Matrik: Jurnal Manajemen, Teknik Informatika, dan Rekayasa Komputer
Vol. 22, No. 1, November 2022, pp. 177∼188
ISSN: 2476-9843, accredited by Kemenristekdikti, Decree No: 200/M/KPT/2020
DOI: 10.30812/matrik.v22i1.2022

Support Vector Machine with Firefly Optimization Algorithm for Apple Fruit Disease Classification

Amien Jafar Makrufi[1], Wikky Fawwaz Al Maki [2]
Telkom University, Bandung, Indonesia

Article Info

Article history:

Received July 21, 2022
Revised September 20, 2022
Accepted November 24, 2022

Keywords:

Apple Disease
Classification
Computer Vision
Support Vector Machine
Firefly Algorithm

ABSTRACT

Fruit is one of the foods that contain vitamins that are beneficial to human health. Therefore, many farmers cultivate various kinds of fruit according to the geographical climate and the target market farmers. However, in the cultivation process, fruit diseases are often encountered which is one of the serious problems faced by farmers because it can threaten their economic results. The main focus of this research object is the identification and classification of diseases in apples. Apples are very susceptible to disease, in general the diseases that usually attack apples are blotch apple, apple rot, and apple scab. For some of these diseases, the type of disease is still using the manual method with the help of human labor. This method certainly has many shortcomings and takes a long time. Meanwhile, proper and rapid disease sorting is needed to anticipate the occurrence of repeated disease attacks. The purpose of the system built by the researcher is to classify the types of apple diseases and healthy apples. This research utilizes computer vision and machine learning technology to solve classification problems. The system developed uses image processing such as augmentation, extraction of dimension reduction features with Principal Component Analysis. The classification algorithm used is the Support Vector Machine (SVM) combined with the Firefly optimization algorithm (FA). The system that has been built can classify the types of diseases of apples and normal apples with the highest accuracy of 94%.

Corresponding Author:

Wikky Fawwaz Al Maki,
School of Computing,
Telkom University, Bandung, Indonesia,
Email: wikkyfawwaz@telkomuniversity.ac.id

How to Cite: W. Fawwaz Al Maki and A. Jafar Makrufi, Support vector machine with a firefly optimization algorithm for classification of apple fruit disease, MATRIK : Jurnal Manajemen, Teknik Informatika dan Rekayasa Komputer, vol. 22, no. 1, pp. 179-190, Nov. 2022.

1. INTRODUCTION

Apples are one of the fruits with the highest number of cultivars in this world, for example in Europes inventory has registered more than ten thousand cultivars. The amount of cultivars varies a lot in quality [1]. There are so many variants of nutrition that are contained in apples, which are antioxidants, vitamins, and fibre. These nutrients are useful for lung health, cholesterol, diabetes, and other diseases. We can say that apple composition is balanced, with the most composition being 84% of water. This fruit contained high glucose, and abundant fibre [2]. Apples are a horticultural crop that is widely cultivated in temperate and tropical regions, every year an apple production request is always one of the highest [3]. Various varieties of apples are cultivated and mass-produced such as Red Delicious, Golden Delicious, Red Roman, and Fuji [4].

An inhibiting factor to producing an apple fruit is because of the diseases that attacked the fruits. Apple got attacked causing a massive loss in the agricultural industry [5]. Some diseases that are found by a farmer in apple fruits are apple scabs, apple blotch, and apple rot [6]. Apple rot has symptoms on the skin that show a brown pattern, apple blotch with black dots, and apple rot the skin would be brown, mushy, and smells bad [7].

Support vector machines are one of the parts of a popular machine learning method to overcome a few classification problems [8]. There are some methods of machine learning that are popular besides SVM, there are random forests and extreme learning techniques [9]. SVM method has already been implemented for a lot of problems in real life, like detecting handwriting, face recognition, detecting disease in the human brain, and so many more [10]. To improve the ability of SVM there are a few modifications that have been done, like combining with the evolving algorithm this to used to increase the classification accuracy and also the parameter optimization value [11]. Optimization algorithms, like Genetic Algorithm (GA) and Particle swarm optimizations (PSO) have been used as a weight selection, feature selection, and optimizing parameter from the SVM algorithm [12], [13].

In these past years, agricultural researchers have used machine learning to solve classification challenges such as disease classification, quality classification, and plant species classification [14]. Similar to Manu Bhagat's research [15], this proposal is for research on the classification of plant leaf diseases using the Grid Search technique as parameter optimization in the SVM method. Hareem Kibriya [16], suggested a study that will identify and categorize plant diseases using CNN and SVM methods. By combining the two deep learning algorithms feature extraction and classification with machine learning, getting the highest accuracy of 98.9%. Akshaya Aruraj [17], proposes research using Local Binary Pattern (LBP) and SVM techniques to detect and classify diseases in banana plants. LBP is used to perform feature extraction which will later be thrown into the SVM classification, the highest accuracy is obtained at 90.9%. KR Aravind [18], proposed a study using a bag of features and multiclass SVM for disease classification in maize. Leaf images are searched for features using a bag of features and classified using SVM, the highest accuracy was achieved at 83.7%.

In [17] and [18], SVM method has been employed in the classification task. However, the proposed models provide the accuracy less than 91%. It is important to increase the performance of SVM in the classification task. Therefore, the accuracy of the classification system becomes higher. In our approach, we did not combine SVM with a deep learning algorithm as suggested in [16]. Deep learning algorithms have demonstrated impressive results in the classification tasks. Therefore, increasing the performance of SVM without involving a deep learning algorithm is more challenging. Our approach is also different from [19]. In [19], the authors employed the Nave Bayes algorithm to classify apple fruit disease. The authors did not modify or optimize the Nave Bayes algorithm to increase the accuracy of the classification system. They proposed a novel thresholding algorithm to increase the accuracy of the classification system. The experimental results demonstrated that the proposed thresholding algorithm can contribute to the higher accuracy than Otsu and Kapur thresholding algorithm can. In the apple fruit disease classification task, optimization of SVM has been studied before. The contribution of our research is the development of optimized SVM. In this connection, the Firefly algorithm is employed to optimize the SVM parameters. Also, another contribution of our research is the use of a dimensionality reduction algorithm in the classification task. By applying a dimensionality reduction algorithm, the number of features can be decreased. In this connection, we employ the PCA as the dimensionality reduction algorithm. Since the number of features is decreased, the classification task becomes faster. In our research, the proposed algorithm will be implemented in the classification of apple fruit disease. There are three different apple fruit diseases discussed in this paper, i.e., scab apple, rot apple, and blotch apple.

2. RESEARCH METHOD

This section discusses the proposed research method for the problem of disease classification in apples, by utilizing one of the machine learning methods, namely the support vector machine. To enhance classification performance, the Firefly Algorithm technique will be used to search for optimization parameters in the SVM classification model. The classification model for apple disease issues that we employ with SVM and FA-SVM is depicted in Figure 1.

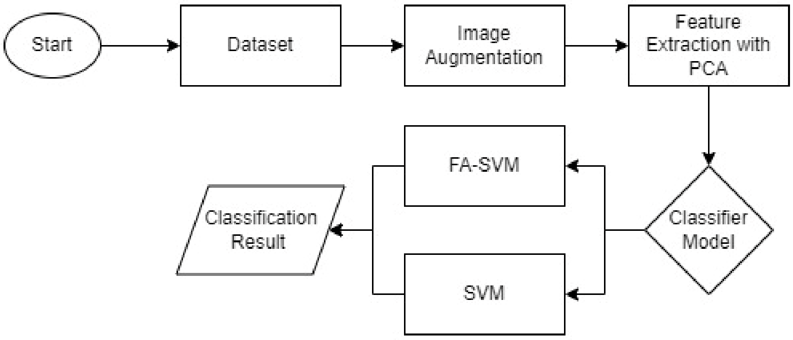

Figure 1. Research Methods

2.1. Dataset

The dataset utilized in this work is an image of an apple divided into four classes, including three disease-related classes: blotch, rot, and scab, as well as a class for normal apples. Researchers used the datasets available on the Kaggle website in the study to classify them using the SVM algorithm and the Firefly optimization algorithm. The dataset for Apple's Blotch includes 146 images, the dataset for Apple's Rot includes 152 images, the dataset for Apple's Scab includes 80 images, and the dataset for Apple's Normal includes 91 images. The photos that follow are examples of each class, as seen in Figure 2.

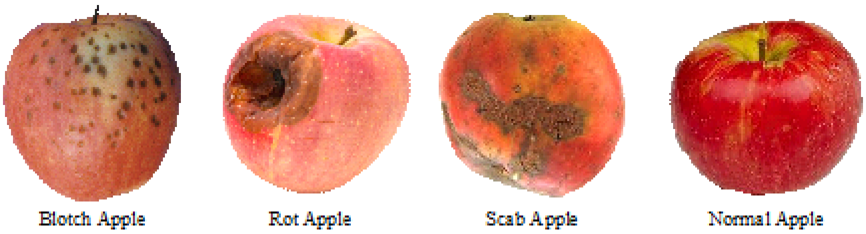

Figure 2. Sample image of the disease class of apples and normal apples

2.2. Augmentation

Data augmentation is a method of modifying the original image into a new sample. This technique is quite useful for machine learning tasks. Data augmentation is frequently used to address problems with a lack of datasets that are required; by doing so, problems can be solved more quickly and the effectiveness of the used classifications is increased [20]. In this study, the researchers used an augmentation technique to increase the amount of data that was used. The parameters used include rotation range of 30, rescale, zoom range of 0.2, horizontal flip, shear range of 0.1, vertical flip and the use of the closest fill mode. After the process of data augmentation was completed, the total number of images was 3356; each class consisted of 1314 blotch, 1364 rot, 315 scab, and 350 normal apple images as shown in Table 1. Each dataset was then resized to 128x128 pixels.

Table 1. Dataset Augmentation

Dataset	Original Data	Augmentation Data
Normal apple	91	350
Blotch Apple	146	1314
Rot Apple	152	1367
Scab Apple	80	315
Amount	469	3346

2.3. Feature Extraction with PCA

The feature extraction process is used to diagnose or determine the value of several problems such as classification, detection, and grouping [21]. In computer vision, feature extraction is the single most important long-term procedure that is used to extract the most important information from the data being used. Principal component analysis (PCA) will be used in this study's investigation to extract the dataset. PCA is a commonly used statistical technique that is based on the teaching of machines, and it can be used to access localized arrays [22]. PCA performs an analysis on data that contains important information to create a new component by reducing the size of the data to its primary component. In this study, the researchers used five different component numbers, 10, 20, 30, 40, and 50, to experiment. Equation (2.3.) contains the covariance matrix calculation.

$$S_t = \frac{1}{N}\Sigma_{i=1}^{N}(x_i - \bar{x})(x - \bar{x})^T \tag{1}$$

$$\text{where } \bar{x} = \frac{1}{n}\Sigma_{i=i}^{n} x_i$$

2.4. Support Vector Machine

Support Vector Machine (SVM) is a machine learning classification method that was first used to analyze neuroimaging data a few years ago. The simple and flexible SVM method, which provides predictions that are consistently balanced, is frequently used to solve classification problems [8]. In the early 1970s, Vapnik developed the SVM algorithm in kernel study, which is frequently used in machine learning, particularly in picture classification. According to the assumptions made, SVM is best used to linearly classify a biner by creating a single limit between two classes. Some problems cannot always be resolved by linear means; therefore, a kernel technique is required to resolve these issues, as shown in Figure 3 by Vapnik [23]. SVM utilizes the hyperplane technique to separate the margin distance between the closest classes to the maximum. SVM is a high-dimensional method in the learning system in the form of a linear function [24].

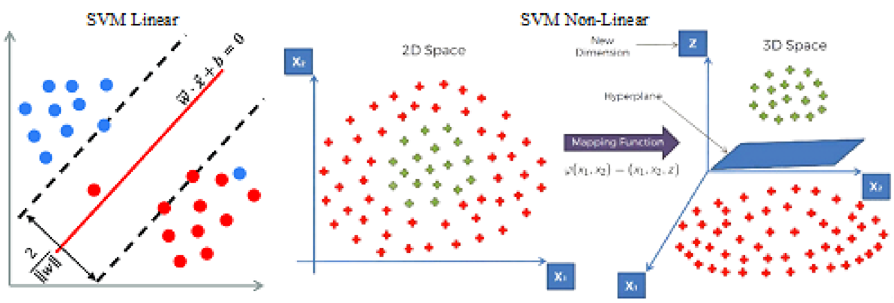

Figure 3. SVM Linear and Non-Linear

2.5. Firefly Algorithm

The Firefly Algorithm (FA) is a technique from swarm intelligence that is frequently used to solve nonlinear optimization problems because it is effective when applied to parameters with minimum and maximum problems [25]. FA is an algorithm optimization that was inspired by an environment that has been actively changing for the past ten years [26]. FA work to find the best solutions using light flashes, which have two characteristics that are all consistently likely to spook potential predators [27]. Equation (2), Equation (3), Equation (4), and Equation (5) are examples of FA interactions.

$$I \propto \frac{1}{r^2} \tag{2}$$

$$I = I_0 e^{-yr^2} \tag{3}$$

$$\beta = \beta_0 e^{-yr^2} \tag{4}$$

$$\beta = \beta_0 e^{-yr^m} \tag{5}$$

According to Equation (2), the inverse square law will cause the light intensity to decrease. If the material allows light to flow through, the light intensity will be at r, where I_0 is the light intensity, as in Equation (3). While the brightness is the same as that in Equation (3), it is generally similar to that in Equation (4). The algorithm of the base firefly:

1:	$t = 0; s* = \phi; \gamma = 1.0;$	gen. counter, best solution, attractiveness
2:	$P^{(0)} = InitializeFA();$	Population
3:	**while** $(t < MAX\_FES)$ **do**	
4:	$a^{(t)} = AlphaNew();$	A new value of a
5:	$EvaluateFA(P^{(t)}, f(s));$	Evaluate s according to f(s)
6:	$OrderFA(P^{(t)}, f(s));$	Sort a according to f(s)
7:	$s^* = FindTheBestFA(P^{(t)}, f(s));$	The best solution
8:	$P^{(t+1)} = MoveFA(P^{(t)});$	Vary the attractiveness accordingly
9:	$t = t + 1;$	
10:	**end while**	

All fireflies are gender-neutral, and their light intensity values determine how attractive they are. The terrain of the fitness function has an impact on this intensity. The population of fireflies is initialized using the function 'InitializeFA,' and this is often done at random. Lines (310) of the while loop part of the aforementioned method function as the firefly's search area. This search is conducted using the MAX FES function's maximum allowed number of evaluations.

2.6. Classification SVM with the Optimizing Parameter FA

In this part, we employ the SVM classification method and optimize the parameters by utilizing the FA technique. The RBF kernel is used in this implementation, and the parameters gamma and cost were found to be optimized. FA will work to find the optimal value for both of them. Parameter settings for FA optimization as shown in Table 2.

Table 2. Parameter Firefly Algorithm

Parameter	Numbers
Population	30
Cost (C)	[1, 100]
Gamma (γ)	[0.0001, 1]
Evaluation	1000

2.7. Evaluation Method

In this study, the evaluation method used is the confusion matrix. The confusion matrix evaluation method is an evaluation method that is often used in classification cases by determining the results of the predicted values. Values generated from the confusion matrix method such as accuracy, precision, recall, and F1-score values [28].

3. RESULT AND ANALYSIS

Four separate dataset comparison tests were run as part of this study. According to the feature extraction section above, each experiment employs a different PCA component value, specifically values 10, 20, 30, 40, and 50. Table 3 displays the experimental model.

Table 3. Model Experiment

Experiment	Percentage	Train	Test
First	90:10	3011	335
Second	80:20	2677	669
Third	70:30	2342	1004
Fourth	60:40	2008	1348

3.1. Dataset Augmentation

The dataset that is owned is carried out by an augmentation process to obtain new data from the modification of the original data. This is done to increase the number of original datasets so that the accuracy of results will increase. The results of the augmentation process are shown as shown in Figure 4. The figure shows some examples of images from the zooming, rotation, flip and resize images of the augmentation process.

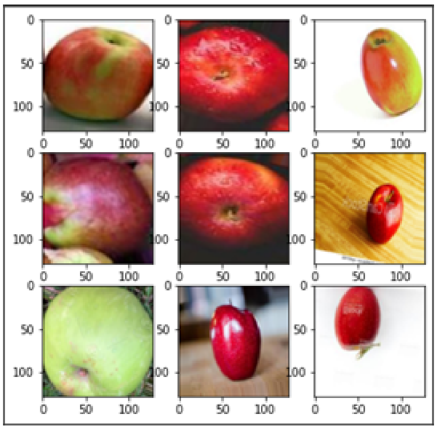

Figure 4. Result of Augmentation Process

3.2. Experiment with Feature Extraction

The experiment was carried out four times as shown in Table 3. This was done to ensure whether the addition of the Firefly optimization algorithm to the SVM always increased accuracy. Each value of the principal component analysis (PCA) is trained using the usual support vector machine (SVM) algorithm and the firefly algorithm-support vector machine (FA-SVM) algorithm. Feature extraction with PCA is used to reduce the feature dimensions of the data set used, so that it will get a data set with more optimal dimensions. All experiments yielded the following accuracy:

1. First Experiment
   In the first experiment the dataset is divided into 90% for training data and 10% for test data. The results show that the FA-SVM algorithm has better accuracy than the SVM algorithm on each PCA component value. The highest results show the highest accuracy of 94% on the component value of PCA 40, as shown in Table 4.

Table 4. First Experiment Result

Algorithm	PCA Component Value				
	10	20	30	40	50
	Accuracy				
SVM	80	82	84	85	85
FA-SVM	93	93	93	94	91

2. Second Experiment
   In the second experiment the dataset is divided into 80% for training data and 20% for test data. The results of this second test obtained FA-SVM accuracy of 93%, an increase of about 9% compared to using ordinary SVM as shown in Table 5.

Table 5. Second Experiment Result

Algorithm	PCA Component Value				
	10	20	30	40	50
	Accuracy				
SVM	79	82	84	85	85
FA-SVM	93	93	92	91	91

3. Third Experiment
   In the third experiment the dataset is divided into 70% for training data and 30% for test data. The result is that the average value of each PCA component gets an accuracy of 90% on the FA-SVM algorithm. The results of this third test the FA-SVM algorithm is still better than the usual SVM algorithm as shown in Table 6.

Table 6. Third Experiment Result

Algorithm	PCA Component Value				
	10	20	30	40	50
	Accuracy				
SVM	79	81	84	84	84
FA-SVM	90	90	90	90	88

4. Fourth Experiment
   In the fourth experiment the dataset is divided into 60% for training data and 40% for test data. Although it experienced a decrease in trances from the previous tests, the use of the FA-SVM algorithm was still better than the usual SVM algorithm. The highest accuracy was obtained at 91% on the FA-SVM algorithm with a component value of PCA 10 as shown in Table 7.

Table 7. Fourth Experiment Result

Algorithm	PCA Component Value				
	10	20	30	40	50
	Accuracy				
SVM	80	82	84	85	84
FA-SVM	91	89	89	89	88

3.3. Result Experiment

Figure 5 shows the results obtained from each test and each component value in the PCA. This graph shows that the FA-SVM algorithm always provides better performance than the regular SVM algorithm in every experiment carried out. The highest accuracy results using the ordinary SVM algorithm is 85% while using the FA-SVM algorithm can achieve an accuracy of 94%.

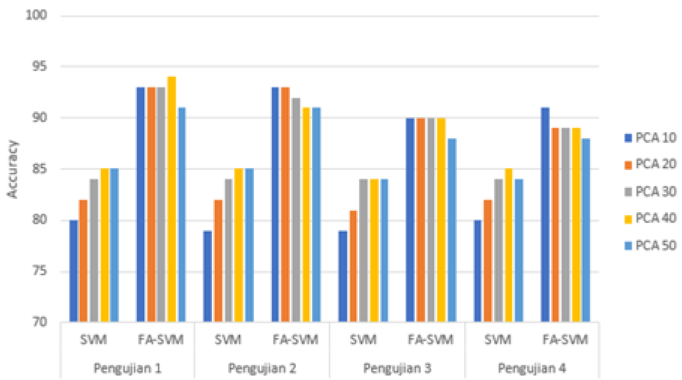

Figure 5. Graph of each Experiment

3.4. Support Vector Machine with Firefly Algorithm

The use of the firefly optimization algorithm (FA) on the support vector machine model will get optimal parameters. In this study, the search for optimal Cost (C) and gamma () parameter values proved to be able to produce an increase in classification accuracy. The combination of the two algorithms is very helpful to improve the accuracy of using the usual SVM algorithm.

The average increase in accuracy from the test scenarios that have been carried out on the ordinary SVM algorithm with the FA-SVM algorithm is 9.6%, 9%, 7.2%, and 6.2%, respectively, as shown in Figure 6. If the accumulation of all test scenarios, the average accuracy increases by 8%.

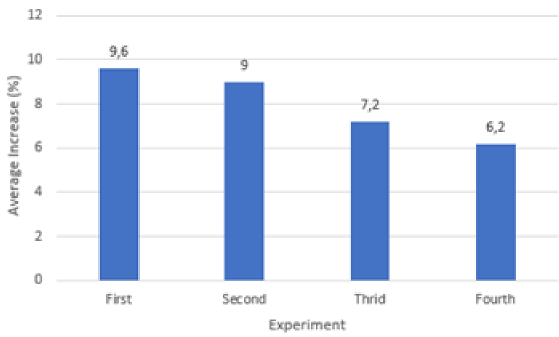

Figure 6. Graph of Average Increase in Accuracy

3.5. Evaluation

The Confusion Matrix is used in this study as an evaluation of the performance of the model used. The results of the four experiments conducted consistently show that the FA-SVM algorithm is better than the standard SVM technique. Trials on five different PCA component values have been conducted to test this, as shown in Table 8. The highest results get an accuracy of 94% which is obtained in test 1 with a value of C 73.16 and gamma (γ) 0.0001.

Table 8. Experiment Result

Experiment	Algorithm	PCA	Accuracy	Precision	Recall	F1-Score
First	FA-SVM	40	94%	93%	90%	92%
Second	FA-SVM	10	93%	92%	90%	91%
Third	FA-SVM	10	90%	89%	85%	87%
Fourth	FA-SVM	10	91%	90%	85%	87%

We also compare our experimental results with those of methods discussed in [19] and SVM. Table 9 shows the performance comparison of some algorithms in apple fruit disease classification task. From Table 9, it is clear that the proposed model can provide a high accuracy. Also, the purpose of our research, i.e. increasing accuracy of SVM by optimizing its parameters (using Firefly algorithm), has been achieved.

Table 9. Performance comparison

No	Algorithm	Accuracy
1	SVM	85%
2	FA-SVM (proposed method)	94%
3	Nave Bayes - Otsu	65%
4	Nave Bayes- Kapur	93.33%

4. CONCLUSION

This study shows the performance of the FA-SVM algorithm which is implemented in the classification of apple diseases. The experimental results demonstrate that the accuracy obtained from the FA-SVM algorithm is always superior to the ordinary SVM algorithm. We have confirmed these results by conducting four different tests, namely dividing the dataset by four comparisons and also trying five component values in different PCA feature extractions. This test proved effective to increase accuracy by adding a firefly optimization algorithm.

Utilization of the optimal parameter selection technique in the use of FA optimization and combined with the SVM classification method will result in a fairly high increase in accuracy. The difference in each test is in the distribution of the number of trains and test data, as well as the use of five different PCA component values. The highest accuracy was obtained in the first experiment with a component PCA value of 40, the results showed an accuracy of 94% for the FA-SVM algorithm. This research has contributed to our findings where we proved that the firefly optimization algorithm (FA) can be used as parameter optimization in the support vector machine (SVM) classification model to maximize accuracy results in image classification, and the results always exceed the usual SVM algorithm.

Compared to experimental results of Nave Bayes-based model, the use of Firefly algorithm in optimizing the parameters of SVM has demonstrated an impressive result in the apple fruit disease classification system. We have not involved any image segmentation algorithm or thresholding algorithm in the proposed classification system. In future works, we will investigate the influence of some segmentation algorithm or thresholding algorithm to the accuracy of the apple fruit disease classification system. Also, we will investigate the influence of different color space models used in pre-processing stage. The proposed algorithm works well with the balanced dataset. Classification task using imbalanced data has not been tackled by the proposed algorithm. This limitation should be studied further and solved in the future.

5. ACKNOWLEDGEMENTS

We would like to express our sincere thanks to Telkom University that has provided a research grant for completing this research. In this research, the first author contributed to data collection, article creation, data analysis and interpretation whereas the second author contributed to the research concept, design, data analysis and interpretation, and revision of the article.

6. DECLARATIONS

AUTHOR CONTIBUTION

Amien Jafar Makrufi contributed to data collection, article creation, data analysis and interpretation whereas Wikky Fawwaz Al Maki contributed to the research concept, design, data analysis and interpretation, and revision of the article.

FUNDING STATEMENT

This work was supported by Telkom University.

COMPETING INTEREST

All authors confirm that there have been no competing personal interests, financial interests, or any involvement that could influence the work reported in this research paper.

REFERENCES

[1] S. Musacchi and S. Serra, "Apple Fruit Quality: Overview on Pre-Harvest Factors," *Scientia Horticulturae*, vol. 234, pp. 409–430, apr 2018.

[2] M. Acquavia, R. Pascale, L. Foti, G. Carlucci, L. Scrano, G. Martelli, M. Brienza, D. Coviello, G. Bianco, and F. Lelario, "Analytical Methods for Extraction and Identification of Primary and Secondary Metabolites of Apple (Malus domestica) Fruits: A Review," *Separations*, vol. 8, no. 7, pp. 1–26, jun 2021. [Online]. Available: https://www.mdpi.com/2297-8739/8/7/91

[3] A. A. Bracino, R. S. Concepcion, R. A. R. Bedruz, E. P. Dadios, and R. R. P. Vicerra, "Development of a Hybrid Machine Learning Model for Apple (Malus Domestica) Health Detection and Disease Classification," in *Proc. IEEE 12th International Conference on Humanoid, Nanotechnology, Information Technology, Communication and Control, Environment, and Management (HNICEM)*. IEEE, dec 2020, pp. 1–6.

[4] Y. Huang, Y. Yang, Y. Sun, H. Zhou, and K. Chen, "Identification of Apple Varieties using a Multichannel Hyperspectral Imaging System," *Sensors*, vol. 20, no. 18, pp. 1–11, sep 2020. [Online]. Available: https://www.mdpi.com/1424-8220/20/18/5120

[5] P. K. Sujatha, J. Sandhya, J. S. Chaitanya, and R. Subashini, "Enhancement of Segmentation and Feature Fusion for Apple Disease Classification," in *Proc. IEEE Tenth International Conference on Advanced Computing (ICoAC)*. IEEE, dec 2018, pp. 175–181.

[6] Sumanto, Y. Sugiarti, A. Supriyatna, I. Carolina, R. Amin, and A. Yani, "Model Naïve Bayes Classifiers for Detection Apple Diseases," in *Proc. IEEE 9th International Conference on Cyber and IT Service Management (CITSM)*. IEEE, sep 2021, pp. 1–4.

[7] A. G. Alharbi and M. Arif, "Detection and Classification of Apple Diseases using Convolutional Neural Networks," in *2020 2nd International Conference on Computer and Information Sciences (ICCIS)*. IEEE, oct 2020, pp. 1–6. [Online]. Available: https://ieeexplore.ieee.org/document/9257640/

[8] D. A. Pisner and D. M. Schnyer, "Support Vector Machine," in *Machine Learning*. Elsevier, jan 2020, pp. 101–121.

[9] I. Ahmad, M. Basheri, M. J. Iqbal, and A. Rahim, "Performance Comparison of Support Vector Machine, Random Forest, and Extreme Learning Machine for Intrusion Detection," *IEEE Access*, vol. 6, pp. 33 789–33 795, may 2018. [Online]. Available: https://ieeexplore.ieee.org/document/8369054/

[10] S. Ghosh, A. Dasgupta, and A. Swetapadma, "A Study on Support Vector Machine based Linear and Non-Linear Pattern Classification," in *Proc. International Conference on Intelligent Sustainable Systems (ICISS)*. IEEE, feb 2019, pp. 24–28.

[11] J. Cervantes, F. Garcia-Lamont, L. Rodríguez-Mazahua, and A. Lopez, "A Comprehensive Survey on Support Vector Machine Classification: Applications, Challenges and Trends," *Neurocomputing*, vol. 408, no. 1, pp. 189–215, sep 2020. [Online]. Available: https://linkinghub.elsevier.com/retrieve/pii/S0925231220307153

[12] P. Tao, Z. Sun, and Z. Sun, "An Improved Intrusion Detection Algorithm Based on GA and SVM," *IEEE Access*, vol. 6, pp. 13 624–13 631, mar 2018.

[13] F. Maulidina, Z. Rustam, and J. Pandelaki, "Lung Cancer Classification using Support Vector Machine and Hybrid Particle Swarm Optimization-Genetic Algorithm," in *Proc. International Conference on Decision Aid Sciences and Application (DASA)*. IEEE, dec 2021, pp. 751–755.

[14] U. Shruthi, V. Nagaveni, and B. Raghavendra, "A Review on Machine Learning Classification Techniques for Plant Disease Detection," in *Proc. IEEE 5th International Conference on Advanced Computing & Communication Systems (ICACCS)*. IEEE, mar 2019, pp. 281–284.

[15] M. Bhagat, D. Kumar, I. Haque, H. S. Munda, and R. Bhagat, "Plant Leaf Disease Classification using Grid Search Based SVM," in *Proc. IEEE 2nd International Conference on Data, Engineering and Applications (IDEA)*. IEEE, feb 2020, pp. 1–6.

[16] H. Kibriya, I. Abdullah, and A. Nasrullah, "Plant Disease Identification and Classification using Convolutional Neural Network and SVM," in *2021 International Conference on Frontiers of Information Technology (FIT)*. IEEE, dec 2021, pp. 264–268. [Online]. Available: https://ieeexplore.ieee.org/document/9701389/

[17] A. Aruraj, A. Alex, M. Subathra, N. Sairamya, S. T. George, and S. V. Ewards, "Detection and Classification of Diseases of Banana Plant using Local Binary Pattern and Support Vector Machine," in *Proc. IEEE 2nd International Conference on Signal Processing and Communication (ICSPC)*. IEEE, mar 2019, pp. 231–235.

[18] K. R. Aravind, P. Raja, K. Mukesh, R. Aniirudh, R. Ashiwin, and C. Szczepanski, "Disease Classification in Maize Crop using Bag of Features and Multiclass Support Vector Machine," in *Proc. IEEE 2nd International Conference on Inventive Systems and Control (ICISC)*. IEEE, jan 2018, pp. 1191–1196.

[19] P. Chithra and M. Henila, "Apple Fruit Sorting using Novel Thresholding and Area Calculation Algorithms," *Soft Computing*, vol. 25, no. 1, pp. 431–445, 2021.

[20] A. Hilal, I. Arai, and S. El-Tawab, "DataLoc+: A Data Augmentation Technique for Machine Learning in Room-Level Indoor Localization," in *Proc. IEEE Wireless Communications and Networking Conference (WCNC)*, vol. 2021-March. IEEE, mar 2021, pp. 1–7.

[21] W. K. Mutlag, S. K. Ali, Z. M. Aydam, and B. H. Taher, "Feature Extraction Methods: A Review," *Journal of Physics: Conference Series*, vol. 1591, no. 1, pp. 1–10, jul 2020. [Online]. Available: https://iopscience.iop.org/article/10.1088/1742-6596/1591/1/012028

[22] S. X. Wu, H.-T. Wai, L. Li, and A. Scaglione, "A Review of Distributed Algorithms for Principal Component Analysis," *Proc. IEEE*, vol. 106, no. 8, pp. 1321–1340, aug 2018.

[23] M. Sheykhmousa, M. Mahdianpari, H. Ghanbari, F. Mohammadimanesh, P. Ghamisi, and S. Homayouni, "Support Vector Machine Versus Random Forest for Remote Sensing Image Classification: A Meta-Analysis and Systematic Review," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 13, pp. 6308–6325, 2020.

[24] N. Putu, N. Hendayanti, M. Nurhidayati, S. Soraya, H. Ratu, and P. Negara, "Community Purchase Decision Modeling in Bali with Non-Linier Methods," *MATRIK : Jurnal Manajemen, Teknik Informatika dan Rekayasa Komputer*, vol. 21, no. 3, pp. 721–734, jul 2022.

[25] X.-S. Yang and A. Slowik, "Firefly Algorithm," *Swarm Intelligence Algorithms*, pp. 163–174, aug 2020.

[26] X. S. Yang and X. S. He, "Why The Firefly Algorithm Works?" *Studies in Computational Intelligence*, vol. 744, pp. 245–259, 2018.

[27] S. L. Tilahun, J. M. T. Ngnotchouye, and N. N. Hamadneh, "Continuous Versions of Firefly Algorithm: A Review," *Artificial Intelligence Review 2017 51:3*, vol. 51, no. 3, pp. 445–492, jul 2017.

[28] A. Mizwar, A. Rahim, A. Sunyoto, and M. R. Arief, "Stroke Prediction using Machine Learning Method with Extreme Gradient Boosting Algorithm," *MATRIK : Jurnal Manajemen, Teknik Informatika dan Rekayasa Komputer*, vol. 21, no. 3, pp. 595–606, jul 2022.