Available online at https://journal.rescollacomm.com/index.php/ijrcs/index

International Journal of Research in Community Service

e-ISSN: 2746-3281
p-ISSN: 2746-3273

Vol. 6, No. 3, pp. 147-155, 2025

Implementation of Machine Learning Model for Pest Classification in Rice Plants

Moch Panji Agung Saputra[1*], Deva Putra Setyawan[2], Muhammad Bintang Eighista Dwiputra[3]

[1] Department of Mathematics, Faculty of Mathematics and Natural Sciences, Universitas Padjadjaran, Sumedang, Indonesia
[2] Master's Program of Mathematics, Faculty of Mathematics and Natural Sciences, Universitas Padjadjaran, Jatinangor, Indonesia
[3] Computer Science Study Program, Faculty of Mathematics and Natural Sciences Education, Universitas Pendidikan Indonesia, Bandung, Indonesia

**Corresponding author email: moch16006@unpad.ac.id*

Abstract

Rice cultivation is a cornerstone of food security in agrarian countries like Indonesia, yet it remains highly vulnerable to pest infestations that can severely impact crop productivity. Manual identification of pests is time-consuming and error-prone, especially when pest species exhibit similar morphological characteristics. This study aims to implement and evaluate the performance of four classical machine learning algorithms Support Vector Machine (SVM), K-Nearest Neighbor (KNN), Random Forest (RF), and Logistic Regression (LR) for classifying rice pests based on image data. The dataset, derived from Kaggle's "Rice Pest Detection Dataset," includes 12 pest classes and underwent a series of preprocessing steps: grayscale conversion, image resizing to 128×128 pixels, feature extraction using Histogram of Oriented Gradients (HOG), label encoding, and class balancing via SMOTE. The experimental setup used 80% of the data for training and 20% for testing. Performance was evaluated using precision, recall, F1-score, and confusion matrices. Among the four models, SVM achieved the most consistent and robust performance, with F1-scores reaching up to 0.98 in several pest classes and an overall balanced classification across the dataset. Random Forest followed closely, particularly excelling in distinguishing classes such as Rice Water Weevil and Yellow Rice Borer, achieving F1-scores of 0.99 and 0.96 respectively. In contrast, KNN showed signs of overfitting, with extreme precision-recall imbalances, while LR was more stable but less accurate in separating visually similar classes like Rice Stem Fly and Thrips. Visual analysis of correct and incorrect predictions revealed that classes 7 (Rice Stem Fly) and 11 (Thrips) were consistently misclassified across all models, likely due to high visual similarity.

Keywords: Classification, machine learning model, rice leaf pests.

1. Introduction

Agriculture is a vital sector in meeting the food needs of the population, especially in an agrarian country like Indonesia. Rice, as a primary crop, plays a crucial role in maintaining national food security. Rice productivity is often hampered by attacks by various pests, which can significantly reduce crop yields. Undetected pest attacks often cause severe damage, necessitating rapid and appropriate detection and management to minimize the negative impact on production (Wang et al., 2025).

Along with the advancement of digital technology, the application of artificial intelligence (AI) and machine learning (ML) in agriculture is growing. This technology offers an efficient and accurate approach to automatically detecting and classifying pests based on image data (Kaushal et al., 2022). By utilizing visual feature extraction techniques such as Histogram of Oriented Gradients (HOG) and machine learning algorithms, the system can recognize visual patterns in images of leaves or plant parts infected with pests. This presents a significant opportunity to reduce reliance on manual detection, which requires expertise and considerable time (Ochango et al., 2022).

Various machine learning algorithms have been developed and used for image classification tasks, including in the context of crop pest detection. Models such as Support Vector Machine (SVM), K-Nearest Neighbors (KNN), Random Forest, and Logistic Regression are some of the commonly used methods due to their ability to recognize patterns in complex data. Each model has its own characteristics and advantages in terms of accuracy, speed, and generalization to new data.

Previous research has shown that the application of machine learning in the classification of plant diseases and pests has attracted the attention of researchers in the field of precision agriculture. Setiawan et al. (2023) studied the application of the Nu-Support Vector Machine (Nu-SVM) algorithm to rice leaf disease classification, specifically in distinguishing between healthy leaves, Brown Spot, and Leaf Blast. By using Hu Moments features and Sobel edge detection on segmented leaf images, this study achieved a moderate level of accuracy (52.12%–53.81%) through 5-fold cross-validation. Although the results indicate challenges in precise classification, this study emphasizes the importance of more sophisticated image processing and feature extraction methods.

Kasinathan and Uyyala (2021) applied a machine vision-based approach to the classification of plant insect pests by combining various feature descriptors such as texture, color, shape, HOG, and GIST. They used several machine learning algorithms, both basic classifiers such as Naive Bayes, SVM, KNN, and MLP, as well as ensembles such as Random Forest, Bagging, and XGBoost. The results of the 10-fold cross-validation test show that the classification accuracy increases significantly when using a combination of features and ensemble methods.

Although machine learning has been widely applied in agriculture, studies specifically comparing various machine learning models for rice pest image classification are still relatively limited. Most studies focus on disease identification or insect classification in general, without emphasizing a comprehensive evaluation of each algorithm's performance in the context of rice pests. Furthermore, visualization of classification results showing which images are correctly and incorrectly classified is still rare, even though this is crucial for understanding model error characteristics and improving system interpretability. Therefore, this study aims to implement and compare several classic machine learning models such as SVM, KNN, Random Forest, and Logistic Regression in image-based rice pest classification.

2. Methodology

2.1. Data Collection

The data used in this study was obtained from the Kaggle platform under the title "Rice Pest Detection Dataset." This dataset is part of the IP102 dataset, specifically filtered for detecting rice pests. The dataset consists of images that have undergone several stages of augmentation and preprocessing to improve data quality and enrich image variety. This dataset comprises 12 classes of rice pests, with the number of images for each class shown in Table 1 below:

Table 1: Number of images per pest class

Label	Pest Name	Number of Images
0	Rice leaf roller	605
1	Rice leaf caterpillar	475
2	Paddy stem maggot	325
3	Asiatic rice borer	745
4	Yellow rice borer	455
5	Rice gall midge	791
6	Brown plant hopper	290
7	Rice stem fly	1110
8	Rice water weevil	1194
9	Rice leaf hopper	686
10	Rice shell pest	480
11	Thrips	580

2.2. Data Preprocessing

Before the data was used in the machine learning model training process, several preprocessing steps were performed to ensure the quality and uniformity of the image data used. The preprocessing steps implemented in this study included:

1) Image Conversion to Grayscale
   Each image was converted to grayscale format to simplify visual information by reducing the color dimension. This allowed the model to focus more on texture and shape patterns relevant for pest detection.
2) Image Resizing to 128x128 Pixels
   The images were then resized to 128x128 pixels so that all images had uniform dimensions and could be processed consistently by the machine learning model (Sundhar et al., 2025).
3) Feature Extraction Using Histogram of Oriented Gradients (HOG)
   Image features were extracted using the HOG method, which aims to capture edge and texture characteristics by dividing the image into small blocks and calculating the gradient orientation of the pixels within them. HOG is effective in recognizing visual patterns and is often used in object recognition (Ramiady et al., 2024).
4) Encoding Class Labels Using LabelEncoder
   The class name or label of each image is converted into numeric form using LabelEncoder so it can be processed by the classification algorithm.

5) Addressing Data Imbalance Using SMOTE
   To address the uneven distribution of images between classes, the SMOTE (Synthetic Minority Oversampling Technique) method is used. This technique generates synthetic data for the minority class to create a more balanced distribution between classes, preventing the model from biasing toward the majority class (Mohanty et al., 2025).
6) Split the Dataset
   The processed data is divided into two parts: 80% training data and 20% testing data. The training data is used to build the classification model, while the testing data is used to evaluate the model's performance on previously unseen data.

2.3. Model Development

This research conducted a classification model development using four different machine learning algorithms: Support Vector Machine (SVM), K-Nearest Neighbor (KNN), Random Forest (RF), and Logistic Regression (LR).

1) Support Vector Machine (SVM)
   Support Vector Machine works by finding the best hyperplane that separates data from different classes in a high-dimensional feature space (Ghaddar and Naoum-Sawaya, 2018).

$$f(x) = sign(w^T x + b) \tag{1}$$

2) K-Nearest Neighbor (KNN)
   K-Nearest Neighbor classifies data based on the majority class of a number of nearest neighbors.

$$d(x \cdot y) = \sqrt{\sum_{k=1}^{n} (x_k - y_k)^2} \tag{2}$$

3) Random Forest (RF)
   Random Forest is a decision tree-based ensemble model that uses multiple decision trees and votes for the final result (Wang, 2022).

$$\hat{y} = mode\{h_1(x), h_2(x), \ldots, h_T(x)\} \tag{3}$$

4) Logistic Regression (LR)
   Logistic Regression is used to predict the probability of data belonging to a particular class based on a linear combination of features.

$$P(y = 1|x) = \frac{1}{1 + e^{-(x^T x + b)}} \tag{4}$$

This training process is carried out by utilizing features generated from the Histogram of Oriented Gradients (HOG) method as a representation of each pest image, as well as labels that have been encoded numerically using LabelEncoder.

2.4. Model Evaluation

The evaluation was conducted using common classification metrics: accuracy, precision, recall, and F1-score. The formula for each metric is as follows:

$$Accuracy = \frac{TP + TN}{TP + TN + FN + FP} \times 100\% \tag{5}$$

$$Precision = \frac{TP}{TP + FP} \times 100\% \tag{6}$$

$$Recall = \frac{TP}{TP + FN} \times 100\% \tag{7}$$

$$F1 - Score = \frac{TP}{TP + FN} \times 100\% \tag{8}$$

where:

TP: True Positive
TN: True Negative
FP: False Positive
FN: False Negative

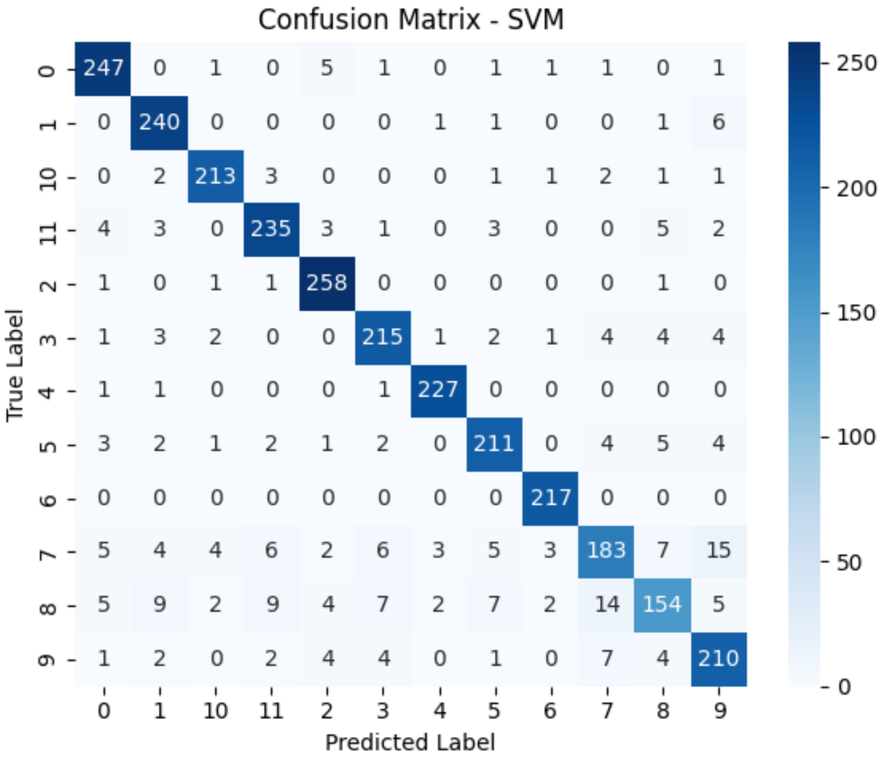
3. Results and Discussion

3.1. Model Performance Evaluation

Table 2: Evaluation of Support Vector Machine

Class	Name	Precision	Recall	F1-score
0	Rice Leaf Roller	0.92	0.96	0.94
1	Rice Leaf Caterpillar	0.90	0.96	0.93
2	Paddy Stem Maggot	0.95	0.95	0.95
3	Asiatic Rice Borer	0.91	0.92	0.91
4	Yellow Rice Borer	0.93	0.98	0.96
5	Rice Gall Midge	0.91	0.91	0.91
6	Brown Plant Hopper	0.97	0.99	0.98
7	Rice Stem Fly	0.91	0.90	0.90
8	Rice Water Weevil	0.96	1.00	0.98
9	Rice Leaf Hopper	0.85	0.75	0.80
10	Rice Shell Pest	0.85	0.70	0.77
11	Thrips	0.85	0.89	0.87

The Support Vector Machine (SVM) model demonstrated excellent performance in detecting most classes of rice pests. Classes such as Brown Plant Hopper (class 6), Rice Water Weevil (class 8), and Paddy Stem Maggot (class 2) had very high precision, recall, and f1-score values, approaching or reaching 0.98–1.00, indicating the model was able to consistently and accurately identify these pests. However, several classes, such as Rice Leaf Hopper (class 9) and Rice Shell Pest (class 10), had lower recall and f1-score values (around 0.70–0.80), indicating that the model still had difficulty identifying these pests with balanced precision and sensitivity.



True \ Predicted	0	1	10	11	2	3	4	5	6	7	8	9
0	247	0	1	0	5	1	0	1	1	1	0	1
1	0	240	0	0	0	0	1	1	0	0	1	6
10	0	2	213	3	0	0	0	1	1	2	1	1
11	4	3	0	235	3	1	0	3	0	0	5	2
2	1	0	1	1	258	0	0	0	0	0	1	0
3	1	3	2	0	0	215	1	2	1	4	4	4
4	1	1	0	0	0	1	227	0	0	0	0	0
5	3	2	1	2	1	2	0	211	0	4	5	4
6	0	0	0	0	0	0	0	0	217	0	0	0
7	5	4	4	6	2	6	3	5	3	183	7	15
8	5	9	2	9	4	7	2	7	2	14	154	5
9	1	2	0	2	4	4	0	1	0	7	4	210

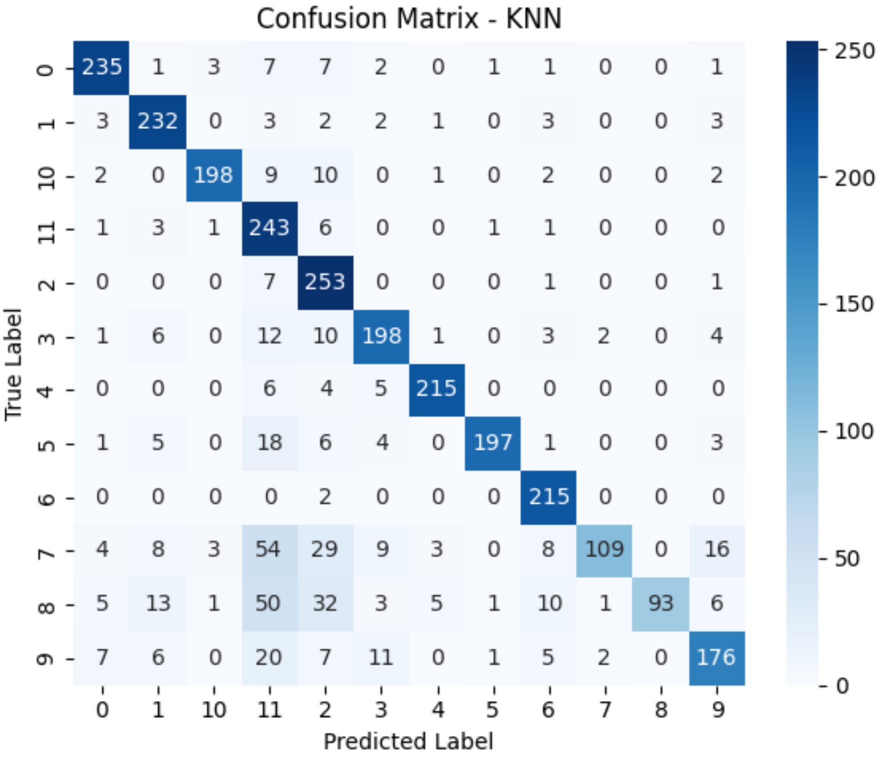
Figure 1: Confusion matrix of SVM

The confusion matrix for the SVM model shows excellent classification performance for most classes. Classes such as Rice Leaf Roller (class 0), Rice Leaf Caterpillar (class 1), and Paddy Stem Maggot (class 2) were predicted almost perfectly with the number of correct predictions reaching more than 240 instances each. However, there were slight misclassifications in several classes, especially in class 7 (Rice Stem Fly), class 8 (Rice Water Weevil), and class 9 (Rice Leaf Hopper), which experienced quite a lot of misclassifications to other classes. For example, in class 7, only 183 of the total instances were correctly classified, while the rest were spread across various classes such as class 6, 8, and 9.

Table 3: K-Nearest Neighbor Evaluation

Class	Name	Precision	Recall	F1-score
0	Rice Leaf Roller	0.91	0.91	0.91
1	Rice Leaf Caterpillar	0.85	0.93	0.89
2	Paddy Stem Maggot	0.96	0.88	0.92
3	Asiatic Rice Borer	0.57	0.95	0.71
4	Yellow Rice Borer	0.69	0.97	0.80
5	Rice Gall Midge	0.85	0.84	0.84
6	Brown Plant Hopper	0.95	0.93	0.94
7	Rice Stem Fly	0.98	0.84	0.90
8	Rice Water Weevil	0.86	0.99	0.92
9	Rice Leaf Hopper	0.96	0.45	0.61
10	Rice Shell Pest	1.00	0.42	0.59
11	Thrips	0.83	0.75	0.79

Based on the K-Nearest Neighbor (KNN) model evaluation table, classification performance varies considerably across classes. Several classes, such as Rice Stem Fly (class 7) and Rice Shell Pest (class 10), exhibit overfitting symptoms, characterized by very high precision values (0.98 and 1.00, respectively), but low recall (0.84 and 0.42, respectively). This indicates that although the KNN model is very confident in its predictions, it can only recognize a small fraction of the actual samples from these classes. Furthermore, for the Asiatic Rice Borer class (class 3), the model has low precision (0.57) but high recall (0.95), indicating that the model often mispredicts this class. However, several classes perform quite well, such as Paddy Stem Maggot (class 2), Brown Plant Hopper (class 6), and Rice Water Weevil (class 8), which have f1-scores above 0.90.



	0	1	10	11	2	3	4	5	6	7	8	9
0	235	1	3	7	7	2	0	1	1	0	0	1
1	3	232	0	3	2	2	1	0	3	0	0	3
10	2	0	198	9	10	0	1	0	2	0	0	2
11	1	3	1	243	6	0	0	1	1	0	0	0
2	0	0	0	7	253	0	0	0	1	0	0	1
3	1	6	0	12	10	198	1	0	3	2	0	4
4	0	0	0	6	4	5	215	0	0	0	0	0
5	1	5	0	18	6	4	0	197	1	0	0	3
6	0	0	0	0	2	0	0	0	215	0	0	0
7	4	8	3	54	29	9	3	0	8	109	0	16
8	5	13	1	50	32	3	5	1	10	1	93	6
9	7	6	0	20	7	11	0	1	5	2	0	176

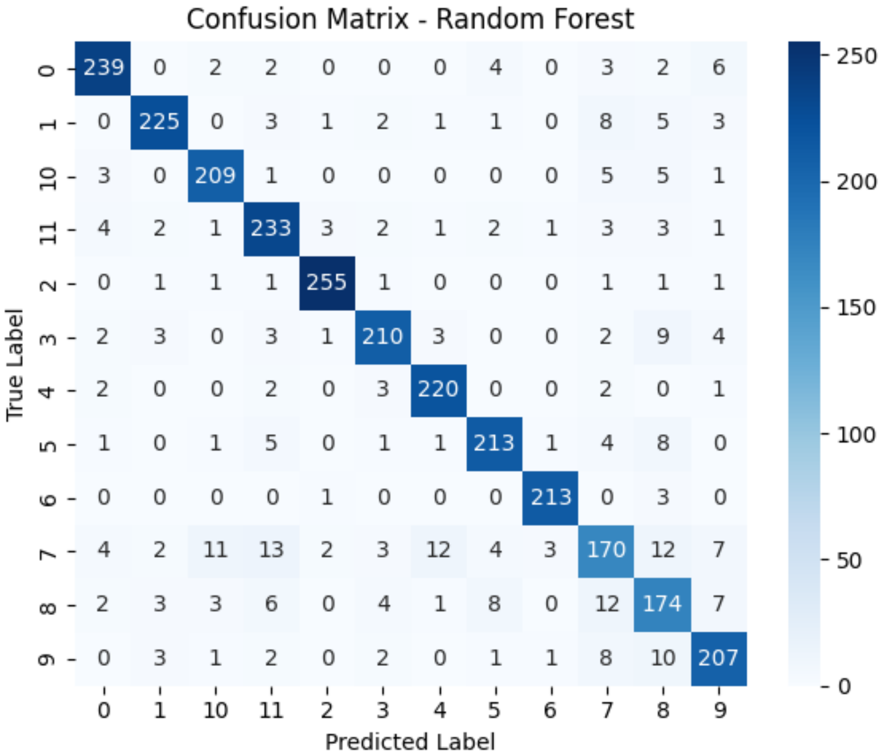
Figure 2: Confusion matrix of KNN

The confusion matrix for the KNN model indicates that it has lower classification performance than SVM, especially for certain classes. While the model is able to classify classes such as Paddy Stem Maggot (class 2) and Rice Leaf Roller (class 0) quite well, with over 230 correctly classified instances, the model appears to struggle to distinguish classes such as Rice Stem Fly (class 7), Rice Water Weevil (class 8), and Rice Leaf Hopper (class 9). For example, in class 7, only 109 instances were correctly classified out of a much larger total, while over 100 other instances were misclassified into various classes such as 8, 6, and even 3. This pattern suggests that KNN is susceptible to similarities between visual features due to its distance-based nature. It also leads to overfitting on certain training data but fails to generalize well to more complex or visually noisy test data.

Table 4: Random Forest Evaluation

Class	Name class	Precision	Recall	F1-score
0	Rice Leaf Roller	0.94	0.91	0.93
1	Rice Leaf Caterpillar	0.91	0.92	0.91
2	Paddy Stem Maggot	0.93	0.95	0.94
3	Asiatic Rice Borer	0.82	0.91	0.87
4	Yellow Rice Borer	0.95	0.97	0.96
5	Rice Gall Midge	0.88	0.84	0.86
6	Brown Plant Hopper	0.94	0.94	0.94
7	Rice Stem Fly	0.93	0.89	0.91
8	Rice Water Weevil	0.99	0.99	0.99
9	Rice Leaf Hopper	0.79	0.71	0.75
10	Rice Shell Pest	0.76	0.80	0.78
11	Thrips	0.88	0.88	0.88

Based on the evaluation results of the Random Forest (RF) model, the classification performance generally shows good consistency with high precision, recall, and f1-score values for almost all classes. Several classes, such as Rice Water Weevil (class 8) and Yellow Rice Borer (class 4), achieved very high f1-scores (0.99 and 0.96), demonstrating the model's ability to detect and classify these categories with high accuracy. However, weaknesses are still visible in the Rice Leaf Hopper (class 9) and Rice Shell Pest (class 10) classes, which only achieved f1-scores of 0.75 and 0.78, respectively. This indicates that although the model is quite reliable, classification challenges remain for classes with high visual similarity or an unbalanced initial data set, even after balancing with SMOTE. There is no indication of overfitting as is the case with KNN, as the difference between precision and recall is relatively balanced.



	0	1	10	11	2	3	4	5	6	7	8	9
0	239	0	2	2	0	0	0	4	0	3	2	6
1	0	225	0	3	1	2	1	1	0	8	5	3
10	3	0	209	1	0	0	0	0	0	5	5	1
11	4	2	1	233	3	2	1	2	1	3	3	1
2	0	1	1	1	255	1	0	0	0	1	1	1
3	2	3	0	3	1	210	3	0	0	2	9	4
4	2	0	0	2	0	3	220	0	0	2	0	1
5	1	0	1	5	0	1	1	213	1	4	8	0
6	0	0	0	0	1	0	0	0	213	0	3	0
7	4	2	11	13	2	3	12	4	3	170	12	7
8	2	3	3	6	0	4	1	8	0	12	174	7
9	0	3	1	2	0	2	0	1	1	8	10	207

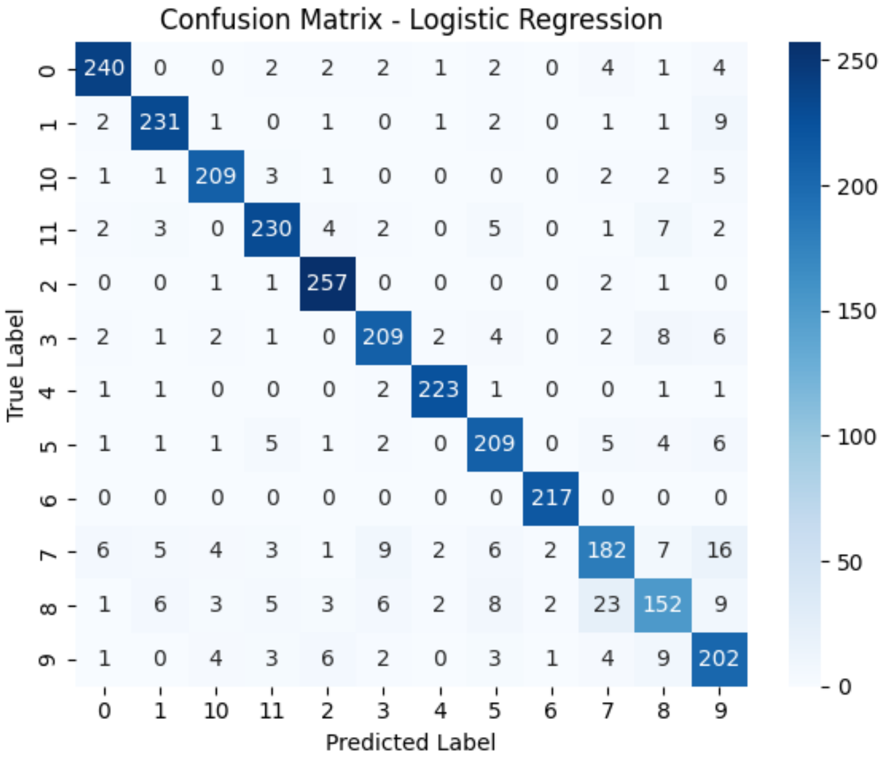
Figure 3: Confusion matrix of Random Forest

The confusion matrix for the Random Forest model showed quite good and relatively balanced performance in classifying most classes. The model was able to recognize classes such as Paddy Stem Maggot (class 2), Rice Leaf Roller (class 0), and Brown Plant Hopper (class 6) with high accuracy each with over 230 correct predictions. However, there was some classification confusion, especially in classes such as Rice Stem Fly (class 7) and Rice Water Weevil (class 8), where the model still made incorrect predictions to some classes such as 6, 5, and 9. For example, out of 240 data points in class 7, only around 170 were correctly classified, while the rest were scattered across similar classes. This shows that although Random Forest is quite robust and stable in handling variations between classes, it still faces challenges in distinguishing classes that share similar visual features.

Table 5: Logistic Regression Evaluation

Class	Name class	Precision	Recall	F1-score
0	Rice Leaf Roller	0.93	0.93	0.93
1	Rice Leaf Caterpillar	0.93	0.93	0.93
2	Paddy Stem Maggot	0.93	0.93	0.93
3	Asiatic Rice Borer	0.91	0.90	0.90
4	Yellow Rice Borer	0.93	0.98	0.96
5	Rice Gall Midge	0.89	0.88	0.89
6	Brown Plant Hopper	0.97	0.97	0.97
7	Rice Stem Fly	0.87	0.89	0.88
8	Rice Water Weevil	0.98	1.00	0.99
9	Rice Leaf Hopper	0.81	0.75	0.78
10	Rice Shell Pest	0.79	0.69	0.74
11	Thrips	0.78	0.86	0.82

The Logistic Regression (LR) model demonstrated strong performance in classifying rice pest images, with an overall accuracy of 89.36%. Several classes, such as Rice Water Weevil (class 8) and Brown Plant Hopper (class 6), achieved very high precision and recall values, even near perfect, indicating that the model was able to recognize these classes with high accuracy and consistency. However, upon closer inspection, there were indications of overfitting in several classes, such as Yellow Rice Borer (class 4) and Paddy Stem Maggot (class 2), which had very high recall values (above 0.95) but were not accompanied by significant variations in precision. This could be the result of the oversampling process using SMOTE, which increases the synthetic data and makes the model too "memorized" of certain patterns. In addition, several classes such as Rice Shell Pest (class 10) and Rice Leaf Hopper (class 9) still show relatively lower performance with f1-score values of 0.74 and 0.78, respectively, indicating that the model is quite often misclassified in these classes.



	0	1	10	11	2	3	4	5	6	7	8	9
0	240	0	0	2	2	2	1	2	0	4	1	4
1	2	231	1	0	1	0	1	2	0	1	1	9
10	1	1	209	3	1	0	0	0	0	2	2	5
11	2	3	0	230	4	2	0	5	0	1	7	2
2	0	0	1	1	257	0	0	0	0	2	1	0
3	2	1	2	1	0	209	2	4	0	2	8	6
4	1	1	0	0	0	2	223	1	0	0	1	1
5	1	1	1	5	1	2	0	209	0	5	4	6
6	0	0	0	0	0	0	0	0	217	0	0	0
7	6	5	4	3	1	9	2	6	2	182	7	16
8	1	6	3	5	3	6	2	8	2	23	152	9
9	1	0	4	3	6	2	0	3	1	4	9	202

Figure 4: Confusion matrix of Logistic Regression

The confusion matrix for the Logistic Regression (LR) model shows that several classes such as Rice Leaf Roller (class 0), Paddy Stem Maggot (class 2), and Brown Plant Hopper (class 6) are classified very well, indicated by a high number of correct predictions (above 240 for class 0 and 257 for class 2). However, classification errors start to appear more significant in classes such as Rice Stem Fly (class 7) and Rice Water Weevil (class 8). For example, only 182 instances in class 7 were correctly classified, with the rest spread across several other classes, especially classes 4, 6, and 9. This indicates that Logistic Regression tends to have difficulty recognizing feature patterns from complex or visually overlapping classes. However, compared to KNN, this model appears more stable with less extreme errors.

3.2. Comparison of wrong and correct predictions

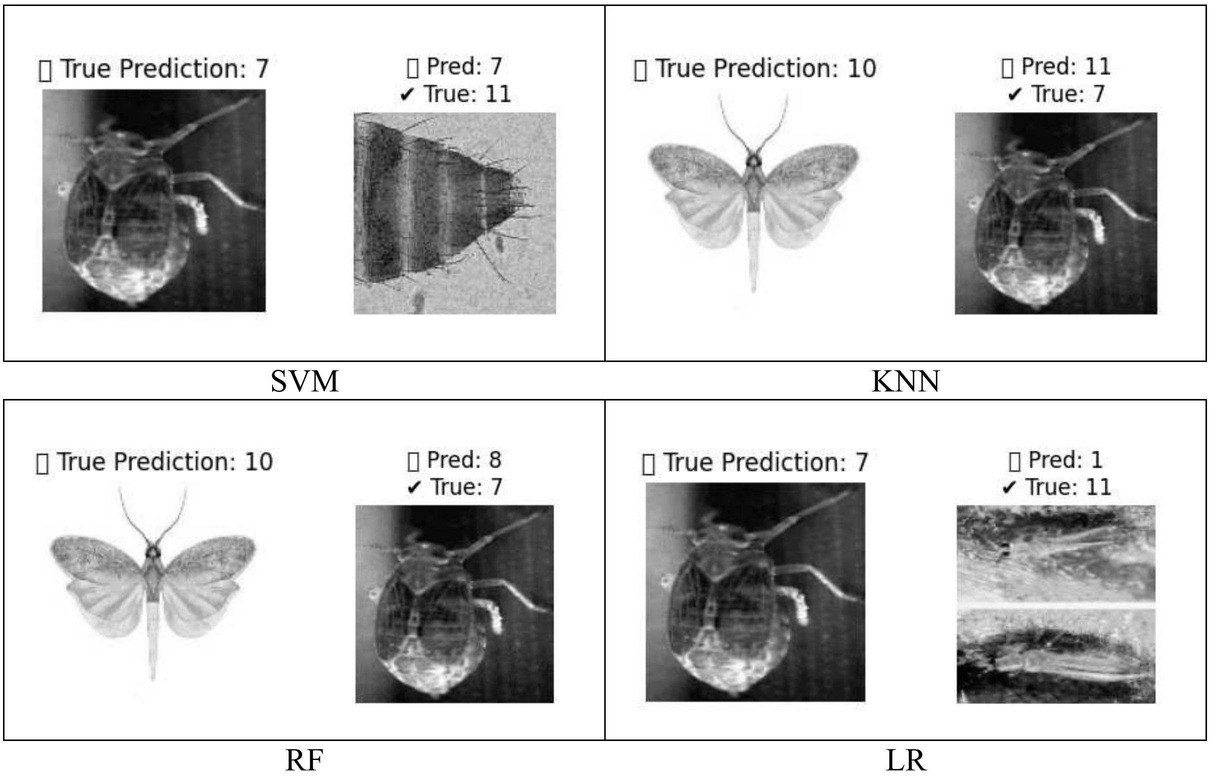


Figure 5: Comparison of correct and incorrect predictions

In the SVM model, it can be seen that one of the pest images in class 7 was successfully predicted correctly, but an error occurred when predicting an image from class 11 which was actually classified as class 7. A similar thing also happened in KNN, where the correct prediction was seen in class 10, but the model incorrectly predicted an image from class 7 as class 11. In Random Forest, the model was able to recognize images from class 10 correctly, but incorrectly classified an image from class 7 as class 8. Meanwhile, Logistic Regression was able to recognize images from class 7 correctly, but experienced an error when predicting an image from class 11 as class 1. Of all the incorrect predictions displayed, it appears that class 7 and class 11 are the two classes that are most often confused between models, indicating that the visual features between the two classes tend to be similar or difficult for the algorithm to distinguish. This error consistently appears in all models, indicating that class 7 (Rice Stem Fly) and class 11 (Thrips) are quite challenging classes to separate visually, possibly due to the similar shape or texture of the insects in the image.

4. Conclussion

Based on the evaluation results of performance metrics (precision, recall, and f1-score) and confusion matrix, it can be concluded that SVM provides the most consistent and balanced performance, with high f1-scores in almost all classes. Random Forest also shows competitive performance, especially in classes that are easier to recognize, although slightly more variable than SVM. In contrast, KNN and Logistic Regression show weaknesses, especially in classes with similar visual features, such as class 7 (Rice Stem Fly) and class 11 (Thrips), where both models often experience misclassification and show symptoms of overfitting in recall values. Confusion matrix analysis strengthens this finding, by showing that the models tend to have difficulty distinguishing certain classes that have visual similarities, such as classes 7 and 11. Visualization of correct and incorrect predictions also shows that most prediction errors occur due to similar pest morphology, which requires models with deeper feature extraction capabilities to accurately distinguish.

References

Ghaddar, B., & Naoum-Sawaya, J. (2018). High dimensional data classification and feature selection using support vector machines. *European Journal of Operational Research*, *265*(3), 993-1004.

Kasinathan, T., & Uyyala, S. R. (2021). Machine learning ensemble with image processing for pest identification and classification in field crops. *Neural Computing and Applications*, *33*(13), 7491-7504.

Kaushal, S., Kumar, S., & Tabrez, S. (2022). Artificial intelligence in agriculture. *Journal of Science and Research, doi10*, *21275*.

Mohanty, N., Behera, B. K., Ferrie, C., & Dash, P. (2025). A quantum approach to synthetic minority oversampling technique (SMOTE). *Quantum Machine Intelligence*, *7*(1), 38.

Ochango, V. M., Wambugu, G. M., & Ndia, J. G. (2022). Feature extraction using histogram of oriented gradients for image classification in maize leaf diseases.

Ramiady, L., Arnia, F., Oktiana, M., & Novandri, A. (2024). Improved Histogram of Oriented Gradient (HOG) Feature Extraction for Facial Expressions Classification. *Jurnal Rekayasa Elektrika*, *20*(3).

Setiawan, R., Zein, H., Azdy, R. A., & Sulistyowati, S. (2023). Rice Leaf Disease Classification with Machine Learning: An Approach Using Nu-SVM. *Indonesian Journal of Data and Science*, *4*(3), 136-144.

Sundhar, S., Sharma, R., Maheshwari, P., Kumar, S. R., & Kumar, T. S. (2025). Enhancing leaf disease classification using GAT-GCN hybrid model. *arXiv preprint arXiv:2504.04764*.

Wang, J., Wang, T., Xu, Q., Gao, L., Gu, G., Jia, L., & Yao, C. (2025). RP-DETR: end-to-end rice pests detection using a transformer. *Plant Methods*, *21*(1), 1-17.

Wang, R. (2022, December). Comparison of decision tree, random forest and linear discriminant analysis models in breast cancer prediction. In *Journal of Physics: Conference Series* (Vol. 2386, No. 1, p. 012043). IOP Publishing.