Surat Keterangan Tidak Plagiat [Kepangkatan]

Nomor: 766.20/II.3.AU/14.00/C/KET/I/2023

Kepada Yth :
Ibu Syamsudduha Syahrorini, ST. MT., Dr.
Di
       Tempat

Assalamua'alaikum Wr. Wb.
Sehubungan dengan adanya permohonan Surat Keterangan Tidak Plagiat dengan rincian:

Judul Artikel	: K-Nearest Neighbor Algorithm to Identify Cucumber Maturity with Extraction of One-Order Statistical Features and Gray-Level Co-Occurrence
Nama Pemohon	: Syamsudduha Syahrorini / TEKNIK ELEKTRO
URL Sinta Pemohon	: https://sinta.kemdikbud.go.id/authors/profile/5993144
Nama Penulis	: Syamsudduha Syahrorini
Tujuan	: Kepangkatan
Tujuan Kepangkatan	: Lektor

Naskah Yang Dimohonkan pengecekan:
https://dosen.umsida.ac.id/modul/publikasi/filesktp/970137/sktp-04-01-2023%2003:43:03-970137.pdf

Artikel tersebut DAPAT digunakan untuk proses kepangkatan.

      Demikian surat keterangan ini kami sampaikan, mohon untuk digunakan sebagaimana mestinya.

Wassalamu'alaikum Wr. Wb.

Mengetahui,
Wakil Rektor 1
Universitas Muhammadiyah Sidoarjo

Hana Catur Wahyuni, ST., MT

Direktur DRPM
Universitas Muhammadiyah Sidoarjo

Dr. Sigit Hermawan, S.E., M.Si

sktp-04-01-2023 03_43_03-970137

by Syamsudduha Syahrorini, St. Mt., Dr.

PAPER · OPEN ACCESS

K-Nearest Neighbor Algorithm to Identify Cucumber Maturity with Extraction of One-Order Statistical Features and Gray-Level Co-Occurrence

View the article online for updates and enhancements.

K-Nearest Neighbor Algorithm to Identify Cucumber Maturity with Extraction of One-Order Statistical Features and Gray-Level Co-Occurrence

Syamsudduha Syahrorini[*], Denny Syamsudin, Dwi Hadidjaja Rasjid Saputra, Akhmad Ahfas

Muhammadiyah University of Sidoarjo, Sidoarjo, Indonesia

Abstract. Determination of the maturity of cucumber fruit after harvest is subjective. The level of thoroughness of each individual's cucumber selection is different. The maturity of cucumbers is seen from the age of the fruit, the resemblance of ripe cucumbers or old, raw or young is difficult to distinguish in terms of the texture of the fruit skin. The use of red, green, blue, and grayscale color modes from the cucumber imagery, is then processed using the extraction of the statistical features of the Order-One and Order-Two GLCM methods. Imagery using a 13 megapixels smartphone camera. Texture parameter values are used in the first order: mean, variance, skewness, kurtosis, entropy. Second order: energy, contrast, correlation, inverse different moments, angular second moments, variance 2, and entropy 2. The classification of the texture parameters of both order uses an algorithm from K-Nearest Neighbors as a comparison of test data values and training data. So that the system is made can identify the maturity of cucumbers old and young. The highest accuracy is found in grayscale imagery with a combination of two skewness and kurtosis parameters and euclidean distance calculation in K-Nearest Neighbor classification of 96.05%.

1. Introduction

Cucumbers, cucumbers, or cucumbers are the fruit of the Cucumis sativus L class derived from pumpkin plant species that are widespread in Indonesia. Cucumbers can be consumed in fresh form as well as processed foods such as vegetables, pickles, hodgepodge, salted, and fresh vegetables as a complement to food. There are three types of cucumbers sold in the market, namely ordinary cucumbers, watang, and wuku. For cucumbers it is usually characterized by the appearance of thin-looking fruit skins. There are bintil-bintil on the texture of the fruit skin, as well as whitish green, light green, and dark green In addition to the maturity of cucumbers by farmers so far still based on the age of the fruit. All plants that have reached the planting period for ±75-85 days or the life of the fruit 30-50 HST will be picked in its entirety. In order for the determination of the maturity of cucumber fruit is not subjective it is necessary to have automatic selection by using the help of digital image analysis. To determine raw cucumbers and ripe cucumbers from the texture of the outer skin of the fruit used a method of extraction of statistical characteristics of the order one and order two. Statistical feature extraction method for the first order with mean parameters, variance, skewness, kurtosis, and entrophy based on histogram characteristics of the imagery and second order of gray-level co-occurrence calculation with angular second moment, variance 2, correlation, contrast, inverse difference moment, and entropy 2 parameters obtained from the calculation of distance probability and orientation

between two images. To classify the values of each parameter is used by the K-Nearest Neighbor method as a gauge of the level of similarity between the two image objects [5][6][7][8][9][10][11][12][13][14][15][16]. The purpose of this study is to understand and know how to classify the maturity of cucumbers using the characteristic extraction method with the parameters of the order one and order of two GLCM based on the value of 11 parameters of each image, as well as identify young or old cucumbers with the K-Nearest Nieghbor method based on the distance of euclidean, cityblock, cosine from the value of K=1, K=2, and K=3 per nearest neighbor [17][18][19][20].

2.   Material And Methods

Implementation of Statistical Characteristic Extraction of Order One and Gray-Level Co-Occurrence To Identify The Maturity of Cucumbers With K-Nearest Neighbor Algorithm through several stages performed. First the extension of cucumber image input *.bmp or *.jpeg is converted from RGB format imagery to red, green, blue, grayscale format. After the imagery with 4 color modes is obtained then the histogram value of each image is extracted using the method of extracting statistical characteristics of the order one and two to be known the characteristic parameters of the image. Classification is done using the K-Nearest Neighbor method with the calculation of the proximity of euclidean distance, Cityblock Distance (Manhattan Distance), and Cosine, shown in Figure 1.
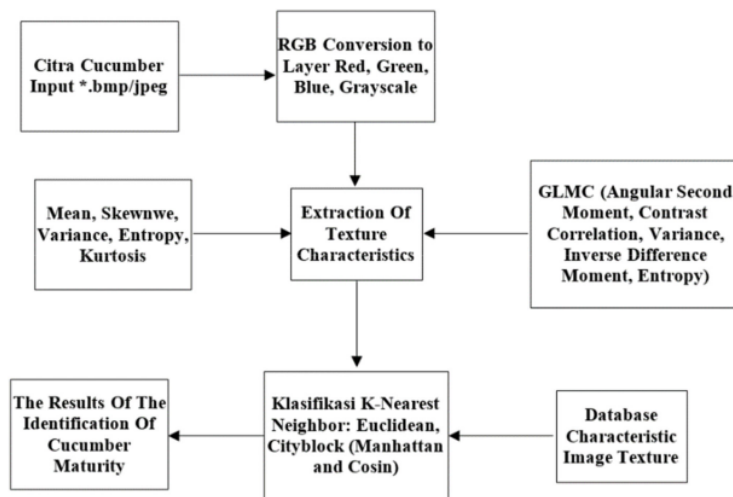
Fig. 1.   System Diagram Blocks

3.   Results and Discussion

3.1. Software Testing

First the cucumber test image is uploaded by pressing Buton 'Browse Image' to then display on the first view with the file name of the test image 'Young(16).jpg'. In this example, we used the 16th image test data of young cucumbers.

Once the image has been successfully input and has been in the next crop the image will be extracted based on the color mode to be tested as well as the feature parameters to be tested. for example Figure 2. is an analysis of the image of cucumbers with Red color mode and the characteristic parameter to be tested is the mean of the order one.

Fig. 2.   Input Image of Young Cucumber and Cropping Image of Young Cucumber

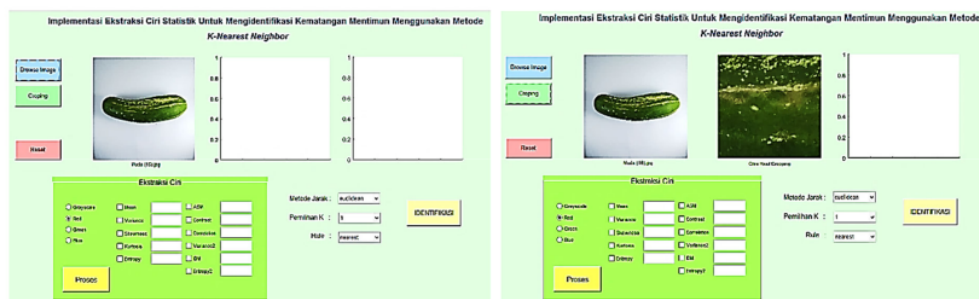Figure 3 of the classification process begins by first specifying which distance method and what value K = you want to specify. For the use of the distance method as well as the value of k by default is K = 1 and the Euclidean Distance method. The resulting identification in the form of a display, in this example, the red image of the young cucumber is in the QUALITY 1 classification, which means the classification of the red image using the mean feature parameter and the Euclidean distance method with the value K = 1 give the correct result.
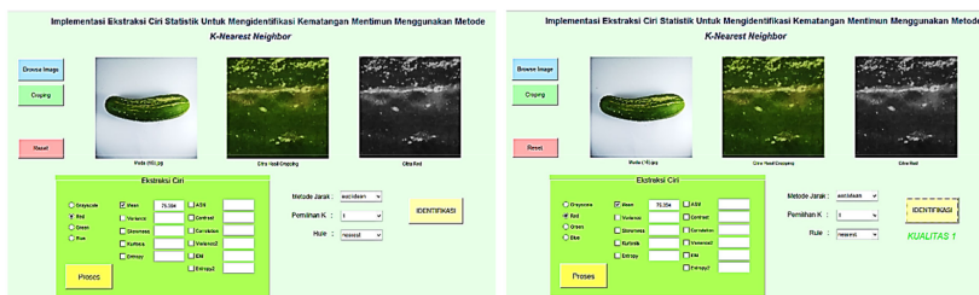

Fig. 3.   Extraction Of One Mean Order Characteristic Parameters and Red Cucumber Image Classification Results With Mean Feature Parameters

3.2. Analysis of Statistical Characteristics of Order One and Two

A feature extraction analysis is performed to obtain the feature value of each parameter as a whole test data against the database. The terms and conditions required are as follows:

1. Total use of cucumber imagery in training as many as 40 imagery. Among them are 20 images of young cucumbers and 20 images of old cucumbers. Meanwhile, the total imagery for the test was 78 images. Among them are 38 images of young cucumbers and 38 images of old cucumbers.
2. The color mode on the image used is red, green, blue, and grayscale color modes.
3. For the extraction of statistical characteristics of the order one characteristic parameter used is mean, skewness, kurtosis, variance, entropy.
4. For the extraction of statistical features of the order two characteristic parameters used are angular second moment, homogeneity/inverse different moment, contrast, correlation, variance 2, and entropy 2.
5. The value of parameter K used in the classification of K-Nearest Neighbor in the extraction of statistical features of order one and two is 1, and the calculation of distance by default is Euclidean Distance.
6. Feature test results that have more than 70% match accuracy will be selected for each other combined.

3.3. Red Image Statistical Feature Extraction Results

Table 1 displays the results of red image extraction that statistical features that have more than 70% accuracy as much as 5 features. In the statistical characteristics of the order one is mean,

skewness, kurtosis, whereas in the statistical characteristics of the order two are correlation and variance 2.

3.4. Green Image Statistical Feature Extraction Results

Table 2 shows the results of geen image extraction that the statistical characteristics of order one and two that have more than 70% accuracy are equal to red imagery which is 5 features, respectively. Feature parameters are also the same as Red imagery consisting of mean, skewness, kurtosis, correlation, and variance 2.

3.5. Blue Image Statistical Feature Extraction Results

Table 3 of blue image extraction results that have more than 70% accuracy as many as 5 features the same as Red and Green imagery. The features in question are mean, skewness, kurtosis, correlation, and variance 2.

3.6. Extraction of Grayscale Image Statistical Features

Table 4 obtained the extraction result of grayscale imagery that has more than 70% accuracy is 5 features. The features in question are mean, skewness, kurtosis, correlation, variance 2. The feature parameters obtained are the same as the previous imagery, Red, Green, and Blue imagery.

Table 1. Red Image Extraction Results

DATA RED				
Ekstraksi Ciri	Jumlah Data Benar		Total	Akurasi %
	Muda	Tua		
Ciri Statistik Orde 1				
Mean	37	25	62	81,6%
Variance	26	20	46	60,5%
Skewness	36	37	73	96,1%
Kurtosis	33	32	65	85,5%
Entropy	17	20	37	48,7%
Ciri Statistik Orde 2				
Angular Second Moment	20	19	39	51,3%
Contrast	24	17	41	53,9%
Correlation	37	25	62	81,6%
Variance 2	37	25	62	81,6%
Inverst Diferend Moment	26	14	40	52,6%
Entropy 2	24	26	50	65,8%

Table 2. Green Image Extraction Results

DATA GREEN				
Ekstraksi Ciri	Jumlah Data Benar		Total	Akurasi %
	Muda	Tua		
Ciri Statistik Orde 1				
Mean	36	27	63	82,9%
Variance	27	19	46	60,5%
Skewness	36	36	72	94,7%
Kurtosis	33	30	63	82,9%
Entropy	22	19	41	53,9%
Ciri Statistik Orde 2				
Angular Second Moment	24	20	44	57,9%
Contrast	27	25	52	68,4%
Correlation	36	27	63	82,9%
Variance 2	36	27	63	82,9%
Inverst Diferend Moment	26	13	39	51,3%
Entropy 2	20	25	45	59,2%

Table 3. Blue Image Extraction Results

DATA BLUE				
Ekstraksi Ciri	Jumlah Data Benar		Total	Akurasi %
	Muda	Tua		
Ciri Statistik Orde 1				
Mean	27	27	54	71,1%
Variance	15	12	27	35,5%
Skewness	26	35	61	80,3%
Kurtosis	31	35	66	86,8%
Entropy	19	24	43	56,6%
Ciri Statistik Orde 2				
Angular Second Moment	28	25	53	69,7%
Contrast	23	28	51	67,1%
Correlation	24	30	54	71,1%
Variance 2	30	28	58	76,3%
Inverst Diferend Moment	25	14	39	51,3%
Entropy 2	25	18	43	56,6%

Table 4. Grayscale Image Extraction Results

DATA GREY				
Ekstraksi Ciri	Jumlah Data Benar		Total	Akurasi % %
	Muda	Tua		
Ciri Statistik Orde 1				
Mean	37	26	63	82,9%
Variance	30	15	45	59,2%
Skewness	36	36	72	94,7%
Kurtosis	32	32	64	84,2%
Entropy	20	22	42	55,3%
Ciri Statistik Orde 2				
Angular Second Moment	23	22	45	59,2%
Contrast	26	27	53	69,7%
Correlation	36	25	61	80,3%
Variance 2	36	26	62	81,6%
Inverst Diferend Moment	26	17	43	56,6%
Entropy 2	17	24	41	53,9%

3.7. Accuracy of Red, Green, Blue, and Grayscale Color Imagery

Figure 4 is the highest accuracy value in an image with grayscale color mode in 2 combinations of 94.74% feature parameters expressed in graphic form. The highest accuracy is found in the combination of the skewness parameter and the kurtosis characteristic parameter is the result of the extraction of the statistical characteristic of the order.

Table 5. It is classified using euclidean distance and cityblock distance methods, because both methods are included in the closest neighbor measurement method of geomatric values between the database vector (training data) and the test data vector, while cosine distance's closest neighbor measurement method is based on the vector set between the database vector and the test data vector.

From the results of experiments and analysis of the cucumber maturity classification system based on the texture of cucumber skin by extracting statistical characteristics of order one and two from the image.

4. Conclusion

From the results of experiments and analysis of cucumber maturity classification system based on the texture of cucumber skin by extracting statistical characteristics of order one and two from imagery and classification with K-Nearest Neighbor, several conclusions are obtained among others: The method of extraction of statistical characteristics of order one and order two using gray-level co-occurrence algorithm approach can be used to identify the maturity of cucumbers with high accuracy accuracy of 96.05%. The parameter of one order feature with a combination of skewness and kurtosis results in the highest accuracy accuracy in grayscale imagery at 94.74%. The grayscale color mode in cucumber imagery has the highest accuracy value of 96.05% because the color in the grayscale image of the texture feature is clearer for the extraction process of the value of each feature parameter. Green imagery with a color structure similar to a cucumber cannot be used as a reference for the classification process, because in the analysis of cucumber skin texture the rate of extraction of green imagery characteristics against the system database is lower than grayscale imagery. K-Nearest Neighbor classification with Euclidean Distance distance approach can produce output as expected which is to have accuracy of 96.05% at the time of K=3.
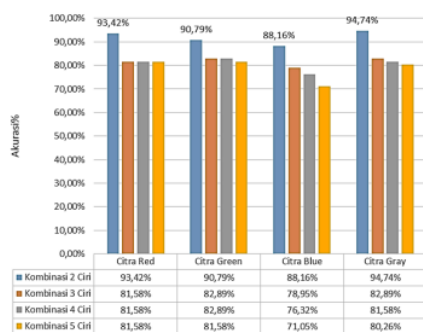
Fig. 4.   Highest Image Feature Parameter Accuracy Graph

Table 5. K-Nearest Neighbor classification accuracy analysis with Euclidean, Cityblock, and Cosine distance methods

Metode Jarak		Jumlah Data Benar		Total	Akurasi %
		Muda	Tua		
Euclidean Distance	K=1	36	36	72	94,74%
	K=2	35	37	72	94,74%
	K=3	36	37	73	96,05%
Cityblock Distance	K=1	36	35	71	93,42%
	K=2	35	37	72	94,74%
	K=3	36	36	72	94,74%
Cosine Distance	K=1	36	31	67	88,16%
	K=2	36	31	67	88,16%
	K=3	36	30	66	86,84%

5. Acknowledgement

6. References

[1]   R. Rukmana, *Budi Daya Mentimun*. Yogyakarta Kanisius, 1994.

[2]   U. Sumpena, *BUDI DAYA MENTIMUN INTENSIF DENGAN MULSA, SECARA TUMPANG GILIR*. 2007.

[3]   F. H. N. Foong, A. Mohammad, and S. J. A. Ichwan, "Biological properties of cucumber (Cucumis sativus L.) extracts," *Malaysian J. Anal. Sci.*, vol. 19, no. 6, pp. 1218–1222, 2015.

[4]   E. K. Eifediyi, S. U. Remison, H. E. Ahamefule, K. O. Azeez, and P. O. Fesobi, "Performance of watermelon (Citrullus lanatus L.) in response to organic and NPK fertilizers," *Acta Univ. Sapientiae, Agric. Environ.*, vol. 9, no. 1, pp. 5–17, 2018, doi: 10.1515/ausae-2017-0001.

[5]   G. Dhingra, V. Kumar, and H. D. Joshi, "Study of digital image processing techniques for leaf disease detection and classification," *Multimed. Tools Appl.*, vol. 77, no. 15, pp. 19951–20000, 2018, doi: 10.1007/s11042-017-5445-8.

[6]   H. Sabrol and S. Kumar, "Recent Studies of Image and Soft Computing Techniques for Plant Disease Recognition and Classification," *Int. J. Comput. Appl.*, vol. 126, no. 1, pp. 44–55, 2015, doi: 10.5120/ijca2015905982.

[7]   J. Siswantoro, H. Arwoko, and M. Widiasri, "Image Based Indonesian Fruit Recognition using

MPEG-7 Color Structure Descriptor and k-Nearest Neighbor," *Int. Conf. Informatics, Technol. Eng. 2019 (InCITE 2019)*, no. August, p. D-84-D-90, 2019.

[8]   V. Sharma, A. A. Mir, and D. A. Sarwr, "Detection of Rice Disease Using Bayes' Classifier and Minimum Distance Classifier," *J. Multimed. Inf. Syst.*, vol. 7, no. 1, pp. 17–24, 2020, doi: 10.33851/jmis.2020.7.1.17.

[9]   M. A. Khan *et al.*, "An implementation of normal distribution based segmentation and entropy controlled features selection for skin lesion detection and classification," *BMC Cancer*, vol. 18, no. 1, pp. 1–20, 2018, doi: 10.1186/s12885-018-4465-8.

[10]   P. K. Sethy, B. Negi, S. K. Behera, N. K. Barpanda, and A. K. Rath, "An Image Processing Approach for Detection, Quantification, and Identification of Plant Leaf Diseases -A Review," *Int. J. Eng. Technol.*, vol. 9, no. 2, pp. 635–648, 2017, doi: 10.21817/ijet/2017/v9i2/170902059.

[11]   R. Widodo, A. W. Widodo, and A. Supriyanto, "Pemanfaatan Ciri Gray Level Co-Occurrence Matrix (GLCM) Citra Buah Jeruk Keprok (Citrus reticulata Blanco) untuk Klasifikasi Mutu," *J. Pengemb. Teknol. Inf. dan Ilmu Komput.*, vol. 2, no. 11, pp. 5769–5776, 2018.

[12]   G. K. Sandhu and R. Kaur, "Plant Disease Detection Techniques: A Review," *2019 Int. Conf. Autom. Comput. Technol. Manag. ICACTM 2019*, vol. 8, no. 4, pp. 34–38, 2019, doi: 10.1109/ICACTM.2019.8776827.

[13]   S. R. Dubey and A. S. Jalal, "Robust approach for fruit and vegetable classification," *Procedia Engineering*, vol. 38. pp. 3449–3453, 2012, doi: 10.1016/j.proeng.2012.06.398.

[14]   A. Bhargava and A. Bansal, "Fruits and vegetables quality evaluation using computer vision: A review," *Journal of King Saud University - Computer and Information Sciences*. 2018, doi: 10.1016/j.jksuci.2018.06.002.

[15]   S. Kumar and R. Kaur, "Plant Disease Detection using Image Processing- A Review," *Int. J. Comput. Appl.*, vol. 124, no. 16, pp. 6–9, 2015, doi: 10.5120/ijca2015905789.

[16]   K. Prakash, P. Saravanamoorthi, R. Sathishkumar, and M. Parimala, "A Study of Image Processing in Agriculture," *Int. J. Adv. Netw. Appl.*, vol. 22, no. 4, pp. 3311–3315, 2017.

[17]   D. D. Prihatin, B. Hidayat, and S. Saidah, "DETEKSI BATIK BOJONEGORO MENGGUNAKAN METODE GRAY LEVEL CO- OCCURRENCE MATRIX (GLCM) DAN NAIVE BAYES DETECTION OF BATIK BOJONEGORO USING GRAY LEVEL CO-OCCURRENCE MATRIX (GLCM) AND NAIVE BAYES Desi Dwi Prihatin [1], Dr . Ir . Bambang Hidayat , DEA .[," vol. 5, no. 3, pp. 4650–4657, 2018.

[18]   L. Veronica, I. H. Al Amin, B. Hartono, and T. Kristianto, "Ekstraksi Fitur Tekstur Menggunakan Matriks GLCM pada Citra dengan Variasi Arah Obyek," *Pros. SENDI_U*, pp. 978–979, 2019.

[19]   C. Jatmoko and D. Sinaga, "Ektraksi Fitur Glcm Pada K-Nn Dalam Mengklasifikasi Motif Batik," pp. 978–979, 2019.

[20]   "119-201-1-PB.pdf." .

sktp-04-01-2023 03_43_03-970137

Disease Image Classifications Using KNN Based On GLCM Feature Extraction", Journal of Physics: Conference Series, 2020