International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 9
***Article Received:** 25 July 2023 **Revised:** 12 September 2023 **Accepted:** 30 September 2023*

Cotton Crop Leaf Disease Detection System Using Machine Learning Approaches to Improve Efficiency”

Kuldeep Pal[1], Ankur Rana [2]

[1] Department of Computer Application, Quantum University, Roorkee
[2]Assistant Professor, Department of Computer Application, Quantum University, Roorkee

Abstract
India's economy heavily depends on agriculture. Over 70% of people in India make their living from agriculture. Accurate and prompt diagnosis of illnesses that harm crops is one of the biggest challenges facing the agriculture sector. Diseases affect crop quality and have the power to destroy whole hectares of agricultural output, costing farmers a lot of money.
Current diagnosis methods need the presence of highly experienced professionals and take a lot of time to examine the damaged crop, understand the symptoms, identify the illness, and provide effective remedies. Due to the limitations of these methodologies, researchers are now looking for other approaches to early illness detection and classification. Addressing food security may be facilitated by smart farming and adequate infrastructure.
In recent years, machine learning has demonstrated tremendous potential in identifying and categorising trends in linked academic disciplines. The goal of the current study is to evaluate the accuracy, precision, recall, and training time of traditional machine learning techniques like the Support Vector Machine (SVM) and random forest against the performance of convolutional neural network (CNN) methods and architectures like Inceptionv3, VGG16, and RasNet50 with data augmentation and transfer learning. The models were trained with the use of a manually gathered database from a farm and a government organisation, which had four distinct classes of photos, including healthy plants. The highest performing model was the Inceptionv3 architecture of CNN with transfer learning, which achieved an overall accuracy of 94 percent and met the demand for a more reliable and effective classification model. Additionally, when the quantity of training data rose, it was shown that the performance of the developed models increased.
The outcomes obtained using transfer learning algorithms on CNN architectures are extremely encouraging, and they may be further refined to create a thorough leaf disease diagnosis system that can function in a real-world environment. As a result, it may enable the agricultural community to recognise problems and start prompt treatment without the intervention of qualified specialists.

Keywords: Transfer Learning, CNN, SVM, Random Forest, Machine Learning

Introduction

Many South Asian countries, including Bangladesh, China, and India, rely heavily on agriculture (Patil & Burkpalli, 2021; Nadiruzzaman et al., 2021). The escalation of crop diseases in specific regions, attributed to climate change and global warming, has resulted in a notable decrease in agricultural output. One crop that is particularly important to the economy is cotton, also known as the "king of fibres" or "white gold" (Khairnar & Goje, 2020). Global trade is currently valued at 40 billion US dollars, but by the end of 2030, that value is predicted to climb to 60 billion US dollars (Meyer et al., 2023). Bangladesh, as the second-largest provider of ready-made clothing globally, has recently exported garments worth 20 billion US dollars worldwide (Mohiuddin, 2008). Notably, cotton plays a crucial role as a raw material in various industrial sectors.
Spinning mills are essential to the manufacture of yarn, which is a vital component utilised in our apparel factories. Bangladesh is endowed with favourable weather, lots of water resources, and rich soil, which makes it possible to use about 500,000 hectares of land for cotton production. Even though cotton is one of the most valuable crops in the world, it is constantly threatened by a variety of pests, illnesses, and climate variations like floods, droughts, and extreme temperatures. Yarn is a crucial component used in our garment industries, and its production depends on spinning machines. Because of its good climate, abundance of water resources, and fertile soil, Bangladesh can produce cotton on almost 500,000 hectares of land. Cotton is one of the most valuable crops in the world, but it is always in danger from a range of pests, diseases, and climate fluctuations like droughts, floods, and extremely high or low temperatures.
Traditional methods are used to identify any pests or illnesses of cotton, mostly by farmers in Bangladesh and India. This procedure frequently entails professionals or specialists visually identifying illnesses, which can be labour- and time-intensive, particularly in poor nations where there is a dearth of specialists in rural areas. The situation could get more

complicated if traditional procedures result in the misdiagnosis of illnesses with comparable symptoms. Furthermore, misdiagnoses that result in the use of highly toxic, inappropriate, or superfluous pesticides could further lower agricultural yields. The agriculture industry is seeing an increase in research interest in creating AI-based, automatic disease detection systems as a solution to these problems. Scholars raise important queries to address the above identified problems: (1) How can pests and illnesses in cotton be efficiently identified? (2) What is the best way to build an automated model for disease detection? (3) How can the performance of the model be assessed using parameters or other techniques?

The automated disease identification method necessitates considerable feature engineering, image processing, and the collecting of images from cotton leaves. With its feature extraction capabilities and high computational capacity, computer vision, and particularly Machine Learning (ML) and Deep Learning (DL), present promising research opportunities (Talukder, Islam, et al., 2023; Saleem et al., 2021; Dhaka et al., 2021; Kundu et al., 2021; Talukder, Hasan, et al., 2023a; Akhter et al., 2023; Sharmin et al., 2023). For example, Kaur et al. (2018) presented an automated system based on k-means principles for the detection of illnesses in soybean leaves. With a 92% accuracy rate, Revathi and Hemalatha (2012) identified cotton disease using the K-nearest neighbour (KNN) classification method. Accurate findings were obtained by Sarangdhar and Pawar (2017) when they presented an SVM-based regression technique for the identification and classification of five distinct cotton illnesses. Mehta et al. (2018) used decision tree and random forest techniques to create a highly accurate cotton disease prediction model. A deep learning method for disease diagnosis was presented by Jenifa et al. (2019) with an astounding 96% accuracy rate. Rapid advances in DL-based technology have been observed in recent years, and these breakthroughs have made a substantial contribution to classification tasks, especially in segmentation and illness detection (Talukder et al., 2022; Uddin et al., 2023; Ahmed et al., 2021).

Producing enough food to meet societal demand is now possible thanks to developed technologies. The food's and the crops' safety and security, however, remained unachieved. Farmers face difficulties due to factors such as climate change, a decrease in pollinators, plant diseases, and other issues. On a priority basis, a crucial foundation for these components needs to be attained [1,2]. Because they can cause food shortages and droughts, illnesses of plants pose a serious threat to human survival. As a result, it results in significant losses where farming is practised for profit. The application of technologies like machine learning (ML) and computer vision aids in the fight against sickness. [3,4]. Healthcare, transportation, business analytics, and agriculture have all seen new heights and developments because of the recent surge in machine learning, neural networks, and computer vision research. Agriculture's use of DL for various item detection and analysis activities has entered a new era thanks to its strengths in image processing and data analysis.

The categorization of cotton leaf diseases using various methods has been the subject of numerous scientific publications (Kumbhar et al., 2019; Jenifa et al., 2019; Revathi & Hemalatha, 2012).

The foundation of the Indian economy is agriculture. Automation and modernization of the conventional procedures and methods of illness detection are crucial for survival in such a changing environment and climate. To address the problems of supply and demand for the specific crop, the agriculture industry must be drastically upgraded. Crop optimal production can be attained by routine crop inspections and prompt disease detection when a crop becomes affected. The main issue with cotton farming in India is leaf diseases. Among the most prevalent are bacterial illnesses (grey mildew), fungal diseases (leaf spot, reddening), and viral infections (leaf curl). Overall product quality and quantity are reduced as a result. Continuous monitoring in agriculture increases output since it allows for the early identification of problems and the implementation of suitable remedies. Plant diseases can be found using a variety of techniques. Some illnesses have no evident signs or symptoms; thus, a thorough evaluation is necessary in these situations.

In this study, Cotton is used to investigate and analyse several methods for illness detection using machine learning techniques. Expert visual observations were traditionally used to identify plant diseases. However, they were time-consuming and came with a 1 high risk of subjective perception. As a result, it is critical to have a reliable and reasonably priced method for detecting crop leaf diseases so that treatment measures may be taken before the illness spreads. Recent advancements in Deep Learning technology and the vast amount of data that has been gathered provide a potential strategy for achieving this objective successfully and economically.

A particular kind of artificial intelligence called machine learning enables software programmes to anticipate events more accurately without having to be explicitly programmed. Every field of study is being affected by machine learning, which is spreading its wings across virtually all academic disciplines. Without comprehending it, nearly everyone uses machine learning on a regular basis. Creating an application that can learn on its own from a given dataset is the aim of machine learning. To improve future predictions or judgements, the model starts the learning process by looking for patterns in the datasets or observations provided. This makes it possible for computerised decision-making systems to learn on their own and make judgements without the need for human input. Machine learning algorithms are divided into two groups as follow Crop leaf diseases have become a problem since they may drastically reduce the crop's quality and yield. As it aids in monitoring huge crop fields and automatically identifies and categorises the illness as it manifests on the leaf, automation in crop leaf disease detection and classification is an essential study issue. The goal of this thesis is to identify crop leaf diseases using a "Comparative analysis of Classical ML Algorithms, Deep learning CNN algorithm, and CNN architecture-based Transfer Learning."

India is well known for a wide variety of crops that may be grown there, including common commodities like potatoes, onions, okra, rice, and wheat. Among these, cotton production is particularly noteworthy, with India being the world's largest exporter of cotton. However, several illnesses brought on by both biological and environmental factors seriously harm numerous crops in India, posing serious problems to the country's agricultural sector. Pests, unexpected temperature swings, environmental stresses, ecological shifts, dietary shortages, and toxins are all included in this list [6, 7].

The work by Sharma et al. (2022) is noteworthy since it provides a thorough assessment of deep learning and machine learning methods in a variety of application domains. Their examination of common practices in these domains provides insightful information that may improve our own process. Aggarwal et al. (2023) presented a proficient system that utilises diverse deep learning methodologies to forecast illnesses in rice leaves. Their suggested solution surpasses current methods by using a variety of machine learning classifiers and pretrained models for feature extraction, resulting in an amazing 94% accuracy rate in identifying rice leaf illnesses. For example, Kalpana et al. (2023) achieved a 96% accuracy rate in cotton leaf disease identification using deep learning-based models, such as VGG-16 and InceptionV3. Reddy et al. (2023) achieved an 89% accuracy rate in their studies using 720 photos, using a CNN-based deep learning model for cotton disease identification. ResNet50, ResNet152v2, and InceptionV3 models were used by Chander and Upendra Kumar (2022) to a dataset of manually gathered cotton plants and leaves. Their suggested approach showed improved classification accuracy rates at lower computational times, with up to 96% efficiency.

DenseNet-121 was used by Arathi and Dulhare (2023) to categorise photos of healthy and diseased plant leaves from the plant village dataset with a 91% accuracy rate.

There are several ways in which the planned research differs from earlier investigations. While Verma et al. (2022) concentrated on using UAVs to automate rubbish detection, the current work tackles the crucial problem of identifying cotton illnesses within the framework of Bangladeshi agricultural production. Additionally, Zekiwos and Bruck (2021) focused on using CNNs to detect pests and cotton disease in Ethiopia. By using Transfer Learning algorithms that have been fine-tuned, the suggested work aims to establish a deep learning-based method for cotton leaf disease identification, in contrast to the previous methodology. The study gets an amazing accuracy rate of 98.70% using models like VGG-16, VGG-19, Inception-V3, and Exception. The Exception model was selected for a web-based application. Its unique focus on cotton leaf disease detection with optimised deep learning techniques sets it apart, potentially improving cotton yield and broadening its applicability to other plant species.

This research highlights the importance of fine-tuned models, in contrast to most previous works that employed default and built-in deep learning models for identification without tuning. The accuracy levels attained by the suggested approach are higher than those of previous studies, proving the applicability of this inquiry. Furthermore, a web-based tool that uses the Xception paradigm has been created to help farmers quickly identify diseases. The results of this study could be advantageous for many aspects of the agriculture industry, providing ways to minimise crop loss by early disease identification and early treatment.

- AI-powered method for identifying diseases in cotton leaves: This proposal presents a deep learning model, which offers the possibility of more accurate identification of leaves affected by the disease. This could help experts diagnose patients more accurately, which would ultimately lead to better harvest results.
- Accelerated Disease Detection: This study offers a method to quickly identify diseases by utilising a web-based application. By utilising IoT devices and integrated cameras in agriculture, the plan seeks to assist farmers in quickly identifying any illnesses.
- Development of Deep Learning Models for Image Classification: Our research advances the development of deep learning models specifically suited to the smart agriculture sector worldwide. The research findings can be utilised to improve forecasts by employing deep learning models.

The establishment of a web-based application to help farmers detect cotton infections intelligently and the development of a deep learning (DL)-based methodology to reliably identify cotton diseases have been the main foci of this paper. In order to mitigate the loss of cotton goods, it is imperative that diseases be accurately and promptly identified. This will allow for the right implementation of actions and remedies. In order to do this, we first use several Transfer Learning (TL) pre-trained models and analyse their performance results. The best model is then incorporated into the web application, depending on how well it performs. As a result, our tool helps farmers identify cotton diseases quickly and precisely. The following are the main findings of this study:

- introduction of a DL-based method for identifying illnesses in cotton leaves by finetuning current TL algorithms.
- Building the Fine-tuning procedure, which involves modifying the layers and parameters of pre-existing models to customise our model for predicting cotton disease.

- Thorough testing of our suggested model utilising a range of performance indicators on the cotton leaf dataset, resulting in the determination of the best model to be implemented on a smart web application for precise cotton leaf disease diagnosis."

Literature Survey

Plant diseases are major issues in agricultural production, as stated by Awad Bin Naeem, Biswaranjan Senapati, Alok Singh Chauhan, Sumit Kumar, Juan Carlos Orosco Gavilan, and Wael M. F. Abdel-Rehim [1]. If not identified quickly, they can reduce both crop yield and quality. As is well understood, early diagnosis and warning being the cornerstones of successful plant disease prevention and control, which are crucial in management and decision-making. However, in many nations and locations, visual observations by professionals or experienced farmers remain the primary method for identifying plant diseases. The manual observations needed on large farms take a long time, and the cost of regular consultations with experts is prohibitive. Therefore, automated plant disease diagnostics is of great practical use since it seeks to recognise the symptoms of plant diseases as soon as they appear on leaves. Agriculture has been a major contributor to India's economic growth. When deciding what to grow, farmers examine factors such soil quality, average annual temperatures, and the crop's potential profit. As a result of rising populations, shifting weather patterns, and political instability, the agricultural sectors began exploring new techniques to increase food production. A farmer will employ pesticides to help with insect control, disease prevention, and increased crop yields. Due to industrial agriculture, poor yield, economic losses, and crop diseases, farmers are having challenges.

The necessity of a correct diagnosis places a premium on accurate sickness identification and assessment. Agriculture in economically developing nations is crucial because it provides food, income, and jobs to people living in rural areas. The greatest factor reducing agricultural productivity is crop loss due to plant diseases, which can be as high as 30 percent. Attempts to prevent such losses by using conventional methods of disease diagnosis have largely failed. Early and precise detection of plant diseases are vital for reducing losses caused by such illnesses. However, without the proper cultivation knowledge, competence, and a sense of disease prediction, such harvests and grains might suffer substantial damage, if not destruction. The farmers and the nation's economy both suffer as a direct result of this. This study therefore aims to apply some AI to agriculture to cut down on crop losses due to leaf diseases. To solve this problem, we used transfer learning models developed with different CNN architectures, including ResNet50, VGG19, InceptionV3, and ResNet152V2. We tested all four approaches on the gold-standard set of cotton leaves to see which one was most effective at identifying diseases.

In [2] Shuyue, the authors described the various graph convolutional neural network architectures. It was programmed to analyse standard EEG data to foresee how the four types of motor fantasies would connect to EEG electrodes. They dealt with their information by shifting it from a two-dimensional to a three-dimensional view. These dimensions measures were applied to the construction. A study suggested using short-term voltage stability to take advantage of deep learning's dynamic approach. They were able to boost reliability by using the clustering approach to achieve short-term voltage stability. It is mentioned in that many mango trees' leaf illnesses were detected using a deep learning technique. Five distinct leaf diseases were employed from a wide range of mango leaf samples, and over 1200 records were analysed. More than 600 images were utilised to teach the CNN structure, with 80% used for training and 20% for testing. Mango leaf illnesses were identified using the remaining 600 photos to test its accuracy and demonstrate its viability in real-time applications. The classification accuracy can be further boosted if more photos in the dataset are provided by modifying the parameters of the CNN model. The mechanism for identifying and classifying rice plant datasets is employed to process the CNN model, according to the study. From the rice experimental field, about 500 photos with illnesses were collected for training. In detection of cotton leafs were handled with image processing. In this case, the datasets are divided using K-means methods. Banana plant illnesses that infect their leaves have been identified, according to the study. Although 3700 photos were used for training in this work, no single class's dataset is evenly distributed. Experiments were conducted in a variety of modes, such as the training mode, where both colour and grayscale picture datasets, as well as a variety of dataset splitting strategies, were used. They achieved 80% and 20% accuracy on the training and validation datasets, respectively, for coloured images.

[3] Deep learning is a subset of machine learning techniques and a distinct approach to ANN education. Using a CNN classification technique, a model was proposed for identifying healthy leaves and 13 different damaged leaves in peach, cherry, pear, Apple, and Grapevine trees. More than 30000 images used in dataset, achieving accuracy between 91% and 98% for separate class test and average accuracy 96.3%. The accuracy of the approach increased to 99.35% with only 20% of the testing data, and to 98.2% with 80% of the testing data, when applied to a public dataset containing 54306 photos of 14 crops and 26 diseases. [19] Developed a model for CNN classifier to distinguish between Septoria, Frogeye, and Downy Mildew on Soybean plants. The accuracy of 99.32% was achieved on a dataset consisting of 12673 photos of leaves divided into four classes. Created a CNN classification method for identifying plant diseases. There are 87848 photos in the dataset, covering 25 plant species and 58 diseases; the detection rate is 99.53 percent.

This section describes several different Crop Disease Detection Systems that have been developed and deployed by other scientists. To learn more about their processes and techniques, you can look up the sources used in this work by consulting the reference page. Using image processing, the authors of employ Support Vector Machines (SVMs) to identify and categorise a wide range of agricultural illnesses. The pictures were taken by captured digitally amid the cotton fields. The image's backdrop is removed using a background removal technique during the preparation step. After the

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 9
***Article Received:** 25 July 2023 **Revised:** 12 September 2023 **Accepted:** 30 September 2023*

backgrounds have been eliminated, a thresholding technique is applied to the images to segment them. Features such as colour, shape, and texture will be extracted from the photos using various segmented images. Once features have been extracted, they will be fed into a classifier. Here, a supervised classification method is employed. Con: The work relied on photographs captured in a cotton field using a digital camera, which added extra expense and time to the project. In this paper, the authors use photos of Cotton Leaf Spot to describe technological solutions to the problem, and then classify the diseases using a neural network. The classifier is educated to accomplish smart agricultural goals like early disease identification in groves, targeted fungicide administration, etc. The suggested work relies on segmentation techniques for enhancing the acquired images, which are used for edge identification in images. Then image segmentation utilising RGB colour feature is carried out to get spots on illnesses. For later use in disease diagnosis and pest management advice, we extract image attributes including border, shape, colour, and texture from images of disease spots. It uses supervised categorization as its foundation. However, in order to detect disease spots, numerous picture features must be retrieved. These features include boundaries, shape, colour, and texture.

In this study, we tested the accuracy of deep neural networks in distinguishing between normal and unhealthy cotton plants and leaves. The VGG16 model was compared to the ResNet50 model and the Mobile Net model. Our research shows that VGG16 and Mobile Net models produce the best results across the board. Metrics used in evaluating models include accuracy, loss, and the number of correct predictions made by the model. Since the advent and widespread adoption of intelligent agricultural systems, there has been a rise in research devoted to the identification and detection of leaf diseases. We used datasets of healthy leaves, apple grey-spot disease, black star disease, cedar rust disease, and healthy leaves to investigate the detection and classification of apple leaf illnesses. SVM classifiers for image segmentation were used for evaluation and improvement alongside ResNet and VGG convolutional neural network models.

A method for the visual identification of plant stem and leaf diseases has been proposed by Al Bashish et al. [5]. The created framework relies on image processing, specifically the K-Means technique for segmenting images, and a trained neural network for processing the acquired image segments. The results show that the proposed method can detect leaf diseases automatically and accurately. The classifier produced which is based on statistical classification had performed well and can correctly classify and detect the diseases. Camargo et al [1] have put out an approach for identification of unhealthy region of plants using image processing. After the image has been captured, the colour adjustment is made. Images that have been modified are improved by applying a Gaussian filter. The region of interest can then be obtained using segmentation. Finding the optimal threshold for the separation is performed. Then by applying SVM classifier the segmented regions are identified and designated as unhealthy or healthy.

Image processing was used to detect cotton leaves [7]. In this case, the datasets are divided using K-means methods. The study [16] revealed the detection of leaf-infecting pathogens in banana plants. Although 3700 photos were used for training in this work, no single class's dataset is evenly distributed. Experiments were conducted in a variety of modes, including the training mode, where different colour and grayscale image datasets and dataset splitting strategies were used. They achieved 80% accuracy in training and 20% accuracy in validation on coloured images.

[8] P. Resized, contrast enhanced, and colour space converted images by Krithika et al. KMeans clustering using GLCM is used for segmentation and feature extraction. Multiple classes were used in the SVM classification. Colour space conversion and improvement were conducted by R. Meena and colleagues. The primary colours of leaves are translated into L∗A∗ ∗B . For this purpose, we employ the K-Means clustering technique. Features are extracted using the GLCM, and classes are determined using the SVM. Bharat et al. used a digital camera to take pictures, and then processed the pictures with a median filter. Segmentation is performed by K-Means clustering. SVM is used for classification. Segmentation is used to obtain the diseased areas, as described by Pooja et al. [13]. Segmentation is performed using a k-means clustering technique, Otsu's detection, and a conversion from RGB to HSI. Preprocessing was carried out by contrast adjustment and normalisation, as done by Rukaiyya et al., [14]. Bi-level thresholding and a YCBCR colour space conversion are carried out. Features are extracted and classified using the GLCM and the HMM.

A comprehensive study was conducted to evaluate various machine learning approaches for disease detection and categorization [9] Mrs. Shruthi U, Dr. Nagaveni V, Dr. Raghavendra B K. Support Vector Machine (SVM) classification, Artificial Neural Network (ANN) classification, K-Nearest Neighbour classification, Fuzzy C-Means classification, and Convolutional Neural Network (CNN) classification were all analysed as part of our study to determine which methods were most effective at identifying plant diseases. The SVM Classifier is a well-liked supervised learning method in machine learning that makes use of previously analysed data to complete classification tasks. Many scientists have relied on the SVM Classifier to identify plant diseases. Citrus fruits like grapefruits, lemons, limes, and oranges were analysed by the SVM Classifier to detect diseases. They zeroed in on the leaf infections caused by canker and anthracnose. The experimental results showed an impressive 95% actual acceptance rate. Like this, [2] scientists used the SVM Classifier to identify powdery mildew and downy mildew in grape plants, with an average accuracy of 88.89% for both illnesses. Proceeding to [3], they employed the SVM Classifier to identify. Going on to [3], they achieved high accuracies of 97% and 95%, respectively,

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 9
***Article Received:** 25 July 2023 **Revised:** 12 September 2023 **Accepted:** 30 September 2023*

using the SVM Classifier for oil palm leaf diseases including Chimaera and Anthracnose detection.

Using a dataset of more than 300 publically accessible photos, the SVM Classifier was used in yet another study [4] to identify potato plant diseases, notably Late blight and Early blight. Here, an astounding 95% accuracy was attained. Lastly, [5] used information from both LAB and HSI colour modes to classify grape leaf diseases, such as Black Rot, Esca, and Leaf Blight, with impressive accuracy.

[10] it is stated that deep learning technique was applied to identify the leaf diseases in different mango trees. The researchers used five different leaf diseases from various specimens of mango leaves, where they addressed nearly 1200 datasets. More than 600 images were utilised to teach the CNN structure, with 80% used for training and 20% for testing. Mango leaf illnesses were identified using the remaining 600 photos to test its accuracy and demonstrate its viability in real-time applications. The classification accuracy can be further boosted if more photos in the dataset are provided by modifying the parameters of the CNN model.

The authors of the Chen et al. (2020) paper presented a framework for disease identification that combines an LSTM unit with a bi-directional Recurrent Neural Network (RNN). The RNN uses climate-related factors to predict the discovery of illness. To predict the presence of diseases and pests, a time series prediction is first formulated. Then, for improved performance, the bi-directional LSTM network (Bi-LSTM) records the long-term dependencies between previous and future contexts in sequential data. Notable metrics that the Bi-LSTM attained were an Area Under the Curve (AUC) of 95%, an F1-score of 88%, an Accuracy and Loss (Acc) of 88%, and an Average Precision (AP) of 90%.

Ankur Das and Dutta (2020) created an automated feature engineering approach for rice leaf disease diagnosis in another study that used deep learning (DL) techniques. Four convolution layers, two fully connected layers, and a SoftMax output layer are all included in the Convolutional Neural Network (CNN) model that analyses photos of damaged leaves. The CNN model exemplifies automated feature engineering by producing a wide range of features devoid of human bias and identifying complex nonlinear correlations between different attributes. Irrelevant features are removed using the dimension decomposition method, and several classifiers are used to categorise diseases. The study used 10,500 photos of infected leaves. The results of the experiments showed that the accuracy of CNN was 91.07%, SVM was 88.73%, DT was 87.87%, RF was 87.09%, KNN was 89.33%, NB was 92.16%, LR was 90.35%, and bagging AdaBoost as 81.58%.

Proposed Methodology

The model and design of our suggested automated cotton disease prediction method are shown in this section. In Fig.1 we show the proposed methodology Architecture.

In this work, we present a framework for the detection of illnesses in cotton leaves using refined deep-learning models, such as VGG-16, VGG-19, InceptionV3, and Xception, which are wellknown models for image categorization. Transfer learning models are extensively trained using a big image dataset to improve detection accuracy. Therefore, the approach starts with gathering a large dataset. Before the raw images are fed into the tailored models for learning, superfluous information is removed by preprocessing. Finally, a different set of photos feature ing both healthy and sick leaves and plants are used to assess the model's performance. Key indicators like accuracy, precision, recall, and F1-score are used to assess the model's performance.

Procedure for the Proposal:

- Data Collection: Compile information from reliable sources to use as the proposal's starting point.
- Data preprocessing: To get ready photos for deep learning model training, tasks include cropping, rotating, sharpening, and augmentation.
- Model Creation and Instruction: Adjust VGG-16, VGG-19, InceptionV3, and Xception, using the pre-processed data (80% for training) to enhance their performance.
- Assessment of the Proposal: 20% of the data should be used to evaluate the proposal's performance using established measures to ascertain how well it detects cotton illnesses.

Data Analysis and Data Preprocessing

The real images in the dataset may contain noise and are not always fresh. Before fitting them into the learning module, preprocessing is necessary to improve accuracy and module performance. Operations include resizing, sharpening, rescaling, shearing, and zooming. These operations aim to enhance image quality, eliminate noise, and prepare the data for effective use in the learning module.

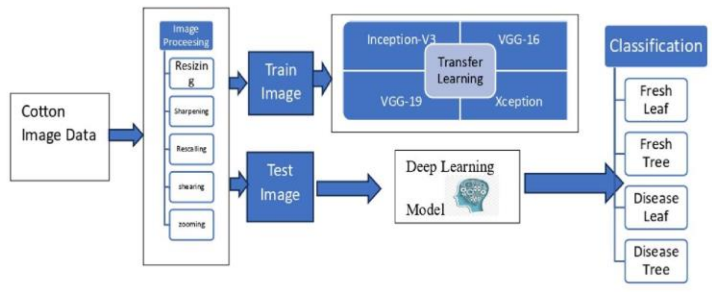


Fig.1 Proposed Model Architecture

Cotton Leaf　　　　　　　　　　　　　　cotton Plant

Fresh cotton leaf and Plant

Diseased Cotton Leaf　　　　　　　　　　　　Diseased cotton plant

In this study, we trained a variety of Transfer Learning (TL) and pre-trained Convolutional Neural Network (CNN) models, such as VGG-16, VGG-19, Inception-V3, and Xception, using preprocessed picture data. These are the specifics of the adopted TL algorithms:

VGG-16: Known for its uniform design and extensive application in image classification, especially in illness detection, VGGNet-16 has 16 convolutional layers. VGG-16 shows how, in certain situations, increasing network depth can lead to improvements in system performance. It is made up of convolution, pooling, and fully connected layers, all of which are essential for performance, information extraction, and spatial size reduction.

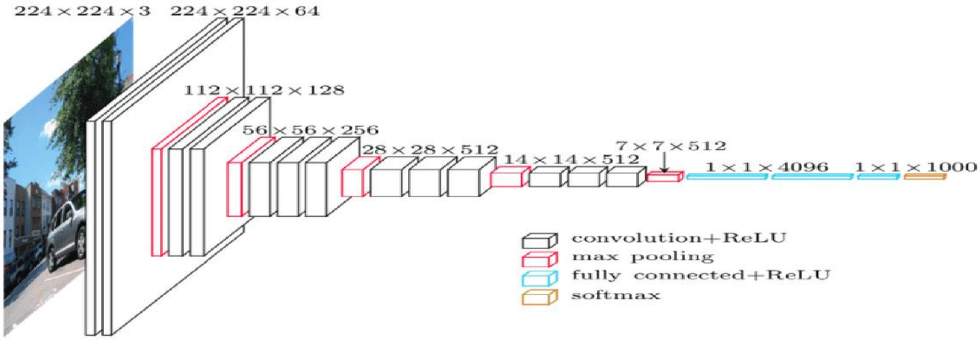
GG-19: First presented by Simonyan and Zisserman in 2014, VGG-19 is a 19-layer structure made up of 3 dense and 16 convolutional layers. It uses multiple 3x3 filters for efficient photo prediction and combines fully linked layers, convolution layers, and max pools to effectively categorise images into 1000 different categories.

Inception-V3: This enhanced version of the original Inception model makes use of simultaneous multi-scale convolutions. It has a classifier, the basic convolutional block, and an improved Inception module. Inception-V3 improves gradient convergence, stabilises learning outputs, and reduces overfitting by adding auxiliary classifiers. 1x1 convolutional kernels are used by the model to minimise feature channels and speed up training.

Xception: An extreme version of an Inception module, Xception maps the spatial relationship of each result unit independently and maps cross-channel interactions using a 1x1 convolution. It employs two convolutions, a depth wise convolution and a pointwise convolution, to optimise

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 9
***Article Received:** 25 July 2023 **Revised:** 12 September 2023 **Accepted:** 30 September 2023*

computing cost and parameter count. It is based on the idea of depth wise separable convolution, much like advanced Inception modules.



Original Architecture of VGG-16

Table 1. Experimental Parameter Setting

Parameters	Values
Learning Rate	.001
Drop Out	0.02
Optimizer	Adam
Shearing	-0.2 to +0.2
Horizontal flipping	True
Rotating	-20 to +20
Zooming	0.8 to 1.5
Validating split	0.2

Table 2. Confusion Matrix

Predicted Result	Actual Positive	Actual negative
Yes	TP	FP TN
No	FN	

Experiment setup and implementation

Accuracy: Indicating the proportion of correct predictions to total samples, accuracy has become the gold standard for gauging the efficacy of classification-based models. It indicates the percentage of test samples successfully classified by the model. It's a crucial indicator that helps us accomplish the aforementioned study goals.

Accuracy = (TP + TN)/(TP + TN + FP + FN)

Precision: Calculating precision involves dividing the number of times the model correctly predicted the test data class by the sum of the times it correctly predicted the test data class and the times it incorrectly predicted the test data class.

$$\text{Precision} = \frac{TP}{TP + FP}$$

Recall: Remember that another metric for judging a model's performance is the fraction of instances in which the model made the correct prediction relative to the total number of instances in which the model was either correct or erroneous in its classification.

$$\text{Recall} = \frac{TP}{TP + FN}$$

F1-Score: Since this is a classification problem with several classes, the F1-score considers the accuracy and the number of correct classifications for each category. The F1 score provides a trustworthy metric by using the harmonic mean of these two numbers to compare different models. The mathematical formula for this is presented in Equation.

$$\boxed{\text{F1 score} = 2\ \text{precision*Recall/precision+Recall}}$$

The activation function for our deep learning models in our proposal is the Rectified Linear Unit (ReLU) function. The ReLU function is preferred because it is good at managing non-linearities, which improves deep neural network performance. The ReLU function is a popular deep learning tool because of its ease of use and high computational efficiency. The mathematical definition of it is as follows: f(x)=max (0, x)

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 9
***Article Received:** 25 July 2023 **Revised:** 12 September 2023 **Accepted:** 30 September 2023*

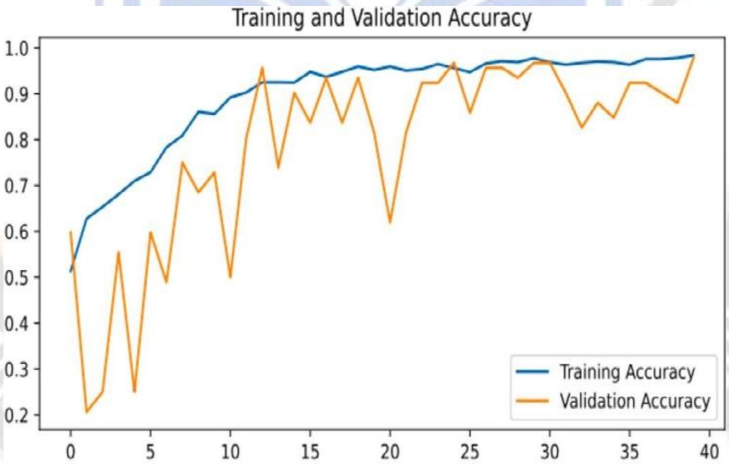
The ReLU function gives the model non-linearity, which makes it able to recognise complex patterns and representations in the input. Its simple thresholding processes are what give it its computing efficiency. We use the ReLU activation function in the hidden layers of our deep learning models as part of our fitness function. Our algorithms can identify and record non-linear connections and traits that are relevant to the diagnosis of crop diseases thanks to this application. The rationale behind the ReLU function's selection is its shown efficacy in a range of deep learning applications and its ability to improve our models' precision in precisely identifying agricultural illnesses. Our deep learning models perform better overall when they can learn and generalise from the given dataset more efficiently, which is made possible by utilising the ReLU activation function.

Result and Discussion

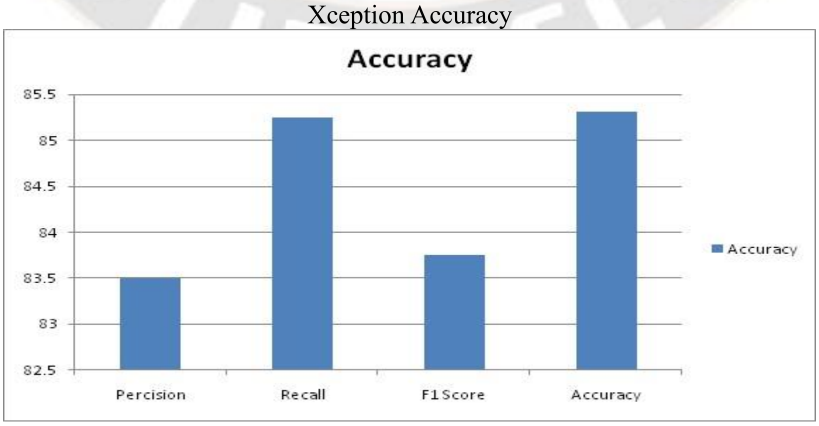
In our experimental setup, we employed several Transfer Learning (TL) models, fine-tuning them for the detection of cotton leaf and plant diseases within our dedicated cotton disease dataset. Performance metrics, including training and validation accuracy and loss, model accuracy, precision, recall, F1-scores, and confusion matrix, were employed for result analysis. Learning curves, representing the learning progress of the model throughout incremental training epochs, served as indicators of its learning capabilities.

Fig. 6 illustrates training and validation accuracy and loss. The loss curve signifies a continual decrease in validation and training losses over time, with minimal temporal gaps. Despite observing slight fluctuations during the validation test, the overall fitting of the curves was favorable. Fig. 7 displays the training and validation accuracy and loss for our proposal. The accuracy curves indicate increasing performance with training time, although the validation curve exhibits some fluctuations. The loss graph, on the other hand, demonstrates a good fit, with closely aligned training and validation loss curves.

Similarly, Figs. 8 and 9 depict the curves for Inception-V3 and Xception. The loss curves of both algorithms show an ideal fit with the dataset, with Inception-V3 displaying some fluctuations in training and validation loss over time. In contrast, the loss curves of Xception are smooth, fitting ideally with minimal disparity between training and validation loss. The training and validation accuracy for the Xception model is 98.15% and 97.26%, respectively. Conversely, Inception-V3 achieves a training and validation accuracy of 98.65% and 96.15%.

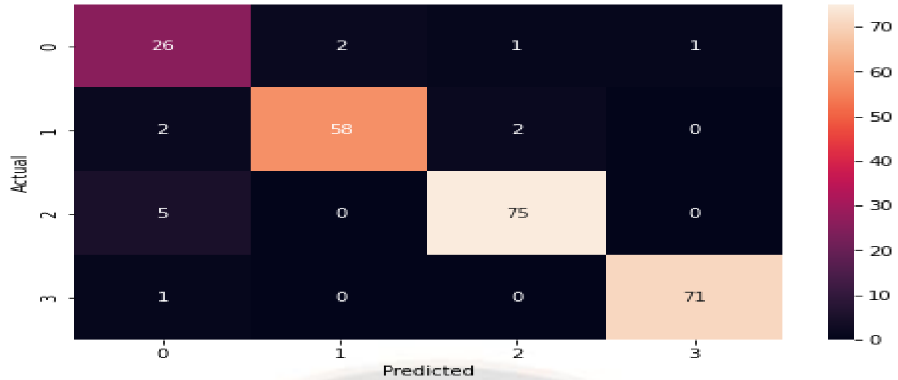


Xception Accuracy

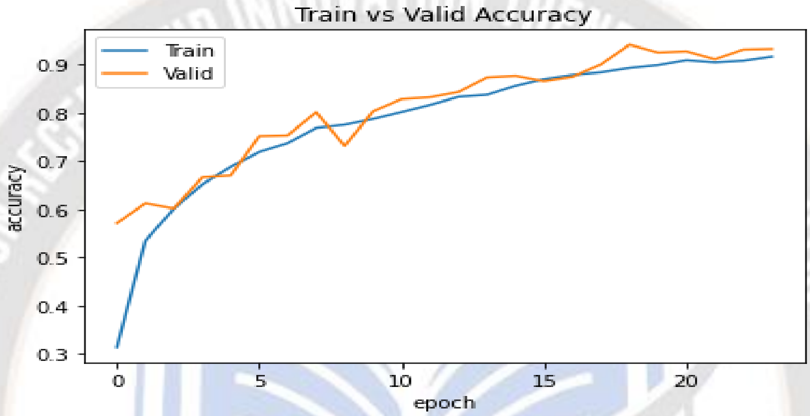


Xception Result

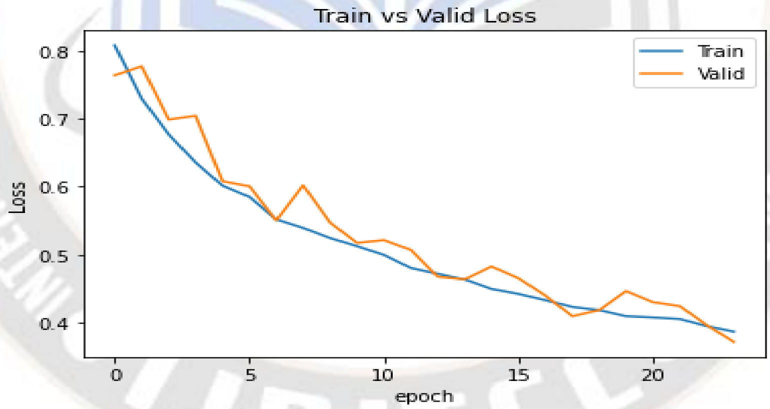
International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 9
***Article Received:** 25 July 2023 **Revised:** 12 September 2023 **Accepted:** 30 September 2023*



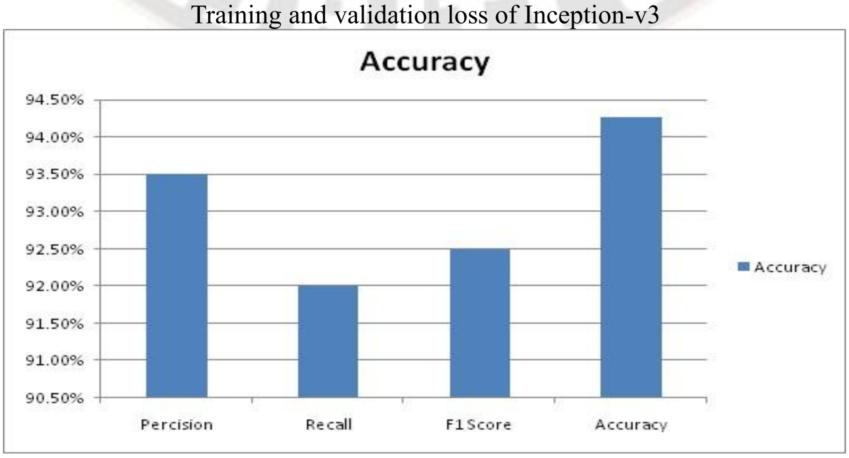
The Xception Confusion Matrix



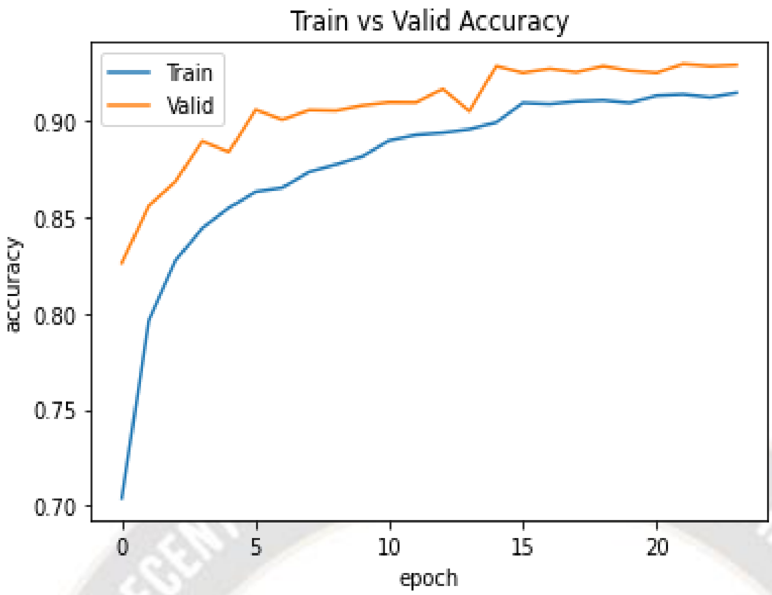
Inception-v3's accuracy during training and validation.

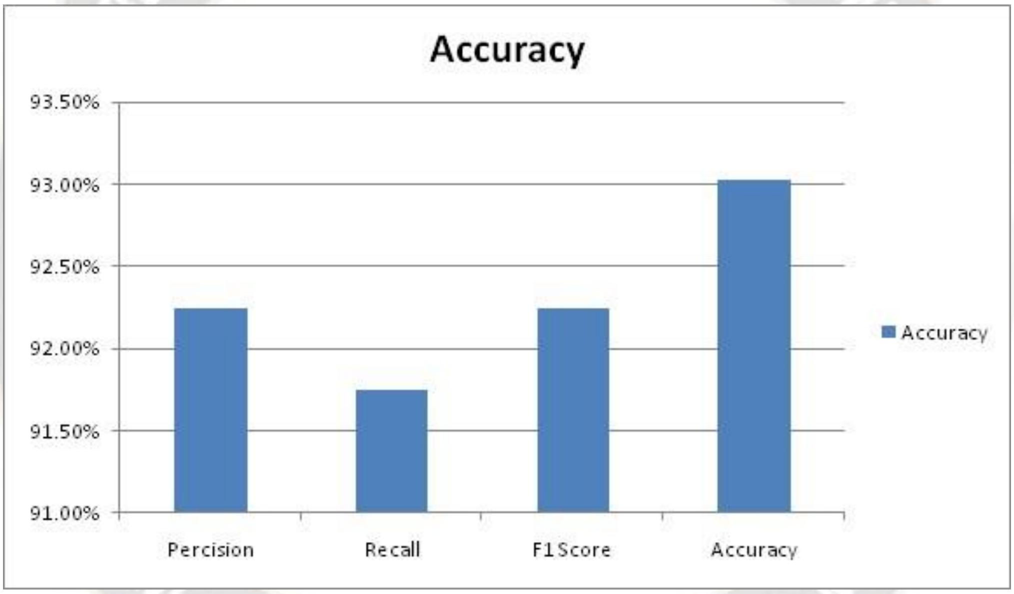


Training and validation loss of Inception-v3

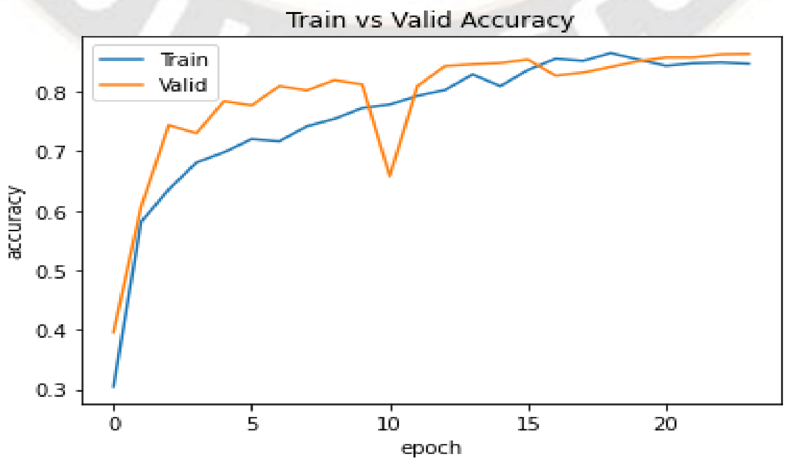


Inception-v3 Outcome



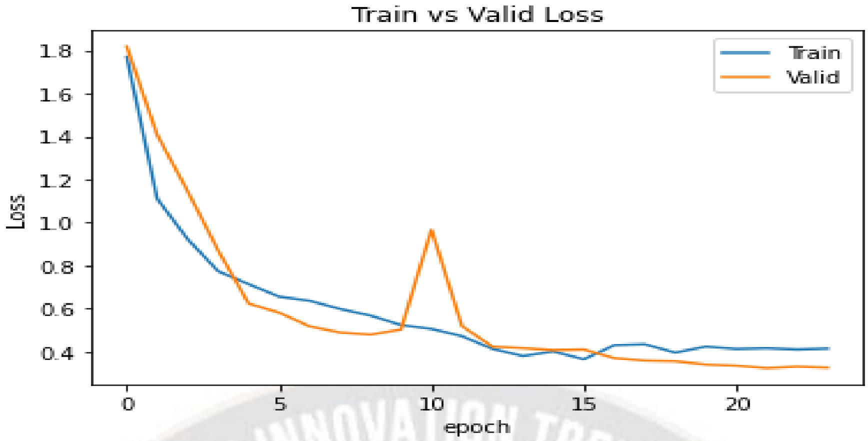


Result of VGG16

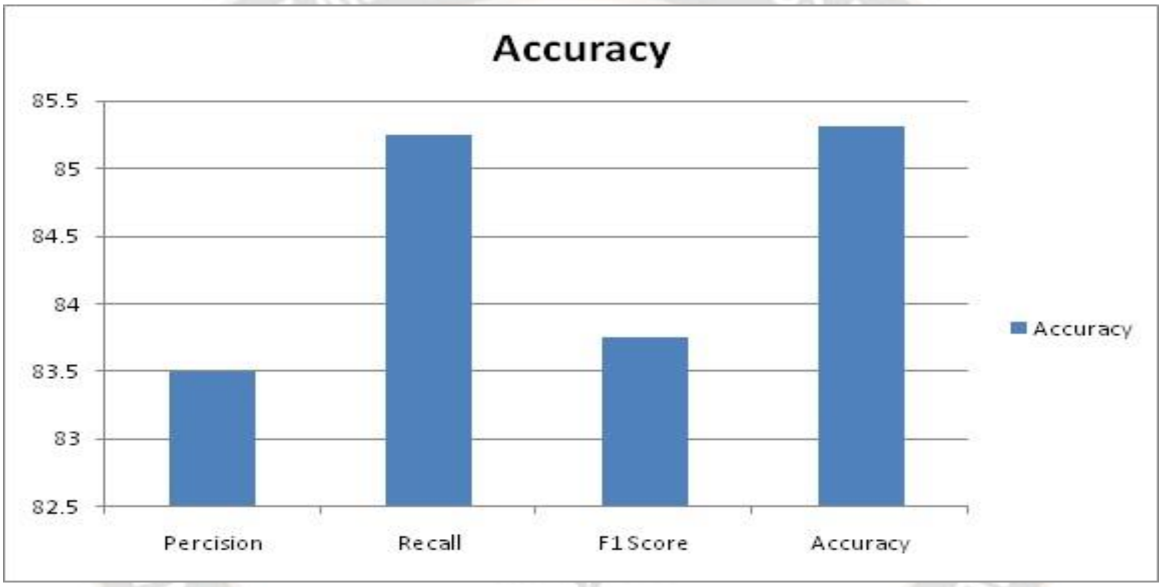


CNN model accuracy during training and validation.

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 9
***Article Received:** 25 July 2023 **Revised:** 12 September 2023 **Accepted:** 30 September 2023*



Loss during CNN model training and validation



Result of CNN

conclusion

We present a deep learning (DL)-based methodology for disease identification in this work, evaluating the performance of different Transfer Learning (TL) algorithms. To make the splitting process easier, the image dataset was first pre-processed. This included resizing, sharpening, rescaling, shearing, zooming, and horizontal flipping. Then, a cotton picture dataset was used to train well-known TL algorithms—namely, VGG-16, VGG-19, InceptionV3, and Xception—for the purpose of illness prediction. Using a test picture dataset, the classifiers' performance was assessed using measures like precision, recall, and F1-measure. In the end, the Xception model—which had the greatest accuracy rate of 98.70%—was selected to create a smart web application for the real-time prediction of cotton illness.

With the ability to facilitate self-validation, consultation, and other agricultural goals, we believe that this tool will be beneficial for agricultural professionals.

Although our DL-based method for detecting cotton disease shows promise, it is not without flaws. Predictions may be impacted by potential problems arising from class imbalance, and more research is necessary to see how well the model adapts to new cases of emerging diseases. Subsequent investigations ought to tackle these constraints to augment the resilience and relevance of the suggested approach.

Limitation of Model

Although we have made great progress with our deep learning (DL)-based method for cotton disease identification, there are still some issues that must be considered. Notably, a class imbalance in the dataset may cause biassed predictions by impeding the model's performance. A further difficulty is the model's flexibility to new disease instances, particularly those that call for segmentation or analysis of a particular leaf region. Moreover, the model may fail to provide critical features required for accurate leaf disease prediction, requiring the creation of feature extraction and selection methods.

Operation in Various Scenarios: The main focus of our research is on the detection of cotton leaf disease using deep learning models. Specifically, we employ VGG-16, VGG-19, Inception-V3, and Xception. Of them, Xception has the

highest accuracy (98.70%). An openly accessible cotton dataset has been used to assess the model's performance.
To showcase its functionality in various situations, we plan to carry out the subsequent actions:
Cross-validation: To confirm the model's resilience and evaluate its performance over various data folds, cross-validation has been carried out on the current dataset. This process helps evaluate the model's ability to adapt to changes in the dataset.
Real-Life Deployment: Testing the model in actual situations is part of a collaboration with farmers and agriculture specialists. Farmers may now upload photos of sick cotton leaves to the web-based smart application and obtain forecasts from the Xception model in the field.

Data collecting from Various Regions: To assess the model's performance in a range of situations, data collecting has been expanded to include a variety of cotton-growing regions with varying environmental conditions and climates. This method aids in the analysis of its capacity to identify illnesses in cotton plants grown in various geographic locations.
Comparison with Other Models: The state-of-the-art disease detection techniques, such as deep learning architectures and conventional methods, have been compared with our suggested Xception model. The purpose of this comparison analysis is to evaluate its superiority and wider application potential.
With these initiatives, we hope to demonstrate the adaptability of our model in a range of contexts, solve its shortcomings, and improve its precision and efficacy for actual cotton disease prediction in agriculture.

References:

[1] Md. Manowarul Islam et al. “A deep learning model for cotton disease prediction using fine-tuning with smart web application in agriculture.” In Intelligent system with Application (Elsevier) 2023.

[2] Mr. Mahesh Shyamsunder Darak “Machine Learning Approach for crop leaf disease detection and classification” Swami Ramananda Teerth Marathwada University Nanded (M.S) India.

[3] P Krithika, S Veni, Leaf disease detection on cucumber leaves using multiclass sup- port vector machine, in: IEEE International Conference on Wireless Communications, Signal Processing and Networking, 2017, pp. 1276–1281.

[4] R Prakash, G P Saraswathy, G Ramalakshmi, Detection of leaf diseases and classification using digital image processing, in: IEEE International Conference on Innovations in Information, Embedded and Communication Systems, 2017, pp. 1–4 S.S. Harakannanavar, J.M. Rudagi, V.I. Puranikmath et al. Global Transitions Proceedings 3 (2022) 305–310

[5] B Mishra, S Nema, M Lambert, S Nema, Recent technologies of leaf disease detection using image processing approach-review, in: IEEE International Conference on Innovations in Information, Embedded and Communication Systems, 2017, pp. 1–5.

[6] C. Puttamadappa, B.D. Parameshachari, Demand side management of small scaleloads in a smart grid using glow-worm swarm optimization technique, Microprocessors Microsyst. 71 (2019) 102886.

[7] Deep Learning Models for Cotton Leaf Disease Detection with VGG-16 Awad Bin Naeem*1, Biswaranjan Senapati2, Alok Singh Chauhan3, Sumit Kumar4, Juan Carlos Orosco Gavilan5, Wael M. F. Abdel-Rehim

[8] V Pooja, R Das, V Kanchana, Identification of plant leaf diseases using image processing techniques, in: IEEE International Conference on Technological Innovations in ICT for Agriculture and Rural Development, 2017, pp. 130–133.

[9] Azath M. , 1 Melese Zekiwos, 2 and Abey Bruck “Deep Learning-Based Image Processing for Cotton Leaf Disease and Pest Diagnosis” Hindawi Journal of Electrical and Computer Engineering Volume 2021, Article ID 9981437, 10 pages Mrs. Shruthi U, Dr. Nagaveni V, Dr. Raghavendra B K “A Review on Machine Learning Classification Techniques for Plant Disease Detection” 2019 5th International Conference on Advanced Computing & Communication Systems (ICACCS) p.282

[10] Azath M., Melese Zekiwos, and Abey Bruck “Deep Learning-Based Image Processing for Cotton Leaf Disease and Pest Diagnosis” in the Article Hindawi Journal of Electrical and Computer Engineering Volume 2021, Article ID 9981437, 10 pages https://doi.org /10.1155/2021/9981437

[11] Jayraj Chopda et al., “Cotton Crop Disease Detection using Decision Tree Classifier” in IEEE conference .

[12] Nikhil Shah et al., “Detection of Disease in Cotton Leaf using Artificial NeuralNetwork” In IEEE 2019

[13] Paramjeet Singh et al., “Cotton Leaf Net: cotton plant leaf disease detection using deep neural networks” The Author(s), under exclusive licence to Springer Science+Business Media, LLC, part of Springer Nature 2023

[14] Smruti Kotian et al., “COTTON LEAF DISEASE DETECTION USING MACHINE LEARNING” in 2nd International Conference on “Advancement in Electronics & Communication Engineering (AECE 2022) July 14-15, 2022.