Journal of Information Technology and Computer Science
Volume 11, Number 2, August 2026, pp. 224-235
Journal Homepage: www.jitecs.ub.ac.id

Comparative Analysis of Random Forest, Support Vector Machine, and K-Nearest Neighbor with Image Feature Extraction for Rice Leaf Disease Detection

Nadia Nafista[1], Wawu Tri Ambodo[2], Analicia[3], Ulfa Siti Nuraini*[4]

[1,2,3,4]Universitas Negeri Surabaya, Surabaya
[1,2,3]{nadianafista.23095, wawu.22035, analicia.22007}@mhs.unesa.ac.id
[4]ulfanuraini@unesa.ac.id
*Corresponding Author

Received 10 January 2026; accepted 30 July 2026

Abstract. Plant diseases and pest infestations have caused a decline in global food production of up to 40%, including in Indonesia, making an efficient and accurate disease detection system essential to support food security. This study proposes a supervised learning approach to detect rice leaf diseases based on image processing. Leaf images are processed through the stages of image segmentation, normalization, Gaussian blur, Canny edge detection, visualization of diseased areas, and hybrid feature extraction. Supervised learning algorithms such as Random Forest (RF), Support Vector Machine (SVM), and K-Nearest Neighbor (KNN) were trained and compared. Test results show that Random Forest delivers the best performance with an accuracy of 97%, outperforming SVM and KNN. These findings indicate that the proposed approach can serve as an effective and reliable solution for the automatic detection of rice leaf diseases.

Keywords: Food Security, Image Processing, Plant Infection, Supervised Learning Model.

1 Introduction

Plant diseases in agricultural crops have become one of the major challenges faced by farmers. A 2021 report by the Food and Agriculture Organization (FAO) states that plant diseases and pest infestations cause a decline in global food production of up to 40% and economic losses estimated to exceed USD 290 billion. Such disease outbreaks can cause serious damage to crops, significantly reduce harvest yields, and potentially trigger a food crisis due to increased vulnerability of agricultural systems [1]. Plant diseases are not only a technical problem in the cultivation process but also threaten agricultural productivity and food security. In Indonesia, the rise in plant disease cases has led to a decline in crop quality, increased production costs due to pesticide use, and reduced farmer income [2].

Digital image processing has emerged as a promising approach for identifying plant diseases through the analysis of leaf images to detect symptoms such as color changes, the appearance of spots, deformities, and damage to the leaf surface [3]. Various studies have been conducted using the same dataset, including a study by Moh. Heri Susanto et al. in 2025, which applied a Convolutional Neural Network (CNN) and achieved a test accuracy of 96.47% [4]. Additionally, Bagus Adi Prastyo et al. in 2023 used a Support Vector Machine (SVM) with a polynomial kernel and achieved a precision of 82%, a recall of 81.75%, and an F1-score of 81.75% [5]. Another study by Bintang Karmila Gulo et al. in 2025 applied a Random Forest combined with feature extraction using an HSV histogram, RGB mean values, and the Gray Level Co-occurrence Matrix (GLCM), achieving a best accuracy of 97.73% [6].

However, previous studies have rarely combined the SMOTE technique, HSV histogram, GLCM, and Canny Edge-based feature extraction with conventional supervised learning models. Therefore, this study evaluates the Random Forest (RF) algorithm, SVM with a polynomial kernel, and K-Nearest Neighbor (KNN) using the proposed hybrid feature extraction framework to determine the most effective model for classifying rice leaf diseases.

2 Method

The research methodology process is presented in the flowchart shown in Figure 1, which illustrates the overall workflow of the proposed system.

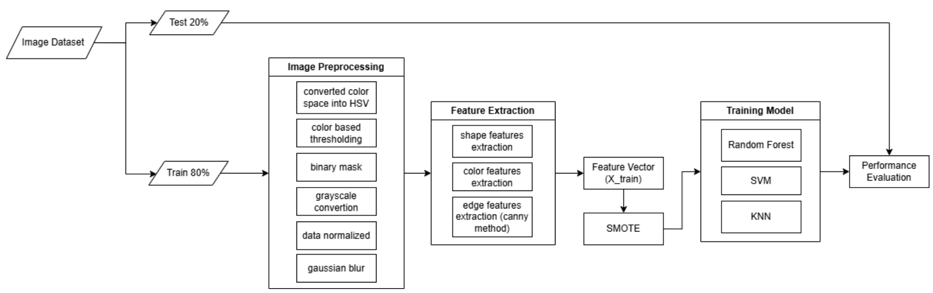

Figure 1. Research methodology and system design flowchart

2.1 Data Collection

This study is use the dataset that was obtained from Kaggle, originally published by Rahmi Burhanuddin, and has previously been utilized in the study "Klasifikasi Penyakit Padi Melalui Citra Daun Menggunakan Metode Naive Bayes" [7]. It consists of 500 rice leaf images belonging to three disease classes, as showed in Table 1. Consequently, the dataset will serve as the basis for feature extraction, model training, and performance evaluation.

Table 1. Class distribution from dataset

Class	Number of Data
Blast	200
Blight	220
Tungro	80

Table 1 shows the class distribution in the dataset that used in this study, consisting of Blast, Blight, and Tungro class with an imbalanced distribution that may bias model training. Figure 2 presents representative images of each class. Blast, caused by Pyricularia oryzae, is characterized by diamond-shaped lesions with dark margins and grayish centers [8]. Blight, caused by *Xanthomonas oryzae*, exhibits elongated yellowish brown lesions that spread along the leaf veins [9]. whereas Tungro, caused by Rice Tungro Bacilliform Virus (RTBV) and Rice Tungro Spherical Virus (RTSV), produces yellow to orange discoloration and stunted leaf growth [10].

Figure 2. Rice leaf diseases

2.2 Image Pre-processing

Image preprocessing is performed to improve image quality and ensure that relevant disease information can be effectively extracted. The initial stage begins by converting all input images from the RGB color space to the HSV color space. This conversion aims to enhance the ability to distinguish color characteristics and reduce sensitivity to lighting variations commonly found in images acquired in the field [11]. Subsequently, color-based thresholding is applied to segment regions that indicate potential disease symptoms. This process produces a binary mask that isolates the infected areas from the background [12]. To maintain numerical stability during the feature extraction stage, pixel intensity values are normalized to a uniform range. Finally, Gaussian Blur filtering is applied to reduce noise and smooth image details, thereby improving the reliability of edge detection and shape analysis in subsequent stages [13].

2.3 Feature Extraction

Feature extraction aims to transform visual disease characteristics into numerical representations suitable for classification. In this study, three categories of features are extracted, such as color, shape, and edge features [14]. Color Histogram is a feature extraction technique that calculates the frequency of pixels in each color range (bin) in a color space, such as HSV. By compiling a histogram of all pixel values in an image, we can obtain a representative color distribution for each image class. The input RGB image was first transformed into the HSV color space.

$$I_{HSV} = f(I_{RGB}) \tag{1}$$

Where I_{RGB} and I_{HSV} denote the RGB and HSV images, respectively. For each HSV channel, a histogram was computed as:

$$H_c(i) = \sum_{x=1}^{M}\sum_{y=1}^{N} \delta\,(I_c(x,y) = i) \tag{2}$$

Where $c \in \{H, S, V\}$. After that, the histogram was then normalized using L2 normalization.

$$\hat{H}(i) = \frac{H(i)}{\sqrt{\sum_{j=1}^{n} H\,(j)^2}} \tag{3}$$

At the end, the color feature vector was obtained by concatenating the normalized histograms.

$$F_{\text{color}} = \left[\hat{H}_H, \hat{H}_S, \hat{H}_V\right] \tag{4}$$

Shape features are then extracted to describe the geometric properties of infected areas, including the structure and spatial distribution of lesions [15]. These features are essential for distinguishing diseases with similar color characteristics but different lesion patterns. Additionally, edge features are extracted using the Canny edge detection method [16]. This method consists of the following stages, it begins with noise reduction using a Gaussian filter to smooth the image and minimize the influence of noise while preserving significant edge information. The Gaussian filter is defined as follows:

$$G_\sigma(x, y) = \frac{1}{2\pi\sigma^2} \exp\left(-\frac{x^2 + y^2}{2\sigma^2}\right) \tag{5}$$

Where σ represents the standard deviation of the Gaussian distribution. After the Gaussian filtering process, edge detection was performed using the Canny method, which involves gradient calculation, non-maximum suppression, double thresholding, and edge tracking by hysteresis to produce accurate lesion boundaries. Next, contour-based shape features, such as area and perimeter, were extracted from the edge detection results. In addition, GLCM is applied to analyze texture by measuring the spatial relationships between grayscale pixels at four orientations: 0°, 45°, 90°, and 135°. From this matrix, texture features, which like contrast, correlation, energy, and homogeneity are calculated and then combined with color and shape features as input for the classification process [17].

$$\text{Contrast} = \sum_{i,j} |i - j|^2 \, p(i, j) \tag{6}$$

Furthermore, a correlation measure is required to determine the extent to which pixel pairs with the same intensity tend to maintain a consistent spatial separation within an image. Correlation values range from -1 to 1, where positive values indicate a positive correlation and negative values indicate a negative correlation.

$$\text{Correlation} = \sum_{i,j} \frac{(i - \mu_i)(j - \mu_j)p(i, j)}{\sigma_i \sigma_j} \tag{7}$$

In addition, to measures the degree of pixel intensity homogeneity in an image, there is an energy to calculate it. The higher the energy value, the more homogeneous the image is.

$$\text{Energy} = \sum_{i,j} p\,(i, j)^2 \tag{8}$$

And last, to measure the extent, homogeneity is used, so that, the pixel intensities in an image approach the same intensity value. Thus, the higher the homogeneity, the more homogeneous the image is.

$$\text{Homogeneity} = \sum_{i,j} \frac{p(i, j)}{1 + |i - j|} \tag{9}$$

By combining color features using HSV histograms, shape features which obtained from Canny Edge Detection and contour analysis, and texture features using GLCM, it is expected that the characteristics of rice leaf diseases can be represented more comprehensively, thereby enhancing the model's ability to distinguish between disease classes with greater accuracy.

2.4 Resampling

Resampling is a nonparametric technique used to address class imbalance by reconstructing the data distribution from the available samples [18]. In this study, the Tungro class contained significantly fewer samples than the other classes, potentially biasing the model toward the majority classes. Therefore, SMOTE oversampling was applied exclusively to the training data. Rather than duplicating existing minority samples, SMOTE generates synthetic samples by interpolating each minority instance (X_i) with one of its randomly selected k-nearest minority neighbors (X_{knn}), where the interpolation is controlled by a random value $\delta \in [0,1]$. The synthetic sample is generated as follows [19].

$$X_{\text{syn}} = X_i + (X_{knn} - X_i) \times \delta \tag{10}$$

2.5 Training Model

After feature extraction and SMOTE-Oversampling, three supervised learning algorithms were trained and compared. The first classifier was Support Vector Machine or SVM, introduced in the early 1990s by Vladimir N. Vapnik, using a polynomial kernel to model the nonlinear relationships among the extracted color, texture, and shape features. The SVM and kernel formulation are presented as follows [20].

$$Main\ equation\ = \frac{1}{2}\|\mathbf{w}\|^2 = \frac{1}{2}(w_1^2 + w_2^2) \tag{11}$$

$$y_i(\mathbf{w} \cdot \mathbf{x}_i + b) \geq 1, \qquad i = 1,2,\dots,N$$
$$\text{Polynomial Kernel} = K(\mathbf{x}_i, \mathbf{x}_j) = \left(\gamma\, \mathbf{x}_i^T \mathbf{x}_j + r\right)^d \tag{12}$$

The Support Vector Machine (SVM) classifier was implemented using a polynomial kernel with the default Scikit-learn hyperparameters ($C = 1.0$, degree $= 3$, gamma = *scale*, and coef0 $= 0.0$), while a random state of 42 ensured reproducibility. The second classifier is Random Forest (RF), developed by Tin Kam Ho in 1995, this method constructs a number of decision trees using bootstrap sampling and random feature selection, with the final prediction determined through majority voting. In this study, the RF is configured with 100 decision trees using the Gini impurity criterion and a random state of 42, defined as follows [21].

$$I(t) = 1 - \sum_{c=1}^{C} \left(\frac{n_c}{N}\right)^2 \tag{13}$$

Next, the optimal split is selected based on the lowest weighted Gini Impurity, as defined by the following Gini Index equation.

$$Gini(t, X) = \sum_{i=1}^{j} \frac{N_i}{N} I(t_i) \tag{14}$$

Other models, such as K-Nearest Neighbor, will also be compared in this study. The KNN algorithm was first introduced by Evelyn Fix and Joseph Hodges in 1951 and is an instance-based classification model that determines the class of a data point based on the majority class of its k nearest neighbors in feature space, with the total number of neighbors typically being an odd number. In this study, the KNN classifier was configured with k = 5. The Euclidean distance is calculated as follows [22].

$$d(a,b) = \sqrt{\sum_{g=1}^{p} \left(X_{ag} - X_{bg}\right)^2} \tag{15}$$

2.6 Model Evaluation

Model performance was evaluated with precision, accuracy, recall, F1-score, and specificity, which were calculated from the confusion matrix components [23]. Accuracy measures the proportion of correctly classified samples and is defined as follows.

$$\text{Accuracy} = \frac{TP + TN}{TP + TN + FP + FN} \tag{16}$$

Sensitivity, also referred to as recall, is used to measure a model's ability to accurately detect positive instances [24]. Sensitivity reflects the model's ability to identify plant leaves infected with disease. Sensitivity is calculated using the following formula.

$$\text{Sensitivity} = \frac{TP}{TP + FN} \tag{17}$$

Specificity is used to measure a model's ability to accurately classify the negative class [25]. This metric indicates the model's ability to identify healthy plant leaves without misclassifying them as infected leaves. Specificity is calculated using the following formula.

$$\text{Specificity} = \frac{TN}{TN + FP} \tag{18}$$

Last, is the confusion matrix to compares the actual and predicted classes, where rows represent the actual classes and columns represent the predicted classes [26].

3 Results and Discussion

3.1 Resampling's Result

The initial dataset exhibits an imbalanced distribution, comprising 176 Blight images, 160 Blast images, and only 64 Tungro images, as shown in Figure 3. Consequently, the classification model may become biased toward the majority classes, thereby reducing its generalization capability. To prevent data leakage, the dataset was first split into training and testing sets at an 80:20 ratio prior to resampling, after which the SMOTE oversampling method was applied exclusively to the training data.

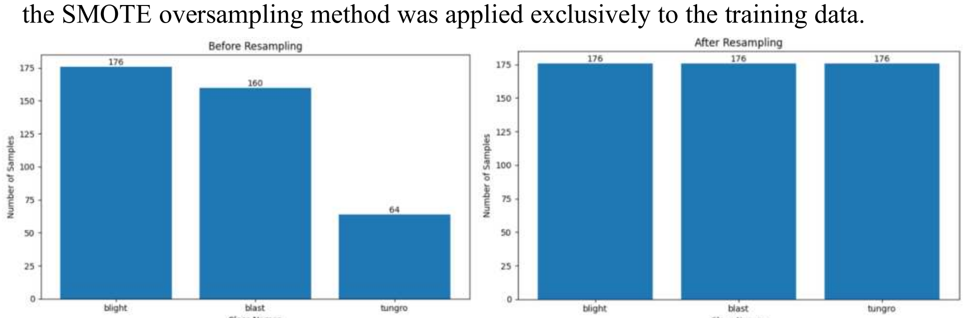

Figure 3. Class of training dataset before and after oversampling

After applying SMOTE, the number of samples in each class of the training data is balanced to 176 images, matching the number of samples in the majority class, as shown in Figure 3. By generating synthetic samples in the feature space, SMOTE is able to reduce class imbalance, help the model learn a more representative decision boundary and improve the model's robustness by minimizing toward the majority class.

Table 2. Number of class from training dataset before and after oversampling

Class	Before Oversampling	After Oversampling
Blast	160	176
Blight	176	176
Tungro	64	176

3.2 Image Processing and Feature Extraction Results

Figure 4 shows the stages of image processing, which include preprocessing, Canny edge detection, and visualization of the diseased areas. Preprocessing highlights the infected areas and reduces background noise, while Canny edge detection clarifies the boundaries of the lesions and the leaf contours. The processed results are then extracted into 26 features, consisting of color (HSV), texture (GLCM), and shape features (Canny Edge). For example, the shape features yield an area value of 2258.5 and a perimeter of 559.35, representing the size and shape of the lesion. So that, these features provides a more comprehensive representation of the image, thereby improving the model's ability to distinguish between Blast, Blight, and Tungro diseases.

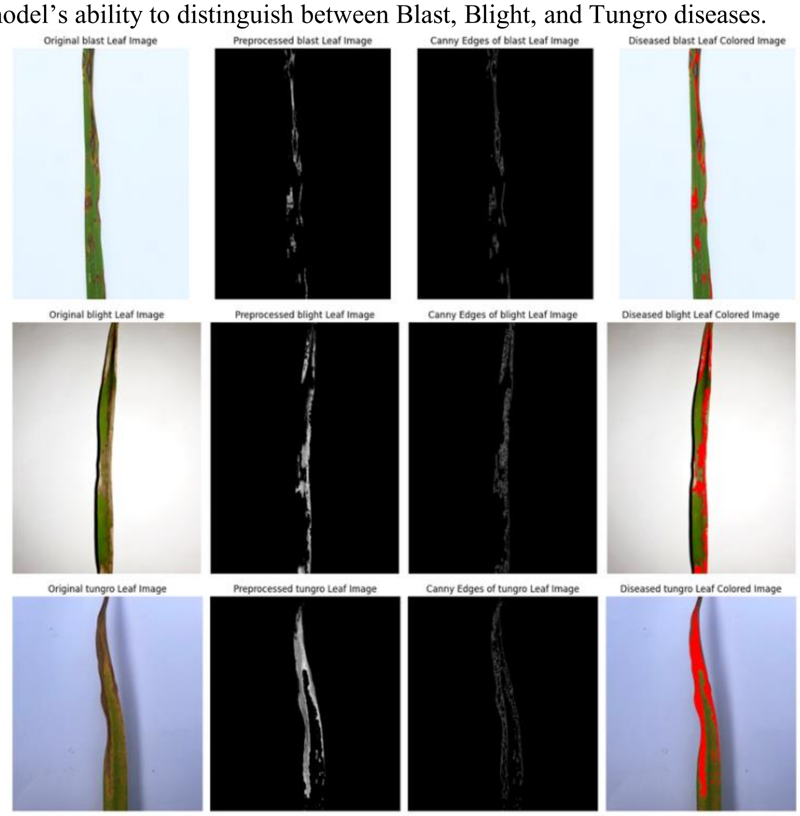

Figure 4. Visualization of preprocessed and extracted images

3.3 Comparison of Model's Performance

The classification models were conducted using the testing dataset. The dataset was divided into 80% training and 20% testing data, as presented in Table 3.

Table 3. The Result of Data Split

Data Group	Data Amount
Training	496
Testing	124

Table 4. Model's Performance Comparison

Model	Acc	F1-Score	Precision	Sensitivity/Recall	Specificity
SVM (Poly)	0.45	0.43	0.62	0.55	0.73
KNN	0.77	0.77	0.78	0.71	0.88
Random Forest	**0.97**	**0.95**	**0.96**	**0.92**	**0.97**

Table 4 shows the evaluation from several models, with Random Forest achieving the best performance at 97% with the most influential features are Hue_2 (0.1495), Hue_1 (0.1258), and Perimeter (0.0751). Outperforming KNN in 77% and SVM in 45%. This performance is supported by the ensemble learning mechanism, which is able to utilize a combination of color, texture, and shape features extracted from HSV, GLCM, and Canny Edge. Conversely, the outperforming of SVM with a polynomial kernel is likely due to its sensitivity to feature scale, making it less optimal for forming decision boundaries. Figure 5 shows that Random Forest produces the fewest misclassifications compared to other models.

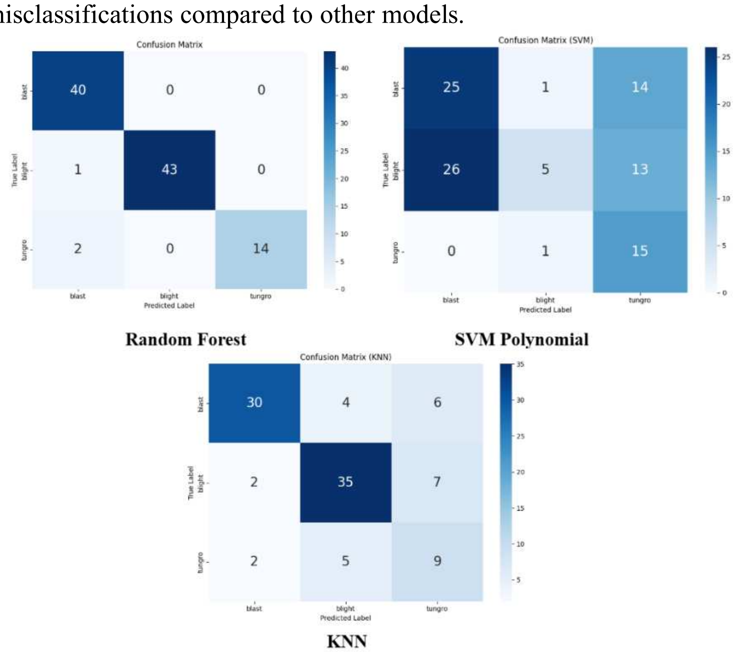

Figure 5. Confusion Matrix of SVM, RF and KNN

Figure 6 shows examples of correct classification and misclassification for each model. Misclassification generally occurs due to similarities in color, texture, and shape characteristics across classes, whereas Random Forest is able to minimize misclassification compared to KNN and SVM.

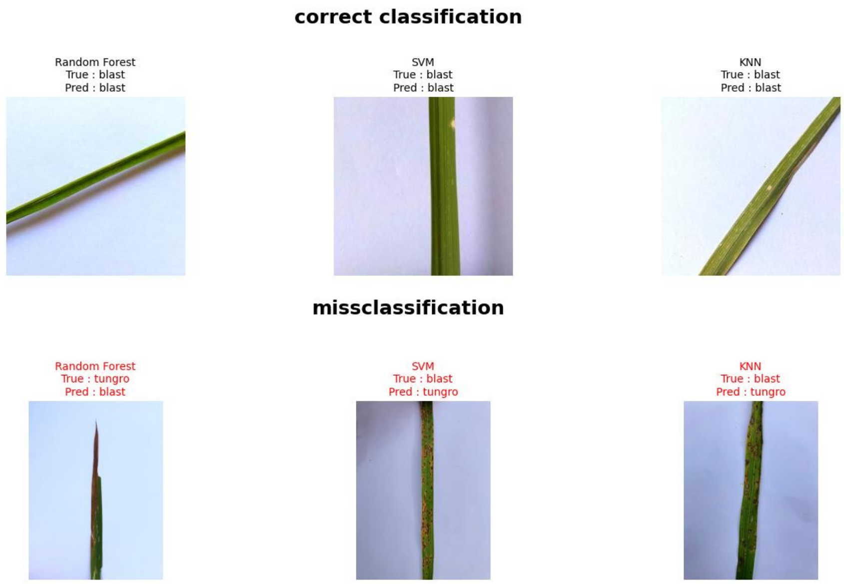

Figure 6. Sample of prediction test

4 Conclusion

This study shows that a hybrid feature extraction framework which combines HSV histograms, GLCM, and contour features based on Canny Edge Detection, and is supported by the SMOTE technique and supervised learning algorithms is capable of delivering good performance in the classification of rice leaf diseases. Of the three models evaluated, Random Forest delivered the best performance with an accuracy of 97%, outperforming KNN and SVM. These results are also comparable to previous studies using CNN (96.47%) and feature extraction based, Random Forest (97.73%), and demonstrate better performance than conventional SVM methods. However, this study still has several limitations, namely a relatively small dataset in 500 images, a focus solely on rice plants, and the use of images taken under relatively controlled conditions, which do not fully represent real world field conditions. Therefore, future research is recommended to use a larger and more diverse dataset, covering various plant species and different environmental conditions. In addition, the developed prototype needs to be implemented and tested on mobile devices through user testing to evaluate its ease of use, system reliability, and effectiveness in supporting real-time plant disease detection.

References

1. N. Ahmad, H. M. S. Asif, G. Saleem, M. U. Younus, S. Anwar, and M. R. Anjum, “Leaf Image-Based Plant Disease Identification Using Color and Texture Features,” *Wirel. Pers. Commun.*, vol. 121, no. 2, pp. 1139–1168, Nov. 2021, doi: 10.1007/s11277-021-09054-2.

2. V. K. Singh, R. Singh, A. Kumar, and R. Bhadouria, "Current status of plant diseases and food security," *Food Security and Plant Disease Management*, pp. 19–35, Jan. 2021, doi: 10.1016/B978-0-12-821843-3.00019-2.
3. F. Aldi and S. Sumijan, "Extraction of Shape and Texture Features of Dermoscopy Image for Skin Cancer Identification," *Sinkron*, vol. 8, no. 2, pp. 650–660, Mar. 2024, doi: 10.33395/sinkron.v8i2.13557.
4. Moh. Heri Susanto, Irwan Budi Santoso, Suhartono, and Ahmad Fahmi Karami, "Klasifikasi Penyakit Padi Menggunakan Convolutional Neural Network (CNN) Berbasis Citra Daun," *Jurnal Nasional Teknik Elektro dan Teknologi Informasi*, vol. 14, no. 3, pp. 181–189, Aug. 2025, doi: 10.22146/jnteti.v14i3.18791.
5. B. Adi Prastyo and A. Yuniar Rahman, "The 6th Conference on Innovation and Application of Science and Technology (CIASTECH 2023) Prefix-RTR Seminar Hasil Riset The 6 th Conference on Innovation and Application of Science and Technology (CIASTECH) KLASIFIKASI PENYAKIT DAUN PADI MENGGUNAKAN SUPPORT VECTOR MACHINE MELALUI TEKSTUR DAN WARNA DAUN DENGAN HSV DAN GABOR FILTER INFORMASI ARTIKEL ABSTRAK".
6. B. K. Gulo, Agustinus Rudatyo Himamunanto, and Jatmika, "Deteksi Penyakit Tanaman Padi (Oryza Sativa L.) Menggunakan Support Vector Machine (SVM) Dan Random Forest Pada Citra Daun," *Bulletin of Computer Science Research*, vol. 5, no. 4, pp. 724–733, Jun. 2025, doi: 10.47065/bulletincsr.v5i4.660.
7. R. R. Burhanuddin, "KLASIFIKASI PENYAKIT PADI MELALUI CITRA DAUN MENGGUNAKAN METODE NAIVE BAYES," *Jurnal Informatika dan Teknik Elektro Terapan*, vol. 12, no. 2, Apr. 2024, doi: 10.23960/jitet.v12i2.4012.
8. V. Arinichev, S. V. Polyanskikh, I. V. Arinicheva, and I. O. Sergeeva, "Applications of convolutional neural networks for the detection and classification of fungal rice diseases," in *IOP Conference Series: Earth and Environmental Science*, IOP Publishing Ltd, Mar. 2021. doi: 10.1088/1755-1315/699/1/012020.
9. R. Sharma, V. Kukreja, R. K. Kaushal, A. Bansal, and A. Kaur, "Rice Leaf blight Disease detection using multi-classification deep learning model," *2022 10th International Conference on Reliability, Infocom Technologies and Optimization (Trends and Future Directions), ICRITO 2022*, 2022, doi: 10.1109/ICRITO56286.2022.9964644.
10. S. Kini, P. K. V. Reddy, and S. N. Pai, "Techniques of deep learning and image processing in plant leaf disease detection: a review," *International Journal of Electrical and Computer Engineering (IJECE)*, vol. 13, no. 3, pp. 3029–3040, Jun. 2023, doi: 10.11591/ijece.v13i3.pp3029-3040.
11. N. Jakse, "Classification Techniques in Machine Learning," *Machine Learning in Geomechanics 1: Overview of Machine Learning, Unsupervised Learning, Regression, Classification and Artificial Neural Networks*, pp. 117–144, Jan. 2024, doi: 10.1002/9781394325634.CH4.
12. P. Kumarivastava, J. Shiney, and P. Shan, "Plant Disease Prediction using Image Processing and Soft Computing Algorithms: A Review," *Proceedings of 2nd IEEE International Conference on Computational Intelligence and Knowledge Economy, ICCIKE 2021*, pp. 143–148, Mar. 2021, doi: 10.1109/ICCIKE51210.2021.9410728.
13. V. K. Vishnoi, K. Kumar, and B. Kumar, "A comprehensive study of feature extraction techniques for plant leaf disease detection," *Multimed. Tools Appl.*, vol. 81, no. 1, pp. 367–419, Jan. 2022, doi: 10.1007/S11042-021-11375-0.

14. Y. Bhajgotar and V. Gandhi, “Comparative Study on Image Feature Extraction Techniques,” *Lecture Notes in Networks and Systems*, vol. 434, pp. 597–605, 2022, doi: 10.1007/978-981-19-1122-4_62.
15. Dr. Santosh K C and Bharathi N, “Image Processing Technique on Identification of Leaf Types and Detection of Diseases Portion on Tomato Leaves,” *International Research Journal on Advanced Engineering and Management (IRJAEM)*, vol. 2, no. 04, pp. 564–570, Apr. 2024, doi: 10.47392/IRJAEM.2024.0078.
16. K. Arunkumar and S. Leninisha, “An effective identification between various plant species using shape descriptors and image processing technique,” *Proceedings of 2023 International Conference on Signal Processing, Computation, Electronics, Power and Telecommunication, IConSCEPT 2023*, 2023, doi: 10.1109/ICONSCEPT57958.2023.10170691.
17. M. Rahmaty, “Machine learning with big data to solve real-world problems,” *Journal of Data Analytics*, vol. 2, no. 1, pp. 9–16, Apr. 2023, doi: 10.59615/JDA.2.1.9.
18. S. A. R. Srg, S. A. R. Srg, M. Zarlis, and W. Wanayumini, “Klasifikasi Citra Daun dengan GLCM (Gray Level Co-Occurence) dan K-NN (K-Nearest Neighbor),” *MATRIK : Jurnal Manajemen, Teknik Informatika dan Rekayasa Komputer*, vol. 21, no. 2, pp. 477–488, Mar. 2022, doi: 10.30812/matrik.v21i2.1572.
19. R. K. Gupta and K. Dutta, “Resampling Strategies for Mitigating Class Imbalance of ASD Dataset on the Performance of Machine Learning Classifiers,” *Lecture Notes in Networks and Systems*, vol. 535 LNNS, pp. 225–235, 2023, doi: 10.1007/978-981-99-4284-8_18.
20. Y. A. Sir and A. H. H. Soepranoto, “Pendekatan Resampling Data Untuk Menangani Masalah Ketidakseimbangan Kelas,” *Jurnal Komputer dan Informatika*, vol. 10, no. 1, pp. 31–38, Mar. 2022, doi: 10.35508/JICON.V10I1.6554.
21. R. Irawan, Y. Herry Chrisnanto, and G. Abdilah, “Perbandingan Fungsi Kernel Pada Algoritma SVM untuk klasifikasi Kredit Macet,” *Jurnal Algoritma*, vol. 23, no. 1, May 2026, doi: 10.33364/ALGORITMA/V.23-1.2938.
22. Suci Amaliah, M. Nusrang, and A. Aswi, “Penerapan Metode Random Forest Untuk Klasifikasi Varian Minuman Kopi di Kedai Kopi Konijiwa Bantaeng,” *VARIANSI: Journal of Statistics and Its application on Teaching and Research*, vol. 4, no. 3, pp. 121–127, Dec. 2022, doi: 10.35580/variansiunm31.
23. L. Rahmawati and A. Dwi Indriyanti, “PENERAPAN METODE K-NEAREST NEIGHBOR (K-NN) UNTUK PREDIKSI PENJUALAN PAKAIAN (STUDI KASUS: UMKM KRESNA),” 2024.
24. Deepa, N. Rashmi, and C. Shetty, “A Machine Learning Technique for Identification of Plant Diseases in Leaves,” *Proceedings of the 6th International Conference on Inventive Computation Technologies, ICICT 2021*, pp. 481–484, Jan. 2021, doi: 10.1109/ICICT50816.2021.9358797.
25. G. Joshi and P. Panse, “Investigating the Feasibility of Machine Learning Models for Early Detection of Plant Diseases,” *International Conference on Self Sustainable Artificial Intelligence Systems, ICSSAS 2023 - Proceedings*, pp. 289–294, 2023, doi: 10.1109/ICSSAS57918.2023.10331640.
26. D. Senanu Ametefe *et al.*, “Enhancing leaf disease detection accuracy through synergistic integration of deep transfer learning and multimodal techniques,” *Information Processing in Agriculture*, vol. 12, no. 3, pp. 279–299, Sep. 2025, doi: 10.1016/J.INPA.2024.09.006.