JTIF : Journal of Technology Information E-ISSN: 3063-0088
Volume 02 No. 03, November 2025 p 08 – 11

CLASSIFICATION OF GUAVA FRUIT QUALITY BASED ON DIGITAL IMAGES USING MACHINE LEARNING

Muhammad Al Qadar Ramadhan

Faculty of Engineering, Putra Indonesia University

Email: ramaqadar@gmail.com

Article Info

Article history:
Received November, 02, 2025
Revised November, 19, 2025
Accepted November, 31, 2025

Keywords:
guava quality,
image
processing,
machine
learning,
KNN, SVM,
RStudio.

Abstract

Guava fruit quality assessment is an important process to ensure the quality of fruit consumed by the public. However, manual quality assessment is still subjective and time-consuming. Therefore, this study aims to classify guava fruit quality based on digital images using computer vision and machine learning approaches. The guava image dataset is divided into two classes, namely good guava and bad guava. Feature extraction is performed by calculating the average values of the RGB (Red, Green, and Blue) color channels from each image. The supervised learning methods used are K-Nearest Neighbor (KNN) and Support Vector Machine (SVM), while the unsupervised learning method uses K-Means Clustering. The system implementation is carried out using RStudio software. The experimental results show that the KNN algorithm achieves an accuracy of 92.39%, while the SVM algorithm achieves an accuracy of 89.17%. Based on these results, it can be concluded that the proposed method is capable of classifying guava fruit quality effectively and has the potential to be used as an automated classification system.

Corresponding Author:
Muhammad Al Qadar Ramadhan
Affiliation author
Email: ramaqadar@gmail.com

1. INTRODUCTION

Guava Fruit Is One Of The Horticultural Commodities That Is Widely Consumed Due To Its High Nutritional Value. Fruit Quality Is A Key Factor That Influences Market Value And Consumer Satisfaction. Guava Fruit Quality Assessment Is Generally Performed Manually Based On Color, Texture, And Physical Condition. However, This Method Has Several Limitations Because It Is Subjective And Highly Dependent On The Assessor's Experience.

The Development Of Image Processing And Machine Learning Technologies Enables Fruit Quality Classification To Be Conducted Automatically And More Objectively. By Utilizing Digital Images, The Visual Characteristics Of Fruits Can Be Analyzed And Used As A Basis For Decision-Making. Several Machine Learning Methods Have Been Widely Applied In Image Classification Tasks, Such As Knn, Svm, And Clustering Methods.

Based On These Problems, This Study Aims To Classify Guava Fruit Quality Using Rgb

DOI: 10.52436/JTIF/ ...

Color Features And Machine Learning Algorithms. The Results Of This Research Are Expected To Provide An Alternative Solution For Automated Guava Fruit Quality Assessment.

2. METHOD

- Data Collection

The data used in this study consist of digital images of guava fruits obtained through a direct image acquisition process. The dataset is divided into two classes, namely good guava and bad guava.

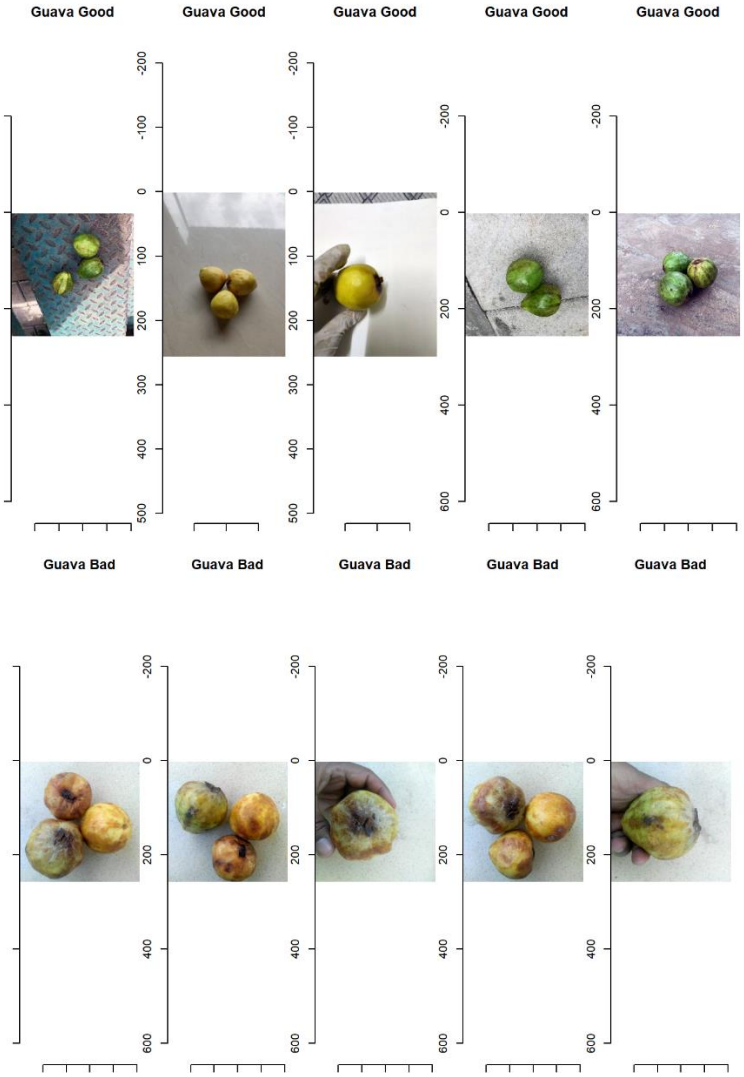

Figure 1. Examples of good and bad guava fruit images

- Image Pre-processing

The image pre-processing stage is conducted to prepare the data before the classification process. This stage includes image resizing and color normalization to ensure uniform image characteristics.

- Feature Extraction

Feature extraction is performed by calculating the average values of the RGB (Red, Green, and Blue) color channels from each image. These average RGB values are used as feature representations that describe the color characteristics of guava fruits.

- Classification Method

The classification methods used in this study include K-Nearest Neighbor (KNN), Support Vector Machine (SVM), and K-Means Clustering. The initial grouping of data is visualized using the K-Means method to observe the distribution patterns of guava fruit quality.

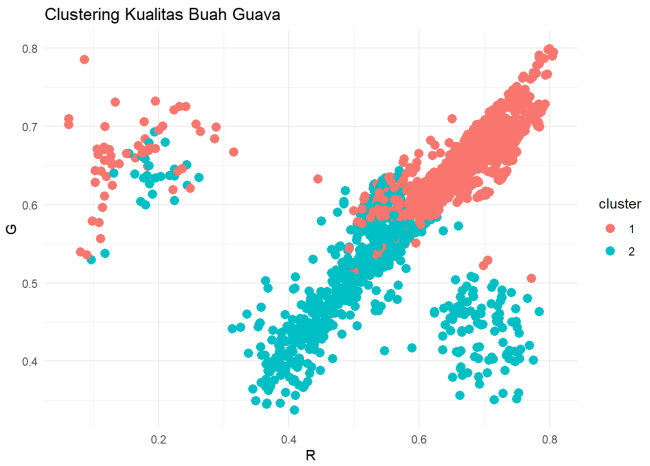

Figure 2. Clustering results of guava fruit quality using K-Means Clustering

- **Evaluation**

Model evaluation is conducted by calculating the accuracy value based on the classification results of the test data.

3. RESULTS AND DISCUSSION

The experimental results show that the KNN algorithm achieves an accuracy of 92.39%, while the SVM algorithm achieves an accuracy of 89.17%. The classification results using the SVM algorithm are visualized in the form of a confusion matrix to demonstrate the model's performance in classifying test data.

Based on the experimental results, the difference in accuracy between the KNN and SVM algorithms is influenced by the characteristics of the image data and the class separation methods used. The KNN algorithm classifies data based on the proximity between data points, allowing it to adapt well to relatively simple RGB color patterns in guava images. Meanwhile, SVM constructs a separating hyperplane that performs optimally on data with specific distributions.

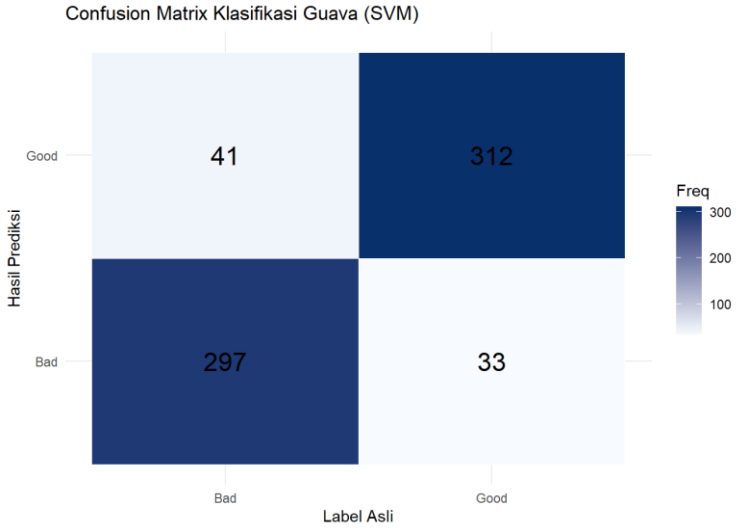

Figure 3. Confusion matrix of guava fruit quality classification using the SVM algorithm

4. DISCUSSION

This section provides a further discussion of the obtained results. The research findings indicate that RGB color features are effective in distinguishing guava fruit quality. Good-quality guava fruits generally exhibit more uniform and brighter color distributions, whereas poor-quality fruits tend to

have darker or uneven color patterns.

The KNN algorithm demonstrates better performance than SVM because it directly considers the similarity between training and testing data. This approach is suitable for datasets with a small number of features, such as average RGB values. Nevertheless, SVM still delivers stable and reliable classification results.

The findings of this study are consistent with previous research, which suggests that machine learning methods based on color features can be effectively applied for fruit quality classification. Therefore, the proposed approach can serve as an alternative solution for automated guava fruit quality assessment systems.

5. CONCLUSION

Based on the conducted research, it can be concluded that machine learning methods using RGB color features are capable of classifying guava fruit quality effectively. The KNN algorithm achieves the highest accuracy of 92.39%, while the SVM algorithm achieves an accuracy of 89.17%. These results indicate that the proposed approach is effective and can be utilized as an automated guava fruit quality classification system.

ACKNOWLEDGEMENTS.

The author would like to express gratitude to the public dataset provider Kaggle for providing the guava fruit image dataset used in this research. The author also appreciates the developers of RStudio, which was utilized as a tool for data processing, analysis, and machine learning implementation.

In addition, the author would like to thank **Mr. Iwan Rizal Setiawan, M.T., M.Kom.**, the lecturer of the **Computer Vision** course, for his guidance, suggestions, and valuable feedback that contributed to the completion of this journal.

REFERENCES

[1] T. Hastie, R. Tibshirani, and J. Friedman, *The Elements of Statistical Learning: Data Mining, Inference, and Prediction*, 2nd ed. New York, NY, USA: Springer, 2017.

[2] I. Goodfellow, Y. Bengio, and A. Courville, *Deep Learning*. Cambridge, MA, USA: MIT Press, 2016.

[3] N. Cristianini and J. Shawe-Taylor, *An Introduction to Support Vector Machines and Other Kernel-Based Learning Methods*. Cambridge, UK: Cambridge University Press, 2018.

[4] T. Cover and P. Hart, "Nearest neighbor pattern classification," *IEEE Transactions on Information Theory*, vol. 13, no. 1, pp. 21–27, 2019.

[5] J. MacQueen, "Some methods for classification and analysis of multivariate observations," in *Proceedings of the Fifth Berkeley Symposium on Mathematical Statistics and Probability*, 2018, pp. 281–297.

[6] Kaggle, "Guava Fruit Image Dataset," 2023. [Online]. Available: https://www.kaggle.com.

[7] RStudio Team, "RStudio: Integrated Development Environment for R," 2023. [Online]. Available: https://posit.co.

[8] R. Gonzalez and R. Woods, *Digital Image Processing*, 4th ed. New York, NY, USA: Pearson, 2018.

[9] S. Alpaydin, *Introduction to Machine Learning*, 3rd ed. Cambridge, MA, USA: MIT Press, 2020.

[10] A. Kadir, *Digital Image Processing: Theory and Applications*. Yogyakarta, Indonesia: Andi, 2019.