EAST WEST UNIVERSITY

B. SC. ENGINEERING THESIS

Plant Disease Classification using Deep Learning

Authors:
Radia Ahmed
(**2016-1-50-008**)
Tariqul Islam
(**2016-1-50-017**)
Mir Siam
(**2013-3-50-009**)

Supervisor:
Muhammad Suhail NAJEEB
Lecturer
ELECTRONICS AND
COMMUNICATIONS
ENGINEERING

A thesis submitted in fulfillment of the requirements
for the degree of Bachelor of Science in Information and Communication Engineering

June 2020

Letter of Acceptance

The thesis entitled "Plant Disease Classification using Deep Learning" submitted by Radia Ahmed (ID: 2016-1-50-008), Tariqul Islam (ID: 2016-1-50-017) and Mir Siam (ID: 2013-3-50-009) to the Electronics and Communications Engineering Department, East West University, Dhaka-1212, Bangladesh is accepted as satisfactory for partial fulfilment of requirements for the degree of Bachelors of Science (B. Sc.) Information and Communications Engineering.

Dr. Mohammed Moseeur Rahman
Assistant Professor Chairperson
Department of Electronics and Communications Engineering
East West University

Muhammad Suhail Najeeb
Lecturer
Department of Electronics and Communications Engineering
East West University

Declaration of Authorship

We, Radia Ahmed, Tariqul Islam and Mir Siam, declare that this thesis titled, "Plant Disease Classification using Deep Learning" supervised by Muhammad Suhail Najeeb and the work presented in it are our own. We confirm that:

- This work was done wholly while in candidature for Bachelor of Science in Information and Communication Engineering degree at this University.
- Where any part of this thesis has previously been submitted for a degree or any other qualification at this University or any other institution, this has been clearly stated.
- Where we have consulted the published work of others, this is always clearly attributed.
- Where we have quoted from the work of others, the source is always given. With the exception of such quotations, this thesis is entirely our own work.
- We have acknowledged all main sources of help.
- Where the thesis is based on work done by ourselves jointly with others, we have made clear exactly what was done by others and what we have contributed ourselves.

Radia Ahmed
2016-1-50-008

Tariqul Islam
2016-1-50-017

Mir Siam
2013-3-50-009

Abstract

In a country like Bangladesh plant disease is a common factor. The timely and accurate diagnosis of plant disease plays a very important role in preventing the loss of productivity and loss of reduce quality of agricultural product. Till now many Machine learning (ML) models have been employed for the detection and classification of plant disease. For advancement Deep learning (DL) has also employed in this research area and it has shown a vital impact in disease detection accuracy. In this paper we classified '15' species of crops from 20,069 images. The dataset have been taken from plant village. We have considered here Support Vector machine which is a supervised learning algorithm for classification and regression.We also used Sequential Model for detection. We have used Train-Test-split model to train the dataset and we achieved 92.5% accuracy.

Acknowledgements

We would like to express our most sincere gratitude to our supervisor Muhammad Suhail Najeeb for his guidance, continuous support and advice. He has always been very helpful and inspired us. We have been extremely lucky to have a supervisor who cared about our work, and who responded to our questions and queries so promptly. He was available for our every question and whenever we needed him. Without his help this thesis would remain unsuccessful. We are truly honoured and to have him as our supervisor.

Contents

List of Figures

Dedicated To Our Parents . . .

Chapter 1

Introduction

1.1 Introduction

Bangladesh is an agricultural- dependent country, wherein about 80% of population depend on agriculture .Plant disease are a noteworthy risk to sustenance security. When plants and crops are affected by pests it affected the agricultural production of a country. The developing country like us are totally depends on crops. Crop disease are a major threat to food security in our country. Plant diseases affect the growth of their respective spices. Therefore, their early detection or identification is important. Usually farmers or experts observe the plants with naked eye for detection and identified the disease. This method can be time processing, expensive and inaccurate. Emergence of accurate techniques in the field of leaf-based image classification has shown impressive results. Many machine Learning models have been employed for the detection and classification of plant disease but the subset of Machine Learning which is called Deep learning appears to have a great potential in terms of increased accuracy. In this paper we are created datasets which has two portion, one portion is the pictures of healthy plants and other one are disease plants. And we are able to identify between healthy and diseased leaf from the dataset we created. Our proposed paper includes various phases of implementation namely dataset creation, feature extraction, training the classifier and classification [1].

The farmers may think it's hard to differentiate which may be available in their harvests. And it is difficult for them to go any agricultural business office and discover what the infection may cause in their plants. So, our main purpose is to distinguish the illness introduce in a plant by identifying its morphology by picture handling and deep learning. Knowledge about the pest management or control and diseases are less in various less developed countries like us. Toxic pathogens, poor disease control, drastic climate changes are one of the main factors which increases in diminished food production .

Various present day advancements have developed to limit postharvest preparing, to brace farming maintainability and to augment the profitability. Various laboratory based like chromatography, thermography etc. have been employed for disease detection. However, this exams are not so cost effective but they are high time consuming. Moreover, for developing country like us it is not easy to find this exams available. And the implementation of this experiments are cost effective for the developing country.

So, to reduce such problems many modern approaches such as machine learning and deep learning algorithm has been employed to disease detection, to increase the recognition rate and the accuracy of the results. Here we demonstrate the technical feasibility through deep learning approach using 20,069 images 10 crop spices with 26 both healthy and diseases that are made available in 'PlantVillage 'project. This research paper shows the research work done to solve this problem in the past years and tried to understand how the methodology of image processing detects at an early stage [2].

1.2 Organization of Thesis

This dissertation shows how a realistic model like deep learning model can help the system not only classify the plant diseases but also can detect the diseases. This organization of thesis is as follows.

Chapter 2 introduces the Literature review of classification and plant related work review which had done earlier. Classification which made work ease where physically it take longer to find output. Here, in classification discusses about the detection of pathologic liver using ultrasound images which detect the image of liver using image classification. Then Document classification which classify the characteristics of given set and predicts the documents using Artificial Neural Network. Then discusses the Fault detection and classification and Automated text classification using Artificial Neural Network methods. We also discuss here plant related research papers which has been done earlier. Our main purpose to show here that various classification is used for detection and classification to get the results.

Chapter 3 introduces the theoretical background of various types of learning methods like Machine learning. Machine learning can be supervised or unsupervised or can be reinforcement learning which discusses here in detail. Here we also give a brief of Deep leaning methods. Here we also discuss about types of Neural Networks. Like feed forward, recurrent, convolutional, De convolutional, and modular neural networks. We also discuss about some function like Sigmoid, Relu, Dropout, Softmax, and Maxpooling functions which has been used in out plant disease code.

Chapter 4 introduces plant disease classification task where we discuss about plant both healthy and diseases leaves. And also we put here our EDA models. We also discuss here the diseases which used in our database.

Chapter 5 which is the main part of research paper discuss about the methodology of our work. The introduction of train and test accuracy and also discuss about our code's train and test accuracy. **Chapter 6** shows the results of our code and also we discuss out result. Last

but not the least in **Chapter 7** gives the summery of this dissertation as well as discusses the direction of our future work. Then we includes the references.

Chapter 2

Literature Review

2.1 Classification

Classification is a supervised learning method in which the computer program learns from input data which is given to it and then uses this learning to classify new observation. The data set can be happened in any object like identifying male or female or it may be multi-class too. There are some examples of classification problems like speech recognition, handwriting recognition, biometric identification or document classification, etc.

The classification made the work ease where physically it takes longer to find out the output. Many algorithms can be used for classification like deep learning, machine learning or neural network, etc. In this chapter we are going to describe some object classification that is prepared for paper.

2.1.1 Detection of pathologic liver using ultra sound images

In this paper, one works on human liver images. The motive of this work to develop procedures for the quantitative analysis of ultrasound images to help the steatosis diagnosis. They work on three approaches are presented and tested with human liver images. The first one addresses the textual analysis of the hepatic parenchyma using five classifiers, 357 features, a feature selector, and classifiers fusion, Accuracy, and area under the ROC curve are the two parameters which are using to measure the performance. The second makes use of the hepatorenal coefficient followed by a statistical analysis to discriminate echogenicity differences between liver and kidney. The third one is based on the acoustical attenuation coefficient evaluated over a line traced in the images with a parallel orientation to the acoustical beam. The use of classifiers fusion has provided better results, it shows the accuracy 0.79. The hepatorenal coefficient proved to be a good parameter for steatosis detection with calculated sensitivity and specificity of 0.90 and 0.88, respectively. It was observed that the hepatorenal

coefficient is not influenced by the ultrasound machine parameters. This job has been done by Jaime M.E Santos who is a Biomedical Engineering [3].

2.1.2 Document classification using ANN

The document classification system in the field of data mining in which the format of data is based on bag of words (Bow) or document vector model and the task is to build a machine which after successfully learn the characteristic of given set, predicts the category of the document to which the word vector belongs. In this approach document is represented by BoW where every single word is used as a feature that occurs in a document. The proposed article presents Artificial Neural Network approach which is a hybrid of n-fold cross-validation and training-validation-test approach for classification of data [4].

2.1.3 Fault Detection and Classification using ANN

Process monitoring is considered to be one of the most important problems in process systems engineering. This problem can be reduced significantly from deep learning techniques. In this paper, deep neural networks are applied to the problem of fault detection and classification to illustrate their capability. At first the fault detection and classification problems are formulated as neural network-based classification problems. Then neural networks are trained to perform fault detection, and the effect of two hyperparameters which is the number of hidden layers and number of neurons in the last hidden layer and data augmentation on the performance of neural networks are perused. The fault classification problem is also equipment using neural networks with data augmentation. Finally, the output obtained from deep neural networks are compared with other data-driven methods to illustrate the advantages of deep neural networks [5].

2.1.4 Automatic text classification using artificial neural network

The rising volume of available documents in the World Wide Web (WWW) has turned the document indexing and searching more and more complex day by day. This complex issue has motivated to development of several researchers in the text classification area. The techniques resulting from these researchers require human intervention to choose more enough parameters to carry on the classification. This article presents a new model for the text automatic classification and it's motivated by such limitation. This model uses a self-organizing artificial neural network architecture, which does not claim previous knowledge on the domains to be classified. The document features are radical frequency format, are presented to such architecture, which generates clusters with similar sets of documents. The model

deals with stages of feature extraction, classification, labeling, and indexing of documents for searching purposes. The classification stage receives the radical frequency vectors, submits them to the ART-2A neural networks that classify them and stores them as patterns in clusters, based on their similarity level. The labeling step is liable for extracting the importance level of each radical for each generated cluster. Such importance is used to index the documents which providing support to the next step, comprehend the document searching. The main contribution provided by the proposed model are proposal distance measures to automate the vigilance parameter responsible for the classification quality, this eliminates the need of human intervention on the parameterization process; proposal for labeling algorithm method that extracts the importance level of each word for each cluster generated by ANN networks and create an automated classification methodology [6].

2.2 Plant related research paper

We know that crops are very important for us because most of the country depends on crops. Plant disease causes major production and economic and also a reduction in both the quality and quantity of agricultural products. We all know it's very difficult for the naked eye to identify the disease and do instant steps for preventing crop diseases. Nowadays, plant disease detection received a lot of attention in monitoring a large field of crops. For well developing, there are lots of papers published on plant-related topics. This paper is related to plant disease or crop detection. So in this chapter we are going to give a short description of some plant-related research papers.

2.2.1 Plant disease detection using image processing

This article is based on detecting plant disease. For this they use image preprocessing, features extraction, and neural-based classification. The objective of this is to concentrate on the plant leaf disease detection based on the texture of the leaf. There are five main steps used for the detection of plant leaf disease. The processing scheme consists of image acquirement through a camera or web camera. Image pre-processing which includes image enhancement and segmentation where the most affected and necessary area is segmentation, feature extraction, and classification. And finally the presence of disease on the plant leaf identified. Early information about plants either it is a healthy leaf or disease leaf detection can facilitate the control of disease through proper management strategies. This method will improve the productivity of plants or crops. This paper also compares the benefits and limitations of these potential methods . This article is done by Ms.Kiran R. Gavhale, prof. Ujwalla Gawande,

Dept. of Computer Technology, Yeshwantrao Chavan College of Engineering, India [7].

2.2.2 Detection of plant leaf diseases using image segmentation and soft computing techniques

This paper presents an algorithm known as Genetic Algorithm for image segmentation method which is used for automatic detection and classification of plant leaf diseases. It also shows a survey that covers different disease classification techniques that can be used to lead to disease detection. Here the segmentation process is based on various features found in the image. This might be color, boundaries,or segment of an image. Here use genetic algorithm for color image segmentation.

Genetic algorithm is an algorithm that generates a solution for optimization problems.An algorithm starts with a set of solution which called population. A solution from one population is chosen and then it used to form a new population.This is done with anticipation that means the new form of the population will be enhanced than the old form of solution. The solution that is selected for new form solutions called offspring. And offspring are chosen according to their fitness means the more appropriate they are, the more probability they have to reproduce.

According to paper the disease identification process includes the four main steps: first for the RGB input, color structure is taken, then using a specific threshold value. Then the green pixels are removed, which followed by segmentation process and for getting useful segments the texture statistics are computed. And finally the classifier is used to classify the feature that are extracted. The robustness of the proposed algorithm is proved by using experimental results of about 500 plant leaves in a database [].

This paper motto to detect plant disease through some automatic technique which is beneficial as it reduces a large work of monitoring in big farms of crops and can detect itself the symptoms of the disease at a very early stage. And image segmentation, which is an important aspect of disease detection in plant leaf disease is done by using Genetic Algorithm.

2.2.3 Deep Neural Networks based Recognition of Plant Diseases by Leaf Image Classification

The latest generation of convolutional neural networks has earned impressive results in the field of image classification. This paper is concerned with a new approach to the development of plant disease recognition models, based on leaf image classification by using deep convolutional networks. The right way of training and the methodology used to simplify a quick and easy system implementation in practice. The developed model can identify 13

different types of plant diseases out of healthy leaves with the ability to distinguish plant leaves from their surroundings. Also this method has been proposed for the first time for plant disease recognition.

In this paper they take all the essential steps. First they collect all the images for the dataset and images were grouped into 15 different classes. 13 classes represented plant diseases that could be visually determined from the leaves. The main goal to present study is to train the network to learn the appropriate features that have been increased. Finally a database containing 30880 images for training and 2589 images for validation has been created. Caffe, a deep learning framework developed by Berkly Vision and Learning Centre was used to perform the deep CNN training. The experimental results on the developed model achieved precision between 91% and 98% for separate tests, on average 96.3%.[8]

2.2.4 Detection and classification of plant leaf diseases by using deep learning algorithm

This paper's goal to find and develop the more suitable deep-learning methodology for their task. They proposed a deep learning-based approach to detection leaf diseases using images of plant leaves in different plants. For this, they consider three main families of detectors and those are Faster Region-based CNN (Faster R-CNN), Region-based Fully Convolutional network (R-FCN), and Single Shot Multibox Detector (SSD), which was used for the purpose of this work. The proposed system can effectively-identified different types of diseases and this system have the ability to deal with complex scenarios from a plant's area [9].

Faster R-CNN is one of the object detection system, which is composed of two modules the first module is a deep fully convolutional network that proposes regions. Then the second module is the fast R-CNN detector which uses the proposed regions. The entire process happens in a single unified network, which admitted the system to share full image features with the detection network, thus enabling nearly cost-effective region proposals.

R-FCN method is much similar to faster R-CNN, bust instead of cropping features from the same layer where region proposals are predicted, crops are taken from the last layer of features prior to prediction. The R-FCN is consists of region proposal and region classification.

.l SSD is based on Feed forward network which is able to deal with multiple feature maps with different resolutions. Also SSD encapsulated the process into a single network.

This paper is successful though they did not give any train and test accuracy. This paper is published by the PG student department of CSE.

Chapter 3

Theoretical Background

3.1 Machine learning

Machine learning is a method of data analysis. This learning method is an application of artificial intelligence that provides system the ability to learn automatically and improve from experience without being explicitly programmed. Machine learning focus on the development of computer programs that can access data and use it for themselves. So the system can learn from data, identify the patterns and also make decision with minimal human invention [10].

Today's machine learning is much more developed than past machine learning method. This method was born from pattern recognition and the theory that computers can learn without being programmed to perform specific tasks. The part of AI wanted to see if computer could learn from data. And the iterative aspect is important because as models are exposed to new data, they can adapt independently. They are able to learn from previous computations to produce reliable, repeatable decisions and results. While artificial intelligence is the broad science of mimicking the human abilities, machine learning is a particular subset of AI that trains a machine how to learn [11].

Machine learning algorithm can build a model which based on sample data, known as training data, to make predictions and decisions without being explicit the program. Machine learning is related to statistics that focuses on prediction. Machine learning is related to optimization that delivers methods, theory and applications. Data mining is related field of machine learning that focuses on exploratory data analysis.

The recent development of machine learning is that the ability to automatically apply complex mathematical calculations to big data-over and over, faster than ever. Besides this, algorithm is developed for many applications as like machine perception, translation, adaptive

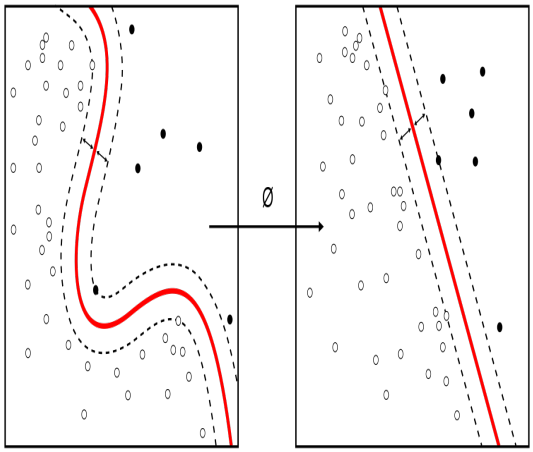

Fig. 3.1. Machine learning and data mining

websites, robot locomotion, sentiment analysis, search engines, speech and voice recognition and so on.

However,the learning approaches of machine learning algorithm go by three criteria, and it given below;

3.1.1 Supervised Machine Learning Algorithm

This learning procedures are simple and supervised. In this section, there is a labeled dataset which the algorithm goes through and gets modify itself until it gets the desired output. Starting from the analysis of a known dataset, the learning produces an inferred function to make predictions about the resultant values. The system can provide targets for any new input after sufficient training. This method can also compare its output values with the correct, intended output and find errors to modify the model accordingly[12].

3.1.2 Unsupervised Machine Learning Algorithm

In this strategy there is no labeled dataset available. Unsupervised learning studies how system can infer a function to describe a hidden structure from unlabeled data. This system does not figure out the right output values but it analyzes the dataset and afterward at that point a cost capacity tells the neural system how distant the target. Then neural network system at that point acclimate the accuracy of the algorithm [13].

There's another learning algorithm which called Semi-supervised machine learning algorithm which fall somewhere in between supervised and unsupervised learning. Usually this

learning is chosen when the acquired labeled data requires skilled and relevant resources in order to train it or learn from it.

3.1.3 Reinforced Machine Learning Algorithm

This system is reinforced for positive result and gets a punishment for negative result. The punishment is forcing the network learn over time. This algorithm is learning algorithm that interacts with its environment by producing actions and discovers errors. Trial and error search and delayed reward are the main characteristics of reinforcement learning. This models allows machines to atomically determine the ideal behavior within specific context for maximizing its performance. Simple reward feedback is required for the machines to learn which action is best [14].

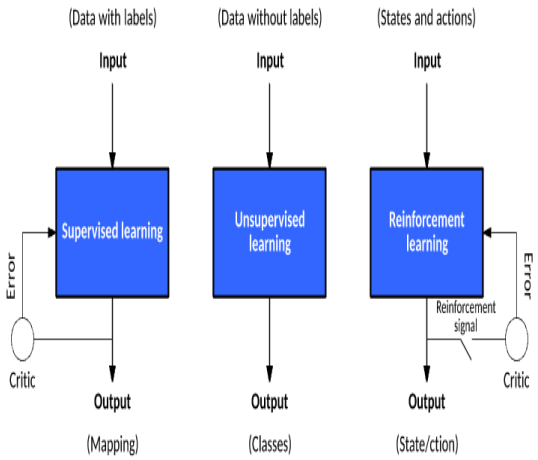

Fig. 3.2. Model showing input and output for supervised,unsupervised and reinforcement learning algorithm

Machine learning enables analysis of massive quantities of data. While it generally delivers faster, more accurate results to identify opportunities as well as dangerous risks. It also require additional time and resources to train the program properly. Combining machine leaning with AI and cognitive technologies can make it even more effective in processing large numbers of information.

3.2 Deep learning

Before doing to solve any problem we need to do know what is Deep learning and how deep learning works?

So, basically Deep learning a sub-field of machine learning which is concerned with algorithms inspired by the structure of the brain called Artificial Neural Networks. Deep learning is a technique that instructs computer to do what comes generally to humans [15].

Deep learning uses multiple layers to gradually eliminate higher level features from initial input. Deep learning is a key technology behind any automated program like driver less car, to distinguish a pedestrian from a sodium light. It's also have the ability to voice control in any kind of devices like phones, tablets, hand free speakers etc.

Deep learning getting huge attention lately and for good reason. Because its achieving the desired outputs that were impossible before. In deep learning, a computer learns how to perform classification tasks directly from any images, sound or text. There are lot of deep learning architectures such as deep neural networks, deep belief networks, recurrent neural networks and convolutional networks .Deep learning have different models which can achieve state of art accuracy. Models are trained by using set of labeled data and those neural network architectures that contain many layers.

In this era, deep learning matters because of its accuracy. Deep learning achieves recognition accuracy at higher than ever before. Two main features that explain us why DL become so useful,

One is deep learning requires large amount of labeled dataset. For an example, when it's about driver less car development these need a big amount of data like millions of images and videos.

The other one is, it requires strong computing power. High performance GPUs have a similar architecture that is efficient for deep learning networks because when combined with cloud computing this enables development to reduce training time from weeks to hours or less.

Now let's have a look in how this deep learning models works –

The term deep refers to the hidden layers in the neural network. But usually neural network contain 2 or 3 hidden layers but where it is about networks it contains large amount of hidden layers as 150 or more. Basically deep networks work as neural networks, so often deep methods called deep neural networks

Deep learning models are trained by using a big amount of sets of labeled data and neural network architectures that learn features directly from the data without the need for manual

feature extraction.

One of the popular deep networks is known as convolutional neural networks in short CNN. CNN eliminate the need for manual feature extraction. By doing so we do not need to identify features used to image classification. The CNN model works extracting the feature directly from image. The relevant features are not pre-determined; they are learned while the network trains on a collection of images. The automated feature extraction makes deep leaning model highly accurate for vision tasks as object classification.

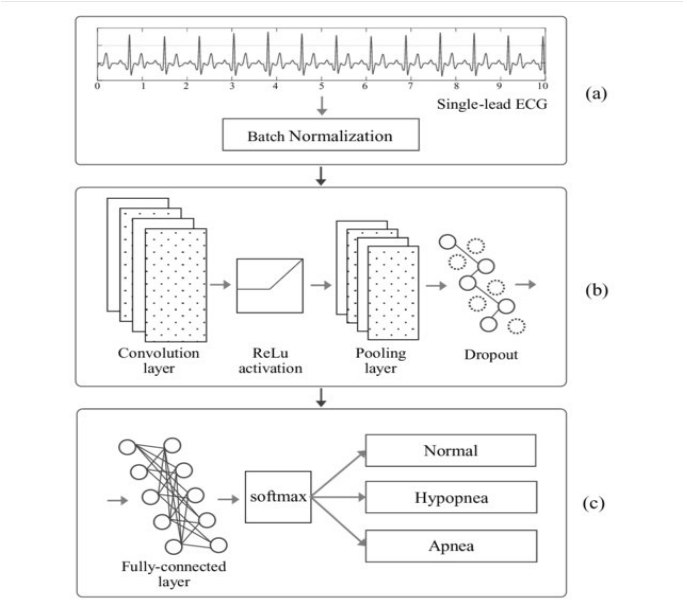

Fig. 3.3. Deep Learning Architecture

There are lots of deep learning applications that are used in industries from automated driving to medical devices [16].

Automated Driving : Automated researchers are using deep learning to atomically detect objects such as stop or traffic signs.

Medical research : Cancer researchers are using deep learning to atomically detect cancer cells.

Industrial Automation : Deep models is helping to worker safety around machinery by automatically detecting when people or objects are within an unsafe distance of machines.

Though deep learning is specialized form of machine learning but there is a difference between them. A machine learning workflow starts with episodic features being manually

extracted with images. Then those features are used to create a model that sort the objects in the image whereas, a deep learning workflow, episodic features are automatically extracted from images.

3.3 Neural Network

A neural network is a kind of artificial model or algorithms that exertions to recognize under the relationship and process a set of data that copy the way human cerebrum operates. A neural network also called Artificial Neural Network [17].Neural network basically are a kind of deep learning technology and all the things falls under the umbrella of AI.

Neural Network was first proposed by Warren McCullough and Walter Pitts in 1944.A neural network works in a system of hardware or software patterned according the operation of neurons in the human brain. The reason why neural network is widely popular is because of its adapting nature. It means neural network can modify themselves as they learn from initial training and posterior runs and provide more information about world. A neural network works exactly like human neuron works. At first it takes the raw input information then training it through processing and gives the best possible result without needing to redesign the algorithm. A neural network takes on layers of interconnected nodes. Each node is a perceptron is similar to a multiple linear regression. A perceptron takes several binary input and produces a single binary output. Perceptron is now more common to use other models of artificial neurons [18].

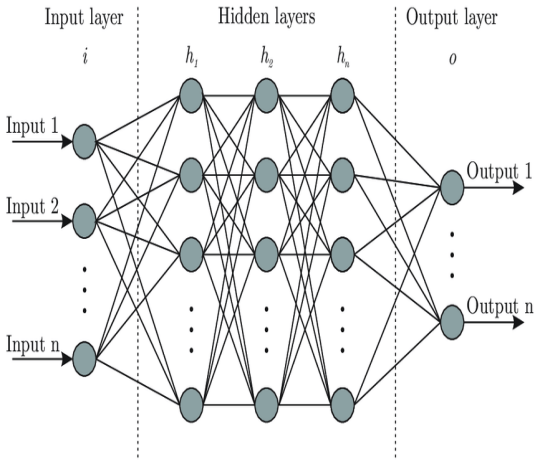

Fig. 3.4. Artificial Neural Network

However, neural network are often described in terms of their intricacy which called model's hidden layers. For this reason neural network is conterminous with deep learning. They can

be described how many hidden nodes or how many input and output node the model has. There are specific types of artificial neural networks and they are given below;

(a) **Feed-forward neural networks**

(b) **Recurrent neural networks**

(c) **Convolutional neural networks**

(d) **Deconvolutional neural networks**

(e) **Modular neural networks**

Neural networks are widely used in commercial purposes like maintain financial operations,trading,enterprise planning ,business analytic and product maintenance .Besides, neural network have earned widely for its adoption in business applications such as marketing research solutions and fraud detection and risk assessment.

Neural network also used for handwriting or pattern recognition, weather prediction and facial recognition.

3.3.1 Feed-forward Neural Networks

FNN are one of the common alternatives of neural networks. This networks takes several inputs and they pass the information through input nodes in one direction. This network may not hidden layers. It passes its information until it makes its desired output node. Basically this networks able to process large amount of noise. This model is used to facial recognition and computer vision [19].

3.3.2 Recurrent Neural Networks

RNN is a more complex class of neural networks where a series of output from the previous step are feeding as input to the current step. Generally in neural networks all the input and output are different or independent from each other but in this case outputs are totally dependent in inputs. Recurrent neural networks are not alike feed-forward networks because RNN allows temporal dynamic behavior. The main feature of RNN is hidden layer. RNN have a memory which remembers all the data about what has been calculated. Basically RNN use their memory to process changeable length sequences of inputs. This makes the RNN

networks more applicable to ought such as handwriting recognition or speech recognition [20].

How RNN works -

Basically RNN converts the independent variable into dependent variable by taking steps the same weights and to all the layers, thus the network reducing the complexity of increasing parameters and also memorizing each previous outputs by giving those outputs as an input to the next hidden layer.

Then these three layers can be joined together and bias of all the hidden layers into a single recurrent layer.

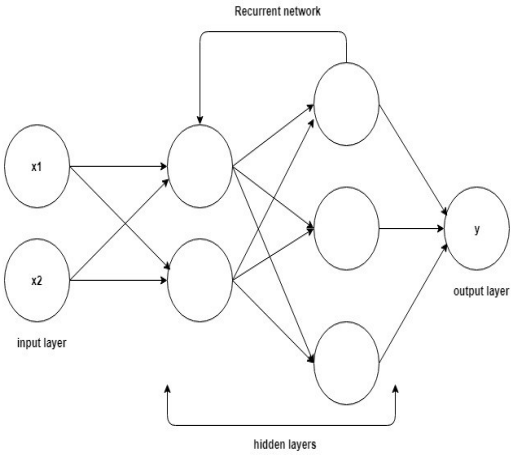

Fig. 3.5. Recurrent model architecture

The main advantage of RNNs that it memorizes each and every information and it is useful in time series prediction, that's why it called Long Short Term Memory.

RNNs is applicable including machine translation, Time series prediction, Speech recognition, Speech synthesis, Human action recognition and so on.

There is another neural networks which is relevant to a recurrent neural network called Recursive neural network [21]. Recursive neural networks are an ordinary form of recurrent neural network. Its can operate on any hierarchical structure. It can parse through input nodes, combining child nodes into parent's nodes and create a tree structure. RNN do the same but the structure is there linear.

3.3.3 Convolutional Neural Networks

From all other networks, CNN are most commonly used networks in deep learning. Because this networks can perform with complexity task with image, sounds, texts and, videos and so on. Convolutional networks basically specialized kind of linear operation. Convolutional networks are generally a class of neural networks that use convolutional in place of common matrix multiplication in at least one of their layers.

A convolutional neural networks basically consist of input, output and hidden layers. This networks share their parameters that's why it also called shift invariant based on their share-parameters architecture. CNNs are orderly versions of multilayer perceptron. It means that they are fully connected to each neurons. The hidden layers are convolve with a multiplication. The Relu layer is an activation function of convolutional networks. This relu layer is followed by addition convolutions such as pooling layers, fully connected layers and normalization layers. All this layers referred as hidden layers and the final part in convolution then it gives the desired result.

Let's describe elaborately the types of layers in below;

Firstly the input layers take the raw input of any 3 dimension image (width, height, depth)

Secondly the convolution layer computes the output by multiplication between all filters and image patch.

Secondly the convolution layer computes the output by multiplication between all filters and image patch.

Third, there some activation function layer like relu, sigmoid applied to activate the function to the output of convolution layer.

Then in pool layer, this layer repeatedly inserted in the convolutional networks and its main feature is to reduce the size of volume which makes the count fast reduces memory and also prevents from overfitting. There are two common types of pool layers are max pooling and average pooling

And finally the fully connected layer, here the layer is general network layer which takes input from the previous layer and gives the desired output.

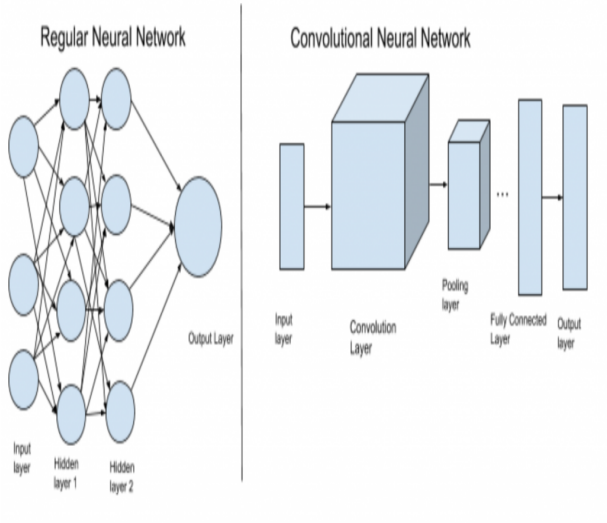

Fig. 3.6. Convolutional Neural Network design

However the convolutional neural networks have applications in image or video recognition, recommender systems, image classification, medical image analysis and neutral language processing[22].

3.3.4 Deconvolutional Neural Networks

This networks basically utilized a reversed CNN models. This models goal to find lost data or features or signals that might have been originally considered less important to the CNN system's task. This model used in image synthesis or analysis.

3.3.5 Modular Neural Networks

This kind of networks contain multiple networks which work separately from each other. This networks do not communicate or interfere with each other while computing the whole process. For this reason a big or complex computational process can be happened more efficiently.

However, recently neural networks are applied in any IOT related topics. Because neural models are so adaptive control. Neural networks have the ability to self-learning. So this self-learning process experience can occur within networks, which can derive them conclusions from a complex dataset and seemingly unrelated set of data.

3.4 Related Concepts

To train deep neural networks with layers is quite challenging because they can be very case sensitive to initial random weights and configuration of the learning algorithm. The distribution of input layers in the networks may change after each epochs when the weights are updated. This happens because the self-learning algorithm is always chase a moving target. This change in the distribution of input layers in the network is known as internal co-variate shift which is a technical name.

Now we will talk about some techniques in below which helps the change in distribution input layers in a network.

3.4.1 Relu Function

In neural network the activation function is responsible for transforming the total weighted input from the node. The rectified linear activation is a linear function that will output the input directly if it is positive value otherwise, it will give the zero output. The Relu function become the default activation function in many neural networks because it is easier to train and most of time it gives better performance [23].

3.4.2 Sigmoid Function

Sigmoid function is a non-linear function and non-linear activation functions are preferred when the nodes are more complex structures in the given data. The Sigmoid function called logistic function and also it is a popular activation in neural networks. The input function is transformed into a value between 0 and 1. When the inputs are larger than 1 then it transformed into the value 1 and when the inputs are smaller than 0 it transformed into the value of 0 [24].

3.4.3 Maxpooling Function

Pooling layers provide a way to down sampling feature maps by summarized the presence of features in patches of the feature map. Basically Maxpooling is a sample emerged discretization process. The object which used is to down sample an input representation. It calculate the maximum value for each patch of the feature map [25]

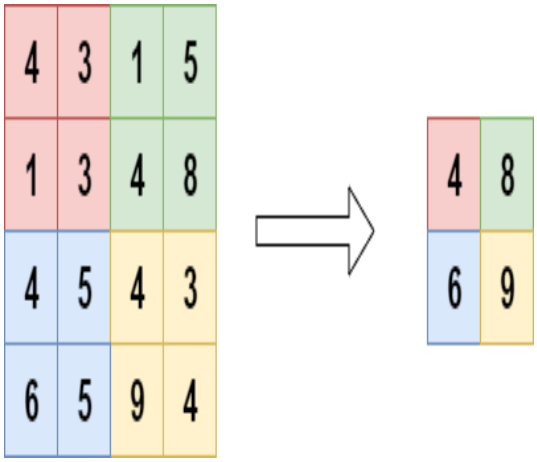

Fig. 3.7. Maxpooling function

3.4.4 Batch-Normalization

To train very deep neural networks a techniques is used called Batch-normalization which standardizes the inputs to a layer for each mini-batch. This has the effect of composing the learning process and reducing the number of training epochs which is required to train deep network models. So, basically batch-norm is technique which standardize the inputs of layers of a network.

The question may arise that why we need to use batch- normalization?

In the upper section we already gave a short description of this topic. So basically for adjusting and scaling the activation we do normalize the input layers. For an example, if we have some features from 0 to 10 and some features from 1 to 1000 then we should normalize the input data to speed up the learning process. By doing this we can see the input layer is benefiting from it, if the input layer is benefitting from it then we can apply the same thing for the values in the hidden layers. But this is challenging all the time and needs to get 10 times more improvement in the training speed.

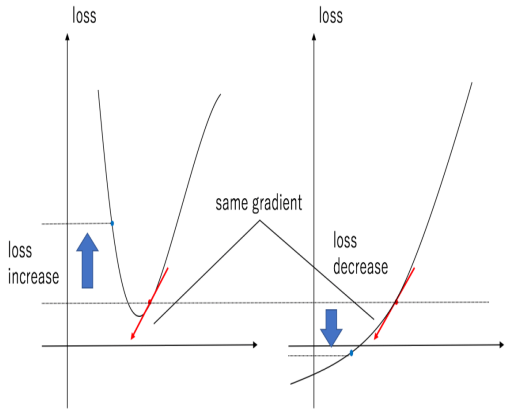

Fig. 3.8. Batch-normalization in Deep learning

Batch normalization reduces the content by what the hidden unit layers values shift around which known as co-variance shift. To explain this if an algorithm has learned X and Y mapping, and if the X distribution changes then we can retain the algorithm by attempting to classify the distribution of X with the distribution of Y.

Moreover, batch-normalization allows each layer to learn by itself more independently from other layers of a network.

How batch-normalization works?

Batch normalization normalizes the output of a previous activation layer by subtracting the mean and dividing the batch standard deviation which increases the stability of a neural network.

After shifting the activation output, the weights in the next layer are no optimal. To minimize the loss function SGD undoes this normalization.

We can use higher self-learning rates because in batch normalization there is no activation's that's gone really high or low. And for this reason, if some previous data that couldn't get to train, it will start to train.

It also reduces overfitting. It adds some error to each hidden layer. So that if we use batch normalization we will use less dropout, which is definitely not a bad thing as we are not going to lose any information [26].

3.4.5 Dropout function

Dropout is a regularization technique that refers to ignoring the units both hidden and visible units in a neural network. Deep neural networks work with a large number of data which is a powerful machine learning system. With this large number of data often over-fitting is a serious problem in these networks. Big and complex networks often slow to use, making it difficult for overfitting [27]. Drop out is a technique for addressing this problem.

This technique randomly drops the units in a neural network during training. More technically, during each training stage, each node is either drop out or kept so that a reduced network is left.

The question may arise that why do we need dropout?

The simple answer can be to these question is to prevent **over -fitting**.

A fully-connected layer occupies most of the parameters and, so neurons are developed co-dependency among each other in the time of training which curbs the individual power of each neurons leading to over-fitting. This prevents units from co-adapting too much [28].

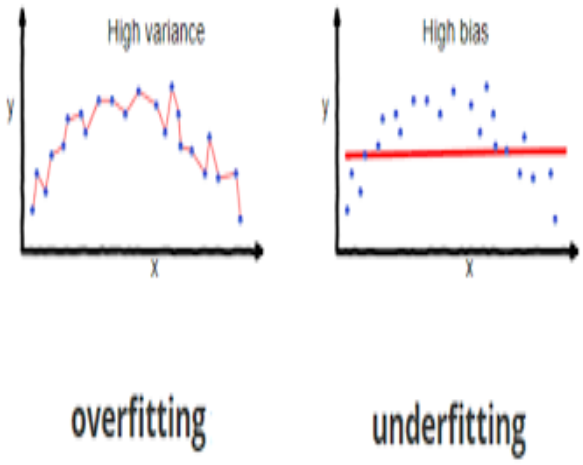

Fig. 3.9. Underfitting and Overfitting in Machine learning

However, dropout improves the performance of neural networks on supervised learning in vision, speech recognition, and document classification, and so on.

3.4.6 Softmax function

Softmax is also an activation function like Relu, Sigmoid, etc. but the sigmoid function can be used easily, the Relu function will not evaporate the effect during the training process. If you want to deal with classification problems, this Sigmoid or Relu function cannot help that much. To be frankly the sigmoid function only can handle two classes which are not what we expect.

The sigmoid function huddles the outputs of each unit to be between 0 and 1 which is similar to Sigmoid function. But it also divides each output so that it added of the outputs is equal to 1.

The output of the softmax function is coequal to a certain probability distribution, it shows the probability that any of the classes are not false.

Softmax often referred to as a multinomial logistic regression because the output values of softmax are between the range [0, 1] which is good as we are able to ignore binary classification and compromise as many dimensions in Neural Network model. In our paper we use softmax function to train our dataset in neural networks[29].

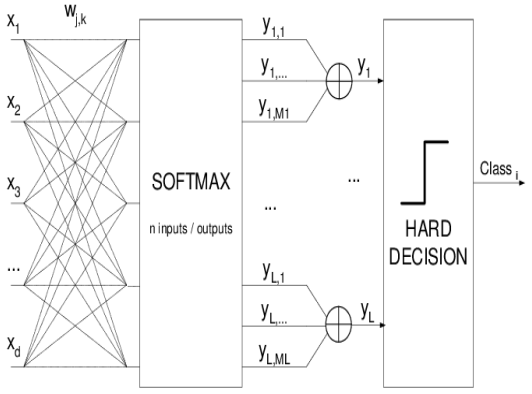

Fig. 3.10. Softmax funtion

Chapter 4

Plant Disease Classification Task

4.1 Plant Diseases

In this paper we classified 15 species of crops. We took this data set from PlantVillage which gave us excess the dataset. Now we are going to give a short description of different plant diseases and also show a comparison between healthy and disease leaf. First of all we will give the visual description that can be shown with the naked eye. We also gave the images which are used from our dataset.

4.1.1 Paper bell bacterial spot

Bell peppers are known as Capsicum annuum are a cultivar group of annual plants in the family Solanaceae grown for their edible fruits. Bell paper may be referred to as red pepper, yellow pepper, or green pepper. Well there are so many common pepper diseases [30].

Bacterial spot symptoms are when Leaf spots that appear on the lower surface of older leaves as small, pimples, and the upper leaf surface as small water-soaked. It also attacks tomatoes. Most probably the spots develop a gray to tan centers with dark borders. When bacterial spot happens leaves may then turn yellow, then brown and drop. Lesions may also develop on stems .bacterial spot have brown or black water-soaked spots on the foliage, sometimes it appears with a yellow halo, generally identical in size.

4.1.2 Tomato bacterial spot

Bacterial spot also attacks tomatoes. Bacterial leaf spot is spread by splashing rain and working with wet which infected plants. This disease can defoliate plants during wet weather. Dry and Hot weather shows the spread of this disease.

Fig. 4.1. Healthy Bell pepper leaf

Fig. 4.2. Bacterial spot pepper leaf

Fig. 4.3. Tomato healthy leaf

Fig. 4.4. Tomato Bacterial spot leaf

4.1.3 Tomato spider mites

The two-spotted spider mites are the most mite species that occur in vegetables particularly tomato. Spider mites prefer the undersides of leaves and it also can occur on both leaf surfaces as well as on the stems and fruits. They suck the sap of plant tissues and infestations in hot and dry conditions .The symptoms are usually clusters of yellow spots on the upper surface of leaves. Feeding by spider mites change the leaf color and turns them bronzer, rusty or yellow-brown color. Spider mites and webbing are present on the lower surface of the leaf, which may appear tan or yellow and have a crusty texture [31].

Fig. 4.5. Two spotted spider mites

4.1.4 Tomato- Septoria leaf spot

Septoria leaf spot is caused by the fungus Septoria lycopersici. The symptoms of septoria spot are quite evident on tomato leaf. It appears on lower mature leaves and causing them to turn yellow and brown and also wither. The diseased leaves show small circular lesions with black and brown margins. The center of these lesions slowly turns into gray, with small black fungal fruiting structures scattered throughout. Lesions may occur on stems and petioles but tend to be more expanded[32].

Fig. 4.6. Septoria spot leaf

4.1.5 Tomato Mosaic Virus

It is a plant pathogenic virus that affects particularly tomatoes and also many other plants. The symptoms of this disease that it affects tomato plants which shows mottling, with alternating yellowish and darker green areas., the latter appearing thicker and raised a blister-like

appearance. The leaves come out a fern-like appearance and also the younger leaves may be twisted[33].

Fig. 4.7. Tomato Mosaic Virus

4.1.6 Tomato Yellow leaf curl virus

The primary host for this virus is the tomato plant. Tomato yellow leaf curl virus (TYLCV) is a DNA virus. TYLCV causes the most destructive disease of tomato. The symptoms of TYLCV infection include severe stunting, reduction of leaf size and upward curling of leaves, chlorosis on leaves. And it also reduction of fruit production[34].

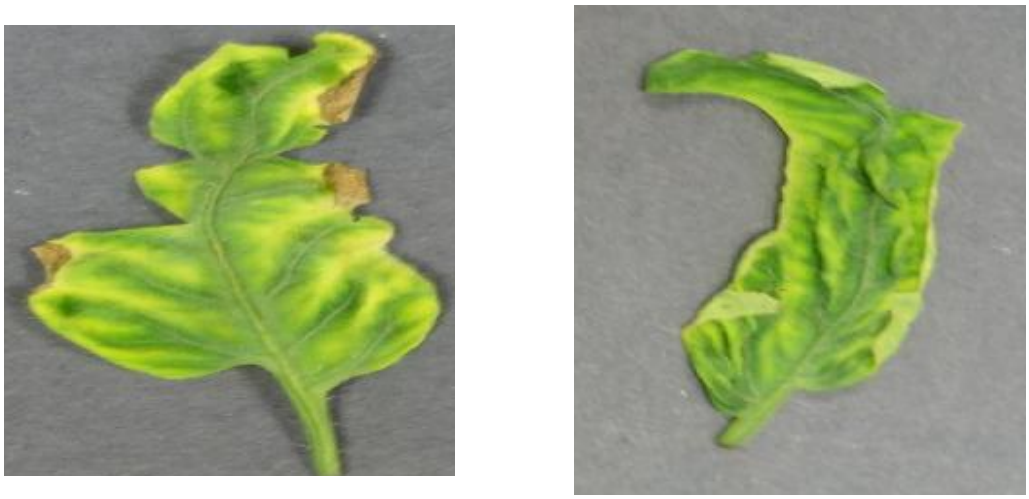

Fig. 4.8. Tomato Yellow leaf curl virus

4.1.7 Early Blight

Early Blight is a plant disease caused by a fungal pathogen known as Alternaria Solani. Early Blight disease affects particularly the tomato and potato plant leaves. The symptoms of early

blight cause dark brown or black spots on the leaves during its initial stages. Later on, the affected leaves will turn into yellow and either dry up or fall off the plant[35].

Fig. 4.9. Tomato leaf early blight

Fig. 4.10. potato leaf early blight

4.1.8 Late Blight

Late blight disease is more devastating than Early Blight and can rapidly destroy the entire tomato and potato plantation. The late blight is caused by Phytophtgthora infestans. The symptoms of late blight is that the lesions are large, firm, irregular and brownish green blotches. Also the lesion surface has a greasy rough appearance[36].

Fig. 4.11. Tomato leaf early blight

Fig. 4.12. potato leaf early blight

4.1.9 Target spot

Target spot is a disease which appear on tomato. Target spot may cause necrotic lesions in a concentric pattern. The symptoms of this appear on all the aboveground parts of the plant. Foliar lesions begin small, water-soaked spots on the upper surface. Gradually it increases in size and becoming circular and pale brown with conspicuous yellow halos [37].

Fig. 4.13. Target Spot (tomato leaf)

4.1.10 Tomato leaf mold

Tomato leaf mold is a well-known fungus that attacks only tomato plants. The symptoms of this disease occurs on foliage. The older leaves are infected first then it moves up towards the younger leaves. The primary symptom appear on the upper surface that infected leaves as small spot green or yellowish and corresponding area of the lower surface. The color of the infected leaves changes to yellowish brown and the leaf begins to curl and dry[38].

Fig. 4.14. Mold Spot (tomato leaf)

So, these are all the diseases that are used to classify in our database. we can successfully classify these diseases using Deep Learning method.

4.2 Dataset

We analyze 20,069 images of plant leaves using Deep learning which is a machine learning process which has interested the machine learning headache engineering [39]. All credit goes to deep learning. Good datasets are needed on all the stages of the object recognition model, from training to testing and accuracy. The dataset was downloaded from the Internet, searched on Kaggle on the name of plant disease detection. Thus, a deep neural network could be trained to differentiate the leaves from the surrounding. We have the dataset where fifteen types of diseases have been tested and trained. Duplicated images taken from different sources were removed by developed python script applying the comparing procedure. We have used the PlantVillage dataset in color then we with a gray-scaled version of the PlantVillage dataset and run the experiments of the PlantVillage dataset where leaves were segmented, hence removing all the extra background information. That might cause the introduction of some inherent bias on the dataset because of the regularized process. The script removed the duplicates by comparing the images' metadata: name, size, and date. After that, we run automated removal and augmented images. Described in this paper, we resize the images to 256 256 pixels, and we perform both the model optimization and predictions on these down-scaled images [40].

The main goal of the presented study is to train the network to learn the features that distinguish one class from the others. Therefore, when using more augmented images, the chance for the network to learn the appropriate features has been increased. This will perform on new unseen data, and also to keep a track of datasets are overfitting, we run all our experiments across a whole range of train-test set splits, namely 80–20 (80% of the whole dataset used for training, and 20% for testing).

Chapter 5

Methodology

The identification of crop diseases can be detected by two methods direct and indirect. Here direct detection is molecular and serological methods that are used for the analysis of high-throughout when there is a large number of samples that are needed to analyze. In these methods, the disease is causing pathogens like bacteria, fungi, or viruses directly detected for providing an accurate number of the disease. Indirect methods identify plant diseases via different parameters like morphological, temperature, transpiration rate, and volatile organic compounds released by infected plants.

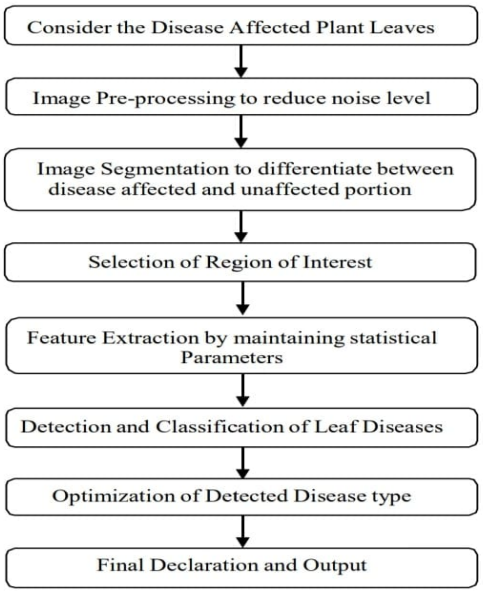

Fig. 5.1. Flowchart and proposed work

Images were downloaded from the Internet where it was in various formats, different resolutions, and quality. To get a better feature from the final images we intended to use as a dataset. We preprocessed the dataset to gain consistency. The procedure of image preprocessing consists of cropping, making the square around the leaves. we have resized the

images 256*256 from the Dataset. Resized to reduce the time of training, which was computed by the written script in Python, using OpenCV framework. Then converted the image into arrays. By using the ImageDataGenerator method array images have been shaped. The features were extracted which is diseased and which is not. The detected with diseases plants will detect which part is detected and what kind of disease got which signs. After the detection of diseased and non-diseased leaves and getting the leaves which got what signs. We have used a sequential method from Keras on our plant disease detection project. In the sequential method, we have added Conv2D, Activation, BatchNormalization, MaxPooling2D, Dropout methods. By using this we get an accuracy of 94.17%.

Support vector machine which is a supervised computing device learning algorithm which can be used for classification and regression. The classifier will assign the label to the photograph and it specifies which category it belongs to, from where the classifier is predefined primarily based upon the feature. This classification is used for each analysis and the trying out phase. SVM makes use of the method referred to as the kernel. Artificial neural networks (ANN) take input as an unstructured image to apply a computational model that works on and converts them into corresponding classification output labels. It requires less preprocessing efforts that can be trained to learn the required features for classification purposes [41].

Chapter 6

Results and Discussion

6.1 Train Test and split

When we are working on a model it is important to train the model because we have datasets. After training, we also want to test the datasets and for this, we need a different dataset that is different from our dataset. Then we need to do split the dataset.

Before going to do train test and split we are giving a short description about train, test and split dataset.

What is training – If a dataset of examples used for learning called a training dataset which is to fit the parameters. For an example, a classifier

What is test dataset - If a dataset is independent of the training dataset called a test dataset. But that dataset follows the same probability distribution as the training dataset. That's means if a model fit to the training dataset also fits the test dataset.

What is validation dataset - A validation dataset is a set of datasets that used to tune the hyperparameters (that is, a parameter whose value is set before the learning process begins) of a classifier.

We usually fit the model on the training set to make predictions on random data that was not trained. The train and test are two important concepts in data science and data analysis and also are used to prevent overfitting.

In machine learning, we split our data into two subsets and that are training data and testing data, and sometimes validate. When we fit a model in training test one of two might be happened. Sometimes we overfit our model or we underfit out model. And we don't want any of these because both of affect the predictability of our model.

In machine learning, Overfitting means that model we trained has trained 'too well', fit too closely to the training dataset. This happens when the model is complex. This model can be very accurate on the training data but may not accurate on untrained data. Because this model is not generalized that is it can generalize the result but not make any interferences on the other data. This leans the noise or error in the training dataset instead of the actual relationships between variables in the data. This noise is not the part of any new dataset and also it cannot be applied [42].

When a model is underfitted, it means that the model does not fit the training data and therefore it misses the trends in the data. It also cannot be generalized to new data. This is usually the result of a very simple model.

So we want to avoid those problems in data analysis. Train test split and cross validation help to avoid overfitting and also uderfitting

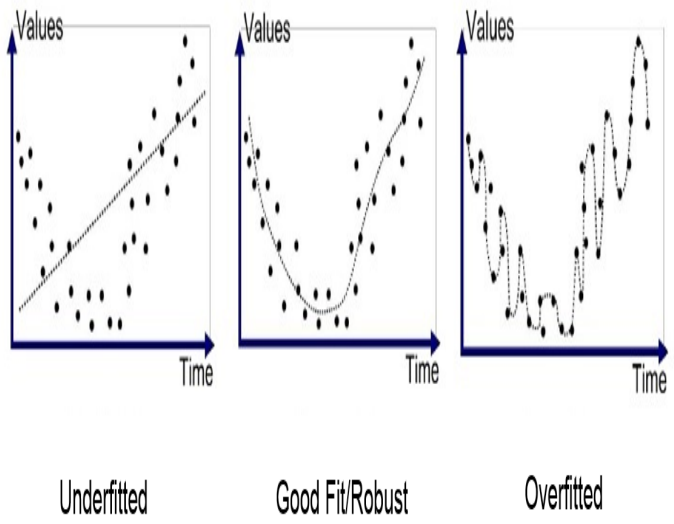

Fig. 6.1. Overfit and Underfit

So the data which we use in our model is usually split into training data and test data. A known output is contained by the training set and the model learns on this training data to be generalized to other data later on. We also have the test data or subset which test our models prediction on this subset.

Now the question comes, how do we split the data?

We have to make sure that we split the data in a random manner and it is possible to manually split ta data into two. For this, there's a library called Model Selection library which is provided by Scikit library. There's a class in the library named train test split. By using this we can easily split the dataset into training data and the testing data. Before we use this class we need to know about some parameters–

Test-size : this parameters decides the size of the data that has to be split as test dataset. The value is given as a fraction. For an example, if we give the value 0.5 then the dataset will be split 50% as the test dataset.

Train-size : if we are not specifying the test-size then we have to specify this parameter. This is quite same as test-size parameter but instead we tell the class what percent of the dataset we want to split as the training set.

Random-size : this parameter pass an integer number instead of passing a fraction number. During split, it act as the seed for the random number generator. Or, we can also pass am instances RandomState class which will become the number generator.

We can also perform Cross-Validation if we don't want to use the train test/split. It is very similar to the train and test method but it's applied to more subsets. That's means, we have to split our data into k subsets and train on k-1 one of those subsets. There is a difference between the validation dataset and test dataset that the validation dataset used to perform the evaluation of models when truing hyperparameters and data preparation and the test dataset is used to describe the evaluation of a final tuned model when comparing it to other final models[43].

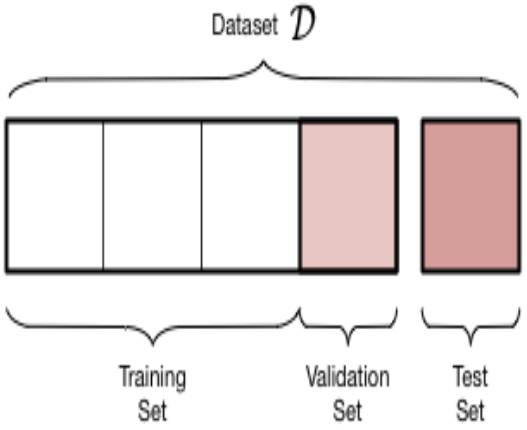

Fig. 6.2. A Visualisation of split

6.1.1 Result of our paper

For getting result we use here a buildup function called train-test-spilt. We also use here a sequential method which helps our data to train. The result we presented in this section is related to training with the database both original and augmented images. As we know that, deep learning algorithms are able to learn features when trained on larger datasets.

After using the train test and split method, an overall accuracy 92.5% was achieved. After training iteration we were achieved high accuracy results with exceedingly reduced loss.

To show effectiveness of the proposed model, a series of experiment have been performed on the available leaf dataset. We randomly select 80% (0.8) of the dataset for training and 20% (0.2) for testing. All the images are doubled and resized to 256*256 pixels. Per-pixel value is divided by the maximum value and subtracted the mean values of the data. In figure 6.1, the blue line shows the training accuracy which has got higher accuracy and the red line which represents validation accuracy.

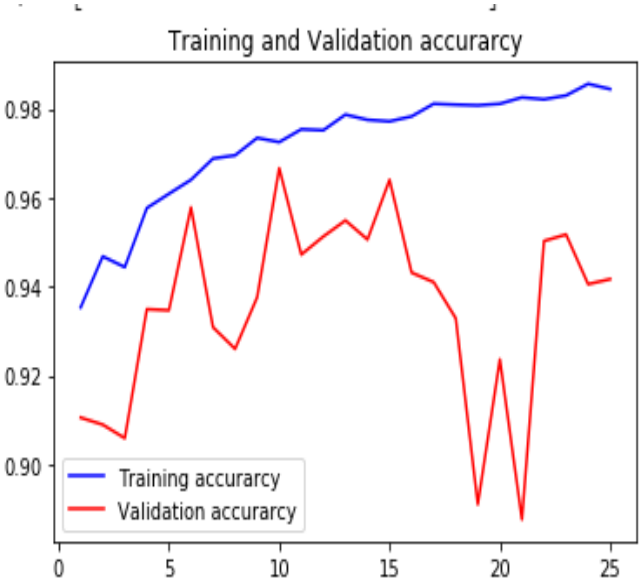

Fig. 6.3. Training and Validation accuracy

In this graph, the blue line represents the loss during the training stage. Through training iterations, the loss was rapidly reduced. The red line in the graph represents the success of the validation test set, through training iterations.

Furthermore, the trained model was tested on each class individually. the test was performed on every image from the validation set. The results are displayed to emphasize how many images from the total of each class are accurately predicted.

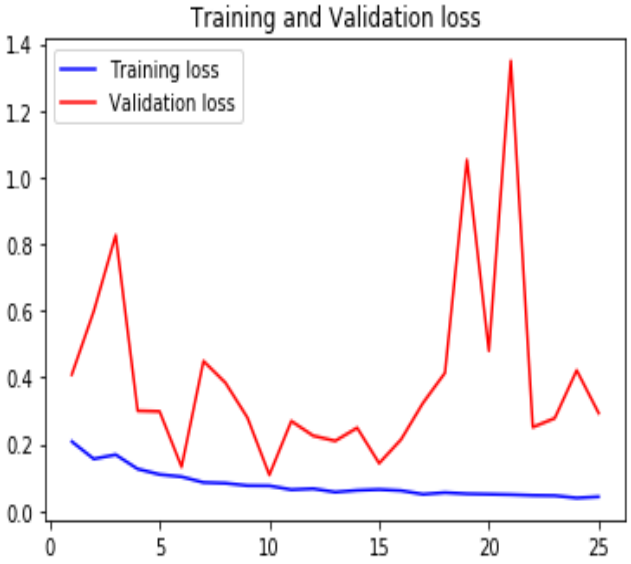

Fig. 6.4. Training and Validation loss

From the results displayed in figure 6.3 , it is notable that trained model accuracy was slightly less for classes with a lower number of images in the training dataset.

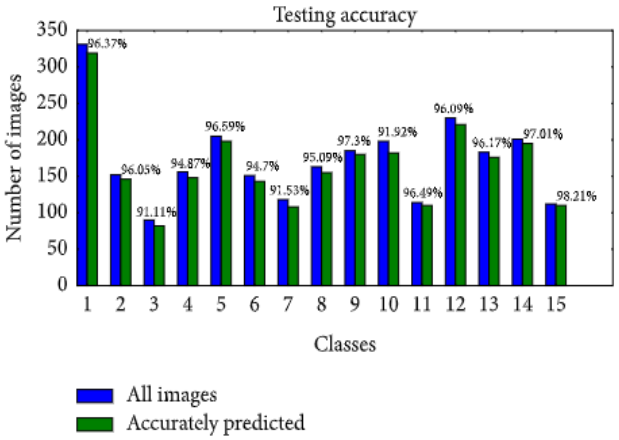

Fig. 6.5. Bar chart of Test accuracy

As a result of extensive review, Deep Learning techniques have shown better results in pattern recognition, in image segmentation and object detection. Presently, there is a commercial solution, Leafsnap, which uses visual recognition in order to identify tree species from their leaves' images but as the Deep Learning method presented in this paper is classifying the plant diseases instead of types of plant, Leafsnap was not used for comparison of the achieved results. Finally, comparing our results with other methods of detecting diseases from leaves images, it can be said that our method provides better results.

Chapter 7

Conclusion and Future Scope

This paper represents the survey, the reviews and the summarizes on different plant disease classification techniques used for Plant Disease Detection using Deep learning techniques that can be used for automatic detection as well as classification of plant leaf diseases. Pepper bell, potato, tomato are some of the spices on which deep learning method is tested. Therefore, the related disease for these plants was taken for identification. With very little computational effort we can obtain the desired result. The method also shows the efficiency in recognition and classification of the leaf disease. The advantage of using this method is that plant diseases can be identified at a very early age which is the main goal of ours. Still, there is a scope of improvement in the existing research. We used here the very least amount of disease.

So our future hope is to work with a big amount of plant leaf. Also to improve the recognition rate in the classification process Artificial Neural Network, Bayes classifier, Fuzzy Logic, and Hybrid Algorithm can be used.

Bibliography

[1] S. P. Mohanty, D. P. Hughes, and M. Salathé, "Using deep learning for image-based plant disease detection," *Frontiers in Plant Science*, vol. 7, no. September, sep 2016.

[2] Google.com, "plant disease detection using deep learning research paper - Google Search." [Online]. Available: https://www.google.com/search?q=plant+disease+detection+using+deep+learning+research+paper&c 8

[3] J. Santos, J. S. Silva, A. A. Santos, and P. Belo-Soares, "Detection of pathologic liver using ultrasound images," *Biomedical Signal Processing and Control*, vol. 14, pp. 248–255, nov 2014.

[4] Researchgate.net, "(PDF) Document Classification using Artificial Neural Networks." [Online]. Available: https://www.researchgate.net/publication/334050771_Document_Classification_using_Artificial_Neu

[5] S. Heo and J. H. Lee, "Fault detection and classification using artificial neural networks," *IFAC-PapersOnLine*, vol. 51, no. 18, pp. 470–475, jan 2018.

[6] R. F. de Mello, L. J. Senger, and L. T. Yang, "Automatic text classification using an artificial neural network," in *IFIP Advances in Information and Communication Technology*, vol. 172. Springer New York LLC, 2005, pp. 215–238.

[7] M. K. R. Gavhale and P. U. Gawande, "An Overview of the Research on Plant Leaves Disease detection using Image Processing Techniques," *IOSR Journal of Computer Engineering*, vol. 16, no. 1, pp. 10–16, 2014.

[8] S. Sladojevic, M. Arsenovic, A. Anderla, D. Culibrk, and D. Stefanovic, "Deep Neural Networks Based Recognition of Plant Diseases by Leaf Image Classification," *Computational Intelligence and Neuroscience*, vol. 2016, 2016.

[9] Ijert.org, "Detection and Classification of Plant Leaf Diseases by using Deep Learning Algorithm – IJERT." [Online]. Available: https://www.ijert.org/detection-and-classification-of-plant-leaf-diseases-by-using-deep-learning-algorithm

[10] Machine_learning, "Machine learning - Wikipedia." [Online]. Available: https://en.wikipedia.org/wiki/Machine_learning

[11] A. M. Tillmann, "On the computational intractability of exact and approximate dictionary learning," *IEEE Signal Processing Letters*, vol. 22, no. 1, pp. 45–49, 2015.

[12] Wikipedia.org, "Supervised learning - Wikipedia." [Online]. Available: https://en.wikipedia.org/wiki/Supervised_learning

[13] T. Hastie, R. Tibshirani, and J. Friedman, "Unsupervised Learning." New York: Springer, 2009, pp. 485–585. [Online]. Available: http://link.springer.com/10.1007/978-0-387-84858-7_14

[14] V. François-Lavet, P. Henderson, R. Islam, M. G. Bellemare, and J. Pineau, "An introduction to deep reinforcement learning," *Foundations and Trends in Machine Learning*, vol. 11, no. 3-4, pp. 219–354, dec 2018.

[15] Mathworks.com, "What Is Deep Learning? | How It Works, Techniques & Applications - MATLAB & Simulink." [Online]. Available: https://www.mathworks.com/discovery/deep-learning.html

[16] L. Deng and D. Yu, "Deep learning: Methods and applications," pp. 197–387, 2013. [Online]. Available: http://research.microsoft.com/pubs/209355/DeepLearning-NowPublishing-Vol7-SIG-039.pdf

[17] Techtarget.com, "What is an Artificial Neural Network (ANN)?" [Online]. Available: https://searchenterpriseai.techtarget.com/definition/neural-network

[18] Techradar.com, "What is a neural network? | TechRadar." [Online]. Available: https://www.techradar.com/news/what-is-a-neural-network

[19] Ieeexplore.ieee.org, "Feed-forward neural networks - IEEE Journals & Magazine." [Online]. Available: https://ieeexplore.ieee.org/abstract/document/329294

[20] Wikipedia.org, "Recurrent neural network - Wikipedia." [Online]. Available: https://en.wikipedia.org/wiki/Recurrent_neural_network#Applications

[21] Towardsdatascience.com, "Recurrent Neural Networks - Towards Data Science." [Online]. Available: https://towardsdatascience.com/recurrent-neural-networks-d4642c9bc7ce

[22] Y. Kim, "Convolutional neural networks for sentence classification," in *EMNLP 2014 - 2014 Conference on Empirical Methods in Natural Language Processing, Proceedings of the Conference*. Association for Computational Linguistics (ACL), 2014, pp. 1746–1751.

[23] Machinelearningmastery.com, "A Gentle Introduction to the Rectified Linear Unit (ReLU)." [Online]. Available: https://machinelearningmastery.com/rectified-linear-activation-function-for-deep-learning-neural-networks/

[24] Deepai.org, "Sigmoid Function Definition | DeepAI." [Online]. Available: https://deepai.org/machine-learning-glossary-and-terms/sigmoid-function

[25] Geeksforgeeks.org, "CNN | Introduction to Pooling Layer - GeeksforGeeks." [Online]. Available: https://www.geeksforgeeks.org/cnn-introduction-to-pooling-layer/

[26] Towardsdatascience.com, "Batch normalization in Neural Networks - Towards Data Science." [Online]. Available: https://towardsdatascience.com/batch-normalization-in-neural-networks-1ac91516821c

[27] N. Srivastava, G. Hinton, A. Krizhevsky, and R. Salakhutdinov, "Dropout: A Simple Way to Prevent Neural Networks from Overfitting," Tech. Rep., 2014.

[28] Medium.com, "Dropout in (Deep) Machine learning - Amar Budhiraja - Medium." [Online]. Available: https://medium.com/@amarbudhiraja/https-medium-com-amarbudhiraja-learning-less-to-learn-better-dropout-in-deep-machine-learning-74334da4bfc5

[29] Towardsdatascience.com, "The Softmax Function, Simplified - Towards Data Science." [Online]. Available: https://towardsdatascience.com/softmax-function-simplified-714068bf8156

[30] Extension.umd.edu, "Bacterial Leaf Spot of Pepper - Vegetables | University of Maryland Extension." [Online]. Available: https://extension.umd.edu/hgic/topics/bacterial-leaf-spot-pepper-vegetables

[31] Gardeningknowhow.com, "Spider Mite Treatment: How To Identify Spider Mite Damage And Kill Spider Mites." [Online]. Available: https://www.gardeningknowhow.com/plant-problems/pests/insects/spider-mite-treatment.htm

[32] Hortnews.extension.iastate.edu, "Tomato - Septoria Leaf Spot | Horticulture and Home Pest News." [Online]. Available: https://hortnews.extension.iastate.edu/2000/8-11-2000/septoria.html

[33] Wikipedia.org, "Tomato mosaic virus - Wikipedia." [Online]. Available: https://en.wikipedia.org/wiki/Tomato_mosaic_virus

[34] Sciencedirect.com, "Tomato Yellow Leaf Curl Virus - an overview | ScienceDirect Topics." [Online]. Available: https://www.sciencedirect.com/topics/neuroscience/tomato-yellow-leaf-curl-virus

[35] Backgarden.org, "Tomato Early Blight Disease - Causes, Symptoms and Treatment." [Online]. Available: https://backgarden.org/tomato-early-blight/

[36] G. L. Schumann and C. J. D'Arcy, "Late blight of potato and tomato," *The Plant Health Instructor*, 2000.

[37] C. Cassiicola, "Tomato Diseases Target Spot," Tech. Rep. [Online]. Available: www.avrdc.org

[38] Wikipedia.org, *Tomato Leaf Mould*. Plantwise Knowledge Bank. [Online]. Available: http://www.plantwise.org/KnowledgeBank/Datasheet.aspx?dsid=24580

[39] Kaggle.com, "PlantVillage Dataset | Kaggle." [Online]. Available: https://www.kaggle.com/emmarex/plantdisease?fbclid=IwAR1cJsCtFUgbCeyic7Nf4j7Jz7fDG_v7TZ

[40] Hindawi.com, "Deep Neural Networks Based Recognition of Plant Diseases by Leaf Image Classification." [Online]. Available: https://www.hindawi.com/journals/cin/2016/3289801/?fbclid=IwAR0BieQZkHpa1HwIkpmeimvTNF o1saJiOOPgBnoW6zd48NesbxhkI

[41] M. H. Saleem, J. Potgieter, and K. M. Arif, "Plant disease detection and classification by deep learning," p. 468, nov 2019.

[42] Machinelearningmastery.com, "What is the Difference Between Test and Validation Datasets?" [Online]. Available: https://machinelearningmastery.com/difference-test-validation-datasets/

[43] Medium.com, "How to split your dataset to train and test datasets using SciKit Learn." [Online]. Available: https://medium.com/@contactsunny/how-to-split-your-dataset-to-train-and-test-datasets-using-scikit-learn-e7cf6eb5e0d