International Journal of Scientific Research & Engineering Trends
Volume 9, Issue 2, Mar-Apr-2023, ISSN (Online): 2395-566X

Predicting Different Types of Paddy Leaf Diseases Using Convolutional Neural Network (CNN)

Assistant Professor Mrs. G.Santhiya , Anusubha k , Jenisha Joy MJ, Reshma V, Snekha D .

Stella Mary's College of Engineering

Abstract: Most of the countries are depends on agriculture, where Tamil nadu is the land of agriculture. Here paddy cultivation is major source of earning. People in Tamil nadu, consumes rice as main meal for three times in a day. Various factors such as diseases on paddy leaf, pest attack etc., the production of paddy will be affected approximately 40stage to protect the paddy because it will destroy the entire farm land. If the diseases are identified in initial stage there is no need to spray a high dose fertilizer on the paddy crops. To overcome this, the proposed system uses pre-processing, transfer learning Inception V3 method, neural network are trained by deep learning based Convolutional Neural Network(CNN) classification algorithm to identify the paddy leaf diseases like bacterial leaf blight, brown spot and rice blast. This method produces good accuracy. Scope of this project is to detect disease on paddy crops and too notify the types of diseases to farmer so that the farmers can take early action to protect the paddy crops.

Index Terms—Paddy cultivation, disease detection, Tamil Nadu agriculture, leaf diseases, bacterial leaf blight, brown spot, rice blast, transfer learning, Convolutional Neural Network, CNN, deep learning, image preprocessing, early disease detection, crop protection

I. INTRODUCTION

Agriculture is the primary occupation in India. Various factors such as climate change, soil condition, nutrient level of plant, various diseases, and insect pests affect crop production. Many farmers struggle to identify diseases in crops, leading to agricultural losses. Image processing techniques can help overcome these challenges by extracting leaf features for easier disease clas- sification and detection [1], reducing farmers' workload. Crop disease detection can be accomplished by observing spots on affected paddy leaves, using image processing combined with deep learning based on Convolutional Neural Network (CNN) algorithms [2]. Tamil Nadu is an agricultural region where people consume rice as their main meal three times daily. Like other crops, rice is susceptible to numerous diseases that vary by region and season [3]. Despite increasing technological implementations in agriculture, farmers in India still rely on manual methods to identify diseases.

A. Diseases on Paddy Crop

Paddy is a major crop for many farmers, grown in two sea- sons: rabi and kharif, each requiring six months for production. Environmental changes can lead to paddy diseases and pest attacks, resulting in crop losses. The major paddy diseases include:

1) Bacterial Leaf Blight: Bacterial leaf blight is caused

by "Xanthomonas Oryzae" and primarily occurs during the monsoon (kharif) season. Symptoms include water-soaked lesions starting at leaf margins and spreading to the leaf base [4]. If not identified early, the affected areas become yellowish to light brown with yellowish borders between dead and green parts of the paddy leaf. The disease is promoted by rainy, cloudy, and stormy climate, as well as high nitrogen fertilizer use. Bacterial leaf blight causes significant losses in Tamil Nadu, Kerala, coastal Andhra Pradesh, and Orissa.

2) Brown Spot: Brown spot is caused by the fungus

"Helminthosporium oryzae" and mainly occurs during the kharif season [5]. Symptoms begin as small brown spots that develop into round or oval brown spots on leaves. The disease transfers from infected paddy seeds and causes fewer seedlings. This disease is commonly found in Orissa, Tripura, and Tamil Nadu.

3) Rice Blast: Rice blast occurs in both kharif and rabi

seasons (November to February) and is caused by the fungus ”Magnaporthe Oryzae.” Symptoms include elliptical spots with light-colored centers and reddish edges [6]. This disease can reduce yields by up to 50% and is found in Tamil Nadu, Kerala, and Karnataka.

Convolutional Neural Network

In neural networks, Convolutional Neural Network (CNN) is a crucial method for image recognition, classification, and object detection [7]. CNN image classifications take images as input, process them, and classify them under certain labels (e.g., blight, blast, brown spot). Computers interpret images as arrays of pixels, with pixel values dependent on image resolution. Based on these values, computers view images as h × w × d arrays (h = Height, w = Width, d = Dimension). For example, an image might be represented as a 6 × 6 × 3 array, where 6 refers to height, 6 to width, and 3 to RGB values [2].

1) Input Layer: The input layer of the model is fed by an

RGB image of size w0 × h0, where w0 is the width and h0 is the height of the image.

2) Convolution Layer: A convolution layer’s primary task is to identify local conjunctions of features from the previous layer and map their presence to a feature map. The model uses three convolution layers with several filters to generate output feature maps. These maps store information about where features appear in the image and how well they match the filter [2]. Each filter is trained spatially regarding its position in the volume and detects specific features from rice leaf disease images.

3) Pooling Layer: Pooling reduces variance and computational complexity by decreasing the number of parameters to learn. It performs down-sampling along spatial dimensions, reducing feature map dimensions and summarizing features that appear in portions of the feature map generated by the convolution layer. This makes the model more robust to variations in rice leaf disease image feature locations [8]. The model employs three pooling layers (Pooling1, Pooling2, and Pooling3) with parameters designed for RGB images of size 256 × 256 and 2 × 2 pool.

4) Dense Layer: The output from the final max pooling layer is flattened into a one-dimensional vector to feed into a fully connected dense layer [2]. This layer produces a one-dimensional vector of size 64, which feeds into a second fully connected dense layer to produce a one-dimensional vector of size 5.

5) Output Softmax Layer: The output layer applies the softmax activation function, which exponentially normalizes the dense layer output and produces a probability distribution across the three different rice leaf disease classes [2].

C. Objectives

• To identify leaf diseases based on training and classification

• To identify the type of disease

• To notify farmers so they can take early action

D. Problem Identification

• Farmers still use human vision-based approaches to detect rice leaf diseases, making expert advice time-consuming and expensive

• Human vision methods have drawbacks, including inaccurate disease type prediction in initial stages, resulting in reduced rice production

E. Motivation

Agriculture is the backbone of India. Each year, approximately 30% to 40% of crops are damaged due to paddy diseases. These diseases vary by region and season, making it difficult for farmers to identify various paddy leaf diseases. This project aims to find and classify paddy leaf diseases to help farmers detect diseases early and take appropriate actions to protect their crops. Early disease identification reduces the need for high-dose fertilizer application. Image processing combined with deep learning methods can help prevent these difficulties.

II. LITERATURE SURVEY

This literature survey examines various approaches to rice disease detection through image processing and deep learning techniques. Ai et al. (2020) proposed a convolutional neu- ral network using the Inception-ResNet-v2 model for crop disease identification, achieving 86.1% recognition accuracy using cross-layer direct edges and multi-layer convolution, though large images reduced efficiency. Archana et al. (2018) focused on automatic rice leaf disease segmentation using threshold value and color index approaches for pre-processing, specifically targeting Bacterial Leaf Blight and Brown Spot diseases that significantly impact crop yield and quality. Chen et al. (2020) employed a deep transfer learning approach using DenseNet pre-trained on ImageNet alongside Inception modules to classify various rice diseases including stack burn, leaf scald, leaf smut, white tip, and bacterial leaf streak, with plans to deploy on mobile devices. Ferentinos et al. (2018) de- veloped CNN models using an open database of 87,848 images representing 58 distinct plant-disease combinations, creating a system with significant potential as an advisory or early warning tool for agriculture. Jiang et al. (2020) combined CNN for feature extraction with SVM for classification, achieving a 96.8% recognition rate with optimal parameters (C=1, g=50), outperforming traditional neural network approaches. Liang et al. (2019) created a novel rice blast recognition method using CNN with a dataset of approximately 2,900 positive and negative samples, demonstrating that CNN features were more discriminative than traditional handcrafted features like LBPH and Haar-WT. Li et al. (2020) used Faster-RCNN with a custom DCNN backbone for video detection of rice diseases, simultaneously identifying rice sheath blight, rice stem borer, and rice brown spot, though very blurry lesions remained problematic. Another study by Li et al. (2021) used CNN to classify 26 diseases across 14 fruit and vegetable crops, achieving 86.67% training accuracy and 84.17% testing accuracy. Lu et al. (2017) proposed a method using sparse auto-encoding to learn features from 500 images of healthy and diseased rice leaves before applying convolutional and pooling layers for classification. Patil et al. (2021) developed a MATLAB-based system using cascaded classifiers and genetic algorithms with nearest neighbor approaches to identify paddy leaf diseases for implementation across multiple platforms. Ramesh et al. (2020) introduced an Optimized Deep Neural Network with Jaya Optimization Algorithm achieving high accuracy (90-92%) for classifying bacterial blight, brown spot, sheath rot, and blast diseases. Chopra et al. (2021) compared eight pre-trained models with transfer learning, finding DenseNet achieved the best results at 99% accuracy. Saleem et al. (2020) implemented three deep learning meta- architectures (SSD, Faster RCNN, and RFCN) for localization and classification of plant diseases, with SSD using Adam optimizer achieving the highest mean average precision of 73.07%. Sethy et al. (2020) extracted deep features from 5,932 contaminated rice leaf images using various architectures (AlexNet, VGG16/19, GoogleNet, ResNet, etc.) before classi- fying with SVM. Sladojevic et al. (2016) developed a model recognizing 13 plant disease types with accuracy around 91% using the Caffe deep learning framework. Sulistyaningrum et al. (2020) implemented multilevel SVM with Fuzzy C- Means segmentation, achieving 86.51% accuracy. Wang et al. (2017) used transfer learning with VGG16 to classify apple leaf diseases by severity stage with 90.4% accuracy. Another study by Wang et al. (2018) compared six CNN architectures using 2,430 natural environment images covering two crop species and eight diseases. Finally, Wang et al. (2021) proposed an ADSNN-BO model based on MobileNet structure with augmented attention mechanisms and Bayesian optimization for parameter tuning.

III. EXISTING SYSTEM

The existing method utilizes high-level features extracted by CNN which are more discriminative and effective than traditional handcrafted features including Local Binary Patterns Histograms (LBPH) and Haar-WT (Wavelet Transform) [9]. Moreover, quantitative evaluation results indicate that CNN with Softmax and CNN with Support Vector Machine (SVM) have similar performances with gained features [10]. SVM method deploys the hybrid layer with the consistence at the certain regional model with connected layer of pre-trained model. In order to increase the quality and yield of the rice crop, many modern and scientific techniques are being tested and adopted by the local farmers each year. Early disease detection is a major challenge in agriculture field. Hence proper measures have to be taken to fight bio aggressors of crops while minimizing the use of pesticides. The techniques of machine vision are extensively applied to agricultural science, and it has great perspective especially in the plant protection field, which ultimately leads to crops management. Our goal is early detection of bio aggressors [11]. A bag- of-features approach is used to automate rapid-throughput taxonomic identification of stonefly larvae. 263 stonefly larvae were collected of four stonefly taxa from freshwater streams in the mid-Willamette valley and Cascade Range of Oregon. Approximately ten photos were obtained of each specimen, which yields 20 individual images. These were then manually examined, and all images that gave a dorsal view within 30 degrees of vertical were selected for analysis [12]. The images were then classified through a process that involves: Identification of regions of interest, representation of those regions as SIFT vectors, classification of the SIFT vectors into a histogram of detected features, and classification of the histogram by an ensemble of logistic model trees. In their work, they have applied three region detectors: Hessian-affine detector and the Kadir entropy detector, including a newly developed Principal Curvature-Based Region (PCBR) Detector [13]. The construction of a codebook was performed by a Gaussian Mixture Model (GMM). The disadvantages of this approach include more human efforts, low area under curve, performance with less accuracy, increased static features, and increased accuracy of maintenance [14].

IV. PROPOSED SYSTEM

This document presents a comprehensive system for rice leaf disease detection using deep learning techniques. The proposed deep neural network effectively classifies crop diseases from leaf images using Convolutional Neural Network (CNN) trained on disease-affected regions [15]. The system utilizes three distinct datasets—bacterial leaf blight, leaf blast, and brown spot—sourced from Kaggle. The workflow begins with image collection, followed by dataset splitting for training and testing purposes, and culminates with disease identification and visualization for user output [16]. The approach employs hybrid models based on deep features and CNN networks, specifically leveraging transfer learning with feature extraction from Inception V3's fully connected layers. Rather than training the convolutional layers, the method maintains pre-trained Inception V3 weights while only training the dense layers from randomly initialized weights, requiring images to be resized to a default dimension before processing [17]. This methodology offers several advantages: reduced processing time, decreased human effort, agricultural accessibility for younger generations, and increased crop yields. The system architecture comprises both training and testing phases. In the training phase, the dataset contains numerous paddy-infected leaf images across three common rice diseases (bacterial leaf blight, leaf blast, and brown spot) stored in separate folders. Pre-processing removes unwanted noise from infected images before feature extraction, with loss function calculation feeding into the CNN classification algorithm. Figure 4.1 shows the Block Diagram of Proposed model. The testing phase utilizes single images to predict disease classi- fication. The workflow follows a systematic process: dataset collection, pre-processing with augmentation, model building using transfer learning approaches, loss function calculation, parameter optimization, and finally disease prediction. The dataset was collected from Kaggle's rice leaf disease detection repository, containing 40 images for each of the three diseases (bacterial leaf blight, brown spot, and rice blast), totaling 120 images. The dataset was loaded as an image data-store with automatic classification based on folder names, then partitioned with 70% for training (157 images) and 30% for validation (84 images). Image pre-processing addresses potential issues like dust, dewdrops, and insect excrement that could appear as noise, as well as distortions from water drops and shadow effects. Figure 4.2 shows the Flow chart of the proposed model. Image resizing reduces pixel count to 50×50, which decreases neural network training time by reducing input nodes and model complexity. Image augmentation artificially expands the dataset to 628 images through techniques like rotation (45° and 75°) and blurring (10×10 pixel size). This approach creates multiple variants of each original image, making the model more robust to variations in image quality and orientation. The training module employs transfer learning with Inception V3, reusing a pre-trained model to reduce training time and enhance performance. Inception V3 offers higher efficiency through a deeper but computationally less expensive network that incorporates auxiliary classifiers as regularizers. The model's architecture has been optimized through factorization into smaller convolutions, spatial factorization into asymmetric convolutions, utility of auxiliary classifiers, and efficient grid size reduction.

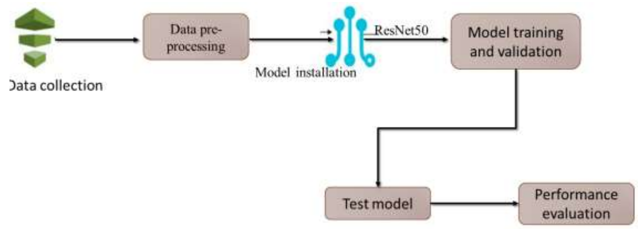

Fig 1. Block diagram for proposed system.

The final CNN model features a 10-layer architecture: input layer, three convolutional layers each followed by max pool- ing, two dense layers, and a softmax output layer. Keras APIutilizes Inception V3 for feature extraction, with ReLU pool- ing to reduce information, a flatten layer to convert matrices to vectors, and dropout layers to prevent overfitting. The dense layer with softmax activation ultimately classifies the paddy leaf diseases. The testing module predicts whether leaves are healthy or affected by bacterial leaf blight, brown spot, or rice blast, verifying the model's classification accuracy.

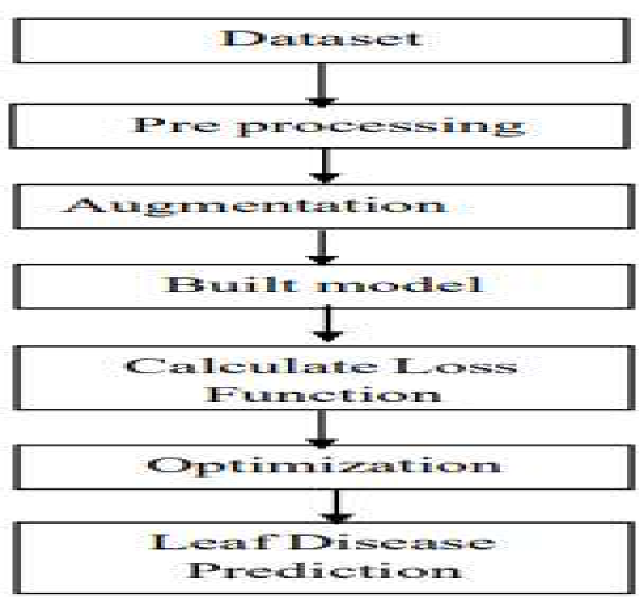

Fig 2. Flow Chart of the Proposed Diagram.

V. SYSTEM ANALYSIS

This system for rice leaf disease detection requires specific hardware and software configurations to operate effectively. The hardware requirements include a Pentium IV 3.5 GHz or later processor, 40 GB hard disk space, a 14-inch color monitor, an optical mouse, and 1 GB RAM. For software, the system runs on Windows 10 operating system, utilizes Python programming language, and employs Jupyter Notebook as the development environment. Windows, developed by Microsoft, dominates the personal computer market with over 90% market share despite losing ground in the overall operating system market to Android in 2014 due to smartphone proliferation. Originally introduced in 1985 as a graphical operating system shell for MS-DOS, Windows has evolved through numerous iterations to become the standard for desktop computing. Jupyter Notebook, a component of Project Jupyter, serves as a web-based interactive computational environment for creating and sharing documents containing live code, equa- tions, visualizations, and narrative text. The project's name references the three core programming languages it supports: Julia, Python, and R. Jupyter Notebook operates through a browser-based REPL built upon popular open-source libraries including IPython, ØMQ, Tornado web server, jQuery, Bootstrap front-end framework, and MathJax. It has gained significant adoption in cloud computing platforms like Amazon's SageMaker, Google's Colaboratory, and Microsoft's Azure Notebook. Python, created by Guido van Rossum and first released in 1991, is the primary programming language used in this system. It is an interpreted, high-level, general-purpose programming language known for its emphasis on code readability through significant whitespace usage. Python supports multiple programming paradigms including structured, object-oriented, and functional programming. Its philosophy emphasizes simplicity and readability, encapsulated in principles like ”Beautiful is better than ugly” and ”Simple is better than complex.” The language features dynamic typing, garbage collection, and a comprehensive standard library that makes it ideal for application development. Python's extensible nature makes it particularly suitable for implementing programmable interfaces in existing applications, while its clean syntax allows developers to express concepts in fewer lines of code than would be possible in languages like C++ or Java.

VI. CONCLUSION

The result and discussion section presents comprehensive findings from the rice leaf disease detection system. The input images for identifying paddy leaf diseases using CNN include JPG and PNG formats of bacterial leaf blight, rice blast, and brown spot diseases. Pre-processing methods standardize input images to 50×50×3 pixels for consistent training. The neural network training utilizes the pre-trained Inception V3 transfer learning model combined with CNN classification, running through 30 epochs. The system achieved 98.44% accuracy on training with a loss rate of 0.0725, validation loss of 0.9895, and validation accuracy of 0.7500. The training progress shows steady improvement from 80.65% accuracy at epoch 5 to 98.44% by epoch 30, while validation accuracy stabilized at 75% from epoch 15 onward. Analysis graphs demonstrate the relationship between training accuracy, validation accuracy, training loss, and validation loss throughout the training pro- cess. For disease prediction, the system accurately identified bacterial leaf blight with 96.2% confidence, brown spot with 99.8% confidence, and rice blast with 97.4% confidence from validation directory images. Testing accuracy was evaluated using a confusion matrix, calculating precision, recall, F1 score, and overall accuracy based on true positives, false positives, and false negatives. When compared with other pre-trained models, Inception V3 outperformed RestNet-50 (52.67% validation accuracy), DenseNet-201 (62.67%), and VGGNet-19 (60.67%) with its 75% validation accuracy after 30 epochs. The conclusion affirms that deep learning methods enable farmers to predict paddy diseases in early stages, allowing for timely intervention before entire farms become af- fected. The proposed framework consisting of pre-processing, feature extraction, and classification using transfer learning with Inception V3 and CNN algorithms effectively identifies the three types of paddy leaf diseases without causing overfit- ting problems. This image processing approach combined with deep learning provides a cost-effective alternative to manual methods. Future work will focus on extending the dataset through real-time unmanned aerial vehicle technology for sim- plified and accurate data collection, as well as implementing various new hybrid algorithms to further improve accuracy rates.

REFERENCES

[1] Ai, Yong, et al. ”Research on Recognition Model of Crop Diseases and Insect Pests Based on Deep Learning in Harsh Environments.” IEEE Access 8 (2020): 171686-171693. [2]Archana, K. S., and Arun Sahayad- has. ”Automatic rice leaf disease segmentation using image processing techniques.” Int. J. Eng.Technol 7.3.27 (2018): 182-185.

[3]Chen, Junde, et al. ”Detection of rice plant diseases based on deep transfer learning.” Journal of the Science of Food and Agriculture 100.7 (2020): 3246- 3256.

[4] Ferentinos, Konstantinos P. ”Deep learning models for plant disease detection and diagnosis.” Computers and Electronics in Agriculture 145 (2018): 311- 318.

[5]Jiang, Feng, et al. ”Image recognition of four rice leaf diseases based on deep learning and support vector machine.” Com- puters and Electronics in Agriculture 179 (2020): 105824.

[6]Kamrul, Mahadi Hasan, Pritom Paul, and Majidur Rahman. ”Machine vision based rice disease recognition by deep learning.” 2019 22nd International Conference on Computer and Information Technology (ICCIT). IEEE, 2019.

[7]Liang, Wan-jie, et al. “Rice blast disease recognition using a deep convolutional neural network.” Scientific reports 9.1 (2019): 1-10.

[8]Li, Dengshan, and et al. ”A recognition method for rice plant diseases and pests video detection based on deep convolutional neural network.” Sensors 20.3 (2020): 578.

[9]Li, Lili, Shujuan Zhang, and Bin Wang. ”Plant Disease Detection and Classification by Deep Learning—A Review.” IEEE Access 9 (2021): 56683- 56698.

[10] Lu, Yang, et al. “Identification of rice diseases using deep convolutional neural networks.” Neurocomputing 267 (2017): 378-384. [11] O’Shea, Keiron, and Ryan Nash. ”An introduction to convolutional neural networks.” arXiv preprint arXiv:1511.08458 (2015).

[12] Patil, Nilam Sachin. ”Identification of Paddy Leaf Diseases using Evolutionary and Machine Learning Methods.” Turkish Journal of Computer and Mathematics Education (TURCOMAT)12.2 (2021): 1672-1686.

[13] Pandey, N., and C. Sharma. ”Different ap- proaches for the management of brown spot (Helminthosporium oryzae) disease in rice (Oryza sativa) in Nepal.” Int. J. Grad. Res. Rev 5 (2019): 190-193.

[14] Ramesh, S., and D. Vydeki. ”Recognition and classification of paddy leaf diseases using Optimized Deep Neural network with Jaya algorithm.” Information processing in agriculture 7.2 (2020): 249-260.

[15] Saleem, Muhammad Hammad, et al. “Image-based plant disease identification by deep learning meta-architectures.” Plants 9.11 (2020): 1451.

[16] Sethy, Prabira Kumar, et al. “Deep feature based rice leaf disease identification using support vector machine.” Computers and Electronics in Agriculture 175 (2020): 105527.

[17] Sethy, Prabira Kumar, et al. ”Image processing techniques for diagnosing rice plant disease: a survey.” Procedia Computer Science 167 (2020): 516- 530.

[18] Shalu Chopra, Chellapandi, Bincy, and M. Vijayalakshmi. ”Comparison of pre- trained models using transfer learning for detecting plant disease.” 2021 International Conference on Computing, Communication, andIntelligent Systems (ICCCIS). IEEE, 2021.