Leaf Disease Detection System Through Deep Learning Using CNN Model

Mrs.P.Sridevi[1], K.Sravika[2], K.Sreelekha[3], A.Apoorva[4], D.Jyothika[5]
[1,2,3,4,5]*Information Technology, GVPCEW,* Visakhapatnam, India

Abstract: Agriculture can be stated as a sector that significantly affects human existence and economic situation. If resources are not used effectively, crop yield might drop significantly. Leaf diseases are harmful to any type of crop since they might attack the leaves and thus, the plant at different stages. Its development and harvest are significantly impacted by this. To ensure a low loss, it is crucial to keep an eye on the crop's development. A subset of deep learning called convolutional neural networks is extensively used for picture segmentation and classification. The primary goal of the proposed model is to build a solution to identify 15 distinct classes of leaf diseases that extracts the region of interest (ROI) through minimum computational resource usage. Rectified Linear Unit (ReLU) is being used as the activation function to classify the input picture into the appropriate disease(s), and neural network models for automating feature extraction. The obtained accuracy is 94.53% through which it can be deduced that the above said technique is applicable under typical and challenging circumstances.

Keywords: Deep Learning, Convolutional Neural Networks (CNN), ReLU, Region of Interest (ROI), feature extraction, Adam optimizer.

I. INTRODUCTION

Food safety and plant health are strongly coupled. The Food and Agricultural Organization of the United Nations (FAO) claims that pests and diseases pose a danger to food security by reducing global food production by 20–40%. (Food and Agriculture Organization of the United Nation, International Plant Protection Convention, 2017). Different issues may contribute to a certain symptom, and these issues may coexist on the same plant. Even pests and dietary shortages can cause symptoms that resemble those illnesses. On vast farms, it is not possible to repeatedly check the status of each plant throughout the growing season. Recognizing plant diseases becomes extremely important in order to prevent any significant reductions in productivity, performance, and the value of agricultural output. Since manual recognition requires a lot of time and is more likely to be inaccurate, improper treatment can result. The "Leaf Disease Detection using CNN model" system that has been proposed focuses on fifteen classes (12 diseased,3 healthy), including pepper bell bacterial spot, pepper bell healthy, potato healthy, potato early blight, potato late blight, tomato target spot, tomato mosaic virus, tomato yellow leaf curl virus, tomato bacterial spot, tomato late blight, tomato early blight, tomato leaf mold, tomato septoria leaf spot, tomato healthy, tomato spider mites. The convolutional neural network (CNN) model with ReLU as an activation function is presented for the development of a model that is operated on the input picture and modifies the input to categorize the output classes. This system is based on the principles of deep learning techniques.

Section 2 describes the literature survey and section 3 contains the methodology of the proposed system along with model architecture and details of the framework for frontend Section 6 contains implementation details of the proposed system and section 7 shows a graph of accuracy of both training and validation datasets. Section 8 contains execution and results followed by conclusion in section 9.

II. LITERATURE SURVEY

A ton of examination has been done somewhat recently on plant illness identification utilizing profound learning and personal computer vision. Traditional personal computer vision calculations such as haar, hoard, filter, surf, picture division, Support Vector Machines (SVM), K-Nearest Neighbors (KNN), K-implies, and Artificial Neural Networks are among the AI techniques (ANN). Plant disease characterization models based on deep learning include the use of a variety of CNN models, such as AlexNet, VGGNet, and others. As a rule, when the dataset size is insufficient, multiclass order with a large number of classes necessitates careful hyper parameter adjustment to avoid the overfitting problem and by using pre-trained models.

[1]In the paper "Rice Leaf Disease Image Classifications Using KNN Based on GLCM Feature Extraction " by R A Saputra and Suharyanto, The authors studied and proposed the GLCM method as feature extraction for text analysis, with five feature values consisting of contrast, energy, entropy, homogeneity, and correlation. KNN (K-Nearest Neighbor) algorithm is used for the classification of rice leaf disease, by finding the maximum k value from the experiment k value ranging from 1 to 20.

[2]In the paper “Plant Leaf Disease Detection using Machine Learning” by Amrita S.Tulshan and Natasha Raul, the authors used classifiers such as KNN to classify data into more than two classes. In the experiment, a total of seven leaf diseases were detected by the system. Furthermore, this system also detects the disease names as well as the accuracy percentage of the affected area, sensitivity, and elapsed time. Diseases that are discussed in the paper were Down Mildew, Early Blight, Mosaic Virus, Leaf Miner, and White Fly.

[3]In the paper “Plant Leaf Disease Detection and Classification based on CNN with LVQ Algorithm” by Melike Sardogan, Adem Tuncer, and Yunus Ozen. The authors proposed a three-channel CNN model based on RGB colors to detect vegetable leaf diseases. As plant leaf images are complex with their background, and the color information extracted from a single-color component is limited, the feature extraction method might produce lower accuracy results. Using different color components is more promising instead of a single one. In the above paper, a CNN model based on RGB components has been developed and Learning Vector Quantization (LVQ) algorithm was chosen as a classifier due to its topology and adaptive model.

[4]In the paper “Plant Infection Detection Using Image Processing” by Dr. Sridhathan C, Dr. M. Senthil Kumar. The authors have proposed a vision based automated plant detection using K- means clustering algorithm for color segmentation and GLCM for disease classification. This system detects the diseases caused by fungi, bacteria and viruses. The diseases identified by the module in this paper are Anthracnose, Cercospora Leaf Spot and Bacterial Blight with the affected area percentage i.e., it also shows how much of the disease is spread.

[5]In the paper “Banana Leaf Disease Detection Using Deep Learning Approach” by Sahil Gandhi, Sachin Walunjkar, Vedant Choudhary, Parth Joshi. The authors have come up with a ML model that helps in early detection of different types of diseases in banana plants and thus minimizes the loss of yield.

[6]In the paper "Lemon Leaf Disease Detection and Classification using SVM and CNN" by Balambigai Subramanian, Jayashree.S,Kiruthika.S, Miruthula.S,for extracting features the authors used GLCM (Grey Level Co Occurrence Matrix) and classification of leaf disease by using MCSVM (Multi Class Support Vector Machine).This project presents the detection and classification of diseases which are done by using the CNN(Convolution Neural Network) technique as the proposed system And so on.

III. METHODOLOGY

A neural network model for leaf disease classification is developed here. By using input leaf images from the user/farmer, this model will be implemented on the web application for real-time detection of plant leaf disease. The whole proposed process is depicted in fig1.

The following steps are carried out in this process:
1. The first step is the collection of the dataset.
2. Pre-processing and data augmentation of the collected dataset are done using Image data generator API by Keras in Python.
3. Building CNN (Convolutional Neural Network) model for the classification of various plant diseases. The model is trained by selecting random batches from dataset.
4. Developed model will be deployed on the web application where the input image is taken from the user to classify the leaf disease.

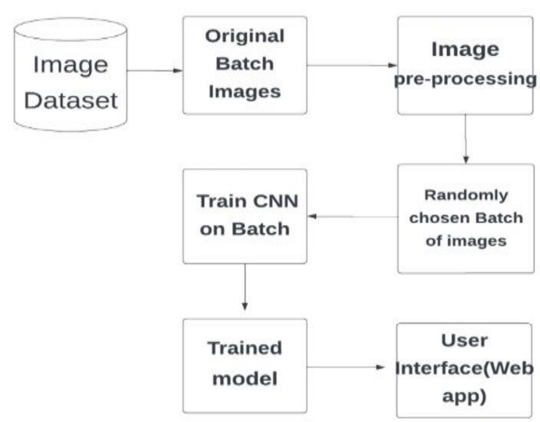

Fig 1: Block diagram of the proposed system

A. Analysis of Dataset

PlantVillage dataset is used, which is openly available from Kaggle.com. The total number of instances is 8072, of which 1,614 (20%) are in the test dataset and 6458(80%) are in the training dataset. The 15 classes included in the dataset are pepper bell bacterial spot, pepper bell healthy, potato healthy, potato early blight, potato late blight, tomato target spot, tomato mosaic virus, tomato yellow leaf curl virus, tomato bacterial spot, tomato late blight, tomato early blight, tomato leaf mold, tomato septoria leaf spot, tomato healthy, tomato spider mites.

B. Data Preprocessing

In this study, conventional methods for data preparation were employed, such as rescaling, resizing, rotation, horizontal flipping, and vertical flipping.

- *Rescaling:*

*tf.keras.layers.Rescaling(scale, offset=0.0, **kwargs).*

This is the pre-processing layer that rescales the image (input values) to a new range. The rescaling is used for both inference and training. Inputs are of the integer or float type, and the layer will output floats by default.

- *Resizing:*

*tf.keras.layers.Resizing(height,width,interpolation="bilinear", crop_to_aspect_ratio=False, **kwargs).*

This pre-processing layer resizes the images to target height and width. (we have chosen 256×256).

A target height and width are applied to an image input by this layer. A 4D (batched) or 3D (unbatched) tensor should be used as the input. Any range of input pixel values, such as [0., 1.] or [0, 255], can be used, as well as integer or floating point data types. The layer will output floats by default.

- *Rotation:*

layers.experimental.preprocessing.RandomRotation(0.2)

This pre-processing layer rotates the input image to a specified angle.

- *Horizontal/vertical flipping:*

layers.experimental.preprocessing.RandomFlip("horizontal_and_vertical").

This pre-processing layer flips both horizontally and vertically the input image to a specified degree of measurement.

C. MODEL BUILDING USING CNN

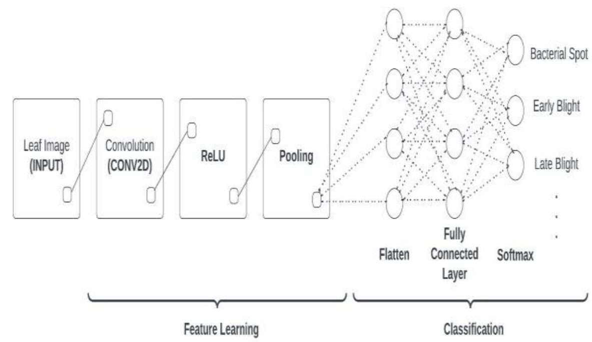

Fig 2: CNN architecture

D. Deep Learning:

Deep learning is an approach to machine learning that tells a computer to filter data across layers. The filtering of information by the human brain is demonstrated via deep learning. The multiple hidden layers found inside the neural network are referred to as "deep" layers. Deep neural networks feature up to 150 hidden layers, as opposed to the 2-3 hidden layers in a standard neural network.

E. *CNN:*

A CNN model has an architecture that is appropriate for processing 2D data, such as images, by combining well-read features with input data and using 2D convolutional layers. CNN eliminates the need for manual feature extraction and classification of images. The CNN model itself pulls features directly from images. The characteristics that are retrieved are well-read when the network is being trained on a small number of picture groups. Convolutional Neural Network (CNN) model comprises the input layer, conv2D layer, pooling layer, fully connected layer, softmax layer, and output layer, which perform the processing of images.

F. *Input Layer:*

Data in the form of images are contained in the input layer. The parameters include the image's height, width, and color information (RGB). Fixed 256 × 256 RGB pictures are taken as inputs.

G. *Conv2D layer:*

A conv2D layer's filter or kernel applies an element-wise multiplication to the 2D input data by "sliding" across it. It will therefore combine the outcomes into a single output pixel. Every region the kernel slides over will experience the identical operation, which changes one 2D feature matrix into another 2D feature matrix.

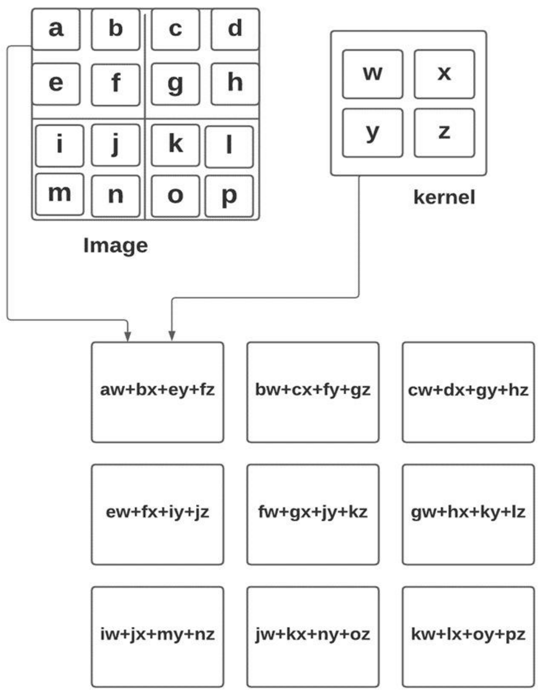

Fig 3: Convolutional operation

H. *Pooling layer:*

The dimensions of the feature maps are decreased by pooling layers. As a result, it lessens the quantity of computation done in the network and the number of parameters to learn. As a result, further operations are carried out on summed-up features rather than precisely positioned features produced by the convolution layer. Here, max pooling is used which takes the maximum value of all the input values.

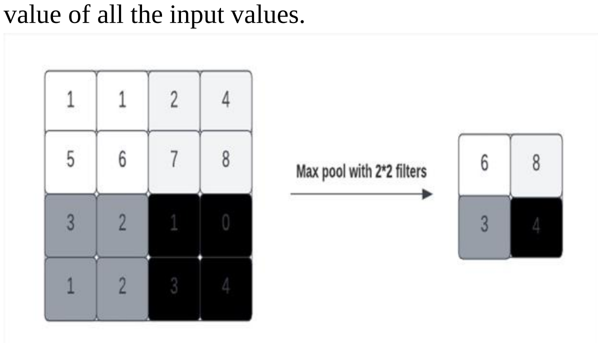

Fig 4: Max Pooling

I. *Fully connected layer:*

A neural network with fully connected layers is one in which each neuron uses a weights matrix to apply a linear transformation to the input vector. As a result, every input of the input vector impacts every output of the output vector, and all layer-to-layer relationships are present.

J. *Softmax Layer:*

The outputs are normalized using the softmax function, which turns weighted sum values into probabilities that add up to one. Each value in the softmax function's output is regarded as the likelihood that a given class would contain that value.

K. *ReLU activation function:*

The rectified linear activation function, often known as ReLU, is a non-linear or piecewise linear function that, if the input is positive, outputs the input directly; if not, it outputs zero.

It is mathematically expressed as $f(x) = \max(0, x)$.

L. *ADAM Optimizer:*

Adaptive moment estimation (ADAM) adjusts the network weights during training, this optimization approach is a further development of stochastic gradient descent(SGD). Adam optimizer modifies the learning rate for each network weight separately, unlike SGD training, which maintains a single learning rate.

M. *System Architecture*:

The proposed system architecture entails data collection from the PlantVillage dataset, processing at several convolutional layers, and finally the classification of plant disease that determines whether the leaf is healthy or not, if not healthy, it classifies the disease type and name. The plant leaf is taken as input from the user into the web application and then the augmentation techniques of resizing, rescaling, horizontal and vertical flipping, and rotation are applied. This input is fed to the trained CNN model and the algorithm recognizes the region of interest and identifies the disease type and name. The result is displayed on the user's dashboard.

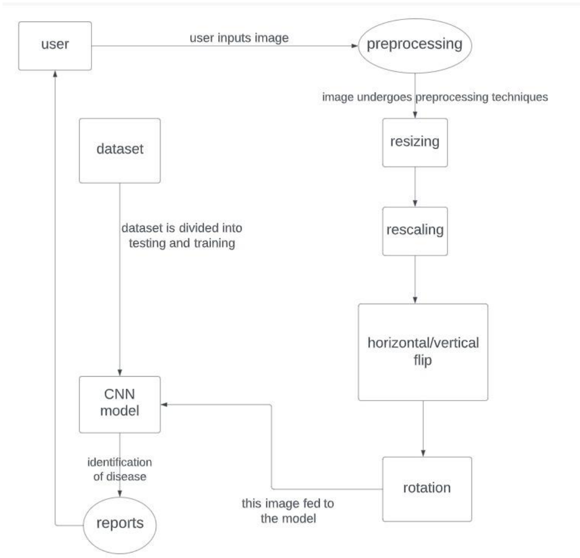

Fig 5: System architecture

N. *Building Frontend by using Django Framework*

Python's Django high-level web framework offers a complete collection of tools and frameworks for quickly and effectively creating online applications. Models are used to define data structures, views are used to manage user requests, and templates are used to produce HTML and other content, all in accordance with the model-view-controller (MVC) architectural paradigm. User authentication, URL routing, database management, and security are just a few of the many built-in features that can be used to handle typical web development tasks. Additionally, it offers an object-relational mapper (ORM) for working with databases, which can greatly ease database operations. Based on the Python web application WSGI standard, Django also includes a lightweight web server for development. Fig 6 shows the login page, which contains a form that allows users to register on the website.

Fig 6: Login Page

O. Metrics:

To classify the leaf disease, Loss and Accuracy are used as metrics.

Loss: Loss is nothing but the squared difference between the expected value and the predicted value. $L = (y_i - \hat{y}_i)^2$.

IV. IMPLEMENTATION DETAILS

A. *Python:*

Python is used for the development of the proposed system. The libraries involved are Keras, TensorFlow, and Numpy. The software platform used to work is Jupyter Notebook.

B. *Jupyter Notebook:*

A server-client program called the Jupyter Notebook App enables editing and executing notebook papers from a web browser. The Jupyter Notebook App may be used locally, without an internet connection, or it can be deployed on a remote server and viewed online.

V. RESULT

A 94.53% accuracy rate was achieved by the proposed model using 60 epochs while training. Fig 7 depicts the visualization of training and validation accuracy. The result of detecting and recognizing a leaf disease is shown in Figure 8.

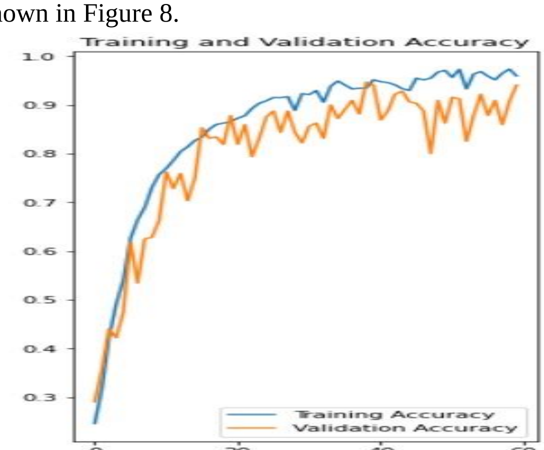

Fig 7*: Accuracy*

VI. EXECUTION

The image in Figure 8 is given as input to the proposed system and it detected that the given leaf has bacterial spot disease.

Fig 8: *Input Image*

Fig 9: Result

VII. CONCLUSION

This research provides a reliable and efficient method for detecting leaf diseases using image processing techniques. For the detection of leaf diseases, Convolutional neural network methods are applied. Farmers cannot individually test each plant in the crop to make sure whether each one is healthy or not and hence the proposed automated system reduces the labor cost and detection time. The farmers can use it to diagnose the disease and then take appropriate corrective action using leaf disease detection using image processing techniques. We will expand our model in subsequent work to identify more leaf diseases.

REFERENCE

[1] Rice Leaf Disease Image Classifications Using KNN Based On GLCMFeature ExtractionTo cite this article: R A Saputra et al 2020 J. Phys.: Conf. Ser. 1641 012080

[2] Plant Leaf Disease Detection using Machine Learning Publisher: IEEE, Amrita S. Tulshan; Nataasha Raul

[3] Plant Leaf Disease Detection and Classification Based on CNN with LVQ Algorithm, Melike Sardogan; Adem Tuncer; Yunus Ozenv

[4] Plant Infection Detection Using Image Processing, Dr. Sridhathan C, Dr. M. Senthil Kumar Associate Professor, Department of ECE, Nalla Malla Reddy Engineering College, Hyderabad, Telangana, India Corresponding Author: Dr. Sridhathan C.

[5] Banana Leaf Disease Detection Using Deep Learning Approach Sahil Gandhi, Sachin Walunjkar, Vedant Choudhary3, Parth Joshi, Computer Engineering Department, AISSMS College of Engineering, Pune, Volume 10 Issue XII Dec 2022.

[6] Lemon Leaf Disease Detection and Classification using SVM and CNN, Balambigai Subramanian, Jayashree.S, Kiruthika.S, Miruthula.S, Volume-8 Issue-4, November 2019.

[7] Paddy crop disease detection using SVM and CNN algorithm, Banda Pavani, Dr.R. Maruthamuthu, Assistant Professor, Department of Computer Applications, Madanapalle Institute of Technology and Science, Angallu, AP, India, Volume: 09 Issue: 08 | Aug 2022.

[8] Leaf Disease Detecting Using CNN Technique, Prof Ramya C N, Naveen G C, Aiswariya Dev S, Sucharitha N, Department of Electronics and Communication Engineering, Atria Institute of Technology, Bengaluru, India Asst Professor, Department of Electronics and Communication Engineering, Atria Institute of Technology, Bengaluru, India.

[9] Automatic Detection of Citrus Fruit and Leaves diseases using Deep Neural Network Model, Asad Khattak, Muhammad, Usamaasghar, College of Technological Innovation, Zayed University, Dubai, United Arab Emirates, Institute of Computing and Information Technology, Gomal University, Dera Ismail Khan (KP) 29220, Information Systems Department, College of Computer and Information Sciences, King Saud University, Riyadh 11362, Saudi Arabia.

[10] Plant disease detection using convolutional neural network, Prakanshu Srivastava, Kritika Mishra, Vibhav Awasthi, Vivek Kumar Sahuand Mr.Pawan Kumar Pal, Information Technology, PSIT College of Engineering, Kanpur Uttar Pradesh, India. Computer Science and Engineering, PSIT College of Engineering, Kanpur Uttar Pradesh, India.