PERFORMANCE ANALYSIS OF K-NN AND SVM IN DIGITAL IMAGE-BASED TEA LEAF DISEASE CLASSIFICATION

P.P.P.A.N.W.Fikrul Ilmi R.H.Zer [1*]; Abdi Rahim Damanik [2]; P.A.M. Zidane R.W.P.P Zer [3]

Informatics Management[1]
Information Systems[2]
Informatics Engineering[3]
STIKOM Tunas Bangsa, Pematangsiantar, Indonesia[1,2,3]
https://stikomtunasbangsa.ac.id[1,2,3]
fikrul@amiktunasbangsa.ac.id[1*], abdirahimdmk@stikomtb.ac.id[2], zidanezer@gmail.com[3]
(*) Corresponding Author

Abstract*—Tea is a commodity with high economic value, but it is susceptible to diseases such as Brown Blight, Red Rust, and Red Spider Mite. The manual identification process currently relies on visual observation, which is time-consuming and prone to error. This research aims to analyze the performance of K-NN and SVM algorithms in classifying tea leaf diseases based on digital images. This research utilized a perfectly balanced dataset of 5,000 images. The research methodology involves image preprocessing and classification using 5-Fold, 10-Fold, and 20-Fold Cross-Validation. The results demonstrate that the SVM algorithm consistently outperforms K-NN across all testing scenarios. Specifically, SVM achieved its highest accuracy of 96.6% using 20-Fold Cross-Validation, whereas the highest accuracy for K-NN was 96.1%. The research concludes that SVM provides superior sensitivity and accuracy for identifying tea leaf diseases, offering a viable solution for automated detection systems in the plantation sector*

Keywords*: Classification, Digital Image, K-Nearest Neighbor, Support Vector Machine, Tea Leaf Disease*

Intisari—Teh merupakan komoditas bernilai ekonomi tinggi, namun rentan terhadap penyakit seperti Brown Blight, Red Rust, dan Red Spider Mite. Proses identifikasi manual saat ini masih mengandalkan pengamatan visual yang memakan waktu dan rentan kesalahan. Penelitian ini bertujuan untuk menganalisis kinerja algoritma K-NN dan SVM dalam mengklasifikasikan penyakit daun teh berbasis citra digital. Penelitian ini menggunakan dataset yang seimbang dengan berjumlah 5.000 gambar. Metodologi penelitian meliputi pra-pemrosesan citra dan klasifikasi menggunakan Cross-Validation 5-Fold, 10-Fold, dan 20-Fold. Hasil penelitian menunjukkan bahwa algoritma SVM secara konsisten mengungguli K-NN di semua skenario pengujian. Secara khusus, SVM mencapai akurasi tertinggi sebesar 96,6% menggunakan 20-Fold Cross-Validation, sedangkan akurasi tertinggi K-NN adalah 96,1%. Penelitian ini menyimpulkan bahwa SVM memberikan sensitivitas dan akurasi yang lebih unggul untuk mengidentifikasi penyakit daun teh, sehingga menjadi solusi yang layak untuk sistem deteksi otomatis di sektor perkebunan.

Kata Kunci*: Klasifikasi, Citra Digital, K-Nearest Neighbor, Support Vector Machine, Penyakit Daun Teh*

INTRODUCTION

Information and Communication Technology (ICT) has influenced many aspects of life, such as agriculture. The existence of Machine Learning (Harakannanavar, Rudagi, Puranikmath, Siddiqua, & Pramodhini, 2022) and Image Processing (Marpaung, 2023) technologies has made plant disease identification faster and more efficient. An example is tea plants, which have high economic value. Tea plants are also part of the national and global beverage industry (Ulfa, 2023). The high public interest in tea is due to its good quality, with raw materials and tea production processes carried out properly (Purba & Sriani, 2024). Tea commodities also have problems such as Brown Blight, Red Rust, and Red Spider Mite diseases that attack tea leaves (Mustikasari, Ramli, & Gibran, 2023; Somesh, Sai, Mude, Surendiran, & Dhakshayani, 2024). The manual disease identification process still relies on visual observation by workers or farmers. This manual technique is time-consuming, results in poor identification, and has the potential to lead to poor

Techno Nusa Mandiri: Journal of Computing and Information Technology
Vol. 23, No. 1 March 2026 | DOI https://doi.org/10.33480/techno.v23i1.8234

tea production. There are many technologies that can identify healthy and unhealthy tea leaves (Li et al., 2024).

One of them is by using digital image-based classification techniques (Harahap, Khairani, & Rismayanti, 2024; Hilmi, Puspaningrum, & Wahanani, 2024). In digital image-based classification, there are two popular methods that have been proven to perform well in pattern recognition, namely the K-Nearest Neighbor (K-NN) (Salsabila, Rozikin, & Adam, 2023; Srg, Zarlis, & Wanayumini, 2022) and Support Vector Machine (SVM) algorithms. These two algorithms have been widely used in various classification cases with digital image data. However, further research on the application of these two algorithms for disease classification in tea leaves is still limited, so research was conducted to discuss the application of these two algorithms.

The purpose of this research was to apply and analyze the performance of the K-Nearest Neighbor and Support Vector Machine algorithms (Syaputra, 2024) on digital images of tea leaves consisting of four classes, namely Healthy, Brown Blight, Red Rust, and Red Spider Mite. In this context, digital images of tea leaves are processed in the stages of preprocessing, feature extraction, model training, and model performance evaluation (Heng, Yu, & Zhang, 2024). The expected result is the best classification model that can be used for an automatic detection system based on digital images. This research is related to the field of Information and Communication Technology (ICT), where Machine Learning (Madhurya, Gururaj, Soundarya, Vidyashree, & Rajendra, 2022) is applied to solve problems in the plantation sector.

The involvement of Information and Communication Technology (ICT) in this research is in the development of digital image-based Machine Learning. This research is planned for TKT 3, which is the proof-of-concept stage, where analysis and experiments are carried out to prove the concept of using the K-Nearest Neighbor and Support Vector Machine algorithms (Assegie, 2021) in the classification of digital images of tea leaf diseases. Thus, this research is expected to produce and provide benefits in the development of ICT science and technology.

MATERIALS AND METHODS

The following are the research methods used to analyze the performance of the K-Nearest Neighbor and Support Vector Machine algorithms :

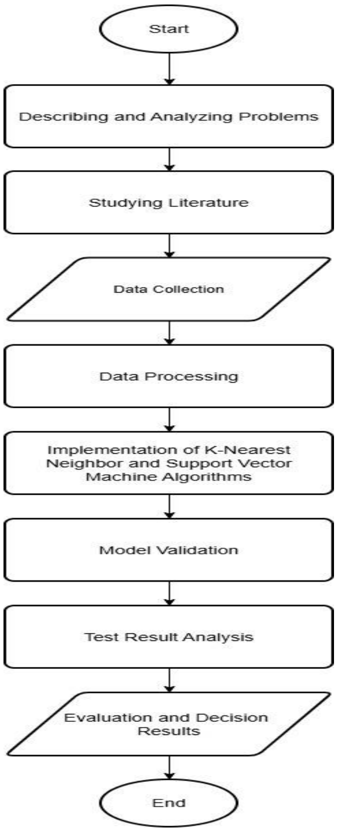
1. Data Type and Source
   The type of data used is Secondary Data, which is obtained from published data that is ready for processing. The data source is obtained from the website www.kaggle.com. We use image data of tea leaf diseases that include four types, Brown Blight, Healthy, Red Rust, and Red Spider Mite. The following is an example of a sample image of tea leaf disease in Table 1.

Table 1. Sample data of tea leaf diseases

Label Name	Picture
Brown Blight	
Healthy	
Red Rust	
Red Spider Mite	

Source : (Research Result, 2026)

2. Research Flowchart
   The following is the research flowchart, which can be seen in Figure 1 :



Source : (Research Result, 2026)
Figure 1. Research Flowchart

Description of the Research Flowchart is as follows:

a. Describing and Analyzing the Problem
This stage explains and describes the problem, namely the identification of tea leaf diseases, which is still done visually by farmers or workers. Next, the problem is analyzed, namely the identification stage, the stage of understanding the data to be used, and the algorithm to be used.

b. Studying the literature
Studying and exploring references related to the problem. The references used were obtained from national journals, international journals, and websites about tea leaf diseases.

c. Data Collection
This stage involves obtaining data from the Kaggle website.

d. Data Processing
The collected data will be processed to prepare it for use as input in identifying tea leaf diseases.

e. Implementation of the K-Nearest Neighbor and Support Vector Machine algorithms
This stage uses the K-NN and SVM algorithms to process the data. In this section, digital images of tea leaf conditions will be studied using the model.

f. Model Validation
After the leaf disease identification model is built, the next step is to validate the model using appropriate techniques.

g. Test Result Analysis
This stage analyzes the results of model testing to determine the best model.

h. Evaluation and Decision Results
The analyzed models will be evaluated based on how well the two algorithms perform in identifying tea leaf diseases. After that, evaluation is carried out by testing based on 5-Fold, 10-Fold, and 20-Fold Cross Validation and getting accuracy. The final step is looking at the results from both methods, and choosing the one with the highest accuracy and best validation. The evaluation results will give decision outcomes that can be used as a reference by the relevant parties.

3. Dataset Description
The dataset utilized in this research consists of a total of 5,000 digital images of tea leaves. The data distribution is perfectly balanced across four classes: Healthy (1,250 images), Brown Blight (1,250 images), Red Rust (1,250 images), and Red Spider Mite (1,250 images). This balanced distribution eliminates the problem of class imbalance during the training phase. The preprocessing and feature extraction stages were executed using the Image Embedding technique. Specifically, the SqueezeNet deep learning model was employed as the embedder. SqueezeNet was selected because it is a highly efficient CNN that achieves AlexNet-level accuracy with 50 times fewer parameters. During this process, the raw images were automatically resized and normalized by the SqueezeNet architecture to extract deep visual features, converting the image data into numerical feature vectors ready for classification.

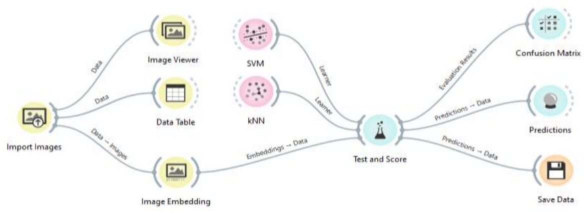
4. Algorithm Configuration and Parameters
To ensure the reproducibility of the experiment, specific parameters were configured for both algorithms using the Orange Data Mining tool :

a. K-Nearest Neighbor (K-NN): The algorithm was set with the number of neighbors $k = 5$. The distance metric used to calculate the similarity between feature vectors was the Euclidean distance, with a Uniform weight applied to all neighbors.

b. Support Vector Machine (SVM): The model was configured as a C-SVM. The regularization parameter or Cost (C) was set to 1.00 to balance the margin maximization and classification error. The Radial Basis Function (RBF) was selected as the kernel to handle the non-linear high-dimensional features extracted by SqueezeNet, with the gamma (γ) parameter set to auto. The optimization process used an iteration limit of 100 and a numerical tolerance of 0.0010.

RESULTS AND DISCUSSION

This research compares the K-NN and SVM algorithms to determine the most accurate and effective performance for application. We used three types of cross-validation to obtain accurate results, namely 5-fold, 10-fold, and 20-fold. The performance of both algorithms was tested using the Orange Data Mining tool. The following is a model for testing the performance of K-Nearest Neighbor and Support Vector on the Orange tool, as shown in Figure 2:



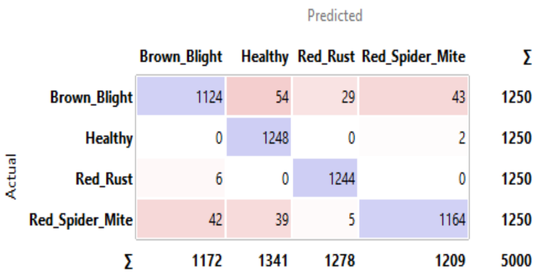
Source : (Research Result, 2026)
Figure 2. Model in the Orange tool

Techno Nusa Mandiri: Journal of Computing and Information Technology
Vol. 23, No. 1 March 2026 | DOI https://doi.org/10.33480/techno.v23i1.8234

Accuracy Results

1. 5-Fold Cross Validation

The results from testing with 5-Fold Cross Validation, using K-Nearest Neighbor and Support Vector algorithms to classify tea leaf diseases can be seen in Figures 3 and 4 below:

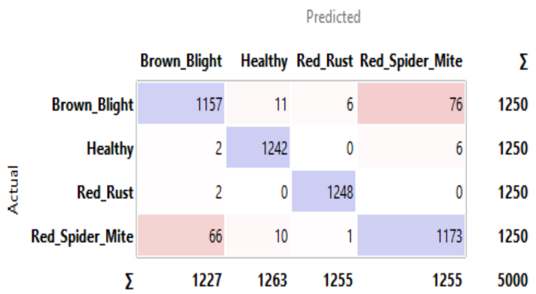


Actual \ Predicted	Brown_Blight	Healthy	Red_Rust	Red_Spider_Mite	Σ
Brown_Blight	1124	54	29	43	1250
Healthy	0	1248	0	2	1250
Red_Rust	6	0	1244	0	1250
Red_Spider_Mite	42	39	5	1164	1250
Σ	1172	1341	1278	1209	5000

Source : (Research Result, 2026)

Figure 3. k-NN Confusion Matrix with 5-Fold

Figure 3 above shows the confusion matrix from the results of testing the K-NN algorithm using 5-fold cross-validation. This model demonstrates the highest prediction accuracy for the Healthy class (1,248 correct predictions) and the Red Rust class (1,244 correct predictions).



Actual \ Predicted	Brown_Blight	Healthy	Red_Rust	Red_Spider_Mite	Σ
Brown_Blight	1157	11	6	76	1250
Healthy	2	1242	0	6	1250
Red_Rust	2	0	1248	0	1250
Red_Spider_Mite	66	10	1	1173	1250
Σ	1227	1263	1255	1255	5000

Source : (Research Result, 2026)

Figure 4. SVM Confusion Matrix with 5-Fold

Figure 4 above shows the confusion matrix from the results of testing the SVM algorithm using 5-fold cross-validation. The Red Rust and Healthy classes recorded the most optimal recognition rates, with 1,248 and 1,242 correct predictions, respectively, out of a total of 1,250 data points per class. Although the overall accuracy is very high, there is still a slight overlap of misclassification errors concentrated between the Brown Blight and Red Spider Mite classes. Based on Figures 3 and 4, the Confusion Matrix results from k-NN and SVM can display can be seen in Table 2.

Table 2. Classification Results with 5-Fold

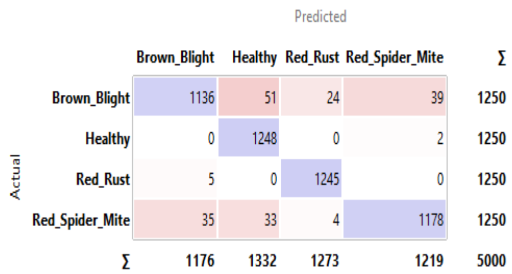
Model	AUC	CA	F1	Prec	Recall
k-NN	0.993	0.956	0.956	0.956	0.956
SVM	0.997	0.964	0.964	0.964	0.964

Source : (Research Result, 2026)

Table 2 presents a comparison of classification results between the k-Nearest Neighbor (k-NN) and Support Vector Machine (SVM) algorithms using a 5-fold cross-validation scenario. SVM achieved a score of 0.964 for accuracy (CA), F1-Score, precision, and recall, as well as a very high Area Under the Curve (AUC) value of 0.997. Meanwhile, the k-NN algorithm achieved a score of 0.956 for accuracy, precision, and recall, with an AUC of 0.993.

2. 10-Fold Cross Validation

The results of testing with 10-Fold Cross Validation using K-Nearest Neighbor and Support Vector to classify tea leaf diseases can be seen in Figures 5 and 6 below:

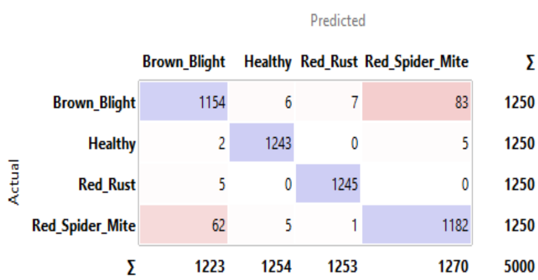


Actual \ Predicted	Brown_Blight	Healthy	Red_Rust	Red_Spider_Mite	Σ
Brown_Blight	1136	51	24	39	1250
Healthy	0	1248	0	2	1250
Red_Rust	5	0	1245	0	1250
Red_Spider_Mite	35	33	4	1178	1250
Σ	1176	1332	1273	1219	5000

Source : (Research Result, 2026)

Figure 5. k-NN Confusion Matrix with 10-Fold

Figure 5 above shows the confusion matrix from the results of testing the K-NN algorithm using 10-fold cross-validation. The K-NN model demonstrates excellent pattern recognition capabilities for the Healthy (1,248 correct predictions) and Red Rust (1,245 correct predictions) classes. Nevertheless, the model experiences slight difficulties in distinguishing between the Brown Blight and Red Spider Mite classes, as evidenced by an increase in misclassifications, where instances were either confused with each other or incorrectly predicted as healthy leaves.



Actual \ Predicted	Brown_Blight	Healthy	Red_Rust	Red_Spider_Mite	Σ
Brown_Blight	1154	6	7	83	1250
Healthy	2	1243	0	5	1250
Red_Rust	5	0	1245	0	1250
Red_Spider_Mite	62	5	1	1182	1250
Σ	1223	1254	1253	1270	5000

Source : (Research Result, 2026)

Figure 6. SVM Confusion Matrix with 10-Fold

Figure 6 above shows the confusion matrix from the results of testing the SVM algorithm using 10-fold cross-validation. The Red Rust and

Healthy classes were recognized with the highest level of precision, recording 1,245 and 1,243 correct predictions respectively out of the 1,250 instances per class. Although the overall performance is highly impressive, the majority of prediction errors (misclassifications) remain concentrated on the feature similarities between the Brown Blight and Red Spider Mite classes, where 83 Brown Blight images were incorrectly predicted as Red Spider Mite, and 62 images experienced the reverse. Based on Figures 5 and 6, the confusion matrix results from k-NN and SVM can display seen in Table 3.

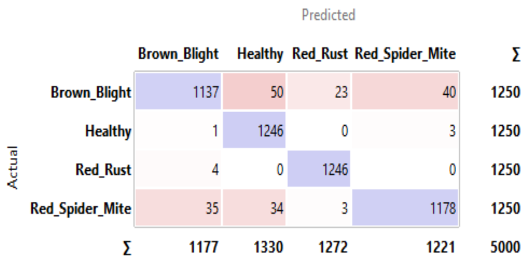
Table 3. Classification Results with 10-Fold

Model	AUC	CA	F1	Prec	Recall
k-NN	0.994	0.961	0.961	0.962	0.961
SVM	0.997	0.965	0.965	0.965	0.965

Source : (Research Result, 2026)

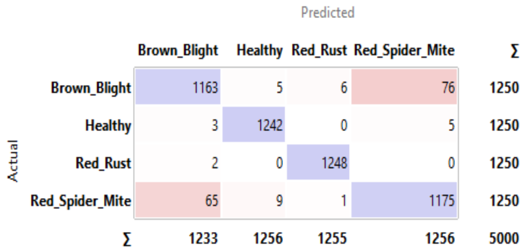
Table 3 presents a comparison of the performance of the k-NN and SVM algorithms based on 10-fold cross-validation testing. The test results show that the SVM model once again outperformed k-NN on all evaluation metrics. SVM recorded a value of 0.965 for accuracy (CA), F1-Score, precision, and recall, and maintained a very high Area Under the Curve (AUC) value of 0.997. Meanwhile, the k-NN algorithm achieved a score of 0.961 for accuracy, F1-Score, and recall, a score of 0.962 for precision, and an AUC of 0.994.

3. 20-Fold Cross Validation
The results of testing with 20-Fold Cross Validation using K-Nearest Neighbors and Support Vector Machine to classify tea leaf diseases can be seen in Figures 7 and 8 below:



Source : (Research Result, 2026)
Figure 7. k-NN Confusion Matrix with 20-Fold

Figure 7 above shows the confusion matrix from the results of testing the K-NN algorithm using 20-fold cross-validation. In this test with a larger number of folds, the K-NN model continues to demonstrate excellent classification capabilities, particularly for the Healthy and Red Rust classes, each of which was correctly identified in 1,246 images out of a total of 1,250 data points per class. However, the model still faced some challenges in extracting feature differences in the Brown Blight class, where misclassification occurred, including 50 Brown Blight images that were incorrectly predicted as healthy leaves (Healthy).



Source : (Research Result, 2026)
Figure 8. SVM Confusion Matrix with 20-Fold

Figure 8 above shows the confusion matrix from the results of testing the K-NN algorithm using 20-fold cross-validation. The Red Rust and Healthy classes recorded the highest recognition rates, with 1,248 and 1,242 correct predictions respectively out of the 1,250 instances per class. Although the overall accuracy reaches its peak in this test, the model still faces minor challenges in differentiating the patterns of the Brown Blight and Red Spider Mite classes. This is indicated by 76 Brown Blight images being incorrectly classified as Red Spider Mite, and 65 images in the reverse case. Based on Figure 7 and Figure 8, the results of the Confusion Matrix from k-NN and SVM can display seen in Table 4.

Table 4. Classification Results with 20-Fold

Model	AUC	CA	F1	Prec	Recall
k-NN	0.994	0.961	0.961	0.962	0.961
SVM	0.998	0.966	0.966	0.966	0.966

Source : (Research Result, 2026)

Table 4 presents a comparison of the performance of the k-NN and SVM algorithms based on 20-fold cross-validation testing. In this maximum fold configuration, the SVM model reaches its peak performance and once again demonstrates an advantage over k-NN. SVM recorded the highest scores across all metrics, achieving 0.966 for Classification Accuracy (CA), F1-Score, precision, and recall, accompanied by a near-perfect Area Under Curve (AUC) value of 0.998. Conversely, the performance of k-NN remained stable but plateaued, yielding scores of 0.961 for accuracy, F1-Score, and recall, 0.962 for precision, and an AUC of 0.994. These results conclusively confirm that the SVM algorithm is the most optimal and robust model for the tea leaf disease classification task in this research.

Vol. 23, No. 1 March 2026 | DOI https://doi.org/10.33480/techno.v23i1.8234

Discussion

The results from the classification tests show that the SVM method produced better classification outcomes compared to the K-NN method. This is supported by the accuracy results from the 5-Fold, 10-Fold, and 20-Fold Cross Validation tests. In the 5-Fold Cross Validation test, the SVM achieved an accuracy of 0.964, which is equivalent to 96.4%, whereas the K-NN model achieved an accuracy of 0.956, or 95.6%. In the 10-Fold test, the SVM achieved an accuracy of 0.965, which is equivalent to 96.5%, whereas the K-NN algorithm reached an accuracy of 0.961, or 96.1%. In the 20-Fold test, the SVM achieved an accuracy of 0.966, which is equivalent to 96.6%, whereas the K-NN algorithm achieved an accuracy of 0.961, or 96.1%. These results explain that the SVM algorithm is better than K-NN in detecting tea leaf diseases for four types of tea leaf conditions. The high Recall value produced, especially by SVM, has crucial practical implications for the tea plantation industry. A high recall value indicates that the system has excellent sensitivity in detecting diseases, thereby minimizing the risk of false negatives, where diseased plants are detected as healthy plants. As for recommendations for further research, it is suggested to add more types of tea leaf diseases to obtain better classification results and accuracy.

The models were assessed through the use of Stratified 5-Fold, 10-Fold, and 20-Fold Cross-Validation. The stratified approach maintains the original class proportions, with 1,250 images per class, in every training and testing split. The use of multiple folds, especially up to 20-fold, was meant to thoroughly evaluate the model's stability, especially considering the large dataset size of 5,000 images. In the 20-fold scenario, the model was trained using 4,750 images and tested on 250 images that had not been seen before in each iteration. The model shows consistent performance across various folds, with accuracy ranging from 96.4% to 96.6% for SVM, indicating it has good ability to generalize and is not overfitting. The features extracted by SqueezeNet were found to be very effective at representing the data, enabling the algorithms to understand and learn the fundamental patterns in the data without simply memorizing the training examples.

Although the absolute difference in accuracy between SVM (96.6%) and K-NN (96.1%) appears relatively small, the performance gap is significant when considering the high-dimensional nature of the embedded features. In addition, the evaluation based on the Area Under the ROC Curve (AUC) metric supports the superior performance of SVM. According to the classification results, the SVM model consistently achieved an almost perfect AUC score of 0.998 across all 20-fold tests, whereas the K-NN model obtained a slightly lower AUC of 0.994. This suggests that SVM is more likely to accurately differentiate between healthy and diseased leaves at different classification levels, which makes it both statistically and practically more reliable for this particular classification task.

CONCLUSION

Based on the analysis and testing conducted, it can be concluded that the Support Vector Machine (SVM) algorithm demonstrates superior performance compared to the K-Nearest Neighbor (K-NN) algorithm in classifying tea leaf diseases. The experimental results using Cross-Validation indicate that SVM consistently achieves higher accuracy, recording 96.4% (5-Fold), 96.5% (10-Fold), and 96.6% (20-Fold), whereas K-NN achieved a maximum accuracy of 96.1%. Beyond accuracy, the SVM model exhibited high recall values, indicating its robustness in minimizing false negatives, which is critical for ensuring diseased plants are not mistaken for healthy ones. For further research, it is suggested to expand the dataset to include a wider variety of tea leaf diseases to further enhance the model's generalization capability.

REFERENCE

Assegie, T. A. (2021). Support Vector Machine And K-Nearest Neighbor Based Liver Disease Classification Model. *Indonesian Journal of Electronics, Electromedical Engineering, and Medical Informatics*, *3*(1), 9–14. https://doi.org/10.35882/ijeeemi.v3i1.2

Harahap, R. I. J., Khairani, S., & Rismayanti. (2024). Implementasi Metode K-Nearest Neighbor Untuk Klasifikasi Penyakit Tanaman Mentimun Pada Citra Daun Implementation Of The K-Nearest Neighbor Method For Classification Of Cucumber Diseases On Leaf Images. *Jurnal Ilmu Komputer Dan Sistem Informasi (JIRSI)*, *3*(2), 135–145. Retrieved from https://jurnal.unity-academy.sch.id/index.php/jirsi/index%0Ahttp://creativecommons.org/licenses/by-sa/4.0/

Harakannanavar, S. S., Rudagi, J. M., Puranikmath, V. I., Siddiqua, A., & Pramodhini, R. (2022). Plant leaf disease detection using computer vision and machine learning algorithms. *Global Transitions Proceedings*, *3*(1), 305–310. https://doi.org/10.1016/j.gltp.2022.03.016

Heng, Q., Yu, S., & Zhang, Y. (2024). A new AI-based approach for automatic identification of tea leaf disease using deep neural network based on hybrid pooling. *Heliyon*, *10*(5), e26465.

https://doi.org/10.1016/j.heliyon.2024.e26465

Hilmi, A. N., Puspaningrum, E. Y., & Wahanani, H. E. (2024). Implementasi Algoritma K-Nearest Neighbor (KNN) untuk Identifikasi Penyakit pada Tanaman Jeruk Berdasarkan Citra Daun. *Router : Jurnal Teknik Informatika Dan Terapan*, *2*(2), 107–117.

Li, Y., Lu, Y., Liu, H., Bai, J., Yang, C., Yuan, H., … Xiao, Q. (2024). Tea leaf disease and insect identification based on improved MobileNetV3. *Frontiers in Plant Science*, *15*(September), 1–12. https://doi.org/10.3389/fpls.2024.1459292

Madhurya, M. J., Gururaj, H. L., Soundarya, B. C., Vidyashree, K. P., & Rajendra, A. B. (2022). Exploratory analysis of credit card fraud detection using machine learning techniques. *Global Transitions Proceedings*, *3*(1), 31–37. https://doi.org/10.1016/j.gltp.2022.04.006

Marpaung, F. (2023). Klasifikasi Daun Teh Menggunakan Metode CNN Dan Arsi Mobilenetv2.

Mustikasari, M., Ramli, A. R., & Gibran, A. K. (2023). Analisis Kinerja Algoritma Machine Learning Untuk Deteksi Penyakit Daun Teh Dengan Particle Swarm Optimization. *Journal of Embedded Systems, Security and Intelligent Systems*, *4*(2), 97–106. https://doi.org/10.59562/jessi.v4i2.708

Purba, S. A. A., & Sriani, S. (2024). Klasifikasi Hasil Produksi Dauh Teh Basah (DTB) Menggunakan Algoritma K-Nearest Neighbor (KNN). *Informatika*, *16*(2), 313–319.

Salsabila, A. P. B., Rozikin, C., & Adam, R. I. (2023). Klasifikasi Motif Batik Karawang Berbasis Citra Digital dengan Principal Component Analysis dan K-Nearest Neighbor. *Jurnal Sistem Dan Teknologi Informasi (JustIN)*, *11*(1), 20. https://doi.org/10.26418/justin.v11i1.46936

Somesh, K., Sai, C. S. S., Mude, N., Surendiran, B., & Dhakshayani, J. (2024). Enhancing Tea Leaf Disease Classification: Leveraging Data Augmentation, Diverse Feature Extraction Techniques, and Ensemble Stacking with Machine Learning Models. *ICDLAIR*, (Icdlair 2024), 149–168. https://doi.org/10.2991/978-94-6463-740-3_14

Srg, S. A. R., Zarlis, M., & Wanayumini, W. (2022). Klasifikasi Citra Daun dengan GLCM (Gray Level Co-Occurence) dan K-NN (K-Nearest Neighbor). *MATRIK : Jurnal Manajemen, Teknik Informatika Dan Rekayasa Komputer*, *21*(2), 477–488. https://doi.org/10.30812/matrik.v21i2.1572

Syaputra, A. (2024). Klasifikasi Penyakit Daun pada Tebu dengan Pendekatan Algoritma K-Nearest Neighbors, Multilayer Perceptron dan Support Vector Machine. *Jurnal Ilmiah Informatika Global*, *15*(3), 131–138. https://doi.org/10.36982/jiig.v15i3.4856

Ulfa, M. (2023). *PERBANDINGAN METODE LEARNING VECTOR QUANTIZATION DAN SUPPORT VECTOR MACHINE DALAM MENGKLASIFIKASI PENYAKIT DAUN TEH MENGGUNAKAN EKSTRAKSI FITUR SIFT*. Retrieved from https://repositori.uma.ac.id/handle/123456789/20323%0Ahttps://repositori.uma.ac.id/jspui/bitstream/123456789/20323/1/178160063 - Mutia Ulfa - Fulltext.pdf