B.Tech (VII sem) Project Work-1

<Tomato Leaf Disease Prediction >

A Synopsis

for

Project Work-1

BACHELOR OF TECHNOLOGY in COMPUTER SCIENCE & ENGINEERING

BY

Divya Purohit – EN20CS303021
Jasnoor Singh Mac – EN20CS303029

Under the Guidance of
<Name of the Guide(s)>

Department of Computer Science & Engineering
Faculty of Engineering
MEDI-CAPS UNIVERSITY, INDORE- 453331

<2023 >

INTRODUCTION

1.1 PREAMBLE

Precision farming is the next step in the evolution of agriculture. Precision agriculture may boost agricultural output by combining science and technology. Precision farming also entails reducing pesticides and illnesses by accurately calculating the number of pesticides needed. Precision farming has improved several agriculture sectors as it transitions from conventional ways to new approaches. Precision farming's only objective is to obtain real-time data to boost agricultural yield and maintain crop quality.

Agriculture is much more than just a means of feeding the world's growing population. Plant diseases have also cost agricultural and forestry businesses a lot of money. As a result, early identification and diagnosis of plant diseases are crucial to take fast action. Plant illness may be detected using a variety of techniques. Certain ailments, on the other hand, are difficult to diagnose early on.

They will have to wait a little longer to figure out what is going on. Advanced analysis, which is often done with powerful microscopes, is necessary under these conditions. Diseases wreak havoc on a plant's overall health, slowing its development. Unfortunately, a plethora of tomato diseases is wreaking havoc on the leaves of the crop.

The proposed study's primary objective is to find a solution to the problem of identifying tomato leaf disease using the most straightforward technique feasible while utilizing the fewest computer resources necessary to achieve results comparable to state-of-the-art alternatives. In addition, to assist in classifying input photos into sickness classifications, automatic feature extraction is applied. Consequently, the suggested system attained an average accuracy of 94%-95%, demonstrating the neural network approach's viability even under challenging scenarios.

Sensors and remote sensing, mapping and surveying, high-precision positioning systems, variable rates, the global navigation satellite system, automated steering systems mapping, computer-based applications, and other technologies are all used in precision agriculture. In addition, precision agricultural concepts based on infrared variation analysis and treatment are also cutting-edge technologies. This examination utilizes a more modest adaptation of the convolutional neural organization model to recognize and analyze messes in tomato leaves.

In other cases, the signals can only be detected in non-visible electromagnetic spectrum areas. This research aims to develop a user-friendly method that will aid farmers in recognizing tomato plant issues without having to consult an expert. We first obtain a picture from the Kaggle dataset, from which we extract characteristics. To extract the attributes, we employ picture conversion and scaling. To diagnose diseases, the transfer learning inception v3 model will be used.

Fig 1.1. ILL Tomato Leaf [12]

1.2 Image Processing in Precision Agriculture

Precision agriculture uses deep learning techniques, and its approach to crop protection effectively boosts crop development. Image analysis may be used to detect the sick leaf and measure and locate the damaged area's border to identify the item appropriately. This study develops an improved deep learning system for determining the state of a tomato crop based on a photo of its leaves.

We all know that the human brain recognizes images far faster than a computer. However, the era has changed with the introduction of Machine Learning. High performance on image identification tasks may be ensured using the model, deep convolution neural network, which can surpass human performance in several domains. By certifying their work against Image Net, researchers have achieved advances in the field of visual recognition.

1.3 SYMPTOMS OF TOMATO LEAF DISEASE:

The plant's color, shape, and function may vary as it responds to the illness. We'll go through the signs and symptoms of these illnesses, as well as what to check for if your plant's development appears to be slow. The appropriate classification and diagnosis of leaf diseases are crucial for reducing agricultural losses. Different plant leaves transmit various diseases and show multiple symptoms.

Leaf Spot on Septoria:

One of the most prevalent tomato plant leaf diseases is the septoria leaf spot. A tiny, round patch with a greyish-white center and black borders is the first sign of this fungus' presence. In the center, tiny black specks may appear. The leaves of sensitive tomato plants become yellow, wither, and fall off due to prolonged exposure to hot, humid conditions.

Early Blight (Alternaria):

Alternaria is a parasite that causes tomato leaf spots and an early curse. On the lower leaves, brown or dark regions with dim edges arise, practically like an objective. Organic product stem closes are exceptionally touchy, creating tremendous, profound dark blotches with concentric rings. A fungus causes this tomato plant disease, which appears after the plants have produced fruit.

Blight in the Late Stages:

Late blight, a tomato plant disease caused by the fungus Phytophthora infestans, arises in perfect, wet conditions after the growth season. Frost damage with uneven green-black splotches emerges on plants. Fruits with large, irregularly shaped black areas can quickly be destroyed. This fungus, which causes tomato plant disease, also affects potatoes and can be transmitted through them. The same precautions should be used as with septoria leaf spot.

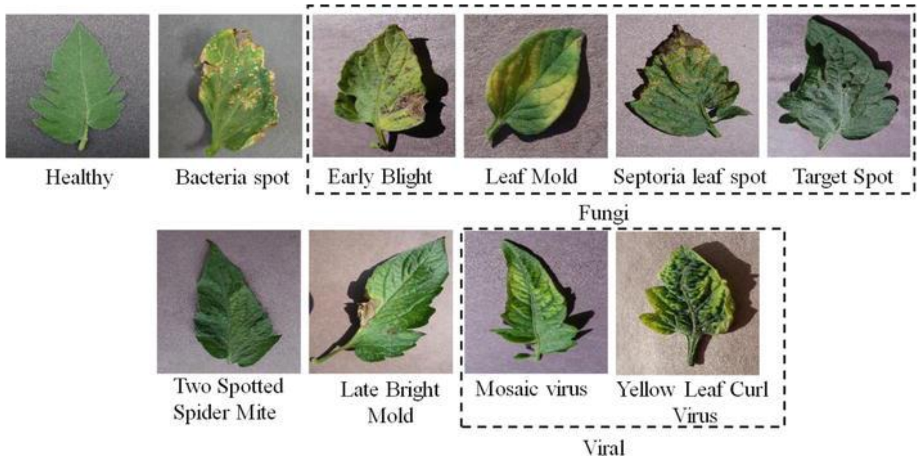


Fig 1.2 Tomato Healthy and Types of Disease [13]

1.4 Machine Learning

Artificial intelligence is a sort of machine learning. Without being coded, a system learns from its past experiences and improves. Instead, it concentrates on developing the computer program such that the data accessed may be used for self-learning.

1.5 Deep Learning

The term "deep learning" refers to a sort of machine learning that excels at identifying patterns in large volumes of data. Object detection from images, for example, is achieved using three or more layers of artificial neural networks, each of which extracts a large number of visual properties.

1.6 Neural Network

A computer model works similarly to neurons in the human brain. Each neuron accepts an input, acts on it, and then passes the findings to the next neuron.

1.7 Deep Convolution Neural Network

As an information processing paradigm, the neural network was inspired by the biological nervous system. It comprises a large number of neurons, which are densely coupled processing components that generate a series of real-valued activations. For example, when given input, early neurons are activated, and weighted connections from previously active neurons activate other neurons.

Depending on how neurons are employed and coupled, large causal chains and links between computational stages may be required. Deep neural networks (DNNs) are neural networks with a large number of hidden layers.

As a result of its growth, Deep Learning has made significant progress in image classification. Deep learning algorithms aim to learn the feature hierarchy automatically. At different degrees of deliberation, this self-learned component permits the framework to dissect complex contributions to yield planning capacities straightforwardly from getting to information without depending on human-made highlights.

1.8 How CNN Works?

A Convolution, Neural Network layers include Convolution, ReLU Layer, Pooling, Fully Connected, Flatten, and Normalization. The photographs would be compared piece by piece using CNN. The item is referred to as a feature or filter. CNN employs the weight matrix to extract particular characteristics from the input picture without losing information about the image's spatial organization. CNN adheres to the following layers:

The layer of Convolution is the first layer. It aligns the feature and picture before multiplying each image pixel by the feature pixel. After completing the multiplication of the associated matrix, CNN adds and divides the result by the total number of pixels, creates a map, and places the filter's value there. The feature is then moved to every other place in the picture, and the matrix output is obtained. Next, the process is repeated for the other filters.

Thus, this layer repositions the filter on the picture in every conceivable location.

1. ReLU Layer:

2. ReLU is an acronym for Rectified Linear Unit

3. Every negative value in the filter pictures will be eliminated and replaced with zeros in this layer. The function only activates a node when the input is zero and the output is zero. However, the dependent variable has a linear connection if the input grows. It enhances the neural network by increasing the amount of training.

4. Polling Layer:

5. This layer reduces the size of the picture by extracting the maximum value from the filtered image and converting it to a matrix. It also keeps overfitting at bay.

6. All Fully Connected Layer

7. This is the final layer of CNN, and it is here that the absolute categorization takes place. First, a single list is created with all of the filtered and compressed photos.

Literature Review :

2.1. REVIEW OF LITERATURE

We used data from Kaggle to evaluate the leaf health identification task of tomato plants. The dataset was created using a leaf picture of a tomato plant, including healthy and unwell photos with various classifications. We will split the data into sections: We will use 70 % of the dataset to train the neural network, and another 30% will be utilized as the test set. In the next stage, we will use the collection of photographs to train the model. So far, we have been successful in identifying the diseased leaf. We will locate illnesses in the future depending on whether they are caused by bacteria, fungi, or viruses and offer treatment recommendations to farmers in the field.

2.2. RELATED LITERATURE

Based on the study of numerous authors' work, the following information regarding the Tomato Plant Disease:

Khirde et al. [1] reviewed different segmentation and feature extraction algorithms that may be used to diagnose plant illnesses using images of their leaves in their literature survey. Manually detecting plant illnesses is extremely tough due to time, understanding of plant diseases, and labor. As a result, the author has divided the entire process of identifying plant leaf diseases into five steps, which are as follows: a) Image acquisition; b) Pre-processing; c) Segmentation; d) Feature extraction; e) Disease classification. For RGB leaf image capture, the transformation structure was utilized. The picture is then pre-processed to reduce noise and improve contrast. Next, segmentation separates an image into multiple feature parts using K-means clustering, otsu filters, and other techniques. This segmented picture is then utilized to extract features before being classified using a variety of classification techniques.

Sannakki et al.[2] developed a feed-forward back-propagation Neural Network-based approach to identify and categorize illnesses in grape leaves. For the diagnosis, the author used photos of a grape leaf with a complicated backdrop. Anisotropic diffusion is then utilized to eliminate the image's noise before it is segmented using k- means clustering. Finally, neural networks are used to observe the outcomes. The results were tested on downy mildew and powdery mildew matrices, with actual positive and false-positive para-author claims to have an accuracy of about 100% when using the color characteristic alone.

Kutty et al.[3]To classify the Downey Mildew and Anthracnose watermelon leaf diseases, researchers used a neural network-based technique. For the efficiency of the suggested approach, the author calculated the valid positive rate, actual negative rate, and total accuracy. This classification is based on the RGB color model's extraction of color features from identified pixels in the region of interest.

K. M. Arjun and colleagues et al. [4] The study's goal is to offer a foundation for monitoring and assessment by presenting critical data regarding the agricultural scenario and its contribution to total GDP.

In this research study, S. Shrivastava and colleagues et al. [5] presented a practical application of digital image processing in agriculture for identifying and categorizing Brown Spot and frog eye. They employ digital image processing to extract the form feature vector, which is then used to classify diseases using a K-NN classifier.

Zhang et al.[6] provide a new technique based on global features and zone-based local features for detecting citrus canker from field leaf photos, which is more challenging than lab leaf images. To begin, an updated Adaboost algorithm is utilized to choose the most critical aspects of citrus lesions for segmentation from the background.

Dasmunshi et al. [8] this study aims to monitor plant health using the NDVI (Normalized Difference Vegetation Index) computation, which can help distinguish between healthy and unhealthy plants by calculating their NDVI values. The photos of the plant were captured using a NIR camera connected to the Raspberry Pi.

Patil et al. [9] Traditional procedures were ineffective and incorrect. As a result of several studies, image processing has been included for reliable disease diagnosis utilizing plant leaves. The disease may be detected by looking for different spots and patterns on plant leaves. The adoption of digital image processing for more accurate findings was a step forward. However, after consulting several reputable IEEE, international conferences, and international journal publications in this sector, it was discovered that none of them provides a cure for the plant ailment.

Mhatre et al. [10] the leaf recognition method is used with the feature recognition system. The Convolutional Neural Network (CNN) is used to identify plant leaves. PCA extracts and improves properties before incorporating them into CNN.

- Problem Definition

In terms of agriculture preservation, our farmers have to check the leaves and illness of the plants by getting the sample itself or checking on the field. There might be a shot at blunder because of the absence of information and numerous different variables. So we need to automate this with the goal that farmers can rapidly build their creation.

This research aims to use a transfer learning inception v3 model that has previously been trained on a significant quantity of data to identify disease in tomato leaves autonomously.

Objectives

The proposed study's primary objective is to find a solution to the problem of identifying tomato leaf disease using the most straightforward technique feasible while utilizing the fewest computer resources necessary to achieve results comparable to state-of-the-art alternatives. In addition, to assist in classifying input photos into sickness classifications, automatic feature extraction is applied. Consequently, the suggested system attained an average accuracy of 94%-95%, demonstrating the neural network approach's viability even under challenging scenarios.

Sensors and remote sensing, mapping and surveying, high-precision positioning systems, variable rates, the global navigation satellite system, automated steering systems mapping, computer-based applications, and other technologies are all used in precision agriculture. In addition, precision agricultural concepts based on infrared variation analysis and treatment are also cutting-edge technologies. This examination utilizes a more modest adaptation of the convolutional neural organization model to recognize and analyze messes in tomato leaves.

In other cases, the signals can only be detected in non-visible electromagnetic spectrum areas. This research aims to develop a user-friendly method that will aid farmers in recognizing tomato plant issues without having to consult an expert. We first obtain a picture from the Kaggle dataset, from which we extract characteristics. To extract the attributes, we employ picture conversion and scaling. To diagnose diseases, the transfer learning inception v3 model will be used.

Methodology

Deep learning increased the learning capacity of the features directly in highly dimensional unprocessed s data, the deep learning algorithms in images and extraction of specific audio segments, and supervised learning in general. Hence, as a strong candidate for classification task modulation, an integrated understanding of deep Learning algorithms solves the central problem as the characteristics of the samples are selected and extracted. Therefore, it shows the combination of simple functions in more efficient and more complex features to achieve the classification more efficiently and complicated. Moreover, deep neural networks have a multilayer structure that can better extract the signal properties by avoiding the lengthy manual selection of data properties. In this project, we create a small-scale deep convolutional neural network (DCNN), assess its performance, and then develop a more advanced deep learning network based on various art data and techniques.

3.2.1 The architecture of DCNN is as follows:

We developed "IDENTIFICATION of tomato leaf disease using transfer learning" using Deep learning. We just took a step and started to collect lots of images of crop leaves. We require space capability to get the correct information. Then, at that point, we pick which calculation is ideal for tackling this issue, and we choose "Convolution Neural Network" not surprisingly (CNN). Be that as it may, we gain less precision utilizing move learning engineering, which admirably prepares and tests datasets.

Pre-processing & feature extraction, we select leaf data such as color, shape, and texture are helpful in pattern recognition, classification.

It gives us more than 81% accuracy on training and validation data set in just 20 epochs. After that, we deployed this model on the flask.

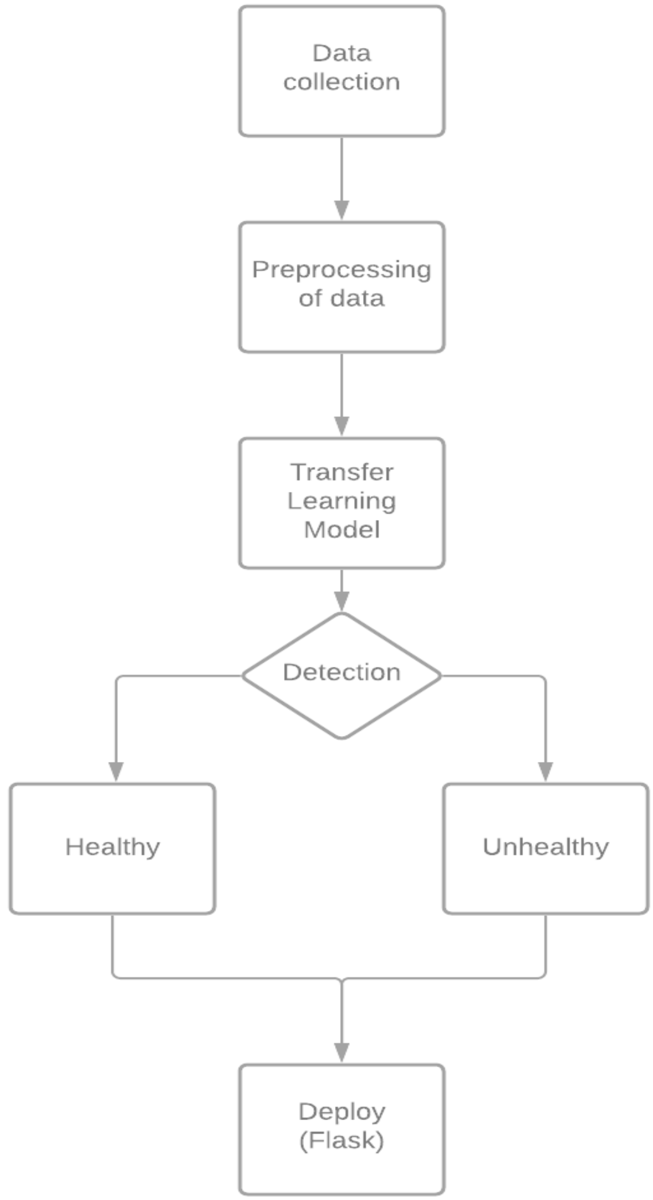


Fig 3.1 Proposed Methodology of the system

3.2.2 TRANSFER LEARNING

Transfer learning is the practice of reusing a previously learned model for a replacement task; it is popular in deep learning since it allows deep neural networks to be trained with a small amount of data. It is instrumental in data science because most real problems do not have many data points marked to coach these complex models. Let us look at transfer learning, how it works, why it is valid, and when it should be used. Several resources for models who have been previously trained in learning transfers are included.

The general idea behind transfer learning is to apply knowledge obtained from projects with many marked data to situations where just-named data is available. Because creating named data is expensive, it is critical to make use of existing datasets whenever possible. The primary goal of a standard machine learning model is to summarise inconspicuous information based on designs derived from preparation data. You use transfer learning to begin this speculative encounter by starting with strategies learned for a separate mission. Essentially, instead of starting the taking-in process with a (usually randomly introduced) transparent sheet, you start with designs that have been planned out to answer an alternate assignment.

Although deep learning is frequently employed in research, many real-world situations lack a large number of named data points on which to create a model. To tune the high number of boundaries in neuronal architecture, profound learning approaches need large amounts of data. This necessitates a large amount of (expensive) labeled information, especially in the case of controlled learning. Although it may seem minor, master information is required to create a large named dataset in Natural Language Processing (NLP).

Transfer learning has both advantages and disadvantages. Understanding these drawbacks is critical for successful AI applications. Information transfer is only possible when it is 'appropriate.' It's difficult to determine the best approaches in this case, so you'll have to try a lot.

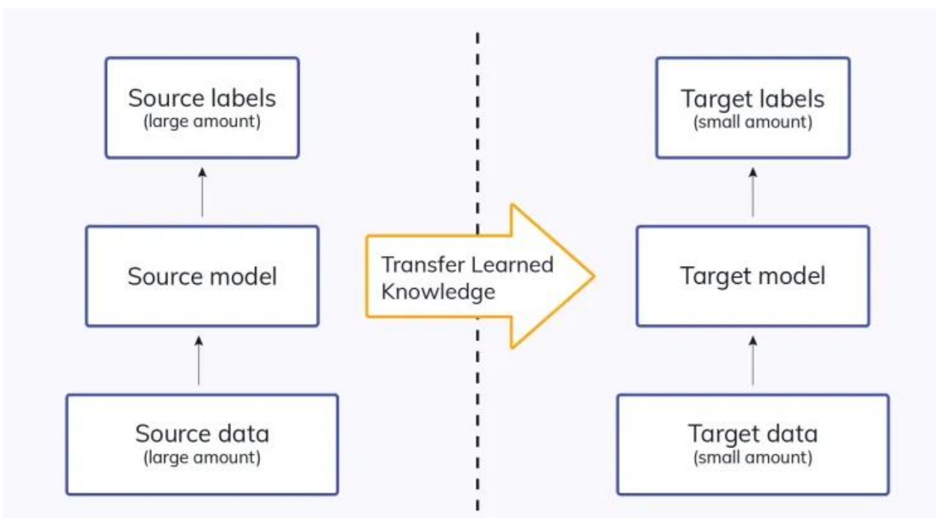


Fig 3.2 Transfer Learning [14]

3.2.3 Commonly used Transfer Learning Models.

Now we will look at some of the most popular and widely utilized transfer learning models. The majority of the models we will talk about later are employed in picture categorization.

- *Inception*

Szegedy proposed the Inception microarchitecture in their work Going deeper with Convolution in 2014. The whole architecture with dimension reduction looks like this:

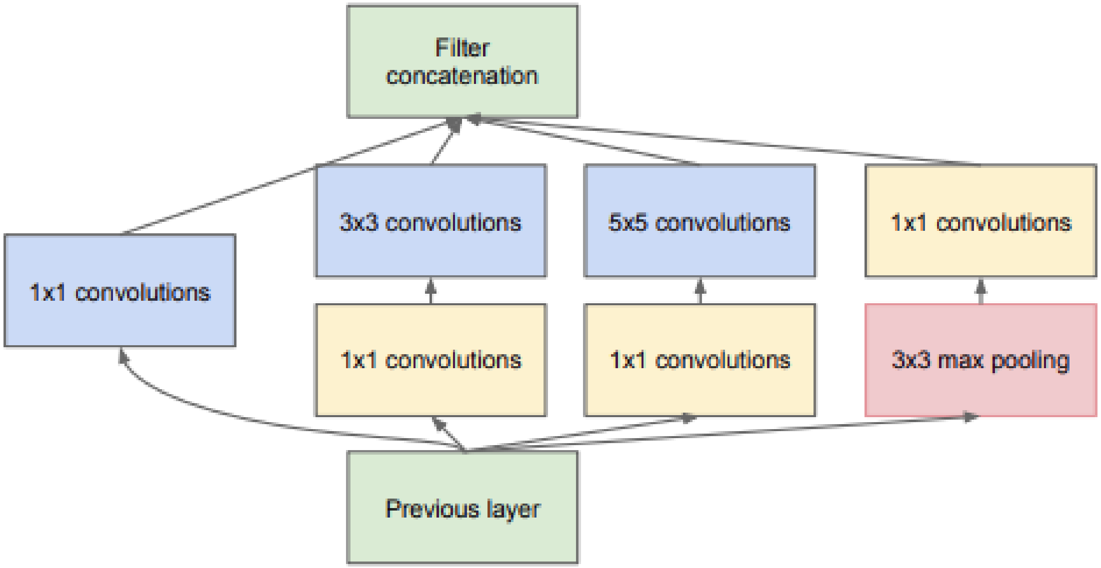


Fig 3.3 Inception Model [15]

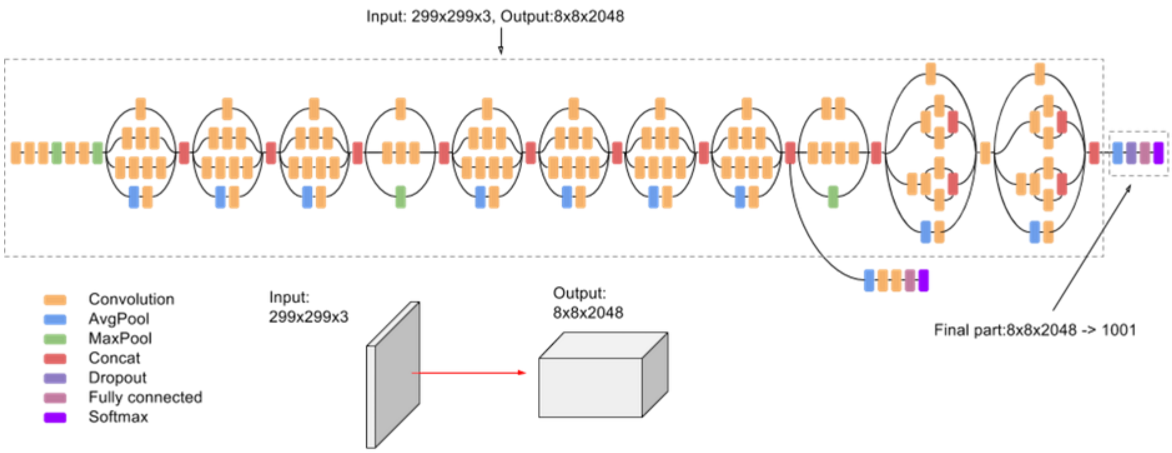
This module operates as a multi-level feature extractor by computing convolution levels 11, 33, and 55.

- **Large Categories' Image Classifier Inception v3:**

Inception-v3, a 48-layer deep pre-trained convolutional neural network, is a version of the network that has previously been trained on over a million pictures from the ImageNet collection. This organization has also been pre-programmed to classify images into 1000 distinct object categories, such as consoles, mice, pencils, and other critters. As a result, the company has acquired a wide range of rich component depictions for various films. In the first stage, the model collects generic characteristics from input photos and then classifies them using those features in the second portion.

On the ImageNet dataset, Inception v3 has been demonstrated to achieve better than 78.1% accuracy and roughly 93.9% accuracy in the top 5 results. The model represents the result of several concepts explored over time by several scholars.

Fig 3.4 Inception V3 Architecture [16]



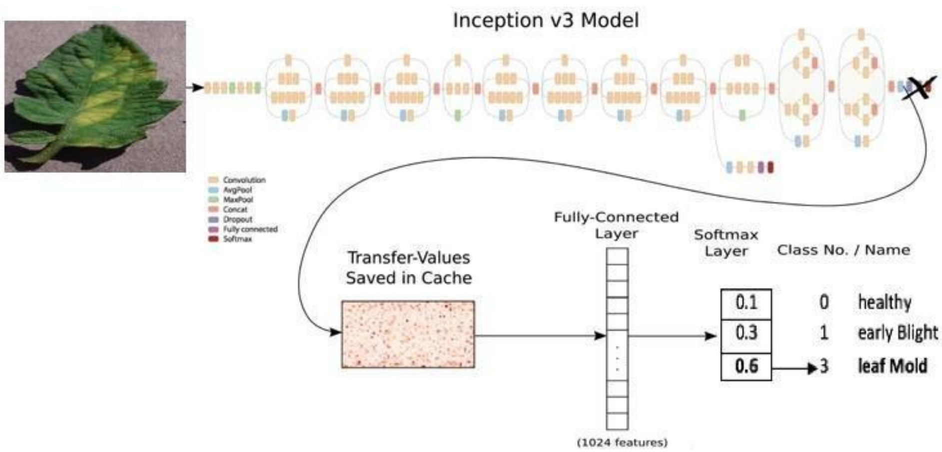


Fig 3.5 IlL Leaf Detection Through Inception V3 [16]

3.2.4 Data Analysis

We gathered the data from Kaggle for the evaluation to detect the tomato leaf illness so that improvement can be made before any lasting harm occurs. In the dataset, there are 21,416 total images in the dataset which we have divided into training dataset of 18,345 & 3,071 images in testing.

Now, the dataset is divided into two parts: one is large set which we used to train the inception model (Transfer learning), and another group is used to test the model. Here, we have divided the data in 80/20 ratio, i.e., 80% for training and 20% testing.

The model training is done with the Keras with TensorFlow as a deep learning library using a TITAN RTX 24G GPU. The Adam optimizer used for the architectures, and the loss function was the categorical cross-entropy function. We used ReLU activation functions for all layers, except the last dense layer, using sigmoid activation functions. We have trained the model up to 20 epochs & we achieved 87.69% Accuracy of model.

3.2.5 Training the Model

For example, in computer vision, neural networks often attempt to detect edges in earlier layers, shapes within the middle layer, and a few features specific to tasks within the following layers. Second grade. It helps leverage the labeled data of the job that it had been trained initially. The model has learned to acknowledge objects, so we'll only retrain the subsequent layers. We attempt to transfer the maximum amount of knowledge possible from the trained model's previous task during transfer learning. This data can take many various forms counting on the matter and therefore the data. For instance, it might be how models are constructed, allowing us to define new objects more easily. Tons of knowledge are required to coach a neural network from scratch but not always have access to that data available - this is often where transfer learning becomes useful. Because the model has been trained beforehand, this is often especially valuable in natural language processing since most of the expertise is required to make large-labeled datasets. Also, training time is reduced because it can sometimes take days or maybe weeks to coach a deep neural network from scratch for a task. After training, true-positive, false-positive, true-negative, false-negative of the test set were recorded successively.

References

1. Agarwal, M., Singh, A., Arjaria, S., Sinha, and Gupta, S., 2020. ToLeD: Tomato leaf disease detection using convolutional neural network. *Procedia Computer Science*, *167*, pp.293-301.

2. Aida Sulinda, Suhaili Beeran, Noor Ezan Abdullah, Habibah Hashim, and Kutty "Using Neural Network Analysis to Classify Watermelon Leaf Diseases." 459464 in IEEE Business Engineering and Industrial Applications Colloquium (BELAC), 2013. 2013 IEEE.

3. Arjun, K. M. (2013). Indian agriculture status, importance and role in the Indian economy. *International Journal of Agriculture and Food Science Technology*, *4*(4).

4. Brahimi, M., Arsenovic, M., Laraba, S., Sladojevic, S.,Boukhalfa, K. and Moussaoui, A., 2018. Deep learning for plant diseases: detection and saliency map visualization. In Human and machine learning (pp. 93-117). Springer, Cham

5. Dash[a], J., Verma[a], S., Dasmunshi[a], S., & Nigam[a], S. (2018). Plant Health Monitoring System Using Raspberry Pi. *International Journal of Pure and Applied Mathematics*, *119*(15), 955-959.

6. Mhatre, R., & Lake, V. COTTON LEAVES DISEASE DETECTION AND CURE USING DEEP LEARNING.

7. Patil, B., Panchal, H., Yadav, M. S., Singh, M. A., & Patil, M. D. (2017). Plant Monitoring using image processing, Raspberry Pi & IoT. *International Research Journal of Engineering and Technology (IRJET)*, *4*(10).

8. Sachin D. Khirade and A. B. Patil, "Plant Disease Detection Using Image Processing." ICCUBEA (International Conference on Computing, Communication, Control, and Automation) is a biennial international conference on computing, communication, control, and automation.), pp. 768-771. 2015, IEEE.

9. Sannakki, Sanjeev S., Vijay S. Rajpurohit, V. Nargund, and Parag Kulkarni are Sanjeev S. Sannakki, Vijay S. Rajpurohit, Vijay S. Rajpurohit, Vijay S. Rajpurohit, Vijay S "Using neural networks to diagnose and classify grape leaf diseases." 2013 Fourth International Conference on Computing, Communications, and Networking Technologies (ICCCNT), pp. 1-5 IEEE, 2013.

10. Shrivastava, S, & Hooda, D. S. (2014). Automatic brown spot and frog eye detection from the image captured in the field. *American Journal of Intelligent Systems*, *4*(4), 131- 134.

11. Zhang, M., & Meng, Q. (2011), Automatic citrus canker detection from leaf images captured in the field. Pattern Recognition Letters, 32(15), 2036–2046.

12. - https://en.m.wiktionary.org/wiki/Alternaria

13. https://www.researchgate.net/figure/Sample-images-of-healthy-and-different-unhealthy-tomato-leaves-from-the-plant-village_fig1_351269745

14. https://images.app.goo.gl/ByPCZeQeNtrkjo6p9

15. https://www.researchgate.net/figure/Inception-network-structure-113-ResNetKaiming-He-Xiangyu-Zhang-Shaoqing-Ren-and-Jian_fig1_335573689

16. https://www.nicepng.com/maxp/u2r5y3o0a9i1w7a9/

Note:

1. The project group should not exceed 3 students.
2. In the body of the text, a reference number should be indicated in parenthesis such as [1].
3. The reference should be given at the end of the project synopsis in order of their use in the text.
4. The reference should be in proper format strictly.

For Example:
The references are given for example only. These are not real citations.

Journal References
[1]. Fuschini F., Falciasecca G., "Experimental investigation of the effects of snowdrifts and ice deposit on linear wire antennas at 2.4 GHz.", IEEE Antennas and wireless propagation letters, Volume 13, pp: 686-689, 2014.

Conference References
[2]. Zhigang D., Qahouq J.A.A., "Modelling and investigation of magnetic resonance couple wireless power transfer system with lateral misalignment", Twenty-Ninth Annual IEEE applied power electronics conference and exposition", pp: 1317-1322, 2014.

Books References
[3]. Balanis C.A., "Antenna Theory: Analysis and design", John Wiley & Sons, Inc., Second edition, pp: 1-4, 1997.

URL References
[4]. Microstrip patch antenna calculator http://www.emtalk.com/mpacalc.php, accessed on 4/8/2022 at 10:30 AM.

****