

Different Types of Beef And Pork Using Histogram Texture Anda K-Means Clustering Method

Heribertus Himawan^{*1}, Widya Wiratama²

*Department of Informatics Engineering, Faculty of Computer, Universitas Dian Nuswantoro,
Semarang, Indonesia ^{1,2}*

*E-mail: himawan26@gmail.com^{*1}, widyawiratama15@gmail.com²*

Abstract - The largest source of protein needed by the human body comes from meat. Basically a high number of buyers and beef is much more expensive than other meats giving chances to irresponsible parties to mix beef with pork, In the Islamic religious teachings, pork is forbidden because consuming pork there is a physical, psychological and very easy to be contaminated by bacteria. Humans have the limitation of identifying beef and pork manually using the senses of human vision are still many weaknesses and less effective. In this research, meat image extraction applies histogram and clustering by applying K-Mean. By applying this method is expected to help to group the image of meat with a high degree of accuracy.

Keywords - clustering, Histogram, K-Means, Image, Meat

1. INTRODUCTION

Meat is the main protein sources of human, there are various kinds of meat to consume like beef, goat, pork and fowl [1]. However, beef has high enough protein that's why most people consume it. From those various kinds of meat, beef is one of the most expensive ones, by this situation many butchers have an idea to do deceitful way by changing beef with pork to get more profit because the texture and color are somewhat similar.

In the Islamic religious teaching pork is forbidden because there are physical, psychological harm and easy contaminated by bacteria. Alquran is Moslem scripture directly sent by Allah Subhanahu Wata'ala and revealed by prophet Muhammad through Jibril using the Arabic language. Alquran is the last revelation from Allah SWT as the guidelines for Moslem people in the world. Recommendation of Islamic people about nutrient consumption lead to physical hygiene and guarantee the health. This was discussed 4 surahs (chapter) forbid pork in Alquran and the importance of physical and psychological harm to consume pork because pork is very easy contaminated by bacteria. Proof of its harm has found in many laboratory types of research. So, those researches proof to not consume pork because of its harmful [1]. Because the cheap price of pork and the amount of beef consumption, make many butchers taking advantage by mixing beef with pork to get more profit. According to that case, the consumer should know how to differentiate which one is beef and pork by using the newest technology. Today, there are many application with high technology to introduce pattern, document classification, economy study, digital image processing [2]. By using application of digital image processing we can make our own computer like human vision which is able to different things around, so this can help people to differ beef and pork. By using K-Means method and feature of histogram texture would be more effective to optimize data collecting and producing a picture. We can collect the data by its characteristic in clustering method, object with dominant similarity level closer to another object so it belongs to the same cluster. K-means

method is clustering that makes data partition in one cluster or more based on the distance to centroid. There are many methods used to classify image and identify its cluster which uses K-means method. By using K-means algorithm the cluster classification would not be hard to do and the opportunity of success [2], besides for processing digital image of K-means algorithm, this also used for making applications like introducing pattern until medical application. According to its proof K-means algorithm used to monitor students' academic score, evaluate the performance of the university by formula different space Euclidian [3,4,5]. Feature extract method of histogram texture is a static method to get texture feature. The excellence of the texture feature method of histogram texture is very useful to know the limitation to determine the background of the object and to determine the percentage of intensity texture composition as image identification [6,7,8]. There are some feature in this method, such as deviation, energy, skewness, average intensity, smoothness and entropy [5,9,10].

2. RESEARCH METHOD

This research was starting with collecting some online journals and materials which reviewed K-means and histogram texture feature about pork and beef image. The activity happened to support this research after all the materials have been got. Later, this research needed data set as supporting sample. The data of meat image found by buying beef and pork, after obtained meat image data then continue the process to change image data into data needed in this research [11,12].

1. Image Data Source

This research will use meat image data like beef and pork obtain, then the data will be processed to change image data into data needed in this research.

Table 1. Data Type and Origin

No	Data type	Origin	Mount Data
a.	Beef image	Market Kobong, Semarang	50
b.	Pork image	Market Kobong, Semarang	50

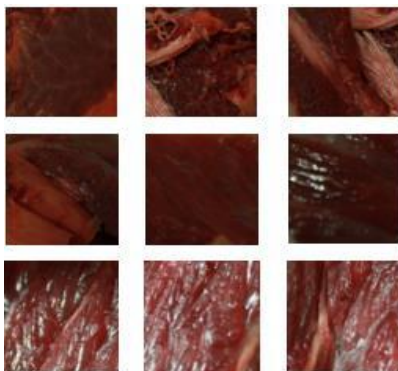


Figure 1. Beef image.

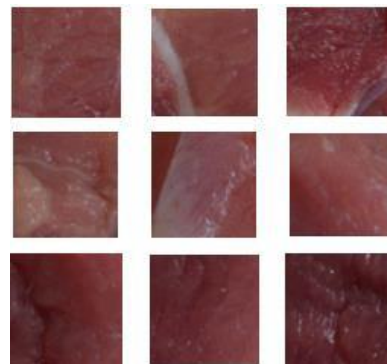


Figure 2. Pork Image

2. Preprocessing Data

There some process to do after obtaining all the data. Here some process to do [13]:

1. Processing beef and pork image into a grayscale image with resolution 400x400 pixel as

- a research source.
2. Using the texture feature extraction method of beef and pork image by histogram base later it would get texture feature data in form of deviation, energy, skewness, average intensity, smoothness, and entropy.
3. The next is processing clustering with K-means clustering.
4. Then it would be identified the kinds of beef and pork from image data result have processed before.

3. Extraction result data

In the extraction result which obtains a feature of deviation, energy, skewness, average intensity, smoothness and entropy into numeric data will be used to count K-means algorithm method with Euclidean Distance count for beef and pork clustering in order to be identified soon [14,15].

Table 2. The example of extraction result pork and beef image texture feature

No	Mu	<i>Deviation</i>	<i>Skewness</i>	<i>Energy</i>	<i>Entropy</i>	<i>Smoothness</i>
1	41,8881	27,16798	0,344749	0,019752	4,247704	0,0112
2	33,91579	21,80658	0,116431	0,027894	3,841348	0,0073
3	81,38337	50,65092	1,291804	0,006193	5,199383	0,038
4	104,6485	47,00503	0,326006	0,0062	5,184319	0,0329
...
97	95,08624	27,41802	0,099267	0,01179	4,549574	0,0114
98	92,72325	28,02915	-0,07626	0,012258	4,544095	0,0119
99	91,41981	31,85261	0,152434	0,011245	4,622468	0,0154
100	66,45654	28,10092	0,201652	0,016333	4,386627	0,012

4. Proposal

There are some procedure and method applied in this research to finish the recent problems. In clustering beef and pork, the researcher applied some methods they are image normalization method, feature extraction method, a classification method, and Davies Bouldin Index by combining the fourth method and some procedures which can solve the problems.

Here are some procedures:

In this research, there are 3 application methods used, the first is the image normalization method, it is original data of pork and beef image changed into greyscale form with 400x400 pixel resolution.

The second method is the feature extraction method, by applying feature extraction based histogram to obtain deviation, energy, skewness, average intensity, smoothness, and entropy. The third method is K-Means algorithm method to clustering pork and beef image that has run through both methods before able to identify.

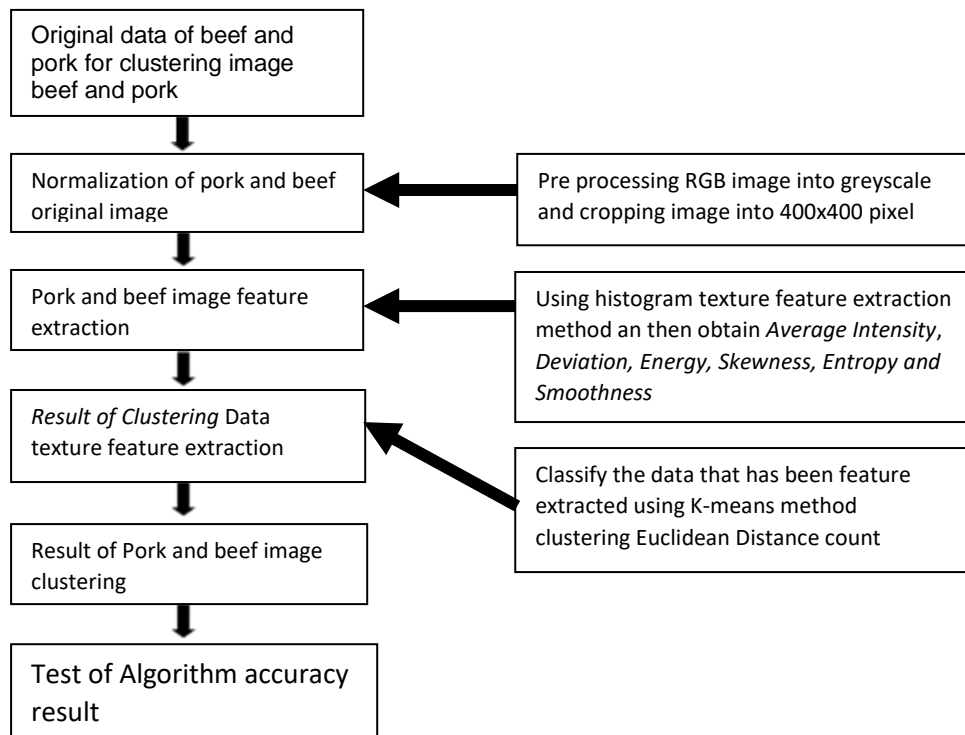


Figure 3. Research method has been applied

3. DISCUSSION AND RESULT

This research used data of raw pork and beef which is not yet processed, there are 100 data with resolution 400x400 pixel obtained from direct butcher in traditional market, it took process taking an image in order to get data into meat image data. Some image data has been used are beef with label 1 as 50 data and pork with label 2 as 50 data. Then, the whole data through normalization and cropping or picture cutting with Adobe Photoshop. All the data are 100 images will use as data training and data testing with 50 data with label 1 and 50 data with label 2.

3.1. Pre-processing

The first image data would run into pre-processing later it would be used to obtain data in this research. In the beginning, the image is in the cropping process into RGB image with 400x400 pixel resolution with the beef image as 50 data and pork image as 50 data. Then, the image will continue to normalization process into greyscale. Because in the method of Histogram Feature Texture extraction needs greyscale to look for the greyscale value that is average intensity or mean.



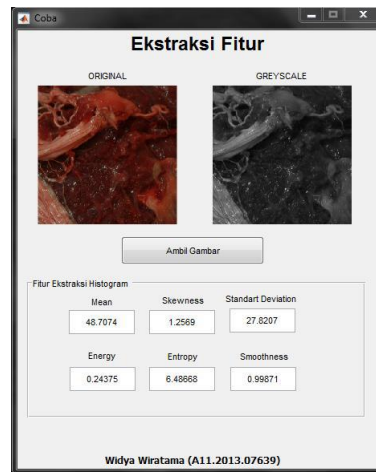


Figure 5. Implementation on system

3.3. K-Means Clustering

Classification data of pork and beef image result by applying clustering K-means algorithm with Euclidean Distance formula. The first step in clustering K-means algorithm is to make sure how much cluster amount. In this opportunity, there are 2 groups of the cluster. Whereas in every centroid defined randomly. The example of pork and beef image texture feature extraction result with 100 data set from table 1, which will define the centroid. Here, centroid will use to count clustering K-means algorithm method with Euclidean Distance formula can be seen on table 2 as follows:

Table 2. Centroid used

	Mu	Deviation	Skewness	Energy	Entropy	Smoothness
C1	112,538	52,52958	0,897023	0,006088	5,244121	0,040708
C2	34,21076	28,6049	0,445507	0,017997	4,354677	0,012427

From example the result of pork and beef image texture feature with 100 dataset in table 1 and centroid data on table 2 has been determined would be used to define D spot (center spot) in clustering K-means algorithm count by Euclidean Distance formula.

After counting by using Euclidean distance formula, the result as in table 3:

Table 3. The result of counting Euclidean distance in the first iteration

No	D1	D2	Cluster	Group
1	75,073	7,812	7,812	2
2	84,427	6,832	6,832	2
...
98	54,492	30,605	30,605	2
99	57,747	26,682	26,682	2

From distance count iteration 1 above get result in table 3., cluster is determined from the lowest value between the first center spot (D1) and the second center spot (D2), next group value can be defined from cluster value comes from which one is center spot, for example, cluster value comes from the first center spot value (D1) it means that group value has value 1, conversely, if cluster value comes from the second center spot value (D2) so the group value is 2. In a group, value 1 means beef label and value 2 is pork label, however, group value is still temporary until the new centroid value is equal with the old centroid value, so the group value

will become final value from counting then do the counting until old centroid is equal with the new centroid.

3.4. Davies Bouldin Index

After counting clustering K-means algorithm with Euclidean Distance formula so the next process is counting distance formula validity by using Davies Bouldin Index. Step to count it as follows:

a. Find out SSW

$$SSW = \sum_{m=1}^m ((X1, C1)) + (X2, C1) + (X3, C1) + (X4, C1) + (X5, C1) + \dots + (, Cn,)$$

Where m is number of class in every iteration and d is the total summary of all classes.

b. Find out SSB

$$SSB = (C1, C2) = \sqrt{(XC1 - XC2)^2 + (YC1 - YC2)^2}$$

Where SSB is root from summary from centroid

c. Find out R

$$R = \frac{SSW1 + SSW2}{SSB}$$

Where R is the result of summary from SSW then divided by SSB

Table 4. The result of summary from SSW then divided by SSB

	Iteration 1	Iteration 2	Iteration 3	Iteration 4	DBI	Presentase
SSW1	26.65979081	14.57668293	14.5520227	14.47889653	0.65436	65%
SSW2	23.50681432	12.66070073	12.55658949	12.57624789		
SSB	192.9537709	209.0375908	205.1090103	205.1090103		
R	0.259992872	0.130298974	0.132166852	0.131906172		

4. CONCLUSION

This research shows how high the percentage accuracy from seeking of Euclidean Distance formula in clustering K-Means method to differ raw meat, in order to identify beef or pork, through histogram texture feature extraction, so can be concluded that Euclidean Distance as distance seeking formula can be used in further research. The more formulas applied, the more apparent different from every image. By using result 6 features such as average intensity, deviation, energy, skewness, smoothness, and entropy can be processed by applying many methods, one of them is K-means clustering. The high method has been applied is 65%.

Suggestion for the further researcher related to this research are:

- Using more sample data, whether they give effect to the accuracy result or not.
- Addition kinds of meat to know whether the result of distance seeking formula will be different or not.
- It can be developed into a system in order to be more useful for users.

REFERENCES

- [1] M. A. Sheikhi dan M. D. Firoozabadi, Pork Meat From The Viewpoints of Quran and Medical, *World Journal of Pharmaceutical Research*. 2015; . IV(8), pp. 1-10.
- [2] B. Venkateswarlu dan P. G. Raju, Mine Blood Donors Information through Improved K-Means Clustering, *Cornell University Library*, 2013.
- [3] N. Dhanachandra, K. Manglem dan Y. J. Chanu, Image Segmentation using K -means Clustering Algorithm and, 2015; *Eleventh International Multi-Conference on Information Processing (IMCIP-2015)*. pp.764 – 771

- [4] A. Y. Isyanto, Measurement of Farm Level Efficiency of Beef Cattle Fattening in West Java Province, Indonesia, *Journal of Economics and Sustainable Development*, 2013; IV(10)
- [5] O. J. Oyelade, O. O. Oladipupo dan I. C. Obagbuwa, Application of k-Means Clustering algorithm for prediction of Students' Academic Performance, *arXiv preprint arXiv: 1002.2425* (2010).
- [6] M. GONDEKOVA, I. BAHLEKA, K. ZAUJEC and P. POLÁK, Evaluation of Quality Acceptance ofnCow's Meat in Respect to Certain Social Aspects of Consumers, *Slovak J. Anim. Sci.* 2014; 47 (2).
- [7] S. W. Martinez, Pork quality and the role of marketing contracts: a case study of the US pork industry, *USDA – Agricultural Economic Report*, 2012; 835 pp. 302- 317.
- [8] A. Kadir dan A. Susanto, Pengolahan Citra dan Teori Aplikasi, Andi Ofset, Yogyakarta, 2013.
- [9] P. Mohanaiah, P. Sathyanarayana dan L. GuruKumar, Image Texture Feature Extraction Using GLCM Approach, *International Journal of Scientific and Research Publications*, 2013, 3(5).
- [10] Anas, Abdur Rahman and T. Lakshmi Sravanthi. "Image Classification in Content-Based Image Retrieval Systems Based on First Order Color Histogram Features." (2014)..
- [11] O. Marques dan B. Furht, Content Based Image and Video Retrieval, *Multimedia Tools and Applications*, 2002; 17, 21–50.
- [12] T. Lijalem, M. Beyan dan S. Banerjee, Quality of beef, chevon and mutton at Hawassa,Southern Ethiopia, 2015, *World Journal of Multidisciplinary and Contemporary Research*, 2015; 1(2), 301-306.
- [13] M.F. Qamar and I. Raza, Scientific Evidences that Pig Meat (Pork) is Prohibited for Human Health, *Scientific Papers. Series D. Animal Science*, 2012; Vol. LV.
- [14] K. Adi, S. Pujiyanto, O. D. Nurhayati dan A. Pamungkas, Beef Quality Identification using Color Analysis and K-Nearest Neighbor Classification, *Hindawi Journal of Food Quality*, 2015; 2017.
- [15] Sridhar and Gowri, Color and Texture Based Image Retrieval, *ARPN Journal of Systems and Software*, 2012; 2(1).