International Research Journal of Innovations in Engineering and Technology (IRJIET)

ISSN (online): 2581-3048

Volume 7, Issue 12, pp 177-181, December-2023

https://doi.org/10.47001/IRJIET/2023.712025

Plant Leaf Disease Feature Extraction Centered on Intensity Permanence

[1]Dr. Shanmuga Rajathi D, [2]Dr. K Chitra

[1,2]School of Computer Studies-PG, Rathnavel Subramaniam (RVS) College of Arts and Science, Coimbatore, Tamilnadu, India

Authors E-mail: [1]shanmugarajathi@rvsgroup.com, [2]chitra.k@rvsgroup.com

Abstract **- The primary organ of the plant, the leaf, is classified automatically using image processing methods. During processing, the leaf images' features are taken out and used for categorization. Research on plants and agriculture benefits from the extraction of leaf features. The manual feature extraction process is less structured and involves measuring features like shape symmetry, color, etc. with accuracy. Advancements in image processing techniques offer a productive means of extracting features from images of leaves, and numerous innovative techniques have been put forth in the literature. These studies rely less on human intuition and instead focus on automatically extracting features from images using a variety of algorithms.**

Keywords: Feature Extraction, Leaf, Automatic Detection, Ann, FEBIP.

I. INTRODUCTION

An object loses its intensity permanence when human experts label it as a feature in the images. Since features must be extracted from the object for intensity permanence, it can be challenging to accurately link data during the manual labeling process. For instance, the largest distance that can exist between two points on a leaf—known as the Long Axis—determines the length of the leaf. Leaves with two or more long axes cannot benefit from such measurements, even though they are ideal for long, slender shape leaves. A high degree of reliability between computer data and expert experimental data is required when designing an agricultural expert system, which is used for plant taxonomy or pest identification. Undoubtedly, manually calculated feature features for plant leaves cannot satisfy such requirements. It is possible to identify the characteristics of plant leaves, providing humans with knowledge about consistently consistent habits.

In addition to providing a fresh approach to feature extraction—such as the symmetry report on shape, color, and texture—feature extraction can resolve the aforementioned issue. FEBIP stands for Feature Extraction Based on Intensity Permanence, a novel approach to feature extraction. Designing parameters for FEBIP can result in data assembly for computer automatic processing as well as human intensity perception. A FEBIP presentation of the plant leaf's color feature extraction is made in order to demonstrate FEBIP. It includes feature description and order correction. Experiments with different colored leaves have produced satisfactory results. This work proposes classifiers using artificial neural networks (ANNs) and decision trees, as well as modified GAs and feature selection based on genetic algorithms (GAs).

II. METHODOLOGY

Decision trees, ANN classifiers, modified GA, FEBIP feature extraction, and GA-based feature selection are all covered in this section.

Feature Extraction using FEBIP

In pattern recognition, feature extraction is the standard relevant step. Characteristics chosen for categorization determine which patterns can be recognized. The intensity of each image of a diseased tomato plant helps identify the pattern of the indicators. Digital evidence, comprising color intensity values of individual pixels, is implied by the digital image. Red, green, and blue—the three primary colors—combine to form the insight of color. As a result, certain colors judiciously depict a vector on the orthogonal axes system, the segments of which are contingent upon the utilized color space. Using intensity-based statistical features, it calculated the accuracy of plant disease recognition for FEBIP recognition. These include the steps shown in ALG 1 for the feature extraction algorithm. To create a feature vector that is FEBIP for grouping, use the steps included in the ALG:

1) The images have a 512 × 512 standard.
2) Used OEM segmentation to create color-segmented images from standard images.
3) Use the color space transformation CI XYZ.
4) From each picture of a healthy plant and one with disease, we computed ten statistical features in total.

For every exacted component—X, Y, and Z—it contains three mean values, three standard deviations, three skewness values, and correlated features between the X and Y components. [CI_XYZColFea] = {Mean(x), mean(y),

ISSN (online): 2581-3048
Volume 7, Issue 12, pp 177-181, December-2023
https://doi.org/10.47001/IRJIET/2023.712025

mean(z), std(x), std(g), std(b), skewness(r), skewness(g), skewness(b), corr(x,y)} is the final feature vector;

The color descriptors as (4.1 to 4.3):

$$\mu = \overline{N}\sum_{j=1}^{f} \tag{4.1}$$

$$o_{i} = J_N - \sum_{j=1}^{N}(f_{ij} - \mu_i) \tag{4.2}$$

$$S_i = I^3 \frac{1}{N}\sum_{j=1}^{N} f_{ij} - \mu_i^3 \tag{4.3}$$

Where f_{ij} is the value of the i-th color component of the image pixel j, and N refers to the number of pixels in the image.

The color correlation coefficient described as (4.4):

$$r = \frac{\sum_X \sum_Y (X_{XY} - \overline{X})(F_{XY} - \overline{F})}{J(\sum_X \sum_Y (X_{XY} - \overline{X})^2(\sum_X \sum_Y F_{XY} - \overline{F})^2} \tag{4.4}$$

As well as where The pairs X and Y represent different correlation vectors that were taken out of the color feature space. Prior to submitting the extracted feature to the final classification, feature normalization is applied to it. Normalization of the image is achieved by adjusting the range of pixel intensity values and applying feature vectors that are taken from plant images that are either disease-free or contain no infection. Improved feature vector quality for effective classification is the goal of each extracted feature vector module. An n-dimensional feature vector called x was extracted from the provided plant images that were either disease-infected or not. Feature normalization was then applied, using Zero- mean and Unit- Variance Normalization (MV).

Methodology 1. Plant disease of tomatoes Highlights Extraction
Data Input: Image dataset for tomato plant disease
Outcome: Feature Vector 1. Image standardization using a 512 × 512 fixed size
2. To create segmented color images, apply OEM segmentation.
3. Converting color segmented images to the CI XYZ color space
4. Feature Extraction: For every picture contaminated with a disease ten statistical features were computed.
Means (X, Y, and Z) for every component that has been eliminated
The standard deviation of every element
//std(X), std(Y), std(Z)
Skewness for each component
//skewness(X), skewness(Y), skewness(Z)
Correlation between two component
// corr2(X,Y)
5.

$$CI\_XYZ\ ColFea = [M_X\ M_Y\ M_Z\ Sd_X\ Sd_Y\ Sd_Z\ S_X\ S_Y\ S_Z\ XY\,|\,Corr]$$

// Feature Vector
Normalized feature vector $X = [CI\_XYZ\ ColFea]$

MV scales all the component xi (i=1,2,….n), i.e., for each X, Y and Z components of CI XYZ color space of x by the following expression to create the normalized feature vector X in (4.5):

$$X_i - X^0_{i=\mu}, i = 1, 2, \ldots n \tag{4.5}$$

Where $X = [CI\_XYZ\ ColFea]$, µ signifies the mean value of the feature vector and o represent it's standard deviation. A random variable with a mean value of zero and a variance of one is created by modifying the feature vector x using the MV technique. The normalized feature vector was eventually formed by it.

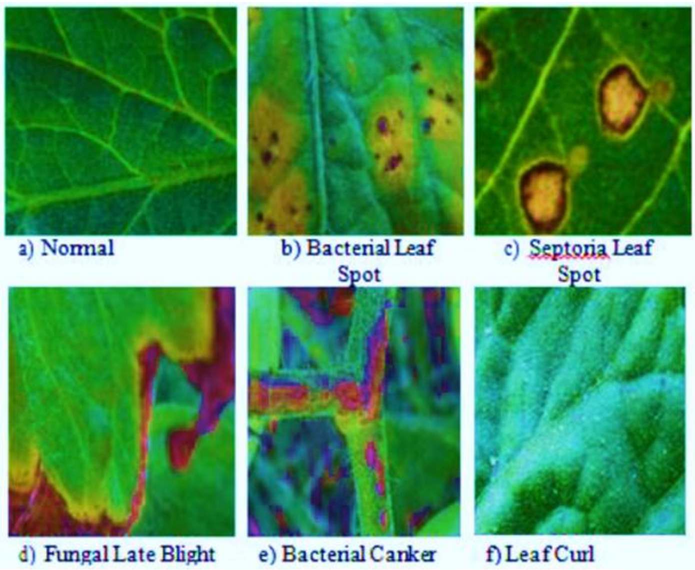

Figure 1: a) Non-affected (Normal) (b) Bacterial Leaf Spot (c) Septoria Leaf Spot (d) Fungal Late Blight (e) Bacterial Canker and (f) Leaf Curl

ISSN (online): 2581-3048
Volume 7, Issue 12, pp 177-181, December-2023
https://doi.org/10.47001/IRJIET/2023.712025

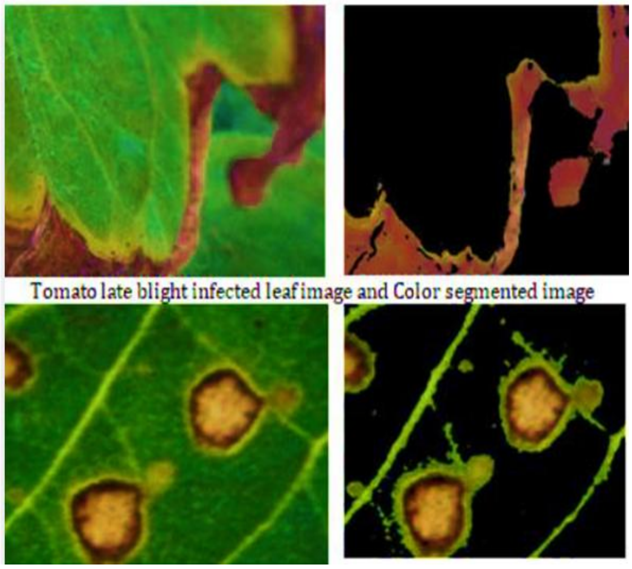

Figure 2: Tomato infected Leaf image and Color segmented image

Utilizing FEBIP for Recognition: The FEBIP is predicated on the notion of a non-parametric supervised learning procedure. This demonstrated how a binary decision tree is made. The greedy technique is the foundation of the supervised approach. There are subsets within the FEBIP that were produced by multiple classes, and these subsets employ the divide and conquer strategy. The training set's tuples and their associated class labels are the first elements in the classification tree, which follows a top-down methodology. The classification tree is generated by recursively splitting the tree into smaller subsets. The Gini index in (4.6) serves as the splitting standard for creating a classification tree.

$$GiniIndex: 1-\sum_{i=1}^{N} P^2 \qquad (4.6)$$

Based on ten statistical features, the root node's Gini index is n = -0.73. Eighty nodes were made in total for classification. An illustration of an if-then rule applied in the suggested method is as follows:

1. If X1<-0.73 then node

2. else if X1>=-0.73 then node

3. else Bacterial Leaf Spot End

    2 if X4<0.71 then node

4. else if X4>=0.71 then node

5. else LeafCurl End

The if-then rules as declared above made for each image by using features extracted from images.

The first node of the tree made using Gini index criteria. For instance, if the feature vector X for image i1 operates a value less than -0.73, it will sub-branched with another vector. If the value neither has the value less than or greater than -0.73, then the image goes to the Bacterial Leaf Spot class.

Table 1 represents the selected features with values measured in bits per pixel.

Table 1: Selected features with values

Features	Values (bpp)
Contrast	1.1
Correlation	0.7
Energy	0.4
Entropy	4.0
Homogeneity	0.8
Kurtosis	3.7
Mean	41.5
Skewness	1.3
Smoothness	1
RMS	9.5
Variance	3234.2

Genetic Algorithm (GA) Based Feature selection

The central claim of feature selection is that the accuracy of classification can occasionally be reduced by multiple redundant features. Optimal feature selection improves accuracy while lowering computation complexity. A fitness function is employed in Genetic Algorithm (GA) based feature selection to evaluate the power of the feature subset. With GA feature selection, it is possible to select discriminative features from the feature vector. A feature selection can be defined as an operator fs that applies the mapping in (4.7) to map from the m-dimensional (input) space to the n-dimensional (output) space.

Modified GA (Stochastic diffusion search (SDS)-GA)

One probabilistic technique for resolving pattern recognition and matching issues is stochastic diffusion search, or SDS. SDS is a distributed computation model that uses a multi-agent population-based global search and optimization algorithm in conjunction with simple agent interaction. Unlike other search algorithms that draw inspiration from nature, SDS has a strong mathematical basis. The minimal convergence criteria, convergence to the global optimum, and resource allocation of the algorithm are used to characterize its behavior. With linear time complexity, the SDS algorithm is a reliable choice.

International Research Journal of Innovations in Engineering and Technology (IRJIET)

ISSN (online): 2581-3048

Volume 7, Issue 12, pp 177-181, December-2023

https://doi.org/10.47001/IRJIET/2023.712025

III. RESULTS AND DISCUSSION

The evaluation of the GA FS-ANN, GA FS-decision tree, FEBIP -ANN, and FEBIP -ANN approaches is presented in this section. As stated in the previous chapter, the same data is used for the investigation. The data in tables 2 through 5 and figures 3 through 6 represent the accuracy, precision, recall, and f measure.

Table 2: Accuracy for GA FS-ANN

	Accuracy
FEBIP - Decision Tree	78.24
FEBIP- ANN	80.2
GA FS-Decision Tree	80.1
GA FS - ANN	82.16

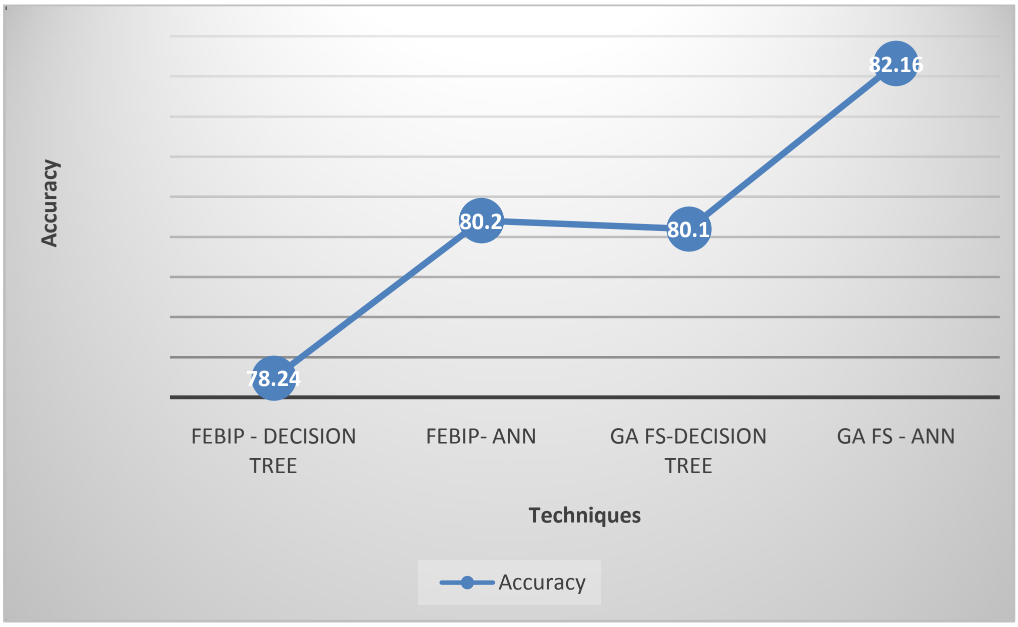

Figure 3: Accuracy for GA FS-ANN

Figure 3 shows that, successively, the GA FS - ANN outperforms the FEBIP - ANN by 4.89%, the FEBIP - ANN by 2.41%, and the GA FS - decision tree by 2.54%.

IV. CONCLUSION

An important step in pattern recognition systems is feature extraction. The usefulness of statistical features for recognition has been evaluated in this work. The FEBIP classifier uses these eliminated features to determine the accuracy of recognition. The outcomes have verified that sufficient credit correctness is achieved by combining the extracted correlated features X and Y with the remaining nine features. The act of the method is dependent on the feature extraction process as well as the various stages of the recognition system, ranging from pre-processing to classification. In order to select features from the input data and arrive at the best classification, the searching capacity of GAs is tested. SDS is a reliable global metaheuristic for swarm intelligence that can solve search and optimization issues quickly. ANNs are a quicker and less time-consuming method of classifying data because they have already received training as feature extractors. Future research in the same field can benefit from the grouping of a few statistical and geometric features with additional classifiers. The GA FS - ANN outperforms the FEBIP - ANN by 4.89%, the GA FS - ANN by 2.41%, and the FEBIP - ANN by 2.54%, according to the results.

REFERENCES

[1] Kumar, S., Sharma, B., Sharma, V. K., Sharma, H., & Bansal, J. C. (2018). Plant leaf disease identification using exponential spider monkey optimization. Sustainable computing: Informatics and systems.

[2] Kumari, C. U., Prasad, S. J., &Mounika, G. (2019, March). Leaf disease detection: Feature extraction with k-means clustering and classification with ann. In 2019 3rd International Conference on Computing Methodologies and Communication (ICCMC) (pp. 1095-1098). IEEE.

[3] Kumari, N., & Sunita, S. (2013). Comparison of ANNs, Fuzzy Logic and Neuro-Fuzzy Integrated Approach for Diagnosis of Coronary Heart Disease: A Survey. International Journal of Computer Science and Mobile Computing, IJCSMC, 2(6), 216-224.

ISSN (online): 2581-3048

Volume 7, Issue 12, pp 177-181, December-2023

https://doi.org/10.47001/IRJIET/2023.712025

[4] Liang, Q., Xiang, S., Hu, Y., Coppola, G., Zhang, D., & Sun, W. (2019). PD2SE-Net: Computer-assisted plant disease diagnosis and severity estimation network. Computers and electronics in agriculture, 157, 518-529.

[5] Liu, B., Ding, Z., Tian, L., He, D., Li, S., & Wang, H. (2020). Grape leaf disease identification using improved deep convolutional neural networks. Frontiers in Plant Science, 11, 1082.

[6] Lv, M., Zhou, G., He, M., Chen, A., Zhang, W., & Hu, Y. (2020). Maize Leaf Disease Identification Based on Feature Enhancement and DMS-Robust Alexnet. IEEE Access, 8, 57952-57966.

[7] Mahlein, A. K. (2016). Plant disease detection by imaging sensors–parallels and specific demands for precision agriculture and plant phenotyping. Plant disease, 100(2), 241-251.

[8] Mukhopadhyay, S., Paul, M., Pal, R., & De, D. (2021). Tea leaf disease detection using multi-objective image segmentation. Multimedia Tools and Applications, 80(1), 753-771.

[9] Muthukannan, K., &Latha, P. (2018). A GA_FFNN algorithm applied for classification in diseased plant leaf system. Multimedia Tools and Applications, 77(18), 24387-24403.

[10] Nowakowski, G., Dorogyy, Y., & Doroga-Ivaniuk, O. (2018). Neural network structure optimization algorithm. Journal of Automation Mobile Robotics and Intelligent Systems, 12.

[11] Oo, Y. M., &Htun, N. C. (2018). Plant leaf disease detection and classification using image processing. International Journal of Research and Engineering, 5(9), 516-523.

[12] Owotogbe, J. S., Ibiyemi, T. S., &Adu, B. A. A Comprehensive Review on various types of noises in Image Processing. and International Journal of Scientific & Engineering Research" ISSN, 2229-5518. International Journal of Scientific & Engineering Research Volume 10, Issue 11, November-2019.

[13] Pantazi, X. E., Moshou, D., & Tamouridou, A. A. (2019). Automated leaf disease detection in different crop species through image features analysis and One Class Classifiers. Computers and electronics in agriculture, 156, 96-104.

[14] Patel, A., & Joshi, B. (2017). A survey on the plant leaf disease detection techniques. International Journal of Advanced Research in Computer and Communication Engineering ISO 3297: 2007 Certified, 6(1).

[15] Pater, L. (2016). Application of artificial neural networks and genetic algorithms for crude fractional distillation process modeling. arXiv preprint arXiv:1605.00097.

[16] Patidar, P., Gupta, M., Srivastava, S., & Nagawat, A. K. (2010). Image de-noising by various filters for different noise. International journal of computer applications, 9(4), 45-50.

Citation of this Article:

Dr. Shanmuga Rajathi D, Dr. K Chitra, "Plant Leaf Disease Feature Extraction Centered on Intensity Permanence" Published in *International Research Journal of Innovations in Engineering and Technology - IRJIET*, Volume 7, Issue 12, pp 177-181, December 2023. Article DOI https://doi.org/10.47001/IRJIET/2023.712025

****