalingam Academy of Research and Education, Tamil Nadu

M Karuppasamypandiyan
Dept of EEE
Kalasalingam Academy of Research and Education, Tamil Nadu

Alan S Puthussery
Dept of Robotics and Automation,
Adi Shankara Institute of Engineering and Technology

***Abstract*— One of the most widely cultivated crops in India is rice, and leaf diseases can significantly affect both yield and quality. Finding rice leaf diseases is crucial since they have a direct influence on the nation's food security and economy. Brown spot, leaf blast, and hispa are the three primary diseases that typically afflict rice leaves. To address this issue, we have looked at several deep-learning techniques to identify the illnesses on their leaves. We have also computed the accuracy, recall, and precision of each approach to evaluate its effectiveness. By identifying illnesses in rice leaves, this study aids rice growers in attaining a robust output. When compared to machine learning techniques, deep learning models exhibit superior performance. After analyzing many deep learning model, we found that the Densenet 201 model performed better in terms of accuracy than the CNN (Convolutional neural networks) VGG19, ResNet, and VGG16 models.**

***Keywords*- Deep learning, Convolutional neural networks, machine learning, Transfer learning**

I. INTRODUCTION

India grows rice all around the country, especially in West Bengal, and its agriculture industry is quite important. Since more than 75% of India's population is employed in agriculture, early illness diagnosis is essential. Although manual disease inspection was once widespread, it had disadvantages. More precise techniques are now available because to developments in deep learning and image processing. While deep learning uses neural networks for disease detection—essential for preserving agricultural output and food security—image processing uses filters and algorithms. In India, the economy depends heavily on agriculture, and progress in disease detection is critical to long-term viability.
Crops can suffer greatly from plant diseases, and early identification is essential for efficient management. Large crops cannot be detected using conventional techniques like visual inspection, which emphasizes the necessity for sophisticated detection techniques. The purpose of this study is to investigate and identify illnesses of rice leaves in order to enable timely measures. In India, rice is an essential crop due to its high yield and high nutritional content. Environmental variables, fungi, viruses, and pathogens are frequently the cause of diseases in rice. The fact that India is a significant rice producer and that its rice-growing region is still developing emphasizes how crucial disease identification and control are to the agricultural industry. [1]
Changes in the climate might be advantageous to pathogens and could impede crop development, especially in the initial phases. Diseases caused by fungi can impede crop growth and reduce yields. Large agricultural areas make it challenging to manually identify ailments, particularly in rice plants where farmers usually struggle to identify issues without expert advice, which adds time and cost. Rice leaves often have LeafBlast, Brown Spot, Hispa, and Healthy diseases. These illnesses need to be promptly detected and efficiently treated to maintain agricultural productivity. [7]
Monitoring big agricultural fields requires the detection of illnesses in leaves since they may seriously impair production and quality by destroying the green layer of the leaf. One such disease is that which affects rice leaves. Disease control depends on early and precise identification, followed by the right intervention. Deep learning networks and digital image processing provide quick, accurate, and time-saving illness diagnosis techniques. Plant protection strategies might benefit from advances in computer vision. There are obstacles, too, such the requirement for reliable data collecting and appropriate rice plant monitoring. Multimedia sensor installation in farms can aid in the periodic recording and monitoring of climate changes. Despite its advantages, this method needs constant upkeep and might be impacted by shadows in images that are taken, which could reduce accuracy.[1]

Healthy leaf Leaf Blast Brown Spot

Figure. 1. Types of diseases on rice leaves

II. LITERATURE REVIEW

Anthonys G. and N. Wickramarachchi's rice disease detection system successfully recognized three common diseases affecting rice: brown spot, rice blast, and rice sheath blight. For categorization, a number of characteristics, including color, texture, and form, were considered. However, they

Published by :
http://www.ijert.org

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 13 Issue 11, November 2024

only managed a 70% classification accuracy, necessitating the usage of three characteristics.

Using a Support Vector Machine (SVM) classifier, A. K. Singh, A. Rubiya, and B. Raja presented K-mean clustering-based segmentation to identify rice blast illness [5]. Their technique yielded an accuracy rate of 82%. However, just one illness was found.

The Sladojevic et al. paper shows how well deep neural networks—more especially, Convolutional Neural Networks, or CNNs—work for classifying leaf images and identifying plant diseases. Their program identified 13 distinct plant disease types with an astounding 96.3% accuracy rate. This high degree of accuracy demonstrates how CNNs may transform agricultural disease diagnosis, providing farmers with a dependable and practical way to track and manage crop health.

Table.1. – An overview of the research that has been done on the use of several classifiers to identify leaf diseases in brief.

Article	Crop	Dataset	Classifier	Accuracy	Advantages	Drawbacks
[9]	Rice	500 lab images	Convolutional Neural Network	95.48%	The proposed The solution is both feasible and effective	Used small dataset for training
[9]	Rice	Kaggle dataset with 1000 images	Image processing techniques and SVM	80%	Easy to train and flexible model	Low performance
[8]	Groundnut	100 lab images	Back Propagation	97.41%	Best Performance	Works withonly one groundnut disease i.e.,
[3]	Corn	1292 lab images	Binary SVMand Multi-Class SVM	85%	Exhibits best performance onthe dataset created	Only suitable to cornfield and itis not suitable for other fields
[1]	Peanut	6029 Lab images	ResNet50 DenseNet121	97.59%	Best prediction performance	It is complicated in field environment
[1]	Groundnut	250 lab images	KNN	92.14%	Best prediction performance	Need to reduce the probability of false classification
[7]	Papaya	160 lab images	Random Forest	70.14%	Easy to train	Low accuracy
[12]	Cotton	2400 lab images	CNN	96.4%	High accuracy	Doesn't suggest remedies

The above literature research makes it clear that different environmental conditions and geographic locations have an impact on the patterns of plant diseases. Deep learning models have been pushed by a number of researchers as a way to improve plant disease detection accuracy. Moreover, it has been observed that deep learning models are advised only in situations when a sizable dataset is available. We provide a collection of 3095 images and investigate a fungal disease that affects rice plants. As a result, we suggest using an automated deep learning model for these illnesses' prediction in the part that follows.

III. METHEDOLOGY

A. DATASET DESCRIPTION

Three different types of sick rice leaf images—Healthy, LeafBlast, and Brownspot—make up the leaf dataset used in this study. This dataset includes 3095 images of rice leaves displaying different illness signs, including 418 images of Brownspot, 623 LeafBlast images, and 1191 Healthy images. All of these images are in JPEG format, with a 1449 * 1449 width and high resolution. Brownspot, LeafBlast, and Healthy are all combined in one leaf dataset. The example images are displayed in Figure 1.

Table 2. Details on the dataset

Disease Type	Train Images	Test images	Validation Images
Healthy	1191	155	297
Leaf blast	623	90	156
Brown Spot	418	60	105

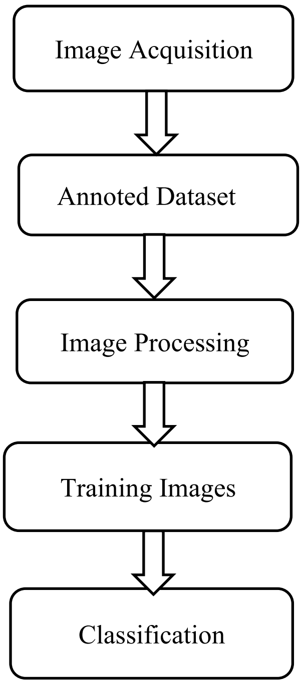

Figure. 2. Block Diagram of leaf disease detection system

The major steps of the rice leaf disease prediction system are listed below. Deep learning, machine learning, and image processing techniques should be used to accomplish these processes.

a) Image acquisition

Four types of damaged rice leaf images—Healthy, Brownspot, LeafBlast, —make up the leaf dataset used in this study. 1600 images of rice leaves with different disease signs

Published by :
http://www.ijert.org

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 13 Issue 11, November 2024

are included in this dataset; there are 400 images of brown spots, 500 images of hispa, 300 images of leaf blast, and 400 images of healthy leaves. All of these images are in JPEG format, with a 1449 by 1449 width and high resolution. Since the backdrop of the image is white as well, background removal is not necessary. A combination of Brownspot, LeafBlast, and Healthy may be found in this leaf dataset.

b) Annoted Data Set

In order to achieve accuracy and dependability, building a leaf identification system requires careful consideration of several factors. To depict various leaf diseases like Leafblast,hispa,brownspot,Healthy. The next stage is to put together a varied picture set for every class, taking pictures of different backgrounds, lighting, and viewpoints to provide more depth to the dataset.

Each image has metadata that includes information of the leaves diseases, and the dataset is carefully arranged into training, validation, and testing sets. Verifying the accuracy and consistency of the annotations is a crucial step in ensuring the dependability of the dataset. Data augmentation techniques such as flips, rotations, and lighting modifications are used to improve dataset variety. The annotations are kept in a file format that is compatible with machine learning frameworks, such CSV, JSON, or XML.

c) Image Pre processing

In data used for analysis, noise, missing numbers, outliers, and discrepancies are commonly discovered in the current global environment. It takes preparation to increase the quality and accuracy of data. It involves transforming high-level data into a lower level for easier computations, cleaning the data to remove noise, and reducing data dimensions without compromising data quality. Data reduction reduces dimensionality without compromising data quality, data transformation facilitates calculations, and data cleaning eliminates noise. One method that aids in effective data preparation for better analysis and decision-making is data cube aggregation.

d) Classification

ResNet requires the use of a deep convolutional neural network, namely ResNet50, for image classification. ResNet gathers hierarchical information from images as they pass through successive layers, honing representations to discover complex patterns. ResNet50 is trained on large-scale datasets such as ImageNet, which allows it to identify discriminative characteristics and facilitate precise categorization. Prior to feeding input images into the trained ResNet50 model, preprocessing is done during inference. The content of the image is identified by the model through the use of learned weights in class probabilities calculations. In the end, the model generates the classifications that are most likely to occur, producing highly accurate categorization results.

1) Densenet

DenseNet is the most sophisticated parameter-light CNN architecture available for visual object identification. Comparably, DenseNet and ResNet are similar, with a few notable differences. ResNet uses an additive attribute (+) to blend the previous layer with the future layers, whereas DenseNet uses the output of the previous layer's concatenated (.) characteristics to integrate it with a future layer. DenseNet Architecture connects everything in an attempt to solve this issue. [13] This study employed the DenseNet-121 [5+(6+12+24+16) × 2)=12 1] architecture among the many DenseNet (DenseNet-121, DenseNet-160, and DenseNet-201) designs. The following are the DenseNet-121's details: There are three transition layers (6, 12, 24), one layer for classification, five layers for convolution and pooling, and two denseblocks (one × 1 and three × 3 convolution). To find the output layers (lth), conventional CNNs usually apply a non-linear transformation Hl (.) to the output of the previous layer Xl−1.

$$Xl = Hl\ (Xl-1) \qquad (1)$$

Rather than actually summing the inputs, DenseNets concatenate the layer output functionality maps with the inputs. DenseNet offers a straightforward model of cross-layer connection to enhance information flow: The layer at the top receives inputs from all features from the ones before it: The equation is then modified one more to:

$$Xl = Hl\ (2)\ [\ (X0, X1, X2, \ldots, Xl-1)], \qquad (2)$$

where the single tensor [X0, X1, X2,..., Xl−1,] is formed by merging the output maps of previous layers. Among the functions, Hl () is a non-linear transformation function. This function consists of three major operations: batch normalization (BN), activation (ReLU), and pooling and convolution (CONV). The DenseNet architecture is seen in Figure 1. However, the growth rate k helps to generalize the lth layer in the following ways:

$$K[l] = (k\ [0] + k\ (l-1)). \qquad (3)$$

if it is believed that the number is k[0].

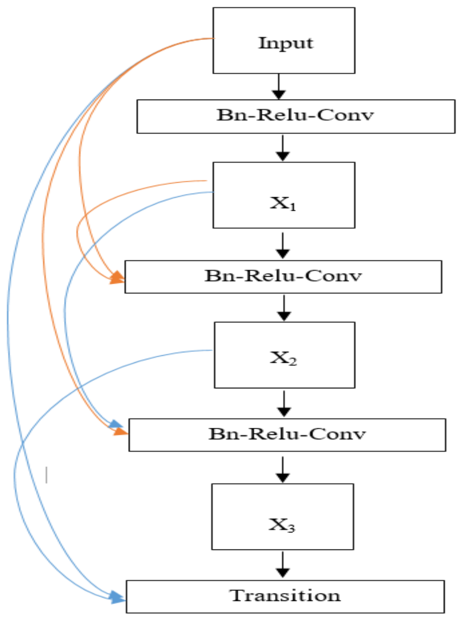

Figure. 3. DenseNet Architecture

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 13 Issue 11, November 2024

2) ResNet50

As an illustration Utilizing Technology for Artificial Neural Networks In this instance, ResNet (Residual Network) has been used to solve the vanishing/exploding gradient issue. We employ a method in this network called skip connections. By skipping many training phases, the skip connection establishes a direct link to the output. The shortcut link leads to a 34-layer basic network design patterned by the VGG-19. These shortcut connections cause the design to eventually degenerate into a residual network. Residual neural networks (ResNets) are artificial neural networksthat are modeled after the cerebral cortex's pyramidal cell architecture. Residual neural networks employ skip connections, often called shortcuts, to go over certain layers. Double- or triple-layer skips with nonlinearities (ReLU) and batch normalization are used in the majority of ResNet models.

3) VGG-19

With 19 layers, VGG-19 is a deep Convolutional Neural Network (CNN). It was trained using the ImageNet database, which comprises over a million images divided into 1000 different object categories, and is frequently used for image classification applications.
A fixed-size RGB picture of 224x224 pixels is provided as input when using VGG-19, and it is represented as a matrix of shape (224, 224, 3). Subtracting from each pixel the mean RGB value obtained for the whole training set is the only preprocessing step.
For convolution operations, VGG-19 employs kernels with a size of 3 by 3 and a stride of one pixel. The image's spatial resolution is preserved by the use of spatial padding. Over a 2x2 pixel frame, maximum pooling with a stride of 2 is performed. VGG-19 introduces non-linearity using the Rectified Linear Unit (ReLU) activation function, which enhances the model's classification performance and processing speed. Previous models employed tanh or sigmoid activation functions.

4) VGG-16

We have used the sixteen-layer VGG-16 model in the implementation. The RGB channels are represented by the number 3, and the input picture size is set to 224x224x3. Next, the input is linked to Conv1, the initial convolutional layer. The task of extracting features from the input picture falls to this layer. For feature extraction and reduction, respectively, a mixture of convolutional and max pooling layers is then used. Eleven convolutional filters are employed in this configuration to extract features from the input picture. The next layers then get these attributes for additional processing. By adding a row and column on each side of the picture, "same padding" is used to guarantee that features from the corners of the image are retrieved. The filter travels one pixel at a time when the stride is set to 1. Five max pooling layers that come after parts of the convolutional layers are used for spatial pooling. By lowering the feature maps' spatial dimensions, max pooling lowers computing cost and guards against overfitting.

5) CNN

Using image analysis, Convolutional Neural Networks (CNNs) play a key role in the detection of illnesses affecting rice leaves. These networks take in pictures of diseased rice leaves as input and use hierarchical characteristics to identify patterns that point to certain diseases. CNNs are made up of pooling layers that downsample while preserving crucial characteristics, activation functions that add non-linearity, and convolutional layers that apply filters to capture features like edges and textures. Afterwards, fully connected layers are used to classify the retrieved characteristics, and the last layer presents a probability distribution of the various illness classifications. CNNs can effectively identify rice leaf diseases by using a collection of annotated pictures as training material. This allows farmers to identify illnesses automatically and effectively, assisting in timely crop management and protection.

IV. RESEACRCH FINDING AND RESULTS

In selecting the optimal deep learning algorithm for our task, it proves beneficial to scrutinize various models applied to the identical rice dataset. These encompass VGG-16, VGG-19, ResNet, Densenet, and CNN. Delving into the accuracy and loss metrics during both training and validation stages aids in gauging their efficacy on the dataset. Validation accuracy and loss metrics shed light on the model's capacity to generalize to unseen data, while training accuracy and loss metrics unveil its performance on the training data.

```
14/14 [==============================] - 19s 1s/step
              precision    recall  f1-score   support

   BrownSpot       0.60      0.62      0.61       118
     Healthy       0.81      0.83      0.82       259
   LeafBlast       0.79      0.64      0.71        70

    accuracy                           0.75       447
   macro avg       0.73      0.70      0.71       447
weighted avg       0.75      0.75      0.75       447
```

Figure 4. The precision of the Brown spot, Healthy, and Leaf-blast tests

The precision attained after each iteration is shown on the y-axis in Figure 5, while the number of iterations is shown on the x-axis. Monitoring the training and validation accuracy trends over iterations is crucial. The model is learning efficiently without overfitting if accuracy increases steadily for both training and validation sets. Conversely, a significant discrepancy in accuracy between training and validation sets might be a sign of overfitting.
The loss at the start of training and the subsequent rise or decrease in loss following each repetition are shown on the y-axis in Figure 6. The iterations are shown on the x-axis. It is important to keep an eye on both training and validation loss, much like accuracy. When the training and validation losses go downward, the model is learning efficiently. On the other hand, overfitting can be indicated if the training loss keeps going down but the validation loss keeps going up. In general, you can use these metrics to evaluate each algorithm and choose the best method for your rice dataset depending on how well it performs in terms of accuracy and loss.

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 13 Issue 11, November 2024

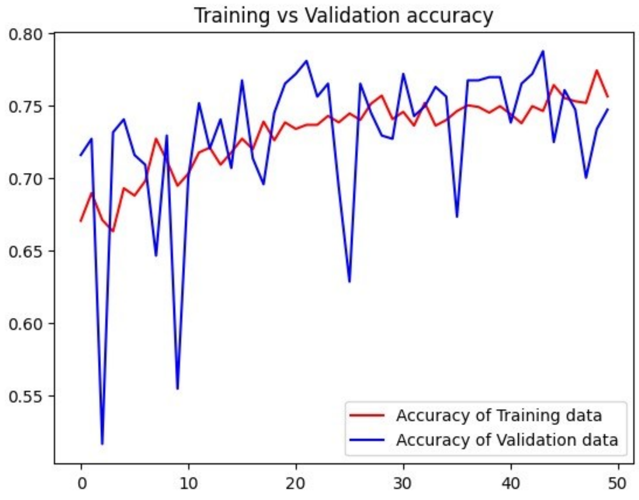

Fig. 5. Graph of Train and Validation Accuracy

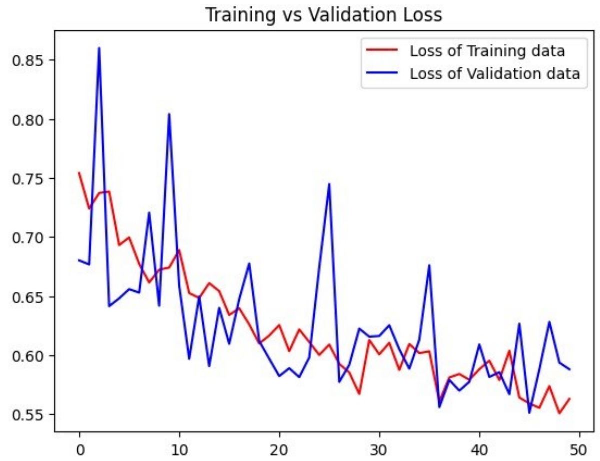

Fig. 6. Graph of Train and Validation Loss

V. CONCLUSION AND FUTURE WORKS

In this work, we employed a range of deep learning (DL) approaches to classify four distinct forms of rice leaf diseases. Using a dataset of broken rice leaves, we used a range of widely used deep learning algorithms, such as VGG19, VGG16, ResNet, Densenet 201 and a convolutional network. Our findings indicate that the Densenet 201 architecture outperformed the other models in the diagnosis of rice leaf diseases. Table 3 illustrates how our recommended Densenet model beat the industry standard deep learning models in terms of accuracy by about 4%. By adjusting training parameters such as learning rate, epoch count, and optimizer techniques, we were able to use a custom model with fewer layers than standard models and achieve significant improvements in accuracy.

Farmers must identify diseases effectively in order to save their crops. In order to improve the breadth, simplicity, and speed of illness identification, we intend to broaden the scope of our study in the future by incorporating other diseases and methods.

Table.3. Accuracy on Different DL Algorithms

ALGORITHMS	ACCURACY
VGG 16	58%
VGG 19	71%
CNN	71%
Densenet201	75%

REFERENCES

[1] A. Jayakumar, "Detection and Classification of Leaf Diseases in Maize Plant using Machine Learning MSc Research Project Data Analytics." [Online]. Available: https://agriculturegoods.com/why-is-agriculture-important/

[2] R Salini, A.J Farzana, B Yamini, (2021). Pesticide Suggestion and Crop Disease classification using Machine Learning , 63(5), 9015-9023.

[3] Mahalakshmi, S. D., & Vijayalakshmi, K. (2021). Agro Suraksha: pest and disease detection for corn field using image analysis. Journal of Ambient Intelligence and Humanized Computing, 12(7), 7375-7389.

[4] H. B. Prajapati, J. P. Shah, and V. K. Dabhi, "Detection and classification of rice plant diseases," Intelligent Decision Technologies, vol. 11, no. 3, pp. 357–373, 2017.

[5] I. H. Witten, E. Frank, M. A. Hall, and C. J. Pal, Data Mining: Practical machine learning tools and techniques. Morgan Kaufmann, 2016.

[6] R. Kaur and V. Kaur, "A deterministic approach for disease prediction in plants using deep learning," 2018.

[7] Ramesh, S., Hebbar, R., Niveditha, M., Pooja, R., Shashank, N., & Vinod, P. V. (2018, April). Plant disease detection using machine learning. In 2018 International conference on design innovations for 3Cs compute communicate control (ICDI3C) (pp. 41-45). IEEE.

[8] Ramakrishnan, M. (2015, April). Groundnut leaf disease detection and classification by using back propagation algorithms. In 2015 international conference on communications and signal processing (ICCSP) (pp. 0964-0968). IEEE.

[9] Rahman, C. R., Arko, P. S., Ali, M. E., Khan, M. A. I., Apon, S. H., Nowrin, F., & Wasif, A. (2020). Identification and recognition of rice diseases and pests using convolutional neural networks. Biosystems Engineering, 194, 112-120.

[10] T. Akter, M. T. Parvin, F. A. Mila, and A. Nahar, "Factors deter mining the profitability of rice farming in bangladesh," Journal of the Bangladesh Agricultural University, vol. 17, no. 1, pp. 86–91, 2019.

[11] Ramakrishnan, M. (2015, April). Groundnut leaf disease detection and classification by using back propagation algorithms. In 2015 international conference on communications and signal processing (ICCSP) (pp. 0964-0968). IEEE.

[12] Ramesh, S., &Vydeki, D. (2018, September). Rice blast disease detection and classification using machine learning algorithm. In 2018 2nd International Conference on Micro-Electronics and Telecommunication Engineering (ICMETE) (pp. 255-259). IEEE.

[13] Hasan N, Bao Y, Shawon A, Huang Y. DenseNet Convolutional Neural Networks Application for Predicting COVID-19 Using CT Image. SN Comput Sci. 2021;2(5):389. doi: 10.1007/s42979-021-00782-7. Epub 2021 Jul 23. PMID: 34337432; PMCID: PMC8300985.