Smart Sugarcane Health Management: Combining Transformer Diagnostics and Adaptive Spraying Strategies

D. Gowri Shankar*[ID], R. Jayaparvathy[ID]

Department of Electrical and Electronics Engineering, Sri Sivasubramaniya Nadar College of Engineering, Chennai 603110, India

Corresponding Author Email: gshankar.rdg@gmail.com

ABSTRACT

Foliar diseases pose a major threat to sugarcane productivity, but timely detection and management remain difficult for farmers. This work presents a real-time decision-support approach that integrates a hybrid CNN-Transformer model with NSGA-II for precision spraying recommendation, based on disease detection and severity estimation. The combination of the CNN and the Transformer module is used to capture fine-grained local lesion features and global disease patterns, respectively. A two-head design of the Transformer module predicts disease type and severity. The proposed hybrid model is benchmarked with a dataset comprising 21,637 sugarcane leaf images, achieving a higher classification accuracy of 98.8% and a better MSE of 0.072 than the CNN-only (89.5%) and Transformer-only (91.2%) models. The proposed hybrid model also outperformed the models including ResNet50, DenseNet201, EfficientNet-B7, and ViT-B/16, with an accuracy of 98.8%, and F1-score of 98.7% respectively, confirming its robustness against 11 disease categories. The NSGA-II module uses disease severity and weather conditions to suggest if spraying is needed, which pesticide to use, dosage level, and the best time to apply it. Thus, making the system useful for farmers as a mobile tool for sugarcane health management.

1. INTRODUCTION

Sugarcane, a principal global crop, significantly contributed 21% to sugar production globally between 2000 and 2020 [1] and is essential for many rural livelihoods, particularly in India where it is cultivated on over 5 million hectares. Though the sugarcane has a high economic importance [2], it faces serious risks from various diseases namely the red rot, grassy shoot, yellow leaf disease, rust, and smut, which can decrease the yield by more than 50% if not dealt properly [3]. The conventional approach to the disease management requires the use of manual inspections and chemical treatment, which may be inefficient, expensive and harmful to the environment [4]. Although precision agriculture tools have emerged, their adoption is limited due to high costs, lack of infrastructure, and insufficient technical expertise among farmers [5].

To identify sugarcane disease accurately from field images, recent developments in deep learning, especially convolutional neural networks (CNNs) and transformer architectures, have found to offer satisfying results [6]. But the current systems mostly concentrate on classification [7] and do not consider other vital aspects like the severity of the disease, environmental conditions, and decision-making based on the factors that are important in disease management. A disease management system should be able not only to identify diseases but also offer recommendations on how the pesticides should be used, their dosage and timing.

In order to overcome these challenges, the work proposes an interconnected decision-support system, which includes a hybrid CNN-Transformer model to detect the diseases and evaluate their severity, as well as a multi-objective optimizer of NSGA-II to create effective spraying recommendations. Its objective is to establish a system that combines visual analysis and decision-making system in order to promote the control of sugarcane diseases.

Management of sugarcane diseases in the field have huge gaps such as the restriction of manual scouting which is not able to identify early-stage infections. There is a lack of a unified disease management system to detect and give treatment recommendations, which is inhibiting effective responses. Also, the necessity to reduce the use of chemicals is acute due to environmental issues and regulations. These issues underscore the need to have a holistic framework that will facilitate accurate diagnosis and environmentally friendly treatment to farmers.

A Hybrid CNN-Transformer architecture has been presented in the proposed work as an effective deep learning model, a combination of CNN as a local lesion feature extractor and Transformer as a global context modeler, this hybrid model can be used to classify diseases and estimate their severity.

The hybrid model uses one image of a leaf to predict both the type and severity of a disease at the same time. Training and validation have been done using an expert-validated data

set (21,637 images) which is a large and various data set and has been manually validated with geographic, temporal and climatic supplementary metadata to achieve better generalization. An NSGA-II multi-objective optimization engine has a spraying recommendation module, which estimates trade-offs between yield, treatment efficiency, and environmental risk to come up with actionable solutions.

2. RELATED WORK

Detection of plant disease, such as sugarcane leaf infection, is now possible by using deep learning, in particular, by the CNN models such as EfficientNet-B7 and DenseNet201. Nevertheless, these models are more likely to utilize small datasets and are difficult to apply in different field conditions [8]. Vision Transformers Fine-tuned Vision Transformers (ViTs) on larger datasets have demonstrated excellent performance, with 96.5% accuracy when compared to ResNet50 and VGG16 [9]. Hybrid CNN-Transformer can classify images with added local and global features, and some of the lightweight architectures can run on a mobile platform [10, 11]. In spite of developments, current systems are more concerned with prediction of diseases and not the severity meaning that a combined model can offer a holistic, severity-informed decision support.

Accurate estimation of disease severity is essential in good spraying recommendations since it is more related with yield loss compared to type of disease. To evaluate the area of lesions on plant leaves, research has used different segmentation techniques, including SLIC super-pixels [12, 13], K-Means clustering [14], Watershed [15], U-net [16], and Mask R-CNN [17]. However, most of the studies are based on small datasets and are not validated in the field, which compromises the reliability.

The temperature and humidity are the crucial factors that determine the outbreak of sugarcane diseases. The maximum temperature, morning humidity, and sunshine are positively related to brown spot disease whereas minimum temperature, evening humidity, and wind speed have a negative correlation [18]. In the case of brown rust, the afternoon humid thermal ratio and the time span of the temperature are important predictors of the severity with a predictive accuracy of between 73-85% [19]. Sugarcane smut is favored by hot and dry environments (30-35°C), where wind assists in spreading the fungal spores, and dry soil improves the survival of the spores [20]. Moreover, steam and humidity enlarge infections by stem borers and improper rain falls interfere with the pest cycles [21].

The use of multi-objective evolutionary algorithms and in particular NSGA-II is critical to the optimization of agricultural systems, combining the economic, ecological, and operational aspects. NSGA-II utilizes fast nondominated sorting process, superior selection and parameter-free comparative operator to help locate a wide variety of solutions that are Pareto-optimal even when confined [22]. The use of

the Fuzzy-Expert-NSGA-II even enhances the agricultural planting strategies by integrating expert rule-based approaches in managing the constraints, fuzzy mathematics in the representation of the objectives, and adaptive searching approaches, which provides better results in the uncertain and complex environment in agriculture [23].

Most systems used to diagnose sugarcane diseases provide only basic classification of the disease and do not provide accurate estimates of the severity of the disease, which limits farmers' ability to manage their fields efficiently. Many of these systems do not validate their results based on expert opinions, and most of them do not use any form of deep learning models to improve the recommendations for the best spray options based on the current disease state. Therefore, an integrated system that provides disease detection with a high degree of accuracy, disease severity assessments, and optimum spray recommendations as a complete package has yet to be developed.

Building upon these limitations, this work proposes an integrated, real-time decision-support system that brings disease detection, severity estimation, and spraying recommendations into one unified framework. The key contributions of this study are as follows:
• A hybrid CNN-Transformer model is developed to not only identify the disease but also estimate how severe it is, offering more meaningful guidance than approaches that stop with classification alone.
• The model is trained on a large and diverse collection of 21,637 leaf images gathered from real fields, helping it performs reliably under practical farming conditions.
• An NSGA-II–based recommendation module uses both the severity level and prevailing weather conditions to suggest whether spraying is needed, the suitable pesticide, and the best time to apply it.
• Finally, the designed system is made to work in real time on mobile devices, giving farmers a tool that can support timely and informed decisions in their daily field activities.

The rest of this work is organized as follows. The methodology section explains the design of the hybrid CNN-Transformer model, the dataset preparation, and the integration of the NSGA-II module. The results and discussion highlight the performance of the proposed system and its advantages under real field conditions. Finally, the conclusion summarizes the key findings and outlines possible directions for future work.

3. MATERIALS AND METHODS

3.1 System overview

The proposed workflow of the proposed system which combines image-based disease detection, severity estimation, climate metadata and selective spraying recommendations through optimization techniques, is represented in Figure 1.
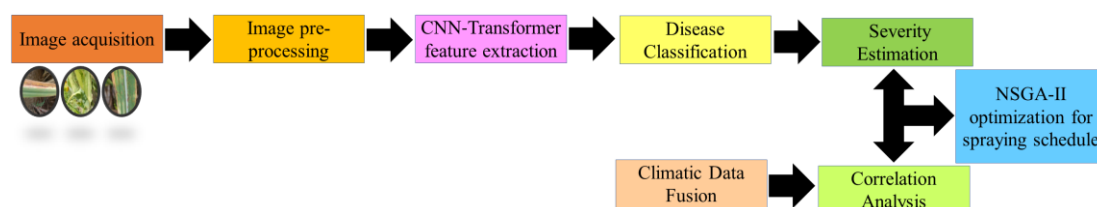

Figure 1. Overview of proposed system

3.2 Datasets

3.2.1 Sugarcane image dataset

The dataset used in this work consist of images from five Mendeley Data repositories [23-27], which ensures the variety in leaf appearance, disease stages, lighting, and backgrounds. The dataset comprises 11 categories (Banded Chlorosis, Mosaic, Ring Spot, Viral Disease, Grassy Shoot, Pokkah Boeng, Rust, Yellow Leaf, Red Rot, Sett Rot), including 10 sugarcane diseases and healthy leaves, totaling 21,715 images. Images were manually reviewed to eliminate duplicates and low-quality samples, and disease labels were verified against published descriptors. The selected dataset provides a representative sample of sugarcane leaves in natural environments, making them suitable for training a robust multimodal classification-severity model. Sample sugarcane leaf images used for this work is shown in Figure 2.
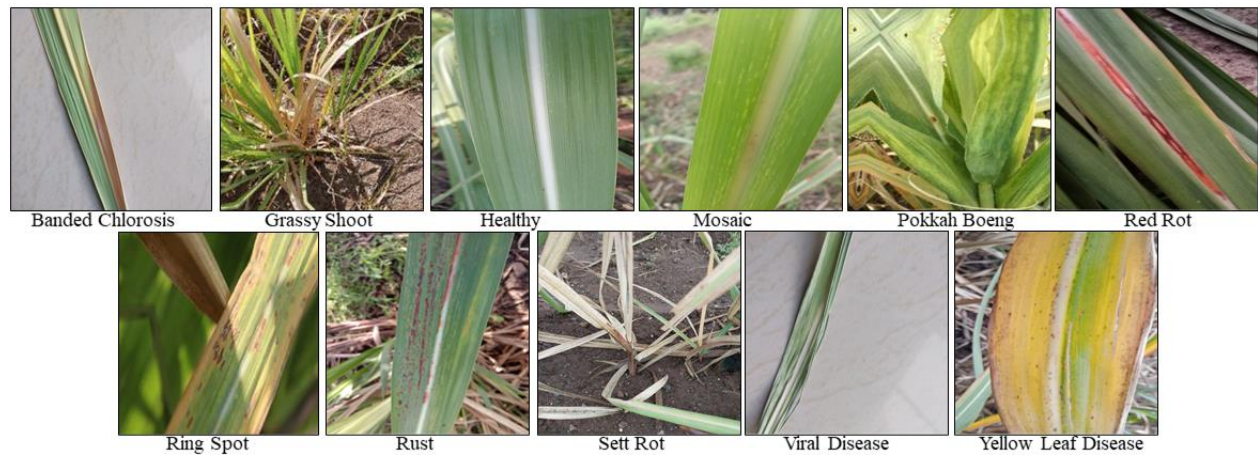

Figure 2. Sample sugarcane images used for this work

3.2.2 Climatic dataset

Important variables like wind speed, temperature, weather, and humidity are included in the climate data used in this study, which was obtained from a public dataset [28]. It offers a representative sample of environmental patterns in important sugarcane-growing regions over a wide geographic and temporal range. The dataset was chosen because it closely reflects the conditions and information that farmers use when deciding how to apply pesticides. This makes it suitable for building a system that also considers environmental factors, which are essential for giving accurate and practical spraying recommendations.

3.3 Data pre-processing

A standardized pre-processing pipeline was used ensuring reliable and uniform quality across the disease datasets and the climatic inputs. The images of sugarcane leaves were resized to 224×224, the intensity was normalized, and the lighting variations were improved using CLAHE [29].

Table 1. Number of images per class in training, testing and validation sets

Class	Total Count	Training (70%)	Testing (15%)	Validation (15%)
Banded chlorosis	1600	1120	240	240
Grassy shoot	1600	1120	240	240
Healthy	2258	1581	338	339
Mosaic	2099	1469	315	315
Pokkah boeng	1600	1120	240	240
Red rot	1736	1215	261	260
Ring spot	2072	1450	311	311
Rust	1600	1120	240	240
Sett rot	1613	1129	242	242
Viral disease	1638	1147	245	246
Yellow leaf	3899	2729	585	585
Total	**21,715**	**15,201**	**3,257**	**3,257**

To improve variability and reduce overfitting, data augmentation methods such as rotation, flipping, scaling, and color jittering were employed. Temperature, humidity, wind speed, and disease severity were all synchronized hourly for analysis after the Kaggle climatic dataset was cleaned and aligned by interpolating missing values and eliminating unnecessary data [30]. Min-Max normalization was applied to continuous climatic variables to ensure model compatibility [31]. The image dataset was split into training (70%), testing (15%), and validation (15%) sets, while maintaining class proportions to avoid bias [32] (Table 1). Controlled augmentation and stratified sampling helped balance strategies and ensure stable learning across sugarcane disease classes.

3.4 Disease severity estimation model

3.4.1 Model architecture overview

The proposed disease assessment framework takes advantage of a dual-head CNN-Transformer architecture [33, 34] to effectively capture fine-grained lesion patterns and broader structural cues from the leaves of sugarcane. While the CNN backbone captures local texture variation such as the necrotic spot and chlorotic patch. The Transformer model captures the global relationships across the leaf surface. This combination increases the model's ability in recognizing early symptoms and distinguishing visually very similar diseases. Figure 3 gives the system-level block diagram showing the process for disease severity estimation.

3.4.2 CNN backbone

To detect the various types of lesions on sugarcane leaves and their colors, the CNN serves as a primary feature extraction method. The images have been resized to a 224×224×3-pixel resolution to ensure that all input images will have equal image quality. The CNN structure consists of five blocks of convolutional layers, which follow a standard sequencing of Conv2D, Batch Normalization, ReLU and

MaxPooling [35]. The backbone uses ResNet skip connections between layers to further help with the improvement of the gradient flow within the network. The network also uses both 3×3 and 5×5 convolutional kernels at different scales, making it easier to identify lesions of varying sizes and shapes.

The features collected at this local spatial level are enhanced, using a global context through the Transformer block module of the architecture. The parameters used in the CNN are given in Table 2.
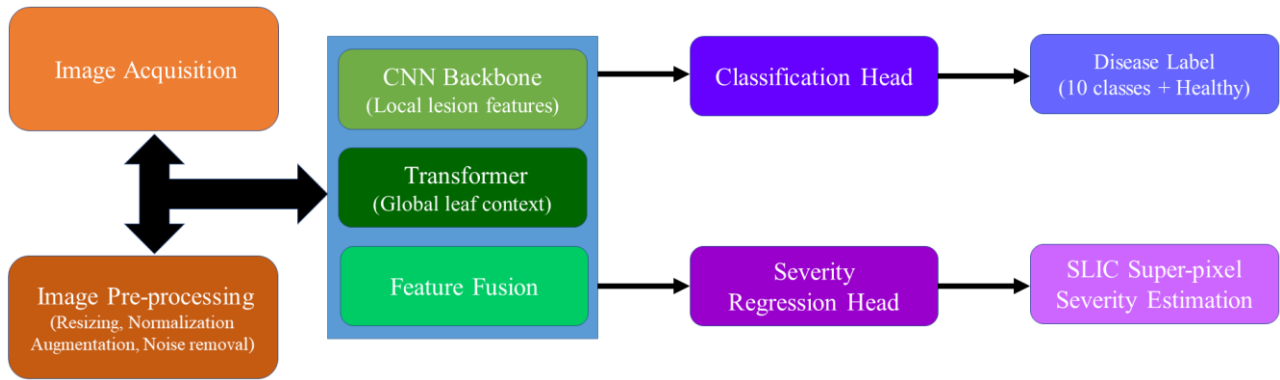

Figure 3. System level block diagram for disease severity estimation model

Table 2. CNN parameters

Parameter	Description / Value
Input Size	$224 \times 224 \times 3$
Pre-processing	Lighting normalization (CLAHE), Gaussian noise removal, resizing
Data Augmentation	Rotation, flipping, scaling, color jittering
Number of Convolutional Blocks	5
Block Structure	Conv2D $\rightarrow$ BatchNorm $\rightarrow$ ReLU $\rightarrow$ MaxPooling
Residual Connections	Enabled (to enhance gradient flow)
Kernel Sizes	Multi-scale: 3×3 and 5×5
Channel Dimensions	[64, 128, 256, 512, 512]
Feature Focus	Fine-grained local features, lesion and necrotic areas

3.4.3 Transformer module

The CNN architecture combined with the transformer module, models long-range spatial relationships across leaf surfaces, offering a wider understanding on lesions. Unlike conventional CNNs that capture lesions locally, the transformer addresses lesions distributed over larger areas or with overlapping characteristics. It segments lesions into patches and uses patch embeddings to create individual representations. Comprising four stacked encoder layers with multi-head self-attention (MHA) and feedforward sub-layers, the Transformer learns global interactions within lesions, enabling identification of infection stages and enhancing contextual understanding beyond local feature extraction.

The CNN model generates a feature map with multiple layers of depth, which is then divided into patches that do not overlap, and the patches are converted into tokens. After the tokenizing of the patches, each of them is processed through 4 layers of multi-head attention for global context modelling. The improved feature representation is then fused with the CNN feature presentation and passed into both the softmax classification output head and the regression output head that is used to determine the severity of the disease.

The Transformer uses the scaled dot-product attention mechanism [36], computed as given in Eq. (1)

$$Attention\ (Q, K, V) = softmax\left(\frac{QK}{\sqrt{d_k}}\right)V \qquad (1)$$

where, Q, K, and V represent the query, key, and value matrices generated from the patch embeddings, and d_k denotes the key dimensionality.

The multi-head attention (MHA) mechanism extends this by performing several attention operations in parallel as given in Eq. (2) allowing the network to simultaneously focus on different lesion features such as color variation, margin sharpness, and necrotic spread. Transformer parameters used in this work is shown in Table 3.

$$MHA\ (X) = Concat(h_1, \cdots\cdots, h_H)W_o \qquad (2)$$

Table 3. Transformer parameters

Component	Parameter / Value
CNN Feature Maps	Partitioned into 16×16 patches
Patch Embedding	512-dimensional vectors
Transformer Layers	4
Multi-Head Attention	8 heads
Feed-Forward Network	GELU activation
Transformer Feature Focus	Global leaf context, vein-aligned and elongated lesion patterns
Fusion Mechanism	Cross-attention
Output Integration	Fused features fed to dual-head prediction (classification + severity)

3.4.4 Feature fusion and dual-head output

A framework using both CNN and transformer features can improve recognition of lesions and provide an estimation of lesion discrimination. A lightweight cross attention system allows the global token from the transformer to focus the important CNN features [37], which allows subtle signs of the early stages of infection to be highlighted. The fused image is then passed through two heads. The classification head contains a softmax classifying 11 categories of diseases while

reducing overfitting through dropout; the regression head produces continuous values from 0 to 1 to indicating the severity of an affected leaf. The dual head structure allows the model to assess both disease diagnosis and severity, helping improve performance through two-way connectivity with agronomy in pesticide application decisions.

3.4.5 Model validation

Five-fold cross-validation was used for model validation on 21,637 images, guaranteeing balanced representation for precise performance metrics. Various architectures such as the ResNet50, DenseNet201, EfficientNet-B7, and ViT-B/16 was trained under consistent conditions and the performance was evaluated. The F1-scores showed clear and consistent improvements, confirming that the model's gains are reliable.

3.4.6 SLIC super-pixel severity estimation

For estimating disease severity, we employed SLIC super-pixel segmentation using 500 super-pixels and a compactness of 20 [12]. The purpose of this module is to generate a lightweight and interpretable measure of disease intensity without requiring pixel-level masks or heavy segmentation models. Compared with Watershed and K-Means, SLIC produced cleaner lesion boundaries and showed fewer failures under challenging conditions such as uneven illumination or background clutter.

SLIC was specifically chosen because it remains reliable under real field constraints such as variable lighting, overlapping leaves, dust, and complex scenes where conventional pixel-wise annotations or fully supervised segmentation networks become impractical to deploy. This makes SLIC a field-ready solution that supports robust severity estimation for downstream decision-making.

3.5 Environmental metadata integration

3.5.1 Climatic variables

The four climatic factors (temperature, relative humidity, wind speed and weather condition) were used as environmental metadata to support in determining the disease severity and to make the recommended spraying decisions. Previous studies [38, 39] have observed that these climatic factors influence the growth of pathogens, movement of spores, development of leaf wetness, and progression of pathogens.

3.5.2 Multimodal data fusion

The integration of climate data and sugarcane images (by linking location) allowed the analysis of disease-related factors to include the environment. This is important to know how temperature, humidity, and wind speed affect disease spread and efficacy of pesticides, are critical for generating a model for accurate spray recommendations. Spearman correlation is used to determine time-dependent relationships between disease severity and climate variables and therefore address the issue of nonlinear biological responses.

3.6 Multi-objective optimization for spray recommendation

To formalize the spray recommendation strategy, a multi-objective optimization (MOO) framework was developed that integrates agronomic constraints, approved chemical guidelines, and disease-yield relationships. The decision vector was structured following standard MOO formulations

[40], ensuring that each variable reflects a practical and controllable field parameter:

$$S = \{T, D, C\} \qquad (3)$$

where, T is the timing of spray after disease detection (hours), D is the dosage level (L ha⁻¹ or % solution), C is the chemical type (active ingredient/formulation).

The optimization problem now follows the canonical MOO form:

$$\min F(S) = [f_1(s), f_2(s), f_3(s)], s \in S \qquad (4)$$

This fixes earlier uncertainties and making sure that the method aligns with standard MOO procedures.

(1) Yield Protection: $f_1(s)$

Instead of using approximate measures, yield protection is estimated from a yield-loss curve based on the predicted disease severity (DS).

$$f_1(s) = 1 - Y(DS, D, T) \qquad (5)$$

where, $Y(DS, D, T)$ represents yield preserved under a specific dose timing strategy. This method matches the way disease severity and yield loss are related, as shown in plant pathology research [41].

(2) Treatment Efficiency: $f_2(s)$

Treatment efficiency represents suppression of lesion progression as mentioned in the previous works [42].

$$f_2(s) = E(D, T) \qquad (6)$$

where, $E(D, T)$ captures fungicide effectiveness as influenced by dosage and spray timing.

(3) Chemical Load: $f_3(s)$

Directly measurable and widely accepted metric to measure chemical load [43] has been used in this study.

$$f_3(s) = D \qquad (7)$$

This allows the optimization to minimize chemical input without relying on unvalidated ecological models. This practice aligns with previous NSGA-II agricultural optimization studies [22]. Yield protection serves as the primary agronomic goal, implicitly capturing the trade-off: higher doses (higher cost) must produce proportionally higher yield benefits [21].

3.6.1 Incorporation of approved chemical types and dosage ranges

For each disease predicted by the CNN-Transformer model, the system automatically maps to a valid treatment option defined by previous studies [44, 45]:

C: ICAR-recommended chemical

D: label-approved dosage

T: recommended timing window

This ensures that optimization operates strictly within realistic and legally compliant boundaries.

Table 4 presents all chemical types and dosage ranges used.

3.6.2 NSGA-II optimization and pareto-optimal outputs

The NSGA-II method has been used to find optimal trade-offs between yield protection, treatment effectiveness and chemical loads through an iterative process of non-dominated sorts and selecting spray schedules by crowding distance,

resulting in a set of Pareto optimal spray schedules. These spray schedules account for climate conditions, disease severity, amount of chemicals allowed, as well as their timing, therefore providing farmers with context relevant and practical spray schedules that provide maximum agronomic value with the least amount of chemicals needed. The workflow for the NSGA-II is given in Figure 4 and Table 5 gives the settings for the NSGA II.

Table 4. Disease and recommended chemical types and dosage ranges used

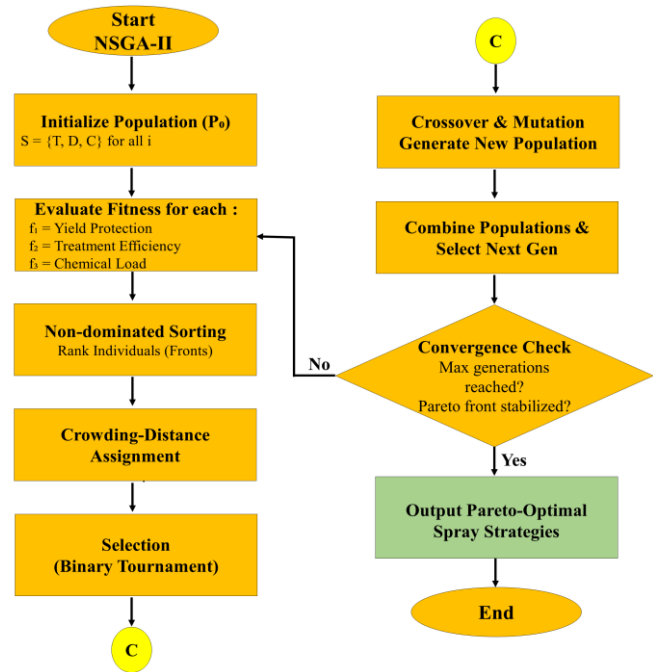| Disease	Recommended Chemical	Dosage Range Used in Model
Rust	Propiconazole 25 EC / Tebuconazole 25 EC	0.05-0.1% (or 250-500 mL/ha)
Pokkah Boeng	Carbendazim 50 WP / Propiconazole 25 EC	0.1% (Carbendazim) or 0.1-0.2% (Propiconazole)
Red Rot	Carbendazim 50 WP (prophylaxis)	1 g/L for sett treatment
Smut	Hot water treatment + fungicide dip (Carbendazim/Thiram)	0.1% fungicide dip
Mosaic, Yellow Leaf Virus	Vector control (Imidacloprid 17.8 SL / Thiamethoxam 25 WG)	0.3-0.5 mL/L (Imidacloprid)
Grassy Shoot	Vector insecticide (Imidacloprid/ Thiamethoxam)	0.3-0.5 mL/L
Ring Spot	Vector control	0.3-0.5 mL/L
Banded Chlorosis	Nutrient spray	0.5-1%
Sett Rot	Carbendazim	1 g/L sett treatment

Figure 4. NSGA-II workflow for generating spray recommendations

Table 5. NSGA II Configuration parameters

Parameter	Value
Population size	100
Number of generations	200
Crossover probability (Pc)	0.9
Mutation probability (Pm)	0.01
Distribution index for crossover (ηc)	20
Distribution index for mutation (ηm)	20
Selection method	Binary tournament

4. RESULTS AND DISCUSSION

4.1 Model performance and generalization

Figure 5 gives the training and validation accuracy for 100 epochs and demonstrate convergence that is stable and generalization capabilities that are favorable. The sharp fluctuations have been reduced by modifying the learning rate and adding more extensive data preprocessing; it is evident from the results achieved that there were final validation accuracies of 94-96% and a corresponding training accuracy of 99%.

The hybrid CNN-Transformer model was trained to improve the reliability of the model. The model can provide high classification accuracy across all 11 diseases as illustrated in the confusion matrix given in Figure 6. The primary reasons for misclassifications of diseases such as red rot and rust or grassy shoot and sett rot are based on visual characteristics that are similar for those particular types of plant disease. Nevertheless, there is evidence from prior research on CNN and transformer models that provide an affirmation to the hybrid model as a current diagnostic tool that can generalize effectively, perform highly reliably during field deployment.
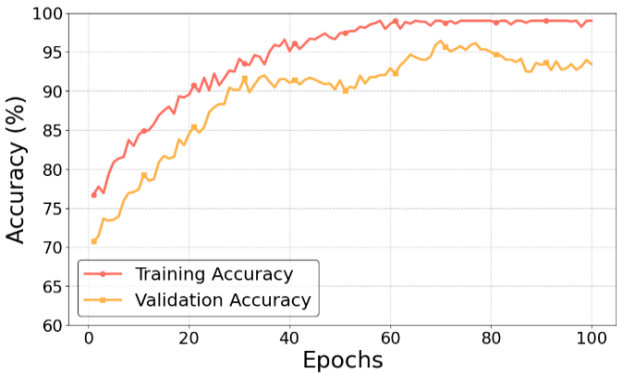

Figure 5. Training and validation curves

During the five-fold cross-validation process, the CNN-transformer model achieved a consistently high level of performance with overall accuracies ranging from 98.4% to 99.1%, producing an average accuracy of $98.8 \pm 0.27\%$. There was a high level of stability across the metrics of mean precision ($98.9 \pm 0.22\%$), mean recall ($98.6 \pm 0.28\%$) and mean F1-score ($98.7 \pm 0.27\%$), indicating that the model developed strong generalizable representation of multiple types of sugarcane disease symptoms through the architecture's stability and level of performance across multiple data sets as indicated by the results in Table 6.

The performance comparison in Figure 7 highlights the superior classification performance of the CNN-Transformer hybrid architecture compared to four common models (ResNet50, DenseNet201, EfficientNet-B7 and ViT-B/16). All of these baseline models produced high levels of average

accuracy (in the 92% to 95% range) and consistent performance through their respective precision, F1 and recall metrics. This is a result of a more pronounced use of

discriminative features from each backbone model in conjunction with global context from the Transformer encoder's attention weights.
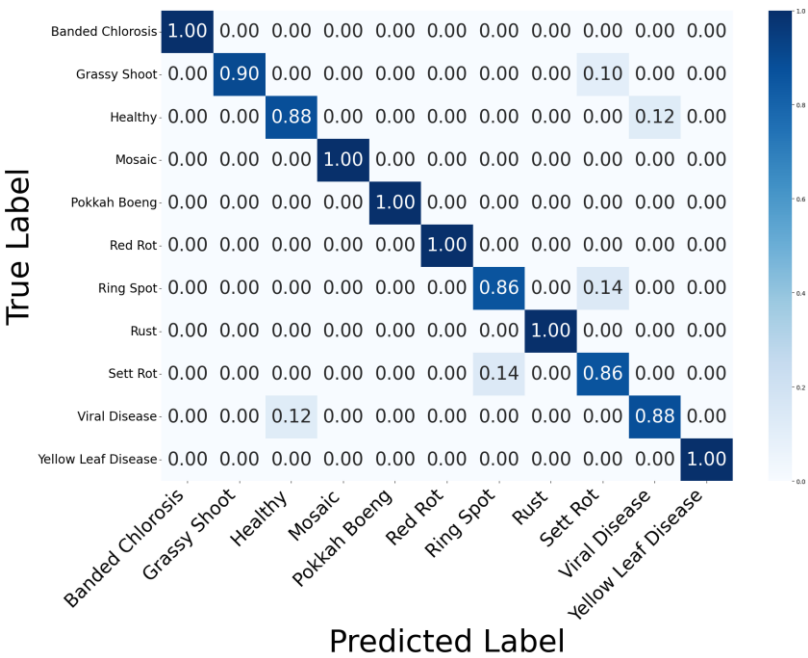

Figure 6. Confusion matrix - Sugarcane disease classification

Table 6. Five-fold cross-validation performance

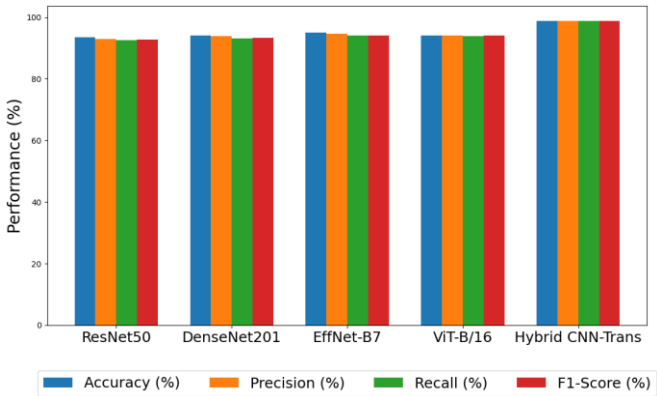| Fold	Accuracy (%)	Precision (%)	Recall (%)	F1-Score (%)
1	98.4	98.6	98.3	98.4
2	99.0	99.1	98.8	98.9
3	98.6	98.7	98.4	98.6
4	99.1	99.2	99.0	99.1
5	98.7	98.8	98.6	98.7
Mean ± SD	**98.8± 0.27**	**98.9± 0.22**	**98.6± 0.28**	**98.7 ± 0.27**

Figure 7. Comparison of performance results between the CNN-transformer and other CNN models

4.2 Ablation and significance testing

The results from the ablation study present in Table 7 indicate that a hybrid CNN-Transformer outperforms stand-alone models in disease classification and severity estimation. Specifically, the CNN model reached an accuracy of 89.5% with an F1-score of 0.86 and the Transformer model achieved an accuracy of 91.2% with a lower F1-score and greater MSE than the CNN model. Conversely, the hybrid method produced

an accuracy of 98.8%, an F1score of 0.987, and the lowest mean squared error (MSE) of 0.072. This demonstrates that the combination of CNNs' ability to capture local detail, and the global context provided by transformers, is highly effective.

Table 7. Ablation study of CNN, transformer, and hybrid CNN-transformer models for sugarcane disease classification and severity estimation

Model	Accuracy (%)	F1-score	MSE
CNN-only	89.5	0.86	0.092
Transformer-only	91.2	0.84	0.108
Hybrid CNN-Transformer	98.8	0.987	0.072

Table 8. Paired t-test results comparing mean F1-scores of the hybrid model with baselines

Comparison	Hybrid F1	Baseline F1	t-statistic	p-value
Hybrid vs ResNet50	98.7	92.8	7.12	< 0.01
Hybrid vs DenseNet201	98.7	93.4	6.94	< 0.01
Hybrid vs EfficientNet-B7	98.7	94.1	6.21	< 0.01
Hybrid vs ViT-B/16	98.7	94.0	6.33	< 0.01

To ensure the performance gains of the hybrid CNN-Transformer model's performance compared to the benchmarked models, statistical analysis was performed using paired t-tests comparing four benchmarked models against the hybrid CNN-Transformer as stated in Table 8. Based on the results of these analyses, it was determined that the hybrid architecture provides a statistical advantage and a verified increase in the hybrid model's F1-score of 98.7% when compared with the other models, which had F1-scores between 92.8%-94.1%. Additionally, t-statistics were produced that ranged from 6.21 to 7.12 with corresponding p-values below $p < 0.01$ confirming that the hybrid model provides a

statistically significant improvement compared with the other models. Therefore, it can be concluded that there are true and statistically significant advantages associated with this Hybrid CNN-Transformer architect compared to traditional deep CNNs and standalone Transformer's architectures.

Such tests confirm the statistically significant nature of observed performance gains as opposed to the effect of sampling noise or data partitioning.

4.3 Segmentation evaluation

Figure 8 represents performance measures of the SLIC super-pixel decompositions and subsequent clustering of 11 sugarcane disease classes. The SLIC grids closely align with the actual lesion boundaries, demonstrate good continuity along the edges and have captured sufficient structural information to reflect similar findings of earlier research on plant phenotyping with super-pixel methods [13]. The K-Means clustered masks of the diseased areas effectively mark the discoloured or necrotic regions, providing precise estimates of lesion area that may be used to estimate levels of severity. The streaky shapes of the red rot, rust and pokkah boeng disease lesions were correctly segmented and the patch-like characteristics associated with periodicity of virus-like diseases are well represented, while healthy samples were minimally misclassified.

The comparative results for pokkah boeng and red rot showcase the superior capability of the SLIC-based segmentation approach. The SLIC method consistently generated coherent, anatomically aligned super-pixels that accurately outlined lesion edges, allowing for reliable identification of symptomatic tissue. In contrast, K-Means, implemented on its own, generated noisy and spatially inconsistent clusters, while Watershed was seen to over-segment to a more severe degree due to its sensitivity to texture and illumination. These findings demonstrate that SLIC can maintain interpretability and robustness under field conditions. The two class examples presented herein Figure 9 are known to be representative of the full set of eleven analyzed disease classes for this study.

4.4 Multimodal effects

Figure 10 illustrates the classification of 11 different sugarcane disease types and shows that the majority of the diseases can be classified accurately, with healthy, yellow leaf, and mosaic diseases receiving the highest classification ratings. Grassy shoot and pokkah boeng received lower classification ratings because these two disease types are the least visually similar and thus difficult to differentiate by the computer model. The accompanying boxplot of the absolute severity estimation error indicates fairly consistent results across the different classes, as all classes have median error values less than 0.12. However, the high variance in the error rates for the classes red rot and viral diseases is likely caused by the typical patterns across lesions for both classes, which may hinder effective segmentation.
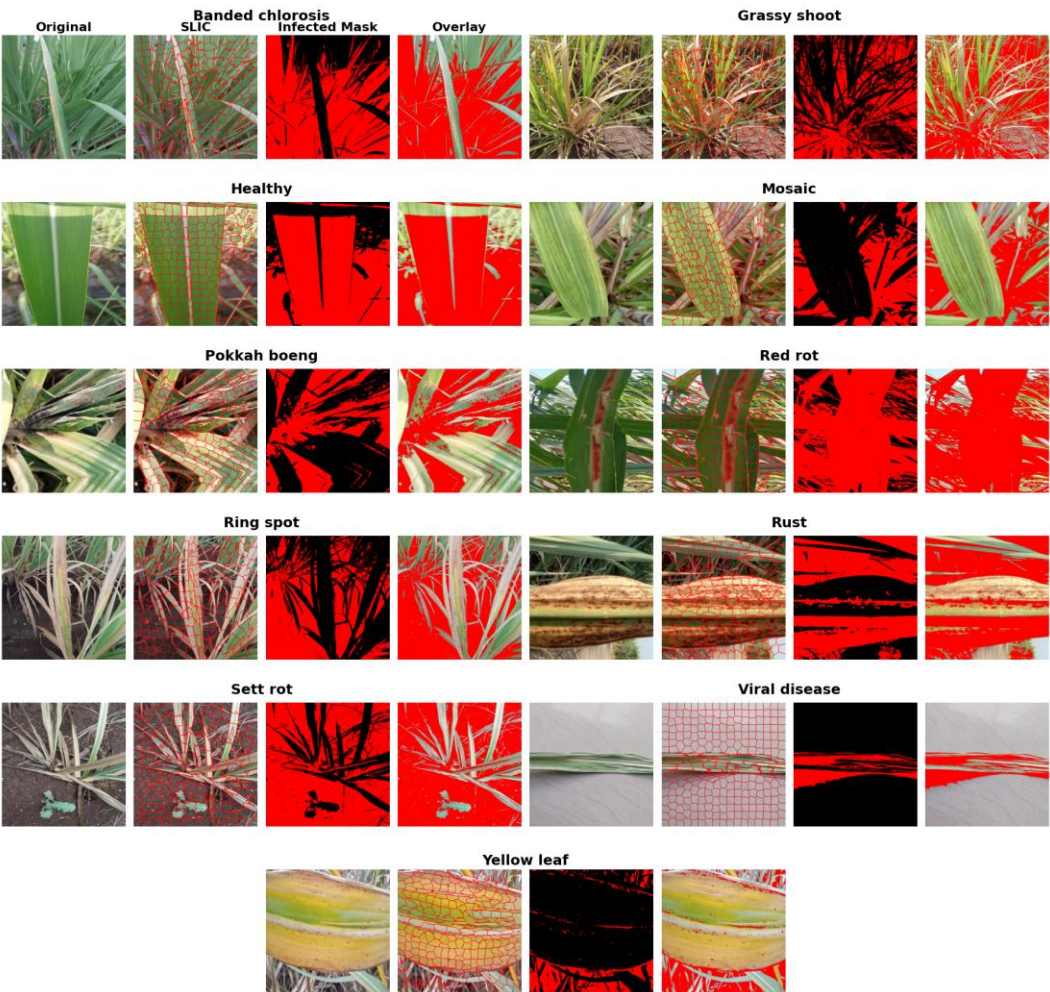

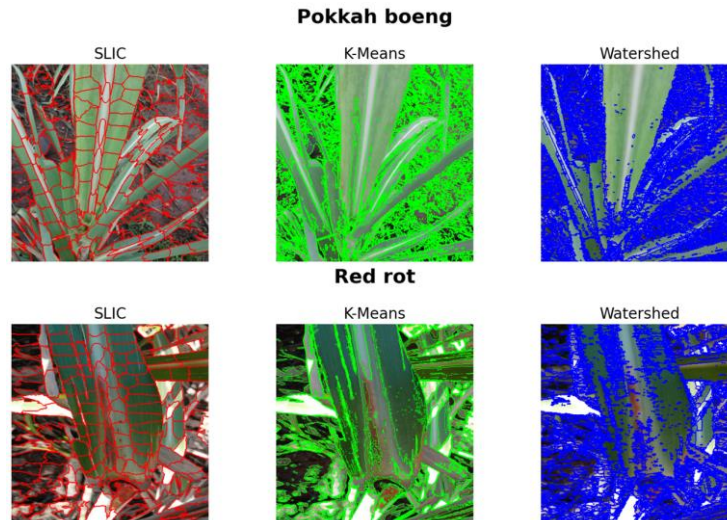**Figure 8.** SLIC super-pixel Segmentation of 11 sugarcane disease classes

Figure 9. Segmentation comparison between SLIC, K-Means and Water Shed
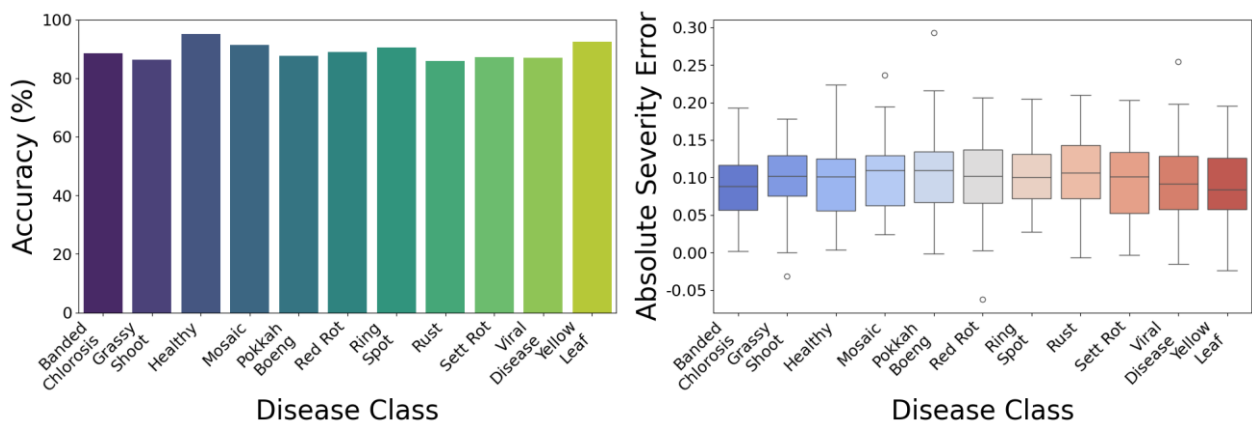

Figure 10. Classification accuracy per sugarcane disease and severity estimation error per disease class
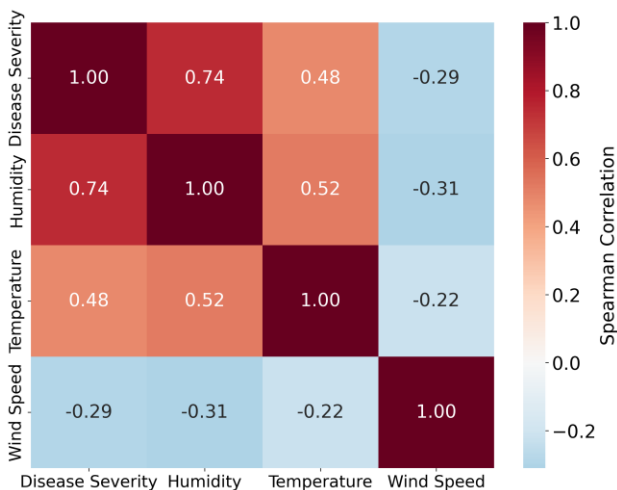

Figure 11. Correlation matrix between disease severity and weather parameters

The correlation analysis of disease severity and weather parameters shown in Figure 11, revealed a very strong positive relationship with relative humidity, temperature, and a weak negative relationship with wind speed. This supports established agronomic observations that high humidity and warm temperatures accelerate the progression of fungal diseases while wind speed can limit fungal establishment and spore retention on leaves.

As can be seen in Figure 11, there was a very strong correlation observed between the severity and weather parameters and weather. There was a very strong correlation between severity and relative humidity, temperature, and a weak negative correlation with wind speed, this is in agreement with agricultural research where high humidity and warm temperatures have been proven to accelerate the progression of fungal diseases and wind speed has been shown to negatively impact the establishment and retention of fungal spores on plant leaf surface. The analysis demonstrates the value of field-based monitoring for microclimatic conditions to high degree the importance of monitoring the microclimate of the field, as environmental conditions will influence both the disease severity and efficiency implementation of disease control strategies.

4.5 Optimization results

The NSGA-II algorithm has generated a three-dimensional Pareto front, in Figure 12, which is well distributed, and it indicates the trade-offs between yield preservation, treatment efficiency, and environmental risk. In accordance with the developed theory of multi-objective optimization [21, 22], there was no dominant solution within the range of solutions, which highlights the fact that agronomic and ecological goals should be considered in decisions regarding sprays. The

general trend in sustainable crop protection studies indicated that high-yield and high-efficiency solutions needed more chemical doses making them more hazardous to the environment. On the contrary, low-risk solutions accumulated on the low-dose region at the expense of yield preservation. The continuous and rounded-off Pareto surface and lack of discontinuities suggests that the optimization frame fully searched the decision space and generated agronomically representative and constant options. The interpretability of the 3D visualization is also enhanced because it demonstrates the response of each axis to the dosage intensity change thus directly responding to the reviewer questions about transparency and model rationality.
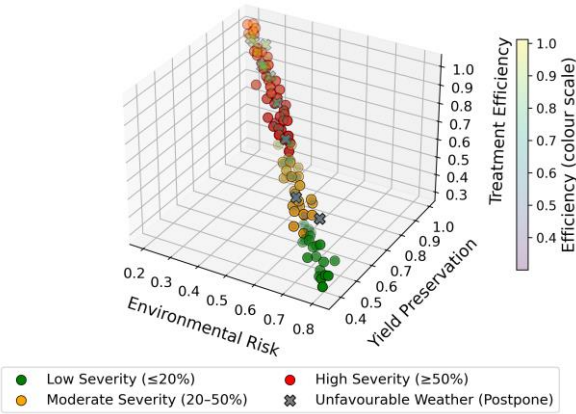
off between efficiency and environmental protection. High-efficiency, high-yield points were also favored in high-severity (> 50%), and therefore, high dose or immediate treatment was encouraged as the best option. Another scenario that contained severe weather constraints saw the system reject the Pareto surface and the decision was categorized as the one that was not appropriate and this showed appropriate integration of operational limits as well as the concerns of the reviewers regarding weather-aware decision-making. In general, this severity-based traversal of the Pareto front offers a systematic rationale of a quantitative nature of spray suggestions, with data-driven decision support being transparent and clear.
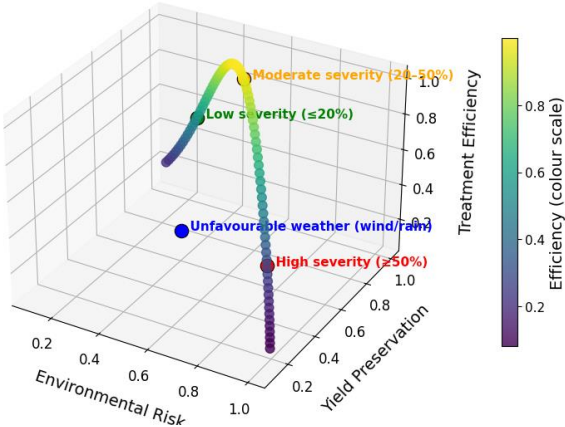

Figure 12. Severity-driven selection with NSGA-II

Figure 13. Disease severity with NSGA-II

Instead of using a fixed rule threshold, Figure 13 shows the dynamical nature of disease severity in determining the recommended strategy in the Pareto surface. The point that was selected changed to areas that best preserved yield and treated as severity escalated as an indication of the necessity of more hard-treatment [44]. When the severity was low (less than 20%), the system repeatedly tended to favour low-risk low-dose zone (which represented preventive or delayed spraying policies). In the case of moderate severity (20-50%), the optimum location was at the midpoint Pareto surface, with moderate-dose spraying providing the most favorable trade-

4.6 Deployment output

Based on the optimization defined in Section 4.5, a web-based dashboard was created to convert the outputs of the model into useful decision support. Although the optimization module creates Pareto-optimal dose-timing strategies that are modified according to the severity of the disease, the conditions of the weather, and the limitations of the risk, the dashboard allows an easy interface to comprehend the results (Figure 14).
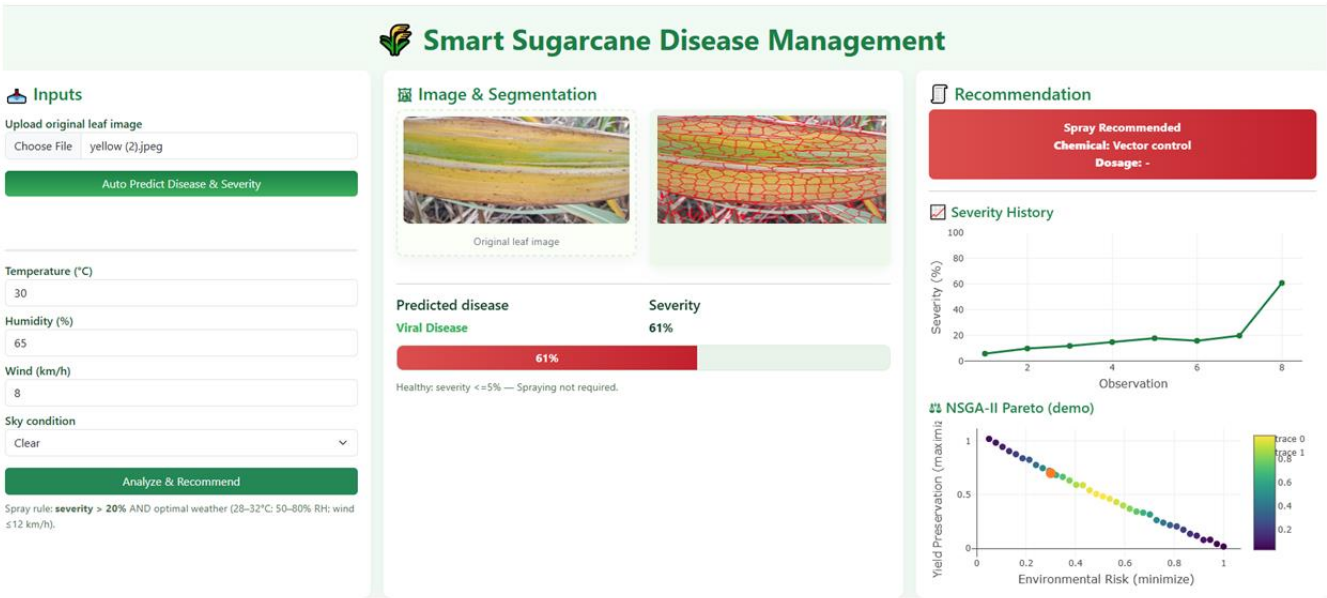

Figure 14. Web-based dashboard implementing the proposed framework

The user is able to input a leaf image and a set of environmental parameters and the system will show four important outputs: (i) the predicted disease type, (ii) severity estimate, (iii) weather suitability, and (iv) the recommended spraying action based on the results of NSGA-II. The dashboard is not a change in the optimization logic, but it is merely a display of the calculated strategies in the form of actionable labels, e.g., low-dose delayed spraying, moderate-dose spraying, and high-dose immediate spraying.

Such implementation shows the operational preparedness of the framework and proves it to be practical application of the framework in decision support in the real world.

5. CONCLUSION

This paper presents a framework that provides an integration of image-based disease detection, severity estimation, climate metadata, and multi-objective optimization to protect sugarcane crop. The suggested CNN-Transformer architecture allows the use of fine-grained lesion textures and global context of the whole leaf to improve the recognition performance, which was more accurate than traditional techniques. To calculate severity, the framework uses an effective SLIC based segmentation model, which is suitable in real field scenarios where there is limited access to computational resources.

The environmental parameters such as temperature, humidity, wind speed, and the general weather conditions are incorporated into the system which allows climate conscious measurements that will guide the effective and safer use of pesticides. The NSGA-II module also assists in decision-making by producing Pareto-optimal spray programs that strike a balance between the application of chemicals and the effectiveness of the treatment and follow realistic agronomic requirements.

In general, the framework represents a realistic way of achieving smart and sustainable disease management. In the future, the system will be expanded to include soil parameters, temporal disease modeling, and autonomous drone-powered spraying in order to increase the scale and predictive potential of the system when used in practice.

REFERENCES

[1] FAO. (2022). World Food and Agriculture-Statistical Yearbook 2022. Rome. https://doi.org/10.4060/cc2211en.

[2] Solomon, S. (2014). Sugarcane agriculture and sugar industry in India: At a glance. Sugar Tech, 16(2): 113-124. https://doi.org/10.1007/s12355-014-0303-8

[3] Mehdi, F., Cao, Z., Zhang, S., Gan, Y., Cai, W., Peng, L., Wu, Y., Wang, W., Yang, B. (2024). Factors affecting the production of sugarcane yield and sucrose accumulation: Suggested potential biological solutions. Frontiers in Plant Science, 15: 1374228. https://doi.org/10.3389/fpls.2024.1374228

[4] Abbas, A., Zhang, Z., Zheng, H., Alami, M.M., Alrefaei, A.F., Abbas, Q., Naqvi, S.A.H., Rao, M.J., Mosa, W.F.A.E.G., Hussain, A., Hassan, M., Zhou, L. (2023). Drones in plant disease assessment, efficient monitoring, and detection: A way forward to smart agriculture. Agronomy, 13(6): 1524. https://doi.org/10.3390/agronomy13061524

[5] Shigueoka, M.Y., Cavichioli, F.A. (2024). Agricultura de precisão na otimização de recursos: An analysis of the use of advanced technologies to maximize agricultural efficiency. Revista Interface Tecnológica, 21(2): 578-587. https://doi.org/10.31510/infa.v21i2.2118

[6] Sharma, P., Berwal, Y.P.S., Ghai, W. (2020). Performance analysis of deep learning CNN models for disease detection in plants using image segmentation. Information Processing in Agriculture, 7(4): 566-574. https://doi.org/10.1016/j.inpa.2019.11.001

[7] Upadhye, S.A., Dhanvijay, M.R., Patil, S.M. (2023). Sugarcane disease detection using CNN-deep learning method: An Indian perspective. Universal Journal of Agricultural Research, 11(1): 80-97. https://doi.org/10.13189/ujar.2023.110108

[8] Srinivasan, S., Prabin, S.M., Mathivanan, S.K., Rajadurai, H., Kulandaivelu, S., Shah, M.A. (2025). Sugarcane leaf disease classification using deep neural network approach. BMC Plant Biology, 25(1): 282. https://doi.org/10.1186/s12870-025-06289-0

[9] Singh, A.K., Rao, A., Chattopadhyay, P., Maurya, R., Singh, L. (2024). Effective plant disease diagnosis using vision transformer trained with leafy-generative adversarial network-generated images. Expert Systems with Applications, 254: 124387. https://doi.org/10.1016/j.eswa.2024.124387

[10] Aswani, S., Sinha, A. (2025). Early diagnosis of Sugarcane leaf diseases through CNN and vision transformer hybrid model. Research Square. https://doi.org/10.21203/rs.3.rs-7232200/v1

[11] Prommakhot, A., Onshaunjit, J., Ooppakaew, W., Samseemoung, G., Srinonchat, J. (2025). Hybrid CNN and transformer-based sequential learning techniques for plant disease classification, IEEE Access, 13: 122876-122887. https://doi.org/10.1109/ACCESS.2025.3586285

[12] Zhang, S., Wang, H., Huang, W., You, Z. (2018). Plant diseased leaf segmentation and recognition by fusion of superpixel, K-means and PHOG. Optik, 157: 866-872. https://doi.org/10.1016/j.ijleo.2017.11.190

[13] Achanta, R., Shaji, A., Smith, K., Lucchi, A., Fua, P., Süsstrunk, S. (2012). SLIC superpixels compared to state-of-the-art superpixel methods. IEEE Transactions on Pattern Analysis and Machine Intelligence, 34(11): 2274-2282. https://doi.org/10.1109/TPAMI.2012.120

[14] Dinesh, P., Lakshmanan, R. (2025). Multiclass semantic segmentation for prime disease detection with severity level identification in Citrus plant leaves. Scientific Reports, 15(1): 21208. https://doi.org/10.1038/s41598-025-04758-y

[15] Kumar, D., Kukreja, V. (2025). Impact of image segmentation and feature sets in automated plant disease classification: A comprehensive review based on wheat plant images. Progress in Artificial Intelligence, 14: 451-504. https://doi.org/10.1007/s13748-025-00389-6

[16] Lei, L., Yang, Q., Yang, L., Shen, T., Wang, R., Fu, C. (2024). Deep learning implementation of image segmentation in agricultural applications: A comprehensive review. Artificial Intelligence Review, 57(6): 149. https://doi.org/10.1007/s10462-024-10775-6

[17] Pradhan, S., Hore, S., Maji, S.K., Manna, S., Maity, A., Kundu, P.K., Maity, K., Roy, S., Mitra, S., Dam, P., Mondal, R., Ghorai, S., Jawed, J.J., Dutta, S., Das, S.,

Mandal, S., Mandal, S., Kati, A., Sinha, S., Maity, A.B., Dolai, T.K., Mandal, A.K., İnce, İ.A. (2022). Study of epidemiological behaviour of malaria and its control in the Purulia district of West Bengal, India (2016-2020). Scientific Reports, 12(1): 630. https://doi.org/10.1038/s41598-021-04399-x

[18] Chaulagain, B., Small, I.M., Shine Jr, J.M., Raid, R.N., Rott, P. (2021). Predictive modeling of brown rust of sugarcane based on temperature and relative humidity in Florida. Phytopathology, 111(8): 1401-1409. https://doi.org/10.1094/PHYTO-02-20-0060-R

[19] Sunarto, D.A., Hidayah, N., Wijayanti, K.S., Riajaya, P.D., Yulianti, T. (2023). Climate change impacts on sugarcane smut disease and its management. Agriculture and Natural Resources, 57(4): 647-656. https://doi.org/10.34044/j.anres.2023.57.4.09

[20] Wakiluzzaman, S., Siddique, M., Bari, K. (2025). Impact of climatic variables on sugarcane stem borer (Scirpophaga excerptalis) infestation in Thakurgaon District, Bangladesh. International Journal of Science and Research Archive, 15(1): 1180-1185. https://doi.org/10.30574/ijsra.2025.15.1.1035

[21] Deb, K., Pratap, A., Agarwal, S., Meyarivan, T.A.M.T. (2002). A fast and elitist multiobjective genetic algorithm: NSGA-II. IEEE Transactions on Evolutionary Computation, 6(2): 182-197. https://doi.org/10.1109/4235.996017

[22] Huang, Z., Pu, Y., Zhang, Q., Wang, Y., Yang, J., Gu, Y. (2025). A fuzzy-expert enhanced NSGA-II approach for sustainable agricultural systems. Scientific Reports, 15(1): 1-18. https://doi.org/10.1038/s41598-025-14488-w

[23] Shikalgar, A., Jadhav, S., Surve, A., Bamane, S., Kamble, A. (2024). Sugarcane Disease Dataset. Mendeley Data, v2. https://doi.org/10.17632/7fbnxcbnhp.2

[24] Daphal, S., Koli, S. (2022). Sugarcane Leaf Disease Dataset. Mendeley Data, v1. https://doi.org/10.17632/9424skmnrk.1

[25] Agarwala, A.G., Rahman, M., Zaman, G.S. (2024). Sugarcane Plant Disease Dataset of Assam. Mendeley Data, v1. https://doi.org/10.17632/x98kpjvxxg.1

[26] Anand, A.K.J. (2025). S2RMCMD. Mendeley Data, v1. https://doi.org/10.17632/rt8sv9t445.1

[27] Thite, S., Suryawanshi, Y., Patil, K., Chumchu, P. (2023). Sugarcane Leaf Dataset. Mendeley Data, v1. https://doi.org/10.17632/355y629ynj.1

[28] Singh, V. (2022). India's State and District-Wise Weather Data. Kaggle. https://www.kaggle.com/datasets/viveksingh2400/indias-state-and-district-wise-weather-data.

[29] Ramtekkar, P.K., Pandey, A., Pawar, M.K. (2024). A comprehensive review of brain tumour detection mechanisms. The Computer Journal, 67(3): 1126-1152. https://doi.org/10.1093/comjnl/bxad047

[30] AlSalehy, A.S., Bailey, M. (2025). Improving time series data quality: Identifying outliers and handling missing values in a multilocation gas and weather dataset. Smart Cities, 8(3): 82. https://doi.org/10.3390/smartcities8030082

[31] Shantal, M., Othman, Z., Bakar, A.A. (2023). A novel approach for data feature weighting using correlation coefficients and min-max normalization. Symmetry, 15(12): 2185. https://doi.org/10.3390/sym15122185

[32] Sambasivam, G.A.O.G.D., Opiyo, G.D. (2021). A predictive machine learning application in agriculture: Cassava disease detection and classification with imbalanced dataset using convolutional neural networks. Egyptian Informatics Journal, 22(1): 27-34. https://doi.org/10.1016/j.eij.2020.02.007

[33] Guerri, M.F., Distante, C., Spagnolo, P., Taleb-Ahmed, A. (2025). Boosting hyperspectral image classification with gate-shift-fuse mechanisms in a novel CNN-transformer approach. Computers and Electronics in Agriculture, 237: 110489. https://doi.org/10.1016/j.compag.2025.110489

[34] Daimary, D., Bora, M.B., Amitab, K., Kandar, D. (2020). Brain tumor segmentation from MRI images using hybrid convolutional neural networks. Procedia Computer Science, 167: 2419-2428. https://doi.org/10.1016/j.procs.2020.03.295

[35] Khan, A., Rauf, Z., Sohail, A., Khan, A.R., Asif, H., Asif, A., Farooq, U. (2023). A survey of the vision transformers and their CNN-transformer based variants. Artificial Intelligence Review, 56(Suppl 3): 2917-2970. https://doi.org/10.1007/s10462-023-10595-0

[36] Venkatachalam, M., Gérard, L., Milhau, C., Vinale, F., Dufossé, L., Fouillaud, M. (2019). Salinity and temperature influence growth and pigment production in the marine-derived fungal strain Talaromyces albobiverticillius 30548. Microorganisms, 7(1): 10. https://doi.org/10.3390/microorganisms7010010

[37] Zhang, Z., Zhao, P., Zheng, Z., Luo, W., Cheng, B., Wang, S., Huang, Y., Zheng, Z. (2025). RT-MWDT: A lightweight real-time transformer with edge-driven multiscale fusion for precisely detecting weeds in complex cornfield environments. Computers and Electronics in Agriculture, 239: 110923. https://doi.org/10.1016/j.compag.2025.110923

[38] Li, B., Zhao, H., Li, B., Xu, X.M. (2003). Effects of temperature, relative humidity and duration of wetness period on germination and infection by conidia of the pear scab pathogen (Venturia nashicola). Plant Pathology, 52(5): 546-552. https://doi.org/10.1046/j.1365-3059.2003.00887.x

[39] Coello, C.A.C. (2011). Evolutionary multiobjective optimization. Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery, 1(5): 444-447. https://doi.org/10.1002/widm.43

[40] Rizal, S., Saha, P. (2025). Deciphering inter-relationship between disease incidence and disease severity and prediction of yield loss in lentil-stemphylium pathosystem. Archives of Phytopathology and Plant Protection, 58(13): 725-741. https://doi.org/10.1080/03235408.2025.2530793

[41] Bhuiyan, S.A., Stringer, J.K., Croft, B.J., Olayemi, M.E. (2022). Resistance of sugarcane varieties to smut (sporisorium scitamineum), development over crop classes, and impact on yield. Crop and Pasture Science, 73(10): 1180-1187. https://doi.org/10.1071/CP21607

[42] Chaulagain, B., Raid, R.N., Dufault, N., van Santen, E., Rott, P. (2019). Application timing of fungicides for the management of sugarcane orange rust. Crop Protection, 119: 141-146. https://doi.org/10.1016/j.cropro.2019.01.007

[43] Bhuiyan, S.A., Croft, B.J., Tucker, G.R. (2015). New method of controlling sugarcane smut using flutriafol fungicide. Plant Disease, 99(10): 1367-1373. https://doi.org/10.1094/PDIS-07-14-0768-RE

[44] Tamil Nadu Agricultural University (TNAU). (2023). Crop Production Guide Agriculture 2020. https://agritech.tnau.ac.in/pdf/AGRICULTURE.pdf.

[45] Indian Council of Agricultural Research-Sugarcane Breeding Institute (ICAR-SBI). (2022). Sugarcane plant protection manual. https://www.icar.org.in/sites/default/files/inline-files/ICAR-SBIVision2050.pdf.