Journal of Innovative Image Processing (ISSN: 2582-4252)
www.irojournals.com/iroiip/

Development and Analysis of CNN based Disease Detection in Cotton Plants

Dr. S. Suriya[1], N. Navina[2]

[1]Associate Professor, Department of Computer Science and Engineering, PSG College of Technology, Coimbatore, Tamilnadu, India.

[2]PG Scholar, Department of Computer Science and Engineering, PSG College of Technology, Coimbatore, Tamilnadu, India.

E-mail: [1]suriyas84@gmail.com, [1]ss.cse@psgtech.ac.in, [2]22mz04@psgtech.ac.in

Abstract

Plant diseases occur due to some organisms like bacteria, viruses and fungi, and has been a problem in agriculture around the world for centuries. Cotton is one of the most highly produced crop in India. Cotton crop help farmers to make good income. The main disadvantage of cotton crop is that it is highly prone to diseases. Early detection and diagnosis of cotton disease is a solution to this problem. Therefore, this research focuses on implementing and evaluating a Machine Learning Algorithm (CNN model) for the analysis and detection of cotton plant diseases. The dataset is pre-processed, the RGB images are converted into grayscale images and the images are resized into a fixed dimension to feed them into the CNN model. The model architecture consists of multiple convolutional layers followed by max-pooling and dense layers. The proposed method significantly contributes to the detection and management of cotton diseases, leading to increased crop yield and economic benefits for cotton farmers.

Keywords: Machine learning, Convolution Neural Network, Convolution Layer, Max Pooling Layer, ReLu, Flattening Layer

1. Introduction

Machine learning (ML) is a field of Artificial Intelligence (AI) that empower a computer framework to learn and improve from experience without being specially designed. The goal of machine learning is to build algorithms and models that can be used to recognize patterns and make predictions based on input. For the machine learning model to learn and

DOI: https://doi.org/10.36548/jiip.2023.2.006
Received: 22.04.2023, received in revised form: 22.05.2023, accepted: 06.06.2023, published: 16.06.2023

evolve, it must be trained on the data of written examples. This dataset is used to teach the model. The more data the model is trained on, the better it becomes at recognizing patterns and make accurate predictions. There are numerous sort of machine learning algorithms, including:

1. Supervised Learning: This model is trained on labeled data, which means data were already labeled as the right answer. The model trains to identify patterns and make predictions based on this labeled data.

2. Unsupervised Learning: This model is trained on unlabeled data, which means the data is not populated with the correct answers. The model learns to recognize patterns and group similar data together.

3. Reinforcement Learning: In reinforcement learning, the model learns by interacting with the environment and receiving feedback through reward or punishment. The model learns to make decisions once a machine learning model is trained. This process is called as inference. Machine learning applications play a major role in many areas, including image recognition, recommendation systems, and predictive modeling.

Convolutional Neural Networks (CNNs) is a kind of machine learning algorithm widely used for image and video analysis. CNNs are based on the structure and function of the human visual system, which can learn and recognize complex visual patterns. Layers such as, convolutional layers, pooling, fully connected layers are the most important layers found in CNN. Input data is passed through each layer sequentially, and input for the next layer is taken from the output of the previous layer.

The first layer in a CNN is typically a Convolutional layer. Features such as edges, textures, and shapes are extracted from the input image by applying a set of filters. Each filter is a small matrix of weights that is learnt during training. The final output of the convolutional layer represents different features present in the input image. After the convolutional layer, a pooling layer is often added, which minimize the spatial size of the feature maps by using a pooling function such as max pooling or average pooling. This helps to minimize the dimensionality of the input data and makes it easier to process in subsequent layers.

The outcome of the pooling layer is then passed through many fully connected layers that execute a non-linear transformation of the data and from the output values that represent the predicted class probabilities. During training, the weights of the CNN are learnt from the minimizing function, which measures the difference between the predicted output and the actual output. This is done using back-propagation, which calculates the gradient of the lossfunction with respect to each weight in the network and updates the weights accordingly. CNNs have proven useful for various tasks, such as object detection, image classification, and face recognition. They are also used for tasks relating to the processing of natural languages such as text classification and sentiment analysis.

Cotton is one of the most highly produced crop in India. Cotton crop help farmers to make good income. Early detection and diagnosis of cotton disease detection and identifying is a very difficult task for the farmers. If the infection or disease on the crops was not identified by the farmers at the initial level, then it will be harmful to the crops as well as for farmers. The main purpose of farming is to avoid such losses, and conventional methods have been used to identify the diseases. However, in order to avoid such losses caused by these diseases, it is very important to diagnose plant diseases early and accurately.

This research deals with the development and analysis of the CNN model for the detection and classification of cotton plant diseases. In order to extract features from a cotton plant image, this proposed study demonstrates the CNN model with convolution layer and max-pooling layer. An image is to be submitted by the user and a digital colour image is obtained from that diseased leaf, which can be used for the prediction of cotton leaf disease using CNN.

2. Literature Review

[1] S.G. Ethiopian coffee plant diseases recognition based on imaging and machine learning techniques:

Coffee is the main crop in Ethiopia, a country is known for producing high-quality Arabica coffee beans. However, coffee plant diseases can have a significant impact on crop yield and quality. The routine disease identification and diagnosis process can be time-consuming and require expertise. Therefore, faster and more accurate diagnosis is needed. In recent years, there has been growing interest in the using graphics and machine learning techniques for the recognition of coffee plant diseases. These processes involve taking images of coffee plants and leaves, then using machine learning algorithms to analyze the images and identify signs of disease. Many studies have been conducted in Ethiopia to identify diseases in coffee plants using image processing and machine learning. These studies have shown promising results with high levels of accuracy in disease detection. The study used the combination of color and texture features from leaf images to identify five coffee plant diseases in Ethiopia: coffee leaf rust, coffee berry disease, anthracnose, root rot, and leaf spots. The study was more than 94 percent accurate in classifying the disease. Another study used CNNs to analyze images of coffee leaves and sort them into healthy or diseased groups. The study had an accuracy of 99.5% in identifying healthy leaves and 94.5% in identifying diseased leaves. Overall, the use of imaging and machine learning techniques for the recognition of coffee plant diseases shows great potential for improving disease diagnosis and management in Ethiopia's coffee industry. However, further research is needed to optimize these methods and develop practical applications for their use in the field.

[2] Computational method for Cotton Plant disease detection of crop management using deep learning and internet of things platforms:

Cotton is an important crop in many parts of the world, and its yield and quality can be affected by many diseases. Traditional methods of disease detection and control methods are time-consuming and costly and do not give accurate results. Therefore, there is a need a faster and more efficient methods of disease detection and management. One promising approach is the use of computational techniques, such as deep learning and Internet of Things (IoT) platforms. This process involves collecting and analyzing data from a variety of sources, such as plants images and environmental sensors, to detect and diagnose plant diseases. The recent research gives a computational approach for detection of cotton plant disease and crop management using deep learning and an IoT platform. The method uses a camera to capture images of cotton plants and then uses a deep learning algorithm to analyze those images to identify symptom of the disease. In addition, environmental sensors are used to gather information about factors such as temperature, humidity, and soil moisture, which can impact disease development. The study achieved good results with a deep learning algorithm achieving an accuracy of over 95% in detecting three common cotton plant diseases: bacterial blight, verticillium wilt, and fusarium wilt. The IoT platform also provided valuable environmental information that can be used to predict disease development and inform crop management decisions. Overall, the application of computational methods to cotton plant disease detection and crop management shows great potential for improving disease diagnosis and management in the cotton industry. However, further research is needed to improve these methods and develop practical applications for their use in the field.

[3] Farmer buddy-web based cotton leaf disease detection using CNN:

Cotton is an important crop that is subjected to many diseases that can affect crop yield and quality. Traditionally, farmers have relied on visual inspection to identify and control crop diseases. However, it may take time and effort to implement and requires specialized knowledge. Therefore, there is a need for a faster and more efficient way to detect and control disease. One of the recent study is a web-based cotton leaf disease diagnostic tool called "Farmer Buddy". It uses CNNs to analyze images of cotton leaves and identify disease symptoms. The system is designed to be accessible to farmers via web interface, allowing them to easily upload crop images for analysis. CNN used has been trained on a dataset of images of disease and disease-free cotton leaves and has been able to detect more than 95% of diseases. The system also provides information on the specific disease detected and recommendations on appropriate control strategies. The study demonstrates that the Farmer Buddy system has the potential to improve disease control in the cotton industry by providing farmers with a quick and easy detection tool. However, the technique's effectiveness in real-world will depend on factors such as image quality and the accuracy of on-site disease diagnosis. Overall, the use of web-based platforms and CNNs for cotton leaf disease detection shows great potential for improving disease management in the cotton industry and reducing the economic impact of crop diseases.

[4] A recommended system for crop disease detection and yield prediction using machine learning approach:

Crop disease detection and yield estimation are important aspects of agricultural management that can have a significant impact on crop production and economic viability. Traditional methods of disease detection and yield prediction can be time-consuming and require specialized knowledge. Faster and more effective methods for the detection of diseases, as well as their prediction of crop yields need to be developed. This research is the detection of crop disease and yield prediction with a help of machine learning approach. The method involves collecting data from various sources, such as images of plants and environmental sensors, and using machine learning algorithms to analyze this data and predict disease occurrence and crop yield. Diseased and non-diseased tomato plant image dataset was used, and CNN was trained to classify the images into healthy or diseased categories. The study also used a Support Vector Regression (SVR) algorithm to detect tomato crop yield depending on environmental factors such as temperature, humidity and soil moisture. The recommended system showed promising results, with the CNN achieving an accuracy of over 95% in disease classification, and the SVR algorithm achieving a mean absolute error of 0.38 in yield prediction. The study showed that the application of machine learning algorithms to detect and predict crops diseases could lead to improvements in crop management, as well as higher yields. However, further research is needed to optimize these methods and develop practical applications for their use in the field.

[5] Cotton Disease Detection Using TensorFlow Machine Learning Technique:

Cotton is an important crop that is vulnerable to various diseases that can significantly impact crop yield and quality. The traditional methods of disease detection and management can be time-consuming, and may not always provide accurate results. Therefore, there is a need for faster and more efficient methods of disease detection and management. The proposed system used TensorFlow machine learning technique to detect cotton disease. The system involves capturing and analyzing images of cotton leaves to detect any signs of disease using a deep CNN. The study employed a dataset on images of cotton leaf that had and did not show disease, which CNN was trained to use for classifying them as benign or diseased. This study achieved an accuracy of over 97% in disease classification, indicating the more potential of the proposed method in detecting cotton diseases. This system's effectiveness in the field will depend on various factors, including the quality of images captured, the accuracy of diagnosis in the field, and the ability to take timely action to manage the diseases. Overall, the use of algorithms, such as TensorFlow, for cotton disease detection, shows great potential for improving disease diagnosis and management in the cotton industry. The proposed method can reduce the economic impact of crop diseases by allowing farmers to take timely action to manage diseases, thus improving crop productivity and quality.

[6] Cotton leaf disease detection & classification using multi SVM:

Cotton is an important cash crop, and diseases can have a significant impact on its yield and quality. In order to effectively manage the disease, it is necessary to detect and classify cotton diseases as soon as possible. Traditional methods of disease detection are often time-consuming, and expert knowledge is required. Therefore, there is a need for more efficient and accurate methods of disease detection and classification. The study suggested a system for the detection and classification of disease in cotton leaf using multi support vector machines. The system involved capturing images of cotton leaves, extracting features, and to classify the images into healthy or diseased categories using multi-SVM classifiers. This study used a dataset of cotton leaf images with and without disease and achieved an accuracy of over 95% in disease classification using the proposed multi-SVM approach. For the detection of different cotton leaf diseases such as bacteria blight, anthrochnose, and leaf curl, it is possible to use this proposed method. The use of multi SVM for detection and classification of cotton leaf diseases has been shown to be an effective and reliable method in detecting these diseases at a timely stage. The proposed method has the potential to improve cotton disease management by allowing for timely action to be taken to control the spread of the disease, leading to increased crop yield and quality.

[7] Machine Learning:

Machine learning is an area of artificial intelligence that involves using algorithms to learn from data and make predictions or decisions that are not explicitly programmed. There's a lot of machine learning algorithms, each with their strengths and weaknesses, so you need to choose the right one for your particular problem. An overview of some of the most widely used machine learning algorithms, including supervised ones such as Decision Trees, K-nearest Neighbors and Support Vector Machines, along with unsupervised methods like clustering and key components analysis is provided in a recent review article. The review discussed the key features, strengths, and limitations of each algorithm, and provided examples of their practical applications in various fields, including healthcare, finance, and image recognition. The review highlighted the importance of choosing the right algorithm for a particular problem, and the need for domain knowledge and expertise in selecting, training, and optimizing machine learning models. Overall, the review provides a useful introduction to machine learning algorithms, their strengths and limitations, and their practical applications. It serves as a helpful resource for researchers and practitioners interested in machine learning, providing a foundation for further exploration and experimentation in this rapidly evolving field.

[8] Survey of machine learning algorithms for disease diagnostic:

The view of its ability to analyze a large amount of data and predict accurately, machine learning algorithms are increasingly being used for disease diagnosis. The survey article provided an overview of various machine learning algorithms used for disease diagnostic. The article examined the main features and applications of a range of machine learning algorithms, which include support vector machines, neural networks, decision trees, quintuple algorithm methods as well as ensembles. The article also highlighted the strengths and limitations of each algorithm and provided examples of their use in different disease diagnostic applications, such as cancer diagnosis, diabetes prediction, and cardiovascular disease detection. The survey emphasized the importance of data quality and quantity, as well as domain expertise, in choosing the appropriate machine learning algorithm and optimizing its performance. The article also noted the need for interpretability and transparency in machine learning models to ensure their clinical relevance and acceptance. Overall, the survey gives a comprehensive overview of machine learning algorithms used for disease diagnostic and highlights their potential for improving disease detection and management. The article serves as a useful reference for researchers and practitioners in the medical field, providing a foundation for further research and development of machine learning-based disease diagnostic methods.

[9] An overview of the research on plant leaves disease detection using image processing techniques:

Plant disease detection is crucial for maintaining plant health and preventing crop losses. The use of these techniques for the detection of plant disease is becoming more and more popular due to new advances in image processing and machine learning. A full summary of the research on detecting plant leaves diseases through image processing techniques was included in the overview article. The article discussed the various steps involved in image processing-based disease detection, including image acquisition, preprocessing, feature extraction, and classification. The article highlighted the importance of selecting appropriate image acquisition techniques to obtain high-quality images and pre-processing techniques to remove noise and enhance image contrast. Different techniques of feature extraction, e.g., color, texture or shape and their suitability to different tasks in the detection of plant diseases have also been discussed. The article also reviewed different classification algorithms used for plant disease detection, including decision trees, support vector machines, and deep learning techniques such as convolutional neural networks. The article highlighted the importance of selecting the appropriate classification algorithm dependence on the nature of the difficult and the available input. Overall, the study on detection of plant leaf disease through image processing methods is summarized in this article. It highlights the potential of these techniques for improving plant disease management and identifies areas for further research and development.

[10] Image processing based approach for diseases detection and diagnosis on cotton plant leaf:

Cotton the most important crops worldwide, and its yield can be significantly impacted by various diseases. To address this issue, researchers have been investigating the use of image processing techniques for cotton plant disease detection and diagnosis. A recent article presented an image processing-based approach for cotton plant disease detection and diagnosis. The proposed approach involved three main steps: image acquisition, preprocessing, and feature extraction. In image acquisition step, images of cotton plant leaves were captured using a digital camera. In the preprocessing step, noise was removed, and the images were enhanced to improve their quality. In the feature extraction step, various features were extracted from the preprocessed images using techniques like as color, texture, and shape analysis. The Support Vector Machine (SVM) classifier is used to extract main features from data and train it, which was able to distinguish between healthy and diseased cotton plant

leaves. The accuracy of the proposed approach was compared to that of other state-of-the-art approaches, and it was found to outperform them in terms of accuracy and processing time. Overall, this article demonstrated the potential of image processing-based approaches to cotton plant disease detection and diagnosis. This proposed approach could be useful for identifying and managing cotton plant diseases in a timely and efficient manner, ultimately leading to improved crop yields.

[11] Early Detection of Cercospora Cotton Plant Disease by Using Machine Learning Technique:

Cercospora is a fungal infection that affects cotton plants, and it can lead to crop losses. Early detection of this disease is crucial for effective disease management. The study proposed a machine learning-based approach for early detection of Cercospora cotton plant disease. The proposed approach involved the use of digital image processing and machine learning techniques. In the image processing step, color images of cotton plant leaves were captured and pre-processed to delete unwanted noise and enhance image quality. The pre-processed images were then subjected to feature extraction using techniques such as texture analysis and wavelet transformation. For the training of Machine Learning models such as Decisions Trees, K-Nearest Neighbours, Support Vector Machines, Classification and Random Forests, extracted features have been applied. This accuracy of the models was evaluated using various performance metrics, including precision, recall, and F1-score. The study found that the random forest classifier performed best, achieving an accuracy of 96.4% in detecting Cercospora cotton plant disease. The proposed approach was found to be more accurate than traditional disease diagnosis methods, such as visual inspection and laboratory testing. Overall, the study demonstrates the potential of machine learning techniques for early detection of Cercospora cotton plant disease. The proposed approach could be useful for farmers and agricultural workers in identifying and managing the disease in a timely and efficient manner, ultimately leading to improved crop yields.

[12] Supervised Learning Algorithms: A Comparison:

For a range of applications, such as the recognition of images, linguistic processing and predictive analytical analysis, supervised Learning algorithms are commonly applied in Machine Learning. The study compared and evaluated the performance of six different supervised learning algorithms: decision trees, random forests, k-nearest neighbors, support vector machines, logistic regression, and naive Bayes. A number of benchmark datasets were considered for algorithm evaluation based on different performance criteria, such as accuracy, precision, recall and F1 score. The results showed that random forests and SVM performed best in terms of overall accuracy, while decision trees and naive Bayes performed the worst. The study also compared the algorithms in terms of their computational efficiency and found that decision trees and Naive Bayes were the fastest, while support vector machines were the slowest. Additionally, the study highlighted the importance of choosing the appropriate algorithm for the specific problem and dataset. Factors such as the common of the data, the number of features, and this amount of training data can all influence the performance of different algorithms. Overall, the study provides a comprehensive comparison of supervised learning algorithms and highlights their strengths and weaknesses. The findings can be useful for researchers and practitioners in selecting the appropriate algorithm for their specific application and dataset.

[13] Machine learning: Trends, perspectives, and prospects:

Machine learning is a rapidly growing field with numerous applications in various sectors, such as healthcare industries, transportation. The article gives an overall study of the current trends, perspectives, and prospects in machine learning. The article notes that the growth of deep learning techniques has led to considerable improvements in many applications, e.g., for image recognition and linguistic processing. Additionally, the article discusses the growing importance of explainable AI, which aims to provide insights into how machine learning models make predictions, and the need for ethical considerations in AI development. The article also discusses the potential impact of machine learning on various industries, such as healthcare, where it can be used for early disease detection and personalized treatment plans, and finance, where it can be used for fraud detection and risk assessment. In addition, this article addresses challenges related to machine learning, such as the need for a large amount of data, potential biases in algorithmic decision making and new information needs to be continuously learned and adapted. The article concludes with a discussion of the future prospects of machine learning, highlighting the potential for advancements in areas such as autonomous systems, reinforcement learning, and quantum computing. This article gives a study of the current state of machine learning, its challenges, and its potential impact on various industries. The insights presented can be useful for researchers, practitioners, and policymakers in understanding the trends and future prospects of machine learning.

[14] Identification of banana fusarium wilt using supervised classification algorithms:

Fungal diseases which affect banana plants and may result in significant crop loss are the fungus fusarium wilt. This study suggested using a machine learning method for the detection of bananas fusarium wilt by unsupervised classification algorithms. The study involved the use of hyperspectral imaging and machine learning techniques. Hyperspectral images of banana plants were captured and pre-processed to remove noise and enhance image quality. Using methods like principal component analysis PCA-8873 and Linear Discriminant Analysis LDA-8873, preprocessed images have been subject to feature extraction. The study found that the random forest classifier performed best, achieving an accuracy of 92.06% in identifying banana fusarium wilt. The proposed approach was found to be more accurate than traditional methods of disease diagnosis, such as visual inspection and laboratory testing. Overall, the research demonstrates the potential of machine learning techniques is identifying banana fusarium wilt. The proposed approach could be useful for farmers and agricultural workers in detecting the disease in a timely and efficient manner, ultimately leading to improved crop yields.

[15] Detection of strawberry powdery mildew disease in leaf using image texture and supervised classifiers:

Strawberry powdery mildew is a fungal disease that affects strawberry plants and can cause significant crop losses. The study proposed an image processing and machine learning-based approach for detecting strawberry powdery mildew disease in leaves. The study involved capturing images of strawberry leaves using a digital camera and processing the images using texture analysis techniques to extract features related to the presence of powdery mildew. The extracted features were then used to train several supervised classification algorithms, including decision trees, k-nearest neighbors, random forests, and support vector machines. Various performance metrics, which included accuracy, precision, recall and F1 score, were used to evaluate the models' accuracy. The study found that the random forest classifier performed best, achieving an accuracy of 98.67% in detecting powdery mildew disease in strawberry leaves. The proposed approach was found to be more accurate and efficient than traditional methods of disease diagnosis, such as visual inspection and laboratory testing. The study demonstrates the potential of image processing and machine learning techniques for detecting strawberry powdery mildew disease in a non-destructive and automated manner. The proposed approach could be useful for farmers and agricultural workers in detecting the disease in a timely and efficient manner, ultimately leading to improved crop yields and reduced crop losses.

[16] CottonLeafNet: cotton plant leaf disease detection using deep neural networks:

The need for timely identification of diseases on cotton plants, which may help farmers implement effective measures to prevent disease transmission and reduce crop loss. The authors start by discussing the significance of cotton as a major cash crop and the challenges posed by diseases that affect cotton plants. The work emphasised that in order to ensure the healthy and productivity of cotton crops, a precise and timely detection of disease is essential. The proposed method, called CottonLeafNet, leverages the power of deep neural networks for automated disease detection. The authors describe the network architecture, consisting of a number of convolution layers in addition to pooling and complete interconnections. For training and evaluation, they are using a large number of images from the cotton leaf that include both normal and infected samples. The results of the tests show that CottonLeafNet can achieve an excellent accuracy, and is better than other techniques for detection of leaf diseases in cotton. The authors have pointed out that this success can be attributed to the network's deep learning capabilities which allow it to infer from input images in a discriminating manner. The work provides a suitable solution for the detection of cotton plant leaf diseases through CottonLeafNet. The accuracy of the diagnosis of cotton leaf disease has been shown to be significantly improved by the proposed Deep Neural Network architecture. The authors claim that this technology, which would result in improved cotton yields and reduced economic loss, can be further developed and deployed as a means of assistance to farmers with monitoring and management of crop health.

3. Proposed Methodology

The purpose of this proposal is to identify leaves and seeds from plants in an extraction method taking account of various characteristics, including shape, colour or texture. CNN is a computer learning technique used to classify tree leaves into healthy or diseased. This is a new system, which helps the user to recognize disease transmission in plants. In order to obtain the most effective results and efficiency, it is first necessary to acquire an image of individual leaves with a higher resolution camera. To extract useful features that will be needed for further analysis, a preliminary processing technique is then applied to these images.

Algorithm:

The Convolutional Neural Network is a kind of neural network, widely used in machine learning for image classification, object recognition, and other computer task. CNN are particularly suitable for certain tasks, as they are designed to work a data with a grid-like topology, such as images, it can learn the raw pixel text and extract features from the data. The key building blocks of a CNN are Convolutional layers, which uses the process of learning filters for input images to extract features, followed by the water layers, which reduces the layer standard output to reduce the dimension of a

feature map. CNNs can also include other types of layers, such as dynamic layer that introduce non-linearity into the model, and fully connected layers, which take the output of the past layers and produce the final classification output. Overall, CNNs have become a powerful tool for image analysis and are widely used in area such as self-driving cars, medical image analysis, and facial recognition.

Algorithm Steps:

- A sample is selected from the given data.

- The first step is the convolution operation used to extract the main features of the image. ReLU shall operate element wise operations and set all negative pixels to 0.
- Max pooling is a feature mapping function that decreases its dimensionality.
- Flattening is used when combining maps into a single, long range 2x2 array in order to transform all the resulting 3D arrays.
- Fully connected layer is one that transforms the output into classes that the network needs.

Manual Tracing with Respect to a Small Sample:

Convolution layer

It's the initial step in the process of extracting essential features from an image. The convolution layer is composed of a number of filters to carry out the convolution operation. Every image shall be treated as a matrix of pixel values.

ReLU layer

The next step is moving them to a ReLU layer as soon as the features are extracted. ReLU performs an element wise operation, setting all negative pixels to 0. It introduces non-linearity to the network, and the generated output is a rectified feature map.

Pooling Layer

Pooling is a way of down sampling feature maps, which decreases their dimensionality. To create a pooled feature map, the rectified feature map goes through a pooling layer.

Flattening

Flattening is an additional step in the process. It is applied for the conversion of all resulting two-dimensional arrays from a set of pooled feature maps that are then one long, continuous square vector.

Fully Connected Layer

A full connection layer is a level in which input from different layers is flattened and sent to the vector. The network translates the output into a list of desired classes.

Dataset Description:

Plants and leaves are divided into two sets: one is diseased and another one is non-diseased. There are 2310 cotton plants and leaves images in the dataset. The training dataset has 1952 images and testing dataset has 108 images.

The diseased cotton leaf samples are shown in Figure 1. The diseased cotton plant samples are shown in Figure 2. The healthy and non-diseased cotton leaf samples are shown in Figure 3. The healthy and non-disease cotton plant samples are shown in Figure 4.

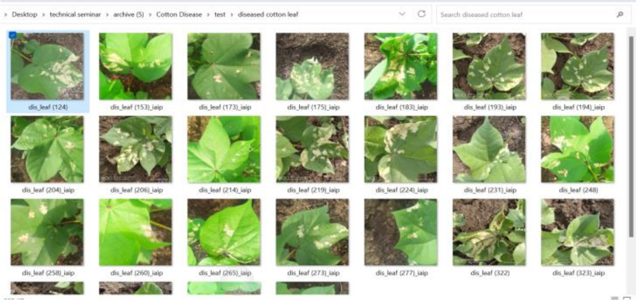

Figure 1. Diseased Cotton Leaf Samples

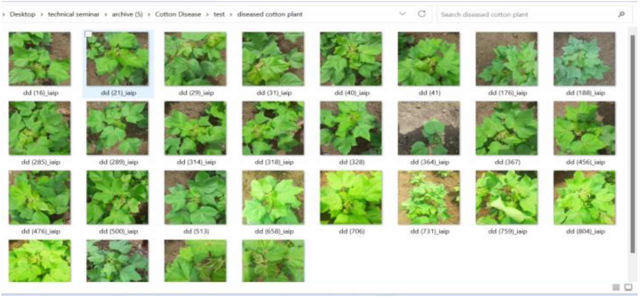

Figure 2. Diseased Cotton Plant Samples

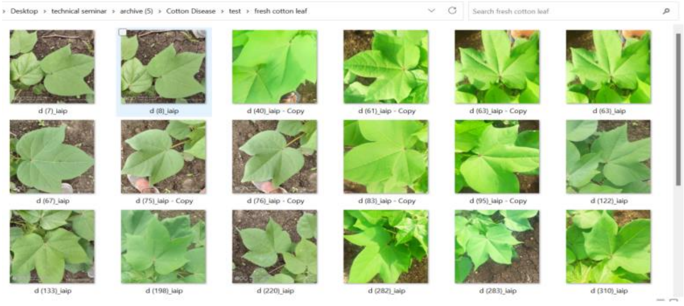

Figure 3. Healthy Cotton Leaf Samples

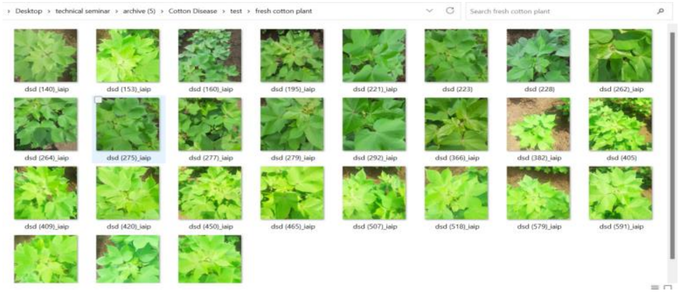

Figure 4. Healthy Cotton Plant Samples

4. Results and Discussion

A) Python

Image data preprocessing is an essential step in preparing image data for analysis or machine learning tasks. Python provides several libraries and tools for image data preprocessing, such as OpenCV, NumPy, and scikit-image, and for making predictions and determining accuracy. The results achieved for the cotton disease dataset are shown below.

Dataset: https://www.kaggle.com/datasets/janmejaybhoi/cotton-disease-dataset

Results:

An InceptionV3 model, which is a pre-trained CNN model for image classification, is created using the Keras library as shown in Figure 5.

Figure 5. Inception V3

New CNN model is created using the Keras library, which is commonly used for image classification task. The Conv2D, MaxPooling 2D, activation function ReLu are used, as shown in Figure 6.

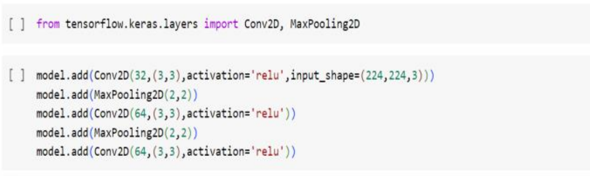

Figure 6. Convolutional Neural Network

This code defines a CNN model with three convolutional layers, each followed by a max pooling layer. The first convolutional layer takes input images with a shape of (224, 224, 3), and subsequent layers operate on the feature maps produced by the previous layers. The number of filters increases from 32 to 64 in the subsequent convolutional layers. The ReLU activation function is applied after each convolutional layer to introduce non-linearity in the network. Max pooling layers down sample the feature maps to capture the most salient information.

```
<ipython-input-32-ce2d88746bbd>:1: UserWarning: `Model.fit_generator` is deprecated and
  history=model.fit_generator(train_set,
Epoch 1/10
61/61 [==============================] - 562s 9s/step - loss: 1.2897 - accuracy: 0.5177
Epoch 2/10
61/61 [==============================] - 263s 4s/step - loss: 0.8149 - accuracy: 0.6868
Epoch 3/10
61/61 [==============================] - 260s 4s/step - loss: 0.7128 - accuracy: 0.7314
Epoch 4/10
61/61 [==============================] - 254s 4s/step - loss: 0.6155 - accuracy: 0.7565
Epoch 5/10
61/61 [==============================] - 264s 4s/step - loss: 0.5578 - accuracy: 0.7960
Epoch 6/10
61/61 [==============================] - 256s 4s/step - loss: 0.5152 - accuracy: 0.8042
Epoch 7/10
61/61 [==============================] - 253s 4s/step - loss: 0.4167 - accuracy: 0.8426
Epoch 8/10
61/61 [==============================] - 252s 4s/step - loss: 0.3968 - accuracy: 0.8560
Epoch 9/10
61/61 [==============================] - 250s 4s/step - loss: 0.3837 - accuracy: 0.8514
Epoch 10/10
61/61 [==============================] - 249s 4s/step - loss: 0.3285 - accuracy: 0.8811
```

Figure 7. Model Compilation

The provided output gives an overview of the model's performance during training. As the epochs progress, changes can be observed in the loss and accuracy metrics. The training loss and validation loss generally decrease, indicating that the model is learning and improving its performance. Similarly, the training accuracy and validation accuracy tend to increase, indicating that the model is becoming more accurate in its predictions. The values of these metrics at the end of the training can be used to evaluate the overall performance of the trained model, which is shown in Figure 7.

```
resnet = ResNet50(input_shape = image_size + [3], weights = 'imagenet', in

Downloading data from https://storage.googleapis.com/tensorflow/keras-applic
ations/resnet/resnet50_weights_tf_dim_ordering_tf_kernels_notop.h5
94773248/94765736 [==============================] - 40s 0us/step
```

Figure 8. ResNet

The resulting `ResNet` object is an example of the ResNet50 model with the specified input shape and pre-trained weights from ImageNet. It can be further customized or used for tasks like feature extraction, transfer learning, or fine-tuning on a specific dataset.

```
61/61 [==============================] - 224s 4s/step - loss: 0.9996 - accuracy: 0.6725 -
Epoch 10/20
61/61 [==============================] - 221s 4s/step - loss: 0.6815 - accuracy: 0.7268 -
Epoch 11/20
61/61 [==============================] - 220s 4s/step - loss: 0.5976 - accuracy: 0.7622 -
Epoch 12/20
61/61 [==============================] - 218s 4s/step - loss: 0.9699 - accuracy: 0.6920 -
Epoch 13/20
61/61 [==============================] - 229s 4s/step - loss: 0.7025 - accuracy: 0.7248 -
Epoch 14/20
61/61 [==============================] - 234s 4s/step - loss: 0.6559 - accuracy: 0.7412 -
Epoch 15/20
61/61 [==============================] - 243s 4s/step - loss: 0.6774 - accuracy: 0.7330 -
Epoch 16/20
61/61 [==============================] - 244s 4s/step - loss: 0.6877 - accuracy: 0.7406 -
Epoch 17/20
61/61 [==============================] - 235s 4s/step - loss: 0.7359 - accuracy: 0.7171 -
Epoch 18/20
61/61 [==============================] - 238s 4s/step - loss: 0.8646 - accuracy: 0.6950 -
Epoch 19/20
61/61 [==============================] - 236s 4s/step - loss: 0.6636 - accuracy: 0.7381 -
Epoch 20/20
61/61 [==============================] - 227s 4s/step - loss: 0.4888 - accuracy: 0.8227 -
```

Figure 9. ResNet Output

Using ResNet gives 82.2% accuracy as shown in Figure 9. The following table shows the comparison of ResNet and Inception V3 models.

Table 1. Comparison of ResNet and Inception V3

ResNet	**Inception V3**
It specifies input shape and pre-trained weights from ImageNet.	It pre-trains CNN model for image classification, using the Keras library.
RseNet gives 82.2% accuracy	Inception V3 gives 88.1% accuracy

B. Discussion

It has been discovered that the accuracy of Convolution Neural Networks is higher than all algorithms when applying machine learning approach to training and testing. The accuracy is calculated with the aid of keras and Inception v3, and is found to be 88.1% accurate.

5. Conclusion and Future Works

The primary purpose of this research is to determine whether the cotton leaf is diseased or healthy, with the help of a convolution neural network. The objective of this algorithm is to recognize abnormalities that occur on plants. The model has been developed by integrating 2310 images of cotton leaves and flowers into the Convolution Neural Network. With an approximate accuracy of 87 percent, the model was able to classify. The introduction of an automatic notification about plant disease, which must be sent to farmers so that they can use the appropriate fertilizers for a specific illness, will further enhance this research.

References

[1] Mengistu, A.D., Alemayehu, D.M. and Mengistu, S.G. Ethiopian coffee plant diseases recognition based on imaging and machine learning techniques. International Journal of Database Theory and Application, vol no 9, pp.79-88, year 2016.

[2] Patil, B.V. and Patil, P.S., 2021. Computational method for Cotton Plant disease detection of crop management using deep learning and internet of things platforms. In Evolutionary Computing and Mobile Sustainable Networks: Proceedings of ICECMSN 2020 (pp. 875-885). Springer Singapore.

[3] Kumbhar, S., Nilawar, A., Patil, S., Mahalakshmi, B. and Nipane, M., 2019. Farmer buddy-web based cotton leaf disease detection using CNN. Int. J. Appl. Eng. Res, 14(11), pp.2662-2666.

[4] Akulwar, P. A recommended system for crop disease detection and yield prediction using machine learning approach, pp.141-163, year 2020.

[5] Kumar, S., Ratan, R. and Desai, J.V., 2022. Cotton Disease Detection Using TensorFlow Machine Learning Technique. Advances in Multimedia, 2022.

[6] Patki, S.S. and Sable, G.S., 2016. Cotton leaf disease detection & classification using multi SVM. International Journal of Advanced Research in Computer and Communication Engineering, 5(10), pp.165-168.

[7] Mitchell, T.M. Machine learning (Vol. 1). New York: McGraw-hill, year 2007.

[8] Fatima, M. and Pasha, M. Survey of machine learning algorithms for disease diagnostic. Journal of Intelligent Learning Systems and Applications, 9(01), p.1, year 2017.

[9] Gavhale, K.R. and Gawande, U. An overview of the research on plant leaves disease detection using image processing techniques. Iosr journal of computer engineering (iosr-jce), 16(1), pp.10-16, year 2014.

[10] Khairnar, K. and Goje, N., 2020. Image processing based approach for diseases detection and diagnosis on cotton plant leaf. In Techno-Societal 2018: Proceedings of the 2nd International Conference on Advanced Technologies for Societal Applications-Volume 1 (pp. 55-65). Springer International Publishing.

[11] Shakeel, W., Ahmad, M. and Mahmood, N., 2020, December. Early Detection of Cercospora Cotton Plant Disease by Using Machine Learning Technique. In 2020 30th International Conference on Computer Theory and Applications (ICCTA) (pp. 44-48). IEEE.

[12] Josephine, P.K., Prakash, V.S. and Divya, K.S., "Supervised Learning Algorithms: A Comparison". Kristu Jayanti Journal of Computational Sciences (KJCS), pp.01-12 ,year 2021.

[13] Jordan, M.I. and Mitchell, T.M., "Machine learning: Trends, perspectives, and prospects", Science, Vol. No 349,Issue No.6245, pp.255-260, year 2015.

[14] Ye, H., Huang, W., Huang, S., Cui, B., Dong, Y., Guo, A., Ren, Y. and Jin, Y. Identification of banana fusarium wilt using supervised classification algorithms. International Journal of Agricultural and Biological Engineering, 13(3), pp.136-142, year 2020.

[15] Mahmud, M.S., Chang, Y.K., Zaman, Q.U. and Esau, T.J. Detection of strawberry powdery mildew disease in leaf using image texture and supervised classifiers. In Proceedings of the CSBE/SCGAB 2018 Annual Conference, Guelph, ON, USA (pp. 22-25), year 2018.

[16] Singh, P., Singh, P., Farooq, U., Khurana, S.S., Verma, J.K. and Kumar, M., 2023. CottonLeafNet: cotton plant leaf disease detection using deep neural networks. Multimedia Tools and Applications, pp.1-26