Leaf Disease detection using Deep Learning

Suneetha Jungili [1] , E .Sateesh [2]

Student[1], Assistant Professor[2]

Amrita Sai Institute of Science and Technology Paritala-521180

Autonomous NAAC with A Grade, Andhra Pradesh, India.

Abstract - Agricultural productivity plays a vital role in the Indian economy, making the contribution of food crops and cash crops crucial for both the environment and human well-being. Plant diseases can significantly impact the quality and production of agricultural products. Inadequate disease diagnosis and the lack of knowledge about symptoms and treatments result in the death of many plants. The field of automatic plant disease detection is gaining prominence, offering advantages in monitoring large crop fields and identifying disease symptoms on leaves. This paper focuses on the detection of plant diseases, aiming to reduce crop losses and enhance production efficiency. Deep Learning is used for detection of plant diseases and Convolutional Neural Network for classifying Leaf images into 39 Different Categories. The model is trained on a dataset consisting of cases. To expand its size, we employed six distinct augmentation techniques. The achieved test accuracy is 98.9%. Various performance metrics are derived for evaluating the model's performance.

I. INTRODUCTION

India, as of April 2020, has a population of around 1.38 billion people. The number of farmers in the country is estimated to be approximately 95.8 million. It is important to note that the agricultural sector contributes to about 18% of India's GDP. Hence, it can be inferred that a revolution in agriculture would greatly benefit the country. This revolution would not only improve the conditions of local farmers but also create employment opportunities and expand the agricultural sector [6]. In India, there has been significant progress in research and development of pesticides, fungicides, and herbicides. However, every year, natural factors lead to the occurrence of known diseases in crops, resulting in the loss of large quantities of produce. Timely detection of plant diseases is crucial to overcome the economic challenges faced by Indian farmers [2][3].Nowadays, technology has positively transformed lives, making almost everything easily accessible through the internet. With a regular camera, it is possible to capture photos of affected plant parts and upload them to a system that can detect specific diseases and provide precise treatment, including the use of pesticides if necessary [1][4]. Many plants are susceptible to various fungal and bacterial diseases, which are influenced by factors such as population growth and climate conditions. Close monitoring of leaves is essential for disease detection [3][5]. Several techniques have been reported by researchers for the detection and monitoring of plant diseases. These diseases can be caused by fungi or fungal-like organisms, while others are caused by viral and bacterial organisms, posing significant threats to food and feed crops [2][6]. Some diseases can be spread from one plant to another, emphasizing the need for timely identification and intervention. Identifying the diseases in plants at an early stage is a challenging task. Diseases can be discerned through an examination of the physical state of leaves, stems, or fruit [4][5].The era of technology and automation. An automated system that can detect plant diseases would be far more effective. Numerous studies have been conducted to fulfil this purpose, with many utilizing traditional machine learning approaches. This paper aims to develop an automated system by utilizing the deep learning techniques to detect the diseases. Deep learning is a sub part of machine learning that offers advantages such as not requiring domain expertise and eliminating the need for feature engineering, instead of conventional machine learning approaches. Our system, utilizes images of plant leaves to detect diseases in the plant leaves. The leaf disease detector is a computer vision based automated diagnostic system that uses the deep learning techniques to accurately identifying both healthy and diseased plants, as well as the specific type of disease. To achieve this, a deep learning network such as Convolutional Neural Network (CNN) can be utilized. CNNs are particularly suitable for extracting features from images and are widely regarded as the most effective for extraction of visual features. The model based on CNN is used for training to detect diseases in plants by using a large dataset of different images consisting of both healthy and diseased plants. The trained CNN model is used to detect diseases in plants by analyzing images of plant leaves.

II. METHODOLOGY

Deep learning is an effective machine learning technique that has overcome the challenges of traditional machine learning, specifically the need for feature engineering. Thanks to deep learning, there is no longer a requirement for domain expertise. The foundation of deep learning lies in Artificial Neural Networks (ANNs), which are mathematical constructs created to imitate the human brain functioning by connecting neurons and synapses. TensorFlow, a widely used library, is instrumental in implementing neural networks. It offers a comprehensive suite of libraries specifically designed for artificial neural networks. Leveraging Tensorflow, one can successfully accomplish classification tasks on both text and image data.

(A) Convolutional Neural Network (CNN):

Convolutional Neural Networks (CNNs) play a crucial role in identifying diseases in plant leaves. CNNs are an advanced version of Artificial Neural Networks (ANNs) and yield superior outcomes when working with images. This is because images exhibit recurring patterns of specific features. Two key functions of CNNs are convolution and pooling. Convolution is employed to detect edge patterns in an uploaded image, while pooling reduces the image's size. For the problem at hand, the following CNN architectures were utilized: (a) Simple CNN(b) Batch Gradient Descent. The training of these models was conducted using Jupyter notebook in conjunction with the Keras API of TensorFlow. Keras serves as a high-level API for TensorFlow, simplifying the process of constructing and training deep learning models. However, this approach requires a significant number of human resources, which is not very efficient in

(B) Dataset Description:

In this research paper an enhanced version of the Plant Village dataset was employed, which consisted of thousands of images obtained from the original dataset comprising 70k images approximately. The augmented dataset encompasses 39 classes of plant disease pairs and it is split into training and validation sets Dataset encompasses a wide variety of images depicting both healthy and diseased plant leaves. The dataset plays an important role in advancing the field of agricultural technology and disease management in various types of crops. Each image within the dataset has a resolution of 256x256 pixels.

Apple_scab	Apple_black_rot
Apple_cedar_apple_rust	Apple_healthy
Background_without_leaves	Blueberry_healthy
Cherry_powdery_mildew	Cherry_healthy
Corn_gray_leaf_spot	Corn_common_rust
Corn_northern_leaf_blight	Corn_healthy
Grape_black_rot	Grape_black_measles
Grape_leaf_blight	Grape_healthy
Orange_haunglongbing	Peach_bacterial_spot
Peach_healthy	Pepper_bacterial_spot
Pepper_healthy	Potato_early_blight
Potato_healthy	Potato_late_blight
Raspberry_healthy	Soybean_healthy
Squash_powdery_mildew	Strawberry_healthy
Strawberry_leaf_scorch	Tomato_bacterial_spot
Tomato_early_blight	Tomato_healthy
Tomato_late_blight	Tomato_leaf_mold
Tomato_septoria_leaf_spot	Tomato_spider_mites_two-spotted_spider_mite
Tomato_target_spot	Tomato_mosaic_virus
Tomato_yellow_leaf_curl_virus	

Figure 1. Sample images of Plant Village dataset

(C)Image processing and Data Augmentation:

The dataset utilized in our study has already undergone augmentation. Therefore, we solely applied the rescaling parameters to normalize the images, and the data splitparameter of the ImageDataGenerator was used to divide the data in to a 80:20 ratio, where 80percent of the data was allocated for training and 20 percent for validation. No additional transformations were applied to the data.

(D) Model Description:

Initially, the dataset undergoes preprocessing in the form of augmentation to enhance its size and improve accuracy. Subsequently, the images are resized to a dimension of 256x256 pixels. Following this, a convolutional neural network (CNN) model is constructed, comprising of several pooling and convolution layers, as well as a dense layer for detection. The architecture of the model includes five convolutional layers that utilize a 3x3 filter size, accompanied by five MaxPooling 2D layers which use a 2x2 filter size for down sampling. Moreover, Batch Normalization is integrated within the model. This technique not only standardizes the input data but also extends its application to various hidden layers, ensuring consistent scaling within the network. Finally, the model is trained on the Plant Village dataset.

III. PROPOSED SYSTEM

The motivation for developing a new system, such as one based on deep learning CNN models, is to address these limitations by providing a more accurate, efficient, and scalable solution for leaf disease detection. The deep learning approach leverages advances in computer vision and machine learning to automate the process, reduce subjectivity, and enable early disease detection.

Data Collection: Gathering an extensive dataset of images containing healthy and diseased plant leaves. These images should cover various types of plants and diseases to ensure the model's generalization.

Data Preprocessing: Clean and preprocessing the dataset by resizing images to a uniform size, normalizing pixel values, and augmenting the data with techniques like rotation, flipping, and brightness adjustments. Data augmentation contributes to model’s performance, robustness by exposing it to different variations of the same image.

Model Architecture (CNN): Design a CNN architecture suitable for image classification. CNNs are particularly effective for image-related tasks due to their ability to capture spatial features through convolutional and pooling layers.

Training: Partitioning the Dataset into training, validation, and testing data sets. Train the CNN model on the training data using techniques like stochastic gradient descent (SGD) or Adam optimizer.

Model Evaluation: Evaluating the trained CNN model on the testing dataset to measure its performance. Metrics such as accuracy, matrix can be used to assess how well the model is identifying healthy and diseased leaves.

Deployment: Once the model demonstrates satisfactory performance, deploy it to a production environment. This could involve integrating it into a mobile app, web application, or even a drone-based system for real-time disease detection in the field.

Continuous Improvement: Collecting user feedback and monitoring the model's performance in real-world scenarios. Periodically retrain the model with new data to adapt to changing disease patterns and maintain its accuracy.

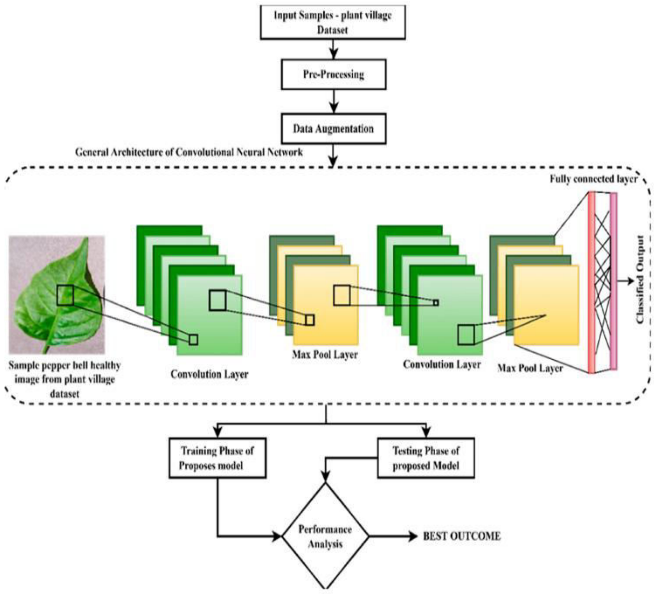

Figure 2. Block diagram of proposed system

IV. DEEP LEARNING USING CNN

CNNs can have multiple convolutional layers, each capturing increasingly complex features. The architecture can also include techniques like dropout (to prevent overfitting), batch normalization (to stabilize training), and more.

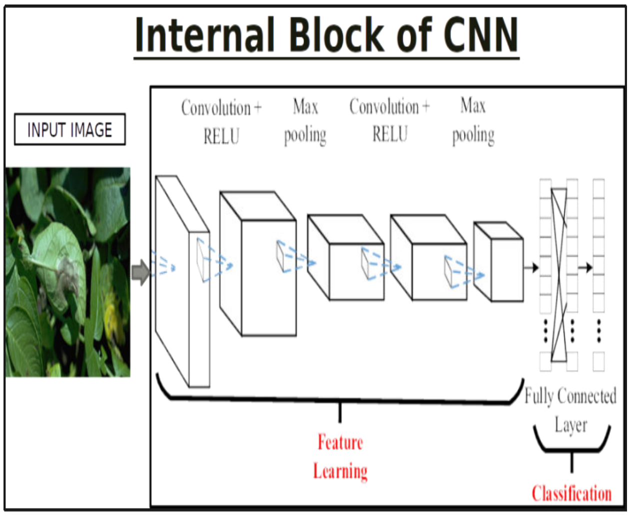

Figure 3. Internal Block of CNN

Convolutional Layer: This is the core building block of a CNN. It involves applying a set of learnable filters (also known as kernels) to the input image. These filters slideover the input data and perform element-wise multiplication followed by summation, creating a feature map that highlights specific features like edges, corners, textures, etc. Multiple filters are used to capture different features.

Activation Function: After the convolution operation, an activation function (commonly ReLU - Rectified Linear Activation) should be applied element-wise for introducing non-linearity to the model. This enables the network to capture complex relationships in the data.

Pooling Layer: Pooling layers (often max-pooling) reduce the spatial dimensions of the feature maps, reducing the computational load and making the network more robust to variations in the input. Pooling essentially takes the highest value in a localised section of the feature map and subsamples it.
Fully Connected Layer: After several convolution and pooling layers, the final feature maps are flattened and connected to a traditional neural network with fully connected layers. These layers make predictions based on the high-level features learned from previous layers.
Output Layer: The final fully connected layer ends with an output layer. The choice of activation function for the output layer depends on the problem being solved. For example, for binary classification, you might use a sigmoid function, and for multi-class classification, you'd use a SoftMax function.

V. MODEL TRAINING AND PERFORMANCE

This section focuses on training the model. It includes data preparation, model architecture, optimizer selection, batching, validation, epochs, testing, and saving models [1][2][5].
Model Performance: The overall accuracy of the model is evaluated using the Adam optimizer [2][4]. The framework automatically detects and preprocesses the given leaf image for prediction [3][5]. The model generates 15 probability values corresponding to 15 different labels, and the label with the highest probability score is considered the predicted disease or result for that particular image [1][4]. The model is trained on 52,595,399 parameters, demonstrating a high-capacity architecture tailored for complex classification tasks [5][6].
Flask:
In this paper Flask frame work is used for user friendly applications. Flask is a python-based web frame work for building the web applications. It is very flexible for user end applications, for building the Restful APIs and many web-based services in IOT for monitoring and controlling.

Layer (type)	Output Shape	Param #
Conv2d-1	[-1, 32, 224, 224]	896
ReLU-2	[-1, 32, 224, 224]	0
BatchNorm2d-3	[-1, 32, 224, 224]	64
Conv2d-4	[-1, 32, 224, 224]	9,248
ReLU-5	[-1, 32, 224, 224]	0
BatchNorm2d-6	[-1, 32, 224, 224]	64
MaxPool2d-7	[-1, 32, 112, 112]	0
Conv2d-8	[-1, 64, 112, 112]	18,496
ReLU-9	[-1, 64, 112, 112]	0
BatchNorm2d-10	[-1, 64, 112, 112]	128
Conv2d-11	[-1, 64, 112, 112]	36,928
ReLU-12	[-1, 64, 112, 112]	0
BatchNorm2d-13	[-1, 64, 112, 112]	128
MaxPool2d-14	[-1, 64, 56, 56]	0
Conv2d-15	[-1, 128, 56, 56]	73,856
ReLU-16	[-1, 128, 56, 56]	0
BatchNorm2d-17	[-1, 128, 56, 56]	256
Conv2d-18	[-1, 128, 56, 56]	147,584
ReLU-19	[-1, 128, 56, 56]	0
BatchNorm2d-20	[-1, 128, 56, 56]	256
MaxPool2d-21	[-1, 128, 28, 28]	0
Conv2d-22	[-1, 256, 28, 28]	295,168
ReLU-23	[-1, 256, 28, 28]	0
BatchNorm2d-24	[-1, 256, 28, 28]	512
Conv2d-25	[-1, 256, 28, 28]	590,080
ReLU-26	[-1, 256, 28, 28]	0
BatchNorm2d-27	[-1, 256, 28, 28]	512
MaxPool2d-28	[-1, 256, 14, 14]	0
Dropout-29	[-1, 50176]	0
Linear-30	[-1, 1024]	51,381,248
ReLU-31	[-1, 1024]	0
Dropout-32	[-1, 1024]	0
Linear-33	[-1, 39]	39,975

Total params: 52,595,399
Trainable params: 52,595,399
Non-trainable params: 0

VI. RESULT AND CONCLUSION

In conclusion, the application of deep learning CNN models in leaf disease detection represents a significant advancement in agricultural technology with far-reaching benefits [1][2][3]. This innovative approach harnesses the power of machine learning to revolutionize the way we manage and protect crops. Throughout the analysis of leaf images, deep learning CNN models have demonstrated their capability to accurately and efficiently identify various plant diseases, paving the way for a more sustainable and productive agricultural future [4][5]. By harnessing the capabilities of artificial intelligence, we are not only enhancing our ability to detect and mitigate plant diseases but also contributing to more sustainable farming practices, increased food security, and a brighter future for agriculture [3][6]. As technology and research continue to advance, the future holds promise for even greater accuracy, real-time monitoring, and multidisciplinary collaborations that can further refine disease detection and management. The used CNN model achieves a highest test accuracy of 98.9% [1].

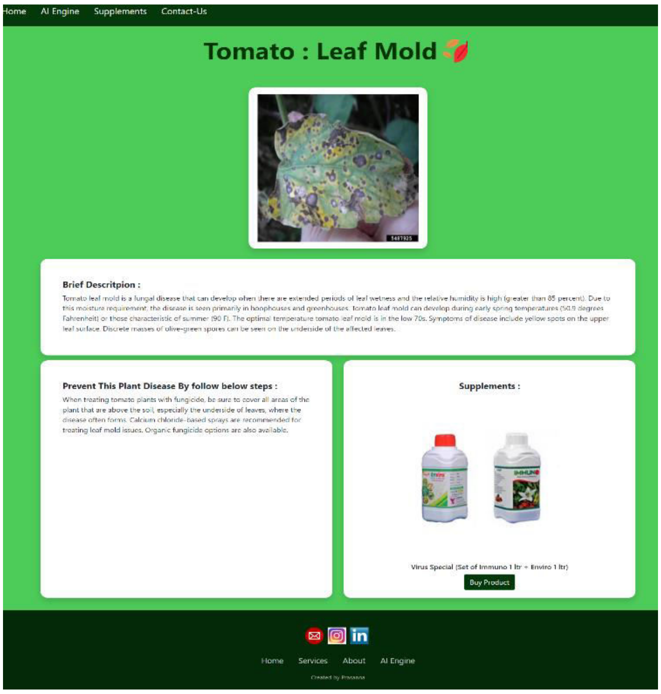

Figure 4:Output of leaf disease detection

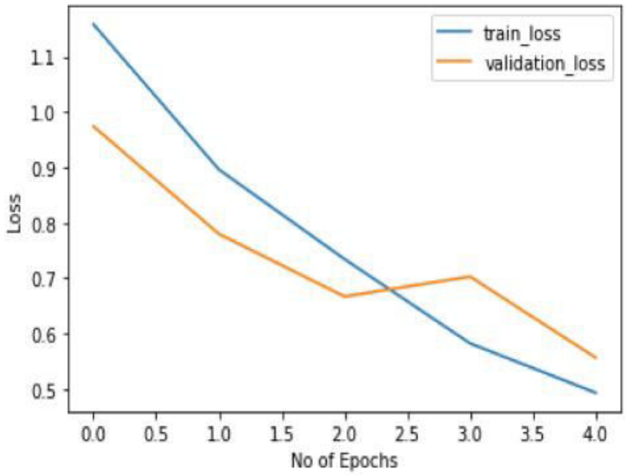

Figure 5:Training and validation loss

VII. FUTURE SCOPE

The future scope of leaf disease detection using deep learning CNN models is quite promising, as advancements in technology and research continue to shape the field of agriculture and plant disease management. Mobile apps can become powerful tools for farmers to easily capture images of leaves, submit them to a cloud-based CNN model for analysis, and receive quick disease diagnosis and recommended actions. As CNN models continue to evolve, their accuracy in disease detection is likely to improve further. Fine-tuning of architectures and utilization of larger and more diverse datasets will contribute to even more precise identification of diseases. Integrating CNN models with IoT (Internet of Things) devices, drones, and remote sensing technologies will enable real-time monitoring of plant health. This can lead to immediate interventions and preventive measures. Deep learning models could potentially provide not only disease diagnosis but also suggest suitable treatments or interventions based on the detected disease and its severity. Aggregating data from various sources can lead to the creation of global disease databases that provide insights into disease prevalence, distribution, and trends on a larger scale.

VIII. REFERENCES

1. Ebrahim Hirani, Varun Magotra, Jainam Jain, Pramod Bide. "Plant Disease Detection Using Deep Learning", 2021 6th International Conference for Convergence in Technology (I2CT), 2021
2. M. A. Hossain, M. Hasan, M. R. Islam, M. S. Islam, M. A. A. Dewan and K. Murase, "A Deep Learning Approach for Detection of Plant Leaf Diseases."
3. Mohanty, S.P., Hughes, D. P.& Salathe, M. in (2016). Using deep learning for image-based plant disease detection.
4. Sladojevic, Arsenovic.M., Anderla.A, CulibrkD., & Stefanovic D. (2016). Using the Deep neural networks based on recognition of plant diseases byleafimage classification. Computational Intelligence and Neuroscience.
5. Teenu Sahasra M, Sai Kumari S, Sai Meghana S, P. Rama Devi (2021). Leaf Disease Detection Using Deep Learning.
6. RinaBora, DeepaParasar, ShrikantCharhate (2021). Plant Leaf Disease Detection using Deep Learning: A Review.