

Watermelon Classification

by Febri Liantoni Rifki Indra

Submission date: 20-Oct-2022 05:10PM (UTC+0700)

Submission ID: 1930459938

File name: Liantoni_2019_J._Phys._Conf._Ser._1175_012114.pdf (887.95K)

Word count: 2472

Character count: 13808

1

PAPER · OPEN ACCESS

4

Watermelon classification using k-nearest neighbours based on first order statistics extraction

1

To cite this article: Febri Liantoni *et al* 2019 *J. Phys.: Conf. Ser.* **1175** 012114

View the [article online](#) for updates and enhancements.



IOP | ebooks™

Bringing you innovative digital publishing with leading voices to create your essential collection of books in STEM research.

Start exploring the collection - download the first chapter of every title for free.

Watermelon classification using k-nearest neighbours based on first order statistics extraction

Febri Liantoni¹, Rifki Indra Perwira², Lusi Dwi Putri³, Rosida Tiurma Manurung⁴, Muhammad Syahrul Kahar⁵, Jehan Safitri⁶, Lusy Tunik Muharlisiani⁷, Dina Chamidah⁷, Abd. Ghofur⁸, Putu Sukma Kurniawan⁹, E. Wityasminingsih¹⁰, Yuniningsih¹¹, Susilo¹², Muhammad Yusuf¹³, Muh Ardian Kurniawan¹⁴, Sumitro¹⁵, Bayu Prabowo Sutjiatmo¹⁶, Nalal Muna¹⁶, Fajria Fatmasari¹⁶, Saroni Widodo¹⁷, Andi Reni Syamsudin¹⁸, Achmad Daeng GS¹⁹, Dudung Ma'ruf Nuris²⁰, Ratna Dyah Suryaratri²¹, Tutut Suryaningsih²², Indrya Mulyaningsih²³, Tari Nawas

¹Information Technology, Adhi Tama Institute of Technology Surabaya, Indonesia

²UPN "Veteran" Yogyakarta

³Universitas Lancang Kuning, Indonesia

⁴Universitas Kristen Maranatha

⁵Universitas Muhammadiyah Sorong, Indonesia

⁶Universitas Lambung Mangkurat Indonesia

⁷Universitas Wijaya Kusuma Surabaya Indonesia

⁸STKIP PGRI Lamongan, Indonesia

⁹Universitas Pendidikan Ganesha, Indonesia

¹⁰Politeknik Piksi Ganesha

¹¹Universitas Pembangunan Nasional "Veteran" Jawa Timur Indonesia

¹²Universitas Negeri Semarang Indonesia

¹³Universitas Malikussaleh, Indonesia

¹⁴Universitas Hamzanwadi, Indonesia

¹⁵Universitas Dayanu Ikhsanudin Indonesia

¹⁶Politeknik APP Jakarta Indonesia

¹⁷Politeknik Negeri Semarang Indonesia

¹⁸Universitas Hasanuddin Makassar

¹⁹Universitas 45 Surabaya, Indonesia

²⁰Universitas Negeri Malang, Indonesia

²¹Universitas Negeri Jakarta, Indonesia

²²STKIP PGRI Tulungagung, Indonesia

²³IAIN Syekh Nurjati Cirebon, Indonesia

nuning.kurniasih@unpad.ac.id

Abstract. Watermelon (*Citrullus Lanatus Tunb / Citrullus Vulgaris Schrad*) is a vine that grows rapidly in Indonesia. Watermelon has many benefits that make watermelon one of the fruits sought by the community. The maturity of watermelon can be distinguished based on watermelon skin texture. The similarity in watermelon skin texture causes people difficulty in



identifying the level of maturity of mature and immature watermelon. Based on these problems, research was carried out for the classification of watermelon maturity based on watermelon skin texture. The first-order statistical feature extraction method is used as a method to recognize watermelon maturity in terms of fruit skin texture in the classification process. The first order statistical characteristic parameters used are mean, variance, skewness, and kurtosis. Based on the extraction value of these statistical features, then used as a reference classification process using the k-nearest neighbour method. The data in this study are 100 images, of which 70 are for training data and 30 for testing data. From the results of testing the maturity classification of watermelon using k-nearest neighbour based on fruit skin texture based on extraction of first-order features, 86.66% accuracy was obtained.

1. Introduction

Watermelon has many benefits that make watermelon one of the fruits sought by the community. The similarity of watermelon skin texture that is ripe with immature fruit makes it difficult for people to identify watermelons. Because humans are subjective in determining the maturity of watermelon fruit, causing an assessment of the maturity of watermelon fruit is different from one person to another. The process of classification of watermelon maturity can be done by identifying the image of watermelon skin texture.

The process of classification of watermelon maturity can be done by identifying watermelon skin texture images. In this way, steps can be taken to identify skin texture patterns by recognizing the structural characteristics of the skin such as watermelon shape and color using histogram equalization. Technological developments for image processing techniques are also growing rapidly. Various techniques were developed to facilitate human work, both as image processors, image analysts and image users for various purposes and needs [1]. Often the image used is not in an ideal condition to be studied due to the many disturbances, it can be in the form of shadows, photos or blurred images, lack of clarity of appearance of objects so that it can cause problems and affect the results of interpellation and will affect the analysis and planning that will be carried out. image processing to obtain the ideal image.

This digital image processing technique is carried out at the pre-processing stage of the image until the structural shapes and characteristics of each watermelon skin are obtained. Histogram equalization is used in the preprocessing stage before feature extraction. The method used in the extraction of this feature is based on first-order statistical characteristics. After that, feature extraction from the watermelon skin image is then used as a watermelon maturity classification data using k-nearest neighbour.

Based on previous research, Novita in 2017 used first-order feature extraction implemented in the classification of batik with backpropagation algorithm. In this study, the parameter values mean, skewness, kurtosis, and entropy are used as the basis for processing the classification process [2]. Guan in 2017 explored the value of diffusion in the apparent diffusion coefficient (ADC). There was a first order statistical feature in the clinical introduction of cervical cancer [3].

Histogram Equalization is a method to improve image quality by changing the distribution of gray levels of images [4]. This is intended so that the gray level distribution is more evenly distributed compared to the original image. Histogram equalization is also used to increase image color contrast. Several studies on contrast enhancement using histogram equalization were carried out by Celik, Chen, and Wang in a previous study. Celik makes image improvements with two-dimensional histogram equalization that effectively improves the quality of different image types [5]. Chen developed a histogram equalization model for image contrast enhancement. The approach is carried out with a variational model containing functional energy to adjust pixel values from an input image directly so that the resulting histogram can be redistributed to be more uniform [6]. Wang also carried out contract improvement techniques based on the local histogram equalization algorithm. The proposal is done by segmenting the image into several sub-blocks using gradient values, this algorithm successfully increases local contrast without adding noise to the image [7].

Research on classification of seeded and seedless watermelons using k-nearest neighbours was carried out by [8]. In the previous study classification of mango plants was carried out using the k-nearest neighbour method and support vector machine based on leaf bone structure [9]. In this research, the maturity classification of watermelon fruit is based on the extraction of first-order statistical features from fruit skin texture using k-nearest neighbour.

2. Methods

In this research, several stages will be carried out including preprocessing, extraction of first-order statistical features, classification using k-nearest neighbour. The stages of the research were carried out as shown in Figure 1.

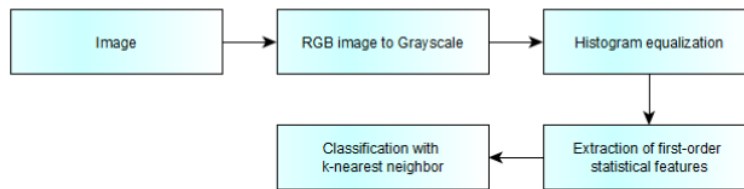


Figure 1. Stage of the research

2.1. Preprocessing

The initial process that is often done in image processing is changing the color image to grayscale. This is done to simplify the image model. In a true color (RGB) image there are three matrix layers, namely R-layer, G-layer, B-layer.

2.2. Extraction of first-order statistical features

First-order statistical feature extraction is a feature retrieval method based on the characteristics of the image histogram by ignoring the relationship between neighbouring pixels. First-order statistical feature extraction by calculating color features from watermelon skin texture images. Color features can be obtained through statistical calculations such as mean, variance, skewness, and kurtosis.

Mean is a measure of image dispersion regarding the color distribution of an image. Variance expressed widely spread distribution. The square of the variance square is called the standard deviation. Variance shows the variation of elements in an image. Skewness or inclination expresses a measure of asymmetry. The distribution is said to be left-leaning if it has a skewness value in the form of a negative number. Conversely, the distribution is said to be leaning right if it has a skewness value in the form of a positive number. Kurtosis is a measure that shows the distribution of data that is tapered or collected.

2.3. K-Nearest Neighbour

The K-Nearest Neighbour method is one of the nearest neighbour (NN)-based methods that can be used to compare values from histograms [10]. The K value used here states the number of closest neighbours involved in determining the prediction of class labels on the test data. From K the closest neighbour chosen is then voting class from the nearest neighbour. The class with the highest number of neighbouring sounds is given as a prediction class label on the test data. K-Nearest Neighbour is one of the methods used in classification.

The working principle of K-Nearest Neighbour is to find the closest distance between the data that will be evaluated with the closest neighbouring K in the training data.

3. Results and Discussion

The test was carried out with a number of 100 images, of which 70 were for training data and 30 for testing data. Examples of watermelon skin images are shown in Figure 1.

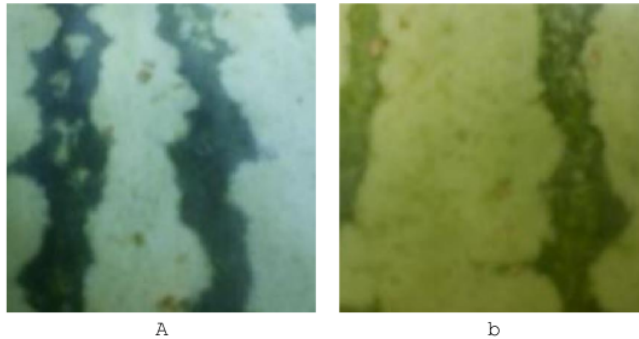


Figure 2. Watermelon skin a) raw fruit b) ripe fruit

The initial stage in this study is to change the RGB image to grayscale. Based on the grayscale image, the histogram equalization process is then carried out. The histogram equalization image results as shown in Figure 2.

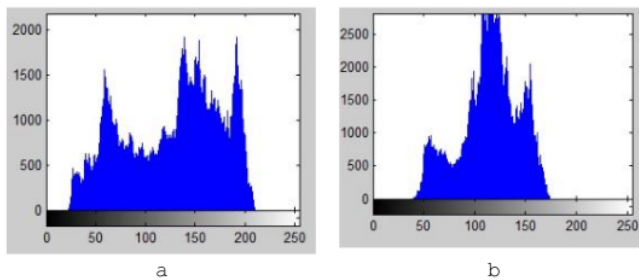


Figure 3. Histogram equalization a) raw fruit b) ripe fruit

First-order statistical feature extraction includes mean values, variances, skewness, and kurtosis. Some examples of the results of the calculation of first-order statistical feature extraction are shown in Table 1.

Table 1. The calculation of first-order statistical feature extraction

Data	Mean	Variance	Skewness	Kurtosis
raw fruit 1	4700	2520191123	-0.56743	-2.92653
raw fruit 2	2355	2162546617	-0.15793	-2.91873
raw fruit 3	1353	1345353445	-0.23446	-2.96454
raw fruit 4	3284	1774514534	-0.49763	-2.90486
raw fruit 5	4354	1985783475	-0.48546	-2.92848
ripe fruit 1	1261	59763	-0.80735	-2.99893
ripe fruit 2	307	158984	-0.70742	-2.99325
ripe fruit 3	875	194334	-0.59623	-2.91546
ripe fruit 4	1112	67488	-0.84533	-2.99877
ripe fruit 5	1174	58453	-0.87463	-2.99856

Based on the testing of the classification process using the value of $K = 3$ as shown in Table 2.

Table 2. The calculation of first-order statistical feature extraction

Test Data	Class		Result	Test Data	Class		Result
	Groundtruth	System			Groundtruth	System	
1	raw fruit	raw fruit	TRUE	16	ripe fruit	ripe fruit	TRUE
2	raw fruit	raw fruit	TRUE	17	ripe fruit	ripe fruit	TRUE
3	raw fruit	ripe fruit	FALSE	18	ripe fruit	ripe fruit	TRUE
4	raw fruit	raw fruit	TRUE	19	ripe fruit	ripe fruit	TRUE
5	raw fruit	raw fruit	TRUE	20	ripe fruit	ripe fruit	TRUE
6	raw fruit	raw fruit	TRUE	21	ripe fruit	ripe fruit	TRUE
7	raw fruit	raw fruit	TRUE	22	ripe fruit	ripe fruit	TRUE
8	raw fruit	ripe fruit	FALSE	23	ripe fruit	ripe fruit	TRUE
9	raw fruit	raw fruit	TRUE	24	ripe fruit	ripe fruit	TRUE
10	raw fruit	raw fruit	TRUE	25	ripe fruit	ripe fruit	TRUE
11	raw fruit	ripe fruit	FALSE	26	ripe fruit	raw fruit	FALSE
12	raw fruit	raw fruit	TRUE	27	ripe fruit	ripe fruit	TRUE
13	raw fruit	raw fruit	TRUE	28	ripe fruit	ripe fruit	TRUE
14	raw fruit	raw fruit	TRUE	29	ripe fruit	ripe fruit	TRUE
15	raw fruit	raw fruit	TRUE	30	ripe fruit	ripe fruit	TRUE

Table 2 shows that the correct results represent the corresponding values based on the actual class and the system, while the wrong values represent the inappropriate class between the actual class and the system class. Based on the results of testing using the k-nearest neighbour method of 30 test images, 26 images were correctly classified according to the actual class. From the results of this test, the system accuracy is $26/30 = 86.66\%$

4. Conclusion

Based on the results of testing carried out for the watermelon maturity classification process based on first-order statistical characteristics using k-nearest neighbour, the system accuracy was 86.66%.

References

- [1] A. Rocha, D. C. Hauage, J. Wainer, and S. Goldenstein, "Automatic fruit and vegetable classification from images," *Computers and Electronics in Agriculture*, vol. 70, no. 1, pp. 96–104, Jan. 2010.
- [2] N. Ningrum, K. Defri, and H. Novi, "Penerapan Ekstraksi Ciri Orde Satu Untuk Klasifikasi Tekstur Motif Batik Pesisir Dengan Algoritma Backpropagasi," *Simetris : Jurnal Teknik Mesin, Elektro dan Ilmu Komputer*, vol. 8, no. 2, 2017.
- [3] Y. Guan *et al.*, "Value of whole-lesion apparent diffusion coefficient (ADC) first-order statistics and texture features in clinical staging of cervical cancers," *Clinical Radiology*, vol. 72, no. 11, pp. 951–958, Nov. 2017.
- [4] N. S. P. Kong and H. Ibrahim, "Multiple layers block overlapped histogram equalization for local content emphasis," *Computers & Electrical Engineering*, vol. 37, no. 5, pp. 631–643, Sep. 2011.
- [5] T. Celik, "Two-dimensional histogram equalization and contrast enhancement," *Pattern Recognition*, vol. 45, no. 10, pp. 3810–3824, Oct. 2012.
- [6] C. Chen and M. K. Ng, "An image pixel based variational model for histogram equalization," *Journal of Visual Communication and Image Representation*, vol. 34, pp. 118–134, Jan. 2016.
- [7] Y. Wang and Z. Pan, "Image contrast enhancement using adjacent-blocks-based modification for local histogram equalization," *Infrared Physics & Technology*, vol. 86, pp. 59–65, Nov. 2017.
- [8] M. Mohd Ali, N. Hashim, S. K. Bejo, and R. Shamsudin, "Laser-induced backscattering imaging for classification of seeded and seedless watermelons," *Computers and Electronics in Agriculture*, vol. 140, pp. 311–316, Aug. 2017.

- [9] F. Liantoni and L. Agus Hermanto, "Adaptive Ant Colony Optimization on Mango Classification Using K-Nearest Neighbour and Support Vector Machine," *Journal of Information Systems Engineering and Business Intelligence*, vol. 3, p. 75, Oct. 2017.
- [10] F. S. Mohamad, A. A. Manaf, and S. Chuprat, "Nearest Neighbour For Histogram-based Feature Extraction," *Procedia Computer Science*, vol. 4, pp. 1296–1305, Jan. 2011.

Watermelon Classification

ORIGINALITY REPORT

15%

SIMILARITY INDEX

%

INTERNET SOURCES

%

PUBLICATIONS

15%

STUDENT PAPERS

PRIMARY SOURCES

1	Submitted to Politeknik Negeri Bandung Student Paper	8%
2	Submitted to Universitas Brawijaya Student Paper	2%
3	Submitted to iGroup Student Paper	1%
4	Submitted to Universitas Negeri Surabaya The State University of Surabaya Student Paper	1%
5	Submitted to Universitas Atma Jaya Yogyakarta Student Paper	1%
6	Submitted to Teaching and Learning with Technology Student Paper	1%
7	Submitted to University College London Student Paper	1%

Exclude quotes On

Exclude matches Off

Exclude bibliography On