International Journal Of Engineering Research And Development
e- ISSN: 2278-067X, p-ISSN: 2278-800X, www.ijerd.com
Volume 20, Issue 6 (June, 2024), PP. 89-98

Image-Based Classification of Apple Freshness Using Wavelet And K-Nearest Neighbors Method

I Gede Sujana Eka Putra
[1]*Department of Computer Informatics, Institute of Business and Technology, Indonesia,*

Ni Luh Putu Labasariyani
Department of Computer Informatics, Institute of Business and Technology, Indonesia,

ABSTRACT:
Detecting the freshness of fresh apples is an important aspect of the agricultural industry and fruit marketing. Moreover, people in Bali, who are predominantly Hindu, use apple fruit for religious ceremonies, so identifying the freshness of apples helps them choose good quality apples. In this research, we propose an apple classification method using the k-Nearest Neighbors method. This research aims to recognize the freshness of apples based on a dataset of apple images. The techniques used include image pre-processing and classification. Fruit images are processed to increase contrast, then important features are extracted using wavelets. These features are used as input for the k-Nearest Neighbors classification algorithm. It is hoped that the research results will show the accuracy of the classification results using the k-Nearest Neighbors algorithm. The application developed uses the Python programming language which can process image datasets to classify new images. This application can be used as a tool for selecting good-quality apples. Testing results of freshness classification of apples using the k-NN algorithm produced better performance at test size training parameter 0.2, an accuracy value of 0.96 was obtained, with a precision value for rotten apples = 1 and fresh apples = 0.93.

Date of Submission: 14-06-2024 Date of Acceptance: 28-06-2024

I. INTRODUCTION

In the world of agriculture, one of the stages before marketing agricultural products is sorting the products based on the quality of the fruit [1]. Currently, the fruit sorting process still uses traditional methods (checking manually/one by one which requires a lot of human power), while this method is often inaccurate and the results vary, due to the different perceptions of each person [2]. Along with the rapid development of technology, the need to obtain information becomes faster. Currently, some applications use color sensors such as digital cameras, spectroscopy, or freshness detection tools. These can determine whether fruit and vegetables are fresh or less fresh and can also detect rotten fruit and vegetables, using color sensors and humidity sensors but the device is quite expensive and not affordable for the public [3]. Moreover, people in Bali, who are predominantly Hindu, use apples as a means of religious ceremonies, so the need for fresh fruit in Bali is very necessary during religious holidays. To overcome this, a system is needed that can help to recognize the quality of apples well, which supports the acceleration of production and streamlines the process after fruit production goes to the fruit marketing stage [4]. This research proposes the development of an application that helps determine the quality of apples to reduce the burden on workers, sorting time, and costs that must be incurred in producing and selling them. This research proposes a method for classifying apples using the k-Nearest Neighbors method. The main aim of this research is to recognize the freshness of apples based on a dataset of apple images. The techniques used include image pre-processing, feature extraction, and classification.

The application developed uses the Python programming language that can process image datasets to classify new images. Based on this, the problem in this research is how to develop an application to classify the freshness of apples using images so that they can be detected as fresh apples and rotten apples. The limitation of the problem in this research is that due to limited research time, the dataset used comes from several images of fresh and rotten apples.

II. MATERIAL AND METHODS

In carrying out this work, the following steps and methods were adopted. The stages are carried out starting from creation to data collection, processing, and classification. The research process flow is shown in Figure 1.

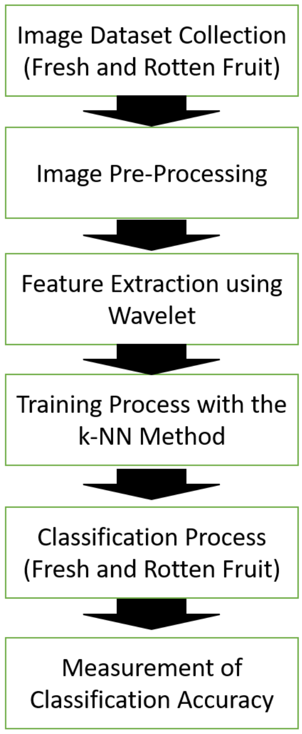

Figure1: Process Flow and Research Stages

The research stages include data collection in the form of image data of fresh apples and rotten apples.

Figure2: Example of Datasets Image of Fresh Apples

Figure 2 shows example of datasets image of fresh apples used for training of classification process.

Figure3: Example of Datasets Image of Rotten Apples

Figure 3 shows example of datasets image of rotten apples used for training of classification process.The datasetsare needed to complete the classification algorithm training process to determine the freshness of new object images. The next stage is image pre-processing, where this activity includes image processing to increase contrast, resize the image to make it clearer, and convert it to HSV color space[13]. The next stage is to extract apple image features using the Wavelet method. This feature extraction functions to obtain image features needed for classification. The next stage is the k-NN algorithm training process using the apple image dataset that was collected in the initial stage. Next, the algorithm that has been trained is tested using images of new apples, to identify whether the algorithm can recognize whether the apple is fresh or rotten fruit. Several test datasets of apple images are used to test the algorithm so that in the end the classification accuracy of the k-NN algorithm can be measured using the confusion matrix.

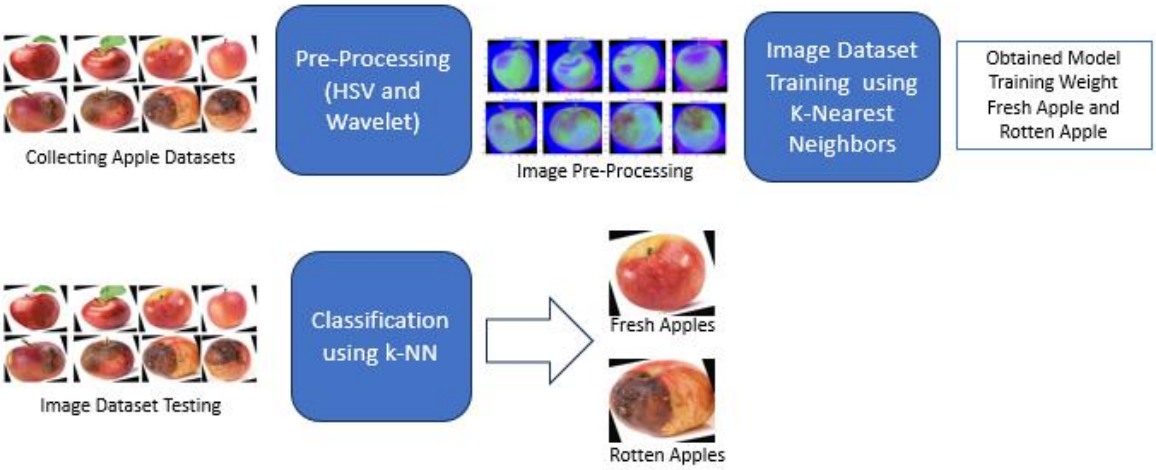

Figure4:Apple Image Classification Process Flow

This section presents the phase-wise description of the developed risk-impact assessment methodology. Figure 4 explains the flow of the Apple image classification process. The initial stage was to collect a dataset of apple images, both fresh and non-fresh apples. The second stage is image pre-processing is carried out for each apple image by converting the color space from RGB color space to HSV color space. The next stage is image feature extraction is carried out using the Wavelet Haar algorithm level 4. Image feature extraction is to obtain features from the apple image to obtain authentic characteristics of the apple image. The image pre-processing results are then trained using the k-Nearest Neighbors algorithm to group fresh apple features and rotten Apple features. Apple image feature training uses 100 epochs with a fresh apple image dataset of 40 images and a non-fresh apple image dataset of 40 images. The results of the training are a weight model for fresh apple features and rotten apple features. After that then continue to test the new apple image. Testing was carried out using a dataset of apple images, with 42 images of fresh apples and 31 images of rotten apples.

In an image, there is a lot of color information, and this information can also be used to simplify image analysis, such as object identification or image extraction[14]. The RGB color space is a color space consisting of 3 colors: red (Red), green (Green), and blue (Blue), which are added in various ways to produce various colors[17]. The main use of the RGB color space is to display images in electronic devices, such as televisions and computers, although it has also been used in ordinary photography.

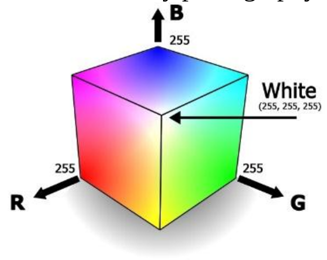

Figure5: RGB Color Cube[17]
(Source: Pamungkas, 2016)

The HSV color space defines colors in terms of Hue, Saturation, and Value. The advantage of HSV is that it contains the same colors as those perceived by the human senses. Meanwhile, the colors formed by other models such as RGB are the result of a mixture of primary colors(Oni et al., 2021). Hue states the actual color, used to distinguish colors and determine redness, and greenness from the color spectrum of light. Saturation states the level of purity of a color, which indicates how much white is given to the color. Value is an attribute that states the amount of light received by the eye regardless of color.

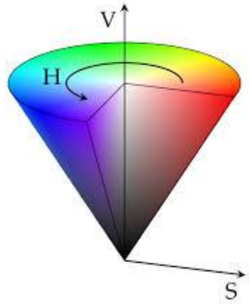

Figure6:HSV Colorspace[17]
(Source: Pamungkas, 2016)

Discrete wavelet transform is usually used for feature extraction because it works efficiently. The feature selection step is used to minimize dimensions by removing irrelevant features [26]. In this transformation, the image signal can be analyzed by passing it through a filtering process consisting of a low-pass filter and a high-pass filter at each decomposition step. Filtering is done on rows and columns. One type of discrete wavelet transform is the Symlet wavelet which is better than other filters in removing noise in images, as a denoising method, and has quite high reliability [15]. The wavelet process means that the original image will be decomposed into 4 images with different frequencies. The original image will be filtered by rows and columns. There are 2 types of filters used, namely low-pass and high-pass [16]. The following filtering illustration can be seen in Figure 7.

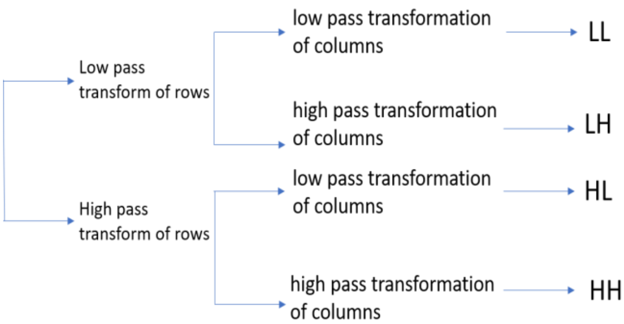

Figure7: Illustration of Wavelet Filtering

There are four types of frequencies resulting from the filtering process. The frequencies that will be produced are Low-Low (LL), Low-High (LH), High-Low (HL), and High-High (HH). Low-low frequency results from low-pass filtering on rows and columns, Low-High Frequency results from low-pass filtering on rows and high-pass filtering on columns, and High-Low Frequency results from high-pass filtering on rows and columns. low-pass filtering on columns, and High-High frequencies resulting from high-pass filtering on rows and columns. The following illustration of the resulting output can be seen in Figure 8.

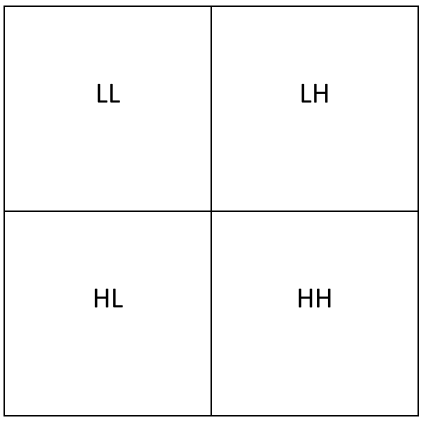

Figure8: Illustration of Wavelet Output

After the image has been decomposed by carrying out the steps above, a new image will be produced with 4 different types of frequencies [17][18]. The output of this process will show that the image is divided into 4 parts with a size of ¼ of the original image as shown in Figure9.

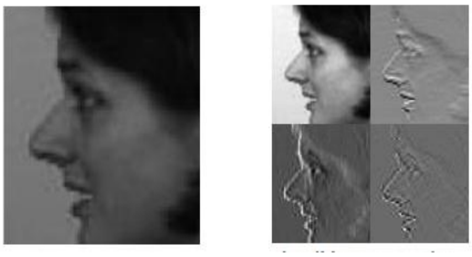

Figure9:Wavelet Results After the Decomposition Process

III. CASE STUDY: LITERATURE REVIEW)

Several previous studies (state of the art) related to fresh fruit classification research, namely Fresh Fruit Classification Using Computer Vision Techniques as seen in Table 1.

Table 1: Summary Reviewed Literatures

Authors	Years	Object Being Tested	Methods	Result
Satrya Darmawan Putra Bahari[1]	2023	Detection of Fruit Quality and Freshness	Computer Vision CNN VGG16	classification accuracy of 98.17%
Prinzky et.al[5]	2022	Fresh and Rotten Fruit Classification	CNN	High classification accuracy
Fani Nurona et. Al[6]	2015	Fresh and Rotten Fruit Classification	Feature Extraction Hu-Moment, Haralick and Histogram	Classification accuracy of 99.6%
Prajatama et.al[2]	2019	Fresh and Rotten Malang Apple Classification	Naive Bayes	Classification accuracy of 63 %
Stifani Napitu1[7]	2023	Fresh and Rotten Orange Classification	Feature extraction RGB and HSV, k-Nearest Neighbors	Classification accuracy of 88,95%

IV. RESULT

4.1 Image Pre-Processing

This research develops an application that processes apple datasets to classify apples as fresh or rotten apples. The initial stage was carried out by collecting datasets of fresh apples and rotten apples as seen in Figure10.

Figure10: Image Datasets of Fresh and Rotten Apples

The next stage is to carry out image processing by converting it into color space from RGB color space to HSV colorspace.

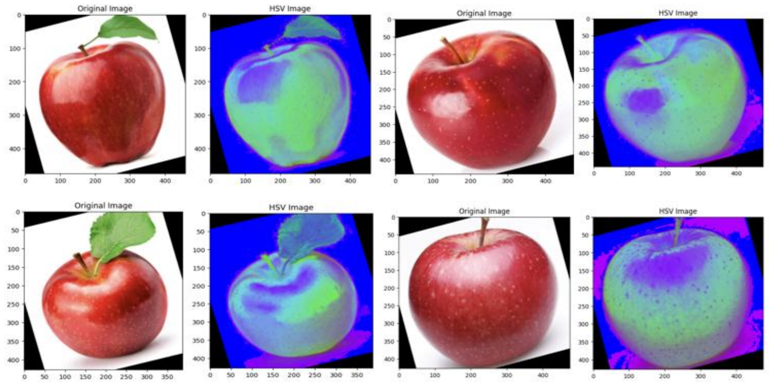

Figure11: Results of Pre-Processing RGB to HSV Color Space from Fresh Apple Image

The preprocessing of datasets fresh apples image includes converting image in RGB (red, green, blue) colorspace into HSV(hue, saturation, value) colorspace as seen in Figure 11. The HSV colorspace allows for effective color saturation improvement, aids in estimating airlight to enhance image contrast, and facilitates efficient image compression by enhancing the value component.

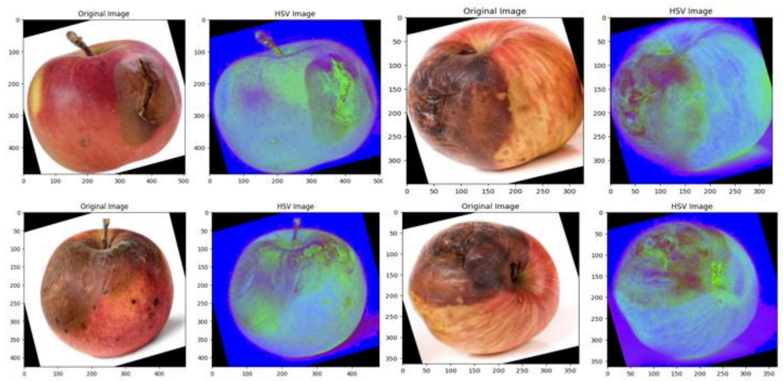

Figure12: Results of Pre-Processing RGB to HSV Color Space from Rotten Apple Image

Likewise applies to rotten apples by converting image in RGB (red, green, blue) into HSV(hue, saturation, value) colorspace as seen in Figure 12.

4.2 Wavelet Transformation

The next stage is to carry out a wavelet transformation of the apple image which has been converted into HSV colorspace. The wavelet transformation used is the Haar wavelet with level 4 transformation as seen in Figure 13.

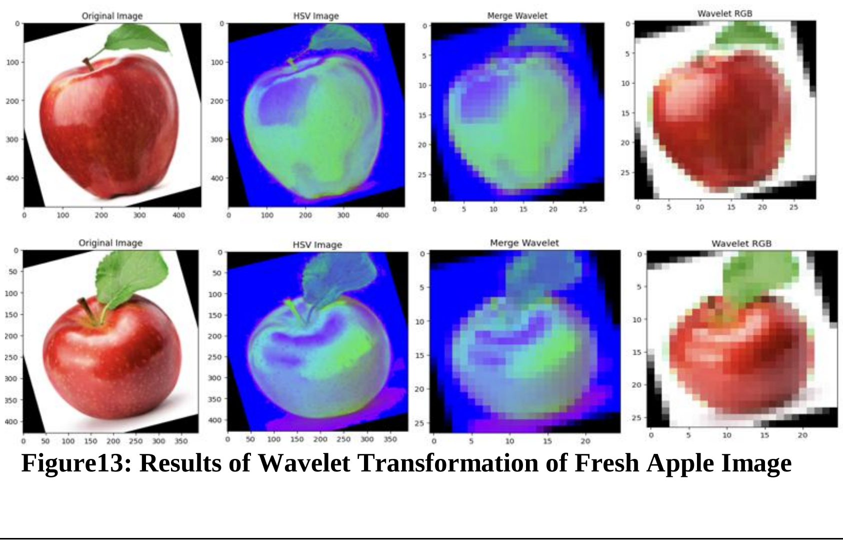

Figure13: Results of Wavelet Transformation of Fresh Apple Image

The Haar wavelet transformation could be used to analyse the localized feature of signals. Due to the orthogonal property of the Haar function, the frequency components of input signal can be analysed.

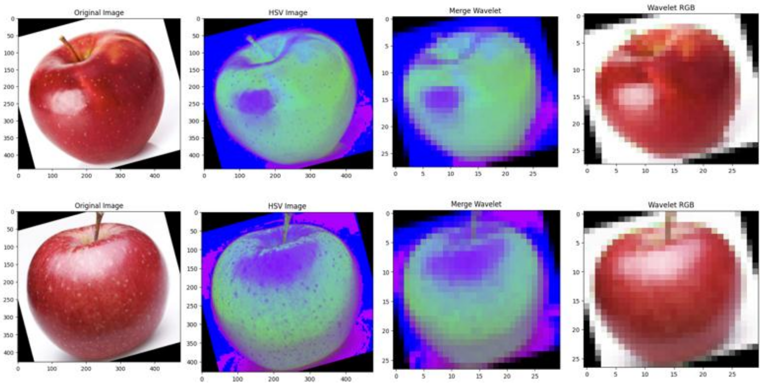

Figure14: Results of Wavelet Transformation of Another Fresh Apple Image

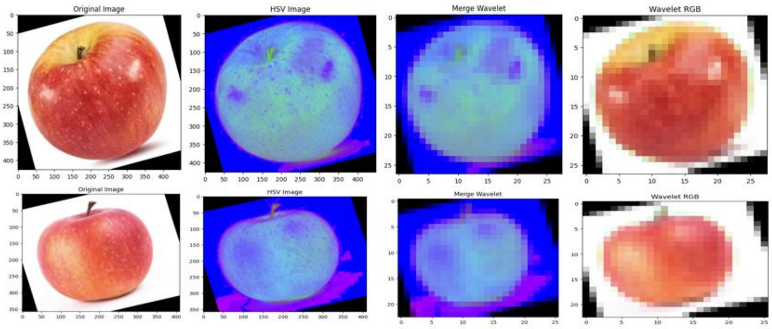

Figure15: Results of Wavelet Transformation of Another Fresh Apple Image

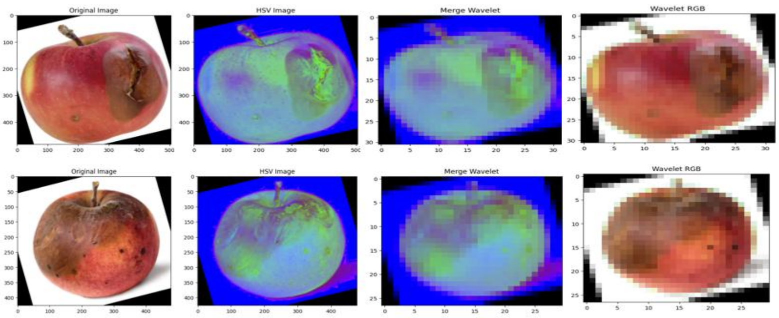

Figure16: Results of Wavelet Transformation of Rotten Apple Image

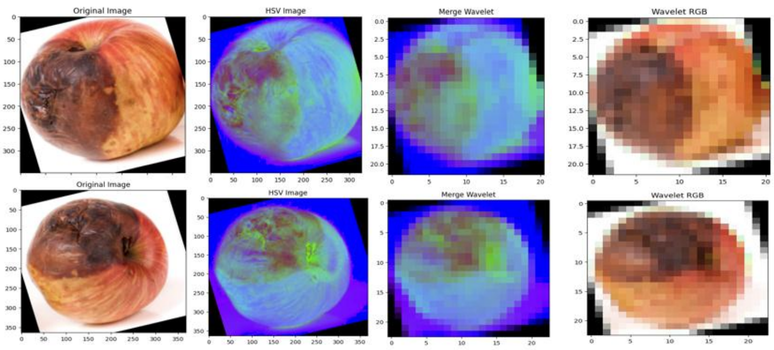

Figure17: Results of Wavelet Transformation of Another Rotten Apple Image

Figure 14 and Figure 15 shows the result of wavelet transformation of fresh apple image, while Figure 16 and Figure 17 shows the result of wavelet transformation of rotten apple image.

4.3 Training Phase

The training stages in the current research use the k-Nearest Neighbors algorithm. The training dataset that has gone through HSV and wavelet pre-processing is then trained using k-Nearest Neighbors with the number k=2. The system was built using the Python programming language.

4.4 Classification Phase

The classification stages in the current research use the k-Nearest Neighbors algorithm. The test dataset that has gone through HSV and wavelet pre-processing is then classified using k-Nearest Neighbors. Next, each image is matched with the weight of the training results, the closest distance indicates that the image is included in that class. The number of fresh apple test datasets is 42 images, while the rotten apple test dataset is 31 images. From the total test images, the performance of the algorithm was then measured using the Confusion Matrix. The results of measuring the performance of the algorithm are shown in the Figure 18.

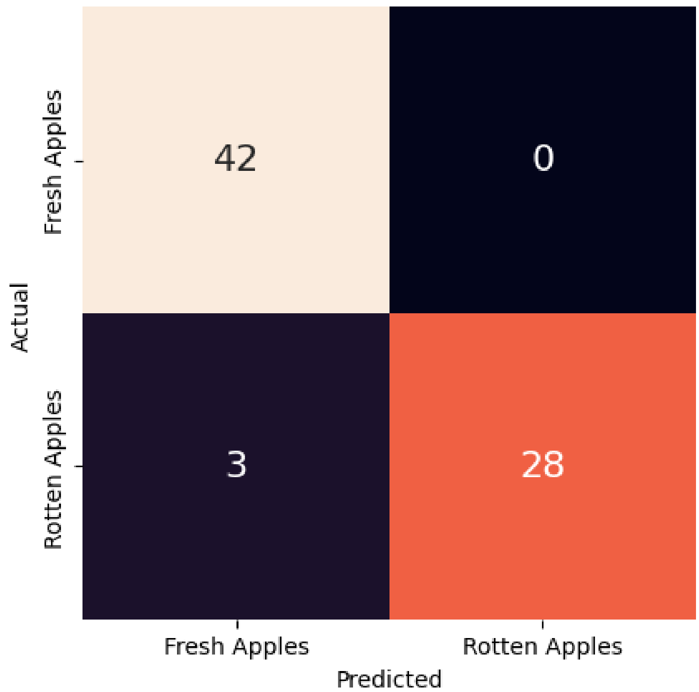

Figure18: Confusion Matrix Performance Measurement Results

Based on the performance measurement results of the Confusion Matrix, the performance classification results is shown in Figure 19.

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

Figure19: Classification Performance Measurement Results

We found that the precision value for rotten apples = 1, the recall value = 0.9 so the F1-Score = 0.95. Figure 19 shows classification performance measurement result. The result shows that the system can correctly recognize the test image for a rotten apple. Meanwhile, for fresh apples, the precision value = 0.93, and recall = 1, so F1-Score = 0.97. This shows that the system can recognize 93% correctly the fresh apple test image.

V. DISCUSSION AND CONCLUSION

5.1 Discussion

We compare accuracy algorithm by adjusting training datasets test_size parameters. We compare dataset training parameter of test_size vary in value 0.2, 0.3, 0.4 and then calculate precision, recall and F-1 score value of classification using k-Nearest Neighbors.

Table 2: 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

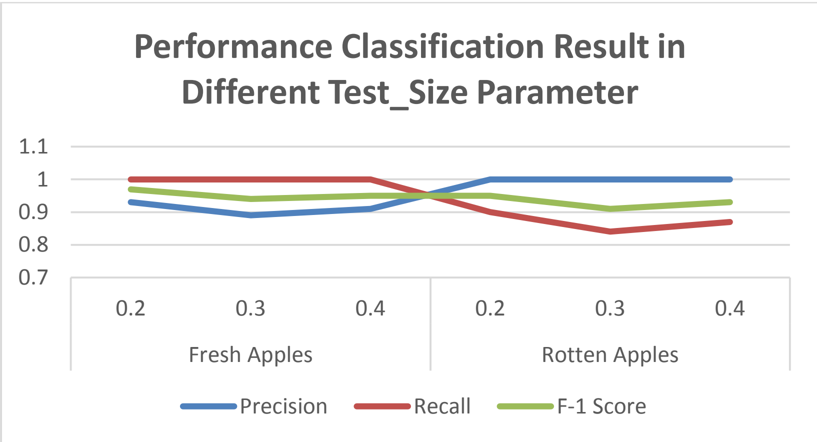

Figure20: Graph of Performance Measurement Resultswith Varying Test_Size Parameter

Based on Figure 20, it found that classification of fresh apples has highest precision 0.93 and highest F-1 score =0.97 for test_size =0.2. It also for rotten apples has highest precision=1, highest recall 0.9 and highest F-1 score 0.95 for test_size = 0.2.

5.2 Conclusion

The research stages include data collection in the form of image data of fresh apples and rotten apples, image pre-processing, image feature extraction, k-Nearest Neighbors algorithm training process, and stages of testing new apple images, to identify whether the algorithm can recognize apples These include fresh or rotten apples. Testing results of freshness classification of apples using the k-NN algorithm produced better performance at testsizetraining parameter 0.2, an accuracy value of 0.96 was obtained, with a precision value for rotten apples = 1 and fresh apples = 0.93.

REFERENCES

[1] U. Satrya Darmawan Putra Bahari, “Klasifikasi Buah Segar dengan Menggunakan Teknik Computer Vision Untuk Pendeteksian Kualitas dan Kesegaran Buah,” *J. Mhs. Tek. Inform.*, vol. 7, no. 3, pp. 1567–1573, 2023.

[2] F. E. N. Kholiq Prajatama, “Deteksi Kualitas Buah Apel Malang Manalagi Menggunakan Algoritma Naive Bayes,” *J. Sist. Inf. DAN Teknol. Inf.*, 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,” *J. Keteknikan Pertan.*, 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,” *J. Tek. Pertan. Lampung (Journal Agric. Eng.*, 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,” *J. Ilmu Komput. dan Sist. Inf.*, 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," *Indones. J. Comput. Inf. Technol.*, vol. 6, no. 1, 2021.

[7] R. P. P. Stifani Napitu, "Klasifikasi Buah Jeruk Segar dan Busuk Berdasarkan RGB dan HSV Menggunakan Metode KNN," *J. SAINTEKOM*, vol. 1, no. 2, 2023.

[8] J. R. Bermúdez, F. R. López-Estrada, G. Besançon, L. Torres, and I. Santos-Ruiz, "Leak-Diagnosis Approach for Water Distribution Networks based on a k-NN Classification Algorithm," *IFAC-PapersOnLine*, vol. 53, no. 2, pp. 16651–16656, 2020, doi: 10.1016/j.ifacol.2020.12.795.

[9] A. M. Ismail, "Cara Kerja Algoritma k-Nearest Neighbor (k-NN)," *Bee Solution Partners*, 2018. https://medium.com/bee-solution-partners/cara-kerja-algoritma-k-nearest-neighbor-k-nn-389297de543e (accessed Jul. 05, 2022).

[10] P. Cunningham and S. J. Delany, "K-Nearest Neighbour Classifiers-A Tutorial," *ACM Computing Surveys*, vol. 54, no. 6. 2021. doi: 10.1145/3459665.

[11] H. Saadatfar, S. Khosravi, J. H. Joloudari, A. Mosavi, and S. Shamshirband, "A new k-nearest neighbors classifier for big data based on efficient data pruning," *Mathematics*, vol. 8, no. 2, 2020, doi: 10.3390/math8020286.

[12] E. Sentosa, H. Armanto, C. Pickerling, and L. Zaman PCSW, "Pengenalan Ekspresi Wajah dengan CNN dan Wavelet," *J. intellegent Syst. Comput.* , vol. 4, no. 02, pp. 69–76, 2022, doi: 10.52985/insyst.v4i2.209.

[13] "SISTEM PENGENALAN KUALITAS IKAN GURAME DENGAN WAVELET, PCA, HISTOGRAM HSV DAN KNN," *Lontar Komput.*, 2015, doi: 10.24843/LKJITI.

[14] S. Khansa, "PENERAPAN EKSTRAKSI CIRI TRANSFORMASI WAVELET DALAM PEMBUATAN MODEL KLASIFIKASI KESEGARAN IKAN SELAR," *Senamika*, 2020.

[15] 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," *Ultim. Comput. J. Sist. Komput.*, vol. 12, no. 1, 2020, doi: 10.31937/sk.v12i1.1631.

[16] 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. Networks*, vol. 16, no. 3, 2020, doi: 10.1177/1550147720911009.

[17] A. Pamungkas, "Model Ruang Warna Pengolahan Citra," 2016. https://pemrogramanmatlab.com/2016/06/08/model-ruang-warna-pengolahan-citra/ (accessed Dec. 22, 2021).

[18] 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