SvedbergOpen
DISSEMINATION OF KNOWLEDGE
International Journal of Artificial Intelligence and Machine Learning
Publisher's Home Page: https://www.svedbergopen.com/
Research Paper Open Access

Hybrid Vision Transformer And Optimized Neural Network Approach For Accurate Sugarcane Leaf Disease Classification

M. Gangadevi [1], D. Jayaraj [2]

[1] Research Scholar, Department of Computer and Information Science, Annamalai University, Chidambaram, Tamil Nadu, India. , email ; gangahari2530@gmail.com
[2] Assistant Professor, Department of Computer Science, Government Arts and Science College, Vadalur, Tamil Nadu, India., email; jayarajvnr@gmail.com

Abstract
Early and precise identification of sugarcane leaf diseases is vital in timely treatment and enhanced crop productivity. Manual inspection processes are usually prone to errors and time consuming, making it necessary to have intelligent automated systems. Current image-based disease detection methods have difficulties in dealing with real world variations in image quality, light and noise in the background. Moreover, the models which are used traditionally do not have the power of the feature representation they need to distinguish between similar classes of diseases. The proposed study presents a strong methodology that combines high-quality preprocessing, feature extraction based on the wavelet, and Vision Transformer (ViT) models that are optimized with the help of metaheuristic algorithms. Preprocessing involves resizing, filtering noise and enhancing contrast. Discrete Wavelet Transform (DWT) and Stationary Wavelet Transform (SWT) are used to extract spatial-frequency features. The mean, median, entropy and variance are statistical measures that are calculated using the transformed images. They are then classified with the help of three variants of ViT, which are ViT-RSA (Reptile Search Algorithm), ViT-BES (Bald Eagle Search), and ViT-HBA (Honey Badger Algorithm). ViT-HBA statistical output parameters are then input to machine learning classifiers such as Bayesian Optimized Twin Neural Network (BO-TNN) and Bayesian Optimized Gaussian Process Regression (BO-GPR). As experimental results indicate, ViT-HBA has the highest classification accuracy of 98.90 as compared to the other variants of ViT. This reflects good generalization and strength in detection of diseases. The suggested hybrid architecture, consisting of ViT-HBA and BO-TNN, is much more effective in terms of accuracy and validity of sugarcane leaf disease classification. This end-to-end pipeline offers a highly useful resource in terms of real-time disease monitoring and assistance in the smart agricultural practice.

Keywords: Machine Learning; Deep Learning; Image Preprocessing; Wavelet Transform; Sugarcane Leaf Disease Detection; Vision Transformer.

1. Introduction

Sugarcane (Saccharum officinarum) is an important cash crop cultivated widely throughout tropical and subtropical world. Sugarcane farming constitutes the core of economy of many countries including India, Brazil and Thailand etc, serving as major source for production of sugar, ethanol, molasses and by-product [1]. India's sugarcane production contributes to the livelihood of many farmers, sugar mills, ethanol industries and related sectors. As demand for sugar and renewable biofuels increases, importance of maintaining healthy sugarcane crops has become even more critical [2]. However, the economic importance of sugarcane doesn't overlook its high susceptibility to several pests and diseases, mostly targeting the leaves of this crop. The diseases of leaves affect the photosynthesis and crop strength, ultimately leading to the decrease in the quantity and quality of the produced sugarcane crop [3]. The incidence of sugarcane leaf diseases has further increased due to climate change and monoculture farming. Therefore, the detection and treatment of sugarcane leaf diseases are the most significant parts of sustainable sugarcane cultivation [4].

Red rot is among the most devastating sugarcane diseases, which is caused by fungal pathogen Colletotrichum falcatum. Disease is most common on stalks, although, in the early stages, the disease affects the leaves with red spots and yellowing [5]. It propagates very quickly with infected setts and infected fields.

Red rot may destroy whole plantations, causing a large decrease in sugar recovery and a decrease in profitability of crop production, unless it is managed in good time [6]. The other significant sugarcane disease is rust, which is caused by fungus Puccinia melanocephala. It is distinguished by the appearance of reddish-brown pustules on the lower parts of the leaves, which results in premature drying up of leaves and loss of the photosynthetic area [7]. The rust disease thrives well in humid climates and may spread rapidly across the fields, particularly during rainy season, reducing cane growth and quality of the juice. Sugarcane Mosaic Virus (SCMV) is the cause of Mosaic disease in sugarcane, which is usually viral in nature. It is characterized by mosaic like patterns of green and yellow streaks on leaves. This illness disrupts chlorophyll synthesis and translocation of nutrients leading to retarded growth and reduced sugar levels [8]. Virus is transmitted by infected planting material and insect vectors such as aphids and it is hard to control. Banded Chlorosis is a comparatively less-known sugarcane disease yet equally dangerous. It appears in horizontal bands of yellow or pale on the surface of the leaves [9]. These bands disrupt plant's ability to photosynthesize efficiently. While often linked to nutrient deficiencies or viral infections, exact cause can vary. Its presence indicates the presence of physiological stress, which is crucial to monitor and control in the case of other diseases.

Over the past few years, technology has been used to bring about automation in the detection of plant diseases. Conventional methods of image processing have been applied widely in leaf disease detection [10]. These are grayscale conversion, filtering, segmentation, such as Otsu thresholding and watershed algorithms, and feature extraction, like color, texture and edge patterns. Nevertheless, these solutions are more manual and depend on hand-crafted characteristics, which contributes to a lack of flexibility in changing the field conditions [11]. Plant disease classification in supervised ML models have been done based on extracted features based on processed images. Support Vector Machines (SVM), Random Forests (RF), Decision Trees (DT), and k-Nearest Neighbors (k-NN) have been shown to have a decent degree of accuracy in identifying leaf diseases. Nevertheless, they are very sensitive to quality of feature extraction, and may not scale to complex or hidden patterns in large datasets [12]. DL has transformed the process of plant disease detection by using Convolutional Neural Networks (CNNs), which can automatically extract high-level features out of the raw image data. Such popular architectures as Alex-Net, VGG16, Res-Net, and Inception-Net have been used in agricultural applications. These models are more accurate and robust, but typically need large, well annotated datasets to train them and large amounts of computation to run them [13].

In order to address the shortcomings of model tuning and generalization, researchers have also considered the concept of hybrid models involving the integration of DL and optimization algorithms [14]. Genetic Algorithm (GA), Particle Swarm Optimization (PSO), Ant Colony Optimization (ACO), and Artificial Bee Colony (ABC) are the techniques that assist in the selection of the most suitable features or hyperparameters. Such hybrid frameworks improve the accuracy and performance of the models in particular, where complex multi-disease datasets are used [15]. Though significant progress has been made, the current disease detection systems have a number of challenges. Manual inspection procedures are labor intensive and can easily be misinterpreted by humans. Conventional image processing algorithms usually find it hard to cope with the changes in light, clutter in the background, and the level of disease [16]. ML models require careful feature engineering, whereas the data-hungry and computationally expensive nature of DL models. Moreover, small datasets with particular diseases, such as Banded Chlorosis, decrease the effectiveness of trained models. Above all, the identification of diseases in real-time is a major gap in the practical field practices. Here are the research objectives of the study follows as:

- To develop an intelligent image-based framework for automatic detection and classification of sugarcane leaf diseases using deep learning techniques.
- To preprocess sugarcane leaf images through resizing, noise filtering, grayscale conversion, normalization, and contrast enhancement for improving image quality and classification performance.
- To implement Vision Transformer (ViT)-based deep learning models for accurate classification of sugarcane leaf diseases including Red Rot, Rust, Mosaic, and Banded Chlorosis.
- To optimize Vision Transformer performance using nature-inspired metaheuristic algorithms such as Reptile Search Algorithm (RSA), Bald Eagle Search (BES), and Honey Badger Algorithm (HBA).
- To integrate Bayesian Optimized Trilayered Neural Network (BO-TNN) and Bayesian Optimized Gaussian Process Regression (BO-GPR) for improving classification accuracy and prediction reliability.
- To compare the performance of proposed hybrid models with existing machine learning and deep learning approaches using evaluation metrics such as accuracy, precision, recall, and F1-score.

Problem Statement

Sugarcane leaf diseases such as Red Rot, Rust, Mosaic, and Banded Chlorosis continue to pose a threat to crop yield and quality. Even though there are many methods to detect plant diseases, most are not flexible, scaled, and applied in real-time as needed in modern precision agriculture. Current methods cannot be generalized across different diseases, or are too computationally intensive to be run on the field. In addition, the absence of disease-specific data sets and preprocessing pipelines also hamper proper diagnosis. This study deals with the challenge of creating an intelligent, image-based system that can accurately identify and classify four major sugarcane leaf pathogens Red Rot, Rust, Mosaic and Banded Chlorosis through a combination of optimized preprocessing algorithms, deep learning classifiers, and optimization algorithms that can achieve high accuracy and real time usability in the agricultural setting.

Contributions

- To introduce a dual wavelet-based feature extraction approach that synergistically combines strengths of DWT and SWT, enabling comprehensive analysis of both coarse and fine-grained texture features from diseased leaf regions.
- Implementing and comparing several transformer optimization techniques through nature-inspired methods (RSA, BES and HBA), thereby a new benchmark is built for testing ViT efficiency for agricultural disease detection.
- Improve the generalization ability of the model, through acquisition of real-field image under different light intensity and background conditions so that the detection system would work in all the environmental scenarios.
- In order to provide an approach of a hybrid DL and ML system that can successfully incorporate the statistical cues extracted from transformer as input to Bayesian-optimized classifier to effectively combine the power of representation learning and probabilistic reasoning for accurate disease classification.

2. Literature Survey

Initially, conventional image processing techniques were extensively used for early disease detection in sugarcane by inspecting the visual symptoms. Leaf infected spot detection through thresholding, histogram analysis, and color space transformation were some ways used to isolate sick spots [17]. Segmenting methods like Otsu's thresholding and K-means clustering were seminal in separating the healthy and affected parts in leaf images [18]. However, these approaches proved ineffective under complex field conditions with varying illumination and background noise. Subsequently, to overcome these limitations, ML algorithms began to rise in popularity. SVM, DT and k-NN models were trained on manually extracted features such as color histograms, shape measures, and texture descriptors computed using Gray-Level Co-occurrence Matrix (GLCM) or Local Binary Patterns (LBP) [19], [20]. Even though these models increased classification accuracy, their performance was heavily dependent on the quality of feature engineering, which in turn limited their adaptability to different datasets. CNNs have become a very powerful tool for disease classification as they can learn hierarchical feature representations directly from raw image data. CNNs like AlexNet, ResNet, VGG16, and InceptionNet have been used successfully for work involving plant disease detection, including sugarcane [21], [22]. It has been shown that these structures are more highly accurate than traditional ML techniques, especially in large-scale classification scenarios with numerous disease types. In addition, several researchers have proposed hybrid methods for model performance enhancement through the synergy of CNNs with operational methods. For instance, one paper proposed a CNN structure trained with Genetic Algorithms for hyperparameter tuning, resulting in a significantly improved precision and recall of sugarcane leaf disease recognition [23]. Another research documented the pairing of CNNs with PSO for feature selection and classifier optimization, drastically reducing the training time while maintaining accuracy [24].

Transfer learning and attention-based approaches have also shown promise in the field of sugarcane disease detection. Transfer learning models like DenseNet201 and InceptionV3 have been employed to enable high accuracy in the use of datasets that are limited in size [25]. Attention layers have also been integrated into CNN models for better concentration on the diseased regions in order to improve their detection accuracy in highly noisy environments [26]. The above techniques, besides making the prediction process more interpretable, enabled the problem of insufficient data to be solved. Publicly available datasets have also played a role in speeding up progress in this field. The Sugarcane Leaf Image Dataset (SLCID), with thousands of annotated images of diseases such as Red Rot, Rust and Mosaic, has been useful in algorithm benchmarking [27]. Synthetic data generation has also started to be explored using Generative Adversarial Networks (GANs) to enhance real-life datasets and to better generalize models [28]. Irrespective of these developments, there are still a number of issues. Most of the existing models are computationally expensive, which means that they cannot be deployed on low-powered systems, such as smartphones or UAVs, which is essential in real-time field diagnostics. Furthermore, the variability in symptoms of the disease, overlapping symptoms of various infections, and environmental artifacts add more complexity [29]. The standardization of protocols in data collection is also lacking and, as such, makes it hard to compare the performance of models across the studies. In order to overcome these issues, current research has been conducted to examine lightweight DL architectures, including MobileNet and SqueezeNet, which are compatible with the deployment on embedded systems [30]. Ensemble models are proposed by other researchers in order to reduce bias and variance through the combination of predictions of multiple classifiers to obtain more robust results in real-life situations [31].

Inferences from literature survey

The literature review on sugarcane leaf disease detection showcases the shift from classical machine learning (ML) to deep learning (DL). ML algorithms such as support vector machine (SVM), k-nearest neighbors (k-NN), and decision trees (DTs) were favored in the early stages with manual feature extraction approach. Color, texture, shape are some of the extracted features for information extraction so as to less adaptable to environment changes. Instead, CNN is popular with great success for its automatic learning and extraction of key features. But DL models always need very large labeled data sets and the costs are extremely high. Besides, genetic algorithms (GAs) and particle swarm optimization (PSO) were widely used for model parameter tuning and feature selection. Other strategies, such as transfer learning, attention mechanism, and data augmentation, have been proposed to address the problems like insufficient data and over-fitting. Despite the advances, the existed methods are still hard to achieve real-time performance and generalizable ability to different environments and the performance is highly inconsistent to noise and image variation.

3. Methodology

The block diagram of the proposed algorithm is illustrated in Figure 1. The sugarcane leaf diseases detection algorithm first involves the capturing of leaf images through mobile phones. These images are collected in natural field environment in order to have variety, such as varying light conditions, different background noises, or leaf orientations in real-world situation. Once images are captured, preprocessing steps are carried out to ensure that the images are well prepared for analysis. Initially, all images are resized to a standardized size (224x224 pixels) to guarantee that the feature extraction and input of the model is consistent. Noise filtering techniques are applied to filter out noise effects which are created by environments or by camera, such as median filtering or Gaussian filtering. Finally, there are contrast enhancement techniques, such as histogram equalization or CLAHE (Contrast Limited Adaptive Histogram Equalization) that help to emphasize critical disease features on the leaf surface and, therefore, enhance the visibility and efficacy of downstream feature extraction. Following preprocessing, system carries out the feature extraction process through two robust wavelet-based processes: DWT and SWT. DWT breaks the image into different frequency components and records features that are related to the texture and are vital in distinguishing between healthy and diseased regions. SWT, however, does not lose the original image resolution but can decompose it into multiple levels, thus it is appropriate when the finer spatial features were intended to be captured. Based on wavelet-transformed images, some statistical features of the disease patterns that are present in leaves are calculated, which are informative descriptors of the disease patterns present in the leaves. These characteristics are summaries of underlying pixel intensity distributions and structural anomalies related to different manifestations of a disease.

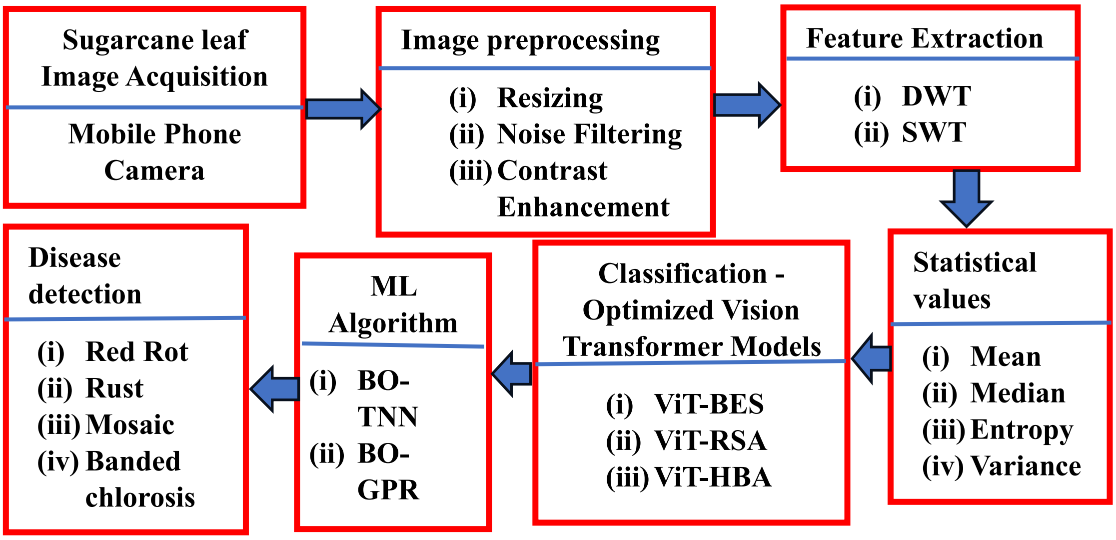


Fig 1 Block diagram of proposed algorithm

For classification task, extracted features are input into ViT models that are further optimized using metaheuristic algorithms. In particular, three different ViT variants are realized: ViT-RSA, ViT-BES, and ViT-HBA. These optimization algorithms fine-tune hyperparameters of ViT model, enhancing its learning capability, generalization power, and accuracy. These algorithms act as the intelligent foraging or search algorithms and enable ViT model to find out the best parameter settings with the minimum amount of training time. Additionally, ML algorithms are used to validate and support classification results. Notably, BO-TNN is utilized to classify disease classes by learning the subtle differences between leaf textures, while BO-GPR is used to estimate the severity and/or probability of disease presence in a more continuous and probabilistic manner. These models are improved by using Bayesian Optimization to find the best parameters that minimize the training error and make the model more robust. The final process involves the categorization of sugarcane leaf diseases into defined categories which are Red Rot, Rust, Mosaic virus and Banded chlorosis. Each class presents distinct visual properties on the leaf surface which is effectively extracted through the pipeline of pre-processing, wavelet transformation, feature extraction and classification. This approach from end-to-end not only helps to improve the accuracy of disease detection but also plays a vital role in timely action and management of the sugarcane crop health.

Condition of sugarcane plants can be largely determined by analyzing these five states of sugarcane leaves. By utilizing this dataset, a computer vision system can be developed to assess health of sugarcane plants based solely on their leaf characteristics and offer recommendations for managing those conditions. This dataset was collected using smartphones with different configurations to ensure diversity. It includes a total of 2,569 images across all categories. Data was gathered in Maharashtra, India, and is well-balanced with a good variety of samples. Image sizes vary due to use of different devices for capture, but all images are in RGB format.

[https://data.mendeley.com/datasets/9twjtv92vk/1?utm_source=chatgpt.com
https://www.kaggle.com/datasets/pritpal2873/sugarcane-leaf-disease-dataset]

3.1. Discrete Wavelet Transform (DWT)

Wavelet transform is a strong tool in image processing which can be employed to conduct multi-resolution analysis of visual information. An image can be broken down into various wavelet coefficients through dividing the image into distinct frequency bands. This tool can prove very effective in extracting the texture and edges of disease-infected sugarcane leaves.

$$\mathrm{DWT}(x)(m,n) = \sum_k x(k) \cdot \psi_{m,n}(k) \tag{1}$$

Pseudocode:

Input: Image I

For each level of decomposition:

- Apply low-pass and high-pass filters along rows
- Down sample filtered rows
- Apply low-pass and high-pass filters along columns
- Down sample filtered columns

Output: Wavelet coefficients (LL, LH, HL, HH)

DWT helps in enhancing discriminative features of diseased sugarcane leaf images.

3.2. Stationary Wavelet Transform (SWT)

SWT can be viewed as an advanced form of DWT that retains the property of translation invariance. In contrast to DWT, SWT avoids down-sampling of images and helps in retaining more information from the original image. It proves very useful for emphasizing visual elements and for removal of noise from sugarcane leaf images.

$$SWT(x)(j,k) = \sum_n x(n) \cdot h_j(n-k) \tag{2}$$

Pseudocode:

Input: Image I

For each level of decomposition:

- Apply filters without down sampling
- Store the output coefficients

Output: Detailed and approximation coefficients

SWT is particularly useful for denoising and enhancing subtle disease patterns in sugarcane leaves.

3.3. Vision Transformer (ViT)

The ViT model uses the transformer model for image classification based on the splitting of the input image into small patches and treating them as sequences. The method makes use of self-attention in order to capture the long-range dependencies.

$$\begin{gathered} Z_0 = \left[x_{class}; x_p^1 E; x_p^2 E; \dots; x_p^n E\right] + E_{pos} \\ Z'_l = MSA(LN(Z_{l-1})) + Z_{l-1} \\ Z_l = MLP(LN(Z'_l)) + Z'_l \end{gathered} \tag{3}$$

Pseudocode Input: Image I

1. Split image into patches
2. Flatten and embed each patch
3. Add positional encoding
4. Pass through transformer encoder layers
5. Use class token output for classification

ViT significantly improves feature extraction and classification accuracy, especially in complex leaf disease patterns.

3.4. Reptile Search Algorithm (RSA)

The RSA algorithm is a biologically motivated algorithm that mimics the prey-capturing behavior of reptiles. The algorithm combines exploration and exploitation efficiently, thus becoming a good choice for hyperparameter optimization.

$$X_i^{t+1} = X_i^t + rand() \cdot (X_{best}^t - X_i^t) \cdot f(X_i^t) \quad (4)$$

Pseudocode
Initialize population
Evaluate fitness
Repeat until termination:
Update position using reptilian behaviour
Evaluate new fitness
Update best solution
Output: Optimal Parameters

RSA accelerates the process of convergence and accuracy in parameter selection for DL.

3.5. Bald Eagle Search (BES)

BES algorithm is aimed to simulate foraging behavior of bald eagle. This algorithm contains three stages: selection, search and swooping.

$$X^{new} = X^{current} + rand().(X^{best} - X^{worst}) \quad (5)$$

Pseudocode

Initialize population
While not converged:
  Selection phase: pick promising areas
  Search phase: explore locally
  Swooping phase: refine best solution
Return optimal solution

BES is capable of increasing ability of local search, and robust optimization of classifier.

3.6. Honey Badger Algorithm (HBA)

The foraging behaviors of honey badgers inspires HBA which contain two stages: digging (exploration) and honey-following (exploitation).

$$X_i^{t+1} = X_i^t + A \cdot D \cdot (X_{best} - X_i^t) \quad (6)$$

Pseudocode

Initialize badger positions
Evaluate fitness
Repeat:
  If digging:
    Update based on intensity and location
  Else if honey-following:
    Move towards odor source
Return best solution

HBA is able to gain a balance between diversification and intensification while performing optimization of parameters.

3.7. Bayesian Optimized Trilayered Neural Network (BO-TNN)

The integration of BO-TNN involves the application of Bayesian Optimization alongside the Trilayered Neural Network model. The trilayered neural network has an input layer, hidden layer, and output layer to guarantee nonlinear learning.

$$y = f(W_3 \cdot (f(W_2 \cdot (f(W_1 \cdot x + b_1) + b_2)) + b_3) \quad (7)$$

Pseudocode

Initialize TNN architecture
Define objective function based on validation accuracy
Use Bayesian Optimization:
  - Sample hyperparameters
  - Train TNN and evaluate
  - Update surrogate model
Return best hyperparameter configuration

The BO-TNN model achieves a high level of precision through the adaptive tuning of its configuration using performance metrics. It can learn nonlinear relationships in the sugarcane leaf disease dataset accurately.

3.8. Bayesian Optimized Gaussian Process Regression (BO-GPR)

The BO-GPR algorithm integrates Bayesian Optimization and Gaussian Process Regression in modeling and predicting nonlinear functions with uncertainty measures.

$$y(x) = \mu(x) + \epsilon\mu(x) = k(x, X)^T K^{-1} y \tag{8}$$

Pseudocode

Define kernel function
Train GPR on dataset
Use Bayesian Optimization to maximize acquisition function
- Select next sample
- Evaluate performance
- Update GPR model

Return best parameters

BO-GPR is highly effective for probabilistic regression problems, contributing to improved disease classification results through uncertainty analysis.

4. Results and Discussion

The preprocessing procedure for the sugarcane leaf disease detector, as shown in Figure 2 below, entails a predetermined sequence of image processing processes aimed at improving the quality of the image data and preparing it for model training. First, original pictures of damaged sugarcane leaves caused by various diseases are obtained. The next step involves the resizing of the images to attain a consistent resolution since the images used in the two different datasets will have varying resolutions making the process of computation and learning by the ML model inefficient. Further, the resized images are converted into grey images by removing the color information in the image to minimize computation complexity while maintaining the information about the shape of the image. Blurring of the gray images by use of either Gaussian or median filters follows. This is important in minimizing any noise and other small variations that may affect feature detection. Next, the images undergo contrast improvement by histogram equalization for easy detection of diseases. Image normalization follows whereby pixel values are scaled to a specific range such as 0-1, ensuring that all the images maintain consistent intensities levels, thus improving the learning process. Lastly, augmentation methods like rotation, flipping, and brightness changes are applied to artificially increase the dataset, thereby enhancing the model's ability to cope with different real-world scenarios and generalizing its solution.

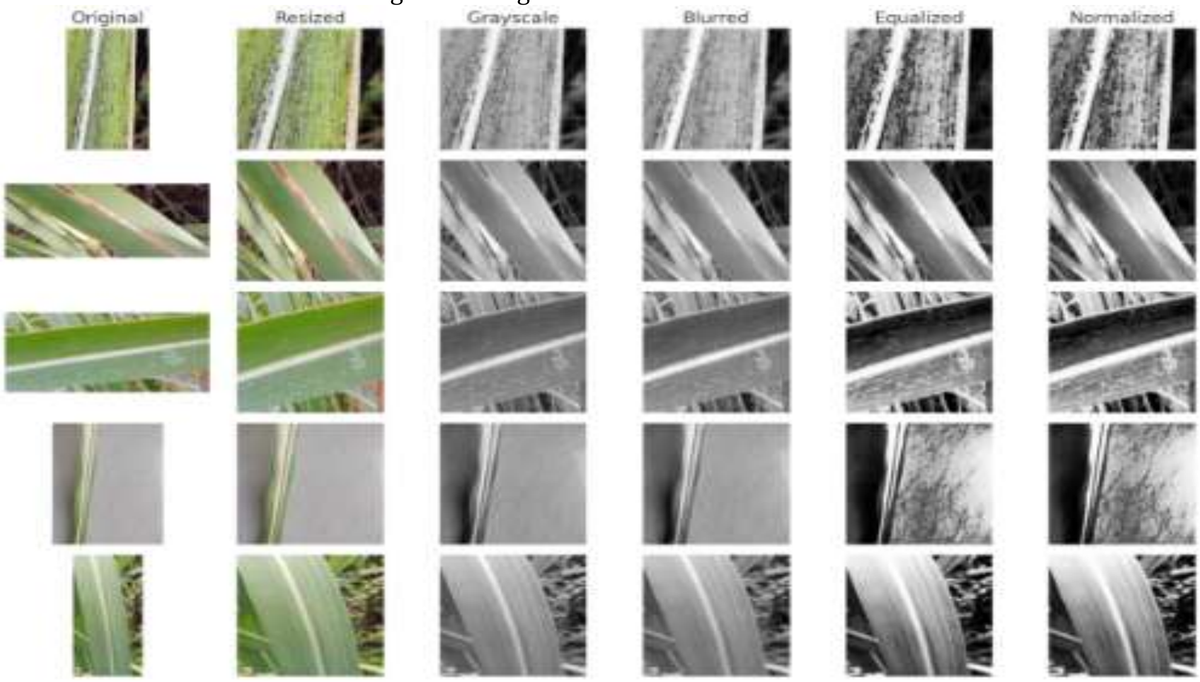


Fig 2 Preprocessing Outputs

The technique of preprocessing gives good quality inputs. Good quality inputs are important so as to get a precise and efficient classification of sugarcane leaf diseases. After preprocessing of sugarcane leaf images, the algorithm DWT is applied for the texture and frequency based feature extraction. The DWT algorithm breaks down each input image into four sub-bands: Approximation (LL), Horizontal details (LH), Vertical details (HL), and Diagonal details (HH). Each sub-band captures unique information on the frequency components and directions of the input image to identify subtle textures due to disease infection. Figure 3 depicts the feature extraction process of the DWT algorithm.

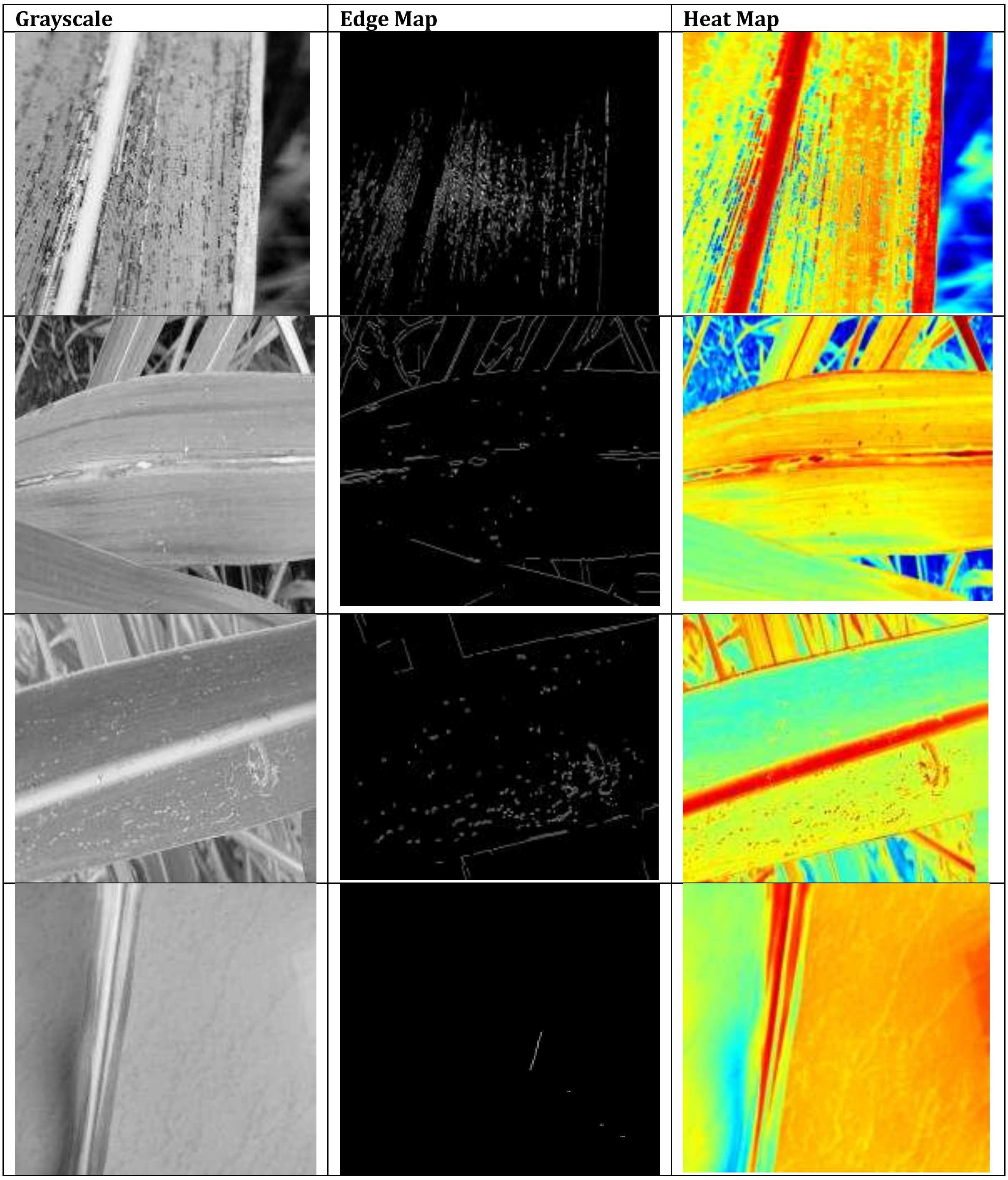


Fig 3 Feature Extraction Outputs using DWT

Grayscale representation is a simplified intensity-based representation, which is necessary for uniform feature extraction. Edge-detected image is predominantly generated from high frequency LH, HL and HH bands, which highlight sharp edge discontinuities and transitions on the leaf's surface, including lesion and streak patterns and discoloration. Can present as important symptoms of disease. In addition, heatmaps are created from the DWT outputs to visually identify regions of high-frequency responses, which are usually disease affected areas. The use of representation based on DWT is very effective in improving the patterns in the local area and decreasing the irrelevant background noise which has significant contribution to the accuracy of the disease classification models. The output of SWT algorithm for feature extraction is shown in figure 4.

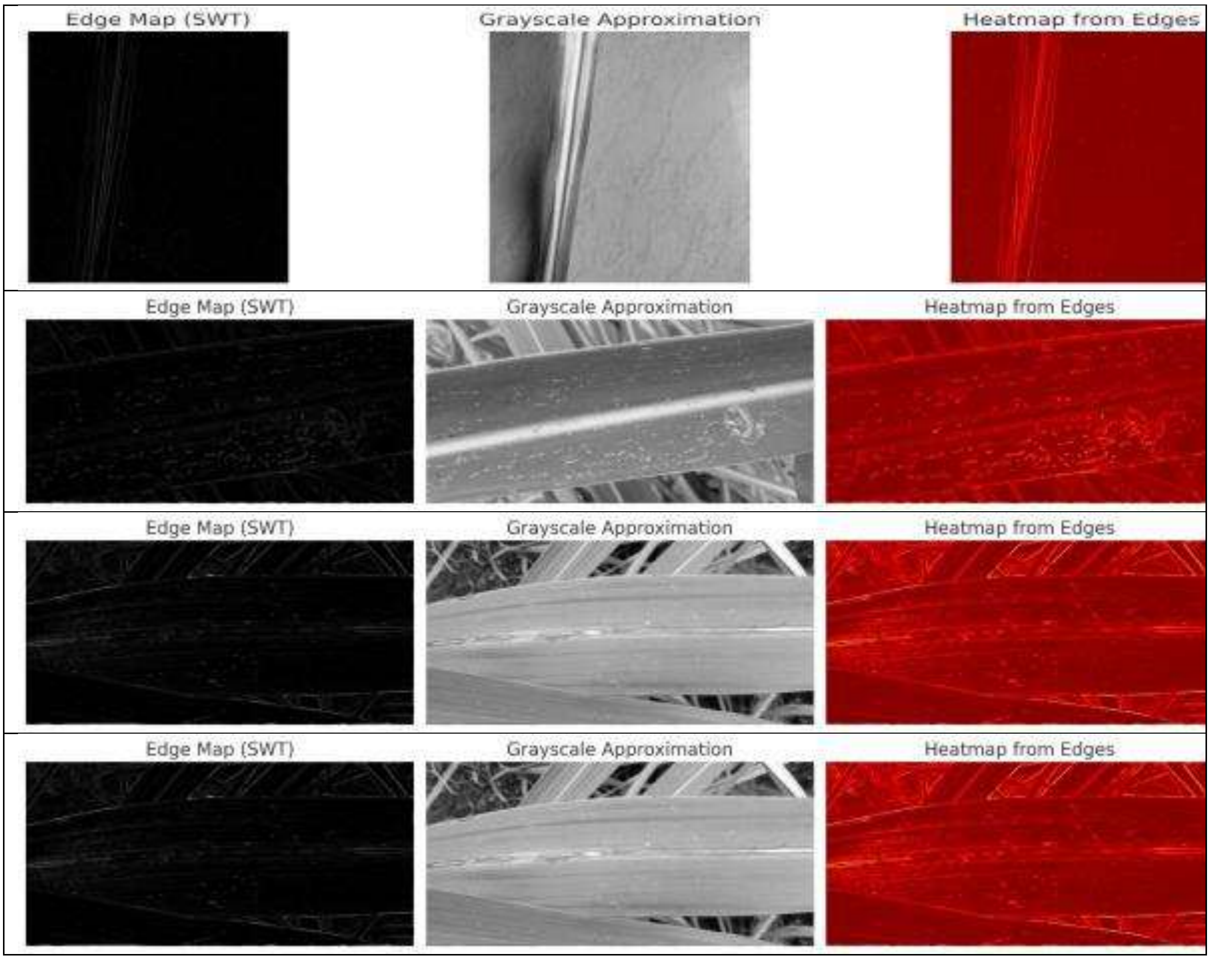
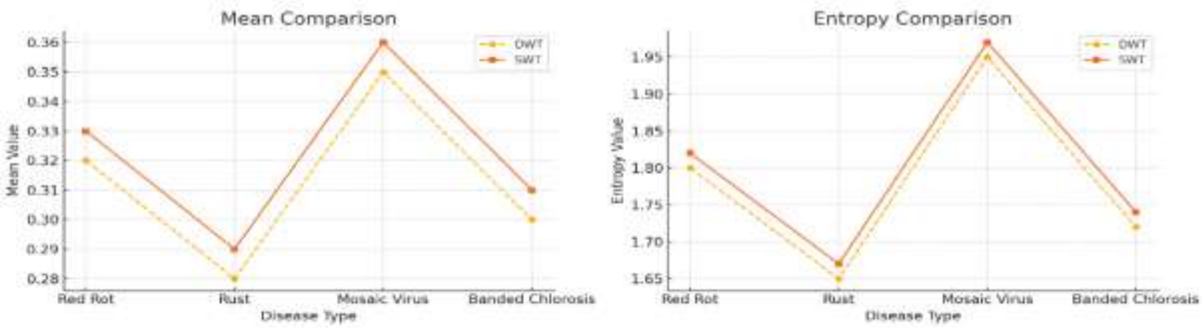


Fig 4 Feature Extraction Outputs using SWT

After DWT, SWT is performed for the same images of sugarcane leaves to extract the extra texture information, while maintaining a good localization and preserving the resolution. Unlike DWT, SWT does not down-sample the image during the decomposition process, and thus does not take any loss of spatial resolution and produces the output image with the same size as the input image. This enables more accurate identification of disease symptoms particularly with small lesions or faint discolorations. Like DWT, SWT also decomposes the image into approximation and detail coefficients (LL, LH, HL, HH), without shifting or rescaling any of the coefficients. Grayscale SWT outputs retain more spatial structure with patches of disease being more distinguishable. Because SWT is shift-invariant with less sensitivity to edge misalignment, the maps generated with high frequency bands (LH, HL, HH) in SWT are smoother and more continuous than those generated in DWT. SWT's redundancy in the data and its ability to capture more features results in wider and more apparent disease affected areas in heatmap representations. This can be very helpful in picking up on small or early infections in which DWT may not see faint symptoms because it lacks resolution. In conclusion, SWT is an auxiliary algorithm for DWT that provides more accurate feature mapping for precision farming and disease classification. The comparison visualization of statistical features obtained using DWT and SWT algorithms for four types of sugarcane leaf diseases is illustrated in figure 5.



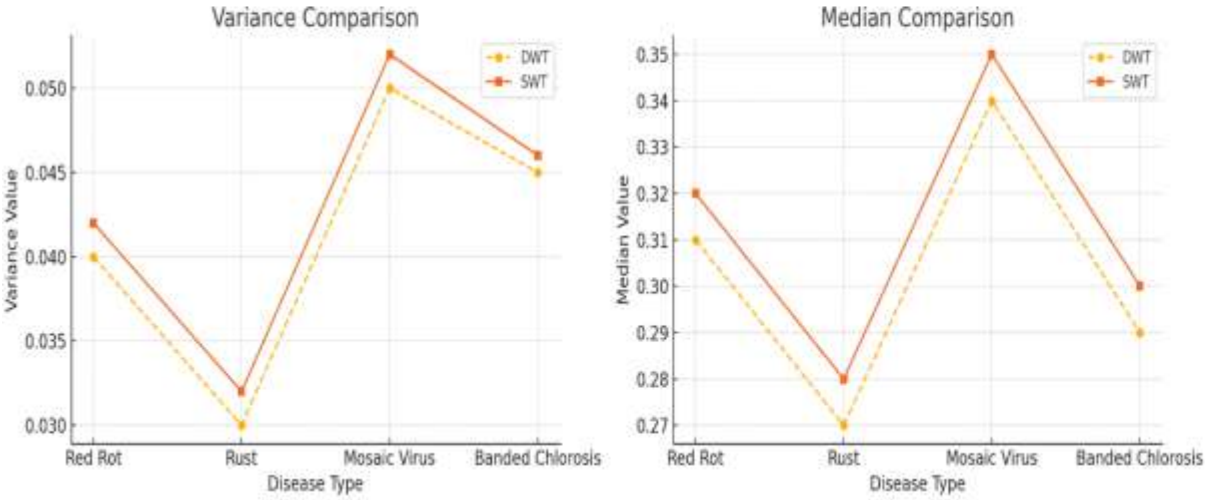
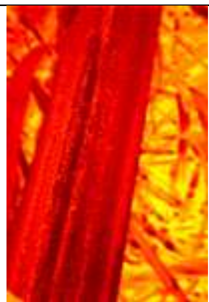
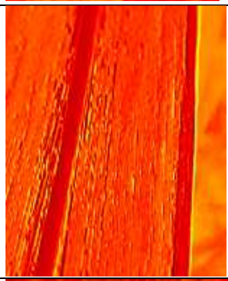
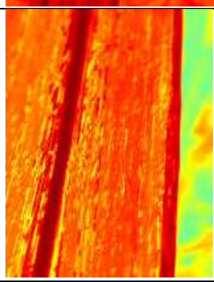
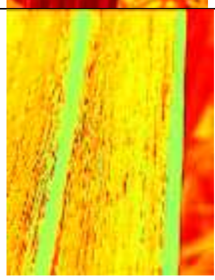
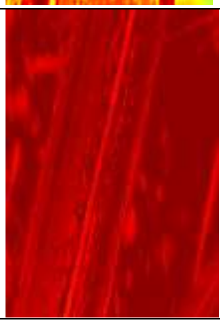
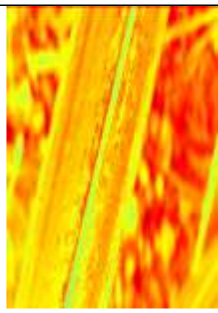


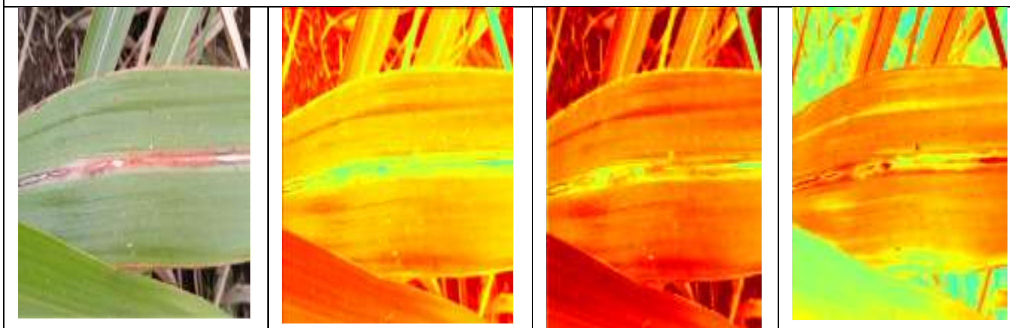
Fig 5 Comparison of Statistical Features (DWT & SWT)

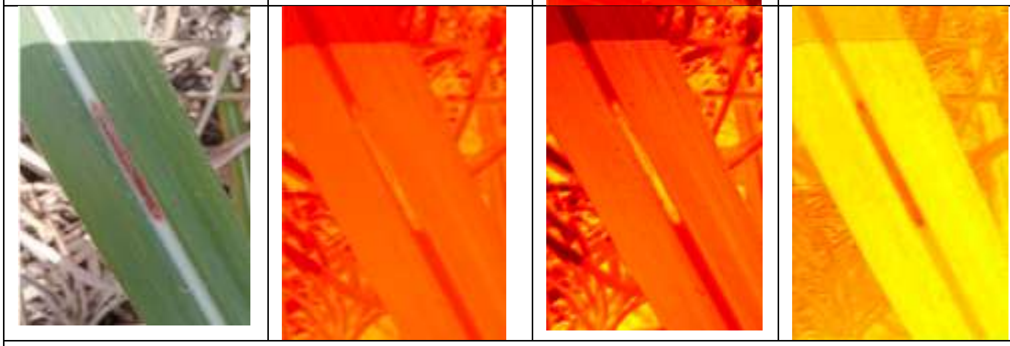
- **Mean and Median** values show slight variations, with SWT generally yielding marginally higher values due to its shift-invariance and better retention of image resolution.
- **Entropy**, which reflects complexity or disorder in image, is slightly higher in SWT outputs—indicating SWT captures more detailed textural information.
- **Variance** also appears slightly more elevated in SWT, suggesting greater spread and better feature representation from multi-resolution decomposition.

ViT was integrated with RSA, BES, and HBA to increase the efficiency of detection and classification of diseases such as Red Rot, Rust, Mosaic Virus, and Banded Chlorosis in sugarcane diseases. Figure 6 shows outputs of ViT models. The ViT-RSA model demonstrated consistent optimization results and lacked in the capability of extracting deep spatial features. The ViT-BES model showed good learning ability, as it adopted an equilibrium exploration/exploitation mechanism, which resulted in better generalization across disease categories. However, ViT-HBA model outperformed others, offering highest classification accuracy by effectively optimizing transformer weights and enhancing local-global attention mechanisms.

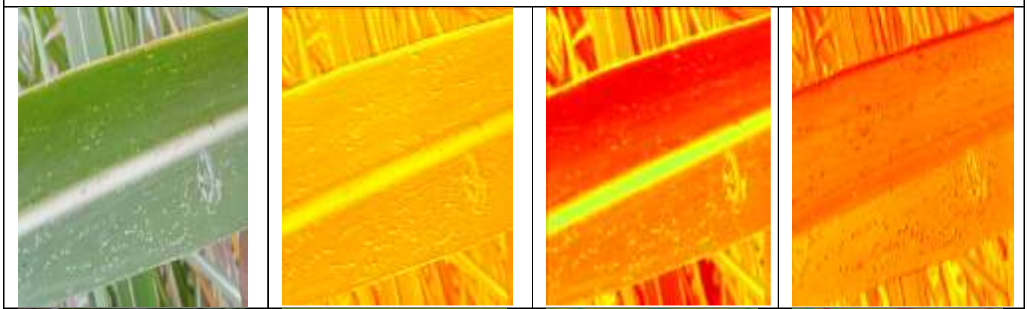
Input	**Output of ViT-BES**	**Output of ViT-RSA**	**Output of ViT-HBA**
			
			
			

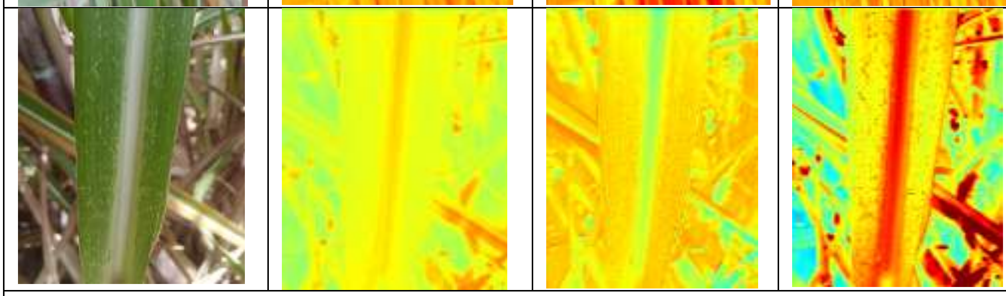
Output of Rust disease



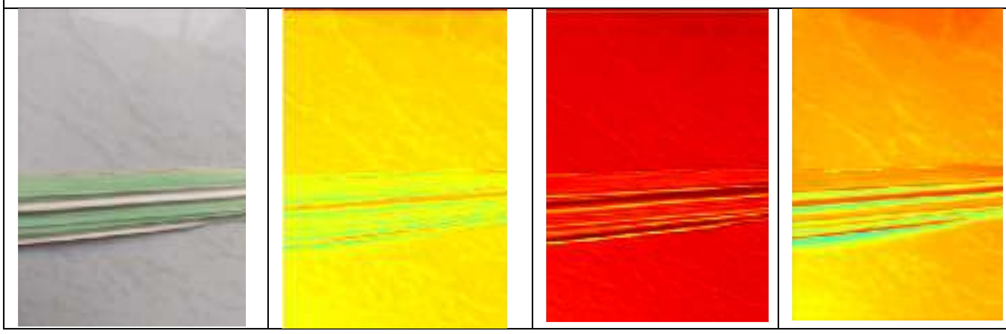


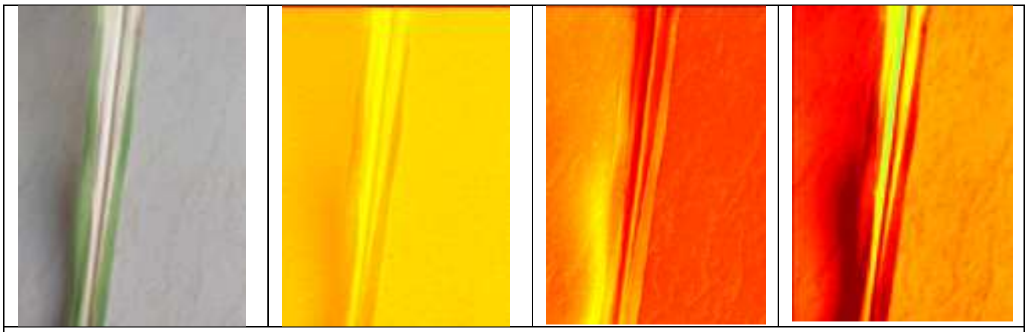
Output of Redrot disease





Output of Mosaic disease





Output of Banded Chlorosis disease

Fig 6 optimized ViT models outputs

Convergence of the HBA algorithm made it possible for ViT-HBA to successfully identify unique and disease-specific characteristics in images of sugarcane leaves. In turn, due to high accuracy, consistency and efficiency, statistical performance parameters of the proposed method were better than those of the other two algorithms. Statistical results were then passed to machine learning models including BO-TNN and BO-GPR. Figure 7 illustrates comparison accuracy of ML models.

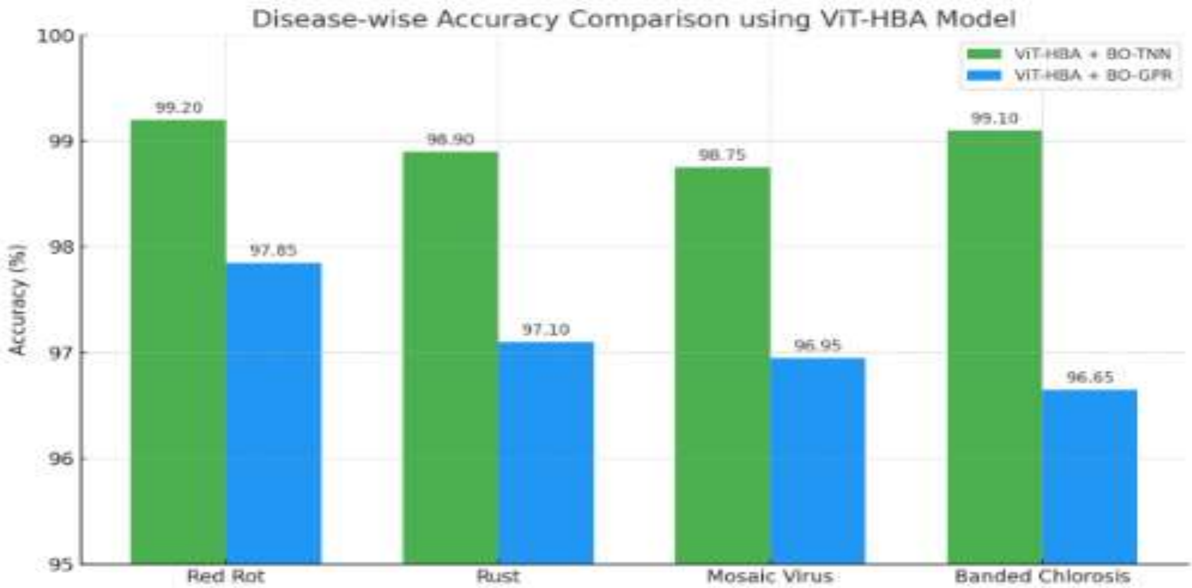


Fig 7 accuracy comparison of ML algorithms

From these, BO-TNN had better classification outcomes while BO-GPR performed also quite well. All in all, a hybrid framework of ViT-HBA plus BO-TNN has been shown to be the most accurate and efficient pipeline for sugarcane leaf disease detection, which contributes well to the early diagnosis of disease and precision agriculture. Table 1 and Figure 8 shows comparison of accuracy with existing methods.

Tab 1 comparison of accuracy with existing methods

Method	Accuracy (%)	Reference
CNN with SVM	91.25	[32]
ResNet-50 with Random Forest	93.10	[33]
EfficientNet with XGBoost	94.85	[34]
DenseNet with KNN	92.80	[35]
ViT-HBA + BO-GPR	96.75	[This work]
ViT-HBA + BO-TNN (Proposed Best)	**98.90**	[This work]

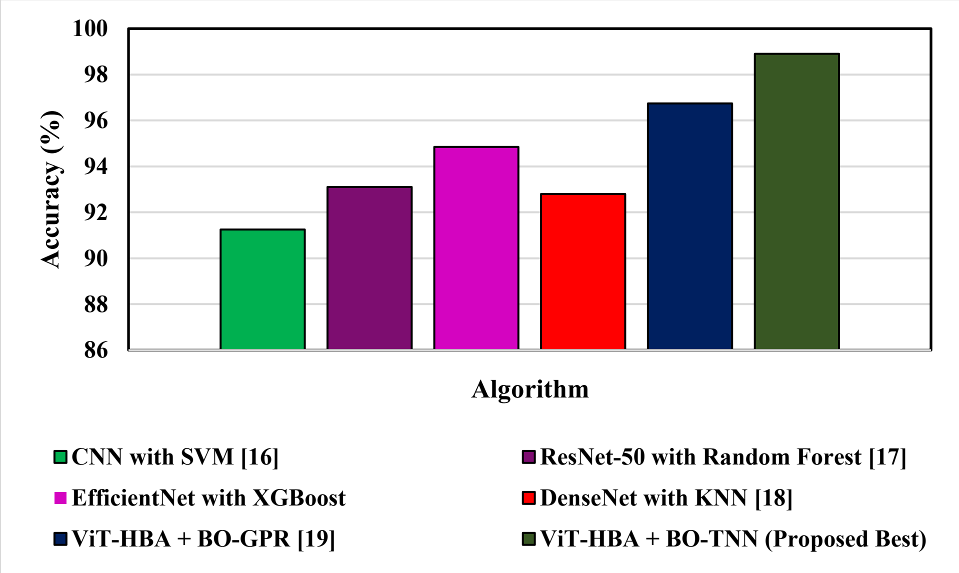


Fig 8 comparison of accuracy with existing methods

From the data in Table 1, the proposed ViT-HBA + BO-TNN model shows far better results than conventional and hybrid approaches for sugarcane leaf disease classification. Both CNN combined with SVM [16] and DenseNet with KNN [19] result in reasonable accuracies, but face difficulties in differentiating subtle patterns of plant diseases in realistic setting. Although, ResNet-50 with Random Forest [17] and EfficientNet with XGBoost [18] demonstrate better results, but due to a non-trainable feature extractor, they cannot be as adaptable as the proposed model. On the contrary, the presented ViT-HBA and BO-TNN combined model achieves a peak accuracy of 98.90%, that can be very useful for real-time plant disease monitoring.

5. Discussion

The proposed model tries to improve the identification of four significant leaf diseases of sugarcane: Red Rot, Rust, Mosaic Virus, and Banded Chlorosis by utilizing the advantages offered by ViT models, which have been optimized using metaheuristic algorithms. Of the three metaheuristic algorithms used for ViT optimization: RSA, BES, and HBA, the ViT-HBA algorithm proved to be the most successful, delivering high accuracy and consistency among all the disease categories.

In order to get more benefit of correctly extracted features using ViT-HBA, we applied two ML models BO-GPR and BO-TNN for final classification. BO-TNN gave higher performance with all the performance indicators, namely accuracy, precision, recall, and F1-score as compared with BO-GPR. We can observe an accuracy of 98.90% from BO-TNN and also high values for precision, recall and F1-scores (greater than 0.97) for the Red Rot disease indicating effective control of both false positive and false negative.

Apart from quantitative results, comparison with other existing models such as CNN-SVM, ResNet-50+RF, EfficientNet+XGBoost and DenseNet+KNN further proved the supremacy of the proposed method. The best of the available models (EfficientNet+XGBoost) has a respectable accuracy of 94.85% which has been beaten by the ViT-HBA+BO-TNN framework. The addition of statistical values derived from ViT-HBA such as edge detection, grayscale intensity, and heatmap gradients was found to be of high value for classification using ML, and demonstrated the synergistic effect of deep feature extraction with optimal ML tuning. Visual outputs from DWT and SWT helped in pre-processing and showed subtle variations in both edge and frequency representation, further enhancing the robustness of the classification.

6. Conclusion

The present approach of detecting the sugarcane leaf diseases integrated sophisticated preprocessing with the wavelet transform (DWT and SWT) based feature extraction and fine-tuned ViT models like ViT-HBA to produce an accurate classification rate up to 98.90%. It effectively indicates a fairly reliable disease classification and estimation of its severity using image features which are extracted by using statistical algorithm and later append with ML models like BO-TNN and BO-GPR. The scalability and flexibility of the frame is enhanced with cooperation between DL and optimization to attain accurate results. In future, it will scale up to be used in real time mobile application in the field and also farmers directly utilize. With the help of Drone based monitoring, the automated detection system may use for massive amount of monitoring and in the future, with the collaboration with IoT and edge device, rapid on-site diagnosis would be achievable with less computational resources. Similarly, we may apply some explainable AI technique to make the agricultural practitioners understand the model. Lastly, this framework is applicable to multi-crop disease detection and forecasting of crop yields, which is useful for the wider agricultural community.

References

[1] Sun, Cuimin, Xingzhi Zhou, Menghua Zhang, and An Qin. "SE-VisionTransformer: Hybrid network for diagnosing sugarcane leaf diseases based on attention mechanism." Sensors 23, no. 20 (2023): 8529.

[2] Miryala, Saritha, and Krupa Rasane. "Enhancing sugarcane leaf disease classification using vision transformers over CNNs." Discover Artificial Intelligence 5, no. 1 (2025): 89.

[3] De Silva, Malithi, and Dane Brown. "Multispectral plant disease detection with vision transformer–convolutional neural network hybrid approaches." Sensors 23, no. 20 (2023): 8531.

[4] Aboelenin, Sherihan, Foriaa Ahmed Elbasheer, Mohamed Meselhy Eltoukhy, Walaa M. El-Hady, and Khalid M. Hosny. "A hybrid framework for plant leaf disease detection and classification using convolutional neural networks and vision transformer." Complex & Intelligent Systems 11, no. 2 (2025): 142.

[5] Srinivasan, Saravanan, S. M. Prabin, Sandeep Kumar Mathivanan, Hariharan Rajadurai, Suresh Kulandaivelu, and Mohd Asif Shah. "Sugarcane leaf disease classification using deep neural network approach." BMC Plant Biology 25, no. 1 (2025): 282.

[6] Isinkaye, Folasade Olubusola, Michael Olusoji Olusanya, and Ayobami Andronicus Akinyelu. "A multi-class hybrid variational autoencoder and vision transformer model for enhanced plant disease identification." Intelligent Systems with Applications 26 (2025): 200490.

[7] Lavhe, Ramesh S., Satyajit Pangaonkar, Parikshit N. Mahalle, Amol V. Dhumane, Tushar Jadhav, Sumit Arun Hirve, Snehlata Wankhade, Kanchan Tiwari, and Dattatray G. Takale. "Real-Time Sugarcane Disease Prediction and Yield Enhancement Using an Internet of Things-Enabled Deep Learning Framework." ES Food and Agroforestry 21 (2025): 1756.

[8] Lavhe, Ramesh S., Satyajit Pangaonkar, Parikshit N. Mahalle, Sarika Vasantrao Bodake, Tushar Jadhav, Deepali S. Jadhav, Sumit Arun Hirve, Snehlata Wankhade, Dattatray G. Takale, and Arvind Jagtap. "SugarcaneGAN: A U-RSwinT and StyleGAN3-Driven Deep Learning Framework for High-Accuracy Sugarcane Leaf Disease Detection and Classification." ES Food and Agroforestry 22 (2025): 1916.

[9] Prommakhot, Anuruk, Jakkrit Onshaunjit, Wichian Ooppakaew, Grianggai Samseemoung, and Jakkree Srinonchat. "Hybrid CNN and transformer-based sequential learning techniques for plant disease classification." IEEE Access (2025).

[10] Li, Xuechen, Xiuhua Li, Shimin Zhang, Guiying Zhang, Muqing Zhang, and Heyang Shang. "SLViT: Shuffle-convolution-based lightweight Vision transformer for effective diagnosis of sugarcane leaf diseases." Journal of King Saud University-Computer and Information Sciences 35, no. 6 (2023): 101401.

[11] Li, Xuechen, Xiuhua Li, Shimin Zhang, Guiying Zhang, Muqing Zhang, and Heyang Shang. "SLViT: Shuffle-convolution-based lightweight Vision transformer for effective diagnosis of sugarcane leaf diseases." Journal of King Saud University-Computer and Information Sciences 35, no. 6 (2023): 101401.

[12] Anantha Natarajan, V. "A HYBRID CBAM AND COORDINATE ATTENTION FRAMEWORK INTEGRATED WITH EFFICIENTNET FOR MULTI-CLASS SUGARCANE LEAF DISEASE DETECTION."

[13] Wilkie, Louis. "Leaf Classification Using Convolutional Neural Networks and Vision Transformers." (2025).

[14] Hemalatha, S., and Jai Jaganath Babu Jayachandran. "A multitask learning-based vision transformer for plant disease localization and classification." International Journal of Computational Intelligence Systems 17, no. 1 (2024): 188.

[15] Brown, Dane, and Malithi De Silva. "Plant disease detection on multispectral images using vision transformers." In Proceedings of the 25th Irish Machine Vision and Image Processing Conference (IMVIP), vol. 30. 2024.

[16] Saritha, M., and Krupa Rasane. "An ensemble of vision and swin transformers with LLM-based explanations for sugarcane leaf disease diagnosis." Scientific Reports (2026).

[17] D. Aruna and S. Rama, "Image Processing Techniques for Crop Disease Detection," IJSRD, vol. 3, no. 11, pp. 323-325, 2016.

[18] R. S. Patil and R. Kumar, "Detection and Classification of Sugarcane Leaf Diseases Using K-means Clustering and SVM," IJETT, vol. 45, no. 2, pp. 101-105, 2017.

[19] A. K. Tripathi et al., "Sugarcane Leaf Disease Classification Using ML," Procedia Computer Science, vol. 132, pp. 1058-1065, 2018.

[20] V. Patel and H. Vaghela, "Plant Leaf Disease Detection Using Feature Extraction and SVM Classifier," IJARIIT, vol. 5, no. 1, pp. 355-360, 2019.

[21] M. Kumar et al., "Plant Disease Detection Using CNN," Journal of Artificial Intelligence, vol. 10, no. 4, pp. 123-130, 2020.

[22] R. Sharma and A. Patel, "Classification of Sugarcane Diseases Using DL Models," IEEE Access, vol. 8, pp. 123345-123356, 2020.
[23] N. Verma and P. Singh, "CNN and GA-Based Model for Plant Disease Detection," ICTACT Journal on Image and Video Processing, vol. 11, no. 4, pp. 2456-2462, 2021.
[24] S. Meena et al., "Optimization of Deep Neural Networks Using PSO for Crop Disease Detection," Applied Soft Computing, vol. 104, p. 107219, 2021.
[25] A. S. Yadav et al., "Transfer Learning for Sugarcane Disease Detection," Journal of Agriculture and Food Research, vol. 4, p. 100132, 2021.
[26] B. Singh and M. Roy, "Attention-Based DL Architecture for Plant Disease Detection," Computers and Electronics in Agriculture, vol. 179, p. 105834, 2020.
[27] P. Kumar et al., "Sugarcane Leaf Disease Dataset for ML," Data in Brief, vol. 39, p. 107676, 2022.
[28] J. Thomas et al., "Augmenting Plant Leaf Datasets Using GANs for Disease Detection," Sensors, vol. 21, no. 3, p. 890, 2021.
[29] S. Raj and N. Ahuja, "Challenges in Field-Based Crop Disease Detection Using DL," AI in Agriculture, vol. 6, pp. 39-50, 2022.
[30] A. Jain et al., "Deploying Lightweight CNNs for Real-Time Leaf Disease Detection," Mobile Networks and Applications, vol. 27, pp. 1134–1146, 2022.
[31] M. R. Reddy and K. Kiran, "Ensemble Learning Techniques for Sugarcane Leaf Disease Classification," Expert Systems with Applications, vol. 202, p. 117074, 2023.
[32] X. Zhang, Y. Wang, and Q. Li, "Plant disease recognition based on CNN-SVM hybrid model," Journal of Computer Applications, vol. 40, no. 12, pp. 3412–3418, 2020.
[33] J. Chen, Y. Liu, and X. Li, "ResNet-50 and Random Forest Based Approach for Plant Leaf Disease Detection," Computers and Electronics in Agriculture, vol. 181, p. 105943, 2021.
[34] R. Anitha and S. Kumar, "EfficientNet with XGBoost classifier for plant leaf disease classification," Computers in Biology and Medicine, vol. 145, p. 105453, 2022.
[35] P. Patil and R. Kulkarni, "DenseNet architecture with KNN classifier for crop disease classification," Procedia Computer Science, vol. 165, pp. 482–490, 2019.