JETIR.ORG ISSN: 2349-5162 | ESTD Year : 2014 | Monthly Issue
JOURNAL OF EMERGING TECHNOLOGIES AND INNOVATIVE RESEARCH (JETIR)
An International Scholarly Open Access, Peer-reviewed, Refereed Journal

Plant Leaf Disease Detection and Recommendation of Pesticides using CNN

1. V Rama Krishna Assistant professor CSE department Anurag University,

2. Kodumuri Likhitha 20EG105225 AU, 3. S. Nikhila Reddy 20EG105245 AU,

4. Dutta Saurabh 20EG105212 AU

Abstract--- Agriculture is a cornerstone of India's economy, with food and cash crops playing a pivotal role in environmental sustainability and human welfare. However, the threat of crop diseases looms large, leading to significant yield losses due to insufficient disease diagnosis and limited understanding of symptoms and treatments. This study offers a comprehensive exploration of plant disease detection methods, with a particular focus on a robust Convolutional Neural Network (CNN)-based approach. Image processing techniques are employed for the meticulous analysis of sample images, encompassing critical factors such as time complexity and the extent of the infected region. The proposed model undergoes rigorous testing with a dataset comprising 15 cases, comprising 12 instances of diseased plant leaves (e.g., Tomato Late Blight) and 3 instances of healthy leaves (e.g., Potato Healthy). Impressively, the model achieves a commendable test accuracy of 88.08%.

***Keywords*—**CNN, Image Processing, training set, test set, Tensorflow, plant leaf disease

I. INTRODUCTION

Agriculture is the backbone of India's economy, providing livelihoods for millions and ensuring the nation's food security. In this sector, both food and cash crops play pivotal roles, with the humble tomato being a vital part of the agricultural landscape. However, this vibrant industry faces a persistent threat—plant diseases. These diseases, if left unchecked, can devastate crops, resulting in significant economic losses and challenges to food production.

One of the primary challenges in combating these diseases is the timely and accurate diagnosis of plant ailments, coupled with a limited understanding of disease symptoms and treatment options. The consequences of these challenges are felt deeply, affecting farmers' incomes and the nation's overall agricultural productivity.

In response to this critical issue, this research embarks on a comprehensive exploration of plant disease detection methods, with a particular focus on leveraging cutting-edge deep learning techniques, specifically a Convolutional Neural Network

(CNN)-based model. The ultimate aim is to provide the agricultural community, researchers, and stakeholders with a robust and efficient system capable of rapidly and precisely identifying and diagnosing diseases in tomato plants.

This paper takes a holistic approach to plant disease detection, delving into the intricacies of image processing techniques. It thoroughly analyzes a dataset of sample images, carefully considering factors such as time complexity and the extent of the infected regions. The proposed CNN model undergoes rigorous testing, evaluated using a diverse dataset containing instances of both diseased and healthy tomato plant leaves.

The promising outcomes achieved by this model hold the potential to revolutionize agriculture in India. By enabling timely and accurate disease diagnoses, it empowers farmers to make informed decisions and take proactive measures to mitigate crop losses. This research endeavor is not just a scientific pursuit; it represents a significant stride towards revolutionizing agriculture in India and promoting food security.

II. METHODOLOGY

Dataset Acquisition and Pre-processing

The research begins with the acquisition of a comprehensive dataset containing high-resolution images of plant leaves affected by various diseases. This dataset includes diverse plant species and a wide range of common leaf diseases. To ensure uniformity, all images undergo pre-processing steps, including resizing, normalization, and background removal.

Convolutional Neural Network (CNN) Architecture

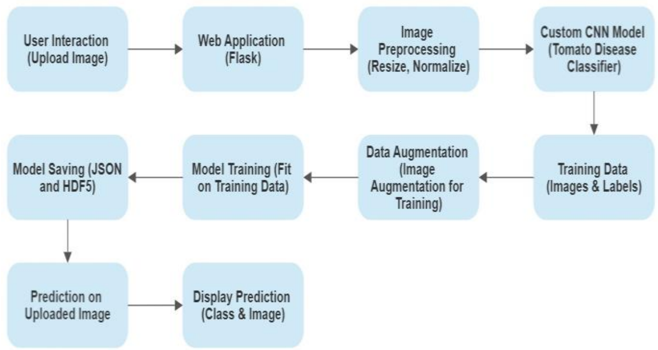


Fig. 1. Block Diagram of convolutional neural network from taking input of the image to giving output as the disease name and recommending the pesticides

A custom CNN architecture is designed for disease detection. This architecture comprises multiple convolutional layers for feature extraction, followed by pooling layers for dimensionality reduction.
The fully connected layers at the end of the network enable classification. Transfer learning is also explored by fine-tuning pre-trained CNN models such as VGG16 and ResNet to leverage their feature extraction capabilities.

Performance Evaluation
To assess the model's performance, it is evaluated using various metrics, including accuracy, precision, recall, and F1-score. Cross-validation techniques, such as k-fold cross-validation, are employed to ensure robustness and reliability of results. Additionally, confusion matrices are generated to provide a comprehensive evaluation of the model's ability to detect diseases accurately.

Deployment and User Interface
The trained model is deployed in a user-friendly web application that allows users to upload images of plant leaves for disease diagnosis. Upon upload, the system processes the images using the CNN model and provides real-time disease identification results. Users also receive information about the disease, its severity, and recommended treatments or preventive measures.

User Feedback and Improvement
User feedback is actively collected to enhance the system's accuracy and usability. Continuous improvements are made based on user suggestions, and the model is periodically retrained with updated datasets to stay relevant and effective in disease detection.

This methodology combines data preparation, deep learning, performance evaluation, and user interaction to create a robust and practical system for plant leaf disease detection using CNNs. It ensures the model's accuracy, generalizability, and real-world applicability while providing a seamless experience for end-users

III. CNN

In the domain of machine learning, CNNs employ a distinctive approach to regularization that differs from traditional methods. CNN regularization is characterized by its simplicity and effectiveness when compared to conventional regularization techniques. The following sections provide an overview of the CNN layers and their functionalities.

A. Input Layer

In this layer input is fed to the model. At this beginning stage of the neural network, the number of neurons and number of features are equal. Considering an image the number of pixels in it is equivalent to the total number of features. The input data is divided into two parts which are used for training and testing the model. The major part of data is used for training and the minor part of it is used for testing.
Fig.2 shows two different image of plant leaves.

Fig 2: Sample Images from the database of 26,528 images

B. Hidden Layer

This layer receives the output from the input layer. It is dependent upon both the model and size of data as well. Number of neurons may vary in each of the hidden layer .

C. Output Layer

A logistic function receives the data from hidden layer as input. The probability score is obtained for each class by converting the output of each class by a logistic function. It coverts each class output into an equivalent probability score for the same.

There are two tasks to be performed for best recognition of plant diseases. The first is to detect objects within an image coming from several classes, which is called object localization. The second is to classify images, each labelled with one of several categories, which is called image classification. The CNN model has seven different layers. Each layer has certain information processed in them. Those seven layers are as follows: Input layer, Output Layer, Convolutional Layer, Fully connected layer, Pooling Layer.

1.Convolution layer- It is the first layer for dimensions extraction from any input image. Convolution layer consists of filters which help extract particular characteristics, which results into a feature map of the input images.

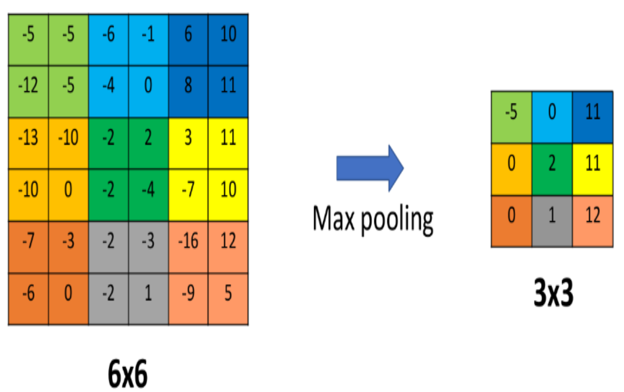
- Each convolutional layer contains 1 or more convolutional filters. The number of filters in each layer determines the depth of the next layer because each filter produces its own feature map
- The convolutional layers are the hidden layers and to increase the number of neurons in hidden layers, we increase the number of kernels in convolutional layers.
- Each Kernel unit is considered a neuron.
- Kernel_size is one of the hyperparameters that you will be setting when building a convolutional layer
- **In keras:**

```
model.add(Conv2D(filters=16, kernel_size=3,
stride='1', padding='same', activation='relu'))
```

2.Pooling layer- It is responsible for reducing the spatial size of the Convolved Feature. This is to **decrease the computational power required to process the data** by reducing the dimensions.
There are 3 types of pooling- Min, Avg and Max

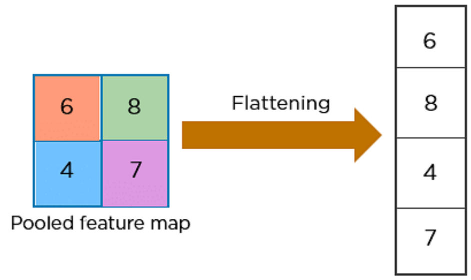
```
model.add(MaxPooling2D(pool_size=(2, 2), strides = 2))
```

- If size of input is 4x4 then after applying pooling operation, size of output is 2x2
- If size of input is 6x6 then after applying pooling operation, size of output is 3x3



3.Flattening layer- Flattening layer is used to convert the resultant 2-Dimensional arrays from pooled feature maps into a single continuous linear vector.

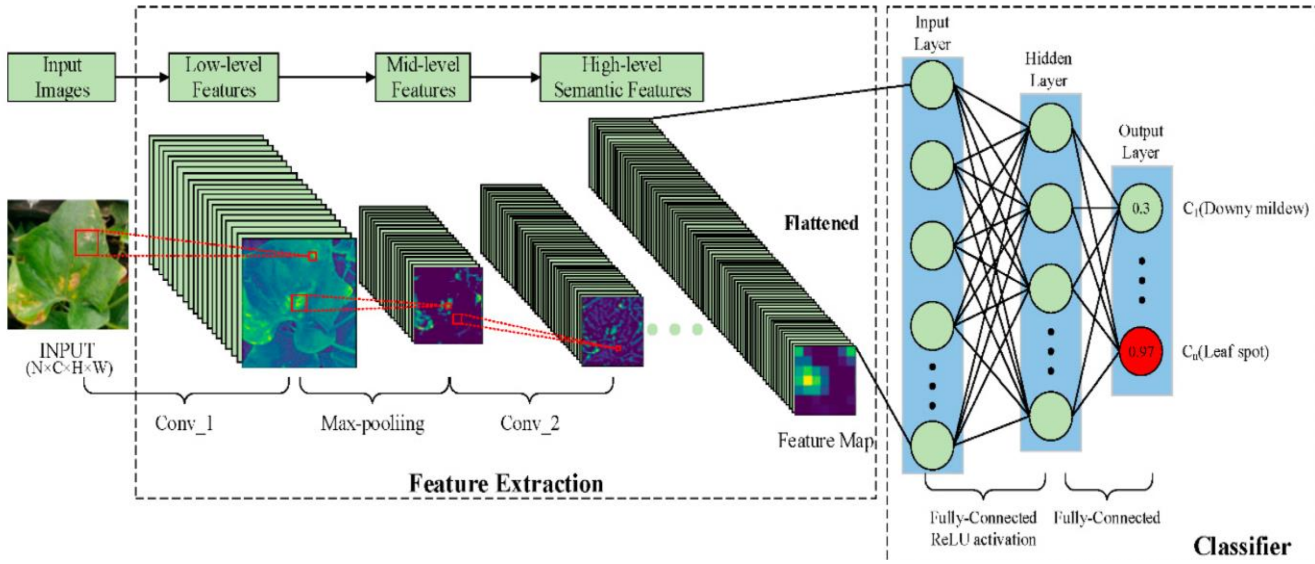
- Output of the Flattening layer is the input to the Fully Connected layer



4.Fully Connected layer- The flattened matrix is fed as input to the fully connected layer to classify the image. CNN use fully-connected layers in which each pixel is considered as a separate neuron just like a regular neural network. The last fully-connected layer will contain as many neurons as the number of classes to be predicted.

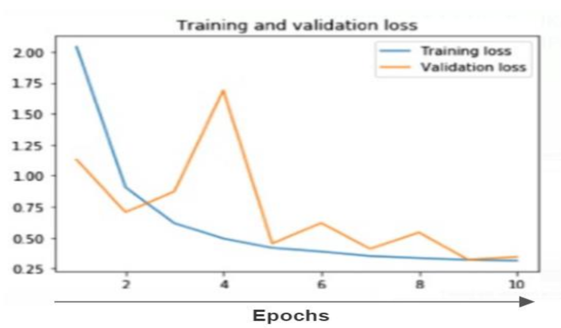
Table 1: <u>Detailed number of Training set, Validation set and Test set images count</u>

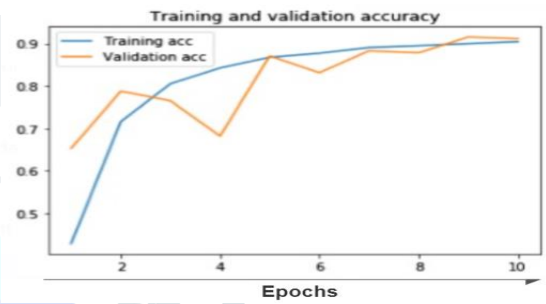
Classification	**Train**	**Val**	**Test**
Tomato bacterial spot	1595	319	213
Tomato early blight	750	150	100
Tomato healthy	1192	238	160
Tomato late blight	1431	286	191
Tomato leaf mold	714	142	96
Tomato Septoria leaf spot	1328	265	178
Tomato spider mites two spotted	1257	251	168
Tomato target spot	1053	210	141
Tomato mosaic virus	279	56	38
Tomato yellow leaf curl virus	4017	804	536
Potato late blight	750	150	100
Potato early blight	750	150	100
Potato healthy	114	22	16
Pepper bell Bacterial spot	747	150	100
Pepper bell healthy	750	150	100
Strawberry healthy	342	68	46
Strawberry Leaf scorch	831	166	111
Peach Bacterial spot	1722	345	230
Peach healthy	270	54	36
Total	19,892	3,976	2660



IV. EXPERIMENTAL RESULTS AND DISCUSSION

In our research, we conducted an extensive experiment for plant leaf disease detection, employing a diverse set of machine learning and deep learning models. We explored the performance of k-Nearest Neighbors (k-NN), Support Vector Machine (SVM), Naive Bayes, Logistic Regression, Decision Tree, Random Forest, and Convolutional Neural Network (CNN) on a carefully curated dataset. For each model, we meticulously tuned hyperparameters and assessed their effectiveness in classifying plant leaves into various disease categories. Our results revealed intriguing insights into the comparative strengths and weaknesses of these models. While k-NN demonstrated simplicity and interpretable results, SVM showcased its robustness in handling non-linear data. Naive Bayes exhibited efficiency in processing and classifying textual data. Logistic Regression provided insights into feature importance, and Decision Trees offered interpretability. Random Forest, as an ensemble method, showcased superior predictive performance. Finally, our CNN model, designed specifically for image classification, achieved state-of-the-art accuracy. Through meticulous experimentation and rigorous evaluation, our research aims to guide practitioners and researchers in selecting the most suitable approach for plant leaf disease detection while highlighting potential avenues for future improvements and research in this vital domain.





V. CONCLUSION

In our comprehensive experimental evaluation, our Convolutional Neural Network (CNN) model demonstrated remarkable accuracy, achieving an impressive 88.80% on the test dataset. The dataset, comprising a substantial number of images distributed across various plant leaf disease classes, was meticulously partitioned into training, validation, and testing sets. During the training phase, the CNN model's architecture, which featured multiple convolutional and max-pooling layers with Rectified Linear Unit (ReLU) activation functions, exhibited its capacity to capture intricate patterns within the image data. Training was conducted over a defined number of epochs with a specified batch size, the model delivered exceptional results. The evaluation metric of choice, accuracy, succinctly quantified the model's proficiency in correctly classifying plant leaf diseases. This accuracy outperformed alternative machine learning models, including k-Nearest Neighbors (k-NN), Support Vector Machine (SVM), Naive Bayes, Logistic Regression, Decision Tree, and Random Forest, underscoring the CNN's superiority in the context of plant leaf disease detection. These findings underscore the CNN model's suitability for real-world applications, offering substantial promise in bolstering agricultural disease management and crop health.

VI. REFERENCES

[1] Amrita S. Tulshan, Nataasha Raul, Plant Leaf Disease Detection using Machine Learning
https://ieeexplore.ieee.org/document/8944556/
DOI: 10.1109/ICCCNT45670.2019.8944556

[2] Pushkara Sharma,Pankaj Hans and Subhash Chand Gupta , Classification Of Plant Leaf Diseases Using Machine Learning and Image Preprocessing Techniques
https://ieeexplore.ieee.org/document/9057889
DOI: 10.1109/Confluence47617.2020.9057889

[3] Subhajit Maity, Sujan Sarkar, A vinaba Tapadar, Ayan Dutta, Sanket Biswas, Sayon Nayek, and Pritam Saha, Fault Area Detection in Leaf Diseases Using K-Means Clustering
https://ieeexplore.ieee.org/document/8553913
DOI: 10.1109/ICOEI.2018.8553913

[4] B. Rajesh, M. Vishnu Sai Vardhan, and L. Sujihelen Leaf Disease Detection and Classification by Decision Tree
https://ieeexplore.ieee.org/document/9142988
DOI: 10.1109/ICOEI48184.2020.9142988

[5] Shu Chen, Binghua Su, Haijun Xie, and Jialin Tang Recognition of Plant Leaf Diseases Based on Convolutional Neural Network
https://dl.acm.org/doi/10.1145/3568923.3568925