



Classification of Pestalotiopsis sp. Leaf Fall Disease Severity in Rubber Plants using UAV Multispectral Vegetation Indices and 1-D Convolutional Neural Networks

Solikin^{1,2}, Yeni Herdiyeni³, Annisa⁴, Lilik Budi Prasetyo⁵, Tri Rapani Febbiyanti⁶, Imas Sukaesih Sitanggang⁷, Sri Nurdianti⁸

¹Doctoral Program in Computer Science, School of Data Science, Mathematics and Informatics, IPB University, Bogor, Indonesia

²Information Systems, Faculty of Informatics and Design, Bina Insani University, Bekasi, Indonesia

^{3,4,7,8}School of Data Science, Mathematics and Informatics, IPB University, Bogor, Indonesia

⁵Department of Forest Resources Conservation and Ecotourism, Faculty of Forestry and Environment, IPB University, Bogor, Indonesia

⁶Indonesian Rubber Research Institute Sembawa, Banyuasin, Indonesia

solikinsolikin@apps.ipb.ac.id

Abstract

Leaf-fall disease caused by *Pestalotiopsis* sp. is a major threat to rubber (*Hevea brasiliensis*) plantations because it suppresses photosynthetic activity, accelerates defoliation, and reduces latex productivity. In operational practice, severity assessment is still dominated by visual field inspection, which is subjective, time-consuming, costly, and difficult to standardize across large plantation areas. This study develops a disease severity classification model for *Pestalotiopsis* sp. using a Convolutional Neural Network (CNN) based on vegetation-index features derived from UAV multispectral imagery. The model classifies disease severity into four levels: L1 (Light Infection), L2 (Moderate Infection), L3 (Severe Infection), and L4 (Very Severe Infection). To represent temporal and biological variability in disease expression, multispectral data were collected from multiple rubber clones over two observation periods. Feature construction focused on NDRE, LCI, CI, NDVI, NDRE Interaction, and GCI Ratio, which capture chlorophyll-related and canopy condition responses to infection. Because severity classes were imbalanced, the Synthetic Minority Over-sampling Technique (SMOTE) was applied before model training. A one-dimensional CNN was then trained to learn nonlinear patterns among index-based predictors for multilevel severity classification. Hyperparameter tuning improved overall accuracy from 85.30% to 90.00%. Class-wise F1-scores changed from 0.91 to 0.94 (L1), 0.83 to 0.84 (L2), 0.75 to 0.88 (L3), and 0.97 to 0.84 (L4), with the largest improvement in L3 recall (0.67 to 0.94). These results indicate that the selected vegetation indices and interaction terms are informative predictors for objective and scalable disease severity classification under heterogeneous plantation conditions.

Keywords: convolutional neural network; leaf fall disease; pestalotiopsis sp.; rubber plant; vegetation index

How to Cite: Solikin, "Classification of Pestalotiopsis sp. Leaf Fall Disease Severity in Rubber Plants using UAV Multispectral Vegetation Indices and 1-D Convolutional Neural Networks", *J. RESTI (Rekayasa Sist. Teknol. Inf.)*, vol. 10, no. 3, pp. 567 - 579, May 2026.

Permalink/DOI: <https://doi.org/10.29207/resti.v10i3.7601>

Received: February 24, 2026

Accepted: April 28, 2026

Available Online: May 28, 2026

This is an open-access article under the [CC BY 4.0 License](#)

Published by [Ikatan Ahli Informatika Indonesia](#)

1. Introduction

Rubber (*Hevea brasiliensis*) is a strategic commodity that sustains natural rubber-based industries and the livelihoods of tropical smallholders. Its productivity, however, is increasingly threatened by leaf fall disease associated with *Pestalotiopsis* sp., which has been reported to reduce leaf area index and canopy capacity,

disrupt physiological functions, and ultimately lower yield [1-4].

At the plantation scale, the epidemiological dynamics and spatial spread of the disease can accelerate canopy loss and intensify production risks [4]. In practice, disease severity is still assessed predominantly through visual field inspection, even though this approach is

subjective, labor-intensive, costly, and difficult to standardize under heterogeneous field conditions and symptom expressions that may resemble other physiological stresses [1-4]. These limitations make objective severity quantification an important measurement and control problem, particularly for large-scale monitoring systems that require consistent, repeatable, and scalable decision support.

Recent studies have shown that remote sensing can provide a more objective basis for plant health assessment, but their contributions differ substantially in scale, purpose, and analytical depth. At broader spatial scales, satellite-based analyses such as Sentinel-2 have been used to characterize phenological and canopy changes in rubber stands during infection, showing that spectral responses can capture disease-related canopy dynamics [5], [6]. These studies are important because they establish that infection leaves measurable spectral signatures at the canopy level, yet their emphasis is primarily on temporal monitoring and landscape-level detection rather than severity discrimination at the plant or symptom level. At finer spatial resolution, UAV-based imaging has become increasingly relevant; for example, multi-angle RGB imagery combined with deep learning has been used successfully for rubber tree recognition, demonstrating the operational value of aerial image analytics in plantation environments [7]. However, recognition of tree objects is analytically different from discrimination of disease severity classes. Other comparable studies have reported that multi-index vegetation modeling from high-resolution imagery can explain spatio-temporal patterns of leaf fall and their relation to climatic drivers [8], while the broader vegetation-mapping literature has shown that vegetation indices and multivariate spectral features can improve machine-learning performance by representing chlorophyll status, canopy vigor, and subtle spectral variation more effectively [9]. Taken together, these main contributors justify the relevance of spectral monitoring, but they also reveal that prior studies have been stronger in detection, mapping, and explanatory canopy analysis than in critically tested severity-level classification based on physiologically grounded index combinations within a unified deep-learning framework.

Despite this progress, an important gap remains in the development of a robust classification model for *Pestalotiopsis* sp. severity that directly integrates physiologically relevant vegetation-index features into a CNN under heterogeneous plantation conditions. In several previous studies, vegetation-index analysis and deep-learning prediction have been treated as loosely connected stages, while classification performance tends to decline when class distributions are imbalanced, especially for minority severity classes [2], [10]. This indicates not only a methodological gap, but also a practical limitation: a model may detect stress patterns without being sufficiently reliable for fine-

grained severity assignment needed in plantation management. Building on prior remote-sensing and machine-learning studies, this study differs by developing and evaluating a CNN-based model for *Pestalotiopsis* sp. severity classification using NDRE, LCI, CI, NDVI_NDRE_Interaction, and GCI_Ratio as predictor features. To strengthen performance across unequal severity distributions, class-imbalance handling is incorporated during model development to improve robustness and sensitivity for underrepresented classes. The contribution of this article is therefore twofold: first, it demonstrates that physiologically grounded vegetation indices can serve as effective direct inputs for CNN-based multilevel severity classification; second, it shows how this contribution is validated throughout the manuscript through feature relevance analysis, class-imbalance treatment, and performance evaluation across four severity categories, namely L1 (Light Infection), L2 (Moderate Infection), L3 (Severe Infection), and L4 (Very Severe Infection). By combining physiologically meaningful spectral descriptors with CNN-based nonlinear pattern learning, this study extends earlier work on spectral detection and mapping toward a more explicit, testable, and operationally relevant framework for severity discrimination under variable field environments.

2. Methods

This study consists of eight stages to obtain the CNN model, as illustrated in Figure 1. There are vegetation index computation, feature selection using Random Forest (RF), Exploratory Data Analysis (EDA), handling class imbalance using SMOTE, disease severity determination using 1-D CNN, and hyperparameter tuning, model evaluation.

2.1 Vegetation Index Calculation

Vegetation indices were computed to characterize plant physiological status from multispectral reflectance values because chlorophyll-related spectral responses are sensitive to stress and disease progression [11]. Before index extraction, the imagery underwent radiometric calibration, geometric correction, and noise reduction to improve the reliability and comparability of pixel values across acquisition conditions [12], [13]. Reflectance values from the required spectral bands were then used to calculate the vegetation indices reported in this study, including NDVI, MCARI, CI, LCI, NDRE, GCI, MCARI_CI_Ratio, NDVI_NDRE_Interaction, and GCI_Ratio, using their respective formulas as presented in Table 1. Each computed index was linked to the corresponding labeled sample and organized into a structured dataset for subsequent analysis. This procedure was chosen because vegetation-index transformation reduces raw spectral complexity while preserving physiologically meaningful information needed for consistent and standardized disease-severity assessment [14].

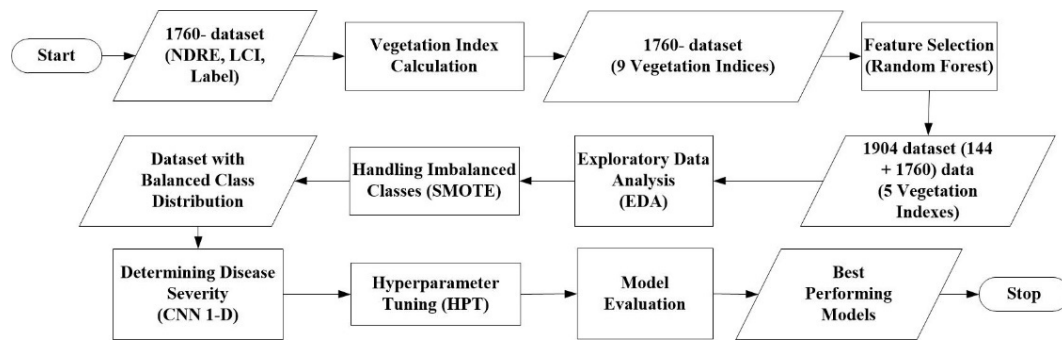


Figure 1. Research Methods

2.2 Feature Selection

Feature selection was performed after vegetation-index construction to reduce dimensionality while preserving the most informative predictors for severity classification. Although several methods such as Recursive Feature Elimination (RFE), Principal Component Analysis (PCA), and RF are commonly used for this purpose [15], [16] RF-based feature importance was adopted in this study because it can rank nonlinear predictor contributions without requiring strict distributional assumptions and can handle correlated vegetation-index variables effectively.

The feature-selection procedure involved training the RF model on the vegetation-index dataset, computing the importance score of each feature, ranking all candidate variables, and retaining the highest-ranked predictors for CNN input. This process improved computational efficiency by reducing input complexity, lowered the risk of overfitting by suppressing redundant information, and supported better generalization during deep-learning training [17-20]. The retained features were then forwarded to the next analytical stages because CNN-based classification using selected vegetation indices has been shown to improve accuracy while optimizing computational resources in agricultural applications [21], [22].

2.3 Exploratory Data Analysis (EDA)

Exploratory Data Analysis (EDA) was conducted before model training to verify data quality and to guide the selection of appropriate preprocessing and modeling strategies. At this stage, descriptive statistics, class-frequency inspection, distribution plots, boxplots, pairwise visualizations, and correlation analysis were used to examine data structure, label balance, inter-variable relationships, and potential anomalies [23], [24].

Outlier inspection was also performed because extreme values can distort feature distributions and bias the training process if left unchecked [25]. Through this procedure, EDA provided practical justification for subsequent methodological choices, particularly the use of feature selection to reduce redundancy and the use of imbalance handling to address underrepresented severity classes. Thus, this stage functioned not merely as a descriptive step, but as a necessary diagnostic stage

to ensure that the downstream CNN model was trained on data that were interpretable, stable, and analytically defensible.

2.4 Handling Imbalanced Classes (SMOTE)

Class imbalance was addressed using the Synthetic Minority Over-sampling Technique (SMOTE) to improve the representation of minority severity classes in the training data. SMOTE works by generating synthetic samples through interpolation between neighboring minority instances in feature space, thereby producing a more balanced class distribution and enabling fairer model learning [26].

In this study, SMOTE was applied only after EDA and feature selection so that synthetic samples were generated from the most relevant vegetation-index predictors rather than from noisy or redundant variables. This order was important because oversampling before variable screening could amplify irrelevant variation. The approach was selected because it has been shown to improve recognition of low-frequency classes and enhance overall classification performance [27], [28], while its known limitations, such as the possible introduction of noisy synthetic observations and the neglect of spatial structure in image-derived data [29-31], were mitigated by applying it to tabular vegetation-index features and by evaluating the resulting model carefully across all severity classes.

2.5 Determination of Disease Severity (CNN 1-D)

Disease severity was determined using a one-dimensional Convolutional Neural Network (1-D CNN) because the model was designed to learn nonlinear patterns from sequential vegetation-index feature vectors rather than from raw 2-D images. In contrast to image-oriented architectures such as ResNet, DenseNet, and EfficientNet, which are widely used for deep feature extraction from spatial image data [32-34], the present study employed a 1-D CNN that is more appropriate for structured spectral-feature input while still enabling hierarchical feature learning [35].

The selected features were arranged as input sequences, passed through stacked Conv1D layers to extract local patterns among vegetation-index combinations, regularized using Dropout layers, reduced through pooling operations, flattened into a dense

representation, and finally classified into four severity levels through fully connected layers. This architecture was chosen because it offers a simpler and more computationally efficient alternative to deeper image-based networks while remaining suitable for multilevel classification of disease severity. The model design therefore reflects a methodological adaptation of standard CNN principles to the specific characteristics of vegetation-index data used in this study [36-39].

2.6 Hyperparameter Tuning

Hyperparameter tuning was conducted to obtain a model configuration that balanced convergence stability, predictive accuracy, and generalization performance. The main hyperparameters considered in this study included learning rate, number of epochs, batch size, optimizer type, and other training-related settings that directly influence gradient updates and model convergence [40], [41].

The tuning process involved training alternative model configurations, comparing their validation performance, and selecting the parameter combination that produced the best trade-off between accuracy and loss. This stage was necessary because an inappropriate learning rate may cause unstable convergence, whereas unsuitable batch size or training duration may either limit learning capacity or increase overfitting risk. Common search strategies such as random search, grid search, and Bayesian optimization are widely recognized [42] accordingly, the tuning strategy in this study was used not as an auxiliary step but as an essential procedure to justify that the final CNN setting was appropriate for the disease-severity classification problem.

2.7 Model Evaluation

In this study, model evaluation was designed to assess multiclass disease-severity classification fairly and transparently, particularly under potentially imbalanced data conditions. The trained model was evaluated using precision, recall, F1-score, accuracy, and AUC-ROC because these metrics jointly reflect class recognition quality, balance between false positives and false negatives, and separability across decision thresholds [1].

In addition, confusion-matrix analysis was used to compare predicted and actual labels and to identify where misclassification occurred, especially between neighboring severity levels that may share similar spectral responses [5]. This evaluation scheme was selected because reliance on accuracy alone can obscure weak performance in minority classes, whereas class-wise metrics provide stronger evidence of operational reliability. A high AUC value was also interpreted as an indicator of stable performance and model comparability, which is important because diagnostic errors directly affect disease-control decision making [27].

3. Results and Discussions

The following section presents the results and discussion of the classification of *Pestalotiopsis* sp. disease severity in rubber plants.

3.1 Vegetation Index Calculation

The dataset construction stage began with 144 ground-truth samples representing disease severity and two key vegetation indices, Normalized Difference Red-Edge (NDRE) and Leaf Chlorophyll Index (LCI), extracted from multispectral UAV imagery in our previous RF-based study. These data were collected from three rubber clones, BPM-24, GT 1, and RRIC 100, in the Sembawa Rubber Research Plantation area, Banyuasin, Palembang, South Sumatra, Indonesia. UAV imagery was acquired in two campaigns: 18–19 January 2022 and 15–16 June 2022. Building on that RF framework, we further developed and expanded the dataset by adding 1760 NDRE, LCI, label records. This augmentation was designed to preserve the spectral patterns and class-label structure learned from the earlier RF model, resulting in an enlarged dataset that remains consistent with the original severity representation [43]. In the context of the present study, this expanded dataset is important because the initial hypothesis assumes that physiologically meaningful spectral information can support stable multi-level severity classification when sufficient variability in field conditions is represented.

Following data acquisition, a disease severity labeling process for *Pestalotiopsis* sp. was conducted. The applied relabeling strategy improved the separation among severity levels (Figure 2), resulting in more specific and distinguishable NDRE and LCI value ranges for each class. This adjustment reduced inter-class overlap and enhanced the spectral characterization of individual severity levels.

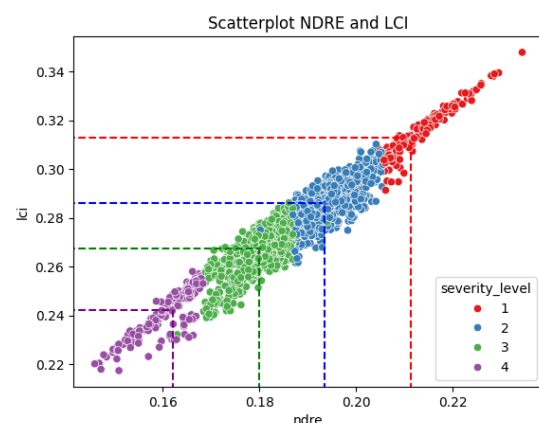


Figure 2. Scatter Plot Showing the Relabeled Dataset Consisting of 1904 Data Points.

The relationship between NDRE and LCI remained strongly positive, indicating that both vegetation indices are consistent and reliable proxies of plant health status. After relabeling, the clustering patterns in the scatterplot became more distinct, facilitating subsequent analyses and supporting more accurate

decision-making for monitoring and managing rubber plant health. This result is consistent with previous remote-sensing studies showing that chlorophyll-sensitive indices are able to capture vegetation stress responses before symptoms become fully visible at the canopy level [44], [45].

Detailing the calculation stages of the multispectral image-based vegetation index (Table 1), such as NDVI, MCARI, CI, LCI, NDRE, GCI, MCARI_CI_Ratio, NDVI_NDRE_Interaction, and GCI_Ratio, which are used to quantitatively represent the physiological condition of rubber trees. This process includes

extracting spectral values from relevant image channels, normalizing reflectance to reduce the influence of lighting differences and environmental conditions, and adjusting the index value scale to ensure consistency between images. The resulting vegetation indices serve as indicators of plant health and stress levels due to *Pestalotiopsis* sp. infection, and are important features in supporting further analysis and CNN-based disease severity classification modeling. Thus, this stage not only generated predictor variables, but also established the physiological basis for the subsequent modeling results and their interpretation.

Table 1. Vegetation Index Calculation

No	Index Abbreviation	Index Full Form	Index Use	Index Formula
1	NDVI	Normalized Difference Vegetation Index	Measuring vegetation density and health	$(\text{NIR} - \text{Red}) / (\text{NIR} + \text{Red})$
2	MCARI	Modified Chlorophyll Absorption Ratio Index	Detecting chlorophyll content by considering the soil background	$((\text{Red Edge} - \text{Red}) - 0.2 \times (\text{Red Edge} - \text{Green})) \times (\text{Red Edge} / \text{Red})$
3	CI	Chlorophyll Index	Estimating chlorophyll content in leaves	$(\text{NIR} / \text{Red}) - 1$
4	LCI	Leaf Chlorophyll Index	Sensitive and precise estimation of leaf chlorophyll content by taking into account spectral differences in the red and red edge bands	$(\text{NIR} / \text{Red Edge}) - 1$
5	NDRE	Normalized Difference Red Edge	More sensitive to chlorophyll changes than NDVI	$(\text{NIR} - \text{Red Edge}) / (\text{NIR} + \text{Red Edge})$
6	GCI	Green Chlorophyll Index	Measuring chlorophyll content with green band reflectance	$(\text{NIR} / \text{Green}) - 1$
7	MCARI_CI_Ratio	MCARI and CI Ratio	Analyzing the relationship between MCARI and CI	MCARI / CI
8	NDVI_NDRE_Interaction	NDVI and NDRE Interaction	Exploring the interaction between NDVI and NDRE	$\text{NDVI} \times \text{NDRE}$
9	GCI_RATIO	GCI and CI Ratio	Ratio for comparing GCI and CI indices to detect chlorophyll content gradation	GCI / CI

3.2 Feature Selection

Feature selection results (Figure 3) indicate that the model predominantly relies on indices representing chlorophyll content and canopy condition, with NDRE

(0.249), LCI (0.227), and CI (0.196) emerging as the three most influential features, followed by NDVI_NDRE_Interaction (0.115) and GCI_RATIO (0.047).

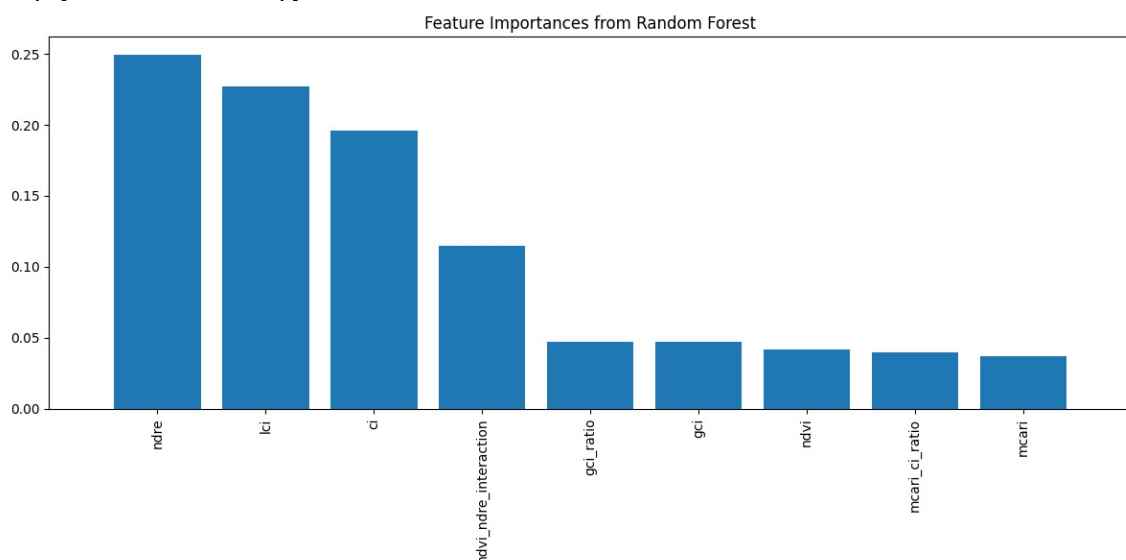


Figure 3. Feature Importances Graph

Collectively, these five indices account for approximately 85% of the total feature importance, suggesting that model decisions are largely driven by

red-edge/NIR signals that are sensitive to chlorophyll depletion and declining canopy vigor. From a physiological perspective, this pattern is expected

because changes in chlorophyll typically precede visually apparent symptoms, while GCI_RATIO helps stabilize chlorophyll-related information derived from the green band and the NDVI_NDRE_Interaction enhances the discrimination between adjacent severity levels. These findings further emphasize the importance of feature selection, as not all vegetation indices contribute equally to representing plant stress responses [46-48]. This result also provides empirical support for the study hypothesis that the most informative predictors would be those directly associated with chlorophyll degradation and canopy stress. Compared with broader vegetation-monitoring studies that often retain a larger number of spectral descriptors, the present result shows that a compact subset of physiologically relevant indices can still capture most of the discriminative information needed for severity classification. In that sense, the contribution of feature selection in this study is not merely computational efficiency, but also interpretability: it clarifies which spectral responses matter most in separating severity levels.

3.3 Exploratory Data Analysis (EDA)

Exploratory Data Analysis (EDA) was conducted to verify that the multispectral features used are genuinely associated with the severity of *Pestalotiopsis* sp. infection in rubber plants. The boxplots reveal a clear physiological pattern: NDRE and three additional vegetation indices (LCI, CI, and NDVI_NDRE_Interaction) decrease progressively from severity level 1 to level 4, indicating declining chlorophyll content and canopy vigor as disease severity increases (Figure 4).

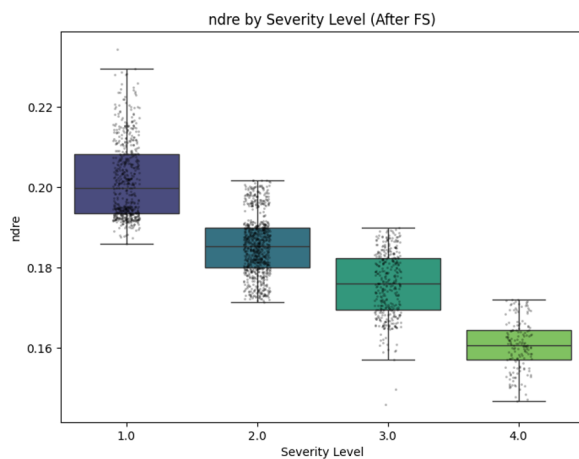


Figure 4. Boxplot distribution: NDRE

In contrast, GCI_Ratio shows an increasing tendency, serving as a complementary indicator when chlorophyll-related indices begin to saturate (Figure 5). The correlation heatmap further confirms a strong negative relationship between the selected features and severity, while also highlighting substantial inter-feature collinearity (e.g., NDRE and CI are nearly identical), which motivates feature selection to reduce redundancy (Figure 6). Finally, the pairplot suggests

that levels 1 and 4 are relatively well separated, whereas levels 2 and 3 overlap most strongly, implying that incorporating ratio and interaction-based features is important to sharpen decision boundaries between adjacent classes (Figure 7).

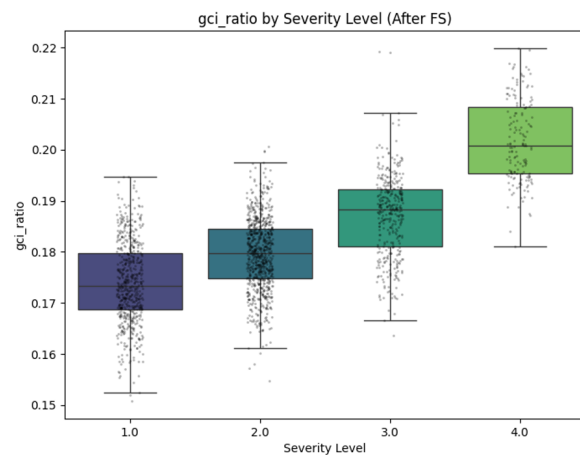


Figure 5. Boxplot distribution: GCI_Ratio

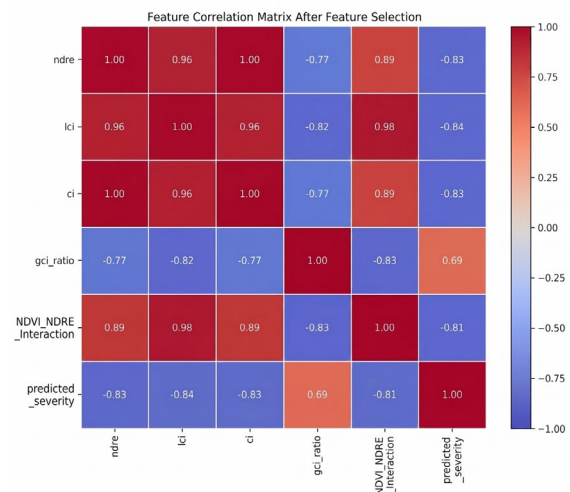


Figure 6. EDA Result: Correlation Heatmap

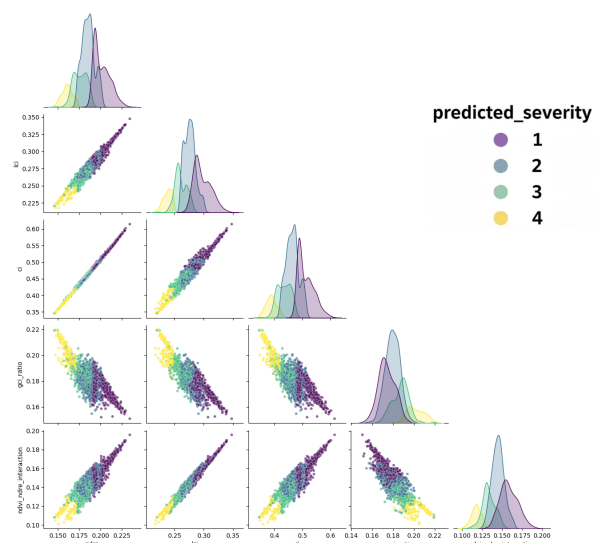


Figure 7. EDA Result: Pairplot predicted

These EDA results are important for the discussion because they explain, before model fitting, why some classes are expected to be easier to classify than others. The clear separation of L1 and L4 is consistent with the later confusion-matrix results, whereas the overlap between L2 and L3 anticipates the main classification difficulty observed in the baseline model. Therefore, the EDA stage does not simply describe the data; it provides an analytical bridge between the physiological behavior of the indices and the classification outcomes reported in the following subsections.

3.4 Handling Imbalanced Classes with SMOTE

Class imbalance was consistently observed across all folds, with L2 dominating at approximately 626–627 samples (41.1%), followed by L1 with about 474–475 samples (31.1–31.2%), while L3 accounted for only 293–294 samples (19.2–19.3%) and L4 was the least represented with 128–129 samples (8.4–8.5%). This distribution poses a risk of bias toward predicting L1 and L2. To address this issue, SMOTE was applied within each fold to balance all classes to match the majority class, resulting in 626 or 627 samples per class (25% each for L1, L2, L3, and L4). Consequently, the total number of samples increased to 2,504 for folds 1–3 and 2,508 for folds 4–5 (Figure 8). Unlike simple data duplication, SMOTE generates synthetic samples via nearest-neighbor interpolation in the vegetation-index feature space, primarily enriching L3 and L4. Potential overfitting was mitigated through RF-based feature selection and 5-fold cross-validation, ensuring robust generalization.

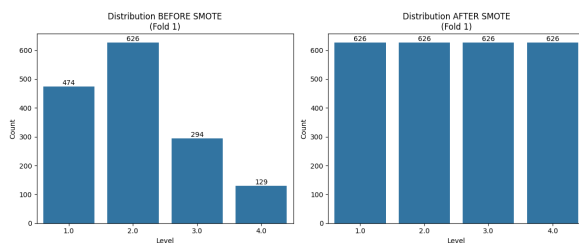


Figure 8. Distribution of severity level data using SMOTE

From a discussion perspective, the use of SMOTE is relevant because the study does not only seek high aggregate accuracy, but also balanced recognition across severity levels. Previous classification studies have noted that minority disease classes are often overlooked when models are trained on skewed distributions, leading to apparently good global performance but weak operational usefulness. The balanced distribution obtained after SMOTE therefore supports the initial expectation that class-sensitive performance, particularly for L3 and L4, could be improved when the model is exposed to more representative training patterns.

3.5 Determination of Disease Severity (CNN 1-D)

The stage of determining the severity level of *Pestalotiopsis* sp. infection was carried out using a 1-D CNN-based deep learning model (Figure 9) that

processes a vector of selected vegetation index features. The adopted 1-D CNN architecture begins with an input layer of size (None, 5, 32), followed by the first Conv1D layer producing an output of (None, 2, 32). A MaxPooling1D layer is then applied without changing the output dimensions, maintaining (None, 2, 32), followed by a Dropout layer for regularization. Subsequently, a second Conv1D layer increases the number of feature channels to 64, yielding an output of (None, 2, 64), and is again followed by Dropout. The extracted feature representation is flattened into a vector of size (None, 128), passed through a Dense layer with 128 neurons and Dropout, and finally terminated by a Dense layer with four neurons (output (None, 4)) for classification into four severity levels.

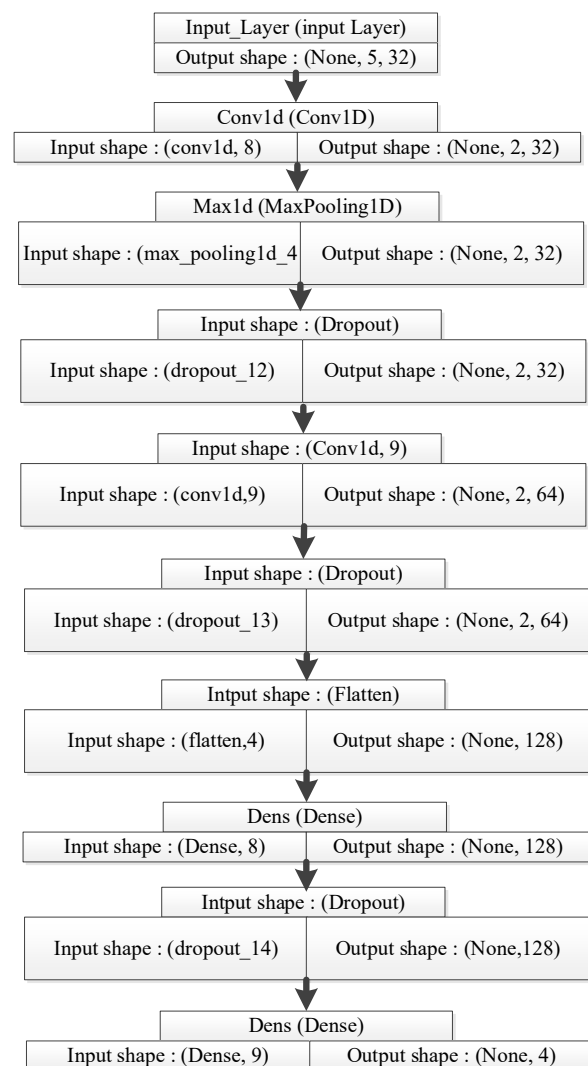


Figure 9. CNN Architecture

Based on Table 2, the 1-D CNN model demonstrates strong and well-balanced overall performance, as reflected by an accuracy of 85.26% and closely matched precision (86.53%) and recall (86.52%). The similarity between precision and recall indicates that the model does not favor a particular prediction strategy, but instead maintains a balance between false positives and false negatives in severity classification. The F1-score

of 86.41%, which is consistent with these metrics, further confirms stable predictive performance across classes. Nevertheless, because disease severity assessment is a multi-class problem with different operational implications for each level, these global metrics should be interpreted in conjunction with the confusion matrix to better understand misclassification patterns, particularly between adjacent severity levels with overlapping symptoms. In relation to the study hypothesis, these results indicate that vegetation-index features processed by a 1-D CNN are sufficiently informative to support robust baseline classification, although not yet equally strong for all severity categories.

Table 2. Evaluation of 1-D CNN Model Performance

Indicator	Value (%)
Accuracy	85.26
Precision	86.53
Recall	86.52
F1-score	86.41

3.6 Hyperparameter Tuning

Hyperparameter tuning (HPT) produces strong and relatively consistent performance across rounds, as

shown in the comparison graph of all rounds (Figure 10). The best validation accuracy per round is 0.859, Fold 2 = 0.859, Fold 3 = 0.756, Fold 4 = 0.870, and Fold 5 = 0.922. The best validation loss per fold is 0.521, Fold 2 = 0.460, Fold 3 = 0.687, Fold 4 = 0.460, and Fold 5 = 0.356. These values indicate that the adjustment not only improves peak accuracy (Fold 5: 0.922 as the highest achievement), but also significantly reduces loss on most folds (Fold 5: 0.356, as well as Folds 2 and 4 around 0.46), indicating better convergence and more stable generalization. The only exception is Fold 3, which remains the most challenging (accuracy 0.756, loss 0.687), resulting in inter-fold variance.

This pattern suggests that the model is responsive to optimization, but not uniformly across all data partitions. The remaining variance is important to acknowledge because it indicates that the proposed framework is promising rather than universally stable under every fold condition. Therefore, the discussion should not overstate the tuning effect; instead, the evidence supports a more careful conclusion that hyperparameter adjustment improves model reliability overall, while some folds still contain more complex spectral overlap or boundary ambiguity than others.



Figure 10. Performance Comparison Across All Folds Before and After Hyperparameter Tuning

3.7 Model Evaluation

The vegetation index vector with inputs (None, 5, 32), which indicates five sequences of 32-dimensional spectral features per sample, is used to classify the

severity of *Pestalotiopsis* sp. disease using the CNN structure shown in Figure 9. When the data faced an imbalance in chlorophyll patterns between severity levels, feature extraction was performed in stages through Conv1D, MaxPooling1D, and dropout. This

also produced initial spatial-spectral patterns. In the next step, a second Conv1D increases the channel depth from 32 to 64 to make the model more sensitive to subtle symptoms such as leaf darkening, edge discoloration, and small necrotic areas. In addition, there is an additional dropout to maintain robustness against lighting noise, soil background residue, and canopy overlap.

After that, the feature map is summarized through Flatten into a 128-dimensional vector, then processed by the Dense layer to form stable multi-class decision boundaries between folds. The first Dense (128 units) is followed by dropout to reduce the dominance of light classes, while the final Dense produces an output (None, 4) as the probability of four severity classes (1-4). Overall, the 32, 64, 128 channel flow forms a hierarchy of abstraction from healthy vegetation, chlorophyll depletion, increased necrosis, to severe damage, so that the physiological progression of rubber leaves can be mapped quantitatively and consistently. This interpretation is aligned with the initial assumption that hierarchical nonlinear learning can translate vegetation-index patterns into operational severity categories.

The 1-D CNN model has good and balanced classification performance, as shown in Table 2. It shows that the model recognizes positive classes consistently without significant bias to false positives or false negatives, with nearly identical precision (86.53%), recall (86.52%), and F1-score values. Although the accuracy rate of 85.26% is slightly lower than other metrics, this model is still suitable for severity mapping. However, to ensure that minority classes are not overlooked, per-class evaluation (such as a confusion matrix) is still necessary. This need for class-wise interpretation is also emphasized in related evaluation studies, which argue that aggregate metrics alone may conceal weaknesses in neighboring or minority classes [49], [50].

Figure 11 illustrates the learning dynamics of the CNN in mapping the severity of *Pestalotiopsis* sp. infection in rubber plants using selected spectral features. Training accuracy increases rapidly from approximately 0.55 to 0.75–0.78 within the first 3–5 epochs, while validation accuracy follows a similar trend, rising to about 0.74–0.77, indicating that red-edge and chlorophyll-related indices (NDRE, LCI, CI) together with the composite NDVI_NDRE_Interaction provide informative physiological signals from the early stages of training. Subsequent improvements occur more gradually, reaching around 0.80–0.81 for training and 0.79–0.80 for validation by approximately epoch 25, reflecting the refinement of decision boundaries between spectrally overlapping severity levels. By epochs 100–150, performance stabilizes, with training accuracy around 0.84–0.86 and validation accuracy around 0.82–0.85, and a small gap of roughly 0.01–0.03, suggesting limited overfitting and good generalization across folds. Toward the end of training (around epoch 200), the

curves plateau, implying marginal gains from additional epochs, while GCI_Ratio acts as a complementary feature that helps discriminate higher severity levels when red-edge responses begin to saturate. In discussion terms, this pattern indicates that the model learns meaningful class structure early and then refines class boundaries incrementally, which is consistent with the observed challenge in separating intermediate severity levels rather than extreme ones.

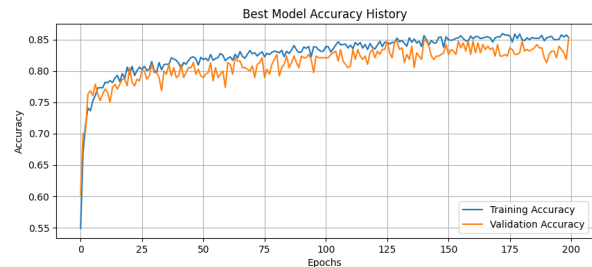


Figure 11. Best Model Accuracy

Before hyperparameter tuning (Figure 12), the confusion matrix indicates that the baseline model was relatively stable for the extreme classes but less consistent for intermediate severity levels. Class L1 was correctly identified in 109 of 119 samples, with most errors shifting to L2 (10 cases). Class L2 achieved 132 correct predictions out of 156, but misclassification was split toward L1 (11) and L3 (13), suggesting overlap in decision boundaries. The largest weakness appeared in L3, with 51 correct predictions out of 73 and substantial confusion into L2 (20), indicating limited separability for adjacent moderate-to-severe symptoms. In contrast, L4 showed perfect recognition (32/32), reflecting strong class distinctiveness at the highest severity stage. Overall, the error structure confirms that model limitations are concentrated in neighboring classes (especially L2-L3), which is consistent with prior evaluation literature emphasizing class-wise confusion analysis rather than relying solely on aggregate accuracy metrics [49], [50].

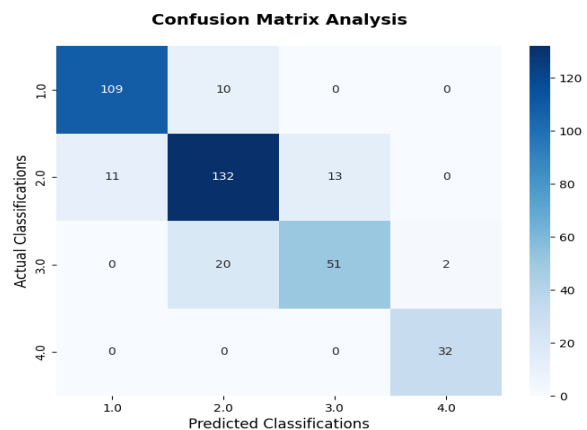


Figure 12. Confusion Matrix before HPT

In the optimized model (Figure 13), the confusion matrix shows a clear improvement in severity-level separability, with most predictions concentrated along the main diagonal: L1 was correctly classified in 33 of

35 samples, L2 in 13 of 15, L3 in 15 of 16, and L4 in 8 of 11. The remaining errors were predominantly local and occurred between adjacent levels, including L1 to L2 (2 cases), L2 to L1 (2 cases), L3 to L2 (1 case), and L4 to L3 (3 cases), with no evidence of extreme jumps across nonadjacent levels. This error structure indicates that parameter optimization improved decision consistency across low-to-moderate severity levels, while the main challenge remains at the L3-L4 boundary, where symptom expression and spectral response are often partially overlapping. Compared with the baseline expectation that tuning would mainly improve general performance, the results show a more specific pattern: the largest benefit occurs in the previously weak intermediate-to-severe transition, rather than uniformly across all classes.

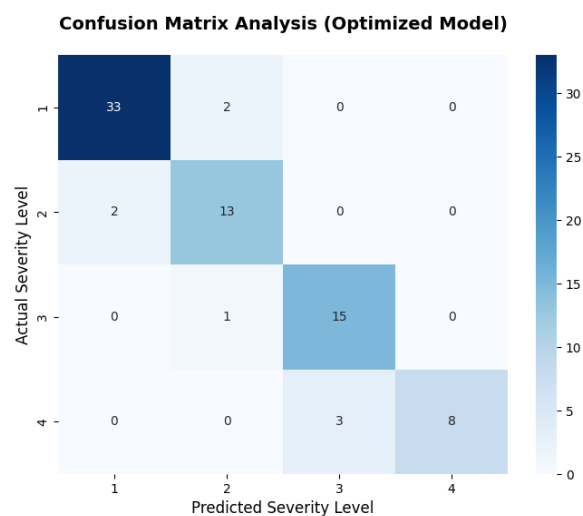


Figure 13. Confusion Matrix After HPT

The impact is most evident in the weaker levels. The largest improvement appears in L3: recall increased from 0.699 to 0.94, F1-score rose from 0.745 to 0.88, while precision showed a modest increase (0.797 to 0.83). For L4, precision also improved (0.941 to 1.00), but recall declined (1.00 to 0.73), resulting in a lower F1-score (0.970 to 0.84), indicating that the model became more conservative when assigning L4 labels (Table 3).

Table 3. Accuracy after Hyper Parameter Tuning

Metric	L1	L2	L3	L4
Precision	0.94	0.81	0.83	1.00
Recall	0.94	0.87	0.94	0.73
F1-score	0.94	0.84	0.88	0.84
Accuracy				0.90

For L1, precision improved (0.908 to 0.94) and recall also increased (0.916 to 0.94), yielding a higher F1-score (0.912 to 0.94). For L2, the improvement was moderate and consistent: recall increased from 0.846 to 0.87 and F1-score from 0.830 to 0.84, although precision slightly decreased (0.815 to 0.81). Overall, hyperparameter tuning produced a more favorable performance profile: total accuracy improved (0.8530 to 0.90), previously under-detected critical levels became substantially better detected (especially L3),

and severe-case assignment became stricter through perfect precision in L4. This trade-off can be operationally beneficial when reducing false alarms is prioritized, but L4 recall still requires improvement to minimize missed very severe cases (Table 4). In relation to the broader implications of the study, these results suggest that the proposed pipeline is particularly valuable for improving discrimination in severity levels that are physiologically adjacent and operationally difficult to separate. At the same time, the decline in L4 recall cautions against overgeneralization: the model is effective, but its deployment for high-stakes surveillance should still consider the risk of missed very severe cases and may benefit from additional samples or calibration strategies in future work.

Table 4. Accuracy before Hyper Parameter Tuning

Metric	L1	L2	L3	L4
Precision	0.908	0.815	0.797	0.941
Recall	0.916	0.846	0.699	1.000
F1-score	0.912	0.830	0.745	0.970
Accuracy				0.8530

4. Conclusions

This study demonstrates that UAV multispectral vegetation indices can do more than describe canopy stress; when organized as physiologically meaningful predictors and learned through a one-dimensional CNN, they can support practical severity discrimination of *Pestalotiopsis* sp. in rubber across four operational classes: L1 (Light Infection; SI = 1), L2 (Moderate Infection; SI = 2), L3 (Severe Infection; SI = 3), and L4 (Very Severe Infection; SI = 4). In this sense, the main contribution of the study is not merely the use of MVIs, but the demonstration that chlorophyll-sensitive Red Edge and NIR information can be translated into a structured severity-classification framework that is relevant for plantation monitoring and disease-control decisions.

The best-performing model was obtained after hyperparameter sampling/tuning, achieving 90.00% overall accuracy and improving upon the pre-tuning baseline (85.30%). The most meaningful gain occurred in L3, where recall increased from 0.699 to 0.940 and F1-score from 0.745 to 0.880, indicating that the proposed approach is particularly valuable in detecting intermediate-to-severe stress patterns that are often difficult to separate under field conditions. L1 and L2 also improved (F1: 0.912 to 0.940 and 0.830 to 0.840, respectively), while L4 became more conservative, with precision reaching 1.00 but lower recall (1.00 to 0.730), reflecting a stricter decision boundary after tuning. These results suggest that the method is already useful for supporting more objective surveillance and targeted intervention, but they also leave an important practical question for the field: improving the sensitivity of very severe-case detection without sacrificing the precision gained through optimization.

Overall, the findings indicate that chlorophyll-driven indices, especially those derived from the Red Edge and

NIR bands, are not only dominant predictors of disease severity but also meaningful variables for building scalable decision-support tools in crop health monitoring. For the research community, this study contributes evidence that physiologically informed spectral features and tuned CNN learning can bridge the gap between spectral detection and operational severity mapping. For plantation management, the approach has potential applications in early warning, block-level prioritization, and precision intervention. Future development should therefore focus on increasing representation of very severe cases, testing model transferability across broader environments and rubber clones, and integrating the classification pipeline into real-time monitoring systems so that its value extends from methodological accuracy to actionable field deployment.

Acknowledgements

The author would like to thank the Institute for Research and Community Service (LPPM), IPB University; the Forest Resource Conservation and Ecotourism (KSHE) Laboratory, IPB University; and the Rubber Research Center, Sembawa, Banyuasin, South Sumatra, Indonesia, for their support in this study.

References

- [1] P. Alchemi and S. Jamin, "Impact of pestalotiopsis leaf fall disease on leaf area index and rubber plant production," *IOP Conf. Ser. Earth Environ. Sci.*, vol. 995, no. 1, p. 12030, 2022, doi:10.1088/1755-1315/995/1/012030.
- [2] N. Damiri, Y. Pratama, T. R. Febbiyanti, S. E. Rahim, D. T. Astuti, and Y. Purwanti, "Pestalotiopsis sp. infection causes leaf fall disease of new arrivals in several clones of rubber plants," *Biodiversitas*, vol. 23, no. 8, pp. 3943–3949, 2022, doi:10.13057/biodiv/d230811.
- [3] M. R. Darajat, S. W. Ardie, F. Oktavia, and Sudarsono, "New leaf fall disease in rubber-pathogen characterization and rubber clone resistance evaluation using detached leaf assay," *Biodiversitas*, vol. 24, no. 4, pp. 1935–1945, 2023, doi:10.13057/biodiv/d240401.
- [4] P. Fitria, I. A. Simangunsong, H. Handoko, Nurliana, and F. A. Barus, "Studi Epidemiologi dan Pola Penyebaran Penyakit Gugur Daun Karet (Pestalotiopsis sp.) Pada Tanaman Karet (Hevea brasiliensis)," *J. Agro Estate*, vol. 8, no. 1, pp. 53–66, 2024, doi:10.47199/jae.v8i1.245.
- [5] Y. Herdiyeni, M. F. Mumtaz, G. F. Laxmi, Y. Setiawan, L. B. Prasetyo, and T. R. Febbiyanti, "Analysis and prediction of rubber tree phenological changes during Pestalotiopsis infection using Sentinel-2 imagery and random forest," *J. Appl. Remote Sens.*, vol. 18, no. 01, pp. 1–17, 2024, doi:10.1117/1.jrs.18.014524.
- [6] P. Han, J. Chen, Y. Han, L. Yi, Y. Zhang, and X. Jiang, "Monitoring rubber plantation distribution on Hainan Island using Landsat OLI imagery," *Int. J. Remote Sens.*, vol. 39, no. 8, pp. 2189–2206, 2018, doi:10.1080/01431161.2017.1420933.
- [7] Y. Liang et al., "Rubber Tree Recognition Based on UAV RGB Multi-Angle Imagery and Deep Learning," *Drones*, vol. 7, no. 9, p. 547, 2023, doi:10.3390/drones7090547.
- [8] N. A. M. Sopian, Supriatna, M. D. M. Manessa, I. P. A. Shidiq, R. Nagasawa, and M. Haidar, "A study on spatio-temporal trend of rubber leaf fall phenomenon using planetscope multi-index vegetation imagery in relations to climatological conditions," *Environ. Mater.*, vol. 2, no. 1, pp. 45–60, 2024, doi:10.61511/eam.v2i1.2024.906.
- [9] A. J. A. Salamanca, R. M. Navarro-Cerrillo, F. J. Bonet-García, M. J. Pérez-Palazón, and M. J. Polo, "Integration of a Landsat time-series of NBR and hydrological modeling to assess Pinus pinaster Aiton. forest defoliation in south-eastern Spain," *Remote Sens.*, vol. 11, no. 19, 2019, doi:10.3390/rs11192291.
- [10] K. Taji, A. Sohail, T. Shahzad, B. S. Khan, M. A. Khan, and K. Ouahada, "An Ensemble Hybrid Framework: A Comparative Analysis of Metaheuristic Algorithms for Ensemble Hybrid CNN Features for Plants Disease Classification," *IEEE Access*, vol. 12, pp. 61886–61906, 2024, doi:10.1109/ACCESS.2024.3389648.
- [11] R. Xu, C. Li, and S. Bernardes, "Development and testing of a uav-based multi-sensor system for plant phenotyping and precision agriculture," *Remote Sens.*, vol. 13, no. 17, p. 3517, 2021, doi:10.3390/rs13173517.
- [12] L. Daniels, E. Eeckhout, J. Wieme, Y. Dejaegher, K. Audenaert, and W. H. Maes, "Identifying the Optimal Radiometric Calibration Method for UAV-Based Multispectral Imaging," *Remote Sens.*, vol. 15, no. 11, p. 2909, 2023, doi:10.3390/rs15112909.
- [13] A. Singh, K. Gaurav, G. K. Sonkar, and C. C. Lee, "Strategies to Measure Soil Moisture Using Traditional Methods, Automated Sensors, Remote Sensing, and Machine Learning Techniques: Review, Bibliometric Analysis, Applications, Research Findings, and Future Directions," *IEEE Access*, vol. 11, pp. 13605–13635, 2023, doi:10.1109/ACCESS.2023.3243635.
- [14] G. Meraj et al., "Assessing the Yield of Wheat Using Satellite Remote Sensing-Based Machine Learning Algorithms and Simulation Modeling," *Remote Sens.*, vol. 14, no. 13, p. 3005, 2022, doi:10.3390/rs14133005.
- [15] H. Feng, J. Chen, Z. Zhang, Y. Lou, S. Zhang, and W. Yang, "A bibliometric analysis of artificial intelligence applications in macular edema: exploring research hotspots and Frontiers," *Front. Cell Dev. Biol.*, vol. 11, 2023, doi:10.3389/fcell.2023.1174936.
- [16] P. Saxena, S. K. Singh, M. Rashid, S. S. Alshamrani, and M. M. Alnfiar, "Efficient deep learning model for classifying lung cancer images using normalized stain agnostic feature method and FastAI-2," *PeerJ Comput. Sci.*, vol. 11, pp. 1–22, 2025, doi:10.7717/peerj-cs.2903.
- [17] Y. Bin, R. Peng, Y. Lee, Z. Lee, and Y. Liu, "Artificial intelligence-assisted capsule endoscopy for detecting lesions in Crohn's disease: a systematic review and meta-analysis," *Front. Artif. Intell.*, vol. 8, 2025, doi:10.3389/frai.2025.1531362.
- [18] H. M. Abubakar et al., "2022 Flood Impact in Pakistan: Remote Sensing Assessment of Agricultural and Urban Damage," in *Proceedings of the AAAI Symposium Series*, 2024, vol. 4, no. 1, pp. 405–410, doi:10.1609/aaais.v4i1.31824.
- [19] R. A. J. Alhatemi and S. Savaş, "A Weighted Ensemble Approach with Multiple Pre-trained Deep Learning Models for Classification of Stroke," *Medinformatics*, vol. 1, no. 1, pp. 10–19, 2023, doi:10.47852/bonviewmedin32021963.
- [20] Z. Song, W. Li, X. Song, and S. Li, "FGA-Corn: an integrated system for precision pesticide application in center leaf areas using deep learning vision," *Front. Plant Sci.*, vol. 16, 2025, doi:10.3389/fpls.2025.1571228.
- [21] J. Cao et al., "Multi-Source Forecast of Solar Cycle Flare Activity Using the Novel Informer-Based Models," *Sp. Weather*, vol. 23, no. 4, 2025, doi:10.1029/2024SW004322.
- [22] S. Ruksakulpiwat et al., "Artificial Intelligence in Nursing Research: A Systematic Review of Applications, Benefits, and Challenges," *Int. Nurs. Rev.*, vol. 72, no. 3, 2025, doi:10.1111/inr.70080.
- [23] M. A. Ariful Furqon, N. F. Najwa, M. Zarkasi, P. Pandunata, and G. W. Fajariyanto, "Critical Exploratory Data Analysis on Stroke Prediction Dataset," *J. Komput. Terap.*, vol. 10, no. 1, pp. 67–77, 2024, doi:10.35143/jkt.v10i1.6307.
- [24] P. A. Riyantoko, T. M. Fahrudin, K. M. Hindrayani, and M. Idhom, "Exploratory Data Analysis and Machine Learning Algorithms to Classifying Stroke Disease," *Ijconsist Journals*, vol. 2, no. 02, pp. 77–82, 2021, doi:10.33005/ijconsist.v2i02.49.
- [25] S. Dhummad, "The Imperative of Exploratory Data Analysis in Machine Learning," *Sch. J. Eng. Technol.*, vol. 13, no. 01, pp. 30–44, 2025, doi:10.36347/sjet.2025.v13i01.005.

- [26] I. R. Saladdin and P. W. Handayani, "Information Technology Governance Implementation Challenges in Healthcare Facilities: A Systematic Literature Review," *SAGE Open*, vol. 15, no. 4, 2025, doi: [10.1177/21582440251369322](https://doi.org/10.1177/21582440251369322).
- [27] F. Yulianto et al., "Spatial and temporal distribution of estimated surface runoff caused by land use/land cover changes in the upstream Citarum watershed, West Java, Indonesia," *J. Degrad. Min. Lands Manag.*, vol. 9, no. 2, pp. 3293–3305, 2022, doi: [10.15243/jdmlm.2022.092.3293](https://doi.org/10.15243/jdmlm.2022.092.3293).
- [28] W. K. Al-Zoubi, "How Sustainable is Environmental Economics? A Review of Research Trends and Implications," *Int. J. Energy Econ. Policy*, vol. 14, no. 2, pp. 319–334, 2024, doi: [10.32479/ijeep.15494](https://doi.org/10.32479/ijeep.15494).
- [29] M. N. A. Birton, M. Maryati, and M. Muttaqin, "Do conferences drive quality improvement in accounting research?," *J. Account. Invest.*, vol. 24, no. 3, pp. 737–756, 2023, doi: [10.18196/jai.v24i3.19045](https://doi.org/10.18196/jai.v24i3.19045).
- [30] X. Li et al., "Morpho-phylogenetic evidence reveals two new species and two new host records of Diaporthe (Diaporthales, Diaporthaceae) from Phellodendron chinense in Sichuan, China," *MycKeys*, vol. 123, pp. 205–233, 2025, doi: [10.3897/mycokeys.123.162866](https://doi.org/10.3897/mycokeys.123.162866).
- [31] D. Jiménez-Islas, M. E. Pérez-Romero, J. García, and A. D. Sánchez, "Scientific Research Trend in Biorefineries in India: Analysis and Systematic Review," *Int. J. Des. Nat. Ecodynamics*, vol. 17, no. 5, pp. 721–727, 2022, doi: [10.18280/ijdne.170509](https://doi.org/10.18280/ijdne.170509).
- [32] S. Karthikeyan, R. Charan, S. Narayanan, and L. Jani Anbarasi, "Enhanced plant disease classification with attention-based convolutional neural network using squeeze and excitation mechanism," *Front. Artif. Intell.*, vol. 8, 2025, doi: [10.3389/frai.2025.1640549](https://doi.org/10.3389/frai.2025.1640549).
- [33] S. Rao, Z. Zeng, and J. Zhang, "Robust Multiclass Pneumonia Classification via Multi-Head Attention and Transfer Learning Ensemble," *Appl. Sci.*, vol. 15, no. 21, p. 11426, 2025, doi: [10.3390/app152111426](https://doi.org/10.3390/app152111426).
- [34] A. González-Briones, S. L. Florez, P. Chamoso, L. F. Castillo-Ossa, and E. S. Corchado, "Enhancing Plant Disease Detection: Incorporating Advanced CNN Architectures for Better Accuracy and Interpretability," *Int. J. Comput. Intell. Syst.*, vol. 18, no. 1, 2025, doi: [10.1007/s44196-025-00835-2](https://doi.org/10.1007/s44196-025-00835-2).
- [35] X. Xu, Y. Li, and X. Ding, "Combined ResNet Attention Multi-Head Net (CRAMNet): A Novel Approach to Fault Diagnosis of Rolling Bearings Using Acoustic Radiation Signals and Advanced Deep Learning Techniques," *Appl. Sci.*, vol. 14, no. 18, p. 8431, 2024, doi: [10.3390/app14188431](https://doi.org/10.3390/app14188431).
- [36] S. Suseela and N. Parekh, "From pixels to prognosis: Attention-CNN model for COVID-19 diagnosis using chest CT images," *IET Image Process.*, vol. 19, no. 1, 2025, doi: [10.1049/ipr2.13249](https://doi.org/10.1049/ipr2.13249).
- [37] B. Nebili, A. Khellal, A. Nemra, and L. Mascarilla, "Augmented Convolutional Neural Network Models with Relative Multi-Head Attention for Target Recognition in Infrared Images," *Unmanned Syst.*, vol. 11, no. 3, pp. 221–230, 2023, doi: [10.1142/S2301385023500085](https://doi.org/10.1142/S2301385023500085).
- [38] M. Alezzi, "Multi-class classification framework for ear imaging using attention-based hybrid dense CNN and non-linear manifold approximation," *Samarra J. Eng. Sci. Res.*, vol. 3, no. 2, pp. 1–23, 2025, doi: [10.65115/7dnryk35](https://doi.org/10.65115/7dnryk35).
- [39] L. S. Chow, G. S. Tang, M. I. Solihin, N. M. Gowdh, N. Ramli, and K. Rahmat, "Quantitative and Qualitative Analysis of 18 Deep Convolutional Neural Network (CNN) Models with Transfer Learning to Diagnose COVID-19 on Chest X-Ray (CXR) Images," *SN Comput. Sci.*, vol. 4, no. 2, 2023, doi: [10.1007/s42979-022-01545-8](https://doi.org/10.1007/s42979-022-01545-8).
- [40] M. M. Sardeshmukh, M. Chakkaravarthy, S. Shinde, and D. Chakkaravarthy, "Crop image classification using convolutional neural network," *Multidiscip. Sci. J.*, vol. 5, no. 4, p. 2023039, 2023, doi: [10.31893/multiscience.2023039](https://doi.org/10.31893/multiscience.2023039).
- [41] S. Shrivastav, V. Jindal, and R. Eswarawaka, "A Fusion Method for Detection and Classification of Diseases in Tomato Plants Using Swarm-based Deep Learning," *Int. J. Exp. Res. Rev.*, vol. 45, no. Special Issue, pp. 135–152, 2024, doi: [10.52756/ijerr.2024.v45spl.011](https://doi.org/10.52756/ijerr.2024.v45spl.011).
- [42] K. C. G. Ganashree, R. Hemavathy, and M. R. Anala, "Land scene classification from remote sensing images using improved artificial bee colony optimization algorithm," *Int. J. Electr. Comput. Eng.*, vol. 14, no. 1, pp. 347–357, 2024, doi: [10.11591/ijece.v14i1.pp347-357](https://doi.org/10.11591/ijece.v14i1.pp347-357).
- [43] M. Iqbal and Y. Herdiyeni, "Pemodelan Keparahan Penyakit Tanaman Karet Berbasis Citra Multispektral Drone Menggunakan Algoritme Random Forest," Institut Pertanian Bogor, 2023, <https://repository.ipb.ac.id/handle/123456789/124008>.
- [44] D. Haboudane, J. R. Miller, E. Pattey, P. J. Zarco-Tejada, and I. B. Strachan, "Hyperspectral vegetation indices and novel algorithms for predicting green LAI of crop canopies: Modeling and validation in the context of precision agriculture," *Remote Sens. Environ.*, vol. 90, no. 3, pp. 337–352, 2004, doi: [10.1016/j.rse.2003.12.013](https://doi.org/10.1016/j.rse.2003.12.013).
- [45] B. Cui, Q. Zhao, W. Huang, X. Song, H. Ye, and X. Zhou, "A new integrated vegetation index for the estimation of winter wheat leaf chlorophyll content," *Remote Sens.*, vol. 11, no. 8, pp. 1–18, 2019, doi: [10.3390/rs11080909](https://doi.org/10.3390/rs11080909).
- [46] R. Saini, "Integrating Vegetation Indices and Spectral Features for Vegetation Mapping from Multispectral Satellite Imagery Using AdaBoost and Random Forest Machine Learning Classifiers," *Geomatics Environ. Eng.*, vol. 17, no. 1, pp. 57–74, 2023, doi: [10.7494/geom.2023.17.1.57](https://doi.org/10.7494/geom.2023.17.1.57).
- [47] E. H. E. Yasin et al., "Crop type classification and spatial mapping in River Nile and Northern State, Sudan, using Sentinel-2 satellite data and field observation," *J. Degrad. Min. Lands Manag.*, vol. 11, no. 3, pp. 5997–6007, 2024, doi: [10.15243/jdmlm.2024.113.5997](https://doi.org/10.15243/jdmlm.2024.113.5997).
- [48] J. Guo et al., "Monitoring of Vegetation Disturbance and Restoration at the Dumping Sites of the Baorixile Open-Pit Mine Based on the LandTrendr Algorithm," *Int. J. Environ. Res. Public Health*, vol. 19, no. 15, p. 9066, 2022, doi: [10.3390/ijerph19159066](https://doi.org/10.3390/ijerph19159066).
- [49] D. M. W. Powers, "Evaluation: from Precision, Recall and F-measure to ROC, Informedness, Markedness and Correlation," *arXiv*, 2020, [Online]. Available: <https://arxiv.org/pdf/2010.16061>.
- [50] D. Chicco and G. Jurman, "The advantages of the Matthews correlation coefficient (MCC) over F1 score and accuracy in binary classification evaluation," *BMC Genomics*, 2021, doi: [10.1186/s13040-021-00244-z](https://doi.org/10.1186/s13040-021-00244-z).