



A Hybrid Framework Combining U-Net, Ant Colony Optimization, and CNN for Rice Leaf Disease Classification under Class Imbalance

Erianto Ongko¹, Asmah Indrawati², Sukiman¹

¹Department of Computer Science, Institut Modern Arsitektur dan Teknologi, Medan, Indonesia

²Department of Agrotechnology, Faculty of Agriculture, Universitas Medan Area, Medan, Indonesia

¹eriantoongko9@gmail.com, ²asmahindrawati@staff.uma.ac.id, ³sukiman.liu@gmail.com

Abstract

Accurate classification of rice plant diseases is essential for early intervention and precision agriculture. However, real-world datasets often suffer from complex backgrounds, high-dimensional features, and severe class imbalances, which compromise classification performance. This study proposes an integrated framework combining image segmentation using U-Net, feature selection via Ant Colony Optimization (ACO), hybrid sampling to handle class imbalance, and final classification using a Convolutional Neural Network (CNN). Segmentation isolates disease-affected areas, ACO optimizes feature subsets, and hybrid sampling balances class distribution using undersampling and SMOTE. The proposed method was tested on four rice leaf disease datasets—Brown Spot, Leaf Blast, Leaf Blight, and Leaf Scald—exhibiting significant class imbalance. Experimental results show that the proposed approach outperforms baseline models (SegNet, PspNet, and E-Net) across multiple metrics: Accuracy, IoU, Precision, and Recall. This indicates the framework's robustness and potential for real-world deployment in precision agriculture. Future work will focus on model compression and real-time implementation in IoT systems.

Keywords: class imbalance; feature selection; hybrid sampling; image segmentation; precision agriculture

How to Cite: E. Ongko, A. Indrawati, and S. Sukiman, "A Hybrid Framework Combining U-Net, Ant Colony Optimization, and CNN for Rice Leaf Disease Classification under Class Imbalance", *J. RESTI (Rekayasa Sist. Teknol. Inf.)*, vol. 9, no. 6, pp. 1474 - 1484, Jan. 2026.
Permalink/DOI: <https://doi.org/10.29207/resti.v9i6.6910>

Received: July 7, 2025

Accepted: December 16, 2025

Available Online: January 11, 2025

*This is an open-access article under the CC BY 4.0 License
Published by Ikatan Ahli Informatika Indonesia*

1. Introduction

Accurate classification of rice plant diseases is vital for sustainable agricultural practices, early intervention, and food security[1]. However, real-world agricultural image datasets often suffer from multiple challenges: complex backgrounds[2], high-dimensional feature representations[3], and severe class imbalance[4]. These challenges can significantly degrade the performance of conventional classification models by introducing noise[5], overfitting[6], and class bias[7]. In this context, earlier studies have attempted to address these challenges through individual methods such as CNN-based classifiers[8], segmentation-only approaches[9], or imbalance handling[10], but these approaches often fail when applied to complex real-world agricultural datasets[11]. While previous feature selection approaches such as filter or PCA-based methods improved computational efficiency, they did not fully address the interpretability and inter-feature dependency challenges[12]. To address these issues,

this research proposes a comprehensive and integrated framework that combines four major components: image segmentation using U-Net, feature selection using Ant Colony Optimization (ACO), class imbalance handling through hybrid sampling, and final classification through a Convolutional Neural Network (CNN). Each method was deliberately chosen to address specific weaknesses commonly found in plant disease detection tasks.

The first crucial step in the proposed framework is image segmentation, which serves to isolate the disease-affected regions of the leaf from irrelevant backgrounds such as soil, shadows, or overlapping objects[13]. Without segmentation, a classifier may learn patterns based on background noise or unrelated leaf structures, reducing its generalization ability[14]. Traditional image classification pipelines that directly process raw images are prone to misclassification due to these unwanted features[15]. Segmenting the images ensures that only the meaningful regions are used for

feature extraction and classification[16]. Among the many segmentation techniques available, the U-Net architecture is selected due to its ability to capture both fine-grained details and contextual information through its encoder-decoder structure and skip connections[17]. U-Net has demonstrated superior performance in biomedical and agricultural domains compared to basic CNNs[18], FCNs[19], or SegNet[20], particularly in scenarios with limited training data[21]. Its architecture allows for precise localization of disease spots, making it ideal for enhancing feature quality in the next stage[22]. Thus, segmentation forms the foundation of many existing works, though often applied without additional optimization techniques[23].

Following segmentation, feature selection becomes necessary to reduce dimensionality and eliminate irrelevant or redundant information[24]. High-dimensional feature spaces, which often result from segmentation and deep feature extraction, can lead to the "curse of dimensionality"[25], where classifiers struggle to distinguish relevant patterns and may overfit to noise[26]. Moreover, processing unnecessary features increases computational cost and training time[27]. While previous feature selection approaches such as filter or PCA-based methods improved computational efficiency, they did not fully address the interpretability and inter-feature dependency challenges[28]. To overcome these issues, Ant Colony Optimization (ACO) is used as a metaheuristic feature selection technique[29]. Unlike statistical or filter-based methods like Chi-Square or Mutual Information that consider features individually[30], ACO evaluates feature combinations and interactions[31]. Compared to PCA, which transforms features into new orthogonal components and may reduce interpretability, ACO retains the original feature semantics[32]. Additionally, ACO offers better global search capabilities than greedy methods, which are prone to local optima[33]. By simulating the natural foraging behavior of ants and updating pheromone trails based on solution quality, ACO efficiently identifies the most informative feature subset, improving both classifier accuracy and efficiency[34].

To further address the issue of class imbalance, which is common in plant disease datasets where some diseases are underrepresented[35], the study implements a hybrid sampling approach[36]. Comparable works on class imbalance handling mainly relied on oversampling or undersampling alone, leading to potential bias or data loss[37]. Our hybrid approach aims to overcome these weaknesses by integrating the strengths of both. This method combines random undersampling of the majority class with SMOTE-based oversampling of the minority class[38]. Unlike pure oversampling, which may duplicate samples and lead to overfitting, or pure undersampling, which risks discarding important data, the hybrid method balances the class distribution while preserving data diversity. This results in a more robust and fair classification model[39].

Finally, the preprocessed, segmented, and balanced data are passed into a CNN-based classifier, which benefits from enhanced input quality. The integration of these four components—U-Net for accurate segmentation, ACO for optimal feature selection, hybrid sampling for balanced learning, and CNN for robust classification—forms the core novelty of this research. Few existing studies combine all these techniques in a unified pipeline tailored specifically for rice plant disease classification. The novelty of this work is not merely in applying each technique individually, but in unifying them into a comprehensive framework tailored for rice leaf disease classification. This contribution is systematically validated in later sections through comparative experiments. The remainder of this article is organized as: Section 2 describes the methodology, Section 3 presents results and discussions, and Section 4 concludes with findings and directions for future work.

In summary, this work highlights the importance of preprocessing and intelligent data preparation in enhancing model performance. The careful integration of segmentation, selection, and balancing techniques makes the proposed framework not only more accurate but also more interpretable and suitable for deployment in real-world precision agriculture systems.

2. Methods

An overview of the research process is presented in Figure 1. Figure 1 illustrates the overall workflow of the proposed framework for rice leaf disease classification, combining established procedures with targeted enhancements to address specific dataset challenges, including class imbalance, high-dimensional features, and complex backgrounds.

The process begins with image preprocessing, where all images were normalized to a 0–1 pixel intensity range and resized to 224×224 pixels to ensure consistency during training. Classification was performed using a Convolutional Neural Network (CNN) composed of convolutional layers with ReLU activation to extract discriminative features, max pooling layers to reduce feature map dimensions and limit overfitting, fully connected dense layers to integrate the learned representations, and a final softmax output layer to generate probability distributions across the disease classes. Model performance was evaluated using widely recognized metrics—Accuracy, Precision, Recall, and Intersection over Union (IoU)—to provide a reliable benchmark for prediction quality.

To strengthen this framework, several additional techniques were integrated. An improved U-Net model was employed for image segmentation, enabling precise isolation of disease-affected regions and reducing background interference. The extracted feature vectors were refined using Ant Colony Optimization (ACO), which efficiently identified the most informative subsets by performing a global search based on cooperative swarm intelligence. Most importantly, to

address the severe class imbalance present in rice disease datasets—where some classes dominate with more than 80% of samples—a hybrid sampling strategy was implemented. This combined random

undersampling of majority classes with SMOTE-based oversampling of minority classes, ensuring balanced training distributions while minimizing both overfitting and information loss.

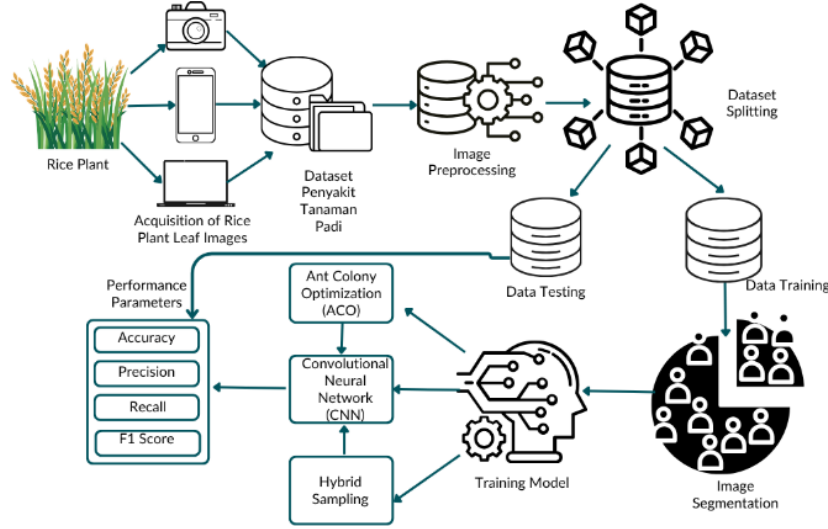


Figure 1. Research Process

2.1 Image Segmentation using an Improved U-Net Model

Image segmentation is the task of assigning a semantic label to every pixel in an image, enabling the identification of regions or objects of interest. In this study, image segmentation is performed using an improved U-Net model, a deep convolutional neural network designed for dense, pixel-wise prediction. The model follows an encoder-decoder architecture with skip connections, which allows the network to balance high-level semantic understanding with precise spatial localization[40].

The segmentation process starts with an input image $I \in \mathcal{H}^{H \times W \times C}$, where H , W , and C denote image height, width, and number of channels, respectively. The input image is first normalized to reduce intensity variability and improve numerical stability. This normalization can be expressed as Equation 1.

$$I_{norm} = \frac{I - \mu}{\sigma} \quad (1)$$

μ and σ represent the mean and standard deviation of pixel intensities. The normalized image is then resized to a fixed resolution to ensure compatibility with the network input.

In the encoder stage, hierarchical feature extraction is performed through a sequence of convolutional layers. At each encoder level l , feature maps are computed as Equation 2.

$$X_l = \text{ReLU}(W_l * X_{l-1} + b_l) \quad (2)$$

W_l and b_l denote the convolution kernel and bias at level l , $*$ represents the convolution operation, and ReLU is the Rectified Linear Unit activation function defined as Equation 3.

$$\text{ReLU}(x) = \max(0, x) \quad (3)$$

To capture contextual information at increasing scales, spatial downsampling is applied using max-pooling, given by Equation 4.

$$X_l^\downarrow = \text{MaxPool}(X_l) \quad (4)$$

The resulting feature maps at each encoder level are stored and later reused as skip connections to preserve spatial detail.

At the bottleneck layer, the deepest feature representation encodes high-level semantic information. This representation is further refined using convolution and non-linear activation to strengthen global feature abstraction before reconstruction.

The decoder stage reconstructs the segmentation map by progressively increasing spatial resolution. Upsampling operations are applied to the bottleneck and subsequent feature maps, which can be expressed as Equation 5.

$$X_l^\uparrow = \text{UpSample}(X_{l+1}) \quad (5)$$

At each decoder level, the upsampled feature maps are concatenated with the corresponding encoder feature maps through skip connections using Equation 6.

$$X_l^{\text{concat}} = \text{concat}(X_l^\uparrow, E_l) \quad (6)$$

E_l denotes the encoder feature map at level l . This concatenation allows the model to combine fine-grained spatial information with high-level semantic context. The merged features are then refined using convolution followed by ReLU activation.

Finally, the output segmentation map is generated using a convolutional layer that projects the refined features into class scores using Equation 7.

$$S = W_{out} * X + b_{out} \quad (7)$$

For binary segmentation tasks, a sigmoid activation function is applied to obtain pixel-wise probabilities using Equation 8.

$$P = \sigma(S) = \frac{1}{1+e^{-S}} \quad (8)$$

For multi-class segmentation, a softmax function is used instead, as shown in Equation 9.

$$P_i = \frac{e^{s_i}}{\sum_{j=1}^K e^{s_j}} \quad (9)$$

K is the number of segmentation classes. The final segmentation mask is obtained by thresholding the sigmoid output in the binary case or selecting the class with the highest probability for each pixel in the multi-class case.

2.2 Feature Selection Using Ant Colony Optimization

Feature selection aims to identify an optimal subset of features that maximizes model performance while reducing redundancy and computational complexity. In this study, feature selection is performed using Ant Colony Optimization (ACO), a population-based metaheuristic inspired by the foraging behavior of ants. In ACO, each candidate feature subset is modeled as a path in a graph, and artificial ants collaboratively explore this graph to discover the most informative combination of features[41].

The feature selection problem is represented as a graph $G = (V, E)$, where each node $v_i \in V$ corresponds to a feature and each edge $(i, j) \in E$ represents a possible transition between features. The algorithm begins by initializing a pheromone value τ_{ij} on each edge, which reflects the desirability of selecting feature j after feature i . Several control parameters are defined, including the pheromone influence factor α , heuristic influence factor β , pheromone evaporation rate p , and pheromone deposit constant Q . The best feature subset P_{best} and its corresponding cost L_{best} are also initialized.

The optimization process proceeds iteratively over a fixed number of iterations. In each iteration, multiple ants independently construct candidate feature subsets. Each ant starts from an initial node and incrementally builds a solution by selecting the next feature based on a probabilistic decision rule. The probability of an ant selecting feature j from feature i is given by Equation 10.

$$p_{ij} = \frac{(\tau_{ij})^\alpha (\eta_{ij})^\beta}{\sum_{k \in N_i} (\tau_{ik})^\alpha (\eta_{ik})^\beta} \quad (10)$$

η_{ij} denotes the heuristic desirability of feature j , often derived from statistical relevance measures or classifier-based feature importance, and N_i is the set of feasible next features. This probabilistic mechanism balances exploration of new feature combinations and exploitation of previously successful ones.

Each ant continues adding features until a stopping condition is met, such as reaching a predefined subset size or exhausting feasible transitions. Once a feature subset is constructed, its quality is evaluated using a cost function L_k , which typically reflects classification error, validation loss, or a weighted trade-off between accuracy and subset size. If the cost of the current subset is better than the best solution found so far, the global best subset P_{best} and its cost L_{best} are updated accordingly.

After all ants complete their solutions in an iteration, pheromone values on all edges are updated. First, pheromone evaporation is applied to prevent premature convergence and encourage exploration, which is expressed as Equation 11.

$$\tau_{ij} \leftarrow (1 - p) \tau_{ij} \quad (11)$$

Next, pheromone reinforcement is applied along the edges that belong to the feature subsets constructed by the ants. For each ant k , pheromone is deposited according to the quality of its solution as shown in Equation 12.

$$\tau_{ij} \leftarrow \tau_{ij} + \frac{Q}{L_k}, \forall (i, j) \in P_k \quad (12)$$

This update mechanism increases the likelihood that features appearing in high-quality subsets will be selected in subsequent iterations.

The algorithm repeats this construction–evaluation–update cycle until the maximum number of iterations is reached. The final output is the best feature subset P_{best} , which represents the optimal or near-optimal set of features identified by the ACO-based feature selection process.

2.3. Class Imbalance Handling using Hybrid Sampling

Class imbalance is a common problem in classification tasks, where the number of samples in one class (majority class) significantly exceeds that of another class (minority class). This imbalance often leads to biased learning, as standard classifiers tend to favor the majority class. To address this issue, a hybrid sampling strategy is employed that combines undersampling of the majority class and oversampling of the minority class in a controlled and systematic manner[42].

Let the original dataset be denoted as Equation 13.

$$D = D_{maj} \cup D_{min} \quad (13)$$

D_{maj} represents the set of majority class samples and D_{min} represents the set of minority class samples. Given a target imbalance ratio r , the objective of the hybrid sampling method is to construct a balanced dataset in which the class distribution better supports effective model training.

The first stage of the hybrid sampling process focuses on undersampling the majority class. A target number of majority samples is defined as Equation 14.

$$\eta_{target} = r \cdot |D_{min}| \quad (14)$$

$|D_{min}|$ denotes the number of minority samples. If the number of majority samples exceeds this target, random undersampling is applied to reduce the majority class to n_{target} samples. Otherwise, the majority class is retained without modification. This step helps eliminate redundant majority samples while preserving a representative subset of the original data.

2.4. Proposed Method

The proposed method is designed as an integrated image analysis pipeline that combines deep learning–based segmentation, feature selection, and class imbalance handling to produce robust classification results. The process begins with a raw image dataset as shown in Equation 15.

$$D = \{(I_i, y_i)\}_{i=1}^N \quad (15)$$

I_i denotes the input image and y_i represents its corresponding class label. The overall objective of this pipeline is to predict class labels accurately while addressing feature redundancy and data imbalance issues.

Initially, each input image is independently preprocessed to ensure consistency and numerical stability. The image is first normalized to standardize pixel intensity values, followed by resizing to a fixed spatial resolution

(H,W) suitable for the deep learning model. The preprocessed image is then passed through a U-Net–based segmentation network, which produces a segmentation map capturing relevant spatial regions of interest. From this segmentation output, discriminative feature representations are extracted to describe structural and semantic characteristics of the image. After processing all images, the extracted feature vectors are aggregated into a unified feature set.

To reduce feature redundancy and enhance discriminative power, a feature selection stage based on Ant Colony Optimization (ACO) is applied. In this stage, the feature selection problem is modeled as a graph, where each feature corresponds to a node and transitions between features are guided by pheromone intensity and heuristic information. Through iterative solution construction and pheromone updates, ACO identifies an optimal subset of features that minimizes a predefined cost function while preserving classification relevance. The resulting selected feature set represents the most informative features for subsequent learning.

The selected features are then partitioned according to their class labels into majority and minority class subsets. To address class imbalance, a hybrid sampling strategy is employed. First, the majority class is undersampled by defining a target sample size proportional to the minority class size. If the majority class exceeds this target, random undersampling is applied to reduce its size; otherwise, it remains unchanged. Next, the minority class is oversampled using the Synthetic Minority Over-sampling Technique

(SMOTE) to generate synthetic samples when additional instances are required. This two-stage strategy ensures a more balanced class distribution while minimizing information loss and overfitting risks.

The balanced dataset is subsequently used for final model learning. Each instance is again normalized and resized before being processed by convolutional layers to extract refined representations. These layers consist of convolution operations followed by ReLU activation functions, enabling the model to learn non-linear and hierarchical feature patterns. This iterative feature refinement improves the model’s ability to distinguish between classes under balanced conditions.

Finally, the learned representations are used to predict class labels for the input data. By integrating segmentation-driven feature extraction, optimization-based feature selection, and hybrid sampling for imbalance correction, the proposed method forms a cohesive and robust image analysis framework. This end-to-end pipeline enhances classification accuracy, improves generalization performance, and ensures stable learning even in the presence of complex visual patterns and skewed class distributions

2.5 Dataset

Four different image datasets were chosen to assess the effectiveness of the classifiers. All datasets are publicly accessible through the Roboflow platform[43].

The dataset comprises four categories of rice leaf diseases, each containing 2,451 samples, as shown in Table 1.

Table 1. Dataset

Dataset	Total Samples	Number of Samples in The Majority Class	Number of Samples in The Minority Class
Brown Spot	2451	2062 (84.13%)	389(15.87%)
Leaf Blast	2451	2020(82.42%)	431(17.58%)
Leaf Blight	2451	2136(87.15%)	315(12.85%)
Leaf Scald	2451	2101(85.72%)	350(14.28%)

All categories exhibit a clear class imbalance, where the majority class significantly outweighs the minority class—typically over 80% of the samples belong to the majority class. This imbalance can lead to biased model training, where the model may perform well on the majority class but poorly on the minority class. Therefore, careful handling through techniques such as data balancing or augmentation is essential to ensure fair and accurate classification.

2.6. Evaluation Metric

Several commonly used metrics are applied to evaluate and report segmentation results. These include Accuracy, Intersection over Union (IoU), Precision, and Recall [17], [44] as shown in Table 2.

The Accuracy, IoU, Precision, and Recall calculations can be seen in Equations 16 to 19.

Table 2. Confusion Matrix

		Predictive Positive Class	Predictive Negative Class
Actual Class	Positive	True Positive (TP)	False Negative (FN)
Actual Class	Negative	False Positive (FP)	True Negative (TN)

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN} \quad (16)$$

$$IoU = \frac{TP}{TP+FP+FN} \quad (17)$$

$$Precision = \frac{TP}{TP+FP} \quad (18)$$

$$Recall = \frac{TP}{TP+FN} \quad (19)$$

The purpose of these metrics is to evaluate how well the model's predictions align with the actual ground truth. Accuracy measures the overall correctness of the model by calculating how often it makes the right predictions across all classes. It gives a general idea of the model's performance but can be misleading in cases of class imbalance, where one class dominates the dataset. Precision refers to how many of the predicted positive cases are actually correct, making it a critical metric when false positives need to be minimized, such as in medical diagnoses or fraud detection. Recall, on the other hand, measures how well the model captures all actual positive cases, which is important in scenarios where missing a positive instance could have serious consequences, like disease detection or safety monitoring. These metrics, along with Intersection over Union (IoU), collectively help evaluate the model's prediction quality from different perspectives. Intersection over Union (IoU) specifically quantifies the degree of overlap between the predicted and actual segmentation areas.

3. Results and Discussion

3.1 Results

To evaluate its effectiveness, the proposed method will be benchmarked against established models such as SegNet[13], PspNet[45], and E-Net[46]. The results can be seen in Tables 3-6. Figures 2 to 5 present heatmap-based visualizations to complement the quantitative results shown in Tables 3–6. These heatmaps provide a qualitative visualization of the semantic segmentation performance, highlighting pixel-wise accuracy, IoU, precision, and recall for each method.

Table 3. Performance for Brown Spot Dataset

	Accuracy	IoU	Precision	Recall
Proposed Method	0.962	0.784	0.889	0.869
SegNet	0.940	0.682	0.812	0.810
PspNet	0.927	0.629	0.761	0.784
E-Net	0.906	0.556	0.688	0.743

The proposed method demonstrates superior performance across all evaluation metrics compared to the baseline models (SegNet, PspNet, and E-Net). It achieves the highest accuracy of 0.962, indicating better

overall pixel classification. Furthermore, it records the highest Intersection over Union (IoU) score of 0.784, reflecting more precise segmentation and overlap between the predicted and ground truth regions. In terms of precision and recall, the proposed method also leads with scores of 0.889 and 0.869, respectively, which implies that it not only produces fewer false positives but also effectively identifies relevant features in the image. In contrast, SegNet, PspNet, and E-Net exhibit lower scores across all metrics, with E-Net performing the weakest, particularly in IoU and precision. These results highlight the robustness and effectiveness of the proposed method in semantic segmentation tasks.

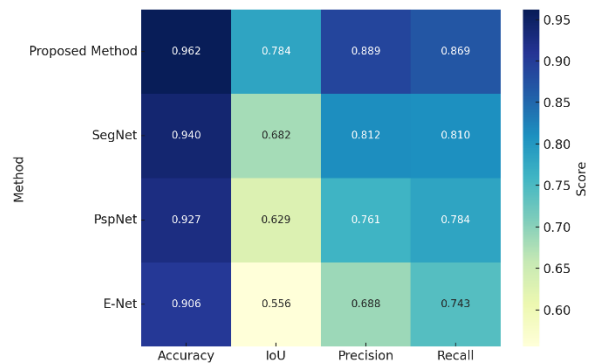


Figure 2. Performance for Brown Spot Dataset

On a Leaf Blast dataset shown in Table 4, the proposed method continues to demonstrate superior performance across all evaluation metrics. It achieves the highest accuracy of 0.947, indicating strong overall classification performance.

Table 4. Performance for Leaf Blast Dataset

	Accuracy	IoU	Precision	Recall
Proposed Method	0.947	0.749	0.813	0.905
SegNet	0.873	0.499	0.620	0.719
PspNet	0.900	0.584	0.683	0.800
E-Net	0.924	0.668	0.743	0.870

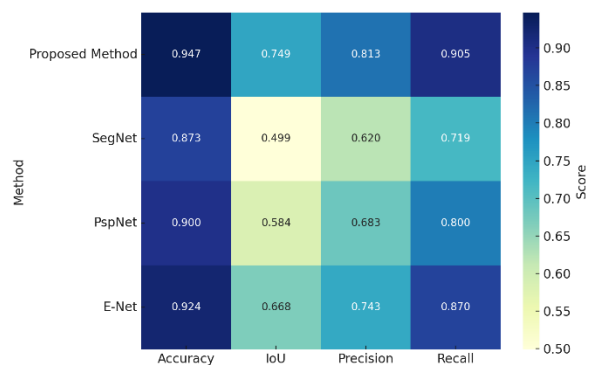


Figure 3. Performance for Leaf Blast Dataset

The method also outperforms others in terms of Intersection over Union (IoU), scoring 0.749, which suggests a high level of overlap between predicted and actual regions. Precision and recall values are also leading at 0.813 and 0.905, respectively, signifying the model's ability to minimize false positives while

effectively detecting relevant segments. In comparison, E-Net shows the second-best performance with an accuracy of 0.924 and recall of 0.870, though it still lags behind the proposed method. SegNet and PspNet perform notably worse, especially in IoU and precision.

On Leaf Blight dataset, the proposed method continues to deliver the best overall performance, achieving the highest accuracy of 0.951. It also maintains a strong Intersection over Union (IoU) score of 0.682, slightly surpassing SegNet, which scores 0.673. While SegNet records a higher recall of 0.857, indicating it captures more relevant segments, the proposed method offers a better balance between precision (0.798) and recall (0.825), suggesting it can detect relevant areas with fewer false positives. PspNet and E-Net lag behind in all metrics, particularly in IoU and precision, highlighting their weaker segmentation performance. Overall, these results show that the proposed method remains highly effective and consistent, even when tested across varying datasets.

Table 5. Performance for Leaf Blight Dataset

	Accuracy	IoU	Precision	Recall
Proposed Method	0.951	0.682	0.798	0.825
SegNet	0.947	0.673	0.758	0.857
PspNet	0.928	0.587	0.693	0.794
E-Net	0.916	0.543	0.643	0.778

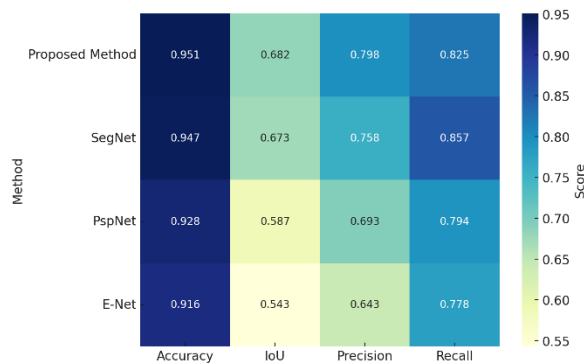


Figure 4. Performance for Leaf Blight Dataset

In Leaf Scald dataset, the proposed method significantly outperforms all baseline models across every evaluation metric. It achieves the highest accuracy at 0.957, indicating excellent pixel-wise classification performance. The Intersection over Union (IoU) score of 0.754 is also the highest, demonstrating superior segmentation quality with a strong overlap between predictions and ground truth. Furthermore, the proposed method records the highest recall at 0.929, meaning it is very effective in capturing relevant regions, and a strong precision of 0.800, indicating relatively few false positives. While PspNet comes second with a recall of 0.886 and IoU of 0.658, it still falls behind in overall performance. SegNet and E-Net perform considerably worse, particularly in IoU and precision. These results reinforce the proposed method's reliability and strong generalization capabilities in handling different segmentation challenges.

Table 6. Performance for Leaf Scald Dataset

	Accuracy	IoU	Precision	Recall
Proposed Method	0.957	0.754	0.800	0.929
SegNet	0.885	0.471	0.580	0.714
PspNet	0.934	0.658	0.719	0.886
E-Net	0.912	0.569	0.654	0.814

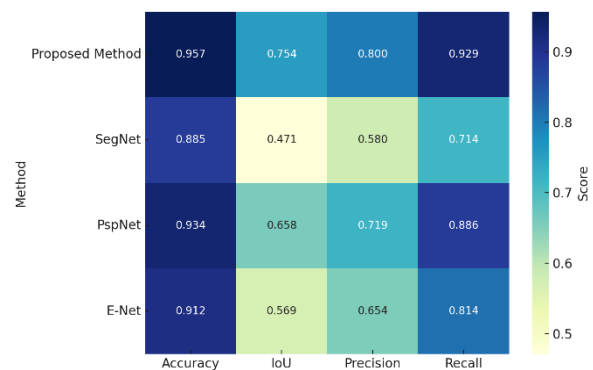


Figure 5. Performance for Leaf Scald Dataset

The performance corresponding to each parameter is presented in Figures 6–9.

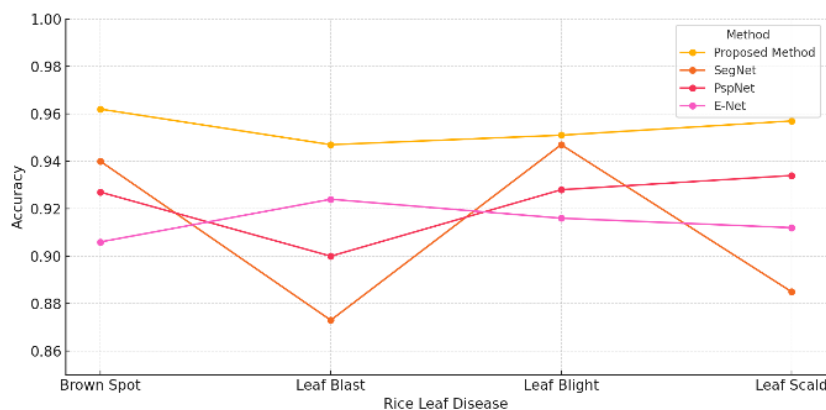


Figure 6. Accuracy for Each Dataset

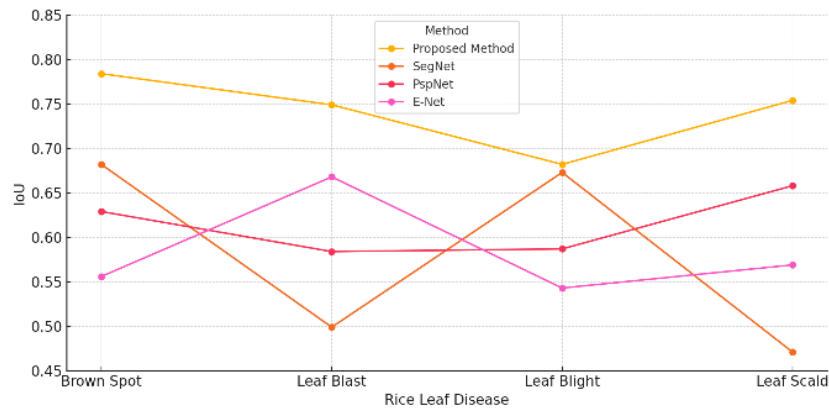


Figure 7. IoU for Each Dataset

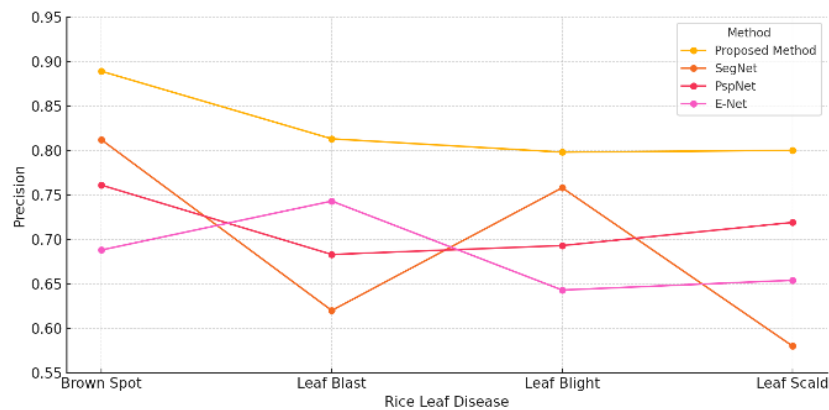


Figure 8. Precision for Each Dataset

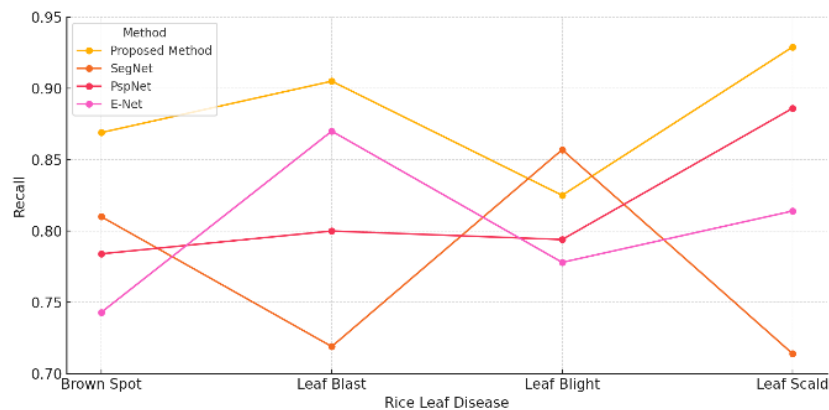


Figure 9. Recall for Each Dataset

3.2 Discussion

The evaluation results demonstrate that the Proposed Method consistently outperforms the other three methods (SegNet, PspNet, and E-Net) across all four metrics: Accuracy, IoU, Precision, and Recall, for the classification of rice leaf diseases: Brown Spot, Leaf Blast, Leaf Blight, and Leaf Scald. The Proposed Method maintains high accuracy (above 0.94) and delivers strong IoU and recall values, particularly for Leaf Scald, where recall exceeds 0.92. In contrast, baseline methods such as SegNet, PspNet, and E-Net exhibit more variability in performance depending on the disease type. For instance, SegNet achieves relatively strong results on Leaf Blight in terms of

recall, but falls behind in IoU and precision on Leaf Scald. Similarly, E-Net tends to underperform across most metrics, particularly in terms of IoU and precision. The superior performance of the Proposed Method indicates its robustness and generalizability across different disease patterns and segmentation challenges. High recall values reflect its ability to accurately detect disease-affected regions, while consistent precision scores suggest fewer false positives. The balance between these metrics makes it a suitable candidate for practical deployment in precision agriculture tools, especially those targeting automated rice disease detection using image-based analysis. Future work may explore enhancing lightweight architectures, reducing computation time.

3.2 Statistical Significance Test

The testing was conducted using Wilcoxon Signed-Rank Test to compare the performance of the Proposed Method with baseline models (SegNet, PspNet, E-Net) using the metrics Accuracy, IoU, Precision, and Recall. The results can be seen in Table 7.

Table 7. Significance Test using Wilcoxon Signed-Rank Test

Metric	Model vs Proposed	p-value
Accuracy	SegNet	0.03125
Accuracy	PspNet	0.03125
Accuracy	E-Net	0.03125
IoU	SegNet	0.03125
IoU	PspNet	0.03125
IoU	E-Net	0.03125
Precision	SegNet	0.03125
Precision	PspNet	0.03125
Precision	E-Net	0.03125
Recall	SegNet	0.15625
Recall	PspNet	0.03125
Recall	E-Net	0.03125

The results of the Wilcoxon Signed-Rank Test presented in Table 7 indicate that the proposed model demonstrates statistically significant improvements over SegNet, PspNet, and E-Net across most evaluation metrics. Specifically, for Accuracy, IoU, and Precision, the p-values are consistently 0.03125, which is below the 0.05 significance threshold, confirming that the performance gains of the proposed method compared to the baseline models are not due to random chance. For Recall, however, the outcomes are mixed: while comparisons with PspNet and E-Net yield significant differences ($p = 0.03125$), the comparison with SegNet produces a p-value of 0.15625, which is above the 0.05 threshold, indicating no statistically significant improvement in this case. Overall, these results suggest that the proposed model provides robust and consistent enhancements in terms of accuracy, intersection-over-union, and precision, while recall improvement is less consistent depending on the baseline model.

4. Conclusions

This study introduces a hybrid framework that integrates U-Net-based segmentation, Ant Colony Optimization (ACO) for feature selection, hybrid sampling for class imbalance handling, and a Convolutional Neural Network (CNN) for rice leaf disease classification. Experimental results across four rice disease datasets consistently demonstrate that the proposed method surpasses established baselines (SegNet, PspNet, and E-Net) in terms of accuracy, precision, recall, and IoU. These findings confirm that combining careful preprocessing, intelligent feature optimization, and balanced learning substantially enhances the robustness and reliability of agricultural image classification systems.

From a practical perspective, the framework offers direct benefits for professionals in agriculture. Farmers and agricultural extension officers could adopt this model to enable early and accurate diagnosis of rice leaf

diseases, reducing reliance on manual inspection and mitigating the risks of delayed treatment. Technology developers can embed the framework into mobile and IoT-based monitoring tools, allowing real-time field deployment in resource-constrained environments. For policymakers, the results highlight the potential of digital agriculture to support national food security strategies, optimize pesticide usage, and reduce crop losses—ultimately contributing to sustainable agricultural practices and aligning with global goals such as the SDGs.

For the broader research community, this work demonstrates the importance of integrated approaches rather than isolated techniques in plant disease detection. The hybridization of segmentation, optimization, and sampling provides a replicable methodology for other domains affected by data imbalance and complex image environments, such as medical diagnostics or environmental monitoring.

Future research may focus on model compression and optimization to support lightweight deployment, extending the framework to multi-label or multi-crop disease classification, and integrating it into IoT-enabled surveillance systems for large-scale agricultural monitoring. Such developments would not only enhance scientific progress but also generate tangible societal benefits by empowering farmers, informing policy, and safeguarding food production against disease-related risks.

Acknowledgements

This work was supported by Ministry of Higher Education, Science, and Technology.

References

- [1] N. Hashim *et al.*, "Smart Farming for Sustainable Rice Production: An Insight into Application, Challenge, and Future Prospect," *Rice Science*, vol. 31, no. 1, pp. 47–61, Jan. 2024, doi: 10.1016/j.rsci.2023.08.004.
- [2] R. C. Joshi, R. Burget, and M. K. Dutta, "AgriDeep-net: An advanced deep feature fusion-based technique for enhanced fine-grain image analytics in precision agriculture," *Ecological Informatics*, vol. 86, p. 103069, May 2025, doi: 10.1016/j.ecoinf.2025.103069.
- [3] Z. Salman, A. Muhammad, and D. Han, "Plant disease classification in the wild using vision transformers and mixture of experts," *Front. Plant Sci.*, vol. 16, June 2025, doi: 10.3389/fpls.2025.1522985.
- [4] W. Albattah and R. U. Khan, "Impact of imbalanced features on large datasets," *Front. Big Data*, vol. 8, Mar. 2025, doi: 10.3389/fdata.2025.1455442.
- [5] L. Wang, X. Gong, J. Zhang, and R. Chen, "Anti-noise face: A resilient model for face recognition with labeled noise data," *Signal Processing: Image Communication*, vol. 134, p. 117269, May 2025, doi: 10.1016/j.image.2025.117269.
- [6] K. Salazar-Serna, S. A. Barona, I. C. García, L. Cadavid, and C. J. Franco, "Addressing overfitting in classification models for transport mode choice prediction: a practical application in the Aburrá Valley, Colombia," *Transportation Letters*, vol. 17, no. 7, pp. 1213–1230, Aug. 2025, doi: 10.1080/19427867.2024.2422717.
- [7] C. Kolling, V. Araujo, A. Veloso, and S. R. Musse, "Mitigating bias in facial analysis systems by incorporating label diversity," *Computers & Graphics*, vol. 116, pp. 173–184, Nov. 2023, doi: 10.1016/j.cag.2023.08.021.

- [8] U. Muñoz-Aseguinolaza, I. Fernandez-Iriondo, I. Rodríguez-Moreno, N. Aginako, and B. Sierra, "Convolutional neural network-based classification and monitoring models for lung cancer detection: 3D perspective approach," *Heliyon*, vol. 9, no. 11, p. e21203, Nov. 2023, doi: 10.1016/j.heliyon.2023.e21203.
- [9] S. Madhu, Dr. K. N. Bhargavi, M. V. S. Ramprasad, S. Gautam, and S. Bhavana, "Accurate diabetic retinopathy segmentation and classification model using gated recurrent unit with residual attention network," *Biomedical Signal Processing and Control*, vol. 102, p. 107348, Apr. 2025, doi: 10.1016/j.bspc.2024.107348.
- [10] H. Vaidya et al., "A class imbalance aware hybrid model for accurate rice variety classification," *International Journal of Cognitive Computing in Engineering*, vol. 6, pp. 170–182, Dec. 2025, doi: 10.1016/j.ijcce.2024.12.004.
- [11] S. Zarbakhsh, F. Fakhrazad, D. Rajkovic, G. Niedbala, and M. Piekutowska, "Approaches and challenges in machine learning for monitoring agricultural products and predicting plant physiological responses to biotic and abiotic stresses," *Current Plant Biology*, vol. 43, p. 100535, Sept. 2025, doi: 10.1016/j.cpb.2025.100535.
- [12] A. Razzaque and D. A. Badholia, "PCA based feature extraction and MPSO based feature selection for gene expression microarray medical data classification," *Measurement: Sensors*, vol. 31, p. 100945, Feb. 2024, doi: 10.1016/j.measen.2023.100945.
- [13] A. V. Jonnalagadda and H. A. Hashim, "SegNet: A segmented deep learning based Convolutional Neural Network approach for drones wildfire detection," *Remote Sensing Applications: Society and Environment*, vol. 34, p. 101181, Apr. 2024, doi: 10.1016/j.rsase.2024.101181.
- [14] C. G. Simhadri, H. K. Kondaveeti, V. K. Vatsavayi, A. Mitra, and P. Ananthachari, "Deep learning for rice leaf disease detection: A systematic literature review on emerging trends, methodologies and techniques," *Information Processing in Agriculture*, vol. 12, no. 2, pp. 151–168, June 2025, doi: 10.1016/j.inpa.2024.04.006.
- [15] M. Schotte, J. Chaumel, M. N. Dean, and D. Baum, "Image analysis pipeline for segmentation of a biological porosity network, the lacuno-canalicular system in stingray tesseræ," *MethodsX*, vol. 7, p. 100905, Jan. 2020, doi: 10.1016/j.mex.2020.100905.
- [16] Y. Yu et al., "Techniques and Challenges of Image Segmentation: A Review," *Electronics*, vol. 12, no. 5, Art. no. 5, Jan. 2023, doi: 10.3390/electronics12051199.
- [17] A. Alweshah, R. Barzamani, F. Hajati, S. S. S. Farahani, M. Arabian, and B. Sohani, "Temporal dependency modeling for improved medical image segmentation: The R-UNet perspective," *Franklin Open*, vol. 9, p. 100182, Dec. 2024, doi: 10.1016/j.fraope.2024.100182.
- [18] C. Zhang, X. Deng, and S. H. Ling, "Next-Gen Medical Imaging: U-Net Evolution and the Rise of Transformers," *Sensors*, vol. 24, no. 14, Art. no. 14, Jan. 2024, doi: 10.3390/s24144668.
- [19] J. Albarracin, F. Cano, E. Romero, and A. Cruz-Roa, "A comparative analysis between two convolutional networks architectures for semantic segmentation of histopathology breast cancer images," in *2023 19th International Symposium on Medical Information Processing and Analysis (SIPAIM)*, Nov. 2023, pp. 1–5, doi: 10.1109/SIPAIM56729.2023.10373477.
- [20] G. Bala, H. Mondal, and A. Sethi, "Brain MRI Segmentation Using U-Net and SegNet: A Comparative Study Across Modalities with Robust Loss Functions," presented at the International Joint Conference on Biomedical Engineering Systems and Technologies, SCITEPRESS, Feb. 2025, pp. 246–254, doi: 10.5220/0013302900003911.
- [21] C. Beeche et al., "Super U-Net: A modularized generalizable architecture," *Pattern Recognition*, vol. 128, p. 108669, Aug. 2022, doi: 10.1016/j.patcog.2022.108669.
- [22] B. Mustapha, Y. Zhou, B. Nawel, S. Chunyan, and X. Zhitao, "Optimized attention U-Net for enhanced lung and area of infection segmentation in chest X-Rays and CT scans," *Journal of Radiation Research and Applied Sciences*, vol. 18, no. 3, p. 101650, Sept. 2025, doi: 10.1016/j.jrras.2025.101650.
- [23] Y. Li, K. Li, and Q. Lu, "Applying segmentation and classification to improve performance of smoothing," *Digital Signal Processing*, vol. 109, p. 102913, Feb. 2021, doi: 10.1016/j.dsp.2020.102913.
- [24] S. Zhao, M. Wang, S. Ma, and Q. Cui, "A feature selection method via relevant-redundant weight," *Expert Systems with Applications*, vol. 207, p. 117923, Nov. 2022, doi: 10.1016/j.eswa.2022.117923.
- [25] L. Tello-Cifuentes, J. Marulanda, and P. Thomson, "Segmentation and feature extraction-based classification of pavement damages using hybrid computer vision and machine learning approaches," *Mach Learn*, vol. 114, no. 8, p. 172, June 2025, doi: 10.1007/s10994-025-06803-3.
- [26] M. Hassan and N. Kaabouch, "Impact of Feature Selection Techniques on the Performance of Machine Learning Models for Depression Detection Using EEG Data," *Applied Sciences*, vol. 14, no. 22, Art. no. 22, Jan. 2024, doi: 10.3390/app142210532.
- [27] E. Oluwasakin et al., "Minimization of high computational cost in data preprocessing and modeling using MP4Py," *Machine Learning with Applications*, vol. 13, p. 100483, Sept. 2023, doi: 10.1016/j.mlwa.2023.100483.
- [28] Q. Zhou, Q. Gao, Q. Wang, M. Yang, and X. Gao, "Sparse discriminant PCA based on contrastive learning and class-specificity distribution," *Neural Networks*, vol. 167, pp. 775–786, Oct. 2023, doi: 10.1016/j.neunet.2023.08.061.
- [29] A. Hashemi, M. Joodaki, N. Z. Joodaki, and M. B. Dowlatshahi, "Ant colony optimization equipped with an ensemble of heuristics through multi-criteria decision making: A case study in ensemble feature selection," *Applied Soft Computing*, vol. 124, p. 109046, July 2022, doi: 10.1016/j.asoc.2022.109046.
- [30] R. Devi Priya, R. Sivaraj, N. Anitha, and V. Devisurya, "Tri-staged feature selection in multi-class heterogeneous datasets using memetic algorithm and cuckoo search optimization," *Expert Systems with Applications*, vol. 209, p. 118286, Dec. 2022, doi: 10.1016/j.eswa.2022.118286.
- [31] S. Kashef and H. Nezamabadi-pour, "An advanced ACO algorithm for feature subset selection," *Neurocomputing*, vol. 147, pp. 271–279, Jan. 2015, doi: 10.1016/j.neucom.2014.06.067.
- [32] H. Huang, H.-B. Xie, J.-Y. Guo, and H.-J. Chen, "Ant colony optimization-based feature selection method for surface electromyography signals classification," *Comput Biol Med*, vol. 42, no. 1, pp. 30–38, Jan. 2012, doi: 10.1016/j.combiomed.2011.10.004.
- [33] I.-S. Oh and J.-S. Lee, "Ant Colony Optimization for Feature Selection in Pattern Recognition," *The Journal of the Korea Contents Association*, vol. 10, no. 5, pp. 1–9, 2010, doi: 10.5392/JKCA.2010.10.5.001.
- [34] A. Z. Ye et al., "High-Dimensional Feature Selection Based on Improved Binary Ant Colony Optimization Combined with Hybrid Rice Optimization Algorithm," *International Journal of Intelligent Systems*, vol. 2023, no. 1, p. 1444938, 2023, doi: 10.1155/2023/1444938.
- [35] B. Chung, "Addressing Data Imbalance in Plant Disease Recognition through Contrastive Learning," in *2024 IEEE 3rd International Conference on AI in Cybersecurity (ICAIC)*, Feb. 2024, pp. 1–6, doi: 10.1109/ICAIC60265.2024.10433841.
- [36] Y. Liu, L. Zhu, L. Ding, H. Sui, and W. Shang, "A hybrid sampling method for highly imbalanced and overlapped data classification with complex distribution," *Information Sciences*, vol. 661, p. 120117, Mar. 2024, doi: 10.1016/j.ins.2024.120117.
- [37] N. Nnamoko and I. Korkontzelos, "Efficient treatment of outliers and class imbalance for diabetes prediction," *Artificial Intelligence in Medicine*, vol. 104, p. 101815, Apr. 2020, doi: 10.1016/j.artmed.2020.101815.
- [38] D. Yilmaz Eroglu and M. S. Pir, "Hybrid Oversampling and Undersampling Method (HOUM) via Safe-Level SMOTE and Support Vector Machine," *Applied Sciences*, vol. 14, no. 22, Art. no. 22, Jan. 2024, doi: 10.3390/app142210438.

- [39] M. Alamri and M. Ykhlef, "Hybrid Undersampling and Oversampling for Handling Imbalanced Credit Card Data," *IEEE Access*, vol. 12, pp. 14050–14060, 2024, doi: 10.1109/ACCESS.2024.3357091.
- [40] Y. Ye et al., "Image segmentation using improved U-Net model and convolutional block attention module based on cardiac magnetic resonance imaging," *Journal of Radiation Research and Applied Sciences*, vol. 17, no. 1, p. 100816, Mar. 2024, doi: 10.1016/j.jrras.2023.100816.
- [41] X. Zhang, K. Xu, L. Zhang, R. Zhao, W. Wei, and Y. She, "Optimal channel dynamic selection for Constructing lightweight Data EEG-based emotion recognition," *Heliyon*, vol. 10, no. 9, p. e30174, May 2024, doi: 10.1016/j.heliyon.2024.e30174.
- [42] L. Chen, J. Jiang, Y. Zhang, and H. Chen, "HSDP: A Hybrid Sampling Method for Imbalanced Big Data Based on Data Partition," *Complex.*, vol. 2021, Jan. 2021, doi: 10.1155/2021/6877284.
- [43] R. Roboflow, "Palm Leaves Disease Detection," Roboflow. Accessed: June 14, 2025. [Online]. Available: <https://universe.roboflow.com/firstproject-ofadf/palm-leaves-disease-detection>
- [44] M. Desai, H. Mewada, I. M. Pires, and S. Roy, "Evaluating the Performance of the YOLO Object Detection Framework on COCO Dataset and Real-World Scenarios," *Procedia Computer Science*, vol. 251, pp. 157–163, Jan. 2024, doi: 10.1016/j.procs.2024.11.096.
- [45] L. Yan et al., "PSP net-based automatic segmentation network model for prostate magnetic resonance imaging," *Computer Methods and Programs in Biomedicine*, vol. 207, p. 106211, Aug. 2021, doi: 10.1016/j.cmpb.2021.106211.
- [46] D. Wicker, M. M. Rizki, and L. A. Tamburino, "E-Net: Evolutionary neural network synthesis," *Neurocomputing*, vol. 42, no. 1, pp. 171–196, Jan. 2002, doi: 10.1016/S0925-2312(01)00599-9.