Engineering and Technology Journal e-ISSN: 2456-3358
Volume 10 Issue 12 December-2025, Page No.-8368-8373
DOI: 10.47191/etj/v10i12.40, I.F. – 8.482

Apple Freshness Classification Based on Images Using Wavelet and K-Nearest Neighbors Techniques

I Gede Sujana Eka Putra[1], Ni Luh Putu Labasariyani[2]
[1,2] Institute of Business and Technology Indonesia

ABSTRACT: An essential component of fruit marketing and the agriculture sector is determining the freshness of fresh apples. Additionally, the majority Hindu population of Bali uses apples for religious rituals, therefore determining the apples' freshness aids in their selection of high-quality apples. In this study, we provide a k-Nearest Neighbors strategy for apple classification. Using a dataset of apple photos, this study attempts to identify the freshness of apples. Pre-processing and classification of images are among the methods employed. After processing fruit photos to improve contrast, wavelets are used to extract key features. The k-Nearest Neighbors classification algorithm uses these features as input. It is expected that the findings of this study will demonstrate the accuracy of the classification results produced by the k-Nearest Neighbors (k-NN) algorithm. The application is developed using the Python programming language, which is capable of processing image datasets to classify fresh fruit images. This application can be used as a tool to assist in selecting high-quality apples. When the k-NN method was used to evaluate the freshness classification of apples, it performed better at test size training parameter 0.2, with an accuracy value of 0.96 and a precision value of 0.93 for fresh apples and 1 for rotten apples.

KEYWORDS: Apple Freshness, Classification, Wavelet, Images, k-Nearest Neighbors

I. INTRODUCTION

In agriculture, one crucial stage prior to marketing products is the sorting of fruits according to their quality[1]. At present, the sorting of fruit is still conducted using traditional methods, which involve manual checking one-by-one; this approach demands a significant amount of labor and often leads to inaccuracies, as results can differ based on individual perceptions[2]. With the rapid pace of technological advancement, there is a growing need for information to be accessed more quickly. Nowadays, some applications employ color sensors like digital cameras, spectroscopy, or freshness detection tools to assess whether fruits and vegetables are fresh or not, as well as to identify rotten produce; however, these devices tend to be quite costly and may not be accessible to many people[3]. Additionally, in Bali, where the majority of the population is Hindu, apples are utilized for religious ceremonies, creating a strong demand for fresh fruit especially during religious holidays. To address this issue, a system is necessary to effectively assess the quality of apples, thereby enhancing production efficiency and streamlining the transition from fruit production to marketing[4]. This research suggests the creation of an application designed to evaluate apple quality, which would alleviate the burden on workers, reduce sorting time, and minimize costs associated with production and sales. The study proposes a classification method for apples using the k-Nearest Neighbors algorithm, with the primary goal being to identify apple freshness based on a collection of apple images. The methodologies implemented consist of image pre-processing, feature extraction, and classification.

The application created utilizes the Python programming language to analyze image datasets for the purpose of classifying new images. In this context, the research problem focuses on how to create an application that can distinguish the freshness of apples by identifying whether they are fresh or rotten. The scope of this research is constrained by the limited time available, resulting in the dataset comprising only a collection of images of fresh and rotten apples.

II. RELATED WORK

Several studies have been conducted on the fruit freshness classification are as follows: research from [1] about classification of fresh fruits using computer vision techniques for quality and freshness detection. The proposed method consists of image preprocessing, feature extraction, and classification stages. Fruit images are processed to reduce noise and enhance contrast, after which important features are extracted using texture-based feature extraction methods. These features are then used as input for classification algorithms based on Deep Learning, employing a Convolutional Neural Network (CNN) model. Based on the experimental results, the accuracy of fruit quality classification reached 98.17%, while the accuracy of fruit type classification reached 95.38%. This application provides significant benefits in accurately identifying fruit types and quality, delivering precise and reliable information. It can be used as a decisionsupport tool for selecting fruits and enhancing the fruit shopping experience, both online and offline.

Study from [5] proposed to design an application that can classify fresh and rotten fruit. The method that will be used in this design is Convolutional Neural Network (CNN), which is architecture that will be used in this design is AlexNet. The fruits that will be classified are apple, banana, grape, guava, jujube, orange, pomegranate, strawberry, mango and tamarillo. The test results on the training data produce an accuracy of 99% and the test on the test data or validation is 98% with the use of the adam optimizer. The confusion matrix shows that the trained model has an accuracy value of 98%, precision of 98%, recall of 98%, and F1-score of 98%. The output of the application is the introduction of fruit names and classification in the form of fresh or rotten.

Study from [6] is about classification of fresh and rotten fruits using Hu Moment, Haralick, and Histogram feature extraction. In this study, feature extraction was performed using Hu Moments, Haralick, and Histogram, and classification was carried out using the Random Forest algorithm. The researchers aimed to classify fruits as fresh or rotten using the Random Forest algorithm. The results showed a very high accuracy of 99.6%, indicating excellent performance. However, to further improve and update this research, future studies may explore different feature extraction methods and classification algorithms to provide comparisons or achieve more optimal results.

Study from [2] is about detection of Malang Manalagi Apple quality using the Naive Bayes algorithm. This study classifies apples into fresh and less fresh or rotten categories. Current technology enables fruit classification into different classes using various algorithms, one of which is the Naive Bayes algorithm. The method presented in this paper can assist in selecting apples that are fresh or not fresh. The freshness level of apples is divided into two categories: fresh and rotten. In the feature extraction method, the researchers used intensity-based features, and the dataset consisted of 130 images. The classification results achieved an accuracy of 63%.

Study from [7] is about classification of fresh and rotten oranges based on RGB and HSV using the k-NN method. In this research, a classification process will be carried out between fresh oranges and rotten oranges based on RGB (Red, Green, Blue) and HSV (Hue, Saturation, Value) color extraction. This study uses the K-Nearest Neighbor classification algorithm with a value of k = 1; 2; 3; 4; 5; 6; and 7. The dataset used consists of 146 training data and 88 testing data. The purpose and benefits of this research are to save time and facilitate classification according to the wishes of fruit growers. The final result of the test accuracy is 88.95%. Based on the test, this system can be said to be quite good at classifying fresh and rotten citrus fruits.

Study from [8] is about fruit classification using colorized depth images, to investigate the use of colorized depth images in fruit classification with four CNN models, namely, AlexNet, GoogleNet, ResNet101, and VGG16, and compare their performance and computational efficiency, as well as the impact of transfer learning. Depth images of apple, orange, mango, banana and rambutan (Nephelium Lappaceum) were manually collected using a depth sensor with sub-millimeter accuracy and subjected to jet, uniform, and inverse colorization to produce three sets of dataset. Results show that depth images can be used to train CNN models for fruit classification with ResNet101 achieving the best accuracy of 96% on the inverse dataset. It achieved 100% accuracy after transfer learning. GoogleNet showed the most significant improvement after transfer learning on the uniform dataset, at 12.27%. It also exhibited the lowest training and inference times. The results show the potential use of depth images for fruit classification and similar computer vision tasks.

III. METHODOLOGY

A. General System Overview

The procedures employed in this study involved a series of steps. These steps encompassed the phases of creation, data gathering, analysis, and categorization. The sequence of the research process is illustrated in Figure 1.

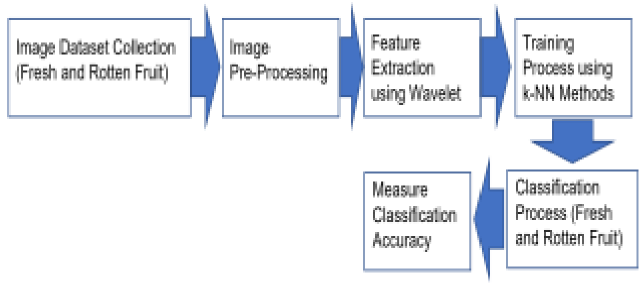

Figure 1. Process Flow and Research Stages

The research stages include data collection in the form of image data of fresh apples and rotten apples. Figure 2 and Figure 3 shows example of datasets image of fresh apples an rotten apples used for training of classification process. The datasets are needed to complete the classification algorithm training process to determine the freshness of new object images.

Figure2. Example of Datasets Image of Fresh Apples

Figure 3. Example of Datasets Image of Rotten Apples

The subsequent phase involves pre-processing images, which includes enhancing contrast, resizing the images for better clarity, and converting them into the HSV color space[8]. Following this, apple image features are extracted utilizing the Wavelet method, a process that aims to gather the necessary features for classification. The next step involves training the k-NN algorithm with the apple image dataset that was compiled during the initial stage. Subsequently, the trained algorithm is evaluated using images of new apples to determine if it can accurately classify the fruit as fresh or rotten. Multiple test datasets containing apple images are utilized to evaluate the algorithm, allowing the k-NN algorithm's classification accuracy to be assessed through the confusion matrix.

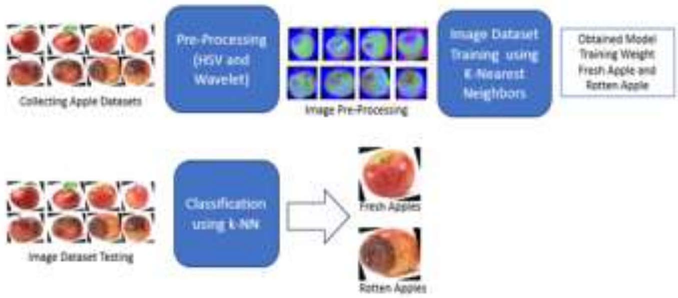

Figure 4. Process Flow for Apple Image Classification

This section outlines the step-by-step approach of the established risk-impact assessment methodology. Figure 4 illustrates the sequence of the apple image classification procedure. The first step involved gathering a dataset of apple images, which included both fresh and spoiled apples. In the second stage, image pre-processing occurs for every apple image by converting the color representation from RGB to HSV color space. The following stage involves extracting image features using the Wavelet Haar algorithm at level 4. Image feature extraction aims to derive features from the apple image that reflect its true characteristics. The outcomes of the image pre-processing are then utilized to train a k-Nearest Neighbors algorithm for categorizing features of fresh and rotten apples. For the training of apple image features, 100 epochs are utilized with a dataset comprising 40 images of fresh apples and 40 images of non-fresh apples. The training results in a weight model distinguishing fresh apple feature from rotten apple features. Subsequently, the process moves on to test a new apple image. Testing is performed using a dataset of apple images that includes 42 fresh apple images and 31 rotten apple images. There is a lot of color information in an image, and this information can also be used to make image analysis tasks like object recognition and image extraction easier[9]. Red (Red), green (Green), and blue (Blue) are the three colors that make up the RGB color space. These colors can be combined in different ways to create other colors[10]. Although it has also been used in regular photography, the RGB color space is mostly employed for displaying images in electronic devices like computers and televisions[11]

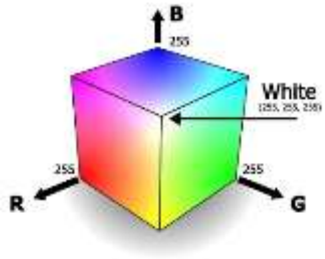

Figure 5. RGB Color Cube

Colors are defined by hue, saturation, and value in the HSV color system. HSV has the benefit of having colors that are identical to what the human senses see. In contrast, a combination of primary colors results in the colors produced by other models, such RGB (Oni et al., 2021). Hue describes the real color and is used to identify redness and greenness in the light's color range. Saturation describes a color's degree of purity, or how much white is added to it. Value is a characteristic that expresses how much light, independent of color, the eye receives.[12]

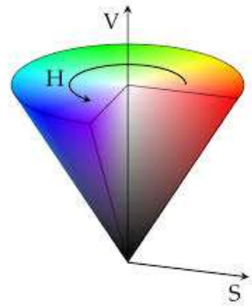

Figure 6. HSV Color Cube

Due to its efficiency, the discrete wavelet transform is typically utilized for feature extraction. By eliminating unnecessary features, the feature selection stage reduces dimensionality. By applying a low-pass filter and a high-pass filter at each stage of decomposition, the picture signal in this transformation may be examined. Rows and columns are filtered. The Symlet wavelet is one kind of discrete wavelet transform that has a high degree of reliability and is superior to other filters in eliminating noise from images as a denoising technique[13]. The original image will be divided into four images with distinct frequencies as a result of the wavelet process. There are 2 types of filters used, namely low-pass and high-pass[14].The following filtering illustration can be seen in Figure 7.

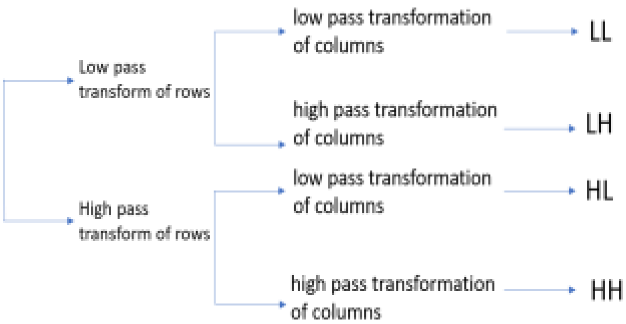

Figure7. Illustration of Wavelet Filtering

A new image with four distinct sorts of frequencies will be created after the image has been broken down using the previously mentioned processes[10][11]. As seen in Figure 8, the result of this operation will indicate that the image has been split into four sections, each of which is ¼ the size of the original image.

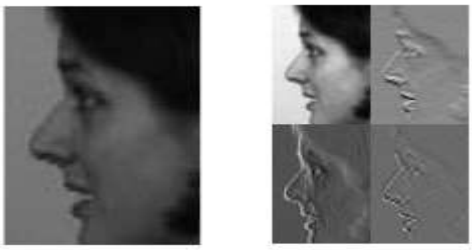

Figure 8. Wavelet Results After the Decomposition Process

IV.RESULT

4.1 Image Pre-Processing

In order to categorize apples as either fresh or rotting, this research creates an application that analyzes apple datasets. As shown in Figure 9, the first step involved gathering datasets of both fresh and rotten apples.

Figure 9. Image Dataset Fresh and Rotten Apples

The next step is to process the image by changing its color space from RGB to HSV.

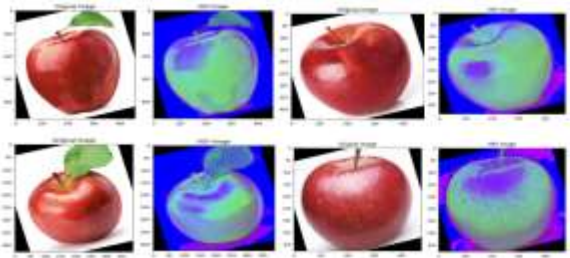

Figure 10. Results of Pre-Processing RGB to HSV Color Space from Fresh Apple Image

As shown in Figure 10, the preparation of the dataset's fresh apple image involves transforming the image from RGB (red, green, and blue) to HSV (hue, saturation, and value) colorspace. Effective color saturation enhancement, airlight estimation to boost contrast, and effective image compression through value component enhancement are all made possible by the HSV colorspace.

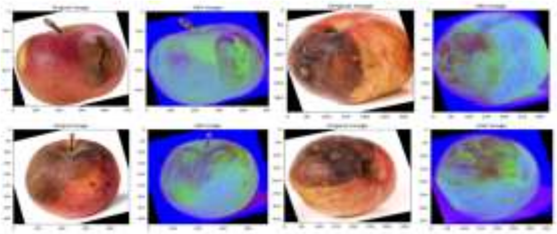

Figure 11. Results of Pre-Processing RGB to HSV Color Space from Rotten Apple Image

4.2 Wavelet Transformation

After converting the apple image to HSV colorspace, the next step is to perform a wavelet modification. As shown in Figure 12, the Haar wavelet with level 4 transformation is the wavelet transformation that was employed.

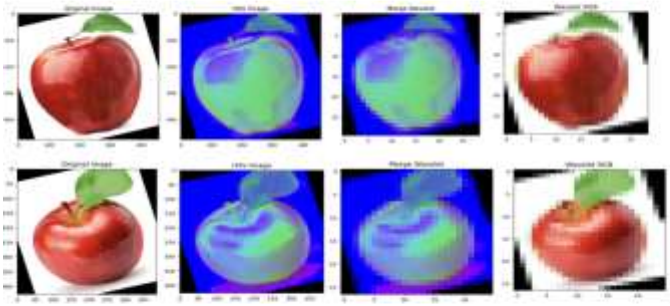

Figure 12. Results of Wavelet Transformation of Fresh Apple Image

The localized characteristic of signals could be analyzed using the Haar wavelet transformation. The Haar function's orthogonal characteristic makes it possible to analyze the input signal's frequency components.

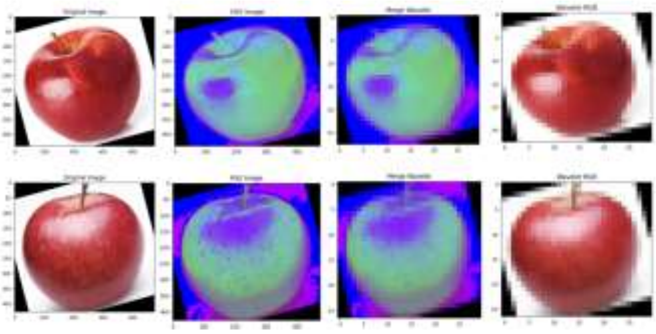

Figure 13. Results of Wavelet Transformation of Another Fresh Apple Image

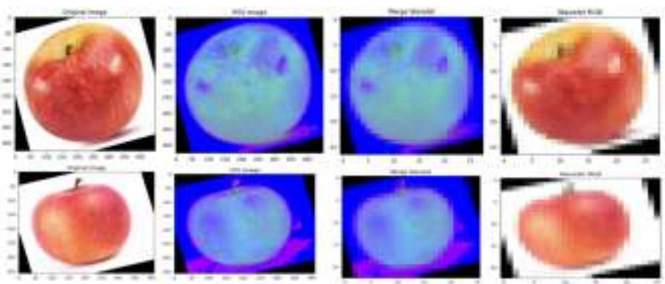

Figure 14. Results of Wavelet Transformation of Another Fresh Apple Image

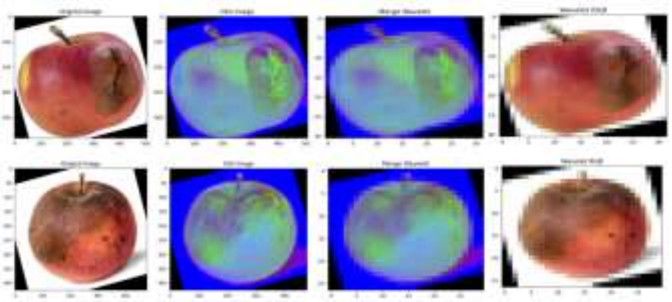

Figure 15. Results of Wavelet Transformation of Rotten Apple Image

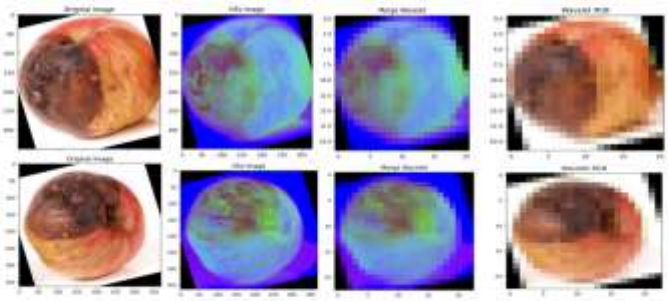

Figure 16. Results of Wavelet Transformation of Another Rotten Apple Image

Figure 13 and Figure 14 shows the result of wavelet transformation of fresh apple image, while Figure 15 and Figure 16 shows the result of wavelet transformation of rotten apple image.

4.3 Training Phase

The k-Nearest Neighbors technique is used in the current study's training phases. Following HSV and wavelet pre-processing, the training dataset is trained using k-Nearest Neighbors with k=2. Python was used in the construction of the system.

4.4 Classification Phase

The k-Nearest Neighbors technique is used in the current study's classification stages. Following HSV and wavelet pre-processing, k-Nearest Neighbors is used to classify the test dataset. The closest distance between each image and the weight of the training outcomes suggests that the image belongs to that class. There are 42 photos in the fresh apple test dataset and 31 images in the rotten apple test dataset. The Confusion Matrix was then used to assess the algorithm's performance based on all test photos. Figure 17 displays the outcomes of evaluating the algorithm's performance.

Figure 17. Confusion Matrix Performance Measurement Results

Based on the performance measurement results of the Confusion Matrix, the performance classification results is shown in Figure 18.

```
Classification Report

               precision    recall  f1-score   support

 Fresh Apples       0.93      1.00      0.97        42
Rotten Apples       1.00      0.90      0.95        31

     accuracy                           0.96        73
    macro avg       0.97      0.95      0.96        73
 weighted avg       0.96      0.96      0.96        73
```

Figure 18. Classification Performance Measurement Results

For rotting apples, we discovered that the precision value was 1, the recall value was 0.9, and the F1-Score was 0.95. The results of the classification performance measurement are displayed in Figure 19. The outcome demonstrates that the test image for a rotten apple may be accurately identified by the system. In contrast, the F1-Score for fresh apples is 0.97 since the precision value is 0.93 and the recall is 1. This demonstrates that 93% of the fresh apple test image can be correctly identified by the system.

V. DISCUSSION AND CONCLUSION

5.1 Discussion

By varying the test_size parameters of training datasets, we compare accuracy algorithms. We use k-Nearest Neighbors to compute the precision, recall, and F-1 score value of the classification after comparing the dataset training parameter of test_size with values of 0.2, 0.3, and 0.4.

Table 1. Performance Classification Result in Different Test_Size Parameter

No	Test Size	Fresh Apples			Rotten Apples		
		Precision	Recall	F-1 Score	Precision	Recall	F-1 Score
1	0.2	0.93	1	0.97	1	0.9	0.95
2	0.3	0.89	1	0.94	1	0.84	0.91
3	0.4	0.91	1	0.95	1	0.87	0.93

Fresh apple classification has the highest precision of 0.93 and the highest F-1 score of 0.97 for test_size = 0.2, according to Figure 19.

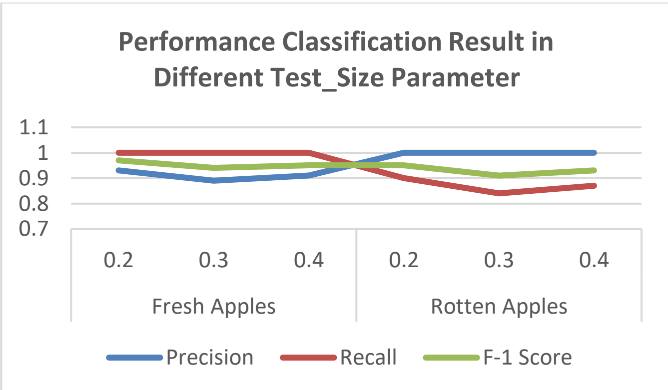

Figure 19. Graph of Performance Measurement Results with Varying Test_Size Parameter

Additionally, it has the highest precision (1), recall (0.9), and F-1 score (0.95) for rotting apples when the test size is 0.2.

5.2 Conclusions

The k-Nearest Neighbors algorithm training procedure, data collection in the form of photographs of fresh and rotten apples, image pre-processing, image feature extraction, and testing phases to see whether the algorithm can identify apples are all included in the study stages. Fresh or rotting apples are among them. When the k-NN method was used to evaluate the freshness classification of apples, it performed better at test size training parameter 0.2, with an accuracy value of 0.96 and a precision value of 0.93 for fresh apples and 1 for rotten apples.

REFERENCES

1. U. Satrya Darmawan Putra Bahari, “Klasifikasi Buah Segar dengan Menggunakan Teknik Computer Vision Untuk Pendeteksian Kualitas dan Kesegaran Buah,” *Jurnal Mahasiswa Teknik Informatika*, vol. 7, no. 3, pp. 1567–1573, 2023.
2. F. E. N. Kholiq Prajatama, “Deteksi Kualitas Buah Apel Malang Manalagi Menggunakan Algoritma Naive Bayes,” *JURNAL SISTEM INFORMASI DAN TEKNOLOGI INFORMASI*, vol. 8, no. 1, 2019.
3. A. Soedarmaji and R. Ediati, “Identifikasi Kematangan Buah Tropika Berbasis Sistem Penciuman Elektronik Menggunakan Deret Sensor Gas Semikonduktor Dengan Metode Jaringan Syaraf Tiruan,” *Jurnal Keteknikan Pertanian*, vol. 25, no. 1, 2011.
4. R. Siskandar, N. A. Indrawan, B. R. Kusumah, S. H. Santosa, I. Irmansyah, and I. Irzaman, “PENERAPAN REKAYASA MESIN SORTIR SEBAGAI PENENTU KEMATANGAN BUAH JERUK DAN TOMAT MERAH BERBASIS IMAGE PROCESSING,” *Jurnal Teknik Pertanian Lampung (Journal of Agricultural Engineering)*, vol. 9, no. 3, 2020, doi: 10.23960/jtep-l.v9i3.222-236.
5. C. L. Prinzky, “Klasifikasi Buah Segar dan Busuk Menggunakan Convolutional Neural Network Berbasis Android,” *Jurnal Ilmu Komputer dan Sistem Informasi*, vol. 10, no. 2, 2022.
6. T. A. M. Fani Nurona Cahya, Rangga Pebrianto, “Klasifikasi Buah Segar dan Busuk Menggunakan Ekstraksi Fitur Hu-Moment , Haralick dan Histogram,” *Indonesian Journal on Computer and Information Technology*, vol. 6, no. 1, 2021.
7. R. P. P. Stifani Napitu, “Klasifikasi Buah Jeruk Segar dan Busuk Berdasarkan RGB dan HSV Menggunakan Metode KNN,” *Jurnal SAINTEKOM*, vol. 1, no. 2, 2023.
8. D. F. K. Gom-os, “Fruit Classification using Colorized Depth Images,” *International Journal of Advanced Computer Science and Applications*, vol. 14, no. 5, 2023, doi: 10.14569/IJACSA.2023.01405106.
9. S. Khansa, “Penerapan Ekstraksi Ciri Transformasi Wavelet Dalam Pembuatan Model Klasifikasi Kesegaran Ikan Selar,” *Senamika*, 2020.
10. A. Pamungkas, “Model Ruang Warna Pengolahan Citra.” Accessed: Dec. 22, 2021. [Online]. Available: https://pemrogramanmatlab.com/2016/06/08/model-ruang-warna-pengolahan-citra/
11. M. Oni, B. Kanata, and D. Ratnasari, “Menentukan Luas Objek Citra Dengan Teknik Segmentasi Berdasarkan Warna Pada Ruang Warna HSV,” *Dielektrika*, vol. 8, no. 2, p. 138, 2021, [Online]. Available: https://dielektrika.unram.ac.id/index.php/dielektrika/article/view/274
12. R. Rahmadewi, G. L. Sari, and H. Firmansyah, “Pendeteksian Kematangan Buah Jeruk Dengan Fitur Citra Kulit Buah Menggunakan Transformasi Ruang Warna HSV,” *JTEV (Jurnal Teknik Elektro dan Vokasional)*, vol. 5, no. 1.1, 2019, doi: 10.24036/jtev.v5i1.1.107560.
13. M. A. Gilbert, N. H. Shabrina, A. Wijaya, and J. P. Wijaya, “Perbandingan Pola Sinyal Penyakit Myocardial Infarction dengan Jantung Normal Menggunakan Metode Wavelet Symlet,” *Ultima Computing : Jurnal Sistem Komputer*, vol. 12, no. 1, 2020, doi: 10.31937/sk.v12i1.1631.
14. A. al-Qerem, F. Kharbat, S. Nashwan, S. Ashraf, and khairi blaou, “General model for best feature extraction of EEG using discrete wavelet transform wavelet family and differential evolution,” *Int J Distrib Sens Netw*, vol. 16, no. 3, 2020, doi: 10.1177/1550147720911009.