

Implementation Strategy of Tomato Plant Disease Detection using Optimized Feature Extraction Method

Seetha Rama Raju Jampana, Dama Tejaswini, Nandini Paleti, Keerthana Sarabu, Prasanth Yalla

2000031306@kluniversity.in, 2000031275@kluniversity.in, 2000031325@kluniversity.in, 2000031606@kluniversity.in,
prasanthyalla@kluniversity.in

Department of Computer Science and Engineering, Koneru Lakshmaiah Education Foundation, Vaddeswaram, Guntur 522302, AP, India

Abstract— Tomato plants normally have a single growing season, during which they develop, bear fruit, and then perish. The species first appeared in Western South America, Mexico, and Central America. In the sixteenth century, they were brought to various regions. They produce self-pollinating yellow blooms. After being pollinated, the blooms turn into fruits, which, depending on the type, might be red, yellow, green, or even purple. Tomatoes are a well-liked element in many recipes, including salads, sauces, and soups. They are high in vitamins A and C, potassium, and antioxidants. They are afflicted by several illnesses that can seriously harm the plant and lower crop output. These illnesses were brought on by a variety of minor inadequacies in the soil, air, and the major. These diseases are produced by a range of mineral deficiencies in the soil, and the air, and their primary causes include insects and fungi. We discovered that machine learning is a potential avenue for detecting these diseases before they spread to the plant. As a result, we thought about using Feature Extraction Methods to optimize the data.

Keywords- Machine Learning, Tomato Plant, Tomato Plant diseases, Feature Extraction, Principle Component Analysis (PCA), CNN.

I. INTRODUCTION

Diseases that ravage the crops are one of a farmer's greatest enemies. Crop production determines a farmer's ultimate fate, and a diminishing crop is a worst-case situation that might happen. All tiers of society are impacted if the food supply is insufficient, and hoarding is a typical reaction to this scarcity. Despite its seeming smallness, farming has a big impact on how a country develops. Farmers frequently lose crops to unidentified diseases, and it takes an excruciatingly long time to learn about a disease infecting a crop, so the damage has already been done. Deep learning was utilized in the classification and identification of the impacted leaves. With the application of artificial intelligence, farmers can quickly identify the disease in its early stages and treat their crops to ensure sustainable growth the most crucial plant that we started with was the tomato plant. It is a staple of many cultures and is widely utilized in many cuisines. It has its roots in western South America and Central America and has a significant influence on Spanish culture. Variations in tomato consumption exist. Although it is a fruit, it is a vegetable ingredient.

We have chosen four ailments that harm tomato plants. The bacterial spot is the source of the initial disease. Four different *Xanthomonas* species generate bacterial spots. [Fig.1]. These species include *Vesicatoria*, *perforans*, *gardneri*, and *Uvesicatoria*. It happens everywhere there are tomato plants. It results in fruit and leaf spots, which decrease

productivity, cause fruit to get sun-scalded, and cause defoliation. Due to the variety of bacterial spot illnesses, this disease can develop at a variety of temperatures and threatens tomato output worldwide. This sickness causes leaf lesions to be round, moist, and surrounded by a hazy yellow halo. These patches are described as dark brown or black. The second illness is the Septoria leaf spot, which is the result of a fungus called *Septoria lycopersici*. It is among the most devastating tomato leaf diseases, and it is especially dangerous in regions where there is rain and humidity. After the first fruit forms on the lower leaves, it typically begins to grow. Circular spots on the leaves have a diameter of 1/16 to 1/4 inch, dark brown margins, and tiny black fruiting structures. There are several dots on each leaf. Tomato mosaic is the third ailment. The impacted plant has alternating yellowish and darker-green regions that are mottled. In the tomato plant, it is fairly typical. Fruits may experience uneven ripening, necrotic patches, yellow blotches that appear on both mature and green fruit, and internal fruit wall browning. The final illness is tomato leaf curl. It is a deadly viral illness brought on by a member of the Geminivirus genus. The tomato plant serves as the disease's primary host. Significant losses may result from it. The amount of fruits produced is drastically reduced. Very Just a handful of the fruits that were on the botanical specimen before infection will set after infection. This virus is not spread mechanically and is not transmitted through seed. Whiteflies are the carriers of the disease. An autonomous system built around artificial intelligence and

methods of image processing that can recognise sick plants and classify tomato illnesses will be extremely beneficial. Visual study of plants may result in incorrect illness diagnosis due to a lack of understanding. This also leads to people implementing inadequate preventive actions. As a result, if a machine can detect ill tomato plants and identify the type of disease, it opens the way to successfully nurturing the remaining crops while taking the required safeguards to minimize loss.

II. INVESTIGATION OF TOMATO PLANT DISEASES

A range of conditions can affect tomato plants, including illnesses, which can have a substantial impact on their growth and output. These are a few of the most prevalent illnesses that can affect tomato plants.

II.1 Early Blight:

Early blight is a frequent and widespread illness that impacts the growth of tomatoes and is brought about by a fungal infection called *Alternaria solani*. It is more common in hot, humid conditions, but it can also occur in cooler locations. Early blight can severely harm tomato crops, resulting in lower yield and quality. The disease often appears on the plant's lower leaves as dark, concentric lesions with a target-like appearance. These lesions gradually enlarge and spread, causing the affected leaves to yellow, wither, and eventually die. As the disease progresses, it can move up the plant, affecting the stems, fruits, and other upper parts. Early Blight thrives under favorable conditions such as high humidity, warm temperatures (around 75-85°F or 24-29°C), and prolonged leaf wetness. The fungus survives the winter in infected crop garbage, acting as a source of the inoculum for additional infections during the next growing season. The spores of *Alternaria solani* can be spread by wind, rain, contaminated tools, and equipment.



Fig.1

Early blight



Fig.2

II.11 Late Blight:

Late blight is a devastating disease that affects tomato crops and is caused by the water mould *Phytophthora infestans*. It is most known for triggering the Irish Potato Famine in the 1840s, but it also poses a serious threat to tomato crops around the world. Late Blight typically occurs in regions with cool and humid climates, thriving in temperatures ranging from 60-80°F (15-27°C) and high humidity or rainfall. The disease is notorious for its rapid spread and ability to cause extensive damage within a short period. Late Blight symptoms first develop as black, water-soaked sores on the leaves, generally at the tips or margins. During humid conditions, these lesions rapidly expand and turn dark brown

or black, with a characteristic fuzzy white or grey mould growth on the bottom. The affected leaves rapidly wither and die, giving the plant a scorched appearance. The disease can also affect the stems, petioles, and eventually, the fruits, causing rotting and decay.



Fig.3

Late Blight



Fig.4

II.111 Fusarium Wilt:

Fusarium Wilt is a dangerous fungus that exists in the soil and kills tomato plants. The fungus *Fusarium oxysporum f. sp. lycopersici* causes it. This illness has the potential to seriously harm tomato crops and challenging to treat once it has been entrenched in the soil. Fusarium Wilt predominantly affects the vascular system of tomato plants, affecting water and nutrient movement inside the plant. The fungus colonises roots and spreads through xylem vessels, causing wilting, yellowing of lower leaves, and stunted growth. As the illness advances, the whole plant might wilt and eventually die. Because *Fusarium oxysporum f. sp. lycopersici* can live in the dirt for several years., crop rotation alone is insufficient for comprehensive control. The fungus can access the plant via wounds on the roots or natural openings like root hairs. It thrives and reproduces inside the plant, causing harm to the vascular tissue. Warm soil temperatures (about 80-85°F or 27-29°C) and high soil moisture favor Fusarium Wilt. The illness has increased prevalence in poorly drained or compacted soils, where the fungus can endure for a very long period.



Fig. 5

Fusarium Wilt



Fig.6

II.1V Septoria Leaf Spot:

The fungus *Septoria lycopersici* causes Septoria Leaf Spot, a widespread foliar disease that harms tomato plants. If allowed unchecked, it can result in defoliation, decreased vigor, and lower yields. Septoria Leaf Spot is a common illness that might have an impact on both field and greenhouse tomato production.

The disease usually manifests itself as small, round to angular patches on the bottom leaf surface of the tomato plant. These patches have a dark brown or black center encircled by a lighter, tan, or greyish halo. As the disease continues, The areas could expand and coalesce, causing the affected leaves to yellow, wither, and eventually drop prematurely.

Septoria lycopersici can be seed-borne and overwinters in infected plant waste or soil. The fungus spreads its spores mostly through splashing water or rain, as well as through tools, equipment, and wind. Warm and humid weather, combined with prolonged leaf wetness, offer ideal circumstances for the disease to thrive and spread quickly



Fig.7

Septoria Leaf Spot



Fig.8

II.V Gray Mold:

Grey mold, additionally referred to as Botrytis grey mold or grey mold rot, is a widespread fungal disease that attacks different areas of tomato plants. It arises from the fungus Botrytis cinerea and can cause substantial harm to tomato crops, especially when growing in humid and cool circumstances.

Gray Mold typically affects mature and overripe tomatoes, but it can also infect flowers, stems, leaves, and green fruits. The disease first appears as a fuzzy gray or brown mold growth on the affected plant parts. As it progresses, the mold can spread rapidly, causing rot, decay, and softening of the tissues.

Botrytis cinerea grows well in cold, humid environments with degrees that range from 68 to 77 degrees Fahrenheit (20 to 25 degrees Celsius) and high relative humidity. It can live in plant detritus, soil, or infected fruits and spreads via airborne spores or personal contact with contaminated plant material. Gray Mold can enter tomato plants through wounds, injuries, or natural openings, such as stomata or lenticels. Once inside the plant, the fungus colonizes the tissues and generates the breakdown-inducing enzymes in the cell walls, facilitating its spread and causing decay.



Fig.9

Gray Mold



Fig.10

II.V1 Verticillium Wilt:

Verticillium wilt is a harmful fungal disease that affects a variety of plants, including tomatoes. It's brought on by soil borne fungi of the genus Verticillium, notably Verticillium dahliae and Verticillium albo-atrum. Because these fungi can live in the ground for a few years, they provide a constant threat to sensitive crops. Tomato plants infected with Verticillium wilt display a wide range of symptoms. Lower leaves may initially turn yellow and wilt, spreading upwards throughout the plant. The entire plant may wither and die as the disease spreads. Verticillium wilt is distinguished by the darkening of vascular tissues, notably in the stem. If you cut the stem of a tomato plant that has been compromised, you will probably notice dark brown streaks or discoloration in the damaged tissues.

Verticillium wilt is primarily transmitted through contaminated soil, tools, or plant debris. Fungi get into the plant using the roots and colonize the vascular system, disrupting water and nutrient absorption. Environmental factors such as high soil moisture, low temperatures, and poor drainage can all contribute to the disease's establishment and spread.

Managing Verticillium wilt in tomato plants can be challenging. Crop rotation is one common practice, where susceptible crops are not planted in the same vicinity for multiple years to decrease the amount of fungi present in the soil. Additionally, using disease-resistant tomato cultivars can help minimize the result of Verticillium wilt. Fungicides are generally ineffective against this disease, so preventive measures and cultural practices play a crucial role in its control.

In conclusion, Verticillium wilt is a significant concern for tomato growers, causing substantial economic losses worldwide. Understanding its symptoms, mode of transmission, and management strategies can aid in minimizing its impact on tomato crops and ensuring a healthy harvest.



Fig.11



Fig.12

Verticillium Wilt

II.V11 Target Spot:

Corynespora leaf spot, often known as target spot, is a type of fungal illness that damages tomato plants. The pathogen Corynespora cassiicola, which can persist in plant detritus and soil, causes it. Target spot is common in tropical and temperate nations because it thrives in humid, warm settings. Small, round lesions appear on the leaves, stems, and fruits of tomato plants. These lesions begin as dark brown patches but evolve into unmistakable concentric rings that resemble a target or bullseye. The lesions may have a yellow halo surrounding them, further aiding in their identification.

Severe infections can lead to defoliation and a decline in fruit quality and yield.

The pathogen spreads through airborne Spores are discharged from infected plant tissues distributed by wind, rain, or physical contact. The spores additionally be transmitted through contaminated tools, equipment, or hands. Once the spores come into contact with susceptible tomato plants, they can penetrate the tissues and initiate infection.

Managing target spots involves a combination of cultural practices, chemical treatments, and preventive measures. To minimize disease incidence, it is important to maintain good plant hygiene by removing and destroying infected plant material. Providing proper spacing between tomato plants and ensuring adequate air circulation can also reduce the favorable conditions for disease development.

Fungicides are available for controlling target spots, but their effectiveness relies on early and regular application. It is essential to select fungicides specifically labeled for the target spot and follow the recommended application rates and intervals. Additionally, planting resistant tomato varieties can offer an effective means of managing the disease.

In summary, the target spot is a fungal disease that poses a significant threat to tomato plants, affecting their foliage and fruit. By recognizing the symptoms, understanding the mode of transmission, and implementing appropriate management strategies, growers can effectively combat target spots and protect their tomato crops.



Fig.13
Target Spot

III. TYPES OF SOILS THAT AFFECT TOMATO PLANT GROWTH

Several types of soils can be suitable for growing tomato plants, depending on the climate, location, and other factors. These are a few of the common types.

III.1 Sandy soil:

Sandy soil is soil that drains water quickly and has large particles, making it easy for roots to penetrate. This type of soil is suitable for tomato plants as it provides good aeration and drainage, however, it might need to be watered more frequently, and fertilization of soils used for tomato plants



Fig,14

Sandy Soil

III.11 Loamy soil:

Loamy soil is a combination of sand, silt, and clay, providing a balance of drainage and water retention. This type of soil is ideal for growing tomato plants as it holds nutrients well and promotes healthy root growth.



Fig.16
Loamy Soil

III.111 Clay soil:

Clay soil is heavy, dense soil that holds water well but can become waterlogged, leading to poor root growth. However, if amended properly, clay soil can be suitable for growing tomato plants as it holds nutrients.



Fig.18
Clay soil

III.1V Sandy loam soil:

Sandy loam soil contains a mixture of sand, silt, and clay, providing good drainage and water retention. This type of soil is ideal for growing tomato plants as it promotes healthy root growth and allows for proper nutrient uptake.



Fig,15
Sandy loam soil

III.V Peat-based soil:

Peat-based soil is a type of soil made from decomposed plant material and is a good choice for growing tomato plants as it provides good aeration, water retention, and nutrient retention.



Fig.17
Peat-based soil

III.VI Compost-based soil:

Compost-based soil is a type of soil made from decomposed organic matter and is rich in nutrients. This type of soil is ideal for growing tomato plants as it promotes healthy plant growth.



Fig.18
Compost-based soil

IV. LITERATURE SURVEY

The authors of the research [14] noted that the following diseases harm tomato plants: Leaf mould, Grey mould, Canker, and Plague. These diseases cause a variety of symptoms, including changes in the colour and texture of leaves, wet markings, and fungus that spreads quickly from plant to plant. These disorders can be caused by humid greenhouse conditions and dietary abnormalities. The methodology used to predict these diseases is based on colour, shape, and texture information extracted from disease photos. Backpropagation, Radial Basis Function, Generalized Regression Network, and Stochastic Artificial Neural Networks are the four neural networks that are used to choose the optimum strategy.

According to the paper [15], tomato plant diseases include early blight, Septoria leaf spot, late blight, Fusarium wilt, Verticillium wilt, Anthracnose, Buckeye rot, and Southern blight. The consequences of these diseases can include significantly restricting yield, defoliation, and tissue death. The reasons for these diseases can include fungi and bacteria

attacking plants through infected soil and seeds. The methodology used for predicting these diseases is not mentioned.

As stated by the paper [18], the following pathogens affect tomato plants, Viruses, Viroids, Fungi, Oomycetes, Bacteria, and Nematodes. The impact of these diseases can include reduced yield and the reduction of product quality. These illnesses may result from several things, like as diseased soil and seeds. The technology utilized to predict these diseases is based on CNN models, which perform exceptionally well in classifying diseases from plant leaf pictures.

According to the paper [19], the following diseases affect tomato plants, Target spot, Bacteria YLVC Virus. The impact of these diseases can include the appearance of lesions and shortened shoots. The reasons for these diseases can include the fungal pathogen *Corynespora cassiicola* and tomato yellow leaf curl virus (TYLCV). The methodology used for predicting these diseases stems from DCGAN-based data augmentation.

According to the plant [20], the authors discussed the impact of weather changes on crop yields and the importance of agriculture in addressing the increasing food demand due to population growth. It focuses on the application of modern technology, notably AI algorithms, to detect and identify diseases in tomato leaves. The idea is to connect farmers to these technologies to be able to reduce crop illnesses. The accuracy of various classification and filtering algorithms, including K-Mean, SVM, MLP, NN, BPNN, and CNN, in detecting tomato leaf illnesses is compared. The proposed architecture achieves 99.4% accuracy. Overall, the paper intends to assist farmers by utilising artificial intelligence technologies to detect and manage infections in tomato plants.

According to the research [16], Anthony recognised the classification of four primary illnesses affecting tomato leaves (bacterial spot, septoria spot, mosaic virus, and yellow-curl) using different feature extraction algorithms. The decision tree classifier and random forest classifier are employed for sickness classification, with 90% and 94% accuracy, respectively. The results reveal that the random forest classifier outperforms the decision tree classifier in terms of accuracy. When compared to existing state-of-the-art techniques, the proposed method offers lower processing time and higher classification accuracy. Future work involves expanding the disease classification to include a broader range of diseases within and across different classes, improving accuracy through the incorporation of additional features such as Visual Bag of Words, SIFT, SURF, and ORB, and developing a system that automatically captures, detects, classifies diseases, and initiates appropriate actions using suitable hardware and software.

The authors of the research [17] proposed an accurate method for detecting tomato leaf diseases using PLPNet, an image-based disease detection system. We use a PAC module to improve the model's feature extraction capability and LRAM

to filter out noise. Furthermore, during feature fusion, SD-PFAN is created to match the similar aspects of tomato leaf diseases. This method efficiently tackles leakage and misdetection concerns, enhancing illness detection accuracy. PLPNet accurately recognizes and matches aspects of tomato leaf disease photos by increasing the performance of the YOLOX model. This research can help growers detect tomato leaf diseases in a timely and reliable manner, allowing for tailored control methods based on the disease type discovered. The testing findings reveal that PLPNet outperforms different models about precision and. However, there are still issues to overcome, such as developing a grading system for leaf illnesses and combining Internet of Things sensor technologies with deep learning to provide complete disease early warning.

V. METHODOLOGY

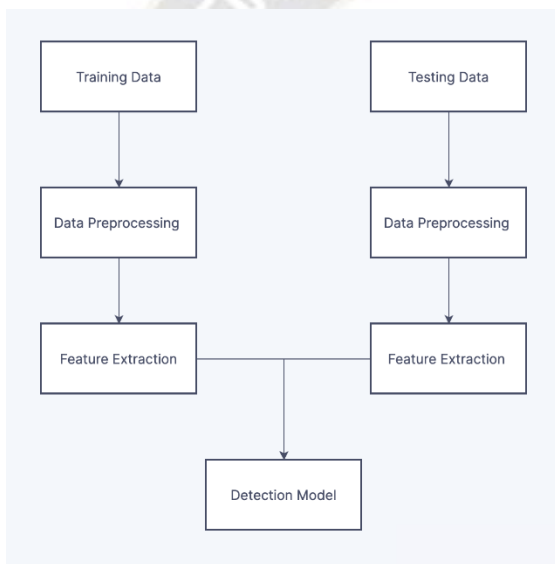


Fig.19

V.1 Training Data:

A series of examples or instances is referred to as "training data" and is used to develop a machine-learning model. Typically, these examples comprise of input data and corresponding output data, which are used to instruct the model on how to forecast or categorize fresh data. Training data can either be labeled or not. A broad and representative sample of the problem space that the machine learning model will encounter is what training data aims to deliver. Because of this, the model may learn from a wide range of events and apply what it has learned to fresh circumstances.

V.11 Testing Data:

Testing data is used in order to evaluate the efficacy or the execution of a software program or machine learning model. This information is distinct from the model's training data and

is normally saved until the model has been trained and optimized. Ultimately, the aim of testing data is to offer a fair assessment of how well the model will function with brand-new, untested data. Developers can find and fix any problems with overfitting or underfitting and make improvements to improve overall performance by assessing the model's performance on testing data.

V.111 Data Preprocessing:

In the pipeline of data processing, a stage known as data preprocessing entails cleaning, converting, and getting raw data for analysis. Tasks like eliminating duplicates or missing numbers, normalizing data, and changing data types might all fall under this category. Preprocessing data aims to improve the consistency, completeness, and analytical suitability of the data. It focuses on cleaning and preparing the data for analysis. Data processing is the broader process of converting raw data into valuable information.

V.1V Feature Extraction:

The concept of feature extraction is used in computer vision and machine learning applications to extract meaningful characteristics or qualities from unstructured data. The purpose of feature extraction is to convert the input data into a group of characteristics that an algorithm can easily comprehend and interpret. Feature extraction techniques in image processing, for example, can help identify objects in a photograph by recognizing lines, edges, corners, and other visual structures. Natural language processing feature extraction strategies may eliminate words or phrases from text files. Characteristics can be retrieved and then supplied into machine learning models for example neural networks, decision trees, and support vector machines to accomplish tasks such as classification, grouping, and regression, among others.

V.V Detection Model:

An image or video's objects can be detected and retrieved using a detection model, an example of a machine learning model. In computer vision applications including autonomous vehicles, facial recognition systems, and surveillance systems, detection models are frequently employed. Region-based models, one-stage models, and two-stage models are only a few of the several kinds of detection models. Region-based models create a list of potential object locations in an image before categorizing each site as either containing or not containing an item. One-stage models predict the object class and bounding box coordinates for each object in a single step, whereas two-stage models predict the object class and bounding box coordinates for each object before performing classification and bounding box regression. Even though running detection models is computationally and time-consuming, they are capable of great levels of accuracy and robustness in object recognition tasks.

VI. CRITICAL ANALYSIS OF FEATURE EXTRACTION METHODOLOGIES.

VI.1 Principle Component Analysis(PCA):

With the help of a number of input features, the widely used linear dimensionality reduction method known as PCA seeks to reduce the dimensions of the original data distribution. PCA succeeds in its objective by maximizing variances and minimizing reconstruction errors through pairwise distances. Despite being an unsupervised learning technique that places a focus on data variation, it could sometimes result in incorrect classification.

VI.11 Independent Component Analysis(ICA):

ICA is a linear dimensionality reduction method created especially for data made up of independent components. Its main goal is to precisely identify and separate each independent component while removing extraneous noise. ICA is widely used in medical fields like fMRI analysis and EEG analysis to separate helpful signals from harmful ones.

VI.111 Linear Discriminant Analysis(LDA):

LDA, Conversely, there is a supervised learning dimensionality reduction method that also functions as a machine learning classifier. Its goal is to minimize spreading within each class while increasing the gap between each class's means. LDA improves classification results by minimizing class overlap in a lower-dimensional space by maximizing the distance between class means. It is critical to remember that LDA assumes the input data has a Gaussian distribution, which may result in poor classification results when applied with non-Gaussian data.

VI.1V Locally Linear Embedding(LLE):

The Manifold Learning-based dimensionality reduction method LLE is created specifically to manage nonlinear interactions between features. Instead of representing items in needlessly higher-dimensional space, it aims to do so in their original dimensions. To achieve this, LLE reconstructs data points in a lower-dimensional space while maintaining local associations. It works especially well with datasets that have intricate nonlinear features.

VI.V T-Distributed Stochastic Neighbor Embedding(T-SNE)

t-SNE nonlinear dimensionality reduction approach is to visualise large datasets. It discovers a low-dimensional representation that preserves pairwise similarities between data points. t-SNE achieves its goal by minimizing the divergence between pairwise probability similarities in the original high-dimensional space and the condensed low-dimensional space. It uses gradient descent and Kullback-Leibler (KL) divergence to optimize. T-SNE models the higher-dimensional space with a Gaussian distribution and the lower-dimensional space with a Student's t-distribution to

eliminate unequal neighboring point distances induced by the reduction.

After thoroughly reviewing the options, it was found that PCA is the most effective feature extraction approach. PCA excels at condensing a large number of variables into a small set while retaining much of the information from the larger set. The variance of the dataset is described by the first component of the major components, which has a number of components equal to the number of variables in the data. Because of this trait, PCA is suitable for a variety of circumstances and data qualities.

VII. COMPARATIVE ANALYSIS OF DETECTION TECHNIQUES

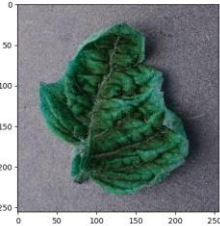
S.NO	Title of the paper	Proposed Model	Accuracy
1.	Plant Infection Detection Using Image Processing	K-Means Clustering	98.27%
2.	Early Detection and Classification of Tomato Leaf Disease Using High-Performance Deep Neural Network	CNN	98.49%
3.	Detection of Leaf Disease Using Principal Component Analysis and Linear Support Vector Machine	SVM	88.67%
4.	Plant Leaf Disease Detection and Classification based on CNN with LVQ Algorithm	CNN	86%
5.	Detection and Classification of Tomato Crop Disease Using Convolutional Neural Network	CNN	88.17%
6.	Identification of Tomato Plant Diseases by Leaf Image Using Squeezenet Model	CNN	93%
7.	Deep Structured Convolutional Neural Network for Tomato Diseases Detection	CNN	95.24%
8.	Tomato diseases recognition based on Faster RCNN	RCNN	90.87%
9.	Automatic recognition of tomato leaf disease using fast enhanced learning with image processing	CNN	99.5%
10.	Plant Leaf Diseases Detection and Classification Using Image Processing and Deep Learning Techniques	CNN	98.02%

From the above comparative analysis, we came to know that CNN is the best-optimized disease detection technique. With an overall accuracy of 98.49%, it is considered the best-optimized technique because the accuracy is purely high when compared to the other detecting techniques.

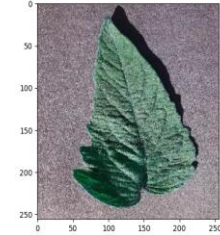
VIII.RESULT ANALYSIS

We have implemented our methodology to identify the features of a leaf to predict the disease.

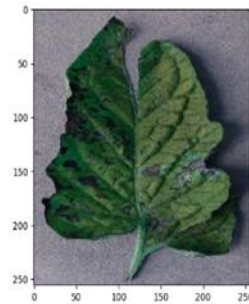
```
In [4]: import cv2
t=cv2.imread('tomato/tomato/train/Tomato___Tomato_Yellow_Leaf_Curl_Virus/cf8dbfba-3cdc-4aa2-9c43-cd6ef7f6d73___YLCV_HREC_4
plt.imshow(t)
img_path = "tomato/tomato/train/Tomato___Tomato_Yellow_Leaf_Curl_Virus/cf8dbfba-3cdc-4aa2-9c43-cd6ef7f6d73___YLCV_HREC_09
img = preprocess_images(img_path)
prediction = loaded_model.predict(img)
predicted_class_index = np.argmax(prediction)
class_labels = ["Tomato___Bacterial_spot", "Tomato___Early_blight", "Tomato___Late_blight", "Tomato___Leaf_Mold",
"Tomato___Septoria_leaf_spot", "Tomato___Spider_mites Two-spotted_spider_mite",
"Tomato___Target_Spot", "Tomato___Tomato_Yellow_Leaf_Curl_Virus", "Tomato___Tomato_mosaic_virus",
"Tomato___healthy"]
predicted_class_label = class_labels[predicted_class_index]
print("Predicted class : ", predicted_class_label)
1/1 [-----] - 0s 464ms/step
Predicted class : Tomato___Tomato_Yellow_Leaf_Curl_Virus
```



```
In [8]: import cv2
t=cv2.imread('tomato/tomato/test/img1')
plt.imshow(t)
img_path = "tomato/tomato/test/img1"
img = preprocess_images(img_path)
prediction = loaded_model.predict(img)
predicted_class_index = np.argmax(prediction)
class_labels = ["Tomato___Bacterial_spot", "Tomato___Early_blight", "Tomato___Late_blight", "Tomato___Leaf_Mold",
"Tomato___Septoria_leaf_spot", "Tomato___Spider_mites Two-spotted_spider_mite",
"Tomato___Target_Spot", "Tomato___Tomato_Yellow_Leaf_Curl_Virus", "Tomato___Tomato_mosaic_virus",
"Tomato___healthy"]
predicted_class_label = class_labels[predicted_class_index]
print("Predicted class : ", predicted_class_label)
1/1 [-----] - 0s 464ms/step
Predicted class : Tomato_healthy
```



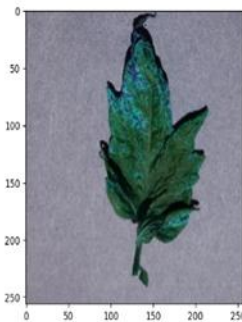
```
In [5]: import cv2
t=cv2.imread('tomato/tomato/test/img3')
plt.imshow(t)
img_path = "tomato/tomato/test/img3"
img = preprocess_images(img_path)
prediction = loaded_model.predict(img)
predicted_class_index = np.argmax(prediction)
class_labels = ["Tomato___Bacterial_spot", "Tomato___Early_blight", "Tomato___Late_blight", "Tomato___Leaf_Mold",
"Tomato___Septoria_leaf_spot", "Tomato___Spider_mites Two-spotted_spider_mite",
"Tomato___Target_Spot", "Tomato___Tomato_Yellow_Leaf_Curl_Virus", "Tomato___Tomato_mosaic_virus",
"Tomato___healthy"]
predicted_class_label = class_labels[predicted_class_index]
print("Predicted class : ", predicted_class_label)
1/1 [-----] - 0s 464ms/step
Predicted class : Tomato_Early_blight
```



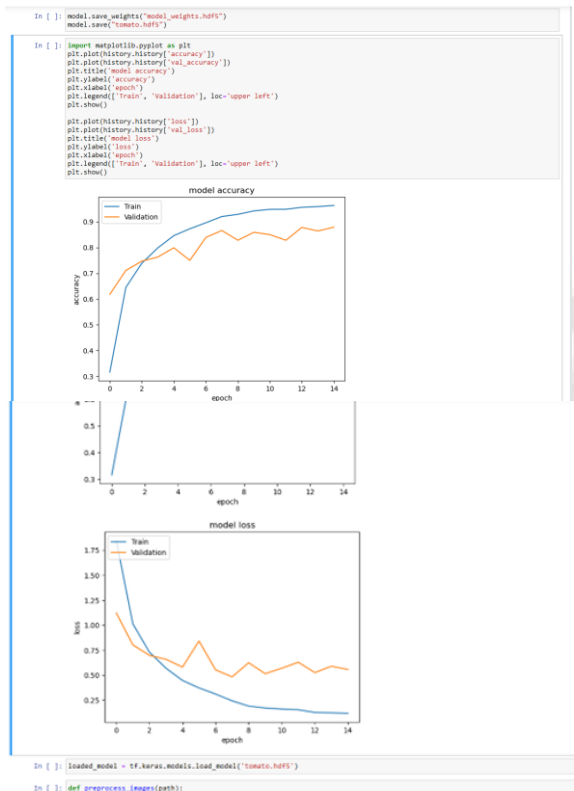
```
In [6]: import cv2
t=cv2.imread('tomato/tomato/test/img1')
plt.imshow(t)
img_path = "tomato/tomato/test/img1"
img = preprocess_images(img_path)
prediction = loaded_model.predict(img)
predicted_class_index = np.argmax(prediction)
class_labels = ["Tomato___Bacterial_spot", "Tomato___Early_blight", "Tomato___Late_blight", "Tomato___Leaf_Mold",
"Tomato___Septoria_leaf_spot", "Tomato___Spider_mites Two-spotted_spider_mite",
"Tomato___Target_Spot", "Tomato___Tomato_Yellow_Leaf_Curl_Virus", "Tomato___Tomato_mosaic_virus",
"Tomato___healthy"]
predicted_class_label = class_labels[predicted_class_index]
print("Predicted class : ", predicted_class_label)
1/1 [-----] - 0s 464ms/step
Predicted class : Tomato_healthy
```



```
In [1]: import cv2
t=cv2.imread('tomato/tomato/test/img1')
plt.imshow(t)
img_path = "tomato/tomato/test/img1"
img = preprocess_images(img_path)
prediction = loaded_model.predict(img)
predicted_class_index = np.argmax(prediction)
class_labels = ["Tomato___Bacterial_spot", "Tomato___Early_blight", "Tomato___Late_blight", "Tomato___Leaf_Mold",
"Tomato___Septoria_leaf_spot", "Tomato___Spider_mites Two-spotted_spider_mite",
"Tomato___Target_Spot", "Tomato___Tomato_Yellow_Leaf_Curl_Virus", "Tomato___Tomato_mosaic_virus",
"Tomato___healthy"]
predicted_class_label = class_labels[predicted_class_index]
print("Predicted class : ", predicted_class_label)
1/1 [-----] - 0s 464ms/step
Predicted class : Tomato___Septoria_leaf_spot
```



According to our methodology we have implemented a feature extraction method to identify the features of the tomato plant and the next step is to implement a detection model to identify the disease of the tomato plant. The detection model we have implemented is Convolutional Neural Networks(CNN).



Conclusion

In this study, we used detection algorithms and feature extraction techniques to address the problem of predicting tomato plant illness. Our results show that Convolutional Neural Networks (CNN) are the best detection model and Principle Component Analysis (PCA) is the best feature extraction technique, with an astounding with highest accuracy rate of 98.49%. We have validated the importance of using PCA for extracting pertinent features from tomato plant photos through our in-depth analysis and experimentation. PCA successfully lowers the dataset's dimensionality while retaining the most crucial data, making disease diagnosis more effective and precise. Our research also demonstrates CNN's supremacy in the field of optimized detection models. CNN successfully learns hierarchical features by utilizing the capability of deep learning.

REFERENCES

1. Ramesh, S., Hebbar, R., Niveditha, M., Pooja, R., Prasad Bhat, N., Shashank, N., & Vinod, P. V. (2018). Plant disease detection using machine learning. At the 2018 International Conference on design innovations for 3Cs Compute communicate control (ICDI3C).
2. Hang, J., Zhang, D., Chen, P., Zhang, J., & Wang, B. (2019). Classification of plant leaf diseases based on improved convolutional neural network. *Sensors (Basel)*, 19(19), 4161. <https://doi.org/10.3390/s19194161>
3. Vamsidhar, E., Jhansi Rani, P., & Rajesh Babu, K. (2019). Plant disease identification and classification using image

processing. *International Journal of Engineering and Advanced Technology (IJEAT)*. ISSN 2249-8958, Vol.8, No.3S.

4. Sannakki, S., & Rajpurohit, V. (2015). Classification of pomegranate diseases based on back propagation neural network. *International Research Journal of Engineering and Technology (IRJET)*, 2(2).
5. Rothe, P., & Kshirsagar, R. (2015). Cotton leaf disease identification using pattern recognition techniques. In *2015 International Conference on pervasive computing (ICPC)*, IEEE, pp. 1–6.
6. Rastogi, A., Arora, R., & Sharma, S. (2015). Leaf disease detection and grading using computer vision technology and fuzzy logic. In *2015 2nd international conference on signal processing and integrated networks (SPIN)*, IEEE, pp. 500–505.
7. Owomugisha, G., Quinn, J. A., Mwebaze, E., & Lwasa, J. (2014). Automated vision-based diagnosis of banana bacterial wilt disease and black SIGATOKA disease. In *International conference on the use of mobile ICT in Africa*, Citeseer
8. Sladojevic, S., Arsenovic, M., Anderla, A., Culibrk, D., & Stefanovic, D. (2016). Deep neural Networks-based recognition of plant diseases by leaf image classification. *Computational Intelligence and Neuroscience*. <https://doi.org/10.1155/2016/3289801>.
9. Mukesh Kumar Tripathi, Dr.Dhananjay, D.Maktedar, "Recent Machine Learning Based Approaches for Disease Detection and Classification of Agricultural Products" *International Conference on Electrical, Electronics and Optimization Techniques (ICEEOT)*, IEEE, ISBN: 978-1-5090-3291-4, DOI 10.1109/ICEEOT.2016.7860043, 2016.
10. R.Meena Prakash, G.P.Saraswathy, G.Ramalakshmi, "Detection of leaf diseases and classification using digital image processing", *2017 International Conference on Innovations in Information, Embedded and Communication Systems (ICIIECS)*, IEEE, ISBN: 978-1-5090-3294-5, DOI 10.1109/ICIIECS.2017.8275915, 2017
11. Dhiman Mondal, Dipak Kumar Kole, Aruna Chakraborty, D. Dutta Majumder, "Detection and Classification Technique of Yellow Vein Mosaic Virus Disease in Okra Leaf Images using Leaf Vein Extraction and Naive Bayesian Classifier", *2015, International Conference on Soft Computing Techniques and Implementations- (ICSCTI)*, ISBN: 978-1-4673-6792-9, DOI 10.1109/ICSCTI.2015.7489626, 2015.
12. S Jafar Ali Ibrahim, Dr.M.Thangamani, "Innovative Drug and Disease Prediction with Dimensionality Reduction and Intelligence Based Random Walk Methods", *International Journal of Advanced Trends in Computer Science and Engineering*, ISSN 2278-3091, Volume 8, No.4, July – August 2019. <https://doi.org/10.30534/ijatce/2019/9384209>
13. Munya A. Arasi, Sangita Babu, "Survey of Machine Learning Techniques in Medical Imaging", *International Journal of Advanced Trends in Computer Science and Engineering*, ISSN 2278-3091, Volume 8, No.5,

- September-October2019.
<https://doi.org/10.30534/ijatcse/2019/398520>
14. Fuentes, A., Yoon, S., Youngki, H., Lee, Y., & Park, D. S. (2016). Characteristics of tomato plant diseases—a study for tomato plant disease identification. In Proc Int Symp Inf Technol Converg (Vol. 1, pp. 226-231).
 15. Sanoubar, R., & Barbanti, L. (2017). Fungal diseases on tomato plant under greenhouse condition. *European Journal of Biological Research*, 7(4), 299-308.
 16. Basavaiah, J., & Arlene Anthony, A. (2020). Tomato leaf disease classification using multiple feature extraction techniques. *Wireless Personal Communications*, 115(1), 633-651.
 17. Tang, Z., He, X., Zhou, G., Chen, A., Wang, Y., Li, L., & Hu, Y. (2023). A Precise Image-Based Tomato Leaf Disease Detection Approach Using PLPNet. *Plant Phenomics*, 5, 0042.
 18. Chowdhury, M. E., Rahman, T., Khandakar, A., Ibtehad, N., Khan, A. U., Khan, M. S., ... & Ali, S. H. M. (2021). Tomato leaf diseases detection using deep learning technique. *Technology in Agriculture*, 453.
 19. Wu, Q., Chen, Y., & Meng, J. (2020). DCGAN-based data augmentation for tomato leaf disease identification. *IEEE Access*, 8, 98716-98728.
 20. Vadivel, T., & Suguna, R. (2022). Automatic recognition of tomato leaf disease using fast enhanced learning with image processing. *Acta Agriculturae Scandinavica, Section B—Soil & Plant Science*, 72(1), 312-324.
 21. Dhinesh, E., & Jagan, A. (2019, December). Detection of Leaf Disease Using Principal Component Analysis and Linear Support Vector Machine. In 2019 11th International Conference on Advanced Computing (ICoAC) (pp. 350-355). IEEE.
 22. Sardogan, M., Tuncer, A., & Ozen, Y. (2018, September). Plant leaf disease detection and classification based on CNN with LVQ algorithm. In 2018 3rd international conference on computer science and engineering (UBMK) (pp. 382-385). IEEE.
 23. Sakkarvarthi, G., Sathianesan, G. W., Murugan, V. S., Reddy, A. J., Jayagopal, P., & Elsis, M. (2022). Detection and Classification of Tomato Crop Disease Using Convolutional Neural Network. *Electronics*, 11(21), 3618.
 24. Hidayatuloh, A., Nursalman, M., & Nugraha, E. (2018, October). Identification of tomato plant diseases by Leaf image using squeezeNet model. In 2018 International Conference on Information Technology Systems and Innovation (ICITSI) (pp. 199-204). IEEE.
 25. Suryawati, E., Sustika, R., Yuwana, R. S., Subekti, A., & Pardede, H. F. (2018, October). Deep structured convolutional neural network for tomato diseases detection. In 2018 international conference on advanced computer science and information systems (ICACSIS) (pp. 385-390). IEEE.
 26. Wang, Q., & Qi, F. (2019, August). Tomato diseases recognition based on faster RCNN. In 2019 10th International Conference on Information Technology in Medicine and Education (ITME) (pp. 772-776). IEEE.
 27. Jasim, M. A., & Al-Tuwaijari, J. M. (2020, April). Plant leaf diseases detection and classification using image processing and deep learning techniques. In 2020 International Conference on Computer Science and Software Engineering (CSASE) (pp. 259-265). IEEE.
 28. Vadivel, T., & Suguna, R. (2022). Automatic recognition of tomato leaf disease using fast enhanced learning with image processing. *Acta Agriculturae Scandinavica, Section B—Soil & Plant Science*, 72(1), 312-324.
 29. Vetal, S., & Khule, R. S. (2017). Tomato plant disease detection using image processing. *International Journal of Advanced Research in Computer and Communication Engineering*, 6(6), 293-297.
 30. Sridhathan, S., & Kumar, M. S. (2018). Plant infection detection using image processing. *International Journal of Modern Engineering Research (IJMER)*, 8(7).
 31. Trivedi, N. K., Gautam, V., Anand, A., Aljhdali, H. M., Villar, S. G., Anand, D., ... & Kadry, S. (2021). Early detection and classification of tomato leaf disease using high-performance deep neural network. *Sensors*, 21(23), 7987.