Beef Image Classification using K-Nearest Neighbor Algorithm for Identification Quality and Freshness

IOP Conf. Series: Journal of Physics: Conf. Series **1179** (2019) 012184 doi:10.1088/1742-6596/1179/1/012184

Beef Image Classification using K-Nearest Neighbor Algorithm for Identification Quality and Freshness

S Agustin[1], R Dijaya[2]

[1]Departement of Informatics Engineering, Faculty of Engineering, Universitas Muhammadiyah Gresik, Gresik, Indonesia.
[2]Departement of Informatics, Faculty of Engineering, Universitas Muhammadiyah Sidoarjo, Sidoarjo, Indonesia.

[1]soffiana@umg.ac.id, [2]rohman.dijaya@umsida.ac.id

Abstract. Nowadays many traders are cheating in selling meat that is not feasible for consumption to get a greater profit by mixing good meat with meat that is not feasible for consumption. So cause anxiety for the community because it has dangerous content. One way to help people in selecting meat through image processing can recognize objects. The purpose of this research is to know the quality of beef which is good and feasible to be consumed using the co-occurrence matrix to classify the meat image with K-NN algorithm. This research can be utilized to be able to distinguish the types of meat based on colour and texture. The data used in this research are 60 image data consisting of 30 images of fresh meat and 30 images of rotten meat. The classification process uses test data in order to distinguish the types of fresh meat and rotten meat. The mean feature value of this method with the feature features of the highest feature extraction is the homogeneity value feature. The results showed that the performance of the system using the KNN method to identify the quality of meat based on colour and texture can detect the type of beef.

1. Introduction

Beef is one type of food that comes from animals that is quite liked by the community to meet the needs of animal protein. This is because meat has been known as one of the perfect food ingredients, because it has complete nutrients and is needed by the body such as protein, energy, water, minerals and vitamins, and has a good taste and aroma. Rotten meat or "daging glonggong" in Bahasa Indonesia has higher levels than fresh beef. The amount of beef that is not suitable for consumption has been widely circulated in the market, causing the selection of good beef is not easy. The characteristics of fresh beef are bright red meat, not pale, elastic, not sticky in the hand and distinctive scented. In order to determine the quality of beef is done by comparing the actual meat represented by each meat image [1]. One way to determine the level of freshness of beef is by processing images with an 8-bit digital data approach on each base color[2]. In the development of quality control computer vision freshness is a cross-approach that aims to estimate colors, morphological features and surfaces that are directly related to food quality and safety characteristics[3]. The aim of this study was to find the quality of beef that is good and can be done to consume using co- occurrence matrix to classify the image of meat using the KNN algorithm. It is expected that this research can be used to find the quality of beef

that is suitable for consumption through image processing using the KNN (K-Nearest Neighbors) method based on the object's HIS color space and GLCM texture of beef image.

2. Related Work

2.1. Beef

Beef is one type of raw food favored by the wider community throughout the world including Indonesia. According to the ministry of beef, health has an energy content of 207 kilocalories, 18.8 grams of protein, 14 grams of fat, 11 milligrams of calcium, 170 milligrams of phosphorus, and 3 milligrams of iron. In addition, beef also has vitamin A, 0.08 milligrams of vitamin B1 and 0 milligrams of vitamin C. Many of the nutritional content in beef and of course is very beneficial for the human body if beef is consumed.one type of raw food favored by the wider community throughout the world including Indonesia. According to the ministry of beef health has an energy content of 207 kilocalorie, 18.8 grams of protein, 14 grams of fat, 11 milligrams of calcium, 170 milligrams of phosphorus, and 3 milligrams of iron. In addition, beef also has vitamin A, 0.08 milligrams of vitamin B1 and 0 milligrams of vitamin C. Many of the nutritional content in beef and of course is very beneficial for the human body if beef is consumed[4].

2.1.1. Fresh Beef.

Fresh meat is new meat which is slaughtered without any treatment. Fresh meat has a water content of 75%, protein 19%, and fat 2.5%. Some factors that can be used as a reference in choosing good fresh meat include[5]:

1. Fresh meat has a bright red and shiny colour.
2. The typical smell of fresh meat is not sour/rotten.
3. Fresh meat has a solid texture, is springy and not stiff, and if pressed, the massage is quickly returned to its original position.
4. Fresh meat is not slimy if it is held it does not feel sticky in the hand and feels wet.

2.1.2. Rotten Beef.

Beef that is slimy or sticky and has a bad odor can be a sign of meat being rotten. Meat that is not safe can endanger the health of humans who eat it (consumers). Some characteristics of meat that are not suitable for consumption include[6]:

1. A rancid or unpleasant odor
2. The texture of the meat is slippery and damaged
3. Bluish even black
4. Slimy
5. Pale red meat

2.2. Gray Level Co-Occurance Matrix (GLCM)

GLCM is a matrix for storing the frequency value of the difference in brightness between one pixel and the surrounding pixels that occur in an image[7]. GLCM then stores information about gray intensity between two pixels separated by q as direction and d as distance. Where the direction of q is expressed at an angle of θ ° (00, 450, 900, 1350). The distance d represents the distance between two pixels that will be reviewed. For more details about direction and distance shown in figure 1. In the figure 1 the pixel in the middle with the symbol (x) is a reference. Pixels 1,2,3, and 4 have the same distance from pixels (x) that is d=1. Pixels 1 have a direction q =00, Pixels 2 has direction q =450, Pixels 3 has direction q = 900, and Pixels 4 has a direction q = 1350.

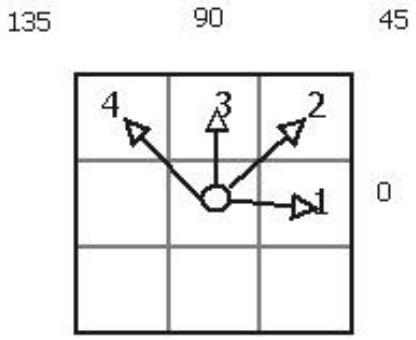


Figure 1. Direction and distance in GLCM

To calculate GLCM, the first thing to do is to decide the distance and direction you want to check. Then the next step forms the matrix framework that will be used. The framework matrix will present the range (distance) intensity of the gray degree that the image has. Then this framework matrix will be filled with the co-occurance information of the pixels in the image. After the adjacent matrix is complete, the values in each element of the matrix must be normalized. The trick is to divide each element in the matrix with the total number of each element. Normalizing each element in the neighboring matrix can be treated as a probability function because the number of elements will be equal to one. The purpose of the normalization is to uniformity the GLCM features to the same scope[8].

2.3. *Model Hue Saturation Intensity (HSI)*

The HSI model is a color system that detects the workings of the human eye. HSI combines information, both color and grayscale from an image. HSI comes from the Hue, Saturation, Intensity[9]. Where Hue describes pure colors like red, blue, or yellow. Saturation describes the degree to which pure colors are softened in white. The following is an explanation of each attribute of the HSI model[10]. Intensity / brightness / luminance that is an attribute that states the amount of light received by the eye regardless of color. The range of values is between dark (black) and bright (white). Hue that is the attribute that states the number of true colors such as red, yellow, and violet. Used to distinguish colors and determine the level of redness (greenness). Hue is associated with the wavelength of light. If we call red, violet or yellow, we actually specify the hue. Saturation is y state the level of purity of the color of the light that indicates how much white is given to the color. Like the red color, 100% saturated color (saturates color), while the pink color is red with a low level of saturation (because there is a white color in it). So, if the hue states the actual color, then saturation states the color size get the value of the HSI color model value the conversion of the RGB to HSI color model involves parameters as input data such as red, green and blue signals for each pixel and three other parameters such as hue, saturation and intensity as output. RGB to HSI color model transformation functions to convert color into a form that is more suitable for image processing[11].

2.4. K Nearest Neighbor (KNN)

The K-Nearest Neighbor method is a method used to classify objects, based on learning data that is close to the object, according to the number of their closest neighbors or k values. The proximity or distance of the neighbor is usually calculated based on Euclidian distance with the following equation.

$$d(x,y) = \sqrt{\sum_{j=1}^{n}(x_j - y_j)^2} \quad \text{.....(1)}$$

with:

d(x,y)	*: euclidien distance between vector x and y*
d	*: distance testing data to training data*
X_j	*: testing data-j, with j = 1,2,...,n*
Y_j	*: training data-j, with j = 1,2,....,n*
n	*: amount of feature.*

The K-NN classification is done by looking for the closest neighbors from the test data and choosing the class with the most members[12]. The steps for K-NN classification are as follows:

1. Determine the parameter K (nearest neighbor).
2. Calculate the square of the Euclidian distance on each object against the given sample data.
3. Sort these objects into groups that have euclantic distances from the smallest.
4. Collecting y categories (nearest neighbor classification)
5. By using the Nearest Neighbor category, the most majority can predict the value of the instance query that has been calculated.

3. Methodology

In developing a meat image classification system based on digital images, a system design is needed to provide a general overview of the process of making the system to be built so that it can solve existing problems. The stages of developing the system are as follows:

3.1. Pre-processing

The initial data processing starts from inputting the RGB image, then converts it to double, to separate RGB channels based on RGB images (Red, Green and Blue)[13]. Then normalize RGB so that it produces the mean value of RGB normalization. then to get the color value, the process of determining the color reference starts from RGB to HSI then select channel H / S / I, channel view H / S / I. in addition to the process, it can also start from conversion to grayscale then grayscale to binary after that it will segment and then enter the channel H / S / I. In the process of taking texture values there are several processes, namely: first, the input image is an RGB image. The image will be converted to a grayscale image to get a gray color image. After that the grayscale image will be converted to binary image. Then the thresholding process will be carried out, namely the separation of pixels based on the gray level owned. Then the next process will be image segmentation or edge detection process on an object in the image to increase the appearance between the boundary line of an area or object in the image.

3.2. Feature Extraction

In the process of taking texture values there are several processes, namely: first, the input image is an RGB image. The image will be converted to a grayscale image to get a gray color image. After that the grayscale image will be converted to binary image. Then the thresholding process will be carried out, namely the separation of pixels based on the gray level owned. Then the next process will be image segmentation or edge detection process on an object in the image to increase the appearance between the boundary line of an area or object in the image. Then proceed by determining the distance and direction they have and determining the range obtained by forming a co-ocurence matrix by calculating the frequency of occurrence of pairs of gray values of neighboring pixels at the specified distance and direction, the final step is to calculate the extraction of GLCM features, energy, entropy , contrast, homogenity, and corelation[14].

4. Result

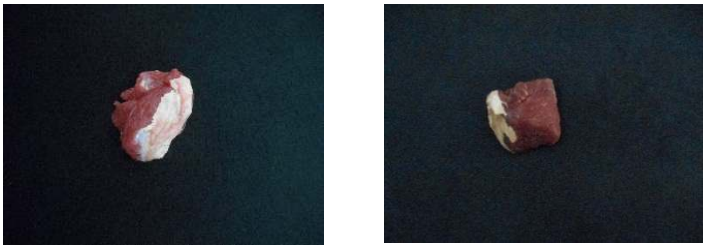
4.1. System Implementation and Testing

The initial data processing

Image data that will be tested in this research are:

a. Images of fresh meat and rotten meat
b. Picture of captured meat using a digital camera ccd (20.1 mega pixels) and use cardboard as background color (background)
c. Image with a size of 640 x 480 pixels.

The image data used is in table 2, there are 60 images consisting of 30 images of fresh meat and 30 images of rotten meat. The training image used is number 1-30 and the test image used is number 31-60 as shown in figure 2.



(a) (b)

Figure 2. (a) Fresh Meat, (b) Rotten Meat

4.2. Testing Value

The results of feature extraction using the GLCM method (Gray Level Co-Occurance Matrix) produce feature values consisting of (x1, x2, x3, x4), where x1 is contrast, x2 is correlation, x3 is energy, and x4 is homogeneity. Feature extraction results are shown in table 1 and the test results from the KNN algorithm with different K values are shown in Table 2.

Table 1. Sample of feature value

File	Class	X1	X2	X3	X4
Meat1	Fresh meat	0.013376	0.85737	0.90655	0.99331
Meat2	Rotten meat	0.011517	0.95532	0.80949	0.99424

Table 2. K Value testing

K Value	Accuracy
K1	91.6%
K3	91.6%
K5	90.0%

5. Conclussion

The results showed that the performance of the system using the KNN method to identify the quality of meat based on colour and texture can detect the type of beef and the amount of accuracy obtained is good enough but needs to be further developed and optimized. In order to improve the classification performance in the same case, other pre-processing approach and other feature extraction methods are needed. Other classification algorithms based on neurons, genetics or evolution can be tried to optimize the classification results of these features.

6. References

[1] Adi K, Pujiyanto S, Nurhayati O D and Pamungkas A 2015 Beef Quality Identification using Color Analysis and K-Nearest Neighbor Classification *International Conference on Instrumentation, Communications, Information Technology, and Biomedical Engineering (ICICI-BME)* pp 180–4

[2] Trientin D and Hidayat B 2015 Beef Freshness Classification by Using Color Analysis , Multi-wavelet Transformation , and Artificial Neural Network *International Conference on Automation, Cognitive Science, Optics, Micro Electro-Mechanical System, and Information Technology* pp 181–5

[3] Grassi S, Casiraghi E and Alamprese C 2018 Fish fillet authentication by image analysis *J. Food Eng.* **234** 16–23

[4] Crichton S O J, Kirchner S M, Porley V, Retz S, Gersdorff G Von, Hensel O and Sturm B 2017 Classification of Organic beef freshness using VNIR hyperspectral imaging *Meat Sci.*

[5] Eduardo K, Nubiato Z, Mazon M R, Silva D, Calkins C R and Naganathan G K 2018 A bench-

top hyperspectral imaging system to classify beef from Nellore cattle based on tenderness *Infrared Phys. Technol.*

[6] Santoso H A 2015 Aplikasi Pendeteksi Tingkat Kesegaran Daging Sapi Lokal Menggunakan Ekstraksi Fitur Warna Dengan Pendekatan ...

[7] Pawening R E 2015 Classification of Textile Image using Support Vector Machine with Textural Feature *International Conference on Information, Communication Technology and System(ICTS)*

[8] Tahir M 2018 Pattern analysis of protein images from fluorescence microscopy using Gray Level Co-occurrence Matrix *J. King Saud Univ. - Sci.* **30** 29–40

[9] Li G A O 2010 Color Edge Detection based on Mathematical Morphology in HSI Space *International Conference on Computer and Information Application*（pp 5–8

[10] Haribabu M, Bindu C H and Veera K 2016 A Secure & Invisible Image Watermarking Scheme Based on Wavelet Transform in HSI color space *Procedia - Procedia Comput. Sci.* **93** 462–8

[11] Dijaya R, Suciati N and Herumurti D 2016 Kombinasi Fitur Bentuk , Warna dan Tekstur untuk Identifikasi Kesuburan Telur Ayam Kampung Sebelum Inkubasi 205–14

[12] Xia M, Lu W, Yang J, Ma Y, Yao W and Zheng Z 2015 A hybrid method based on extreme learning machine and k-nearest neighbor for cloud classification of ground-based visible cloud image *Neurocomputing* **160** 238–49

[13] Kaur D and Kaur Y 2014 Various Image Segmentation Techniques: A Review *Int. J. Comput. Sci. Mob. Comput.* **3** 809–14, date accessed: 18/05/2016

[14] Saban O 2018 Application of Feature Extraction and C...sing GLCM, LBP, LBGLCM, GLRLM and SFTA.pdf *Procedia - Procedia Comput. Sci.* **132** 40–6