ACCSCIENCE PUBLISHING

International Journal of AI for Materials and Design

ORIGINAL RESEARCH ARTICLE

Automated fruit sorting system integrating image processing and support vector machine techniques

Babatunde Olayinka Oyefeso[1], **Oluwaseun Emmanuel Oyewande**[1], **and John Audu**[2*]

[1]Department of Agricultural and Environmental Engineering, Faculty of Technology, University of Ibadan, Ibadan, Oyo State, Nigeria

[2]Department of Agricultural and Bio-systems Engineering, College of Engineering, Joseph Sarwuan Tarka, University Makurdi, Benue State, Nigeria

Abstract

Traditional fruit grading methods are mostly time-consuming and subjective, thereby limiting efficiency in the agricultural sector. To address these problems, this paper presents the design and implementation of an automated fruit sorting system for classifying certain fruits, namely oranges, tomatoes, and mangoes, using image processing and support vector machine (SVM) techniques. An ESP32 camera was used to capture images of the fruits, which were later passed through algorithms in Python. Extracted features were then fed into a SVM model for the classification process of fruits. The model demonstrated excellent performance, achieving an accuracy of 100%, a precision of 96%, a recall of 92%, and an F1 score of 89%. The results indicated that incorporating multiple features significantly increases the accuracy of the classification. Moreover, the performance was optimized by selecting an appropriate regularization parameter during the training of the model and the use of polynomial kernel functions. Finally, the whole automated system was assembled to physically sort the classified fruits into different containers. This research highlights the potential of integrating image processing and machine learning technologies to revolutionize fruit classification processes, thereby improving both efficiency and quality control in agriculture.

Keywords: Image processing; Fruit classification; Support vector machine; Automated sorting; Feature extraction

***Corresponding author:**
John Audu
(audu.john@uam.edu.ng)

Citation: Oyefeso BO, Oyewande OE, Audu J. Automated fruit sorting system integrating image processing and support vector machine techniques. *Int J AI Mater Design*. 2025;2(2):79-90. doi: 10.36922/IJAMD025150011

Received: April 8, 2025

1st revised: May 9, 2025

2nd revised: May 14, 2025

3rd revised: May 18, 2025

Accepted: May 22, 2025

Published online: June 20, 2025

1. Introduction

Fruit image classification techniques are continually being developed due to their vital roles in agriculture and food analysis within the food industry. In the agricultural and food industries, imaging technology streamlines operations by enhancing quality control and optimizing the process. Mango production in Nigeria, the ninth most produced fruit globally, is hindered by several challenges resulting from outdated technology.[1-9]

It was reported in previous studies that the existing methods for manual classification of fruits are somewhat inefficient, ineffective, slow, and prone to bias.[10-13] Developments in image analysis have provided an efficient, reliable, and accurate system of fruit categorization without causing much harm to the fruits.[14-17] In contrast, manual sorting of fruits by professional personnel undergoes physical handling, thus potentially damaging the fruits and affecting their value.[18-21] This research focuses on creating an image processing system for the classification of fruits through machine learning to enhance precision and productivity in the agriculture business as well as the food chain. Similar approaches utilizing image processing and machine learning to detect mangoes, tomatoes, and oranges were also performed by other researchers. Image processing methods and machine learning algorithms have been widely used to classify mangoes, tomatoes, and oranges, achieving classification accuracies ranging from 80% to 100% across these fruit types.[22-37] Research on fruit classification using image processing and machine learning is constrained due to its focus on single fruit type, small datasets, inconsistent image acquisition methods, and the lack of deep learning approaches. Future investigations should prioritize standardized data collection and the application of deep learning algorithms.

The proposed fruit classification system leverages image processing to achieve maximum accuracy and minimal time expenditure. It is trained on a large database comprising mangoes, tomatoes, and oranges. Such a classification system has significant potential to enhance the quality of both the agricultural sector and the food industry by monitoring product quality, minimizing wastage, and adding value to the product. Furthermore, it can be integrated into other automated systems and mobile applications.

The novel integration of real-time image processing with real-time mechanical fruit sorting, powered by artificial intelligence machine learning optimization techniques using Python programming, represents the novelty of this study.

2. Materials and methods

2.1. Image processing system to process fruit images

Four steps were included in developing the image processing system to classify fruits (mango, oranges, and tomatoes): Image acquisition, pre-processing, feature extraction, and feature selection. Each step played a key role in ensuring the accuracy and effectiveness of the classification system.

2.1.1. Image acquisition

This step was crucial to capture high-quality images of the fruits, which we needed for later analysis. Crucial factors to consider during image acquisition included:

(i) Lighting conditions: Good lighting was vital in showing the visual traits of the fruits. An ESP32 camera provided the best lighting and clarity.
(ii) Camera specifications: The camera's resolution and color accuracy had a big impact on image quality. We fine-tuned these factors to ensure the fruits' features were displayed accurately.

2.1.2. Pre-processing

Pre-processing prepared the captured images for feature extraction by enhancing important properties and reducing noise. The methods used include:

(i) Resizing: The system resized all images to uniform dimensions while preserving their aspect ratio. This step was essential to ensure consistency across datasets and to enhance computational efficiency.
(ii) Histogram equalization: This technique enhanced image contrast by spreading out pixel intensity values. It standardized the appearance of images, making key features more distinguishable during the later processing stage.
(iii) Thresholding: Thresholding splits pixels into object and background areas based on a predefined value. It effectively isolated the fruit from the background and reduced noise, thereby improving feature extraction.

2.1.3. Feature extraction

At this point, the system extracted various features from the pre-processed images. These features included color, shape, texture, and size, which were essential in differentiating fruits.

2.1.4. Feature selection

After the features were extracted, the system evaluated the significance of each feature. The features that varied widely between fruit types were retained, while those with insignificant variation were excluded. This step enhanced the sorting system by focusing on the most distinctive features, which resulted in better accuracy and reduced computational work.

2.2. Support vector machine (SVM) model to classify fruit images

The SVM model helped classify fruits by their appearances. This part explained the key steps in building the SVM model, which included standardizing the features, training the model, and applying different kernel functions.

2.2.1. Standardizing the features

Standardizing the features was a key step before starting the modeling process. It ensured that all the features selected for model creation contributed to the fruit-sorting process. In this study, normalization was used to scale all feature values into the same range. This prevents any single feature from having too much influence over the model's performance.

2.2.2. Training the model

To train the model, the system splits the entire dataset into two parts: One for training and one for testing. It used 70% of the data in training the model and kept the other 30% to test the training efficiency. Then, the SVM model was trained using the selected features. The study also used kernel functions to transform the feature space, which helped the model capture non-linear relationships.

2.2.3. Kernel functions

Kernel functions play a crucial role in SVMs as they allow data to be transformed into higher dimensional spaces, where it can be linearly separable. The following kernel functions were implemented:

(i) Linear kernel

If the relationship between features was linearly separable, then the linear kernel was utilized. It is defined in Equation I[38]:

$$K(x, y) = x^T y \quad \text{(I)}$$

Where x and y are input feature vectors and $x^T y$ represents the dot product of the transpose of x and y.

(ii) Polynomial kernel

The polynomial kernel enabled the SVM to handle non-linear relationships between features. It mapped the input data into a higher-dimensional feature space using polynomial functions, allowing SVM to learn more complex decision boundaries. The polynomial kernel is defined in Equation II[39]:

$$K(x, y) = (\gamma \times (x^T y) + r)^d \quad \text{(II)}$$

Where γ is the scaling factor, r is a constant, and d is the degree of polynomial.

(iii) Radial basis function (RBF) kernel

The RBF kernel was used to capture non-linear relationships in the data. It assigned lower weights to more distant points and higher weights to closer points, allowing the SVM to identify local patterns effectively. The RBF is defined in Equation III[40]:

$$K\left(x, y\right) = e^{\left(-\gamma * \|x - y\|^2\right)} \quad \text{(III)}$$

Where γ is a constant, e is the base of the natural logarithm, and is the Euclidean distance between x and y.

This is an established method of SVM model development, which ensures accuracy and efficacy in fruit classification using different kernel functions.

2.3. Comparison of kernel functions

The different kernel functions were benchmarked against each other based on their performance in efficiently classifying the fruits. This evaluation utilized several performance metrics, including accuracy, precision, recall, and F1 score, all derived from confusion matrix analyses.

2.3.1. Confusion matrix method

One of the most effective approaches for evaluating the performance of the trained classification model is the confusion matrix. In this study, the confusion matrix provided the number of correct and incorrect classifications among the three fruit classes: tomato, mango, and orange. A standard confusion matrix table is presented in Figure 1.

(i) Model accuracy

Model accuracy measured the proportion of correct predictions made by the classifier. It was calculated based on a ratio of total true predictions to the total prediction value, which provided a simple sense of the model's performance as indicated in Equation IV.[41]

$$\text{Accuracy} = \frac{TP + TN}{TP + TN + FP + FN} \quad \text{(IV)}$$

Where TP is true positive, TN is true negative, FP is false positive, and FN is false negative.

(ii) Model precision

Model precision measured the accuracy of positive predictions. In this study, it referred to the proportion of predicted positive instances that were actually positive. This provided insight into the reliability of the model in making positive classification, as shown in Equation V.[42]

$$\text{Precision} = \frac{TP}{TP + FP} \quad \text{(V)}$$

(iii) Model recall score

Model recall, also known as sensitivity, was used to evaluate the model's ability to correctly identify

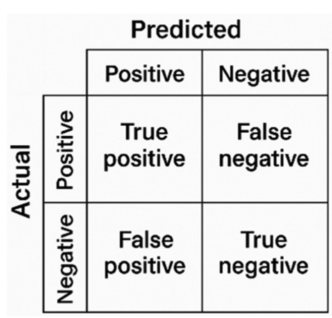

Figure 1. A typical confusion matrix table used to evaluate classification performance

positive instances. It indicated how many actual positive observations were correctly identified among all the true positives, thereby showing the model's effectiveness in capturing relevant cases, as shown in Equation VI.[43]

$$\text{Recall} = \frac{\text{TP}}{\text{TP} + \text{FN}} \tag{VI}$$

(iv) Model F1 score

The F1 score is the harmonic mean of precision and recall, providing a balanced metric that considers both aspects. It was particularly useful for evaluating model performance on imbalanced datasets, as it offered a better overview than individual metrics. The F1 score was calculated using Equation VII.[44]

$$\text{F1} = \frac{2}{\dfrac{1}{\text{Precision}} \quad \dfrac{1}{\text{Recall}}} = \frac{2 \times (\text{Precision} \times \ \text{Recall})}{(\text{Precision} \quad \text{Recall})} \tag{VII}$$

In such a structured approach, the performance metrics of different kernels were compared directly to identify the most effective kernel for fruit classification.

2.4. Model Optimization

It was essential to optimize a model for its performance and accuracy toward fruit classification. The techniques applied in model development were also used to tune the model, with a focus on the regularization parameter (C).

The C value controlled the trade-off between minimizing training error and testing error.

(i) Impact of the value of C: A lower C value applied stronger regularization by shrinking the coefficients less aggressively, allowing a larger margin of error, which may result in higher misclassification rates. On the other hand, a higher C value reduced the error margin, thereby lowering misclassification rates.
(ii) Optimal C by cross-validation: Cross-validation was utilized to identify the optimal value of C. This involved testing various C values using the model to determine which value achieved the best classification performance while avoiding overfitting.

2.5. Automated image sorting mechanism

The automated sorting mechanism played a crucial role in the fruit classification system. It physically sorted the classified fruits into their respective containers based on the output classifications identified by the SVM model. The system was designed to ensure efficient and accurate sorting of the fruits.

The process began by placing fruits on a slanting plane that acted as a conveyor. As the fruits rolled down the inclined plane, images were captured by a camera for classification according to their classes, as implemented by the SVM model. The camera used was an ESP32-cam (Espressif Systems, China), controlled by the Arduino Uno R3 (Microchip, United States), with the entire setup powered by a 9 V battery. The captured images were transmitted to the processing unit for classification using the trained SVM model.

Based on the classification results, the Arduino board (Arduino, Italy) sent commands to three servo motors (MG996r, TowerPro, China). These motors controlled a mechanical arm responsible for guiding the fruits into different containers. Once the fruits reached the end of the inclined plane, the servo motors activated the mechanical arms to direct the fruits into their respective containers. Collection containers were placed at the base of the inclined plane to collect the sorted fruits.

In this project, an efficient and accurate automated sorting mechanism was developed. Figure 2A shows the conceptual drawing of the mechanism, Figure 2B shows the side view, and Figure 2C shows the front view of the mechanism, clearly depicting the system components and their arrangement. Figure 3 displays the complete experimental setup for the system, while Figure 4 illustrates the overall framework for the automated fruit sorting process.

This setup incorporated an automated sorting mechanism integrated into the Arduino board, which controlled the classification results from the SVM model and ensured efficient and accurate sorting of fruits into their respective containers. The system was specifically designed to handle the fruits with the least damage, thereby maintaining the quality of the fruits during the sorting process. All program codes that automated this system were displayed in Programs S1–S6 (in Supplementary File).

3. Results and discussion

3.1. Experimental results for model detection of fruit image features

Table 1 shows the results of the fruit classification system. The results highlight key differences in the efficacy of various features used in the image processing approach. This was evident in the performance metrics, including accuracy, precision, recall, and F1 score, all of which demonstrated that feature selection significantly impacts the overall results of classification.

One of the important metrics is accuracy, which represents the proportion of correct predictions made by the classifier. The accuracy obtained in this study was 90% with combined features. The result demonstrates

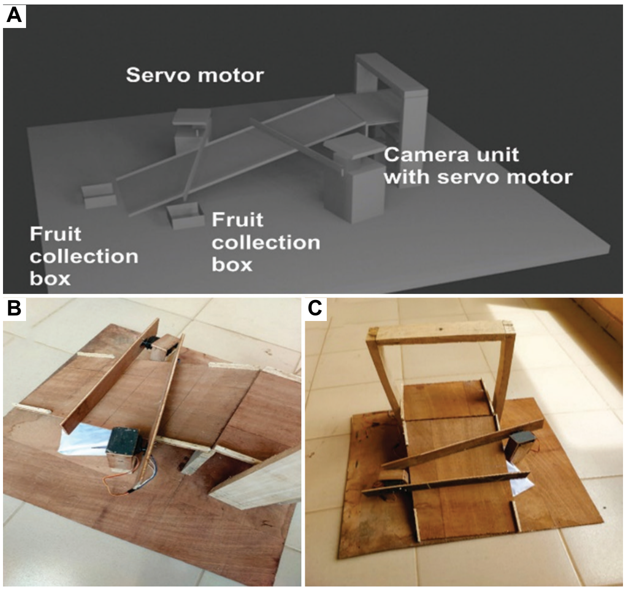

Figure 2. Automated fruit sorting system developed in this study. (A) The conceptual drawing of the system. (B) The fabricated side view of the system. (C) The fabricated front view of the system. Image produced by the authors.

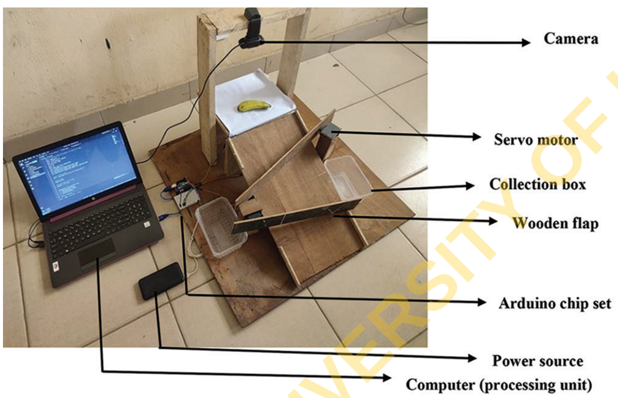

Figure 3. Complete experimental setup of the system. Image produced by the authors.

that combined features provide more comprehensive information about the fruit characteristics, leading to improved prediction outcomes.

The results were also favorable in terms of precision, which measures the accuracy of positive predictions. In this study, the model achieved a precision score of 88%, indicating high reliability in its positive classifications. The accuracy is particularly important in agricultural applications, where precise identification of fruit types can significantly influence sorting efficiency and marketing decisions.

Recall (or sensitivity) measures the effectiveness of a model in identifying relevant instances. In this study, using combined features, the model achieved a maximum recall score of 85%, efficiently capturing most of the true positive cases. This measure is important in ensuring that no fruit

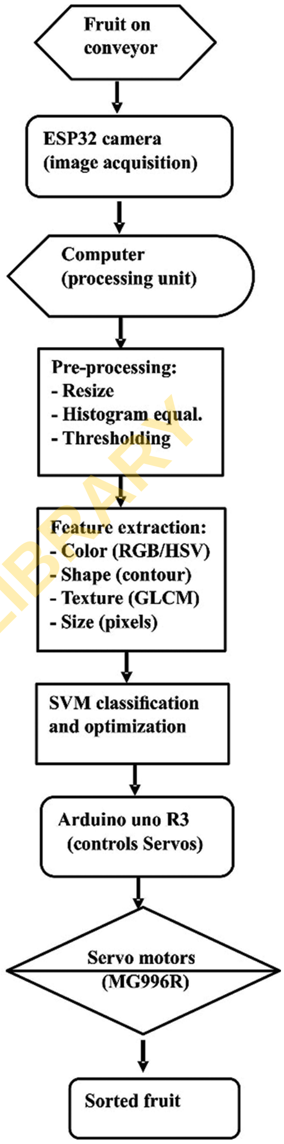

Figure 4. An overall framework of the method for the automated fruit sorting system.
Abbreviation: SVM: Support vector machine.

Table 1. Summary of the experimental results for fruit classification models based on image features

Feature	Accuracy (%)	Precision (%)	F1 score (%)
Color feature	85.0	82.0	81.0
Shape feature	78.0	75.0	72.5
Texture feature	80.0	78.0	76.5
Size feature	76.0	74.0	73.0
Combined features	90.0	88.0	86.5

doi: 10.36922/IJAMD025150011

is left unclassified by the system, thereby maximizing the efficiency of the sorting process.

The F1 score balances the measure between precision and recall, further underscores the performance of a model. Using combined features, an F1 score of 86.5% was observed, demonstrating an effective balance between precision and recall. This result highlights the robustness of the classification framework and its suitability for reliable fruit sorting.

In summary, the results show that the fruit classification system performed significantly better when using a comprehensive set of features. The findings underscore the importance of effective feature selection in achieving high accuracy, precision, recall, and F1 scores, all of which contribute to a reliable and efficient automated fruit sorting mechanism.

3.2. Confusion matrix for fruit classification

The confusion matrix shown in Table 2 indicates the performance of the model across the three fruit classes: Tomato, mango, and orange. The entries along the main diagonal represent correct classifications, while the off-diagonal entries indicate misclassification.

For example, of the 100 actual tomato instances, the model correctly classified 85 of them as tomatoes, misclassifying 10 of them as mangoes and five of them as oranges. The average model accuracy was calculated to be 86.67%, which aligned well with the experimental results of 90% accuracy when using combined features.

The confusion matrix provides an overview of the model's performance and highlights potential areas for improvement. For example, a higher misclassification rate between tomatoes and mangoes immediately indicates that these two fruit classes are similar. Therefore, additional or more distinctive features might be required to enhance the model's ability to distinguish between them.

3.3. Effectiveness comparison of the selected features on image processing system accuracy

This study critically assessed the effectiveness of each feature in enhancing the accuracy of the image processing system in fruit classification. Figure 5 shows the impact of each feature on the performance metrics, including accuracy, precision, recall, and F1 score.

The classification accuracy for the model using different features is shown in Figure 5. It can be seen that combined features improve the model in classifying fruits with reasonable accuracy. Other influential features were color, shape, and texture, with color features showing the highest score. This aligns with a previous study by Chithra and

Table 2. Confusion matrix for fruit classification

Fruit type	Predicted tomato	Predicted mango	Predicted orange
Actual tomato	85	10	5
Actual mango	12	85	3
Actual orange	8	2	90

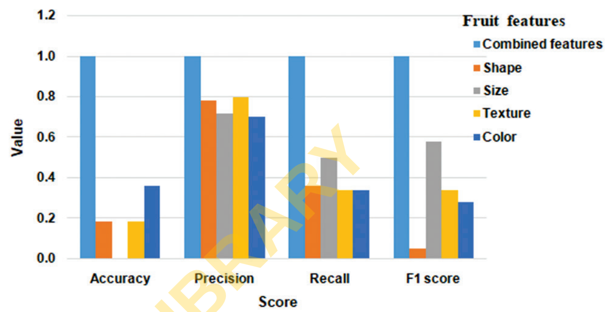

Figure 5. The score of each feature on the performance metrics

Henila,[45] which emphasized the necessity of changing the image format from Red, Green, and Blue to Hue, Saturation, and Intensity to acquire more quantifiable color values. In terms of precision scores, Figure 5 indicates that all features contribute positively to the model's precision score. However, the combination of features yields the highest precision score, confirming that the model is most reliable when using a combined feature set. Similarly, combined features resulted in the highest recall score, indicating the model's ability to capture the maximum proportion of true positive cases. This is particularly important in agricultural applications, where accurate identification of fruit types can directly impact sorting efficiency and marketing decisions. Finally, the combined features resulted in the best F1 score across all features, further reinforcing the conclusion that utilizing all available features significantly enhanced the overall performance of the classification system.

Hence, the comparison of selected features indicates that combined features improve the accuracy of this image-processing system. These findings further highlight the feature selection step in optimizing the performance of the fruit classification model to ultimately contribute to a reliable and efficient automated sorting mechanism.

3.4. Choice of the optimal regularization parameter value in SVM model

One of the most important steps to improve the performance of an SVM model for fruit classification is selecting an appropriate C value. This parameter controls the balance between minimizing training errors and testing errors. A smaller C value allows a greater margin of error, which may lead to more misclassifications. In contrast, a higher C value reduces this margin, thereby reducing misclassification rates.

Figure 6 shows the performance scores for various C values. The results demonstrate that the optimal value of C is influenced by the numerical scale of the input feature values. For example, when feature values ranged from 0 to 100, a C value of 1 yielded strong model performance. In contrast, when the features were scaled between 100 and 1,000, optimal performance was achieved with much higher C values of 100.

This observation underlines the importance of tuning the C parameter in accordance with the scale of feature values. Through cross-validation, the study evaluated model performance across a range of C values to identify the optimal parameter that maximized the classification accuracy while minimizing the risk of overfitting. The results underscore that careful adjustment of the C value is a critical factor in optimizing the SVM model performance and improving accuracy in fruit classification.

This selection of optimum C value is a key part of the SVM training process. The results indicate that properly setting the C parameter with respect to the nature of the feature values is essential. It significantly enhances the efficiency of the image processing system used for fruit classification.

3.5. Comparative analysis of the accuracy of SVM owing to impacts of different kernels

The current research investigated whether significant differences exist in the performance of various kernel functions, aiming to identify the most effective kernel for SVM-based fruit classification. Figure 7 illustrates the accuracy scores for the SVM model using three kernel functions: Linear kernel, polynomial kernel, and RBF kernel.

These results clearly show that the accuracy of the model is highly dependent on the choice of the kernel function. Among them, the polynomial kernel achieved the highest accuracy score. This aligns with the fact that polynomial kernels are efficient in handling non-linear relationships within data, enabling the SVM to generate higher-order, non-linear decision boundaries. In contrast, while the linear kernel performs well on linearly separable data, it did not perform as effectively in this context. This outcome reflects the underlying complexity of the fruit classification task, where features can be non-linearly separable.

The RBF kernel also delivered strong performance, although slightly lower than that of the polynomial kernel. This may suggest that while the RBF kernel is effective in

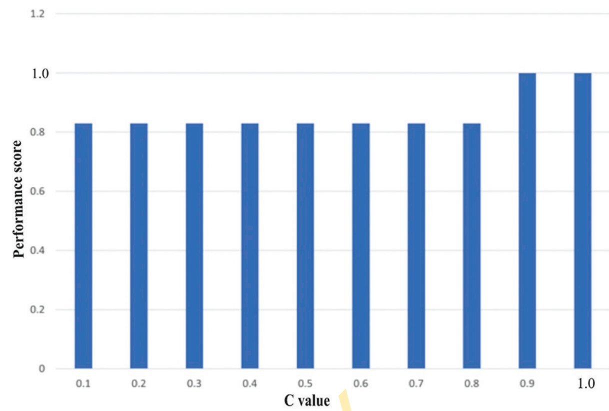

Figure 6. The score of the model across various C values

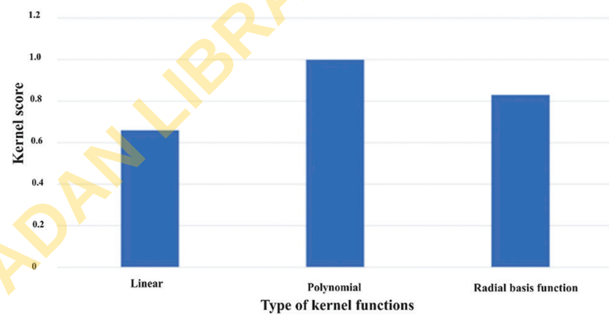

Figure 7. Effect of different kernels on the accuracy of the support vector machine model

capturing local patterns in the data, the polynomial kernel's strength in modeling non-linear relationships contributes to its superior accuracy.

The comparison underscores the importance of selecting a kernel function based on the specific nature of the data and classification task. The results support the conclusion that the polynomial kernel offers a significant advantage in improving accuracy for SVM-based fruit classification applications.

3.6. Hyperplanes constructed from the classification of selected fruits

A critical aspect of the SVM model's functionality lies in its ability to generate a classifying hyperplane to distinguish between fruit classes. Figures 8 and 9 illustrate examples using tomato, mango, and orange data to train the SVM model in generating hyperplanes for their classification.

Figure 8 shows a separating hyperplane in a two-dimensional feature space. This example demonstrates how the model distinguishes between fruit classes within a two-dimensional feature space. The hyperplane serves as a decision boundary, and the clear separation of classes suggests that the

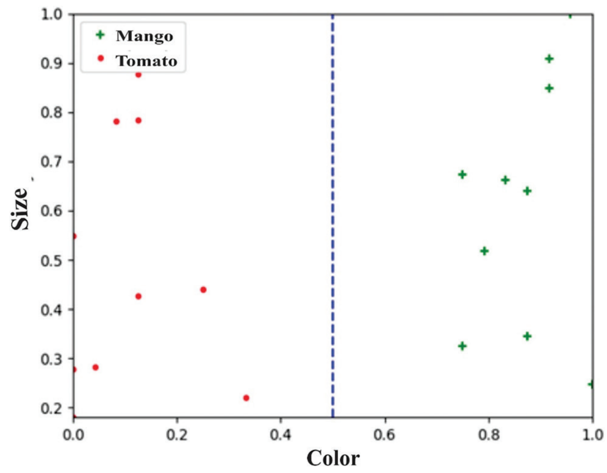

Figure 8. Scatter plot of fruit classification between mango and tomato

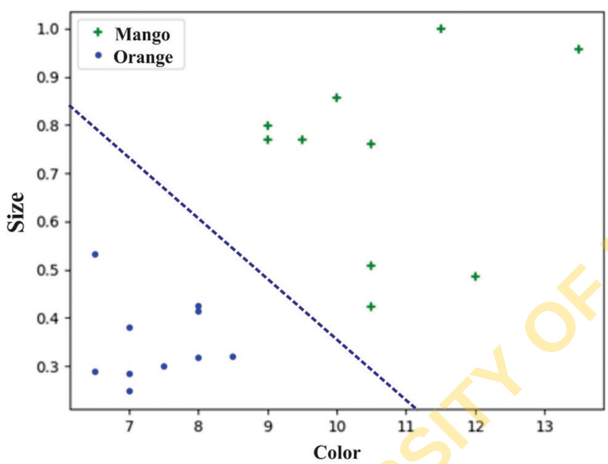

Figure 9. Scatter plot of fruit classification between mango and orange

SVM model efficiently captures the underlying patterns in the data, enabling accurate classification.

Figure 9 provides another visualization of hyperplanes in a two-dimensional feature space. In contrast to Figure 8, where there is a distinct separation between mango and tomato based on the color feature but not size, resulting in a vertical hyperplane. Figure 9 shows a clear separation between mango and orange. Here, both color and size features contribute to the distinction, resulting in a diagonal hyperplane.

3.7. Performance of the SVM model

The performance of the SVM model was evaluated using several standard metrics, as depicted in Figure 10, which shows the model's accuracy, precision, recall, and F1 score in the context of fruit classification.

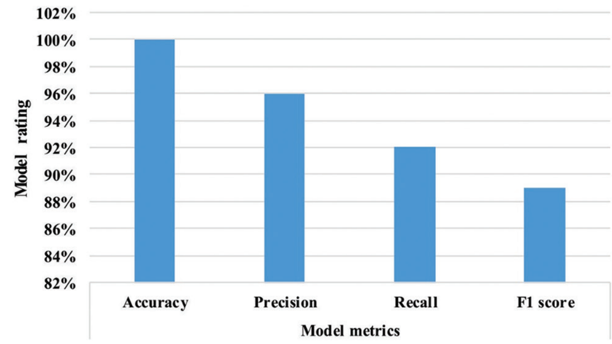

Figure 10. Results obtained from different feature metrics evaluated on the system

The results indicate that the SVM model performed well in classifying the fruits, achieving high scores across all metrics. This is attributed to the effective feature extraction, the selection of an optimal kernel, and careful tuning of the C value.

Therefore, the performance evaluation has identified the capability of the SVM model to classify fruits with reasonably high accuracy using only visual features. This reinforces the potential of image-processing techniques in agricultural applications. The findings also emphasize the importance of selecting appropriate kernels, optimizing model parameters, and use of comprehensive feature sets to further improve the classification accuracy of the system.

4. Discussion

The use of image processing techniques in fruit classification solves all the problems associated with the traditional method of classification, such as subjectivity in classification and potential damage to the fruit. Various features can be utilized in the classification system; while some features allow highlighting the differences across fruit classes easily and accurately, the selection of the most impactful features helps reduce the computational power required.

There are various machine learning models available. Naskar and Bhattacharya[46] utilized artificial neural networks and achieved an accuracy above 90%. Similarly, Bahaghighat *et al.*[47] proposed the use of neural networks and reported reasonable results. Bahaghighat *et al.*[47] demonstrate that the SVM model is comparable in effectiveness to other models while offering the advantage of requiring a significantly smaller dataset.

The SVMs are primarily designed for linearly separable data. However, when dealing with non-linear data, a kernel function can be applied. The kernel effectively projects the input data into a higher-dimensional feature space, where the data are considered linearly separable.[48]

According to Yekkehkhany *et al.*,[48] the linear kernel is the simplest and fastest to process, but the polynomial kernel, though more time-consuming and complicated, often yields more accurate and reliable results. In this study, the polynomial kernel achieved a much greater accuracy than a linear kernel. This suggests that when more than two features are used for fruit classification, utilizing the linear kernel may negatively affect the results obtained.

Color features played a significant role in the classification of the fruits. As such, the quality of the camera and the lighting conditions during image capture are crucial. Poor image quality or inconsistent lighting may distort color representation, particularly due to background interference, which may affect classification results.

5. Conclusion

This paper presents an image processing system developed for classifying selected fruits, including oranges, tomatoes, and mangoes. This system demonstrated tremendous potential in enhancing both the efficiency and accuracy of fruit sorting within associated processes. The system achieved impressive performance metrics, with an accuracy of 100%, precision of 96%, recall of 92%, and F1 score of 89%, as shown in Figure 10. These results demonstrate the effectiveness of the implemented techniques and the robustness of the model in accurately classifying fruits.

The findings highlight the importance of feature selection in the classification process. With the inclusion of multiple features, the accuracy of the system increased significantly. This indicates that certain features contribute more strongly to overall classification performance compared to others. For the SVM model, various kernel functions were evaluated. The results show that the polynomial kernel outperformed both the linear kernel and the RBF kernel, demonstrating its effectiveness in handling non-linear relationships within the data.

Furthermore, it is essential to optimize the C value. This study suggests that the optimal C value is tightly connected with the various features' values. This indicates the need for dataset-specific parameter tuning to achieve optimal performance.

The integration of the Arduino board with the automated sorting mechanism, guided by the classification results from the SVM model, sorted all fruits into their respective containers. This integration of image processing with mechanical sorting provides a proof of concept for a fully automated fruit classification and sorting system.

The image processing system developed in the study demonstrates a reliable and efficient approach to fruit classification, with the potential to enhance quality control and reduce labor costs in the agricultural sector. The findings further emphasize the potential of advanced technologies to replace or improve traditional fruit classification methods. Future research should focus on extending the system to accommodate a wider variety of fruits, along with the integration of machine learning algorithms for real-time fruit classification and sorting.

Acknowledgments

None.

Funding

None.

Conflict of interest

The authors declare that they have no competing interests.

Author contributions

Conceptualization: All authors
Data curation: Oluwaseun Emmanuel Oyewande, John Audu
Formal analysis: All authors
Investigation: Oluwaseun Emmanuel Oyewande
Methodology: All authors
Project administration: Babatunde Olayinka Oyefeso, John Audu
Resources: All authors
Software: Oluwaseun Emmanuel Oyewande, John Audu
Supervision: Babatunde Olayinka Oyefeso, John Audu
Validation: Babatunde Olayinka Oyefeso, John Audu
Visualization: Oluwaseun Emmanuel Oyewande, John Audu
Writing – original draft: Oluwaseun Emmanuel Oyewande
Writing – review & editing: John Audu

Ethics approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

Availability of data

The code, analysis scripts, and datasets supporting this article have been included as part of the supplementary material.

References

1. Ibeawuchi II, Okoli NA, Alagba RA, *et al*. Fruit and vegetable crop production in Nigeria: The gains, challenges and the way forward. *J Biol Agric Healthc*. 2015;5:194-208.

2. Tian H, Wang T, Liu Y, Qiao X, Li Y. Computer vision technology in agricultural automation-a review. *Inf Process Agric*. 2020;7(1):1-19.

   doi: 10.1016/j.inpa.2019.09.006

3. Li D, Song Z, Quan C, Xu X, Liu C. Recent advances in image fusion technology in agriculture. *Comput Electron Agr*. 2021;191:106491.

   doi: 10.1016/j.compag.2021.106491

4. Gupta S, Tripathi AK. Fruit and vegetable disease detection and classification: Recent trends, challenges, and future opportunities. *Eng Appl Artif Intell*. 2024;133:108260.

   doi: 10.1016/j.engappai.2024.108260

5. Kaushal S, Tammineni DK, Rana P, Sharma M, SridharK, Chen H. Computer vision and deep learning-based approaches for detection of food nutrients/nutrition: New insights and advances. *Trends Food Sci Technol*. 2024;146:104408.

   doi: 10.1016/j.tifs.2024.104408

6. Pauzi NAM, Mustaza SM, Zainal N, Zaman MHM, Moubark AM. Artificial Intelligence in precision agriculture: A review. *J Kejuruter*. 2025;37(2):1025-1047.

   doi: 10.17576/jkukm-2025-37(2)-38

7. Iano Y. Image processing-based via agridrones for agriculture fields. In: *Intelligent Designs, Innovations and Sustainability in Agriculture 4.0*. United States: CRC Press; 2025. p. 214.

   doi: 10.1201/9781003380801

8. Amin R, Amin S. Machine learning techniques for increasing the productivity of high-quality food products. In: *Artificial Intelligence in the Food Industry*. United States: CRC Press; 2025. p. 239-259.

   doi: 10.1201/9781032633602

9. Kheiralipour K, Kazemi A. A new method to determine morphological properties of fruits and vegetables by image processing technique and non-linear multivariate modeling. *Int. J. Food Prop*. 2020;23(1):368-374.

   doi: 10.1080/10942912.2020.1729177

10. Han B, Lu Z, Dong L, Zhang J. Lightweight non-destructive detection of diseased apples based on structural re-parameterization technique. *Appl Sci*. 2023;14(5):1907.

    doi: 10.3390/app14051907

11. Wang W, Zhu A, Wei H, Yu L. A novel method for vegetable and fruit classification based on using diffusion maps and machine learning. *CRFS*. 2024;8:100737.

    doi: 10.1016/j.crfs.2024.100737

12. Shen C, Wang R, Nawazish H, Wang B, Cai K, Xu B. Machine vision combined with deep learning-based approaches for food authentication: An integrative review and new insights. *Compr Rev Food Sci Food Saf*. 2024;23(6):e70054.

    doi: 10.1111/1541-4337.70054

13. Usman U, Yunita F, Ridha MR. Improving classification accuracy of local coconut fruits with image augmentation and deep learning algorithm convolutional neural networks (CNN). *JADS*. 2025;6(1):1-19.

    doi: 10.47738/jads.v6i1.389

14. Safari Y, Nakatumba-Nabende J, Nakasi R, NakibuuleR. A Review on automated detection and assessment of fruit damage using machine learning. *IEEE Access*. 2024;12:21358-21381.

    doi: 10.1109/ACCESS.2024.3362230

15. Anjali JA, Bamola A, Mishra S, *et al*. State-of-the-art non-destructive approaches for maturity index determination in fruits and vegetables: Principles, applications, and future directions. *Food Prod Process Nutr*. 2024;6(1):56.

    doi: 10.1186/s43014-023-00205-5

16. Ercan U, Kabas O, Kabaş A, Moiceanu G. Classification of dragon fruit varieties based on morphological properties: Multi-class classification approach. *Sustainability*. 2024;17(6):2629.

    doi: 10.3390/su17062629

17. Barbosa JMR, Santos RGD, Sales LDA, Vargas RBS, Deltsidis A, Oliveira LPD. Image-based and ML-driven analysis for assessing blueberry fruit quality. *Heliyon*. 2025;11(3):e42288.

    doi: 10.1016/j.heliyon.2025.e42288

18. Pathare PB, Opara UL, editors. *Mechanical Damage in Fresh Horticultural Produce: Measurement, Analysis and Control*. Germany: Springer Nature; 2024.

    doi: 10.1007/978-981-99-7096-4

19. Hofman PJ, Bower J, Woolf A, Defilippi BG, Olivares D, García-Rojas M. Harvesting, packing, postharvest technology, transport and processing. In: *The Avocado: Botany, Production and Uses*. United Kingdom: CABI; 2024. p. 622-683.

    doi: 10.1079/9781800621824.0015

20. Khatun T, Nirob MAS, Bishshash P, Akter M, Uddin MS. A comprehensive dragon fruit image dataset for detecting the maturity and quality grading of dragon fruit. *Data in Brief*. 2024;52:109936.

    doi: 10.1016/j.dib.2023.109936

21. Dutta SK, Bhutia B, Misra T, Mishra VK, Singh SK, Patel VB. Application and prospects of artificial intelligence (AI)-

based technologies in fruit production systems. *Appl Fruit Sci*. 2025;67(1):1-18.

doi: 10.1007/s10341-024-01223-4

22. Supekar AD, Wakode M. Multi-parameter based mango grading using image processing and machine learning techniques. *NFOCOMP JCS*. 2020;19(2):175-187.

23. Worasawate D, Sakunasinha P, Chiangga S. Automatic classification of the ripeness stage of mango fruit using a machine learning approach. *AgriEngineering*. 2022;4(1):32-47.

doi: 10.3390/agriengineering4010003

24. Doan TN, Le-Thi DN. A Novel mango grading system based on image processing and machine learning methods. *IJACSA*. 2023;14(5):1118-1129.

doi: 10.14569/IJACSA.2023.01405115

25. Tripathi MK, Neelakantappa M, Mahalle PN, Channapattana SV, Deshmukh G, Prashant G. Detection and classification of mango fruit-based on feature extraction applying optimized hybrid LA-FF algorithms. In: *Data-Centric Artificial Intelligence for Multidisciplinary Applications*. Boca Raton, FL, USA: Chapman and Hall/CRC; 2024. p. 177-185.

doi: 10.1201/9781003461500

26. Kiran PR, Aradwad P, Arunkumar TV, Parray RA. Enhancing mango quality control: A novel approach to spongy tissue inspection through image clustering and machine learning models via X-ray imaging. *J. Food Process Eng*. 2024;47(6):e14664.

doi: 10.1111/jfpe.14664

27. Azeez TB. An automatic mango quality grading system in smart agriculture using novel adaptive feature vector and ensemble learning. *Multimed Tools Appl*. 2025;84:1-26.

doi: 10.1007/s11042-025-20688-3

28. Siddiquee A, Islam MS, Ud Dowla MY, Rezaul KM, Grout V. Detection, quantification and classification of ripened tomatoes: A comparative analysis of image processing and machine learning. *IET Image Processing*. 2020;14(11):2442-2456.

doi: 10.1049/iet-ipr.2019.0738

29. Ropelewska E, Piecko J. Discrimination of tomato seeds belonging to different cultivars using machine learning. *Eur Food Res Technol*. 2022;248:685-705.

doi: 10.1007/s00217-021-03920-w

30. Bai Y, Mao S, Zhou J, Zhang B. Clustered tomato detection and picking point location using machine learning-aided image analysis for automatic robotic harvesting. *Precis Agric*. 2023;24(2):727-743.

doi: 10.1007/s11119-022-09972-6

31. Vo HT, Mui KC, Thien NN, Tien PP. Automating tomato ripeness classification and counting with YOLOv9. *Int J Adv Comput Sci Appl*. 2024;15(4):1120-1128.

32. Moya V, Quito A, Pilco A, Vásconez JP, Vargas C. Crop detection and maturity classification using a yolov5-based image analysis. *Emerg Sci J*. 2024;8(2):496-512.

doi: 10.28991/ESJ-2024-08-02-08

33. Bhargava A, Bansal A. Automatic detection and grading of multiple fruits by machine learning. *Food Anal Methods*. 2020;13:751-761.

doi: 10.1007/s12161-019-01690-6

34. Ismail N, Malik OA. Real-time visual inspection system for grading fruits using computer vision and deep learning techniques. *Inf Process Agric*. 2022;9(1):24-37.

doi: 10.1016/j.inpa.2021.01.005

35. Chakraborty SK, Subeesh A, Dubey K, *et al*. Development of an optimally designed real-time automatic citrus fruit grading-sorting machine leveraging computer vision-based adaptive deep learning model. *Eng Appl Artif Intell*. 2023;120:105826.

doi: 10.1016/j.engappai.2023.105826

36. Raza SM, Raza A, Babeker MIA, Haq ZU, Islam MA, Li S. Efficient citrus fruit image classification via a hybrid hierarchical CNN and transfer learning framework. *J Food Meas Charact*. 2025;19(1):356-377.

doi: 10.1007/s11694-024-02973-1

37. Huong PT, Hien LT, Son NM, Tuan HC, Nguyen TQ. Enhancing deep convolutional neural network models for orange quality classification using MobileNetV2 and data augmentation techniques. *J Algorithms Comput Technol*. 2025;19:1-22.

doi: 10.1177/17483026241309070

38. Lord M. Validation of an invariant embedding method for Fredholm integral equations. *Appl Math Comput*. 1980;7(1):1-7.

doi: 10.1016/0096-3003(80)90030-2

39. Nantomah K, Ege I. A lambda analogue of the gamma function and its properties. *Res Math*. 2022;30(2):18-29.

doi: 10.15421/242209

40. Coulembier K. Homological kernels of monoidal functors. *Sel Math New Ser*. 2023;29(2):24.

doi: 10.1007/s00029-023-00829-y

41. Singh P, Singh N, Singh KK, Singh A. Diagnosing of disease using machine learning. In: *Machine Learning and the Internet of Medical Things in Healthcare*. United States: Academic Press; 2020. p. 89-111.

doi: 10.1016/B978-0-12-821229-5.00003-3

42. Obi JC. A comparative study of several classification metrics and their performances on data. *WJAETS*. 2023;8(1):308-314.

doi: 10.30574/wjaets.2023.8.1.0054

43. Sathyanarayanan S, Tantri BR. Confusion matrix-based performance evaluation metrics. *Afr J Biomed Res*. 2024;27:4023-4031.

doi: 10.53555/AJBR.v27i4S.4345

44. Humphrey A, Kuberski W, Bialek J, *et al*. Machine-learning classification of astronomical sources: estimating F1-score in the absence of ground truth. *MNRAS Lett*. 2022;517(1):L116-L120.

doi: 10.1093/mnrasl/slac120

45. Chithra PL, Henila M. Fruits classification using image processing techniques. *IJCSE*. 2019;7(5):131-135.

doi: 10.26438/ijcse/v7si5.131135

46. Naskar S, Bhattacharya T. A fruit recognition technique using multiple features and artificial neural network. *Int J Comput Appl*. 2015;116:20.

doi: 10.5120/20453-2808

47. Bahaghighat M, Abedini F, S'hoyan M, Molnar AJ. Vision inspection of bottle caps in drink factories using convolutional neural networks. In: *2019 IEEE 15th International Conference on Intelligent Computer Communication and Processing (ICCP)*. United States: IEEE; 2019. p. 381-385.

doi: 10.1109/ICCP48234.2019.8959737

48. Yekkehkhany B, Safari A, Homayouni S, Hasanlou M. A comparison study of different kernel functions for SVM-based classification of multi-temporal polarimetry SAR data. *ISPRS Arch*. 2014;40:281-285.

doi: 10.5194/isprsarchives-XL-2-W3-281-2014, 2014