IAES International Journal of Artificial Intelligence (IJ-AI)
Vol. 13, No. 3, September 2024, pp. 2934~2945
ISSN: 2252-8938, DOI: 10.11591/ijai.v13.i3.pp2934-2945

A hybrid hue saturation lightness, gray level co-occurrence matrix, and k-nearest neighbour for palm-sugar classification

Mila Jumarlis[1], Ida Mulyadi[2], Mirfan[3], Irmawati[4], Mardiah[5], Muhammad Faisal[2], Hairin Anisa[3]
[1]Islamic Communication and Broadcasting Study Program , Department of Ushuluddin, Adab and Da'wah, STAIN Majene, Majene, Indonesia
[2]Department of Computer Science, STMIK Profesional, Makassar, Indonesia
[3]Department of Informatics Engineering, Faculty of Computer Science, Universitas Handayani Makassar, Makassar, Indonesia
[4]Department of Computer and Network Engineering, Politeknik Negeri Ujung Pandang, Padang, Indonesia
[5]Department of Computer System, STMIK Tidore Mandiri, Tidore, Indonesia

Article Info

Article history:

Received Nov 10, 2023
Revised Feb 22, 2024
Accepted Mar 5, 2024

Keywords:

Classification
Gray level co-occurrence matrix
Hue-saturation-lightness
K-nearest neighbor
Palm sugar

ABSTRACT

In recent years, there has been an increasing demand for high-quality raw materials driven by consumers and the food industry. This study aims to build a model to predict the type of palm sugar using a hybrid method of hue-saturation-lightness (HSL), gray level co-occurrence matrix (GLCM), and K-nearest neighbor (KNN). The price of palm sugar is determined based on the type and ingredients used. However, due to the lack of public knowledge in distinguishing the types of palm sugar, there is the potential for price manipulation that can harm the community. The accuracy rate of 97.6% of the palm sugar type prediction results shows that the model that was built has worked very well. The results have practical implications, such as developing automated systems to classify palm species in specific industries to benefit economics and operational efficiency. Future research directions may explore the integration of advanced machine-learning techniques and real-time image processing for further improving classification performance and scalability in industrial applications.

Corresponding Author:

Muhammad Faisal
Department of Computer Science, STMIK Profesional
St. AP. Petarani No.27, Panakukang, Makassar, Indonesia
Email: muhfaisal@stmikprofesional.ac.id

1. INTRODUCTION

In recent years, there has been an increasing demand for high-quality raw materials driven by consumers and the food industry [1]. The worldwide food sweetener market is expected to be worth more than $100 billion by the end of the decade, with the Asia-South Pacific region witnessing the highest market share growth [2]. Palm sugar is one of the raw materials that are in demand by the public and industry because it can be used in various types of dishes and beverages to improve taste and texture. Good quality palm sugar generally has a bright and uniform colour, smooth texture, and does not clump. Palm sugar staining is one of the key attributes that matter to consumers when purchasing, but models exploring brown sugar colouring are still rare in scientific research [3]. Palm sugar is a sweetener made from the sap or nectar collected from different varieties/species of palm trees. It is also a valuable non-centrifugal sugar and a traditional sweetener with a unique taste and flavour [4]. Palm sugar is made from sap or nectar collected from the flowers of various types of palm, including coconut, aren, and lontar [5]. Palm sugar can be used as an alternative to cane sugar due to its popularity and high production in Southeast and South Asia, thus providing economic benefits [6]. Traditional markets have many palm sugar types, with prices varying based on the type. However, the lack of public knowledge in distinguishing types of palm sugar causes some traders to manipulate prices. In contrast, palm sugar products need to be marketed, targeted, and positioned correctly so that the product is economically viable for those involved in its production and supply [2].

Image processing with potential quantum computing hardware requires quantum colour models to capture and manipulate image colour information [7]. Research by Kadhim and Abbas in [8] proposes a preference for using hue, saturation, and lightness (HSL) and hue, saturation, value (HSV) color spaces to deliver superior performance over using dedicated HSV to detect white. The results proved to be very effective in detecting straight paths with an accuracy ratio of 96.06%. Research by Noman *et al.* in [9], automatic lane detection of lane lines on various roads to help drivers use gradient threshold on algorithm HSL in binary images. The proposed system achieves 96% accuracy in detecting lane lines on various roads, and its performance is assessed using data from several road image databases under various lighting conditions.

Research related to applying the gray level co-occurrence matrix (GLCM) algorithm to recognise human walking activities from a series of colour images processed first using resizing approaches, median filters, and contrast stretching [10]. Better performance results are obtained to predict gait events effectively. Another study related to the methodology for analyzing the inner surface of steam pipes using digital image processing on the GLCM algorithm was carried out by analyzing all attributes of each image texture used, where the classification results achieved effectiveness more significantly than 92% [11]. In the domains of data mining, machine learning, and pattern recognition, there is a substantial emphasis on conducting research related to classification tasks due to their fundamental importance in advancing and evolving the field [12]. A study proposes the k-nearest neighbors (KNN) method to evaluate crop pest images and filter highly informative data to solve pest recognition tasks with efficient data [13]. Based on the comparative experiments conducted, the results show that the proposed KNN method can effectively distinguish high and low informative data. Previous research applied image processing and machine learning in identifying and classifying machine surface textures using the GLCM extraction feature [14], where the test results show a reasonable accuracy rate of 91.3%.

Based on previous research on extraction techniques and image classification, no research has been found using hybrid HSL, GLCM, and KNN. This research proposes to solve the problem of identifying various types of palm sugar by applying the HSL-GLCM-KNN hybrid technique to classify palm sugar types. By identifying images, it becomes possible to gather information regarding various palm sugar types, including aren palm, coconut palm, and lontar palm, which can serve as guidelines for determining appropriate pricing. The research design is shown in Figure 1. Based on the illustration in Figure 1, it is known that this study involves phase training through the formation of a model used in phase testing so that the results of the classification of palm sugar types are known.

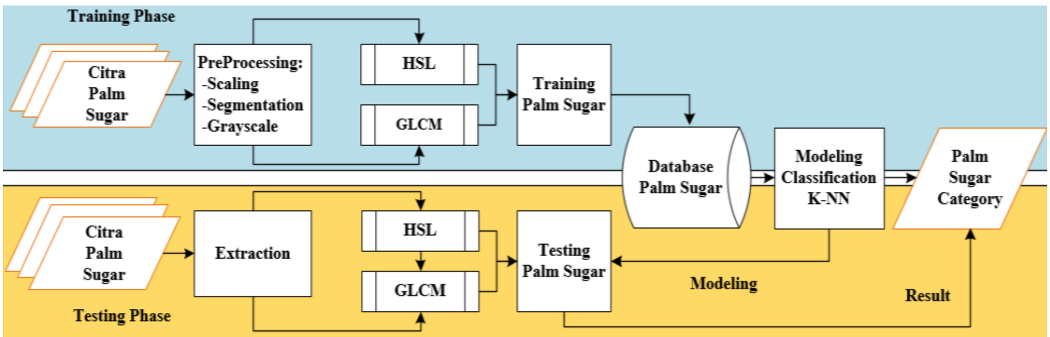


Figure 1. Research design

2. METHOD

2.1. Palm sugar

Palm Sugar is a sugar that has a solid form with a reddish-brown to dark-brown colour. Palm sugar is usually sold in semicircular shapes printed using coconut shells or bamboo. Chemically, sugar is the same as carbohydrates. Generally, sugar refers to carbohydrates that have a sweet taste, are small in size and can dissolve [15]. This study divides palm sugar in Indonesia into three categories, as shown in Figure 2. In Figure 2, the types of palm sugar are shown, including Figure 2(a) shows aren (*Nypafruticans*), which has reddish-brown characteristics and dry and clean texture; Figure 2(b) shows coconut/*Cocos nucifera linn* has dark brown characteristics with a compact structure and texture, and Figure 2(c) shows lontar (*Borassus flabellifer*) has light brown characteristics and dry texture and rough surface.

(a) (b) (c)

Figure 2. Palm sugar category consists of (a) aren, (b) coconut, and (c) lontar

2.2. Hue, saturation and lightness

Lightness in HSL models is a brightness value in colour, a component value that is white to black. When the lightness value is 0, then the colour has a very low brightness or causes the colour to darken. Meanwhile, if the lightness value is 1, the brightness level of the colour becomes very bright or white. Lightness is a relative scale measure of the lightness or darkness of a colour, generally measured as a percentage from 0% (black) to 100% (white). The algorithm of colour conversion from RGB to HSL is written as follows [16]:

$$\mathrm{H} = \mathrm{arctan2}(\mathrm{B}', \mathrm{R}') * (180/\pi) \quad (1)$$

$$\mathrm{L} = (\max(\mathrm{R}', \mathrm{G}', \mathrm{B}') + \min(\mathrm{R}, \mathrm{G}, \mathrm{B})) / 2 * 100\% \quad (2)$$

$$\mathrm{S} = ((\max(\mathrm{R}', \mathrm{G}', \mathrm{B}') - \min(\mathrm{R}, \mathrm{G}, \mathrm{B})) / (1 - |\mathrm{L} * 2 - 1|)) * 100\% \quad (3)$$

The immutability in each formula is the range of values possessed by lightness, which is 0 to 1, and the range of values in saturation. Meanwhile, the range of values owned by hue is 0 to 360. The HSL colour model has a different colour space model than RGB.

2.3. Gray level co-occurrence matrix

GLCM is a 2D image analysis technique that provides quantitative data on image textures [17]. The co-currency matrix is one of the statistical methods used for texture analysis formed from an image by looking at paired pixels with a certain intensity. This method is based on the hypothesis that there be looping of greyish configurations or pairs of levels in a texture. A co-currency matrix expresses the spatial distribution between two neighbouring pixels with intensities i and j, a d distance between them, and an angle θ between them. The co-currency matrix is expressed by Pdθ(i, j) [18]. A neighbouring pixel with a distance of "d" between them can be located in eight directions. In the co-currency matrix, texture characteristics can be obtained from an image used as a differentiator between images with certain classes and other classes [19]. Extraction of characteristics from a digital image from the angle of neighbourliness of pixels contained in GLCM [20]:

The working contrast shows the spread (moment of inertia) of elements on the image matrix, calculated using in (4):

$$\text{Contras} = \sum_i \sum_j [i-j]^2 . p(i,j) \quad (4)$$

Furthermore, the correlation characteristic is used as a measure of the linear dependence between greyish level values in the image calculated using (5):

$$\text{Correlation} = \sum_i \sum_j \frac{(i-\mu i)(j-\mu j)p(i,j)}{\sigma i \sigma j} \quad (5)$$

Furthermore, the energy feature serves to represent a measure of uniformity in the image, calculated using the (6):

$$\text{Energy} = \sum_{i,j} p(i,j)^2 \quad (6)$$

Finally, the homogeneity feature is the similarity or homogeneity of variations of the concurrency matrix in the observed image [21], which is written in (7):

$$\text{Homogeneity}= \sum_i \sum_j \frac{p(i,j)}{1+[i-j]^2} \quad (7)$$

2.4. K-nearest neighbors technique

KNN is a popular supervised machine learning algorithm that can be effectively used for multiclass classification and regression utilizing multiple dataset features [22]. Calculation of neighbourly distance using Euclidean algorithm as shown in (8):

$$\text{Euclid}=\sqrt{((a_1\text{-}b_1)^2+...+\left(a_n\text{-}b_{n)^2}\right)} \quad (8)$$

Where a=a1,a2, .., an, dan b=b1, b2, .., bn represents n attribute values from two records. Among classic data mining algorithms, the KNN-based methods are practical and straightforward solutions for classification tasks [23].

The proposed method is divided into three main parts. Firstly, which involves preprocessing and feature extraction, begins with image acquisition to obtain high-quality palm sugar sample images. In the preprocessing step, the images are resized, undergo noise reduction, and have their color normalized. For feature extraction, the images are converted to HSL to calculate the mean, standard deviation, and skewness of the HSL. GLCM is also computed to extract contrast, homogeneity, dissimilarity, energy, entropy, and correlation. These HSL and GLCM features are then combined in the feature concentration step. If needed, principal component analysis (PCA) is used to reduce the dimensionality of the feature vectors. Secondly, focuses on palm sugar classification, starting with selecting the distance metrics based on Euclidean and Manhattan theory. Next, the number of nearest neighbors (k) is specified. A KNN model is trained using labeled data, and the trained model is used to classify new images. Thirdly, evaluation and validation occur. The model's accuracy, precision, recall, and F1-score are measured using test data. Cross-validation is then performed to ensure the model's generalizability. The next section will provide a detailed explanation of each stage to clarify the design and implementation process.

3. RESULTS AND DISCUSSION

This section explores the hybrid methodology involving HSL, GLCM, and K-means clustering. These techniques are integrated to analyze and classify palm sugar samples accurately. Each stage is interrelated to ensure a seamless process, linking preprocessing, feature extraction, and classification for accurate and reliable outcomes. The following subsections provide detailed insights into the methodology's architecture, design, and each stage's implementation, showing how the components work together to achieve precise results.

3.1. Proses preprocessing

The training model allows the system to automatically recognize and classify objects in images without human intervention [24], thus increasing efficiency in various palm sugar image processing applications. The image entered into the system is an image that has been cropped and resized, containing images of palm sugar, aren, coconut and lontar types. The preprocessing technique aims to transform the dataset into suitable inputs for machine learning [25]. In this study, the preprocessing stage is carried out in cropping, resizing and image colour conversion, where each image be cut by comparison 1:1. Figure 3 displays the cropping process of palm sugar images in relevant sections for analysis to prepare training process datasets.

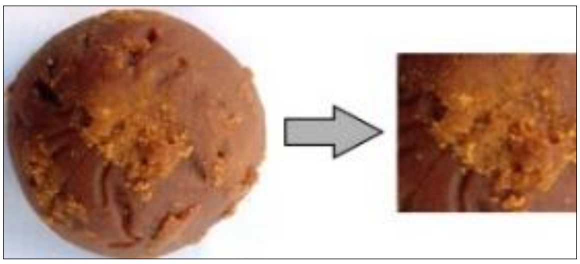


Figure 3. Cropping of image

3.1.1. Scaling and segmentation

Image decomposition generally separates image components or information based on parameters such as colour, orientation, or scale [26]. In this study, the image used as input with the original size of 6,000×4,000 pixels was cropped to 256×256 pixels. The purpose of reducing pixels is to limit the value of objects compared to make the classification process easier. Image segmentation is essential in various computer vision applications by grouping image pixels into pixel-filled segments on an object [27]. Segmentation is the first step in pattern recognition, where the characteristics of segmented objects can be used to recognize patterns or perform classification tasks. The results of segmentation on the processed image are shown in Figure 4. Based on the results of segmentation in Figure 4, better image quality is produced so that it is easier to identify and understand an object.

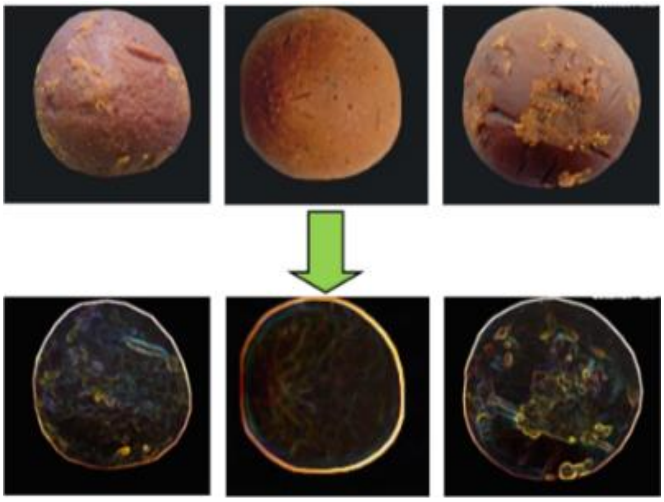


Figure 4. The results of palm sugar segmentation types of aren (left), lontar (middle), coconut (right)

3.1.2. Grayscale

Grayscale images support a more focused analysis process on a component without the distraction of colour information [28]. The grayscale mode in the image processing method is needed to represent each pixel in an image that has one greyish value so that it is known to what extent the pixels are in light or dark conditions. In this study, the resizing process was carried out to change the resolution or horizontal and vertical size of an image and convert RGB images to grayscale to form a matrix of grayish values of an image. The image conversion RGB to Grayscale is shown in Figure 5.

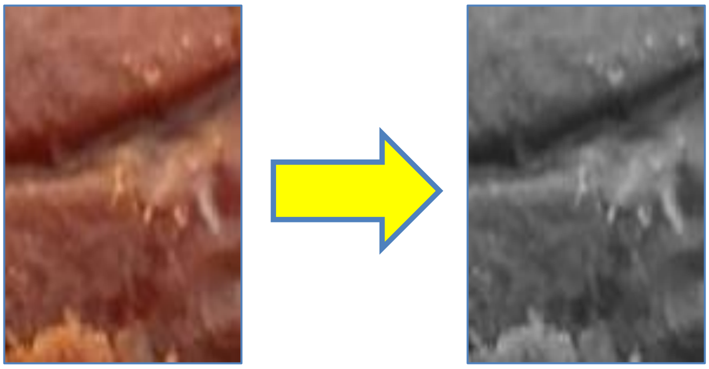


Figure 5. Conversion of RGB to grayscale mode

The grayscale mode requires only one colour channel per pixel, significantly reducing data complexity compared to colour (RGB) images. RGB to the grayscale conversion of multiple sugar datasets consisting of 3 categories is shown in Figure 6. After converting from RGB to grayscale, then calculate the HSL calculating as follows:

$R = 176, G = 111 , B = 77$

Grayscale $= 0.296R + 0.312B + 0.113G$

The value of RGB to grayscale conversion results is as follows :

$= 0.296(176) + 0.312(111) + 0.113(77)$

$= 52.096 + 34.632 + 8.701 = 95.429.$

The calculation of the HSL value is carried out by calculating the normalization of RGB values as follows:

$R = 176 / 255 = 0.690$. $G = 111 / 255 = 0.435$. $B = 77 / 255 = 0.302$

Based on the normalized value, the Hue value is determined using (1),

$H = \arctan2(0.302, 0.690) * (180/\pi)$. $H = 27.37°$

Next, calculate the Saturation value using (2),

$L = (\max(R, G, B) + \min(R, G, B)) / 2 = (0.690 + 0.302) / 2 = 0.496$

$S = ((\max(R, G, B) - \min(R, G, B)) / (1 - |L * 2 - 1|)) * 100\%$

$S = ((0.690 - 0.302) / (1 - |0.496 * 2 - 1|)) * 100\%$. $S = 39.11\%$

Finally, calculate the Lightness value using (3),

$L = (0.690 + 0.302) / 2 * 100\%$. $L = 49.6\%$

Based on the calculation results, it is set in Figure 3 to have RGB values of 176, 111, and 77. Then, it is decided HSL values consist of 27.37°, 39.11%, and 49.6%.

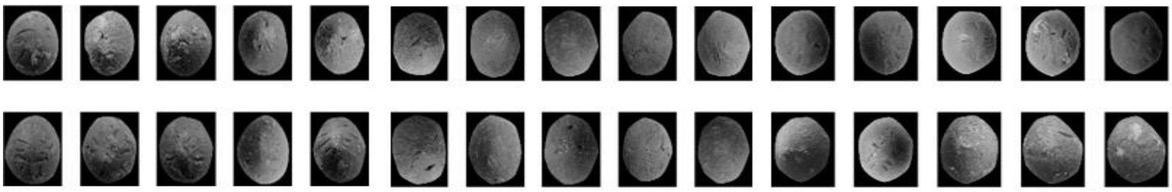
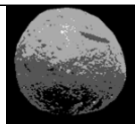
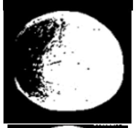
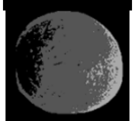
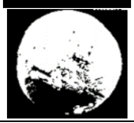
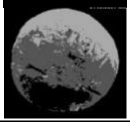
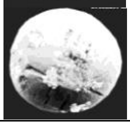


Figure 6. Palm sugar for each category in grayscale mode

3.2. Extraction

The extraction of this research feature uses the HSL and GLCM method, as stated that the advantage of the image extraction process is that it increases precision based on pixel values or histograms [29]. Image inputs that have gone through the pre-processing process be extracted features. Feature extraction begins by forming a co-occurrence matrix of each image then, this matrix calculate GLCM extraction features, namely entropy, energy, homogeneity, and contrast, with each angle used 0°, 45°, 90°, and 135°. This study proposes a colour correction methodology based on the linguistic definitions of RGB colour images. The idea is based on correcting both hue and saturation channels according to the colours that appear in the colour image without changing their RGB values. The objective of the correction stage is to minimize the number of objects in the H channel to be processed faster and more efficiently. In Table 1, the results of extracting images are displayed.

Table 1. Extraction result for each palm sugar image type

Palm Sugar	RBG	Hue	Saturation	Lightness
Aren				
Lontar				
Coconut				

The GLCM feature extraction results, shown in Table 1, are calculated using distinct equations for each feature. In (4) is used to compute the accumulated contrast values, providing insights into the difference in intensity. In (5) measures the accumulated correlation, which indicates the linear dependency of pixel intensities. In (6) calculates the accumulated energy to reflect the uniformity of pixel distribution, while (7) determines accumulated homogeneity, offering a measure of the distribution similarity across the image. The results for each feature are then detailed comprehensively in Table 2.

GLCM features extracted the entire training image dataset of 300 images and 12 tested images. The features used are contras, correlation, energy, and homogeneity. The results of the extraction of training image characteristics are stored in a database for use in the classification and identification process.

Table 2. Extraction sample for each palm sugar type

No	Type	Contras				Correlation				Energy				Homogeneity			
		0°	45°	90°	135°	0°	45°	90°	135°	0°	45°	90°	135°	0°	45°	90°	135°
1	Aren	3.25	5.72	7.16	6.58	-0.007	-0.012	-0.092	-0.008	0.054	0.08	0.10	0.09	0.35	0.32	0.143	0.32
2	Coconut	3.20	6.41	7.37	6.41	-0.009	-0.008	-0.095	-0.008	0.013	0.09	0.429	0.08	0.28	0.28	0.140	0.28
3	Lontar	3.40	5.79	6.75	6.88	-0.007	-0.002	-0.08	-0.007	0.040	0.08	0.263	0.08	0.32	0.30	0.135	0.30

3.3. Classification

In this study, feature values are obtained in the image classification process using the GLCM method. Subsequently, these feature values are utilized in the classification process by comparing them with test and training data values through the KNN method. Training is carried out to obtain the weight of the training sample [30], [31]. It is further used to classify the image of palm sugar types. As is known, the combination of KNN enables accurate feature extraction and efficient classification, effectively overcoming computational challenges that are common to deep learning models [32]. The training result database contains the weight of the training sample generated from the training stage with KNN containing the palm sugar type model.

All data in Table 3 are used in training data to determine the type of palm sugar based on the dataset using the KNN method by applying (8), where the calculation results are displayed in Table 4. Each entered value is sorted according to the value of "K" in ascending order, and then the KNN algorithm looks for neighbouring values, which determine the results of the new value classification. Based on Table 4, it can be concluded that the closest distance obtained by each distance is shown in Table 5.

Based on Table 5, it can be concluded from 10 test data used with five K values, namely K=3, K=5, K=7, K=9, and K=10, the following percentages are produced:

Aren $= 100\% \ \frac{5}{10} = 50\%$

Lontar $= 100\% \ \frac{3}{10} = 30\%$

Coconut $= 100\% \ \frac{2}{10} = 20\%$

The calculation results show that the new data is in the aren class with a percentage of 50%.

Based on repeated testing, limitations were found related to the application of the methods used. Firstly, KNN performance heavily depends on the training data's size and quality, which a small or poorly labelled dataset can lead to overfitting and unreliable results. Secondly, variations in lighting, camera angles, and sample preparation can affect the features extracted from HSL and GLCM, potentially reducing classification accuracy. Finally, HSL and GLCM capture specific aspects of colour and texture but might not represent all relevant visual characteristics for accurate classification.

Table 3. Training data

Id	HSL			GLCM				Classification
	H	S	L	Contrast	Correlation	Energy	Homogeneity	
1	13	41	40	68.466364	0.903730	0.000693	0.254166	Lontar
2	14	64	31	70.379293	0.903020	0.000686	0.251907	Aren
3	97	36	37	125.59757	0.906657	0.000432	0.218463	lontar
4	24	54	40	162.27747	0.799944	0.000668	0.216504	Coconut
5	23	41	43	358.25626	0.639698	0.000310	0.146045	Coconut
6	20	35	54	54.198990	0.881672	0.001328	0.288104	Aren
7	21	71	31	48.793838	0.966512	0.000689	0.290730	Aren
8	84	22	50	126.14404	0.906464	0.000437	0.221047	Lontar
9	22	70	33	90.537980	0.887026	0.000520	0.240710	Lontar
10	12	53	33	35.575960	0.995470	0.002028	0.520886	Aren
11	10	50	25	81.968990	0.813505	0.007690	0.229342	Aren
12	21	57	31	202.01212	0.807464	0.000393	0.188507	Lontar
13	14	48	37	57.681414	0.960930	0.000576	0.275649	???

Table 4. Calculation of euclidean distance

Id	Euclidean Distance		Class
1	$=\sqrt{(68.46636-57.681414)^2+(0.90373-0.960930)^2+(0.000693-0.000576)^2+(0.2541663-0.275649)^2}$	= 10.7851	Lontar
2	$=\sqrt{(70.379293-57.681414)^2+(0.90302-0.960930)^2+(0.000686-0.000576)^2+(0.251907-0.275649)^2}$	= 12.9785	Aren
3	$=\sqrt{(125.597576-57.681414)^2+(0.906657-0.960930)^2+(0.000432-0.000576)^2+(0.218463-0.27565)^2}$	= 67.9161	Lontar
4	$=\sqrt{(162.277475-57.681414)^2+(0.799944-0.960930)^2+(0.000668-0.000576)^2+(0.216504-0.27565)^2}$	= 104.596	Coconut
5	$=\sqrt{(358.256263-57.681414)^2+(0.639698-0.960930)^2+(0.000310-0.000576)^2+(0.254166-0.27565)^2}$	= 300.574	Coconut
6	$=\sqrt{(54.198990-57.681414)^2+(0.881672-0.960930)^2+(0.001328-0.000576)^2+(0.288104-0.275649)^2}$	= 3.61600	Aren
7	$=\sqrt{(48.793838-57.681414)^2+(0.966512-0.960930)^2+(0.000689-0.000576)^2+(0.29073-0.2756490)^2}$	= 4.68710	Aren
8	$=\sqrt{(126.144040-57.681414)^2+(0.906464-0.960930)^2+(0.000437-0.000576)^2+(0.221047-0.27565)^2}$	= 8.88790	Lontar
9	$=\sqrt{(90.537980-57.681414)^2+(0.887026-0.960930)^2+(0.00052-0.000576)^2+(0.240710-0.275649)^2}$	= 32.8566	Lontar
10	$=\sqrt{(35.575960-57.681414)^2+(0.995470-0.960930)^2+(0.002028-0.000576)^2+(0.520886-0.275649)^2}$	= 22.1068	Aren
11	$=\sqrt{(81.968990-57.681414)^2+(0.813505-0.960930)^2+(0.00769-0.000576)^2+(0.229342-0.275649)^2}$	= 24.2880	Aren
12	$=\sqrt{(202.012121-57.681414)^2+(0.807464-0.960930)^2+(0.000393-0.000576)^2+(0.188507-0.27565)^2}$	= 144.331	Lontar

Table 5. Result of classification

Id	Euclidean Distance	Class	K=3	K=5	K=7	K=9	K=10
6	3.61600	Aren	Lontar	Aren	Aren	Aren	Aren
8	8.88790	Lontar					
1	10.7851	Lontar					
2	12.9785	Aren					
10	22.1068	Aren					
11	24.2880	Aren					
9	32.8566	Aren					
3	67.9161	Lontar					
4	104.596	Coconut					
12	144.330	Coconut					

3.4. Evaluation

Based on the hybrid technique that has been built, testing of expert needs was carried out through respondents spread across several villages in South Sulawesi - Indonesia. It is known that a confusion matrix with multifaceted views serves a fundamental role in evaluating classification performance [33]. The palm sugar images were tested to evaluate the results of hybrid HSL, GLCM, and KNN against 300 data. In this study, several palm sugar images with different qualities were tested to determine the accuracy of the classification results obtained through the hybrid HSL, GLCM and KNN approaches, as shown in Figure 7.

Based on the testing and recording results on each sample image, the classification process works well on images with a high level of sharpness, as shown in Figures 7(a) to 7(c). Meanwhile, in Figure 7(d), the classification accuracy still decreases due to the low quality of image sharpness. Test results are evaluated using the confusion matrix testing method to determine the accuracy, recall, precision and F1 score value. Table 6 shows testing data from respondents.

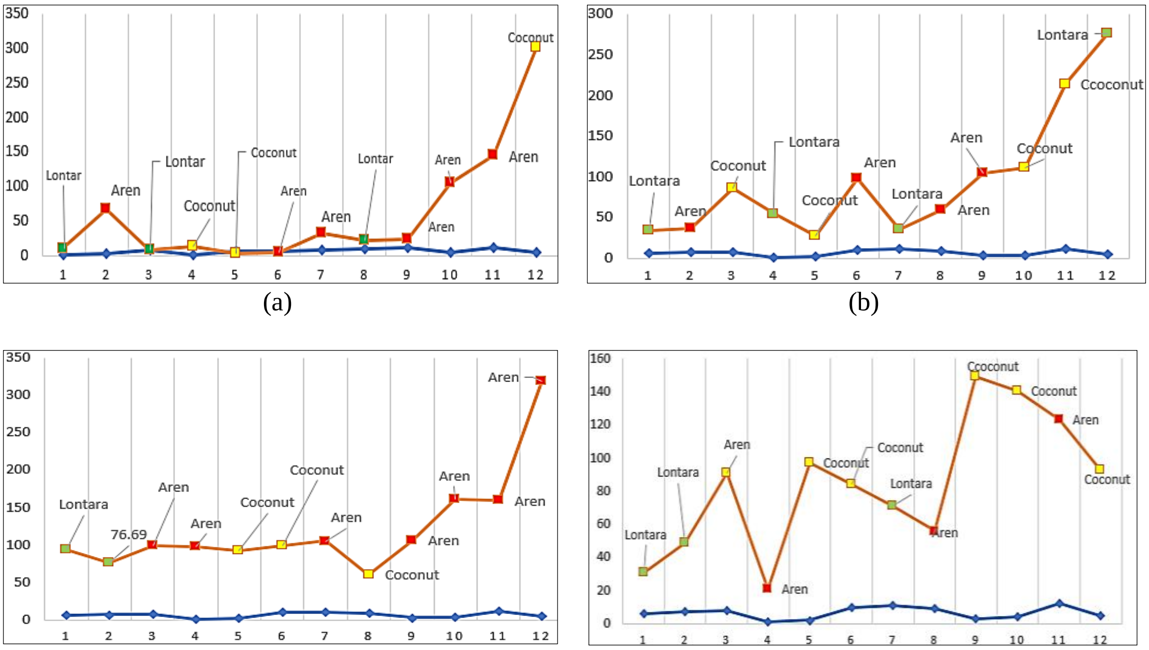
Based on the data in Table 6, a validation of the recommendation results with actual data was carried out using the confusion matrix method as follows [34]:

$$\text{Precision} = TP/(TP+FP) \tag{9}$$

$$\text{Recall} = TP/(TP+FN) \tag{10}$$

$$\text{F1 Score} = 2(\text{Precision x Recall}) / (\text{Precision} + \text{Recall}) \tag{11}$$

$$\text{Accurate} = (TP) / (TP + FP + FN) \tag{12}$$



(a) (b)

(c) (d)

Figure 7. Result of analytics palm sugar based on image in level (a) low sharpness, (b) moderate sharpness, (c) medium sharpness, and (d) high sharpness

Table 6. Testing data

	Aren	Coconut	Lontar	Total
Aren	105	1	1	107
Coconut	1	90	0	91
Lontar	2	2	98	102
Total				300

Table 7 shows an accuracy value of 0.976, which indicates that the research results perform well in classifying data. Accuracy level testing is carried out by k-fold cross-validation, dividing the subset in several iterations with one of the k subsets used as test data. At the same time, the other k-1 subset is used as training data. Then the results of the palm sugar classification assessment in the system and the actual palm sugar classification assessment results are compared using the confusion matrix shown in Table 8.

Based on eight experiments conducted with 300 test data, the highest accuracy of 0.90 on fold-1 and the lowest accuracy of 0.79 on fold-8, it can be concluded that the quality of the model built is good. The contribution of this research is to utilize the role of technology in terms of the ability to distinguish various types or qualities of palm sugar so that it can produce the application of automation systems for the introduction and classification of palm sugar in industry or used by the community.

Table 7. Result of confusion matrix

Indicator	Aren	Coconut	Lontar
TP	105	90	98
FP	2	1	4
FN	3	3	1
Precision (9)	21	22.5	19.6
Recall (10)	0.972	0.967	0.989
F1 Score (11)	1.858	1.855	1.884
Accurate (12)		0.976	

Table 8. Result of K-Fold validation

	Fold-1	Fold-2	Fold-3	Fold-4	Fold-5	Fold-6	Fold-7	Fold-8
Accurate	0.90	0.82	0.87	0.85	0.89	0.87	0.84	0.79

In this study, a comparative analysis is carried out between the proposed technique and the results of previous relevant research cases using the same algorithm and technique to know the accuracy value. A comprehensive comparison is provided in Table 9. Implication the hybrid approach through the incorporation of colour representation HSL, texture features GLCM, and KNN classification algorithms is the ease of identifying types of palm sugar to expand opportunities for the development of automated systems capable of distinguishing different types of palm sugar efficiently and accurately for the image processing-based food industry.

Table 9. Result of the comparison accuracy of relevant algorithms

Practical	Methodology	Accuracy (%)
Detecting Straight Paths in an Image [8]	HSV and PCA	87.36
Automatic Lane Detection for Driver Assistance[9]	Gradient Threshold and Hue-Lightness-Saturation	96
Human Walking Activity Recognition[10]	GLCM and LSTM	96
Textural Analysis[11]	GLCM	92
Classification of Machined Surfaces [14]	ML techniques applied to GLCM	91.3
Palm Sugar Classification	HSL, GLCM, and KNN	97.6

4. CONCLUSION

This research concludes that the result of hybrid HSL, GLCM, and KNN methodology shows the effectiveness of the proposed options for classifying palm sugar. The classification performance is highly dependent on the image quality, where parameters on the HSL and GLCM matrices strongly influence the quality of informative features. The accuracy rate of 97.6% of the palm sugar type prediction results shows that the model that was built has worked very well. The results have practical implications, such as developing automated systems to classify palm species in specific industries to benefit economics and operational efficiency. Future research directions may explore the integration of advanced machine learning techniques and real-time image processing for further improving classification performance and scalability in industrial applications.

ACKNOWLEDGEMENTS

The authors would like to express their gratitude to STMIK Profesional Makassar for providing significant financial support for this study under research scheme funding number 001/SP/IK/II/2024.

REFERENCES

[1] J. E. D. M. Paz, A. M. Dantas, D. D. D. S. Fernandes, and M. J. C. Pontes, "Classification of sugar using digital imaging and pattern recognition techniques," *Journal of Food Composition and Analysis*, vol. 125, Jan. 2024, doi: 10.1016/j.jfca.2023.105796.

[2] A. Upadhyaya and S. K. Sonawane, "Palmyrah palm and its products (Neera, Jaggery and Candy)—a review on chemistry and technology," *Applied Food Research*, vol. 3, no. 1, Jun. 2023, doi: 10.1016/j.afres.2022.100256.

[3] V. Alves, J. M. D. Santos, E. Pinto, I. M. P. L. V. O. Ferreira, V. A. Lima, and M. L. Felsner, "Digital image processing combined with machine learning: A new strategy for brown sugar classification," *Microchemical Journal*, vol. 196, Jan. 2024, doi: 10.1016/j.microc.2023.109604.

[4] J. Liu, P. Wan, W. Zhao, C. Xie, Q. Wang, and D.-W. Chen, "Investigation on taste-active compounds profile of brown sugar and changes during lime water and heating processing by NMR and e-tongue," *LWT*, vol. 165, Aug. 2022, doi: 10.1016/j.lwt.2022.113702.

[5] T. Sarkar, M. Mukherjee, S. Roy, and R. Chakraborty, "Palm sap sugar an unconventional source of sugar exploration for bioactive compounds and its role on functional food development," *Heliyon*, vol. 9, no. 4, Apr. 2023, doi: 10.1016/j.heliyon.2023.e14788.

[6] A. D. Saputro, D. V. D. Walle, and K. Dewettinck, "Palm sap sugar: a review," *Sugar Technology*, vol. 21, no. 6, pp. 862–867, Dec. 2019, doi: 10.1007/s12355-019-00743-8.

[7] F. Yan, N. Li, and K. Hirota, "QHSL: A quantum hue, saturation, and lightness color model," *Information Sciences*, vol. 577, pp. 196–213, Oct. 2021, doi: 10.1016/j.ins.2021.06.077.

[8] H. J. Kadhim and A. H. Abbas, "Detect lane line for self-driving car using hue saturation lightness and hue saturation value color transformation," *International Journal of Online & Biomedical Engineering*, vol. 19, no. 16, pp. 4–19, Nov. 2023, doi: 10.3991/ijoe.v19i16.43359.

[9] M. A. A. Noman *et al.*, "A computer vision-based lane detection technique using gradient threshold and hue-lightness-saturation value for an autonomous vehicle," *International Journal of Electrical and Computer Engineering*, vol. 13, no. 1, Feb. 2023, doi: 10.11591/ijece.v13i1.pp347-357.

[10] E. Anbalagan and S. M. Anbhazhagan, "Deep learning model using ensemble based approach for walking activity recognition and gait event prediction with grey level co-occurrence matrix," *Expert Systems with Applications*, vol. 227, Oct. 2023, doi: 10.1016/j.eswa.2023.120337.

[11] J. I. Fajardo, C. A. Paltán, L. M. López, and E. J. Carrasquero, "Textural analysis by means of a gray level co-occurrence matrix method. Case: Corrosion in steam piping systems," *Materials Today: Proceedings*, vol. 49, pp. 149–154, 2022, doi: 10.1016/j.matpr.2021.07.493.

[12] Y. Ma, R. Huang, M. Yan, G. Li, and T. Wang, "Attention-based local mean k-nearest centroid neighbor classifier," *Expert Systems with Applications*, vol. 201, Sep. 2022, doi: 10.1016/j.eswa.2022.117159.

[13] Y. Li and S. Ercisli, "Data-efficient crop pest recognition based on KNN distance entropy," *Sustainable Computing: Informatics and Systems*, vol. 38, Apr. 2023, doi: 10.1016/j.suscom.2023.100860.

[14] G. Prasad, G. S. Vijay, and R. Kamath C., "Comparative study on classification of machined surfaces using ML techniques applied to GLCM based image features," *Materials Today: Proceedings*, vol. 62, pp. 1440–1445, 2022, doi: 10.1016/j.matpr.2022.01.285.

[15] A. A. B. A. Mohammed *et al.*, "Effect of sugar palm fibers on the properties of blended wheat starch/polyvinyl alcohol (PVA) -based biocomposite films," *Journal of Materials Research and Technology*, vol. 24, pp. 1043–1055, May 2023, doi: 10.1016/j.jmrt.2023.02.027.
[16] D. Qiu, H. Liang, Z. Wang, Y. Tong, and S. Wan, "Hybrid-supervised-learning-based automatic image segmentation for water leakage in subway tunnels," *Applied Science*, vol. 12, no. 22, Nov. 2022, doi: 10.3390/app122211799.
[17] P. Sferrazza, "Grey level co-occurrence matrix and learning algorithms to quantify and classify use-wear on experimental flint tools," *Journal of Archaeological Science: Reports*, vol. 48, Apr. 2023, doi: 10.1016/j.jasrep.2023.103869.
[18] M. H. Daneshvari, E. Nourmohammadi, M. Ameri, and B. Mojaradi, "Efficient LBP-GLCM texture analysis for asphalt pavement raveling detection using eXtreme gradient boost," *Construction and Building Materials*, vol. 401, 2023, doi: 10.1016/j.conbuildmat.2023.132731.
[19] S. Zhang *et al.*, "MM-GLCM-CNN: a multi-scale and multi-level based GLCM-CNN for polyp classification," *Computerized Medical Imaging and Graphics*, vol. 108, Sep. 2023, doi: 10.1016/j.compmedimag.2023.102257.
[20] F. Gapsari *et al.*, "Analysis of corrosion inhibition of Kleinhovia hospita plant extract aided by quantification of hydrogen evolution using a GLCM/SVM method," *International Journal of Hydrogen Energy*, vol. 48, no. 41, pp. 15392–15405, May 2023, doi: 10.1016/j.ijhydene.2023.01.067.
[21] D. Das and R. Naskar, "Image splicing detection using low-dimensional feature vector of texture features and Haralick features based on gray level co-occurrence matrix," *Signal Processing: Image Communication*, vol. 125, 2024, doi: 10.1016/j.image.2024.117134.
[22] S. Adhikary and S. Banerjee, "Introduction to distributed nearest hash: on further optimizing cloud based distributed KNN variant," *Procedia Computer Science*, vol. 218, pp. 1571–1580, 2023, doi: 10.1016/j.procs.2023.01.135.
[23] N. B. Oğur *et al.*, "Detection of depression and anxiety in the perinatal period using marine predators algorithm and kNN," *Computers in Biology and Medicine*, vol. 161, Jul. 2023, doi: 10.1016/j.compbiomed.2023.107003.
[24] Y. Cui *et al.*, "AGTML: a novel approach to land cover classification by integrating automatic generation of training samples and machine learning algorithms on Google Earth engine," *Ecological Indicators*, vol. 154, Oct. 2023, doi: 10.1016/j.ecolind.2023.110904.
[25] T. Hayashi, D. Cimr, H. Fujita, and R. Cimler, "Image entropy equalization: a novel preprocessing technique for image recognition tasks," *Information Science*, vol. 647, Nov. 2023, doi: 10.1016/j.ins.2023.119539.
[26] G. Xu, X. Xu, and X. Wang, "Generalized multi-scale image decomposition for new tone manipulation," *Digital Signal Processing*, vol. 135, Apr. 2023, doi: 10.1016/j.dsp.2023.103945.
[27] C. M. Hoang and B. Kang, "Pixel-level clustering network for unsupervised image segmentation," *Engineering Applications of Artificial Intelligence*, vol. 127, Jan. 2024, doi: 10.1016/j.engappai.2023.107327.
[28] M. Hagara, R. Stojanović, T. Bagala, P. Kubinec, and O. Ondráček, "Grayscale image formats for edge detection and for its FPGA implementation," *Microprocessors and Microsystems*, vol. 75, Jun. 2020, doi: 10.1016/j.micpro.2020.103056.
[29] S. K. M. Abujayyab *et al.*, "Integrating object-based and pixel-based segmentation for building footprint extraction from satellite images," *Journal of King Saud University-Computer and Information Sciences*, vol. 35, no. 10, Dec. 2023, doi: 10.1016/j.jksuci.2023.101802.
[30] M. Faisal and T. K. A. Rahman, "Determining rural development priorities using a hybrid clustering approach: a case study of South Sulawesi, Indonesia," *International Journal of Advanced Technology and Engineering Exploration*, vol. 10, no. 103, pp. 696–719, Jun. 2023, doi: 10.19101/IJATEE.2023.10101215.
[31] M. Faisal, T. K. A. Rahman, I. Mulyadi, K. Aryasa, Irmawati, and M. Thamrin, "A novelty decision-making based on hybrid indexing, clustering, and classification methodologies: an application to map the relevant experts against the rural problem," *Decision Making: Applications in Management and Engineering*, vol. 7, no. 2, pp. 132–171, Feb. 2024, doi: 10.31181/dmame7220241023.
[32] V. Selvakumar and C. Akila, "Efficient diabetic retinopathy diagnosis through U-Net – KNN integration in retinal fundus images," *Automatika*, vol. 64, no. 4, pp. 1148–1157, Oct. 2023, doi: 10.1080/00051144.2023.2251231.
[33] J. Xu, Y. Zhang, and D. Miao, "Three-way confusion matrix for classification: A measure driven view," *Information Sciences*, vol. 507, pp. 772–794, Jan. 2020, doi: 10.1016/j.ins.2019.06.064.
[34] M. Faisal *et al.*, "Faster R-CNN algorithm for detection of plastic garbage in the ocean: a case for turtle preservation," *Mathematical Problems in Engineering*, vol. 2022, pp. 1–11, 2022, doi: 10.1155/2022/3639222.

BIOGRAPHIES OF AUTHORS

Mila Jumarlis is Bachelor of computer (S.Kom.) in informatics, Master of Computer (M.Kom.) in computer system, besides several professional certificates and skills. She is currently lecturing at the Department of Islamic communication and broadcasting State of Islamic College of Majene. She is a member of Indonesian Association of Informatics Experts (IAII). Her research areas of interest include image processing, artificial intelligent, and teaching media. She can be contacted at email: mila.jumarlis@stainmajene.ac.id.

Ida Mulyadi received her S.Kom. degree from the Department of Information Systems, Diponegara University, Makassar, Indonesia, at 2012, and M.T. degree from the Masters in Teknik Elektro, Postgraduate School, Hasanuddin University, Makassar, Indonesia in 2016. She is currently a lecturer and researcher at the Department of Information Systems, STMIK Profesional of Makassar, Indonesia. His current research is big data. She can be contacted at email: idamulyadi@stmikprofesional.ac.id.

Mirfan is Bachelor of computer (S.Kom.) in informatics, Master of Engineering (M.T.) in Electrical and Informatics, Master of Computer (M.Kom.) in computer system, and a Doctoral student in Faculty of Teacher Training and Education of Muhammadiyah University Makassar-Indonesia, besides several professional certificates and skills. He is currently lecturing with the Department of informatics Engineering at Handayani University Makassar. He also a member of Indonesian Association of Informatics Experts (IAII). His research areas of interest include image processing, artificial intelligent, and teaching media. He can be contacted at email: mirfan@handayani.ac.id.

Irmawati a bachelor of engineering in Electrical Engineering Hasanuddin University, and M.T. degree from master in information technology, Postgraduate School of Hasanuddin University. She is currently a lecturer and researcher at Politeknik Negeri Ujung Pandang, Department of Computer and Network Engineering. Her current research interests include data engineer, big data, integration system, and digital supply chain management. She can be contacted at email: irmawati@poliupg.ac.id.

Mardiah received her S.Kom. degree from the Department of Information System, STMIK Mandiri Tidore, Tidore, Indonesia, at 2011, and M.T. degree from the Masters in Information Technology, Postgraduate School, Hasanuddin University, Makassar, Indonesia in 2014. He is currently a lecturer and researcher at the Department of Computer Science, STMIK Tidore Mandiri, Indonesia. His current research interests include mobile programming, data mining, applications, and image processing. She can be contacted at email: mardiah@stmik-tm.ac.id

Muhammad Faisal received her S.SI degree from Department of Information System, STMIK Profesional, Makassar, Indonesia, at 2011 and M.T degree from the Masters in Information Technology, Postgraduate School, Hasanuddin University, Makassar, Indonesia in 2014. He is currently a lecturer and researcher at the Department of Computer Science, STMIK Profesional Makassar, Indonesia. His current research interests include mobile programming, data mining, decision support system, expert system, web applications, and image processing. He can be contacted at email: muhfaisal@stmikprofesional.ac.id.

Hairin Anisa a bachelor of computer (S.Kom.) in informatics. He is a member Indonesian Association of Informatics Experts (IAII). His research areas of interest include image processing, artificial intelligent, and instructional media. She can be contacted at email: nurilmi@handayani.ac.id.