EARLY DETECTION OF ROT IN THAI PAPAYA (CARICA PAPAYA) USING THE K-NN METHOD

Agus Prayitno[1*]

Information Technology[1]
Universitas Widya Kartika Surabaya, Surabaya, Indonesia[1]
http:widyakartika.ac.id[1]
agus.prayitno.sby@gmail.com[1*]
(*) Corresponding Author

Abstract— *Determining the category of a plant or fruit involves several criteria. One of the easiest methods to use is morphological criteria, which entails studying the external structure that can be directly observed. However, this approach cannot be regarded as a fixed standard since people's interpretations may vary. To address this, a system was developed to assess the ripeness of Thai papaya fruit, utilizing image processing and the K-Nearest Neighbor (KNN) method. This study analyzes a data set to detect rotten papaya fruit, which is expected to help consumers recognize papaya fruit that is purchased in a perfectly ripe condition, not ripe with certain parts that are rotting. The indicator used to determine the category is the color of the skin of Thai Papaya fruit with an ROI of 600 pixels x 300 pixels by finding the mean RGB value and then calculating it using the Euclidean distance formula. From the results of these calculations, it is expected to get a classification using K-Nearest Neighbor (KNN) to get an image pattern of the level of rottenness on the surface of the papaya. Therefore, by improving the RGB image eliminating noise in the papaya image, and using the K-NN classification of the image pattern obtained from the research results from the sampling data, an accuracy level of 80% was obtained with a range of mean R values: 130,671-169,630, mean G: 106,891-131,895, and mean B: 61,119-100,776 which came from 120 data.*

Keywords: *classification, k-nn, rotten papaya, thai papaya.*

Intisari—Menentukan kategori suatu tumbuhan ataupun buah-buahan tentunya ada suatu kriteria tertentu. Kriteria morfologi merupakan salah satu kriteria yang paling mudah digunakan karena untuk mengetahui kriteria perlu mempelajari struktur bagian luar yang dapat dilihat secara langsung. Namun hal ini tentunya tidak bisa dijadikan sebagai kriteria tetap karena pandangan setiap orang yang berbeda. Oleh sebab itu dibuatlah sistem pengecekan tingkat kebusukan buah Pepaya Thailand yang dapat dilakukan dengan konsep pengolahan citra dengan menggunakan metode *K-Nearest Neighbor* (KNN). Penelitian ini menganalisis sekumpulan data untuk mendeteksi buah pepaya busuk, yang diharapkan dapat membantu konsumen mengenali buah pepaya yang dibeli dalam kondisi matang sempurna, belum matang dengan bagian tertentu yang membusuk. Indikator yang digunakan untuk menentukan kategori adalah warna kulit buah Pepaya Thailand dengan ROI 600 piksel x 300 piksel dengan mencari nilai rata-rata RGB kemudian menghitungnya menggunakan rumus jarak Euclidean. Dari hasil perhitungan tersebut diharapkan mendapatkan klasifikasi menggunakan K-Nearest Neighbor (KNN) untuk mendapatkan pola citra tingkat kebusukan pada permukaan buah pepaya. Oleh karena itu, dengan melakukan perbaikan citra RGB dengan menghilangkan noise pada citra pepaya, serta menggunakan klasifikasi K-NN pada pola citra yang diperoleh dari hasil penelitian data sampling, diperoleh tingkat akurasi sebesar 80% dengan rentang nilai mean R: 130.671-169.630, mean G: 106.891-131.895, dan mean B: 61.119-100.776 yang berasal dari 120 data.

Kata Kunci: *klasifikasi, k-nn, pepaya busuk, pepaya thailand.*

INTRODUCTION

Papaya is a fruit that is rich in benefits. There are 3.65 milligrams of vitamin A and 78 milligrams of vitamin C in every 100 grams of papaya. Papaya can be made into jam by adding granulated sugar and citric acid. (Dihni, V. A. ,2022). The result is a delicious, smooth, and shiny jam. One of the advantages of papaya is its ability to continue to

Techno Nusa Mandiri: Journal of Computing and Information Technology
Vol. 22, No. 1 March 2025 | DOI https://doi.org/10.33480/techno.v22i1.6522

bear fruit regardless of seasonal changes. Papaya production continues to increase from year to year, according to data from the information system of the Central Bureau of Agricultural Statistics and the Ministry of Agriculture of the Republic of Indonesia. In 2018, Indonesia's papaya production was 887,580 tons, then increased to 986,991 tons in 2019 and to 1,016,388 tons in 2020 (Dihni, V. A. ,2022).

The largest papaya fruit production in Indonesia in 2020 was in the province of East Java (Agustina & Sukron, 2022). Fruit is a type of food that has a very important role type of human health. Various kinds of vitamins can be found in fruit. One type of fruit that is often consumed by people is papaya fruit. This plant is easy to cultivate and can be found in almost all tropical areas because this plant is a plant that knows no seasons. The fruit is delicious, contains lots of vitamins, can be processed into other preparations, and has many benefits for humans.

This makes papaya fruit very potential as a food for daily consumption, whether to be eaten directly when it is ripe or processed again. Indonesia itself is one of the largest papaya-producing countries in the world. Because this fruit can grow at any time regardless of climatic conditions, many people grow this plant either for personal consumption or to produce it on a large scale. However, before using the papaya plant, of the process it is necessary to carry out an identification process to obtain information from the fruit which of course will be very useful to help in the grouping process.

The process of grouping certainly has certain criteria. The easiest criteria to use are morphological criteria because these criteria study plant structures that can be seen directly or easily observed so that variations can be assessed quickly. It is with these criteria that people usually determine the level of ripeness or rottenness of fruit, especially Thai papaya fruit. Of course, this cannot be said to be a fixed criterion because of each person's vision and perception. For example, Mr. Rudi thinks that the Thai Papaya fruit is rotten because white fungus has appeared, but Mr. Budi thinks that it is still ripe because only a small amount of fungus has grown.

Therefore, we need a fixed criterion or system that functions to help the public determine the level of rottenness of papaya fruit more easily and more accurately. To carry out the process of checking the level of rottenness of papaya fruit itself, an image processing concept is needed to check it because the easiest indicator to see is the fruit skin color indicator. One tool or device that can produce quite good images today is a smartphone camera. With the help of the camera on a smartphone, of course, it is easy to identify the level of rotten papaya fruit through image processing using RGB color selection and with the help of the K-Nearest Neighbor (KNN) classification algorithm so that it can increase the level of accuracy in checking.

This research to check the level of rottenness of Thai Papaya fruit itself was of course carried out by reviewing several previous existing studies, but with different aims and objectives that had been carried out by other researchers. Research conducted previously (Afifah, L. ,2020) obtained the same average accuracy of 75% with several K-Neighbors of 5 and 7. Research conducted by Muhammad Ezar Al Rivan and Jessica Suherman was entitled "Determining the Quality of California Papaya Fruit (Carica Papaya L.) Using Fuzzy Mamdani". The results of the research obtained were the highest accuracy obtained, namely 75% using the Centroid, Bisector, and MOM defuzzification methods. The LOM and SOM defuzzification methods provide results of 70%. By using the confusion matrix it can be seen that the average accuracy is 80% to 83%.

Based on the description above, it is also necessary to research to check the level of rottenness of papaya fruit in a way that is relatively easy and has a good level of accuracy. It is also hoped that this research can be used as additional information regarding checking the level of fruit rot through all related scientific fields.

MATERIALS AND METHODS

To support the creation of a system, of course, a large collection of data is needed. Collecting the data needed to create a "Thai Papaya (Carica Papaya L.) Rotten Level Checking System Using the K-Nearest Neighbor (KNN) Method Classification Algorithm" is: Carrying out literature studies by collecting information obtained from reading related books, and related reference journals, and collecting information related to the objects and methods that will be used for research.

Analyzing system requirements by studying how the system works based on system design. This step was carried out to find out what methods will and can be used to create this system. Apart from that, analysis was also carried out on related research references so that the creation of this system could be by the expected goals. The method used to develop this system is the waterfall method which explains the system's working stages regularly or systematically and sequentially so that the steps in the process will be more optimal and efficient. The working system is as follows :

a. Needs Analysis

This stage is carried out by the data collection process by determining a certain distance on the papaya object, the angle of image capture, and by using lighting with room conditions. Collection of papaya image data taken in good condition and will rot. Thus, a data set of papaya fruit with good and rotten conditions will be obtained which will be used as test data and training data for the image processing process in the study.

b. System Design

This stage is carried out to provide an overview or explanation of how papaya and papaya image capture works as a data set. Papaya is obtained from several traditional markets in Surabaya. Papaya images are primary data taken directly. Thus, correct image data will be obtained. Modeling the image pattern of the initial decay level in papaya fruit can be used to facilitate the analysis process and can be used to build a further system (Margaretha & Voutama, 2023).

c. Implementation

After getting the papaya image, the next step is to process the original image obtained from the cellphone camera. Papaya image data of 30 papayas with 4 sides of the capture so that the total image obtained is 120 ripe data set images and papaya images that are about to rot. The data is divided into two parts, namely 80% of the training data and 20% of the testing data by labeling each papaya image one by one. The next step is to process the sRGB image which is converted into RGB image format to facilitate further calculations. The RGB image data set obtained will be processed with statistical calculations by finding the mean value in the image. From the division of the data, classification is carried out using K-NN to obtain the expected results, namely finding the pattern of the level of decay of the papaya surface.

d. Integration

At this stage, the Research created and developed will undergo a thorough re-testing to find out whether the classification results of the test data and training data work by the initial design objectives. In addition, this test is carried out to identify potential problems or deficiencies that may not have been previously detected, so that improvements can be made to classify and provide a model for papaya that has found a pattern of image levels of decay in papaya.

e. System maintenance

Digital imagery is a field of science that studies how an image is formed, processed, and analyzed to produce information that can be understood by humans. Digital imagery represents the light intensity function in a two-dimensional discrete form. The image is composed of a collection of pixels (picture elements) and has coordinates (x, y) and amplitude f (x, y). The coordinates (x, y) indicate the location of the pixel in the image while the amplitude f (x, y) indicates the color intensity value of the image. The larger the pixel size, the more precise the image obtained. Based on the combination of colors in pixels, images have three types, namely RGB (Red, Green, Blue) images, grayscale images, and binary images. In this research, it is expected to obtain the results of the rotten papaya color classification pattern.

Literature study is conducted with the aim of collecting theoretical ideas related to the research subject. Literature study is an important component in the research planning stage that focuses on extraction using deep learning methods at the feature learning stage and image detection using the K-NN method. By conducting a literature study, researchers can understand the strong theoretical foundations that support the development of the method.

In addition, in this effort, references from books, journals, online articles, and relevant sources will be used as key reference materials to understand the concepts related to image processing, pattern recognition, and deep learning techniques that underlie this research. In this way, the literature study will ensure that the research is at the forefront of knowledge and can provide significant contributions to solving the intended problem. The K-Nearest Neighbor (KNN) algorithm is a method used to classify objects based on the training records that are closest to the object being tested, where *k* represents the number of nearest neighbors considered in the classification process (Malik Namus Akbar, 2024.).

KNN is a simple algorithm that works by calculating the distance difference between the test data and the training data in the system. This data is then processed by finding the most similar value, which is then used for classification. KNN can be used to insert new data (test data) into a group of data that is close to the existing training data. Therefore, this method can be used to classify test image data according to the appropriate image data group. K-Nearest Neighbor will group the calculation results with the training data that has the most relations within the specified range of values (Afifah, L, 2022). In the process, the object will be classified into the nearest neighbor class, with the value of *K* being a parameter that has been previously set by the user. The steps of the K-Nearest Neighbor algorithm are as follows:

1. Determine the parameter K (the number of closest neighbors).

Techno Nusa Mandiri: Journal of Computing and Information Technology
Vol. 22, No. 1 March 2025 | DOI https://doi.org/10.33480/techno.v22i1.6522

2. Calculate the square of the Euclidean distance of the object to the given training data.

$$d(x,y) = \sqrt{\sum_{i=1}^{n}(x_{i-Y_i})} \quad (1)$$

$$d(x,y) = \sqrt{(X_1 - X_2)^2} + \sqrt{(Y_1 - Y_2)} \quad (2)$$

Description:
X = Test data value
Y = Training data value

3. Then sort the results of no. 2 in order from the highest to the lowest value.
4. Collect the Y category (nearest neighbor classification based on the k value).
5. By using the most majority K-nearest neighbor category, new objects can be predicted.

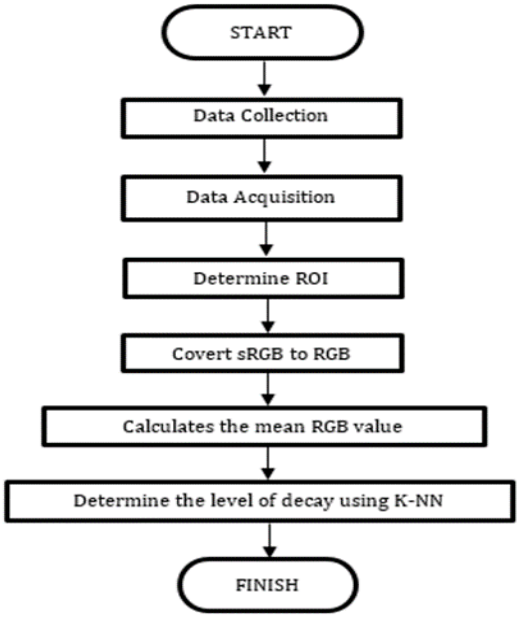
Classification involves the process of identifying patterns or functions that can be used to distinguish between different classes. In other words, classification is a systematic arrangement used to separate an object being studied based on certain rules that have been established or standardized. The main purpose of classification is to determine the class of an object whose class is not yet known. (Afifah, L, 2022). Thus, the same data or object will belong to the same group (Putra et al., 2022). There are two processes in data classification, namely the training process and the testing process. The training process is the process of training data used in making predictions. The testing process is the process of testing the dataset to determine the level of accuracy of the model.

RGB is a method for representing colors. In the RGB color space, the image shows the primary spectral components of red, green, and blue, arranged based on the cartesian coordinate system ((Rabbani et al., 2021). Usually, RGB is used for computer graphics. Used for televisions, computers to other electronic layers. The abbreviation of RGB itself is Red (R), Green (G), and Blue (B). Each color component uses 8 bits (value range 0 to 255), which when presented is around 16,777,216. It's just that the human eye can only capture approximately 7 million colors. As a result, many colors appear the same to the human eye.

This leads to some colors being rarely used. RGB is ideal for digital images because when light is added simultaneously with these three colors, it extends the wavelength of the color. Longer wavelengths of color allow for more accurate and complex color mixing interactions due to the creation of new colors.

RESULTS AND DISCUSSION

The following are the stages carried out in the research which can be described in Figure 1.



Source : (Research Results, 2024)
Figure 1. Flow of Knowledge Discovery in Database Methodology

based on figure 1, the following are the details related to each process in figure 1:

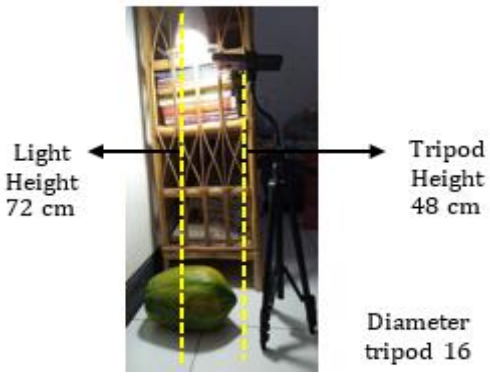
a. Data Collection

Collecting 30 photos of papaya fruit data which are divided into 3 parts, namely half-ripe, mature and rotten. In this research, a dataset comprising a collection of papaya images will be used, specifically focusing on Thai papaya. The images were captured by documenting each papaya individually in a market setting. Image acquisition took place in a closed room, resulting in a total of 30 data sets. This dataset includes images of half-ripe, ripe, and rotten papayas. A smartphone camera was utilized for the image acquisition, with the camera positioned 30 cm away from the subject. The images were taken under room lighting and are of 2 sRGB megapixel quality. Papaya is placed on a white base as a background, then set the phone stand with a shooting distance of approximately 30 cm from the object and a ring light set with 6x brightness, 8 watts. This image data collection was obtained from the right and left sides of the papaya skin. This shooting was carried out in the morning around 8-9 every day. For shooting using the supporting properties of this ring light, it can add lighting to the camera object to be more even, so that the photo results are clearer and more detailed.

Source: (Research Results, 2024)
Figure 2. Ripe Papaya Prediction

b. Data Acquisition

Image data acquisition is the process of sampling and quantifying visual data (images) from the real world and converting it in Image data acquisition is the process of capturing and quantifying visual data (images) from the real world and converting it into a digital format that can be processed by a computer. This process involves the use of sensors, such as digital cameras or scanners, to capture images and transform them into numerical data that can be stored and analyzed. Image data acquisition is the process of capturing and quantifying visual data (images) from the real world and converting it into a digital format that can be processed by a computer. This process involves the use of sensors, such as digital cameras or scanners, to capture images and transform them into numerical data that can be stored and analyzed to a digital form that can be processed by a computer.



Source: (Research Results, 2024)
Figure 3. Image Data Retrieval Techniques

Source: (Research Result, 2024)
Figure 4. Image Data Retrieval Results

This process involves using sensors, such as digital cameras or scanners, to capture images and convert them into numerical data that can be stored and analyzed. From the results of the data image capture, the next step is data preparation. For data processing, namely eliminating excessive light noise, cropping, converting sRGB image data to RGB, and finding the mean RGB value to the dataset. After data collection, data acquisition is carried out, and determining the Region of Interest (ROI) in the image with a size of 600 pixels x 300 pixels. The pixel size is important because it represents the smallest unit of a computerized digital image and plays a crucial role in the further analysis process. Therefore, this size is chosen as a standard to facilitate the next steps in the analysis process (Putra, 2023).

c. Determine ROI

This stage, several processes are carried out including data cleaning, data integration and data reduction so that the results of the process carried out are more accurate and can be accounted for. The Region of Interest (ROI) further, builds on the previous explanation. Determining the ROI effectively is crucial for accurate papaya image analysis because it is the area that can interpret or provide relevant information that needs to be explored further (Mufida, 2022). Here's a more detailed look at the process, considering different scenarios and techniques. Thai Papaya fruit with an ROI of 600 pixels x 300 pixels.

d. Convert sRGB to RGB

sRGB image conversion to RGB is expected to obtain an optimal image for the image processing process. The RGB value of the image obtained will obtain an optimal value, of course, the image processing has been carried out to eliminate noise when taking images for the data set.

$$RGB_{lin} = f(x)\begin{cases} \frac{sRGB}{12,92}, & if\ sRGB \leq 0,04045 \\ \left(\frac{sRGB+0,055}{1,055}\right)^{2,4}, & if\ sRGB \geq 0,04045 \end{cases} \quad (3)$$

e. Calculates the mean RGB value

The middle value of RGB is used to find the most optimal value. Thus, this value is used to find out how big the color of the ripe papaya skin changes from the color of the rotten papaya skin.

f. Determine the level of decay using K-NN

Classification Algorithm The K-Nearest Neighbor Method (K-NN) is a classification algorithm in data mining that is easy and widely used because it works quite easily by taking values at the nearest neighbor (K) distance. The following is an example of calculating the k-nn method with an existing dataset.

Techno Nusa Mandiri: Journal of Computing and Information Technology
Vol. 22, No. 1 March 2025 | DOI https://doi.org/10.33480/techno.v22i1.6522

Table 1. Mean RGB Value From The Thai Papaya Fruit Dataset

No	Data	Value			Category
		R	G	B	
1	Second fruit first day first side	133.412	126.611	75.873	Almost Ripe
2	Second fruit first day second side	143.488	126.318	70.139	Almost Ripe
3	Second fruit second day first side	143.195	123.783	71.006	Ripe
4	Second fruit second day second side	148.464	119.816	66.162	Ripe
5	Second fruit of the sixth day second side	152.227	113.179	74.980	Rotten

Source: (Research Result, 2024)

1. Determine the K value. The K value from the data above is 4.
2. Calculate the distance value using the Euclidean distance formula.

Tabel 2: Calculate the distance value using the euclidean distance

Data	Calculate proses
1	$\sqrt{(133.412-144.174)^2+(126-611-215)^2+(75.873-82.794)^2}$=13.886
2	$\sqrt{(143.488-144.174)^2+(126.318-121.215)^2+(70.139-82.794)^2}$=13.496
3	$\sqrt{(143.195-144.174)^2+(123.783-121.215)^2+(71.006-82.794)^2}$=12.104
4	$\sqrt{(148.464-144.174)^2+(119.816-121.215)^2+(66.162-82.794)^2}$=17.233
5	$\sqrt{(152.227-144.174)^2+(113.179-121.215)^2+(74.980-82.794)^2}$=13.801

Source: (Research Result, 2024)

Table 3. Mean RGB Value from the Thai Papaya fruit dataset after sorting the distance values

No	Data	Value			Category
		R	G	B	
1	Second fruit first day first side	147.469	112.945	74.778	Rotten
2	Second fruit first day second side	143.195	123.783	71.006	Ripe
3	Second fruit second day first side	143.448	126.318	70.139	Almost Ripe
4	Second fruit second day second side	152.227	113.179	74.980	Rotten
5	Second fruit of the sixth day second side	133.412	126.611	75.873	Almost Ripe

Source : (Research Result, 2024)

Based on the Euclidean distance values which have been sorted from smallest, because the K value used is 4 (four) in the closest data in rows 1, 2, 3, and 4, the result of the category that appears most frequently is Rotten, so the category for the new data (Eleventh fruit of the fifth day on the first side) is Rotten.

Table 4: Mean RGB Values of Thai Papaya Fruit Photo Dataset

No	Data To	Value		
		R	G	B
1	First fruit, first day, first side	135.494	125.483	72.672
2	First fruit, first day, second side	129.043	123.054	77.702
3	First fruit, first day, second side	113.252	110.262	69.582
4	The first fruit of the second day of the first day	113.252	110.262	69.582
5	The first fruit of the second day of the second day	113.454	103.611	61.605
6	First fruit, second day, third day	127.739	118.227	68.476
7	First fruit, third day, first side	151.549	121.708	71.010
8	First fruit, third day, second side	137.971	118.117	79.773
9	Buah kesatu hari ketiga sisi ketiga	139.868	118.712	73.595
10	The first fruit of the fourth day of the first day	156.727	122.539	73.398

Source : (Research Result, 2024)

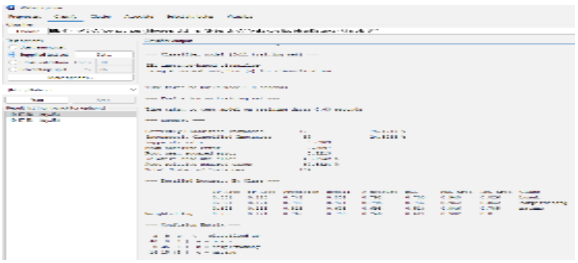
Table 5 : K-NN Model Training Results

No	Data To	Value			Direct Observation	K - NN
		R	G	B		
1	Sampling day one side one	134.161	123.781	72.881	Almost Ripe	Almost Ripe
2	Sampling on the first day and the second day	124.437	119.919	74.233	Almost Ripe	Almost Ripe
3	Sampling on the first day and the third day	125.874	118.644	75.217	Almost Ripe	Almost Ripe

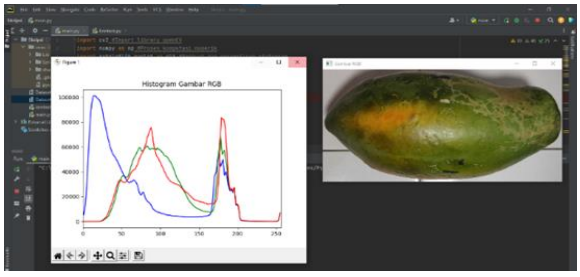
No	Data To	Value			Direct Observation	K - NN
		R	G	B		
4	Sampling on the second day of the first day	145.946	123.538	68.972	Ripe	Ripe
5	Sampling on the second day of the second side	135.610	119.758	70.264	Ripe	Ripe
6	Sampling on the second day and the third day	133.280	116.297	68.587	Ripe	Almost Ripe
7	Sampling on the third day, side one	153.571	127.742	69.714	Ripe	Ripe
8	Sampling on the third day of the second side	141.460	121.042	71.699	Ripe	Ripe
9	Sampling on the third day, third side	133.457	114.407	70.373	Ripe	Ripe
10	Sampling on the fourth day, side one	148.819	123.024	72.794	Ripe	Ripe

Source : (Research Result, 2024)

Detection results from testing this system, there are 3 different data out of 30 existing data. By carrying out trial results on existing sampling data, the percentage accuracy level of this system is 80%, where the 6 existing data correspond to the 30 existing data.

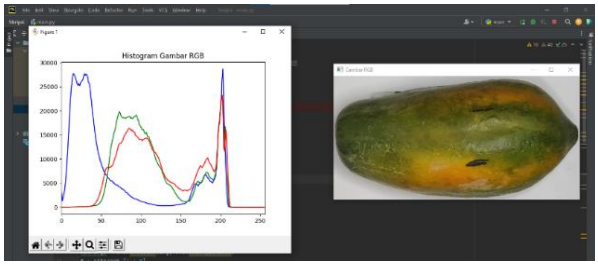


Source: (Research Result, 2024)
Figure 5. Dataset classification results

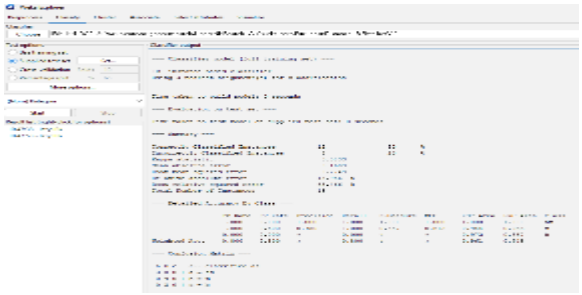


Source: (Research Result, 2024)
Figure 6. The results of the first day's first side third Fruit dataset trial

In addition to using direct comparison on the system that has been created, a comparison is also carried out with the help of the WEKA application to test the accuracy of the existing dataset and sampling data. After testing with the K-NN classification method with a value of k = 4 for the dataset, the percentage of accuracy was obtained at 75%.

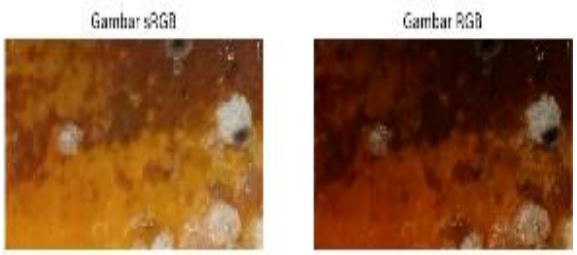


Source: (Research Result, 2024)
Figure 7. The results of the first day's first side third Fruit dataset Sampling

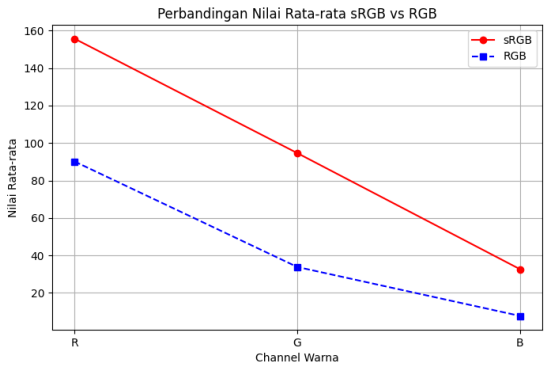


Source: (Research Result, 2024)
Figure 8. Sampling Data Result

Mean while, the percentage for sampling data obtained an accuracy level of 80%.



Source: (Research Result, 2024)
Figure 9. Sampling Data Result sRGB to RGB format image of rotten papaya



Source: (Research Result, 2024)

Figure 10. Sampling Data Result sRGB to RGB grafik format image of rotten papaya

In this study, data testing was conducted on 11 Thai Papaya fruits, 3 sides of which were taken according to the number of natural side dividers that exist and the 3 sides are sufficient to represent all the sides that exist. The shooting itself was carried out until around the fourth to sixth day due to the emergence of different indicators of rottenness levels on each fruit. Indicators of rottenness levels that can be observed directly or appear physically include:

1. Softer texture level.
2. Increasingly brown skin color.
3. Sweeter fruit odor or aroma.

For this research, the indicator of rottenness level used was the condition of the skin color. The total dataset obtained from the 11 existing fruits was 159 images. From each image or image data that exists, it will be processed using the RGB color method. From the results of the tests carried out, an accuracy level of 80% was obtained for testing data sampling where all images had been determined Region of Interest (ROI) with a size of 600 pixels x 300 pixels.

CONCLUSION

In this research, we demonstrate the recognition of early patterns of papaya fruit decay with RGB image processing classified using the K-NN machine learning algorithm. The values used in this study seek the form of Euclidean Distance values on the color of the healthy papaya surface with the color pattern of papaya that is starting to rot. From the data set, we label the image pattern with (ROI) 600 pixels x 300 pixels and 120 data. We get quite significant results when comparing healthy and rotten papaya image patterns. The distribution of the training data set 80% and the testing data set 20%, then the accuracy level is 80% which has a range of values: • Mean R-value = 130.671 - 169.630 • Mean G value = 106.891 - 131.895 • Mean B value = 61.119 - 100.776 • Based on the results of the trial on the sampling data, if there are only 2 (two) values that fall into the RGB value range above, then the fruit will be included in the rotten category, if only 1 (one) is included then it will be included in the almost ripe or ripe category. In conclusion, the study to determine the level of rottenness of papaya fruit with RGB image format and using the Euclidean Distance value using the K-NN method can be used to compare ripe papaya fruit and rotten papaya.

REFERENCE

Afifah, L. (2020, November 23). K-Nearest Neighbor (KNN) *Algorithm* for Classification. Retrieved June 11, 2021, from: https://ilmudatapy.com/algoritma-k-nearest-neighbor-knn-untuk-klasifikasi/

Dihni, V. A. (2022, June 7). Indonesia Enters the List of Countries Producing the Largest Papaya Fruit in the World. Retrieved March 8, 2022, from:https://databoks.katadata.co.id/datapublish/2022/06/07/indonesia-masuk-daftar-negara-penghasil-buah-pepaya-terbesar-di-dunia

Malik Namus Akbar, F. (2024, f February 1). *Metode KNN (K-Nearest Neighbor) untuk Menentukan Kualitas Air*. *18*(1).

Margaretha, J., & Voutama, A. (2023). Perancangan Sistem Informasi Pemesanan Tiket Konser Musik Berbasis Web Menggunakan Unified Modeling Language (UML). *JOINS (Journal of Information System)*, *8*(1), 20–31. https://doi.org/10.33633/joins.v8i1.7107

Mufida, W. (2022). Analisis Letak Tracking Di Area Arteri Karotis Internal Dan Arkus Aorta Pada Pemeriksaan Ct Scan Angiografi (Cta) Serebral Di Instalasi Radiologi Rsud Salatiga. 1.

Oktriwina, A. S. (2021, December 15). What is a Class Diagram and Its Function in Programming. Retrieved November 11, 2021, from: https://glints.com/id/lowongan/class-diagram-adalah/#.Yw4N7HZBzDd

Papaya Content and Benefits for Health and Beauty. (2020, August 10). Retrieved June 19, 2021, from:https://www.chilibeli.com/blog/sayur-dan-buah/kandungan-manfaat-buah-pepaya

Putra, A. S. (2023). Identifikasi Kematangan Buah Jeruk Medan Menggunakan Metode KNN (K-Nearest Neighbor) Berbasis Red, Green, Blue (RGB). 10.

Putra, P., H Pardede, A. M., & Syahputra, S. (2022). ANALISIS METODE K-NEAREST NEIGHBOUR (KNN) DALAM KLASIFIKASI DATA IRIS

BUNGA. *Jurnal Teknik Informatika Kaputama (JTIK)*, *6*(1).

Rabbani, H. A., Rahman, M. A., & Rahayudi, B. (2021). *Perbandingan Ruang Warna RGB dan HSV dalam Klasifikasi Kematangan Biji Kopi* (Vol. 5, Issue 6). http://j-ptiik.ub.ac.id

Rezkia, S. M. (2021, July 13). Types of Machine Learning Classification Algorithms That Are Important to Know. Retrieved May 15, 2022, from:https://www.dqlab.id/macam-algoritma-klasifikasi-machine-learning-yang-penting-untuk-diketahui

Rivan, Muhammad Ezar Al, & Suherman, Jessica (2020, October 1). Determination of California Papaya Fruit Quality (Carica Papaya L.) Using Fuzzy Mamdani. Retrieved April 12, 2022, from: https://media.neliti.com/media/publications/357698-penentuan-mutu-buah-pepaya-california-ca-01d5a036.pdf

Setiawan, R. (2021, October 30). What is Data Mining and How is the Method? Retrieved May 10, 2022, from:https://www.dicoding.com/blog/apa-itu-data-mining

What is PyCharm? Getting to Know PyCharm *Python*, Its Disadvantages, and Advantages. (2021, May 24). Retrieved February 20, 2022, from:https://appkey.id/pembuatan-aplikasi/mobile-programming/pycharm-python/

Wibowo, P. T. (2021, November 9). What is UML? Retrieved April 7, 2022 from: https://wartaekonomi.co.id/read373179/apa-itu-uml