# An Analytical Research on Artificial Intelligence based Plant Disease Detection

**Ramesh Kumar**[1]
[1]Assistant Professor, Department of Computer Science & Engineering
B. P. Mandal College of Engineering, Madhepura, Bihar, India
**Rakesh Kumar Roshan**[2]
[2]Assistant Professor, Department of Computer Science & Engineering
RRSDCE, Begusarai, Bihar, India

**Abstract-**
One of the biggest reasons for financial losses in the agriculture sector is plant disease, which is described as an abnormal condition that interferes with a plant's natural growth. Increasing agricultural crop output requires early detection of plant diseases. In this study, a novel robust hybrid classification model that permits real-time disease classification in tomato, grape, and apple plants has been created. It is based on swarm optimization-supported feature selection and includes machine learning and deep learning methods. This will allow for the early diagnosis of the plant illness and the application of the proper remedy. This study presents the development of a new hybrid plant leaf disease classification model with low computational complexity and high accuracy. The model consists of a convolutional neural network (CNN) classifier and the wrapper approach, which includes the flower pollination algorithm (FPA) and support vector machine (SVM) and a wrapper-based feature selection approach using metaheuristic optimisation techniques. Using wavelet families including biorthogonal, Coiflets, Daubechies, Fejer–Korovkin, and symlets, a two-dimensional discrete wavelet transform (2D-DWT) was used to extract the characteristics from the picture dataset, which included apple, grape, and tomato plants. 10684

***Keywords***— *Detection of Plant Diseases, Machine Learning, Deep Learning Methods, Swarm Optimization-Supported Feature.*
**DOI Number: 10.48047/nq.2022.20.8.nq221094** **NeuroQuantology 2022; 20(8): 10684-10692**

## INTRODUCTION

Reducing the incidence of plant diseases, which are a major global source of hunger and food insecurity, and preventing them early on are critical to raising agricultural production. Determining the type of plant disease and detecting it early are crucial steps in choosing the right course of therapy. Algorithms that can automatically categorise plant disease with high accuracy and low computing cost have to be developed since visual examination, which is commonly used to determine plant disease and kinds, is laborious and prone to human mistake. The current number of the world's population and its rapid growth need optimal use of the limited amount of agricultural areas as well as efficient production of agricultural goods. Early plant disease identification is essential for the effective use of fertiliser, spraying, and weed detection techniques in the production of agricultural goods. Plant diseases are known to cause yield losses in agriculture, with losses ranging from 20% to 40%. There will be fewer consumers of that commodity on the market as a result of this significant yield loss. When plant illnesses caused by viruses, pathogens, or plague go undiagnosed, additional pesticide dosages must be applied to the affected plant, which lowers agricultural yield. Detecting and correctly classifying plant diseases at an early stage not only contributes to improving the quality of agricultural products but also allows for the reduction of undesirable chemical spray applications such as fungicides and herbicides. In this study, we propose a robust hybrid plant disease classification model based on metaheuristic optimization-assisted feature selection that includes machine learning and deep learning algorithms for real-time early detection of diseases in apple, grape, and tomato plants. The performance of the model embedded in the NVIDIA Jetson Nano developer kit on an unmanned aerial vehicle (UAV) has been experimentally tested on apple, grape, and tomato plants. Plant disease, which is one of the main causes of economic losses in the agricultural industry in the world, is defined as an abnormal state that disrupts the normal growth of the plant. Symptoms related to the abnormal state can usually be seen in the leaf, stem, and root parts of the plant. Leaf images are a good source of information in the classification of plant diseases. Therefore, researchers focus on the classification of plant diseases through leaf images. In this study, plant diseases are classified using leaf images of apple, grape, and tomato plants. Although traditional methods based on visual inspection are used in the detection of plant diseases in small-scale agricultural lands, it is very difficult to apply these methods in large-scale agricultural lands as they require tiring and continuous monitoring. Especially in diagnosing both the species and the disease of plants with similar leaves, this process becomes more difficult and leads to visual errors. To cope with these challenges, various image processing and artificial intelligence techniques such as machine learning and deep learning are used in the real-time classification of plant diseases based on plant leaf images. In the literature, there are image processing-based studies for real-time prediction of plant diseases and disease levels on many plant species. In the severity level of white-tip disease (Aphelenchoides besseyi Christie), commonly seen in paddy crops, is estimated based on image processing techniques. The diseased area in the paddy crop is determined with the help of a color image segmentation algorithm based on chromatic aberration. The severity of the disease is estimated by proportioning the area of the determined region to the entire area of the leaf. Based on the estimation results obtained, a variable rate chemical spray system has been developed for the precise application of agrochemicals in real time. In, a ΔE (Delta E) segmentation-based feature extraction algorithm is proposed to classify diseases in images of citrus plants. The diseased region is identified by the ΔE method and hue saturation value (HSV), local binary patterns (LBP), and red green blue (RGB) histogram models are extracted from this region. The features obtained with the HSV, LBP, and RGB descriptors are combined, and a hybrid feature set is created. Principle component analysis (PCA) is used to reduce dimensionality. Fine k-nearest neighbor (kNN), cubic support vector machine (SVM), boosted tree, and bagged tree ensemble classifiers are tested. It is stated that the best performance is achieved with the bagged tree ensemble classifier. In, the temperature difference between the healthy and infected parts of the cucumber leaf is measured and downy mildew disease is detected using Fourier transform infrared (FTIR) spectroscopy. It is reported that the FTIR technique is effective in the pre-symptomatic detection of downy mildew disease on cucumber leaves. In, feature extraction and segmentation algorithms are used together for detection and classification of early scorch disease in banana leaves, fungal disease in bean leaves, sunburn disease in lemon leaves, and bacterial disease in roses. Segmentation of diseased regions has been carried out using a genetic algorithm (GA). It is also emphasized that the optimization algorithm used reduces the computational complexity. In the image processing approach, the classification accuracy of plant diseases directly depends on the performance of feature extraction and segmentation.

## STRUCTURE OF THE ALGORITHM FOR DETECTING PLANT DISEASES

It presents the framework of our model, which is capable of accurately and real-time classifying tomato, grape, and apple plant leaf diseases. The biorthogonal, Coiflet, Daubechies, Fejer–Korovkin, and symlet wavelet families are used to the picture dataset of tomato, apple, and grape plants in our suggested model to extract features for each family. Using the wrapper technique, which consists of the FPA and SVM algorithms, the features that offer the best model performance and the least amount of model complexity are chosen. Plant leaf diseases are classified by the CNN classifier using specific characteristics, which helps it solve the model hyperparameter problem. The next subsections include a detailed description of every approach we employed in our model.

**(i) Discrete Wavelet Transform**- The Fourier transform (FT), which has been successfully applied for stationary signals, cannot be used for non-stationary signals. Since the spectrum changes with time, FT is insufficient to reveal the correct spectrum for non-stationary signals. Non-stationary signals are divided into sufficiently small pieces by short-term Fourier transform (STFT), and these small pieces are considered stationary. A time-frequency representation is needed to know what frequency components are present at different times and how they change as time passes. The sinusoidal frequency and phase content of a time-varying signal are found by STFT, and the relationship between frequency and time variation is defined with the help of a moving window. However, this technique is quite complicated to study non-stationary signals. Wavelet transform (WT) offers an alternative approach to overcome such challenges of FT and STFT with its variable-size windowing techniques.

**(ii) Multiresolution Analysis**- Not only the time domain features of the signals, but also the frequency domain features can be extracted and classified. DWT-based multi-resolution analysis is very useful in feature extraction applications from image signals. Undesirable components, such as noise and trend in the signal, can be separated by multiresolution analysis (MRA).

**(iii) Wavelet Families**- The biorthogonal, Coiflet, Daubechies, Fejer–Korovkin, and symlet wavelet families used in the calculation of high and low pass filter coefficients of the DWT are introduced in the following subsections.

**(iv) Flower Pollination Algorithm**- The FPA introduced by Xin-She Yang in 2012 is a new metaheuristic inspired by the reproductive process of flowering plants to achieve the best result in the shortest time in solving global optimization problems. Briefly, the algorithm mimics pollination in plants, including all the mechanisms for pollinators such as pollen transfer insects, bats, birds, wind, and water. The main purpose of flower pollination is to provide optimal vitality and the optimum biological reproduction phase. Pollination and other factors interact best to reproduce plants. Flowers need pollinators to carry out their pollination guidelines. There are two types of pollinators. The first of these are biotic pollinators, such as bats, flies, and bees. These pollinators can carry the pollen of flowers over great distances. In addition, the flight paths of these pollinators can be modeled with the Lévy distribution. The second type of pollinator is abiotic pollinators, such as water and wind.

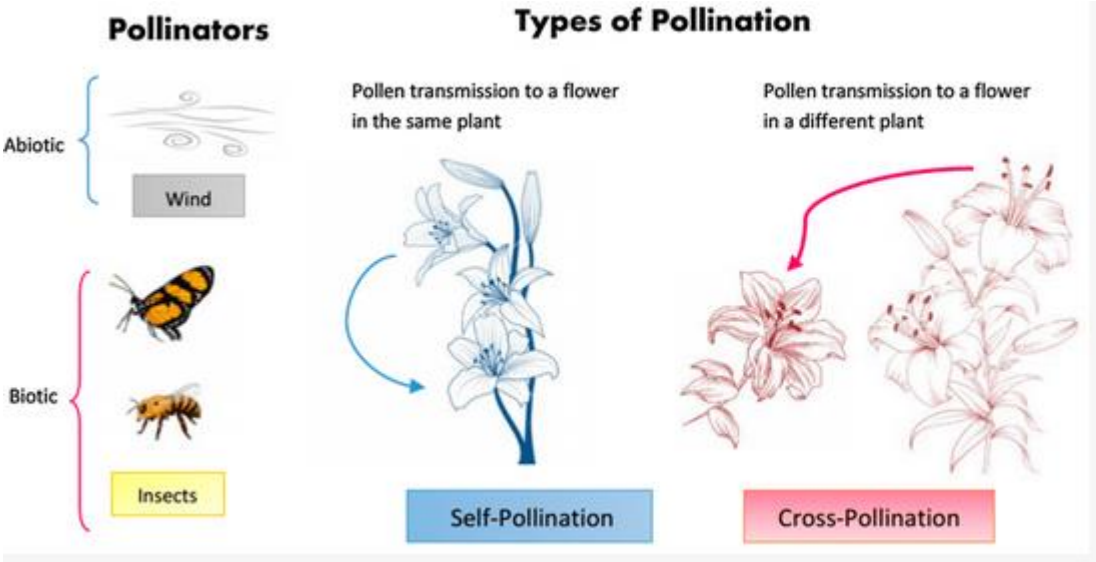


**Figure 1- Pollination types and pollinators in the FPA**



## FRAMEWORK OF THE PROPOSED METHODOLOGY

In this study, a new robust hybrid model whose features are extracted with 2D-DWT and selected with the FPA-SVM wrapper approach and include a CNN classifier is proposed to classify plant leaf diseases in real time with high accuracy, low computational cost, and low parameter complexity. The framework of the proposed approach for real-time classification of plant leaf diseases is presented in Figure 2. The phases of the proposed approach are briefly summarized below: In the data preparation phase, the image dataset consisting of apple, grape, and tomato plant diseases is randomly divided into two independent datasets, 80% and 20%, respectively, for the training and validation phases.

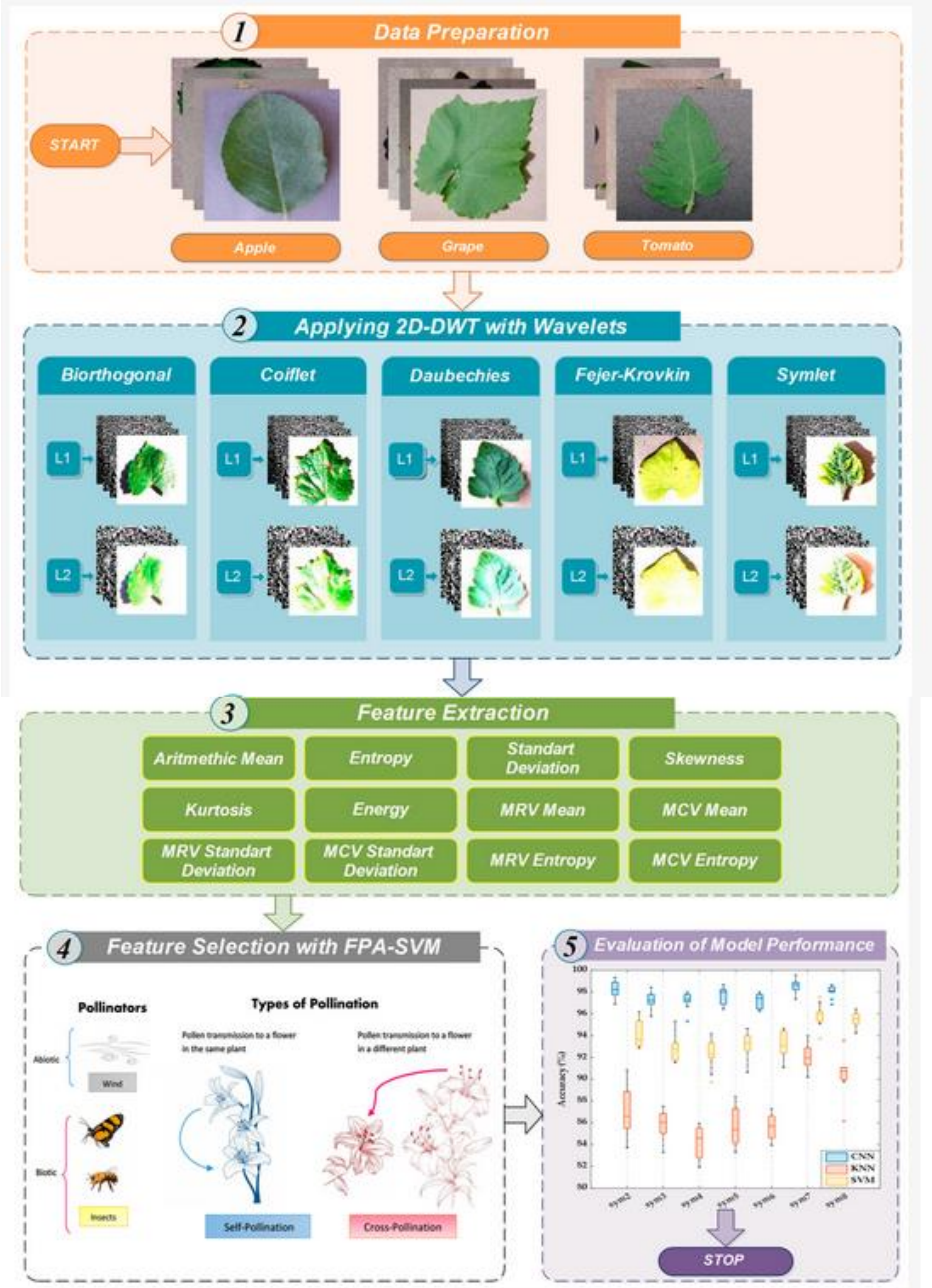




**Figure 2- Framework of the proposed robust hybrid model to classify plant leaf diseases in real time**

At the phase of applying 2D-DWT with wavelets, the distinguishing features defining the characteristic features of plant leaf diseases are extracted with the 2D-DWT using wavelet families such as biorthogonal, Coiflets, Daubechies, Fejer–

Korovkin, and symlets. In the feature extraction phase, energy and statistical-based features are extracted from the vertical, horizontal, diagonal, and approximate matrices of 2D-DWT. Six features, namely the arithmetic mean, entropy, standard deviation, skewness, kurtosis, and energy, are applied to these four matrices. Moreover, the column vector, which is the maximum value of the columns of any matrices, is expressed with MCV, and the row vector, which is the maximum value of the rows of them, is expressed with MRV. Additionally, six properties, namely arithmetic mean, standard deviation, and entropy of both MCV and MRV, were applied to the four matrices. The same process is also repeated for the second level of decomposition in 2D-DWT. At the end of this phase, a total of 96 features are extracted. At the feature selection with the FPA-SVM phase, the most suitable ones among the normalized features for each wavelet family are selected with the help of the wrapper approach consisting of the FPA and SVM algorithms. In addition, the fitness function, which takes into account both the number of features used in the model and the model performance, is determined in order to keep the model complexity and computation cost to a minimum level. In the evaluation of model performance phase, CNN, SVM, and KNN classifier performances are measured with the help of performance metrics and the model with the highest performance is determined.

**RESULTS AND DISCUSSION OF THE EXPERIMENT**

The studies showing the classification efficiency of the proposed hybrid model for real-time classification of plant leaf diseases on a dataset consisting of leaf images of healthy and diseased apple, grape, and tomato plants are presented. The proposed model includes the 2D-DWT signal processing method based on the “sym7” wavelet family, the wrapper approach consisting of FPA and SVM, and the CNN classifier. The efficiency of the proposed optimization algorithm is also compared with the PSO algorithm. The performance of the proposed CNN classifier for the hybrid model is compared with the performances of the SVM and KNN classification algorithms, and its effect on the performance of the plant leaf disease classification model is examined.

**(i) Dataset**- In this study, a data set consisting of leaf images of apple, grape, and tomato plants with a size of 256×256 pixels was used. The image dataset consisting of apple, grape, and tomato plant diseases used in the study was randomly divided into two independent datasets, 80% and 20%, respectively, for the training and validation phases. The dataset consisting of apple plant leaf images consists of healthy, black rot, cedar rust, and scab disease classes seen in Figure 3. In the training phase of the models in the study, a total of 1100 apple plant leaf images, 275 images for each class, were used, and in the validation phase, a total of 220 apple plant leaf images, 55 images for each class, were used.




**Figure 3- Leaf image examples belonging to healthy and diseased classes of apple plants**

**(ii) Extraction of Statistical and Entropy-Based Features-** Plant leaf disease is diagnosed by tests such as enzyme-linked immunosorbent assay (ELISA) and polymerase chain reaction (PCR) performed in the laboratory environment, as well as visual inspections made by experienced individuals, be they a botanist or farmer. While this is the right approach, it is a costly and highly labor-intensive process as it requires the installation of laboratory equipment. However, these traditional methods based on experience and laboratory testing are not suitable for real-time detection of plant leaf disease as they are time-consuming and allow expert error under heavy workload. Since microscopic evaluation and diagnostic experiments such as ELISA and PCR do not allow real-time detection of plant leaf disease, the distinguishing features of diseases have been extracted. Extracted features define the characteristic structure of plant leaf diseases, unlike the features obtained by visual image analysis techniques.



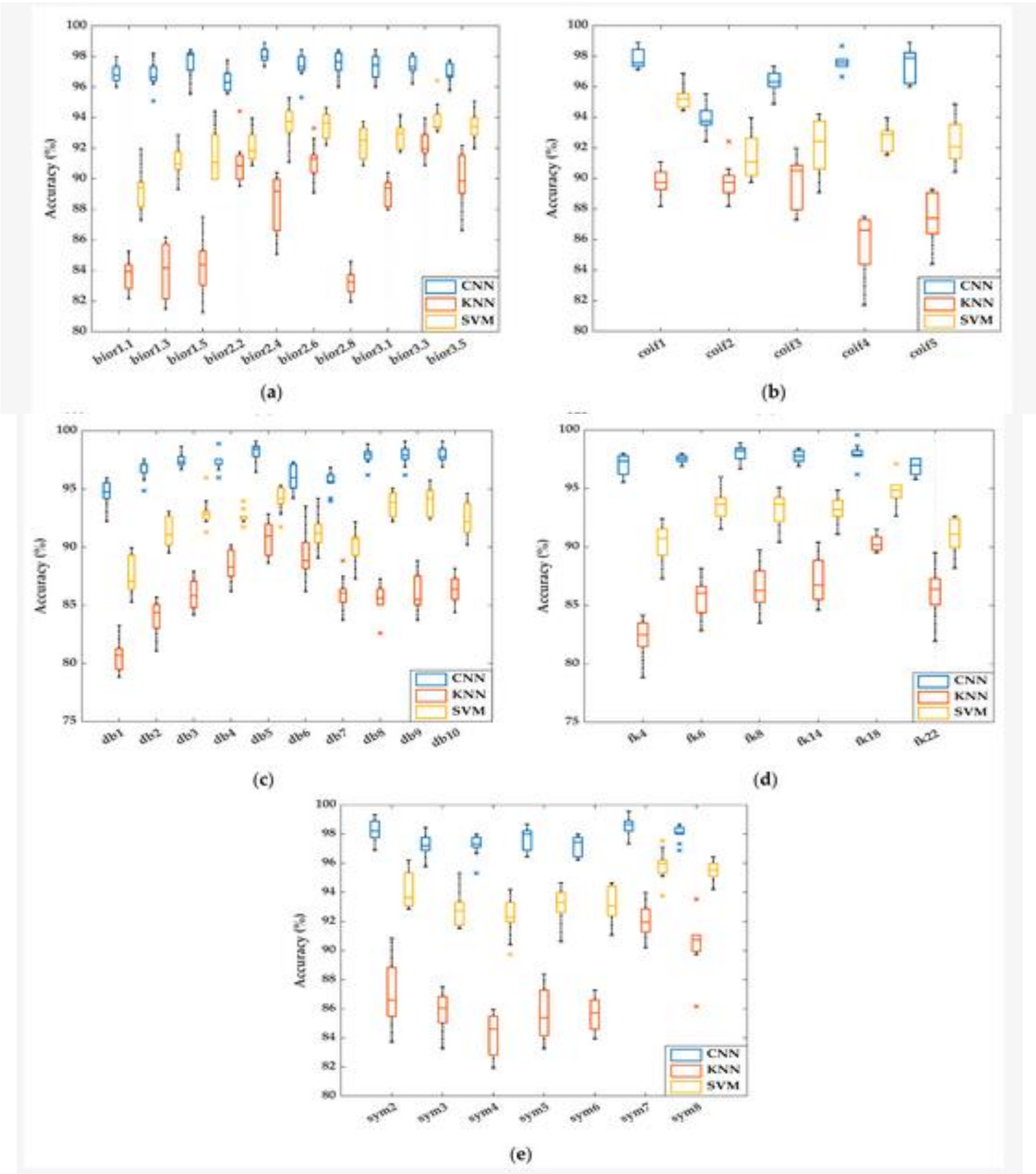

**Figure 4- Performances of classifiers used in the study for the (a) biorthogonal, (b) Coiflet, (c) Daubechies, (d) Fejer–Korovkin, and (e) symlet wavelet families**

**(iii) Applying 2D-DWT with Wavelet Families**- The features of plant leaf diseases were extracted by 2D-DWT using the biorthogonal, Coiflet, Daubechies, Fejer–Korovkin, and symlet wavelet families for filter lengths shown. A two-level decomposition was applied to the original image matrix of 256×256 pixels. As a result of this decomposition, vertical, horizontal, diagonal, and approximation image matrices were obtained for each wavelet family. In the first level of decomposition, four image matrices with 130×130 pixel size and four image matrices with 67×67 pixel size in the second level of decomposition, a total of eight image matrices were obtained.

**(iv) Evaluation of Plant Leaf Disease Classification Models and Discussion**- All models mentioned in the study of plant leaf disease images were performed on a personal computer with an Intel Core i7–10875H processor, an 8 GB NVIDIA RTX 3070 graphics card, and 16 GB of RAM. All codes for the models were compiled by MATLAB 2021b. The models created in the study were tested on 896 pieces of data, including 12 plant leaf disease classes, taken with the camera on the UAV. All models created were run 50 times, and the performance of the models was computed as the mean and standard deviation. In the study, the effects of both optimization and classifier algorithms on model performance were examined. The performances of the CNN, KNN, and SVM classification models created with the features selected by both the FPA-SVM and PSO-SVM wrapper approaches were measured in terms of accuracy metrics

**CONCLUSION**

Early detection of plant diseases is crucial for the proper course of treatment. Detecting and classifying plant leaf diseases automatically, as opposed to visually, increases the yield of agricultural products. This paper proposes a novel robust hybrid classification model to accurately and real-time diagnose illnesses in tomato, grape, and apple plants using swarm optimization-supported feature selection, machine learning, and deep learning methods. The CNN classifier, the wrapper technique using FPA-SVM, and the 2D-DWT signal processing method based on the "sym7" wavelet family are all included in the suggested model. Features collected from several wavelet families (e.g., biorthogonal, Coiflets, Daubechies, Fejer–Korovkin, and symlets) are utilised with the 2D-DWT approach in order to make the model robust. In order to minimise the computing complexity of the model, a wrapper method utilising FPA and SVM is used to choose the features that maintain excellent classifier performance for each wavelet family. The fewest features possible to maintain good classification performance are used in the creation of the CNN classification model. The CNN classifier, which has only one classification layer and no feature extraction layer, is used to reduce model complexity and solve the hyperparameter issue. We evaluate the real-time performance of our proposed plant leaf disease classification model incorporated in the NVIDIA Jetson Nano AI development kit on the UAV.



**REFERENCES**

[1] F. Martinelli et al., “Advanced methods of plant disease detection. A review,” Agron. Sustain. Dev., vol. 35, no. 1, pp. 1–25, 2015, doi:10.1007/s13593-014-0246-1.

[2] S. Sladojevic, M. Arsenovic, A. Anderla, D. Culibrk, and D. Stefanovic, “Deep Neural Networks Based Recognition of Plant Diseases by Leaf Image Classification,” Comput. Intell. Neurosci., vol. 2016, 2016, doi: 10.1155/2016/3289801.

[3] S. R. Maniyath et al., “Plant disease detection using machine learning,” Proc. - 2018 Int. Conf. Des. Innov. 3Cs Comput. Commun. Control. ICDI3C 2018, pp. 41–45, 2018, doi: 10.1109/ICDI3C.2018.00017.

[4] J. G. A. Barbedo, “Factors influencing the use of deep learning for plant disease recognition,” Biosyst. Eng., vol. 172, pp. 84–91, 2018,doi: 10.1016/j.biosystemseng.2018.05.013.

[5] A. Lowe, N. Harrison, and A. P. French, “Hyperspectral image analysis techniques for the detection and classification of the early onset of

plant disease and stress," Plant Methods, vol. 13, no. 1, pp. 1–12, 2017, doi: 10.1186/s13007-017-0233-z.

[6] A.-K. Mahlein, "Present and Future Trends in Plant Disease Detection," Plant Dis., vol. 100, no. 2, pp. 1–11, 2016, doi: 10.1007/s13398-014-0173-7.2.

[7] D. I. Patrício and R. Rieder, "Computer vision and artificial intelligence in precision agriculture for grain crops: A systematic review," Comput. Electron. Agric., vol. 153, no. June, pp. 69–81, 2018, doi: 10.1016/j.compag.2018.08.001.

[8] K. Golhani, S. K. Balasundram, G. Vadamalai, and B. Pradhan, "A review of neural networks in plant disease detection using hyperspectral data," Inf. Process. Agric., vol. 5, no. 3, pp. 354–371, 2018, doi: 10.1016/j.inpa.2018.05.002.

[9] J. G. A. Barbedo, "A review on the main challenges in automatic plant disease identification based on visible range images," Biosyst. Eng., vol. 144, pp. 52–60, 2016, doi: 10.1016/j.biosystemseng.2016.01.017.

[10] K. G. Liakos, P. Busato, D. Moshou, S. Pearson, and D. Bochtis, "Machine learning in agriculture: A review," Sensors (Switzerland), vol.18, no. 8, pp. 1–29, 2018, doi: 10.3390/s18082674.

[11] S. P. Mohanty, D. P. Hughes, and M. Salathé, "Using deep learning for image-based plant disease detection," Frontiers in Plant Science, vol. 7, p. 1419, 2016.

[12] F. Martinelli, R. Scalenghe, S. Davino et al., "Advanced methods of plant disease detection. A review," Agronomy for Sustainable Development, vol. 35, no. 1, pp. 1–25, 2015.

[13] Ankit Narendrakumar Soni (2018). Data Center Monitoring using an Improved Faster Regional Convolutional Neural Network. International Journal of Advanced Research in Electrical, Electronics and Instrumentation Engineering, 7(4), 1849-1853. doi:10.15662/IJAREEIE.2018.0704058

[14] J. Amara, B. Bouaziz, and A. Algergawy, "A deep learning-based approach for banana leaf diseases classification," in Proceedings of the Datenbanksystemefür Business, Technologie und Web (BTW '17) - Workshopband, 2017.

[15] Karunakar pothuganti (2013) „An Efficient Architecture for Lifting Based 3D-Discrete Wavelet Transform‟ International Journal of Engineering Research & Technology (IJERT), Vol. 2 Issue 12, December – 2013 ISSN: 2278-018

[16] Ankit Narendrakumar Soni (2018). Feature Extraction Methods for Time Series Functions using Machine Learning. International Journal of Innovative Research in Science, Engineering and Technology, 7(8),8661-8665. doi:10.15680/IJIRSET.2018.0708062

[17] A. Ramcharan, K. Baranowski, P. McCloskey, B. Ahmed, J. Legg, and D. P. Hughes, "Deep learning for image-based cassava disease detection," Frontiers in Plant Science, vol. 8, p. 1852, 2017.

[18] E. Fujita, Y. Kawasaki, H. Uga, S. Kagiwada, and H. Iyatomi, "Basic investigation on a robust and practical plant diagnostic system," in Proceedings of 2016 15th IEEE International Conference on Machine Learning and Applications (ICMLA), pp. 989–992, 2016.