Journal of Artificial Intelligence and Engineering Applications

Websites:https://ioinformatic.org/

15th October2023. Vol.3. No.1; e-ISSN :2808-4519

Identification of Longan Species Based on Leaf Shape Texture and Color Using KNN Classification

Setia Adiyasa Lubis[1*], Yani Maulita[2], Milli Alfhi Syari[3]

[1,2,3]*Informatics Engineering*, STMIK Kaputama
setiaadiyasa@gmail.com[1*], yanimaulita26@gmail.com[2], milli.fhisya@gmail.com[3]

Abstract

This study aims to identify the type of longan based on the shape, texture and color of the leaves using KNN classification. With a method that can identify the type of longan automatically, farmers and researchers can obtain information more quickly and accurately about the type of longan that is being cultivated or studied. This can help in choosing the right variety, more efficient maintenance, and improve the quality and productivity of longan plants. This research is an experimental research consisting of eight steps, namely preparation, theoretical studies, data collection, data analysis and processing, testing and implementation and the last is the final stage. Based on research conducted at UD Mitra Tani on Jalan Madura No. 81 Kebun Lada, Kec. Binjai Utara, Binjai City, North Sumatra, the results of data analysis from longan leaves show that the most common type of longan found in the nursery is Red longan. This study was conducted to identify the dominant longan species in the population and gain a deeper understanding of the diversity of longan varieties in the region.

Keywords: Identification, KNN Classification, Longan, Type

1. Introduction

Longan is a type of fruit that is very popular in Southeast Asia, especially in Indonesia, Thailand and Vietnam. The longan plant is one of the most popular fruit crops and has high economic value in several countries, including Indonesia. Longan has several different varieties, including local varieties and imported varieties. Longan identification is important in the maintenance and breeding of this plant. However, the process of manually identifying longan species can be a complicated task and requires in-depth knowledge of the plant's characteristics. In addition, manually identifying longan species can indeed be a difficult and time-consuming task, especially if done in large quantities [1].

In recent years, the use of image processing and machine learning techniques has experienced rapid development. These techniques are able to extract visual features of objects, such as shape, texture, and color, to perform automatic recognition and classification. In this case, identification of longan species based on leaf features becomes possible automatically with the help of machine learning algorithms. The use of image processing technology to identify longan species based on leaf features such as texture and color allows us to save time and money in the identification process. Digital image processing technology, namely classification techniques, can generate features that can be used to classify longan species automatically [2]. One of the most important stages in the image recognition process is feature extraction. In some studies, the feature extraction used is the color value of RGB images, or HSV images and grayscale images.

Research entitled "Classification of Longan Types Based on Leaves Using Convolutional Neural Network Multilayer Perceptron" develops a system for recognizing types of longan based on leaves by increasing the extraction of longan leaf features based on characteristics, shape, or texture. The image used to identify the type of longan is a 128x128 color image. The types of longan studied are itoh, diamond river, and local longan with a total of 150 training data images and 30 test data images. The classification results of longan types using multilayer perceptron are good [3].

Meanwhile, in his research entitled "Classification of Longan Plant Types Based on Leaf Texture Characteristics Using the Adaptive Neuro Fuzzy Inference System (AFIS) Method" explains that not everyone can identify plant species with characteristics such as leaves, stems, and fruit. One of the characteristics of the two longan fruits is as image data for classification and feature extraction to identify the characteristics of longan plants. In this study, the classification method used to classify longan leaf images is the Adaptive Neuro Fuzzy System (ANFIS) with the extraction of longan leaf texture features using the Gray Level Co-occurance Matrix (GLCM) extraction method. The types of longan used are diamond river, new crystal, matalada and puang rai longan leaf images [4].

Features of leaf shape, texture, and color are characteristics that can be used to distinguish litchi types. Shape features can be used to identify leaf shapes, texture features can be used to identify leaf patterns and structures, and color features can be used to identify color variations in leaves. Digital image processing techniques such as GLCM (Gray Level Co-occurrence Matrix), HSV (Hue, Saturation, Value), and Eccentricity Metric can be used to extract shape, texture and color features of leaves. Classification methods such as KNN (K-Nearest Neighbor) can be used to compare unknown leaf images with known leaf images in the database and determine the appropriate type of longan [5].

This research also has relevance in the context of agriculture and longan plant breeding. With a method that can automatically identify the type of longan, farmers and researchers can obtain information more quickly and accurately about the type of longan that is being cultivated or studied [6]. This can help in choosing the right variety, more efficient maintenance, and improve the quality and productivity of longan

plants. KNN (K-Nearest Neighbors) is one method that is quite simple but effective in classifying data based on the features that have been extracted. KNN is also very popular in the field of image processing because of its ability to handle differences in data scale, the ability to handle data that is not normally distributed, and can overcome the problem of outliers [7].
Based on the above problems, the authors are interested in developing research that uses image processing technology to identify longan types based on leaf characteristics such as shape, texture and color, we can produce an accurate and efficient longan identification system, so that it can speed up and facilitate the process of identifying longan types with the title "Identification of Longan Types Based on Leaf Shape, Texture and Color Using KNN Classification".

2. Research Methods

Research method is a method or steps used to carry out research in a systematic and structured manner with the aim of collecting information or data needed to answer research questions or achieve predetermined research objectives. This research is an experimental research consisting of eight steps, namely preparation, theoretical studies, data collection, data analysis and processing, testing and implementation and the last is the final stage [8].

In the preparation stage, the researcher ensures the purpose of the research, the method to be used and the tools and materials needed. In the theory study stage, researchers describe previous research and write down the literature related to the research title. At the data collection stage, researchers determine what methods will be used to collect data, in this study researchers used observation, interviews and literature studies as data collection methods. At the data analysis and processing stage, the data is identified based on the shape, texture and color of the leaves, of course requiring image data of the longan leaf type.

3. Results and Discussion

In the identification system to recognize the type of longan based on the shape, texture and color of its leaves, of course, requires image data of the type of longan leaves. Thus the data is collected by taking photos of longan leaves based on their type which can be seen in the picture below.

Figure 1: Longan Aroma Durian

Figure 2: Longan Red

Figure 3: Longan New Crystal

Figure 4: Longan Pimpong

Figure 5: Longan Puangray

After identifying the types of longan based on the leaves as mentioned in the picture above, data on extraction from longan leaves will be presented.The data used in the identification process using the KNN classification are image data of longan leaves. Then the data is extracted features to be used as training data. The following is the training data class used for testing longan leaves based on the shape, texture and color of the leaves as shown in the table below.

Table 1: Training Data

No	Feature Extraction	Mark	Class	Longan Type
1	Color	Hue	0.23643	Aroma Durian
		saturation	0.44707	
		Value	0.29289	
	Shape	Eccentricity	0.94318	
		Metrics	0.53216	
	Texture	Means	124,579	
		Variances	6340,9658	
		Skewness	0.40163	
		kurtosis	-1.6583	
		Entropy	6.5801	
		ASM	0.0032204	
		Contrast	52.5197	
		Correlation	0.99585	
		Variances	6305.0156	

		IDM	0.37487	
		Entropy	10,398	
2	Color	Hue	0.21297	Aroma Durian
		saturation	0.7202	
		Value	0.34678	
	Shape	Eccentricity	0.89842	
		Metrics	0.64943	
	Texture	Means	113.6758	
		Variances	4421,4451	
		Skewness	0.7688	
		kurtosis	-1.2328	
		Entropy	6.4652	
		ASM	0.0024506	
		Contrast	25,804	
		Correlation	0.99707	
		Variances	4395,5787	
		IDM	0.4065	
		Entropy	10.0381	
3	Color	Hue	0.11714	Red
		saturation	0.33813	
		Value	0.35096	
	Shape	Eccentricity	0.92265	
		Metrics	0.64238	
	Texture	Means	132,712	
		Variances	4971.0164	
		Skewness	0.7716	
		kurtosis	-1.2088	
		Entropy	6.5359	
		ASM	0.0020898	
		Contrast	56.5917	
		Correlation	0.99429	
		Variances	4925,2663	
		IDM	0.3262	
		Entropy	10.5928	
4	Color	Hue	0.10204	Red
		saturation	0.28106	
		Value	0.32987	
	Shape	Eccentricity	0.94069	
		Metrics	0.58536	
	Texture	Means	131.6494	
		Variances	4727,6812	
		Skewness	0.68869	
		kurtosis	-1.3386	
		Entropy	6.3268	
		ASM	0.0025208	
		Contrast	29.7787	
		Correlation	0.99684	
		Variances	4696,779	
		IDM	0.40056	
		Entropy	9.9038	
5	Color	Hue	0.2356	New Crystal
		saturation	0.33275	
		Value	0.37659	
	Shape	Eccentricity	0.96925	
		Metrics	0.45179	
	Texture	Means	123.6288	
		Variances	4212,1077	
		Skewness	0.85684	
		kurtosis	-1.0155	
		Entropy	6.4141	
		ASM	0.0027583	
		Contrast	75.3552	
		Correlation	0.99102	
		Variances	4156,1884	
		IDM	0.35496	
		Entropy	10.4315	

6	Color	Hue	0.22536	New Crystal
		saturation	0.36937	
		Value	0.35596	
	Shape	Eccentricity	0.93989	
		Metrics	0.59227	
	Texture	Means	115,879	
		Variances	3893,7019	
		Skewness	1.0346	
		kurtosis	-0.71213	
		Entropy	6,185	
		ASM	0.0043528	
		Contrast	41.5468	
		Correlation	0.99464	
		Variances	3856,2326	
		IDM	0.42506	
		Entropy	9.7557	
7	Color	Hue	0.32325	Pimpong
		saturation	0.24091	
		Value	0.2571	
	Shape	Eccentricity	0.8427	
		Metrics	0.69318	
	Texture	Means	94.4561	
		Variances	5373,754	
		Skewness	0.72293	
		kurtosis	-1.3662	
		Entropy	6.1189	
		ASM	0.0041314	
		Contrast	36.4864	
		Correlation	0.99659	
		Variances	5327,1643	
		IDM	0.48336	
		Entropy	9.3123	
8	Color	Hue	0.30335	Pimpong
		saturation	0.22606	
		Value	0.32831	
	Shape	Eccentricity	0.83352	
		Metrics	0.70217	
	Texture	Means	103.9002	
		Variances	4667,5162	
		Skewness	1.0865	
		kurtosis	-0.67133	
		Entropy	5.9328	
		ASM	0.0045577	
		Contrast	39.3704	
		Correlation	0.99575	
		Variances	4612,8325	
		IDM	0.44746	
		Entropy	9.2928	
9	Color	Hue	0.24136	Puangray
		saturation	0.49482	
		Value	0.34151	
	Shape	Eccentricity	0.93666	
		Metrics	0.54235	
	Texture	Means	111.3514	
		Variances	4411,4498	
		Skewness	0.8379	
		kurtosis	-1.1197	
		Entropy	6.0373	
		ASM	0.0062751	
		Contrast	29.7775	
		Correlation	0.99661	
		Variances	4376,1555	
		IDM	0.55484	
		Entropy	8.8612	
10	Color	Hue	0.24599	Puangray
		saturation	0.49096	

		Value	0.31615	
	Shape	Eccentricity	0.94648	
		Metrics	0.51174	
	Texture	Means	103.4131	
		Variances	4459.2194	
		Skewness	0.87328	
		kurtosis	-1.0568	
		Entropy	5.9534	
		ASM	0.0064789	
		Contrast	35,348	
		Correlation	0.99602	
		Variances	4419.4359	
		IDM	0.56615	
		Entropy	8.7326	

Calculation of the K-Nearest Neighbor Method for all types of samples is done manually using the following equation:
First,Determined the value of K = 3
Second, Calculate the distance of test data and training data withEuclidean distance formula.

$$d(p,q) = \sqrt{\sum_{i=0}^{n}(q_i - p_i)^2} \quad (1)$$

$p, q \quad = \text{two points in space} - \text{n euclidean}$ (2)

$q_i\ , p_i \quad = \text{vektor euclidean, starting from the origin of the space (starting point)}$ (3)

$n \quad = \text{space} - n$ (4)

After manual calculation; withK-Nearest Neighbor method of all types of samples take the three closest distance data classes from the calculation results in table 2.

Table 2: Closest Distance Data

No	Euclidean Distance	New Data Distance	Class
1	1602,79417	9	Aroma Durian
2	1105,43192	6	Aroma Durian
3	341.61293	2	Red
4	675.77632	3	Red
5	1422,26018	8	New Crystal
6	1859,45203	10	New Crystal
7	231.96471	1	Pimpong
8	778.00279	4	Pimpong
9	1126,12249	7	Puangray
10	1061.85887	5	Puangray

After taking the three classes with the shortest distance, then take the most neighboring classes in table 3.

Table 3: Most Class

No	EuclideanDistance	New Data Distance	Class
1	231.96471	1	Pimpong
3	341.61293	2	Red
4	675.77632	3	Red

the results with the most majority are the Red class shown in table 4.

Table 4: The final result

No	Feature Extraction		Mark	Class
1	Color	Hue	0.12483	Red
		saturation	0.33972	
		Value	0.34671	
	Shape	Eccentricity	0.91938	
		Metrics	0.63369	
	Texture	Means	136.0822	
		Variances	5211.0042	
		Skewness	0.62313	
		kurtosis	-1.4243	
		Entropy	6.5883	
		ASM	0.0023442	
		Contrast	55.6036	
		Correlation	0.99465	
		Variances	5168,3565	
		IDM	0.32991	
		Entropy	10.6116	

Evaluate the accuracy and effectiveness of the image processing system and KNN classification model developed on different longan leaf image data used in a system design. The designed system contains a system front page menu, which is the initial menu that only displays the features of the training, testing and image processing menus, after the main menu there is a system training menu, this menu is a menu that displays training data that will be trained with existing training data using the KNN classification with the K parameter that has been determined and the trained data model is stored and can be used in the testing menu and image processing menu. After the training menu there is a testing menu, this menu displays test data that will be tested with previously stored training data models using KNN classification. The image processing menu is a menu that displays 1 test data that will be processed with a pre-stored training data model using KNN classification to determine the value of feature extraction and classification results. Flowchart menu or program logic for image retrieval, classification process, classification results and recommendations.

4. Conclusion

Based on research conducted at UD Mitra Tani on Jalan Madura No. 81 Kebun Lada, Kec. Binjai Utara, Binjai City, North Sumatra, the results of data analysis from longan leaves show that the most common type of longan found in the nursery is Red longan. This study was conducted to identify the dominant longan species in the population and gain a deeper understanding of the diversity of longan varieties in the region.

Acknowledgments

The authors would like to express their deepest gratitude to all those who have provided support and assistance during the implementation of this research. Especially to the owner of the UD Nursery Business. Mitra Tani on Jalan Madura No. 81 Kebun Lada, Kec. North Binjai, Binjai City, North Sumatra, for his willingness to provide a research location and become a very valuable resource person in this research. Also to the supervisors who with full direction and guidance have helped in the implementation of research and writing this report to completion. Thank you all for your valuable contributions and support.

References

[1] Mariana, B. D., Sugiyatno, A., & Supriyanto, A. (2016). Genetic Diversity of Local Accessions of Dimocarpus longan Revealed By ISSR Markers.*Buletin Plasma Nutfah*, 17(1), 25.

[2] Kaya, J. U. (2022). Klasifikasi Motif Kain Tenun menggunakan K-Nearest Neighbor Berdasarkan Gray Level Co-occurrence Matrix. *Jurnal J-COM (Jurnal Informatika dan Teknologi Komputer)*, Vol. 03 No. 02 73 – 76.

[3] Fauziah, W. N. (2015). "Uji aktivitas Antimikroba Ekstrak etanol Daun, Kulit Dan Biji Kelengkeng Terhadap Pertumbuhan Saccharomyces Cerevisiae Dan Lactobacillus Plantarum Penyebab Kerusakamm Mira Siwalan". *Skripsi, Malang Program Pasca Sarjana Sains.*

[4] Firdaus R, A. (2022). Klasifikasi Jenis Tanaman Kelengkeng Berdasarkan Ciri Tekstur Daun Menggunakan Metode Adaptive Neuro Fuzzy Inference System (AFIS). *Jurnal Explore IT 14.*

[5] Mardhyath, A., Latipa , H. S., & Julita, R. (2020). Algoritma K-Nearest Neighbour dalam Mengklasifikasi Penilaian Peserta Didik. *JSAI : Journal Scientific and Applied Informatics*, Vol. 3, No. 3, hal. 149-155.

[6] Hanafi, M. H., Fadillah, N., & Insan, A. (2019). Optimasi Algoritma K-Nearest Neighbor untuk Klasifikasi Tingkat Kematangan Buah Alpukat Berdasarkan Warna. It Journal Research and Development, 4(1), 10–18. https://doi.org/10.25299/itjrd.2019.vol4(1).2477 Hanafi, M. H., Fadillah, N., & Insan, A. (2019). Optimasi Algoritma K-Nearest Neighbor untuk Klasifikasi Tingkat Kematangan Buah Alpukat Berdasarkan Warna. *It Journal Research and Development*, 4(1), 10–18

[7] Nafi'iyah, N., & Prasetyo, R. (2022). Klasifikasi Jenis Kelengkeng Berdasarkan Daun Menggunakan Convolution Neural Network Multilayer Perceptron. Jurnal Telematika, Vol. 17 No. 2.

[8] Salsabila, A., Yunita, R., & Rozikin, C. (2021). Identifikasi Citra Jenis Bunga menggunakan Algoritma KNN dengan Ekstrasi Warna HSV dan Tekstur GLCM. *Technomedia Journal*, 6(1 Agustus), 124–137.