IOSR Journal of Electrical and Electronics Engineering (IOSR-JEEE)
e-ISSN: 2278-1676, p-ISSN: 2320-3331, Volume 17, Issue 6 Ser. II (Nov. – Dec. 2022), PP 06-16
www.iosrjournals.org

Plant Disease Detection Using Image Processing

L.W CHEN[1], Hidayat Ullah[1]
[1](Ubiquitous Perception and Multi-sensor Intelligence, Fusion Research Institute/ Fujian University of Technology Fuzhou, China.)

Abstract:
Background: *The agriculture sector is extremely crucial for economic growth in every country. Thus, there is a need for disease detection in crops, which is vital in the agricultural area, as a disease in vegetation is a naturally occurring phenomenon. If adequate attention is not taken in this region, it can have critical costs for plants, thus impacting the quality of the output, volume, and productivity. For example, a disease known as "tiny leaf disease" is a dangerous illness that affects pine trees.*
Materials and Methods: *A plant diagnostic system using an automated technique is beneficial since it decreases the time of supervision required in large crop fields and identifies illness at an early stage.*
Results: *This research paper provides an image processing technique that, when combined with algorithms, can be utilised to forecast disease accurately. Consequently, we will focus on picture segregation, image fragmentation-processing data, identification, and recognition of characteristics as some of the ways utilised in the identification of infections.*
Conclusion: *The project aims to identify the most frequent disorders on a tomato leaf, such as bacterial spots, early blight, and curl, utilising image processing techniques and advanced technologies like machine learning. In layman's terms, the farmer could precisely diagnose the sort of disease a certain plant is suffering from by examining the image of the crop.*
Key Word: *Plant Disease, Image Processing, Image Acquisition, Disease Detection.*

Date of Submission: 04-12-2022 Date of Acceptance: 16-12-2022

I. Introduction

It is critical to introduce technical advances in the domain of crop yield to nations such as India and China, where population growth is relatively high. In the significant subject of qualitative farming, scientific endeavours and preliminary study processes are focused on enhancing yield and agricultural production standards cheaply, with a higher monetary consequence. The agricultural construction model results from a complex integration of soil, seeds, and growth-promoting chemicals. Fruits and vegetables are currently one of the most important agricultural products. Product value evaluation and enhancement have always been critical in acquiring excess and desirable items. Diseases impair the plant's normal condition, affecting or preventing vital functions like photosynthesis and propagation [1].

Pathogens such as fungi, bacteria, and viruses cause illnesses due to poor environmental conditions. As a result, the preparation phase for plant disease diagnosis is a critical effort. Farmers require expert inspections regularly, which may be excessively expensive and time-consuming. As a result, finding agile, less expensive, and more exact approaches to identifying infections from signs that appear on the plant's leaves is critical. This paper proposes a technique for determining the specific illness that a plant may be suffering from. A key problem is identifying the type of infection that an essential crop like tomatoes can have by adopting modern techniques like image recognition, which graphically reflects the software's functioning [2]. It is also a significant reason for the widespread adoption of technology.

Machine vision is also reinforced, which allows for image-based automation in supervision, assessment, and robot guidance. Plant pathogen recognition by naked human eyesight is uneconomical, time-wasting, and yields poor results. Automatic detection requires fewer efforts, is quicker, and is more precise. "Brown and yellow spots, early and late scorch, and fungal, viral, and bacterial infection" are examples of common plant diseases. Image processing is employed to spot the variation in coloration of the afflicted section and to quantify the damaged area. The technique of dividing or categorising a picture into various portions is known as "image segmentation," which can be accomplished in a variety of ways, spanning from simple thresh holding to complex colour image segmentation approaches. These components are usually associated with anything that humans can effortlessly distinguish and perceive as independent things. Because computers lack the ability to recognise elements intelligently, numerous diverse methods for segmenting photos have been devised. The segmentation procedure is founded on the image's numerous attributes. This could be colour information, the boundaries of an image, or a fragment. For colour picture segmentation, we employ a genetic algorithm [2-5].

In agriculture, different individuals and technology organisations are working to boost output and productivity. Many strategies have been employed in the past to overcome problems with disease transmission in plants. For instance, the tomato plant diagnostic system has gotten easier and more accurate as technology has evolved. In this system, a different technique is adopted, namely the KNN algorithm. Recently, various approaches have been designed to estimate the types of plant pathogens. Some of them involve the research and study of biochemical analytical methods to diagnose plant illnesses, as well as indirect methods, including the use of physical methods such as leaf spectroscopy and scanning to get data on plant features. Because it combines statistical machine learning and analysis algorithms, the suggested method for plant disease identification is operationally less costly and takes less time to forecast than existing deep learning-based systems [4-7].

II. Problem Statement

Crop infections have become an issue since they considerably diminish the quantity and quality of produce. Automatic disease detection is an essential research issue since it has the potential to help vastly screen fields of crops and, as a consequence, identify symptoms as early as possible on a leaf image. This initiative attempts to provide knowledge to local farms and farming specialists. Growers will be more inclined to use this approach when concerned about time. Therefore, they will obtain all of the current updates on their crops thanks to machine learning and image analysis technologies. The technology uses photos of infected plants and a machine-vision algorithm to diagnose infection.

Additionally, plant pathogens have been a pressing issue for growers in recent years. Because they do not even understand what infections contaminated plants have, they tend not to know what agrochemicals to use to treat them. Consequently, the wrong chemicals are applied, resulting in crop damage and reduced yield [5].
Some of these diseases include;

Early and late blight of potato and tomato

It is reported that late blight disease is particularly important with regard to potatoes and tomatoes. They illustrate this by reporting losses of up to 100%. This sentiment is supported by the work of [6], which showed that the incidence and prevalence of the causal pathogen Alternaria sp. were widespread in Algeria. The disease appears as dark brown or black lesions with concentric rings on leaves, producing a "target spot" effect; these symptoms appear on plants that are already under physiological and mechanical stress [7]. In potato and tomato late blight diseases, a similar pathogen causes a similar disease.

Late blight of potatoes and tomatoes

Late blight is caused by Phytophthorainfestans, toward which significant attention has been directed. Since it is not only resistant to fungicides but is becoming more virulent and capable of infecting resistant varieties [8, 9].

Alternaria leaf spot

Alternaria has also been implicated in causing spots on the leaves and fruits of pomegranates [10]. The symptoms include black spots on leaves that appear irregular in shape and are surrounded by yellow haloes [11]. Similar to the other mentioned pathogens, Alternaria leaf spots can also cause considerable economic losses [12].

Anthracnose of Colletotrichum

Anthracnose leaf spots are caused by several phytopathogens within the genus Colletotrichum [13]. The spots appear as brown to dark lesions on the leaves or cankers on the stems of infected plants, and it is not rare to observe the total wilting and dying of infected plants [14], [15]. There is evidence that the prevalence of this disease is increasing [15-18].

Bacterial blight

As a case study, bacterial blight caused by several species of Xanthomonas sp. has had devastating effects on cotton fields [19] and rice paddies [20]. It can cause widespread crop damage and yield losses of 30–50% [21].

To tackle this problem, we created a strategy for developing a system that can swiftly detect various common infections that harm tomato crops by merely glancing at their foliage.

Plant infections pose a serious risk to the agriculture system, with the ability to drive the entire human population to famine if not recognized. Using a machine learning framework in the field of plant pathology will make the diagnosis of plant infections easier and less expensive, assisting many producers in the early identification of plant infections, minimising waste, and avoiding the transmission of pathogens from afflicted to healthy crops. Many studies on plant disease detection give a comparative analysis utilising multiple machine learning algorithms; however, they fail to elaborate on the forecasts that their models make [7-9]. This study aims to compare the functioning of a basic and sophisticated model and offer interpretability for the models' projections. The inclusion of explain ability is motivated by the fact that most machine learning techniques frequently employed in this sector are black-box frameworks, which result in consumers not relying on or knowing how their models create their projections. The introduction of explainable artificial intelligence approaches de-mystifies these black-box systems, allowing users to efficiently comprehend their model forecasts and decide whether to believe them [9-11]. The use of XAI aids greatly in implementing plant infection detection since the visibility and interpretability of the algorithms used are critical in obtaining the faith of farmworkers, whose existence is based on the growth of healthy crops.

III. Literature Review

In the past, a great deal of effort has been spent on identifying leaf infections using image processing, and this topic continues to draw researchers. In recent years, automatic crop infection identification utilising machine learning and image processing has gained popularity. This literature focuses on analysing past studies related to this topic to understand the extent to which the effectiveness of the proposed technology has been employed. Disease identification is one of the most important aspects of agriculture that must be addressed. Even though numerous techniques have been developed and applied to address this issue, swift and accurate disease identification remains a work in progress. Machine learning aids in diagnosing and characterising threats, which helps mitigate the condition on a larger scale.

Erdogan M. et al. in 2018 [19] proposed a framework with an integrated learning vector quantization and convolutional neural networks (CNN) for classifying and identifying infections in a tomato leaf. The proposed model was tested on 500 leaf samples from four tomato disease classes. CNN was employed for extracting the main characteristics from the image, along with its categorization. On the same topic, Walleligm, S. et al. (2018) [20] analysed the feasibility of the CNN framework for categorising different plant infections with the help of leaf images. This model was employed by embracing LeNet, one of the most popular CNN models for infection categorization in soybean crops. The leaf images of approximately 12 species and 763 samples were acquired from the standard database regarded as Plant Village. This model achieved a precision level of 99.32%, showing the feasibility and suitability of CNN when applied to classifying plant infections using a leaf's images.

Sladojevic, S., et al. [3] focused on developing a technological framework for identifying different illnesses in 13 crops from photographs of healthy plant leaf surfaces in 2016. The researchers employed a deep learning model known as Caffe for data training. The findings were generated from the Caffe model with a precision level of around 91–98%.

Fuentes, A. et al. (2017) [4] recommended a mode that can be employed in two phases. First, the meta-models of SSD, R-FCN, and Faster R-CNN would be integrated into a single model. Eventually, other approaches, such as ResNet-50 and VGG-19, were included to obtain the characteristics in more depth, and the efficiency of this framework was approximated. After contrasting them with other frameworks, the recommended model was better. Arivazhagan, S., and Ligi, S. V., in 2018 [5], suggested a model based on automatic deep learning for the classification and recognition of different infections of mango trees. The data used for this model entailed 1,200 images that encompassed both healthy and infected leaves from mango trees. The precision recorded from the proposed model was 96.67%.

In 2020, Uguz, S., and Uysal, N. [6] compared a transfer learning simulation with CNN models like VGG-16 and VGG-19, as well as recommended CNN designs, in the situation of olive plant infections. The dataset contained 3400 photos of olive plant leaves, which were used to create the architecture. A deep learning mechanism was used in this approach to increase the dataset's size. The accuracy before data augmentation was around 88 per cent, and after data augmentation, it was around 95 per cent. M. Agarwal et al. (2020) [7] established a tailored model for disease detection in tomato plants based on CNN. In addition, the suggested model was compared to machine learning algorithms and VGG-16. The suggested framework had a 98.4 per cent precision, the KNN system had a 94.9 per cent accuracy, and the VGG-16 design had a 93.5 per cent accuracy [8, 9]. The database of tomato leaf pictures used in this approach was taken from the Plant Village database.

In 2017, P. Moghadam et al. [10] showed how hyper spectral imaging could be employed to detect crop diseases. The study used VNIR and SWIR spectra as well as a k-means algorithm for clustering in the spectral sphere for the leaf clustering. They also recommended a central grid removal algorithm to eliminate the grid from the hyper spectral images. The researchers achieved an 83 per cent accuracy level with foliage indices in VNIR spectral scope and a 93% accuracy level with full spectrum. Although the recommended model attained high accuracy, it required the use of a hyper spectral camera with 324 spectral bands, which resulted in an expensive solution.

In their research, Zahid Ullah et al. [11] provided a precise methodology for contrast enhancement. There are a variety of instances in which some or all of the image's valuable information is destroyed, altered, or deleted. As a result, the author suggested that we employ filtering. To remove some noise, the photograph is passed through a median filter. For effective analysis, an image's contrast must be high. All of the image's objects are apparent because of the high contrast. The authors of this study propose a very successful method for histogram equalization by reducing contrast. This method is superior to histogram equalisation since it eliminates the problem of over-amplification.

Waghmare and R. Kokare (2016) [12] developed a paradigm for identifying plant infections based on leaf textures. The system receives a coloured file as input, subdivided to detect the contaminated area in the photograph, and a special portion of the leaf is acquired. A texture-based model is created depending on the attributes. Every new type of leaf disease has a distinct leaf surface. This is the information that the SVM classifier uses to determine the condition. In this study, the authors employed a multiclass SVM classifier to characterise and diagnose disease in photos of grapevine leaves. The imaging trend is then classified as a multiclass SVM categorization in either normal or unhealthy subgroups. The research focused on black and red downy mildew diseases, two of the most popular and worst-affecting diseases. The proposed system provides 96.6 per cent accurate expert advice to tenants in a timely manner.

Shima Ramesh and Ramachandra Hebbar et al. (2018) [13] developed effective methods for identifying leaf-based infections. Random forest is used in this suggested study to create a database by detecting healthy or contaminated leaf photos. The suggested work is divided into four phases: dataset identification, feature extraction from leaf images, database identification, function generation, classifier extraction, and categorization. The datasets generated for contaminated and healthy leaves are combined and trained using the Random Forest classifier for the classification of infected and healthy films. The histogram-oriented gradient (HOG) has been used to derive valuable features from photographs. In general, using machine learning to train large datasets offers a straightforward and effective method for detecting numerous plant infections. The model in this study was trained on 160 papaya leaves using the Random Forest Classifier technique. With a 70% accuracy rate, the algorithm could be used to classify diseases. With such a high number of photos and other global and regional characteristics, the framework's accuracy can be improved.

Jitesh P. Shah et al. (2016) [14] suggested a method for examining diseased leaves employing an image processing method and a machine learning system to identify infected plant leaves via photographs. The authors not only examined the plants but also addressed essential ideas in digital image processing and machine learning for pathogen detection and diagnosis. This planned work was carried out in depth, with the author conducting 19 tests on multiple disease classes. This study was also incorporated into a survey for future studies and system upgrades. This paper offers taxonomy of several paddy plant infections. Mckay et al. (2008) [15] have classified the algorithm into three components. Firstly, any image collected from any resource must be demodulated. The picture is then subjected to wavelet evaluation and neighbourhood pixel assessment. The assessment of pixels, or pixel contrast with neighbouring pixels, is performed in this step to ensure that the picture improvement is accurate. This is performed to see if any significant data is lost during the image smoothing procedure.

Zahid Ullah et al., 2020 [16], have developed a precise image enhancement method. There are a variety of situations in which some or all of the image's valuable information is destroyed, altered, or destroyed. As a result, the first author suggests that we employ filtering. To remove some noise, the image is passed through a median filter. For efficient evaluation, an image's contrast must be high. All of the image's objects are apparent because of the high contrast. The author of this study proposes a very successful method for histogram equalization by reducing contrast [17]. This approach is superior to histogram equalisation since it eliminates the problem of over-amplification.

Numerous models and studies have been established using deep learning models to identify and classify different categorical maladies in a particular plant. Deep learning could also identify and classify macronutrients in a given plant. For example, Tran, T. T., et al. [18] suggested a system founded on a deep learning simulation in 2019 that provides a monitoring framework that monitors many phases from germination to yielding to obtain a higher production rate. The suggested model was tested with a dataset of 571 photos, including tomato leaf and fruit photographs from different phases of the crop's lifecycle. The auto-encoder and inception-ResNet v2 achieved a precision of 87.27 per cent and 79.09 per cent, respectively. The influence of transfer learning on the identification and categorization of plant infections using leaf pictures is demonstrated in this research review.

In 2020, S. Hernandez et al. [19] raised a fundamental worry concerning prediction uncertainty. Making a prognosis on an unknown sample where the algorithm has not been trained is challenging and unpredictable. This can be evaluated by including uncertainty in the prediction. The author proposes that Bayesian deep learning methods be used to solve this challenge. In this case, the misclassified output can be considered a source of uncertainty. The system is trained using a deep CNN architecture to identify the diseased portion. This article utilises three optimization algorithms: stochastic gradient MCMC, stochastic gradient descent, and MC dropout. From these three, the SGD and MC dropout algorithms produced overly confident projections, while SGLD produced less confident prognostications based on stochastic entropy. Among the three, it is the most accurate. The major parameter utilised to train the deep CNN was image entropy.

IV. Methodology

The study will also entail several steps that are integral to collecting the necessary data for analysis. The initial stage is acquiring the images that will be used for the analysis. This stage entails capturing images using an iPhone or digital camera with a high resolution. Afterward, skewers filters and Kurtosis will be used to smooth the captured images [19]. Thirdly, the research technicians will conduct image segmentation by using an inverse difference approach to isolate the areas affected by the disease from the healthy ones. This stage is crucial because it ensures that there are both healthy and affected leaves. After the images have been segmented, the extraction of features is performed. Segmentation is the procedure whereby the colour and texture of the image are reconsidered to acquire special features from it.

a. Datasets

Massive amounts of raw data were trained in order to extract useful information from the dataset. The system's dataset is a collection of images of various plant leaves that are either normal or unhealthy. To provide the best reliability and precision, the infection diagnosis methodology necessitates a huge database. Every one of these pictures is resized and enhanced to a similar size and quality in order to build a uniform dataset. This dataset will be used to train the program's digital image processing element in the next stage.

Figure 1 Sample leaves with diseases and pathogens for different conditions.

b. Image Pre-Processing and Labelling

Image pre-processing was employed to alter or enhance the raw files that were required to be analysed by the CNN classifier prior to training the algorithm. Creating a successful model necessitates a thorough examination of both the network's design and the input data's structure. The researchers pre-processed the dataset such that the suggested framework would be able to extract the relevant traits from the image. The initial step was to scale the image to 256 by 256 pixels and equalise its size [19]. The photos were then converted to grey scale. For the explicit learning of the training data features, this pre-processing phase necessitates a large quantity of training data. The next stage was sorting the tomato leaf photos by category, and then labelling each with the condition's abbreviation. In this scenario, the database revealed ten test collections and training classes.

c. Training Dataset

The initial step in processing a pre-existing dataset was to prepare it. During this phase, image data was loaded into a CNN process, ultimately forming a model that evaluated performance. Figure 2 shows the normalisation stages for tomato leaf pictures.

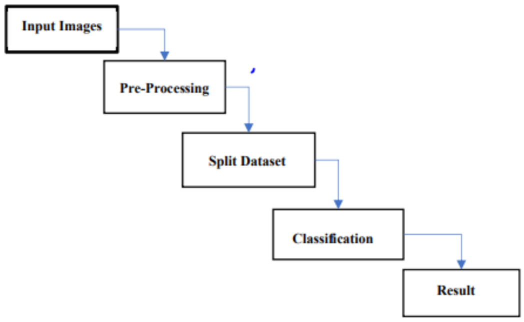

Figure 2 Normalization stages for tomato leaf pictures.

d. Convolutional Neural Network

The CNN is a neural network innovation frequently used to handle or train image content. The convolution matrix structure is intended to screen the images. For data training, the CNN employs each tier, including the convoy layer, input layer, dropout layer, fully connected pooling layer, and finally the linked dataset categorization layer. This could map a sequence of computations to every layer's input testing set. Figure 3 depicts the entire architecture.

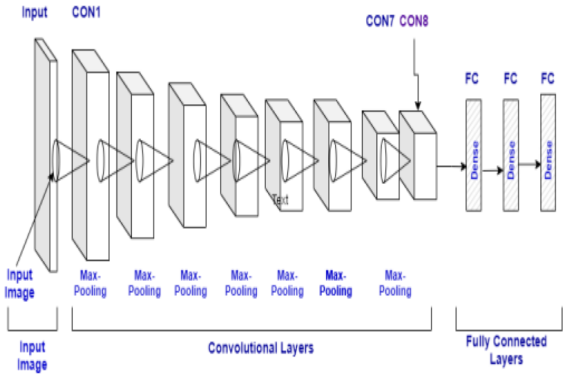

Figure 3 The CNN model architecture

e. Convolutional Layer

By using the convolution technique with the presentation layer, a convolution layer is utilised to map attributes. Every map function is paired with a variety of input parameters. CNNs are built on the foundation of convolution, a two-function procedure [20]. A map, or 2D function map, is formed once each filter is transformed into every part of the input data. Because of the model's intricacy, substantial layer convolutional function tuning is required.

f. Pooling Layer

To optimise and enhance accuracy, the pooling layer progressively increases the number of variables. Moreover, as the parameters increase, the area of the maps decreases. The pooling layer reduces the convolution layer's overall output. It decreases the number of trainable parameters by considering the spatial features of a country-wide area.

g. Computer image processing

The image processing system is a complicated system with several algorithms and stages. The goal of this stage, when considered as a whole, is to train on the photos that have been pre-composed, collected, and recorded as the dataset. Perform tests on the image of the plant leaf we captured to see if it's diseased. The information gathered is used to build training data for the image processing algorithm. Afterward, the model is saved and used to evaluate the camera's collected images.

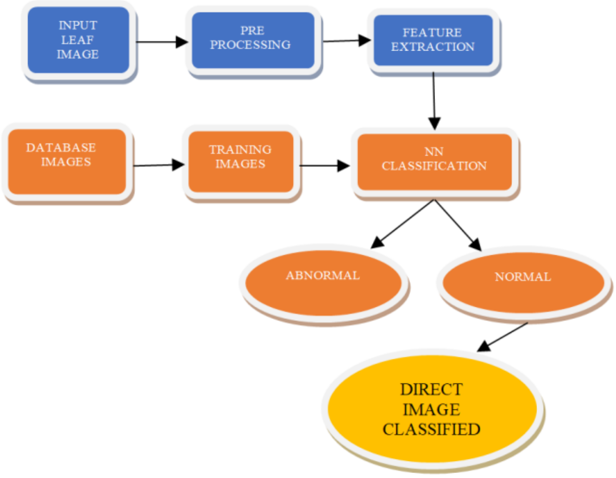

Figure 4 Stages and algorithms used in Image processing

h. Feature extraction

It is a dimensional reduction approach that helps in the representation of the interesting aspects of a photograph into a succinct feature vector. Feature extraction is an approach that employs a wide variety of resources to characterise a large amount of data accurately. The "minimum distance" classifier and the supporting "vector machine classifier" are used one at a time in the proposed classification technique.

$$P_{i,j} = \sum_{z=0}^{M-1} \sum_{y=0}^{N-1} C[I(x,y) = i \text{ and } I(x \pm d\emptyset_1, y \pm d\emptyset_2) = j]$$

This strategy is particularly beneficial when large image sizes and simpler feature illustrations are required to complete tasks quickly. In the system, we use the following feature extraction techniques: Figure 5 shows a histogram (historically oriented gradients). To gather characteristics, it estimates the impact of the gradient direction. It divides the image into different sections and creates a HOG orientation distribution across all of the photographs [20]. HOG is used in a variety of object recognition applications, including face recognition and, in this case, tomato leaf recognition. The oriented gradient histogram is employed in different research centres with considerable effectiveness.

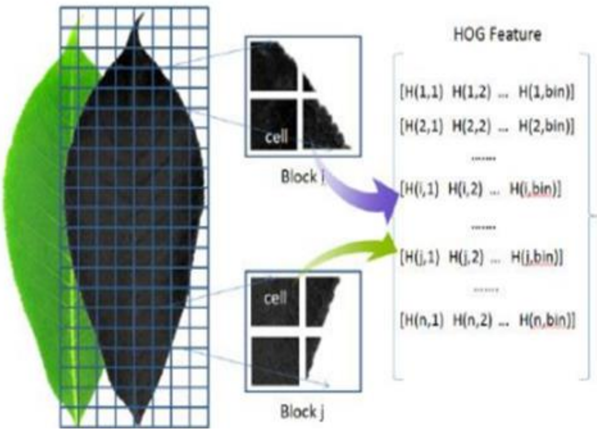

Figure 5 Histogram of oriented gradients with their features

i. **Image Segmentation**

The next procedure is picture segmentation, part of the data collection process. Typically, an image is surrounded by unwanted items or a backdrop. This necessitates the use of image segmentation methods. Image segmentation is a computerised processing method that separates a digital image into multiple relevant or focused pieces necessary for the issue area. It usually works with image pixels of comparable quality. It was used to separate a photograph's background and foreground. When the extraction is conducted, this method aids in extracting the proper characteristics [21]. Machine-driven thresh holding approaches called Otsu's methodology are used in this study for segmentation based on the results of several experiments.

Grey Scale Co-occurrence Matrix

The statistical texture impression visible in the image histogram is displayed and calculated using the texture filter. They are necessary for giving appropriate texture effects in a picture, but they cannot provide shape metadata, which would be the spatial correlation of image pixels. Feature extraction needs a reduction in the resources needed to manage a large volume of data.

Figure 6 GLCM extracted textual features

Classification: KNN Classification Algorithm

The datasets that were trained and evaluated against the training model make up the last stages of our image processing cycle. This classification model employs the KNN approach. The KNN method is also regarded as a "supervised machine learning algorithm," since it may be used to tackle both regression and classification tasks. This algorithm is based on the idea that extremely similar items exist in close proximity or that similar things exist in immediate contact with one another. We started by inserting data and setting K to the required number of neighbours in the KNN algorithm (Figure 7). For every example in the database, the distance between the query and the present example is determined. The data were organised in order of variation in ascending order, and the first k elements from the organised collection were selected along with their markings [21]. The algorithm returns the mean when there is regression, but the method returns the mode when there is categorization. Consequently, this program processes our photographs and sorts them into disease classes. This is a very precise and accurate method that produces high-quality results.

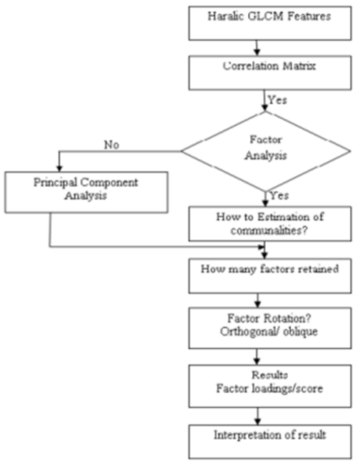

Figure 7 The KNN Classification Algorithm

Training

The process of acquiring kernels in weights and convolution layers in entirely connected layers to decrease disparities amongst output projections and predefined ground truth labels on training data is known as "training a network." The study used 70 per cent of the data for training in this step. The network trains by identifying characteristics in plant leaf infection photos and learning from them so that each image can be discriminated on its own basis.

V. Results And Discussion

Our suggested technique combines naive Bayes with a technique for picture recognition. The initial level of verification provides a modest forecast of plant pathogens, whereas the second-level forecast is complete. The leaf collection for image identification is shown in Figure 1.

The tests are run on a computer with MATLAB R2018b, 4 GB of RAM, and a Core i5 4300U CPU running at 2.50 GHz. The researchers trained the network and preserved the trained network such that the training procedure was not duplicated and the time spent training was recorded [22]. The network is then evaluated by examining data and demonstrating its correctness. The last step is where plant-leaf infections are identified and categorized. For each photograph, a random choice is produced; here, we've picked a sample of tomato plant paper; after putting it into the system, selecting the trained network, and pressing the detection key, the infection and kind of the afflicted crop will be disclosed.

Multiple photos of each condition are taken to build datasets. Photographs are acquired from numerous camera sources and at different places within a 100-kilometer radius, and certain standard pictures from agriculture-related institutes are also utilised in the dataset preparation. The ratio of successfully identified picture samples to the overall number of test sample photographs is used to calculate the percentage accuracy [22]. Because of the time constraints and continuing pandemic, the machine learning algorithms used in this research were optimised to the highest standards possible. For instance, the value of "K" was selected by validating the model on all figures between 1 and 25, and the "K" that was eventually selected generated the highest precision and lowest margin of error. Likewise, the model parameters of the CNN model, which achieved 98.5 per cent accuracy, were tested several times until a steady model was generated that was neither over fitting nor under fitting.

After training the model, illness-affected photos are acquired, and different procedures such as pre-processing, extraction of features, segmentation, feature fusion to merge all characteristics, and disease categorization are performed. Researchers employed state-of-the-art machine learning algorithms such as SVM and KNN for the categorization challenge [23]. For the experiment, the confusion matrix is employed as an assessment variable. Based on the confusion matrix, several computations are done, such as those for precision, recall, accuracy, and selectivity.

In agricultural settings, machine learning-based approaches for detecting plant infections and pests contain three key connections: data labelling, model inference, and model training. Model inference receives increased attention in real-time agricultural settings. Most plant conditions and pest detection approaches now rely on identification accuracy. The effectiveness of model inference goes unrecognized. Deep, differentiated convolutional structure architecture for leaf disease detection was presented to increase the model's efficiency and satisfy real agricultural demands.

In contrast to traditional image processing techniques that focus on plant infection detection activities in several stages and connections, machine learning-based plant pest and pathogen detection systems consolidate them into end-to-end extracting features with extensive development opportunities and enormous promise. While plant infection detection technology is rapidly evolving, shifting from scientific work to agricultural applications, there are still some improvements required to make it available in the real world, and some issues must be resolved.

In the diagnosis of crop diseases and pathogens, machine learning has made some headway. Different image recognition algorithms have been improved and expanded, laying the groundwork for detecting specific infections and plagues on a theoretical level. Past studies, on the other hand, collected picture samples primarily from identifying disease patches, insect visual features, and classifying pest species and foliage.

The CNN's efficiency in object detection and picture classification has advanced gradually in recent years. Formerly, image classification issues were tackled using hand-engineered elements such as HoG, SURF, and others, followed by some learning algorithms in these image features. As a result, the effectiveness of these methods was heavily dependent on their fundamental, predefined characteristics. Feature engineering is a time-consuming and complicated procedure that must be addressed whenever the situation at hand or the accompanying dataset evolves meaningfully [24]. This challenge arises with all the recent image-processing techniques that can detect plant diseases because they rely primarily on locally available technologies, picture augmentation approaches, and various other complicated and labour-intensive procedures.

Furthermore, classic machine learning techniques for disease categorization often concentrate on a restricted classifier, typically within a specific crop. To categorise tomato powdery mildew versus normal tomato leaves, a feature classification and extraction pipeline employing thermal and stereo images was used, and to identify powdery mildew in unconstrained conditions, RGB images were used. In detecting orange huanglongbing, infrared spectral patterning and plane-based sensors using fluorescence imaging spectroscopy were employed. Identifying tomato yellow leaf curl infection employs a series of traditional feature extraction procedures, accompanied by categorization using an SVM pipeline and a variety of additional techniques [23]. The work in this area is thoroughly discussed in a recent study on the application of machine learning to crop morphological characteristics. While CNN was previously used to identify plant diseases, the method needed the pictures to be represented using a hand-picked set of textural attributes before the neural network could categorise them.

The study's societal implications may be divided into two categories: favourable and harmful. The beneficial effects of this research include that it enables local farmers to be self-sufficient and not reliant on professional expertise; it allows for inexpensive, large-scale diagnosis of plant diseases; and it allows local farmers to benefit from their healthy produce. Having the diagnosis of plant diseases available to everybody and everyone aids in meeting human civilization's food demands, assisting society in avoiding famine and health concerns.

The negative repercussions of this study include its influence on many community members, such as specialists who used to physically diagnose plant illnesses or companies that hire workers to detect plant infections. As new illnesses arise, one possible remedy for this detrimental impact is to incorporate expert knowledge into refining the model's projections [25]. As a result, the specialists' job security and the systems' on going improvement are reinstated. It is crucial to note that the societal implications of this study have not been evaluated and that the societal implications listed above are conceivable.

VI. Conclusion

The suggested technique concentrates on developing an effective and sophisticated system that simplifies the process of producing a good return of tomatoes for growers. The project aims to identify the most frequent disorders on a tomato leaf, such as bacterial spots, early blight, and curl, utilising image processing techniques and advanced technologies like machine learning. In layman's terms, the farmer could precisely diagnose the sort of disease a certain plant is suffering from by examining the image of the crop [24]. The recommended model is founded on four main modules: segmentation, classification using KNN, feature extraction, and pre-processing.

In this investigation, we compare our system against pre-existing solutions using a suitable approach and execution. The suggested system's capabilities outperform the existing infection detection system since it produces more accurate and precise results while being faster and easier to deploy. Its goal is to make farming simpler. The team will benefit the agriculture industry by advancing farm productivity and control, as farming is a key contributor to the country's economic growth.

In a future study, researchers want to enhance the model to incorporate specific abiotic illnesses caused by nutrient deficiencies in crop leaves. The long-term goal is to acquire more unique data and amass a large amount of information on plant diseases. Future studies will use advanced technologies to improve accuracy.

References

[1]. Erdogan, M., Tuncer, A., and Ozen, Y.: Plant Leaf Disease Detection and Classification Based on CNN with the LVQ Algorithm. In: 3rd Int. Conf. Comput. Sci. Eng. (2018) 382– 385

[2]. Wallelign, S., Poland, M., and Buche, C.: Soybean plant disease identification using a convolutional neural network. In: Proc. 31st Int. Florida Artif. Intell. Res. Soc. Conf. FLAIRS 2018 (2018), 146–151

[3]. Sladojevic, S., Arsenovic, M., Andela, A., Culibrk, D., and Stefanovic, D.: Deep Neural Networks Based Recognition of Plant Diseases by Leaf Image Classification. Comput. Intell. Neurosci. 2016(2016)

[4]. Fuentes, A., Yoon, S., Kim, S. C., and Park, D. S.: A robust deep-learning-based detector for real-time tomato plant diseases and pests recognition. Sensors (Switzerland) 17 (2017).

[5]. Arivazhagan, S. and Ligi, S. V.: Mango Leaf Diseases Identification Using Convolutional Neural Network. Int. J. Pure Appl. Math., 120(2018) 11067–11079

[6]. Uğuz, S. and Uysal, N.: Classification of olive leaf diseases using deep convolutional neural networks. Neural Comput. Appl., 5 (2020)

[7]. Agarwal, M., Gupta, S. K., and Biswas, K. K.: Development of Efficient CNN model for Tomato crop disease identification. Sustain. Comput. Informatics Syst., 28 (2020) 100407

[8]. S. D. Khirade and A. B. Patil, "Plant Disease Detection Using Image Processing," 2015 International Conference on Computing Communication Control and Automation, 2015, pp. 768-771, doi: 10.1109/ICCUBEA.2015.153.

[9]. S. C. Madiwalar and M. V. Wyawahare, "Plant disease identification: A comparative study," 2017 International Conference on Data Management, Analytics, and Innovation (ICDMAI), 2017, pp. 13-18, doi: 10.1109/ICDMAI.2017.8073478

[10]. P. Moghadam, D. Ward, E. Goan, S. Jayawardena, P. Sikka, and E. Hernandez, "Plant Disease Detection Using Hyperspectral Imaging," 2017 International Conference on Digital Image Computing: Techniques and Applications (DICTA), 2017, pp. 1-8, doi: 10.1109/DICTA.2017.8227476

[11]. Ullah, Z., Farooq, M. U., Lee, S. H., &An, D. (2020). A hybrid image enhancement-based brain MRI images classification technique. Medical Hypotheses, 143(April), 109922. https://doi.org/10.1016/j.mehy.2020.109922

[12]. H. Waghmare, R. Kokare, and Y. Dandawate, Detection and classification of diseases of Grape plant using opposite color Local Binary Pattern feature and machine learning for automated Decision Support System 2016, 3rd Int. Conf. Signal Process. Integ. Networks, 513–8.

[13]. S. R. Maniyath et al., Plant disease detection using machine learning 2018, Proc. 2018 Int. Conf. Des. Innov. 3Cs Comput. Commun. Control., 41–5.

[14]. J. P. Shah, H. B. Prajapati, and V. K. Dabhi, A survey on detection and classification of rice plant diseases 2016, 2016 IEEE Int. Conf. Curr. Trends Adv. Comput.

[15]. Chohan M, Khan A, Chohan R, Kaspar SH, Mahar MS. Plant disease detection using deep learning. International Journal of Recent Technology and Engineering. 2020;9(1):909-14.

[16]. Ullah Z., Farooq MU., Lee SH., An D. A hybrid image enhancement based brain MRI images classification technique. Medical hypotheses. 2020 Oct 1;143:109922.

[17]. Konstantinos P. Ferentinos. "Deep learning models for plant disease detection and diagnosis Computers and Electronics in Agriculture, 145 (2018) 311318.

[18]. Tran, T. T., Choi, J. W., Le, T. T. H. and Kim, J. W.: A comparative study of deep CNN in forecasting and classifying the macro-nutrient deficiencies on development of tomato plant. Appl. Sci., 9 (2019).

[19]. Hernández, S., and Juan L. López. "Uncertainty quantification for plant disease detection using Bayesian deep learning." Applied Soft Computing 96 (2020): 106597.

[20]. Mohanty SP, Hughes DP, Salathé M. Using Deep Learning for Image-Based Plant Disease Detection. Front Plant Sci. 2016 Sep 22;7:1419. doi: 10.3389/fpls.2016.01419. PMID: 27713752; PMCID: PMC5032846.

[21]. A. Camargo and J. S. Smith, "An image-processing based algorithm to automatically identify plant disease visual symptoms," Biosyst. Eng., vol. 102, no. 1, pp. 9-21, 2008.

[22]. S. Sankaran, A. Mishra, R. Ehsani and C. Davis, "A review of advanced techniques for detecting plant diseases," Computers and Electronics in Agriculture, vol. 72, no. 1, pp. 1-13, 2010.

[23]. P. Chaudhary, A. K. Chaudhari, and S. Godara, "Color Transform Based Approach for Disease Spot Detection on Plant Leaf," Int. J. Comput. Sci. Telecommun., vol. 3, no. 6, pp. 65-71, 2012.

[24]. S. Arivazhagan, R. N. Shebiah, S. Ananthi, and S. Vishnu Varthini, "Detection of the unhealthy region of plant leaves and classification of plant leaf diseases using texture features," Agric. Eng. Int. CIGR J., vol. 15, no. 1, pp. 211-217, 2013.

[25]. S. D. Khirade and A. B. Patil, "Plant disease detection using image processing," Proceedings-1st International Conference on Computing Communication Control and Automation ICCUBEA 2015, pp. 768-771, 2015.

L.W CHEN, et. al. "Plant Disease Detection Using Image Processing." *IOSR Journal of Electrical and Electronics Engineering (IOSR-JEEE)*, 17(6), 2022, pp. 06-16.