Retraction

Retracted: An IoT and Machine Learning-Based Model to Monitor Perishable Food towards Improving Food Safety and Quality

Journal of Food Quality

This article has been retracted by Hindawi following an investigation undertaken by the publisher [1]. This investigation has uncovered evidence of one or more of the following indicators of systematic manipulation of the publication process:

(1) Discrepancies in scope

(2) Discrepancies in the description of the research reported

(3) Discrepancies between the availability of data and the research described

(4) Inappropriate citations

(5) Incoherent, meaningless and/or irrelevant content included in the article

(6) Peer-review manipulation

The presence of these indicators undermines our confidence in the integrity of the article's content and we cannot, therefore, vouch for its reliability. Please note that this notice is intended solely to alert readers that the content of this article is unreliable. We have not investigated whether authors were aware of or involved in the systematic manipulation of the publication process.

Wiley and Hindawi regrets that the usual quality checks did not identify these issues before publication and have since put additional measures in place to safeguard research integrity.

We wish to credit our own Research Integrity and Research Publishing teams and anonymous and named external researchers and research integrity experts for contributing to this investigation.

The corresponding author, as the representative of all authors, has been given the opportunity to register their agreement or disagreement to this retraction. We have kept a record of any response received.

References

[1] A. Batool, S. Ganguli, H. A. Almashaqbeh et al., "An IoT and Machine Learning-Based Model to Monitor Perishable Food towards Improving Food Safety and Quality," *Journal of Food Quality*, vol. 2022, Article ID 6302331, 6 pages, 2022.

WILEY | Hindawi

Research Article

An IoT and Machine Learning-Based Model to Monitor Perishable Food towards Improving Food Safety and Quality

Amina Batool [iD],[1] **Souvik Ganguli** [iD],[2] **Hashem Ali Almashaqbeh** [iD],[3] **Muhammad Shafiq** [iD],[4] **A. L. Vallikannu** [iD],[5] **K. Sakthidasan Sankaran** [iD],[5] **Samrat Ray** [iD],[6] and **F. Sammy** [iD][7]

[1]*School of Automation, Beijing Institute of Technology, Beijing, China*
[2]*Department of Electrical and Instrumentation Engineering, Thapar Institute of Engineering and Technology, Patiala 147004, India*
[3]*Okan University, Istanbul, Turkey*
[4]*School of Artificial Intelligence, Neijiang Normal University, Neijiang, Sichuan, China*
[5]*Department of ECE, Hindustan Institute of Technology and Science, Chennai, India*
[6]*Sunstone Eduversity, Gurugram, India*
[7]*Department of Information Technology, Dambi Dollo University, Dembi Dolo, Welega, Ethiopia*

Correspondence should be addressed to F. Sammy; sammy@dadu.edu.et

Increased quantities of the same sort of item are not nearly as critical to client happiness as a high-quality product. The requirements and expectations of the consumer have an impact on the overall quality of a product or service. The term "quality" may also be defined as the sum total of all the features that contribute to the production of goods and services that are satisfactory to the consumer. Certain imported commodities have lately seen an improvement in quality thanks to efforts by importing nations. Additionally, it safeguards food imported from other nations by confirming that it is safe for human consumption before it is released. This article describes a technique for monitoring perishable goods that is based on the Internet of Things and machine learning. Pictures are recorded using high-resolution cameras in this suggested architecture, and then these images are sent to a cloud server using Internet of Things devices. When uploaded to a cloud server, these photos are segmented using the K-means clustering method. Then, using the principal component analysis technique, features are extracted from the photos, and the images are categorized using machine learning models that have been trained. This proposed model makes use of the Internet of Things, image processing, and machine learning to monitor perishable food.

1. Introduction

Things of the same type are not as important to the customer's satisfaction as high-quality products [1]. The quality of a product is influenced by the needs and expectations of the customer. Another way to define quality is to think of it as the sum total of all the characteristics that go into generating things that the customer is satisfied with [2]. Importing countries have recently improved the quality of certain imported items. Additionally, it protects food imported from other countries by ensuring that it is safe for consumption. Agricultural products' external quality is the primary indicator of their immediate sensory quality. Product's quality is generally judged by its appearance, including its color, texture, size, shape, and imperfections [3]. Food manufacturing company administrators take into consideration the product's significance, the social setting, and the difficulties experienced by farmers in accomplishing their agricultural tasks [4].

The discovery of faults prior to the sale or export of items is one of the most essential parts of quality assurance [5]. When it comes to determining the quality of mangos, human operators have long relied on their eyes. This is not the case anymore. They take a long time, are tedious, and are

inconsistent [6]. In agricultural engineering studies [7, 8], this method has been applied.

Technological advances and the incorporation of these advances into a sustainable farming system have assisted agricultural research. Technology has automated every stage of the agricultural process, making it a long-term economic model. An intermediary who buys cheap from farmers and then resells it at a profit is removed from the equation. Computational intelligence approaches have recently been used to handle agricultural system decision modelling challenges. Technologies like artificial intelligence and the Internet of Things are already making a difference in agriculture. Machine learning and the Internet of Things may be used to keep tabs on perishable commodities, which is an essential application. Food safety and quality are improved as a result [9].

The Internet of Things (IoT) is a hot issue right now. Every element of our everyday life is being influenced by the Internet of Things. The Internet of Things has numerous applications and is making people's lives better (IoT). IoT technology enables smart cities, smart homes, smart healthcare, smart agriculture, smart farming, smart logistics, and smart production [10, 11].

Agriculture is essential to the country's economy since it feeds its inhabitants. This means that all of the country's major corporations are linked and interrelated. A country is considered rich if its economy and society are mainly dependent on agriculture. Agriculture employs the vast majority of people in the vast majority of countries. The upkeep of plants and animals on large farms may entail the hiring of additional workers. The bulk of these huge farms has close processing facilities where they generate and process their agricultural goods.

Automated machine learning techniques can be used to analyze the complex interaction between agricultural systems' inputs and outputs by employing nonlinear and time-dependent analytical methodologies as well as different unknown components [12, 13].

This article presents an Internet of Things and machine learning-based model to monitor perishable foods. In this proposed model, first images are captured using high-resolution cameras, and then these images are forwarded to a cloud server through IoT devices. On the cloud server, these images are segmented using the K-means clustering algorithm. Then, features are extracted by the principal component analysis algorithm. After that, the images are classified using trained machine learning predictors. This proposed model makes use of the Internet of Things, image processing, and machine learning to monitor perishable food.

2. Literature Survey

The term "machine learning" (ML) refers to an emerging subject in which algorithms can learn and develop on their own without the need for explicit coding. There are two main approaches to machine learning: unsupervised and supervised. Classification and regression are the two main types of inference. The training data for classification must be

labeled, while the training data for regression must be unlabeled (clustering) [14].

Identifying the traits that make something unique, or "target attributes," is the first step in determining how to classify it. Decision trees and support vector machines, as well as Bayesian classifiers, are commonplace in many fields [15]. They are used in machine learning as well as in neural networks and classifiers. One at a time, they will be covered in detail. Algorithms for learning form the basis of each approach.

Decision trees may be useful in situations when classification is a problem. A decision tree sorts events into categories based on the values of the attributes they contain. Each branch in the decision tree indicates a possible value for a node in the classification. Feature values are used to rank instances in the hierarchy, whereas root nodes are used to classify them.

When it comes to input attributes, it may not be possible to accurately describe them. While input attribute set values match those in the training dataset, class variables are nondeterministic. This is due to a lack of consideration for noise and confusion in the data processing. There are many ways to tell if you have cardiovascular disease in your daily life.

Even those who follow a heart-healthy diet and engage in a regular physical activity run the risk of getting heart disease if they smoke, drink alcohol, or have a genetic predisposition. It is not always possible to get reliable information from the well-known symptoms of cardiac disease. In these cases, the Bayesian classifier uses attributes to determine the class label [16].

In order to create animal brains, a biological neural network is used. ANN is a connectionist system due to the fact that it is made up of connected nodes and directed links. Communication between nodes is carried out by means of connections, each of which has its own weight. Prior to being transferred to the next node, signals must be processed at each node's end.

Nonlinear functions are applied to the output of each neuron in most artificial neural network implementations. As the process of learning continues, the strength of the signal changes based on the weight and connections of the artificial neurons [17].

There are two ways to build a learning model in ML classification. These models begin learning as soon as they have access to the training data. A separate model can recognize a test case if the properties of the test case exactly match any of the training instances. Students like these are labeled "lazy" [18].

When it comes to categorization, each sample is considered an individual data point in the context of all other samples. A clear separation between the test case and the rest of the data has been formed. A data point's neighbors include every other data point in its immediate proximity.

The data point is classified in accordance with the classifications of its immediate neighbors. Data points having a large number of classified neighbors are given the class label that is most frequently used by other data points. The precise value of K-nearest neighbors must be

determined in order to complete the task. Using a modest value for k can lead to incorrect classification due to noise in the training data. It is possible that data points that are outside of the test attribute's immediate area can be found in the number of closest neighbors that are greater than k. An error in classification could be the outcome.

Initially, random forest is a machine learning technique that uses many decision trees constructed from random vector inputs to produce a forest of judgments. Classification and regression problems can be solved using this method. The random forest's forecasts will be more accurate if the number of trees in the random forest is increased. It is critical to realize that a forest and a decision tree are not interchangeable.

For the root node to be found, randomly separating feature nodes is the fundamental difference between decision trees and random forests. Because of its advantages, random forest classification is frequently employed. One of its numerous advantages is that it can be utilized for both classification and regression analysis. It also avoids the problem of overfitting if there are many trees available. Random forest classification can also be used to describe categorical data and deal with missing data.

Retail, healthcare, and financial businesses can benefit from random forest classifiers that can be employed in a range of industries. When it comes to banking, random classifiers are frequently employed to distinguish between legitimate consumers and potential scammers. Random forest is used in medicine to find the optimum pharmaceutical combinations and to diagnose illness based on the medical history of a patient. Random forest is the swankier way of tracking a stock's performance and calculating its profit or loss. You may use random forest to predict what products customers will buy based on their preferences in an online store.

The finest tool for categorizing is the support vector machine (SVM). As a result, excitement is high among the employees. In vector space, it is feasible to distinguish between instances of different categories using an SVM model. When they are mapped into the specified vector space, new samples are given a label based on which side of the gap they fall into [19]. Nonlinear classification can be accomplished using an SVM by utilizing the kernel technique.

3. Methodology

This section presents an Internet of Things and machine learning-based model to monitor perishable foods. In this proposed model, first, images are captured using high-resolution cameras, and then these images are forwarded to a cloud server through IoT devices. On the cloud server, these images are segmented using the K-means clustering algorithm. Then, features are extracted by the principal component analysis algorithm. After that, the images are classified using trained machine learning predictors. The block diagram of the model is shown in Figure 1.

Image segmentation is the process of identifying and grouping together parts of a picture based on shared characteristics. Two kinds of segmentation approaches are
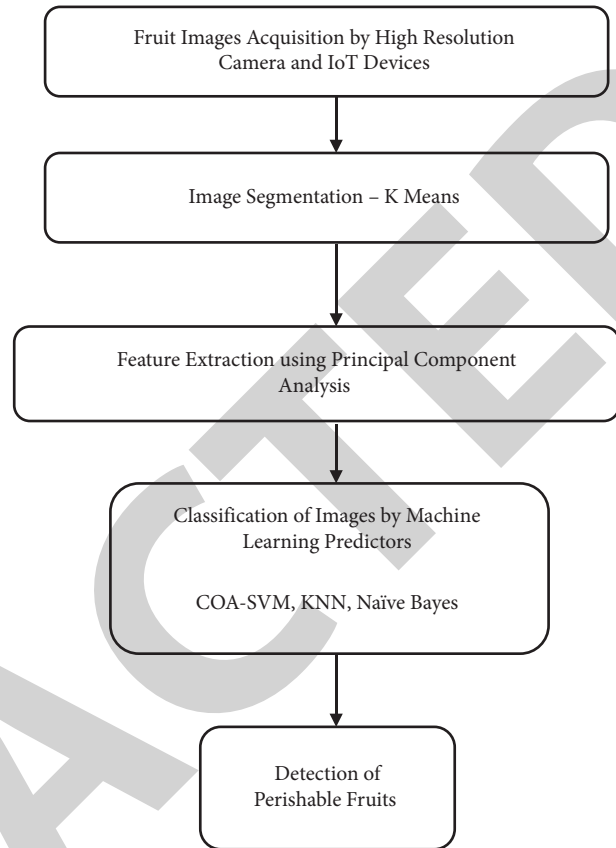

Figure 1: Internet of things, image processing, and machine learning-based model for monitoring perishable food items.

region-based and edge-based. Region-based segmentation may be used to classify anatomical or functional features into groups based on patterns in intensity values around a cluster of neighboring pixels.

Based on the unique textures and patterns seen in each area, this research uses regional segmentation to split the ROI. The local mean is utilized as a clustering pattern to assign each unique observation to one of the K-means clusters. This program looks for data clusters based on the total number of groups specified by the variable k. On-the-fly data point proximity is determined by calculating squared Euclidean distances. Data points are allocated to one of the k categories according to the input characteristics. In order to group data points, a similarity measure is utilized [20].

Using a white or black backdrop, the number K shows that clustering is effective. But if you take a picture of an object with a live backdrop and precisely segment the ROI, it fails to work. K shows that clustering is unable to recover the ROI area from a picture with a moving backdrop.

Classification data is classified using KNN's basic classifier, which utilizes the tuple that is closest to the unknown [21]. Using likeness measurements such as Euclidean distance, Manhattan distance, Pearson correlation, and others, it predicts new cases. For categorical values, the hammering distance must be determined and standardized using the numerical variables 1 and 0. For the majority of datasets, a K value of 3–10 is ideal.
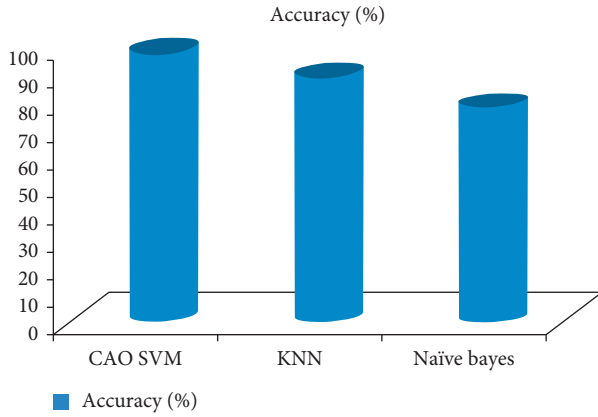
Accuracy (%)

Figure 2: Accuracy of classifiers for perishable food identification.

Sensitivity (%)

Figure 3: Sensitivity of classifiers for perishable food identification.

Specificity (%)

Figure 4: Specificity of classifiers for perishable food identification.

COA-SVM (cuckoo optimization algorithm) support vector machine is the best option for nonprobabilistic binary linear classification. One or more target classes can be identified using this method. A single point is used to represent each piece of data. The clear differences among the various groupings cause them to expand in width. Those new instances are remapped to different target classes based on where they fall in the gap. Nonlinear classification is possible when the input datasets are unlabeled. The support vector machine employs an unsupervised learning technique to classify the data because the instances cannot be allocated to target classes. More instances are added when the clusters based on functions have been formed. A nonlinear support vector machine-based recommendation system has been demonstrated to be effective. Nonlinear support vector machine approaches are the most frequently utilized way for dealing with unlabeled data [22].

Naive Bayes is a subcategory of Bayes' theorem. It is a monitored, classified system. A simple approximation of iterative parameters is not required when building a Naive Bayesian model [23]. The posterior likelihood of $P(c|x)$ can be calculated using the Bayes theorem, which takes into account the prior likelihood of $P(c)$, $P(x)$, and $P(x|c)$. The posterior probability can be calculated by first generating a frequency table for the desired property. A Naive Bayesian formula is then used to calculate each group's posterior likelihood based on the frequency tables. The category with the highest posterior likelihood is the result of the calculation.

4. Results and Discussion

First, images are captured using high-resolution cameras, and then these images are forwarded to a cloud server through IoT devices. On the cloud server, these images are segmented using the K-means clustering algorithm. Then, features are extracted by the principal component analysis algorithm, after that the images are classified using trained machine learning predictors.

For research purposes, 500 images of bananas and oranges have been gathered. There are pictures of both healthy and unhealthy fruits. Bananas are the subject of 300 pictures,
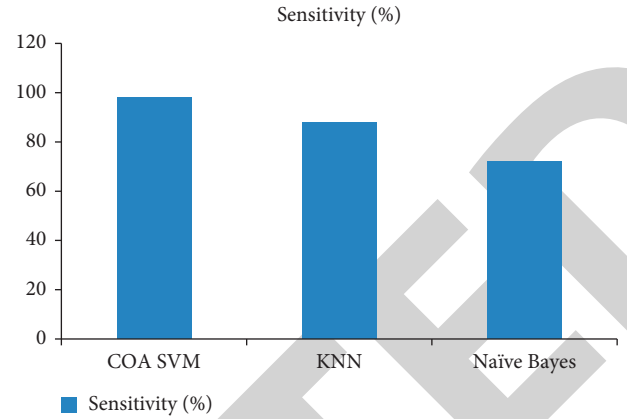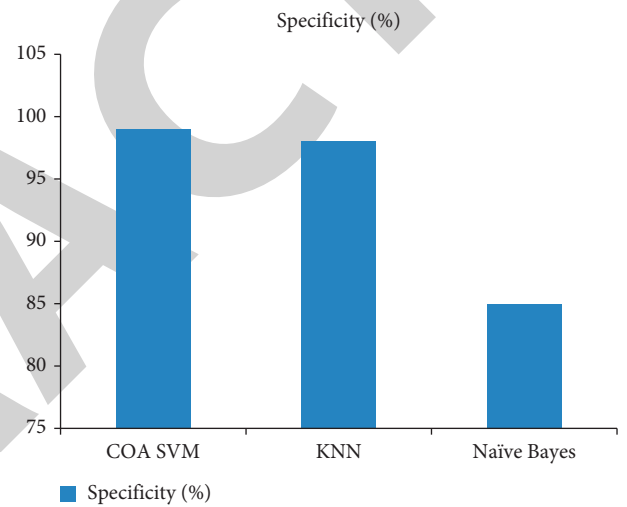
including 158 of them while they are still fresh and 142 of them after they are rotting. Oranges are the subject of 200 photographs. There are 120 photographs of oranges in their prime and 80 images of decaying mangoes in their worst state. The K-means clustering technique is used for image segmentation. The fruit photos are then classified using COA SVM, KNN, and Naive Bayes machine learning algorithms. These algorithms determine if a piece of fruit is excellent or bad. Accuracy, sensitivity, and specificity are shown in Figures 2–4.

Parameters used in the experimental study are discussed as follows:

$$\text{accuracy} = \frac{TP + TN}{TP + TN + FP + FN},$$

$$\text{sensitivity} = \frac{TP}{TP + FN}, \quad (1)$$

$$\text{specificity} = \frac{TN}{TN + FP},$$

where TP = true positive; TN = true negative; FP = false positive; and FN = false negative.

5. Conclusion

Monitoring of perishable food items is an important research area as far as smart agriculture is concerned. It takes the Internet of Things, image processing, and machine learning to accurately monitor perishable food items. Increased quantities of the same type of item are not nearly as critical to client happiness as a high-quality product. The consumer's needs and expectations influence the overall quality of a product or service. The term "quality" can alternatively be defined as the sum of all the characteristics that contribute to the production of consumer-satisfactory goods and services. Certain imported commodities have recently improved in quality as a result of efforts by importing countries. Furthermore, it protects food imported from other countries by ensuring that it is safe for human consumption before it is released. This article covers a perishable commodity monitoring system based on the Internet of Things and machine learning. In this proposed architecture, images are captured with high-resolution cameras and relayed to a cloud server via Internet of Things devices. When these photographs are uploaded to a cloud server, they are segmented using the K-means clustering algorithm. Then, using the principal component analysis technique, features are retrieved from the photographs, and the images are classified using trained machine learning models. The COA-SVM algorithm has better accuracy in the monitoring and detection of perishable food items.

Data Availability

The data shall be made available on request.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

References

[1] M. Tripathi and D. D. Maktedar, "A role of computer vision in fruits and vegetables among various horticulture products of agriculture fields: a Survey," *Information Processing in Agriculture*, vol. 7, no. 2, pp. 183–203, 2019.

[2] T. Laikko-Roto, S. Makela, J. Lunden, J. Heikkilä, and M. Nevas, "Consistency in inspection processes of food control officials and efficacy of official controls in restaurants in Finland," *Food Control*, vol. 57, pp. 341–350, 2015.

[3] H. Calvo, M. A. Moreno-Armendáriz, and S. Godoy-Calderón, "A practical framework for automatic food products classification using computer vision and inductive characterization," *Neurocomputing*, vol. 175, pp. 911–923, 2016.

[4] J. Zhou, Z. Yan, and K. Li, "Understanding farmer cooperatives' self-inspection behaviour to guarantee agri-product safety in China," *Food Control*, vol. 59, pp. 320–327, 2016.

[5] D. P. Cavallo, M. Cefola, B. Pace, A. F. Logrieco, and G. Attolico, "Non-destructive automatic quality evaluation of fresh-cut iceberg lettuce through packaging material," *Journal of Food Engineering*, vol. 223, pp. 46–52, 2018.

[6] P. Sharma, Y. Paul Singh Berwal, and W. Ghai, "Performance analysis of deep learning CNN models for disease detection in plants using image segmentation," *Information Processing in Agriculture*, vol. 7, no. 4, pp. 566–574, 2019.

[7] M. Ji, L. Zhang, and Q. Wu, "Automatic grape leaf diseases identification via united model based on multiple convolutional neural networks," *Information Processing in Agriculture*, vol. 7, no. 3, pp. 418–426, 2019.

[8] X. Guo, X. Zhao, Y. Liu, and D. Li, "Underwater sea cucumber identification via deep residual networks," *Information Processing in Agriculture*, vol. 6, no. 3, pp. 307–315, 2019.

[9] L. F. Santos Pereira, S. Barbon Jr., N. A. Valous, and D. F. Barbin, "Predicting the ripening of papaya fruit with digital imaging and random forests," *Computers and Electronics in Agriculture*, vol. 145, pp. 76–82, 2018.

[10] A. Raghuvanshi, U. K. Singh, G. S. Sajja et al., "Intrusion detection using machine learning for risk mitigation in IoT-enabled smart irrigation in smart farming," *Journal of Food Quality*, vol. 2022, Article ID 3955514, 8 pages, 2022.

[11] V. Hemamalini, S. Rajarajeswari, S. Nachiyappan et al., "Food quality inspection and grading using efficient image segmentation and machine learning-based system," *Journal of Food Quality*, vol. 2022, Article ID 5262294, 6 pages, 2022.

[12] A. Raghuvanshi, U. Singh, and C. Joshi, "A review of various security and privacy innovations for IoT applications in healthcare," *Advanced Healthcare Systems*, Wiley Online Library, Hoboken, NJ, USA, pp. 43–58, 2022.

[13] V. D. P. Jasti, A. S. Zamani, K. Arumugam et al., "Computational technique based on machine learning and image processing for medical image Analysis of breast cancer diagnosis," *Security and Communication Networks*, vol. 2022, pp. 1–7, Article ID 1918379, 2022.

[14] K. G. Liakos, P. Busato, D. Moshou, S. Pearson, and D. Bochtis, "Machine learning in agriculture: a review," *Sensors*, vol. 18, no. 8, p. 2674, 2018.

[15] B. Gulzar and A. Gupta, "DAM: a theoretical framework for SensorSecurity in IoT applications," *International Journal of Next-Generation Computing*, vol. 12, no. 3, 2021.

[16] K. Phasinam, T. Kassanuk, and M. Shabaz, "Applicability of internet of things in smart farming," *Journal of Food Quality*, vol. 2022, Article ID 7692922, 7 pages, 2022.

[17] A. Kaya, A. S. Keceli, C. Catal, H. Y. Yalic, H. Temucin, and B. Tekinerdogan, "Analysis of transfer learning for deep neural network based plant classification models," *Computers and Electronics in Agriculture*, vol. 158, pp. 20–29, 2019.

[18] J. Wang, H. Yue, and Z. Zhou, "An improved traceability system for food quality assurance and evaluation based on fuzzy classification and neural network," *Food Control*, vol. 79, pp. 363–370, 2017.

[19] P. Ratta, A. Kaur, S. Sharma, M. Shabaz, and G. Dhiman, "Application of Blockchain and Internet of Things in healthcare and medical sector: applications, challenges, and future perspectives," *Journal of Food Quality*, vol. 2021, Article ID 7608296, 20 pages, 2021.

[20] M. N. Reza, I. S. Na, S. W. Baek, and K. H. Lee, "Rice yield estimation based on K-means clustering with graph-cut segmentation using low-altitude UAV images," *Biosystems Engineering*, vol. 177, pp. 109–121, 2019.

[21] Z. Deng, X. Zhu, D. Cheng, M. Zong, and S. Zhang, "Efficient k NN classification algorithm for big data," *Neurocomputing*, vol. 195, pp. 143–148, 2016.

[22] E. Suryawati, H. F. Pardede, V. Zilvan et al., "Unsupervised feature learning-based encoder and adversarial networks," *Journal of Big Data*, vol. 8, no. 1, p. 118, 2021.

[23] D. Mondal, D. K. Kole, and K. Roy, "Gradation of yellow mosaic virus disease of okra and bitter gourd based on entropy based binning and Naive Bayes classifier after identification of leaves," *Computers and Electronics in Agriculture*, vol. 142, pp. 485–493, 2017.