International Journal of Early Childhood Special Education (INT-JECS)
DOI:10.9756/INTJECSE/V14I7.52 ISSN: 1308-5581 Vol 14, Issue 07 2022

Early-Stage Prediction & Detection of Guava LeavesDisease Using Genetic Algorithm

Running Title: - Early-Stage Prediction & Detection

Kartika Mehta[1]*, Himashu Kothari[2], Dr.Vikas Somani[3], Dr. Awanit Kumar[4]

*[1] M.Tech. Scholar dept. of computerscience Sangam University Bhilwara, Rajasthan, India
E-mail:- kartikamehta022@gmail.com

[2] Assistant Professor Dept. Of computerscience Sangam University Bhilwara, Rajasthan, India
E-mail:- himanshu.kothari@sangamuniversity.ac.in

[3] Associate Professor Dept. Of computerscience Sangam University Bhilwara, Rajasthan, India
E-mail:- vikas.somani@sangamuniversity.ac.in

[4] Assistant Professor Dept. Of computerscience Sangam University Bhilwara, Rajasthan, India
E-mail:- awanit.kumar@sangamuniversity.ac.in

***Corresponding Author: - Kartika Mehta**

*[1] M.Tech. Scholar dept. of computerscience Sangam University Bhilwara, Rajasthan, India
E-mail:- kartikamehta022@gmail.com

Abstract
Agriculture is both a source of income and a means of subsistence in India's economy. As a result, early detection of disease and analysis of fruit plants are more crucial for the growth of a healthy fruit corporation. If bacteria infect plant leaves, it will ultimately damage crop yieldand result in significant agricultural losses. It may also result in a decline in crop quality. As a result, identifying plant and leaf diseases is critical. Pesticides are commonly used on plants to kill microorganisms that are harmful to human health. Farmers' broad observation system Perhaps time-consuming, costly, and often erroneous. In tropical locations such as India, Indonesia, Pakistan, Bangladesh, and South America, Psidium guajava [1] is a popular fruit. Biological activity of guava leaf extracts (GLs) has been examined, including anticancer, antidiabetic, antioxidant, antidiarrheal, antibacterial, lipid- lowering, and hepatoprotection. So, we propose an appropriate Genetic Algorithm for automatically detecting guava leaf disease for this purpose. Rust, algal leaf spot, mummification, and canker diseases will have been identified in this research. We recommended a disease identification approach based on this dataset and calculated the experimental results, which showed high accuracy on test data.

Keywords: - Leaf Disease Detection, Guava leaves, Optimal solution, Genetic Algorithm

I. INTRODUCTION

Agriculture is India's backbone. India's economy is primarily basedon agriculture because it is a developing country. Agriculture employs more than 70% of rural households and accounts forroughly 17% of total GDP. India has a wide range of climates. From tropical in the south to temperate in the north, the climate is diverse. Because of the unpredictability, due to climatic changes, crops are lacking in nutrients and minerals. This results in Deficiency illnesses have a negative impact on crop productivity.[2]

Guava, also known as Psidium guajava [1], is a popular tropical fruit that may be found all year around in India. According to a recent report conducted by the National Institute of Nutrition, guava has the highest concentration of anti-oxidant levels, making it more nutritious and beneficial to human health. Several disordersof guava plants, on the other hand, are significant factors in reducing the amount and quality of production, as well as theeconomy. The difficult aspect is detecting and diagnosing guava leaf disease, which is difficult to do manually. Image-based automated systems are more effective than traditional systems,thanks to the widespread utilization of smartphones and digitalcameras.

In this paper we mainly focused on four diseases namely canker, dot, mummification and rust. A 'canker' is actually a sign of an injury, most commonly associated with an open wound infected with a fungal or bacterial pathogen.[3] Small black spots appear onthe leaves as the disease progresses. The fungus doesn't harm your plant, but it does prevent it from performing photosynthesis by blocking sunlight from reaching the leaves.[4] Rust diseases are fungi disease that affect plants and belong to the Pucciniales order.[5]

The optimum solution to this issue is to identify the plant's disease so that preventative measures can be implemented to protect it. Theconcept of using a genetic algorithm to identify leaf diseases in the guava leaf is implemented in this paper. Genetic algorithms are a type of evolutionary algorithm that generates answers to problems involving optimization. The population is the starting point for an algorithm. One population's solutions are chosen and then used to create a new population. Thisis done in the hopes that the new population will exceed the previous one. Solutions that are chosen to form new solutions(offspring) are chosen based on their fitness; the more suitable theyare, the more likely they are to reproduce.[6]

II. LITERATURE REVIEW

The mother of all nations is agriculture. Agriculture research aims to improve the quality and quantity of the output while reducing costs and increasing profits. Plant disease may affect the quality of the agricultural product. Pathogens, such as fungus, bacteria, and viruses, cause these diseases. As a result, detecting disease at an early stage is a difficult process. Farmers require skilled monitoring on a regular basis, which can be costly and time intensive. Many systems have been proposed to solve or at least mitigate the difficulties, depending on the applications. We've selected a few studies that deal with detecting leaf disease using various methodologies, and some of them are included below.

An automated system was presented by Khan et al. [7]. Deep learning is being used to treat a variety of fruit disorders. The details provided here is Disease segmentation is the first of two keyprocesses in the approach. Using a correlation-based technique andfruit classification diseases using pre-trained CNNs to extract deep properties, VGG16 and AlexNet are two examples of models. [2018]

Anand H. Kulkarni et al. [8] use artificial neural networks (ANN)and a variety of image processing techniques to describe a methodology for detecting plant illnesses early and accurately. Because the suggested method uses an ANN classifier for classification and a Gabor filter for feature extraction, it produces better results, with a recognition rate of up to 91%. An ANN-based classifier classifies distinct plant diseases and recognizes them using a combination of textures, colors, and characteristics. [2015]

Ormani et al. [9] developed an automated method for diagnosingapple disorders. The procedure described here is made up of three steps. The first phase entails the K-Means clustering approachwhich is used to segment lesion regions. The features of color, GLCM, and wavelet are then extracted from segmented spots that have been identified using a support vector regression (SVR). The SVR algorithm has a significant benefit that is to reduce the error rate and improve classification accuracy. [2014]

Smita Naikwadi uses [10] Histogram matching to identify plant disease. As disease develops on leaves in plants, histogram matching is performed using an edge detection technique and a color feature. The training method employs a layer’s separation technique, which includes the training of these samples, which separates the layers of an RGB image into red, green, and blue layers, as well as an edge detection technique, which detects the layered images' edges. The color cooccurrence texture analysisapproach is based on spatial Gray-level dependence matrices. [2013]

The main processes of image processing is to identify and classify disease in plants were outlined by Sachin Khirade and A.

B. Patil [11]. Image acquisition, image pre-processing, image segmentation, feature extraction, and classification are all stepsinvolved. Methods for segmentation include otsu's approach, transforming RGB images to HIS models, and k-means clustering. The k-means clustering approach is the most accurate of them all. Following that, features such as color, texture, morphology, and edges are extracted. Among these, the extraction of morphological features yields the best results. Following feature extraction, classification is carried out using Artificial Neural Network and Back Propagation Neural Network methods. [2015]

For the classification of cotton leaf disease data, Bhog and Pawar [12] used the neural network approach. K-means clustering was applied for segmentation. Red spot, white spot, yellow spot, Alternaria, and Cercospora are some of the diseases that affect cotton leaves. The MATLAB toolbox was used to conduct the experiments. The recognition accuracy of the K-Mean Clustering technique with Euclidean distance is 89.56 percent, and the execution time is 436.95 seconds. [2016]

III.PROPOSED SYSTEM

The goal of this research is to present a methodology for classifying guava leaf diseases and recommending the best

DOI:10.9756/INTJECSE/V14I7.52 ISSN: 1308-5581 Vol 14, Issue 07 2022

solutionfor dealing with them. Genetic Algorithm is successfully deployed and as well as free source software Python programming language is used. The proposed procedure depicts below in Figure 1.
There are following terminology in the process of geneticalgorithm and these are illustrated as below.

A) Population
It's a subset of all the feasible (encoded) solutions to theproblem.

B) Chromosome
A chromosome (also known as a genotype) is a set of parameters that define a possible solution to the problem that the genetic algorithm is attempting to solve.

C) Genes
An individual is characterized by a set of parameters (variables) known as **Genes**.

D) Genetic Algorithm
A genetic algorithm is a search heuristic based on Charles Darwin's natural selection theory. This algorithm

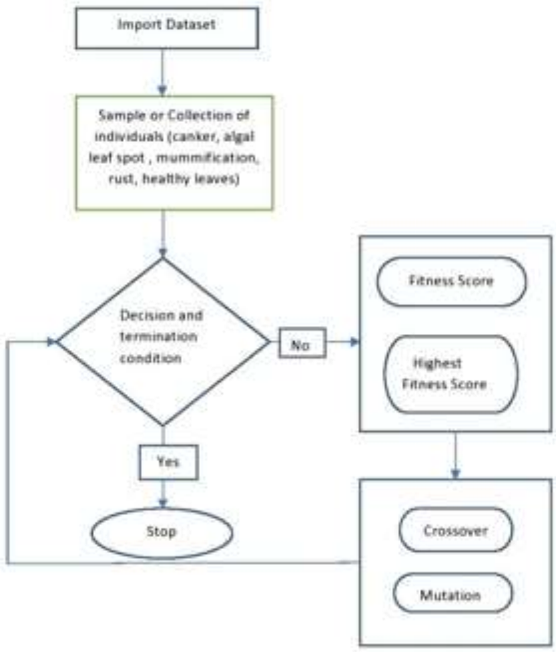

Fig. 1 process diagram of proposed method

Mimics natural selection, in which the fittest individuals are chosen for reproduction in order to create the next generation's offspring.[13]

E) Population
It's a subset of all the feasible (encoded) solutions to the problem.

F) Chromosome
A chromosome (also known as a genotype) is a set of parameters that define a possible solution to the problem that the genetic algorithm is attempting to solve.

G) Genes
An individual is characterized by a set of parameters (variables) known as **Genes**.

H) Genetic Algorithm
A genetic algorithm is a search heuristic based on Charles Darwin's natural selection theory. This algorithmmimics natural selection, in which the fittest individuals are chosen for reproduction in order to create the next generation's offspring.[13]

DOI:10.9756/INTJECSE/V14I7.52 ISSN: 1308-5581 Vol 14, Issue 07 2022

I) Fitness Score

The fitness function determines a individual level of fitness. It assigns each person a fitness score. The fitness score determines the chances of a person being chosen for reproduction.[13].

J) Crossover

Crossover is a probabilistic process in which two parent chromosomes exchange information in order to generate child chromosomes.[14]

K) Mutation

A mutation is a minor random change in the chromosome that results in a new solution. It's used to keep and add diversity to a genetic population.[14]

ALGORITHM 1:

1) Initialize populations at random p
2) Determine population fitness
3) Repeat until convergence occurs:
a) Choose parents from the overall population
b) Crossover, resulting in a new population.
c) Mutate a new population
d) Determine fitness in a new population

IV. EXPERIMENTAL PROCESS

In this section, we described experimental environments and introduced a brand-new database with detailed information including some parameters. Finally, the outcomes of the experiments are presented.

Dataset Description:

To evaluate the performance of any genetic algorithm, we'll need a suitable dataset. This dataset contains 1842 images divided into five groups, four of which depict diseased guava leaves and the remaining represent healthy leaves. The sample image for each category is shown in Figure 2. All images are collected from kaggle.com. The following is a detailed discussion of this dataset.

1) **Healthy Leaf**: We gathered 1032 healthy leaf images from distinct guava species and grouped them into two sections. For the training system, we used 773 images and for the testing system, we chose 259 images.

2) **Algal Leaf Spot**: Algal leaf spot is a widespread disease of guava leaves that affects a large number of guavaplants. We collected 143 images from species to form our database, with 106 images chosen for training andthe rest for validation.

3) **Rust**: Rust is another major guava leaf disease that reduces the plant's productivity. We collected 224images as a sample for our model, of which 167 are used for training and rest are used for testing.

4) **Canker:** 223 images are collected in which 166 are used for training and 57 are used for testing purpose.

5) **Mummification:** 220 images are used in which 165 for training and 55 for testing. However, in order to improve recognition accuracy, this dataset was divided into several phases and implemented. The overview of this dataset is shown in Table I.

TABLE 1: Overview of Dataset

Name	Total Images	Training Images	Testing Images
Healthy leaf	1032	773	259
Algal leaf spot	143	106	37
Rust	224	167	57
Canker	223	166	57
Mummification	220	165	55
Total	**1842**	**1377**	**465**

DOI:10.9756/INTJECSE/V14I7.52 ISSN: 1308-5581 Vol 14, Issue 07 2022

Figure 2: Example of guava leaf 1. Guava Healthy, 2. Algal leafspot, 3. Rust, 4. Canker, 5. Mummification

Table II: List of abbreviations andtheir description

Abbreviation	Description
GA	Genetic Algorithm
ITN	Iteration
NPN	No. of Population
CR	Crossover Rate
MR	Mutation Rate
FF	Fitness Function
CHR	Chromosomes
CF	Color Feature
GL	Guava leaf

Table III. Detail information of experimental environment

Name of the environment	Parameters
Operating System	Windows 10 Home
CPU	i5 Intel Core (TM)
GPU	NVIDIA GEFORCE GTX 1070 8GB
RAM	16GB
Hard Disk	1TB
Environment	Anaconda navigator, Tensor Flow, Keras 2.8.0
Language	Python 3.10

ALGORITHM 2:

1) Input Dataset (canker, rust, mummification, healthy,algal leaf spot)
2) Initialize No. of Iteration (ITN), ITN=600
3) Initialize No. of Population (NPN), NPN=30
4) Initialize Crossover Rate (CR) CR=0.5
5) Initialize Mutation Rate (MR) MR=0.002
6) Denote the color feature and spotted vector
7) Perform Correlation
8) Feature length measure of pixel.
9) Set Iteration=10
10) Set population=pi
11) Evaluate Pi
12) Select parent for population
13) Perform Crossover (Coff)
14) Perform Mutation (Mmut)
15) Evaluate offspring
16) Calculate population from replacement
17) Perform Fitness function (FF)
18) End

DOI:10.9756/INTJECSE/V14I7.52 ISSN: 1308-5581 Vol 14, Issue 07 2022

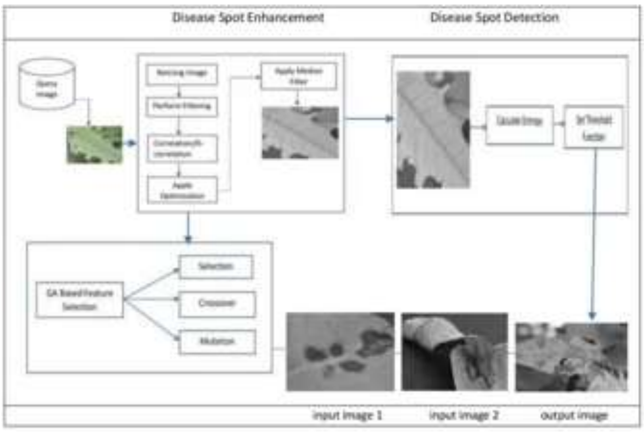

Figure 3: Flow Diagram of Proposed Model

As given FV (i) is feature vector of dimension A B, given A- the index of extracted feature and B- the length of optimized feature vector. Initialization, crossover, mutation, selection, and fitnessfunction are the five major steps of the GA in general. During the initialization phase, the chosen number of iterations (ITN = 600), the number of populations (NPN = 30), crossover rate (CR = 0.5), mutation rate (MR=0.002). We choose the uniform crossover strategy, which has a crossover rate of 0.5. This is how the crossover function is defined:

$$\varphi CR = CROSS\,(h1, h2)$$
$$h1 = \propto \times f1 + (1-\propto) \times f2$$
$$h2 = \propto \times f2 + (1-\propto) \times f1 \qquad - (1)$$

Where h1, h2 indicates the input features and f1, f2 denotes the newoffspring individuals following the crossover process. The offspring are then mutated using uniform mutation with MR =0.002 in the next stage. Crossover produces some unique individuals, which are subjected to the mutation operation. After that, offspring are evaluated, and the actual population is chosen byreplacing individuals.

V. RESULT

The average performance measures of guava leaf disease detection is shown in the below table which can be depicted as follow:

Table IV: Average performance measure of Guava leafdisease detection (Proposed method)

S.No.	Different type of Guava leaves disease	Accuracy (%)
1	Healthy	99%
2	Algal Leaf Spot	97.66%
3	Rust	96.88%
4	Canker	97.23%
5	Mummification	95%

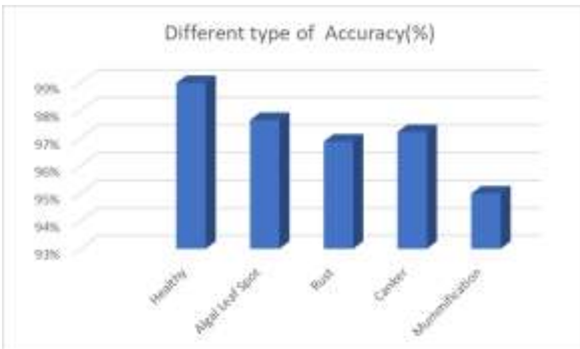

Figure 4: Average performance of guava leaf disease detection (Proposed Method)

The proposed method is compared to Deep CNN [1]. The proposedsystem's performance Over the previously mentioned, the approachprovided a high accuracy rate because genetic algorithm applies to give best optimal solution and compares them to the previously trained dataset image. The comparative metrics for the proposed approach and the other method are shown in Table V. It is determined that the proposed strategy improves the D-CNN algorithm's capacity to represent deep convolutional neural networks and effectively classify disease of guava leaves with alow error rate. Table V compares average performance measures for detecting Guava leaf disease.

Table V: Comparison average performance of guava leaf diseaseDetection

S.No.	Method	Accuracy (%)
1	Deep-Convolutional Neural Network	98.74
2	Guava leaf disease detection using Genetic algorithm (Proposed algorithm)	99

REFERENCES

1. Md. Rasel Howlader, Umme Habiba, Rahat Hossain Faisal and Md. Mostafifijur Rahman “Automatic Recognition of Guava Leaf Diseases using Deep Convolution Neural Network” 2019 International Conference on Electrical, Computer and Communication Engineering (ECCE), 7-9 February, 2019.
2. Vaishnavi Monigari, G. Khyathi Sri , T. Prathima “Plant Leaf Disease Prediction” International Journal for Research in Applied Science & Engineering Technology (IJRASET**)** ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.429 Volume 9 Issue VII July 2021.
3. P. Perumal, Kandasamy Sellamuthu , K.Vanitha, and V.K.Manavalasundaram “Guava Leaf Disease Classification UsingSupport Vector Machine” Turkish Journal of Computer and Mathematics Education Vol.12 No.7 (2021),1177-1183
4. B. Srinivas, P. Satheesh, P. Rama Santosh Naidu, and U Neelima “Prediction of Guava Plant Diseases Using Deep Learning”.
5. Monzurul Islam, Anh Dinh, Khan Wahid, and Pankaj Bhowmik. Detection of potato diseases using image segmentation and multiclass support vector machine. In Electrical and Computer Engineering (CCECE), 2017 IEEE 30th Canadian Conference on, pages 1–4. IEEE, 2017.
6. Youssef Es-saady, Ismail El Massi, Mostafa El Yassa, Driss Mammass, and Abdeslam Benazoun. Automatic recognition of plant leaves diseases based on serial combination of two svm classifiiers. In Electrical and Information Technologies (ICEIT),2016 International Conference on, pages 561–566. IEEE, 2016.
7. Khan, Muhammad Attique, Tallha Akram, MuhammadSharif, Muhammad Awais, Kashif Javed, Hashim Ali, and TanzilaSaba. "CCDF: Automatic system for segmentation and recognition of fruit crops diseases based on correlation coeffificient and deep CNN features." Computers and Electronics in Agriculture 155 (2018): 220-236.
8. Anand.H.Kulkarni, Ashwin Patil R. K.” Applying image processing technique to detect plant diseases”, International Journal of Modern Engineering Research, Vol.2, Issue.5, Sep-Oct. 2012 pp-3661-3664.
9. Omrani, E., Khoshnevisan, B., Shamshirband, S., Saboohi, H., Anuar, N. B., & Nasir, M. H. N. M. (2014). Potential of radialbasis function-based support vector regression for apple disease detection. Measurement, 55, 512-519.
10. Smita Naikwadi, Niket Amoda” ADVANCES IN IMAGE PROCESSING FOR DETECTION OF PLANT DISEASES”,
International Journal of Application or Innovation in Engineering and Management, Volume 2, Issue 11, November 2013.
11. Sachin D. Khirade, A. B. Patil, “Plant Disease DetectionUsing Image Processing,” IEEE, International Conference on Computing Communication Control and Automation, Pune, pp 768-771, 2015.
12. Bhong, Vijay S., and B. V. Pawar. "Study and Analysis of Cotton Leaf Disease Detection Using Image Processing." International Journal of Advanced Research in Science, Engineering and Technology, 2016.
13. Jenna Carr “An Introduction to Genetic Algorithms” May 16, 2014
14. K. Beulah Suganthy “Identification of Disease in Leaves using Genetic Algorithm” International Journal of Trend in Scientific Research and Development (IJTSRD) Volume: 3 | Issue: 3 | Mar-Apr 2019