ELINVO (Electronics, Informatics, and Vocational Education), November 2023; vol 8(2): 156-163
ISSN 2580-6424 (printed), ISSN 2477-2399 (online) DOI: https://doi.org/10.21831/elinvo.v8i2.64810

Classification of Beef and Pork Images Based on Color Features and Pseudo Nearest Neighbor Rule

Ahmad Awaluddin Baiti[1,2], Muhammad Fachrie[3*], Saucha Diwandari[3]
[1]Department of Electrical Engineering, Southern Taiwan University of Science and Technology
[2]Department of Electronics Engineering Education, Universitas Negeri Yogyakarta
[3]Faculty of Science and Technology, Universitas Teknologi Yogyakarta
*E-mail: muhammad.fachrie@staff.uty.ac.id

ABSTRACT

This research is motivated by the need for halal foods in Muslim society with the purpose of avoiding non-halal foods, such as pork, that are sold in the market. Although beef and pork basically have different characteristics, not all Muslims know the differences. Moreover, people nowadays sell beef mixed with pork to obtain more profits. Hence, this paper proposed the implementation of the Pseudo-Nearest Neighbor Rule (PNNR) in classifying images of beef and pork slices based on color features. Based on the image dataset that has been collected, the very significant difference that can be identified visually between beef and pork is the color. The color features were extracted from the image using a color histogram from two different color channels, RGB and HSV. As the result, PNNR that used color features from the RGB channel achieved up to 87.43% accuracy, while using the HSV channel, it can reach up to 93.78% of accuracy. Additionally, this paper evaluates the stability of the proposed method by assessing the variance of classification accuracy across different values of k. It is also noticed that PNNR's performance is relatively consistent for various values of k compared to the traditional kNN algorithm.

Keywords: pseudo nearest neighbor rule, classification, color features, beef and pork, halal food

INTRODUCTION

A Muslim has obligation to eat only the halal foods, although in certain emergency eating non-halal foods is permitted to maintain his life. Generally, food of animal originating needs more concern since there are several restrictions about which animal that cannot be consumed by Muslim. Pork is one of the foods that is prohibited. In numerous Ayahs, the Quran specifies precisely which meals are considered haram (unlawful) and halal (lawful). For example, Surah An-Nahl 16:115 lists the following foods as banned to Muslims: blood, swine flesh, and dead animals. There some health impact for human if they eat pork because inside of the meat and the behaviors of the pork life [1] Unfortunately, people in certain area, even in Muslim-populated country, sell the beef mixed with pork slices, so that people difficult to recognize whether the meat is beef or pork. Even though beef and pork basically have different characteristic, but not all Muslim knows the difference.

While the differences in pork and beef color can be discerned by the human eye under specific circumstances, the application of computer-based systems offers significant advantages. Firstly, in scenarios involving mixed meats or complex food products, distinguishing color disparities can pose challenges to human perception. Computer-based systems aid in surmounting this challenge by providing an objective analysis grounded in robust color data. Secondly, the utilization of computer-based systems enables automation and scalability in the identification process, which proves particularly advantageous in large-scale food industries with stringent quality control requirements. Thus, the adoption of computer systems can enhance the reliability and efficiency of halal food recognition.

Several research have been conducted to develop systems that can recognize or classify between beef and pork. There are two different approaches that were used in previous work, that is image-based and aroma-based approach. An aroma-based approach usually utilizes a certain sensor that can sense and analyze any aroma. From [2]–[4] it can achieve good performance when using appropriate classification technique, even though in other work it just have quite good performance as in [5]. Compared to aroma-based approach, image-based is much simpler because it does not need any certain sensor instead of a camera to capture the meat slice.

In image-based approach, many of previous works used Gray Level Co-occurrence Matrix (GLCM) to extract the image feature as in [4], [6], [7] .This texture-based feature extraction is working by describing the characteristic of object based on the pattern repetition in the adjacent pixels [8]To apply this method, the color image should be converted into grayscale format [9]. GLCM shows good performance in many cases, but this algorithm is complex enough due to several mathematical equations that must be calculated. Several works also utilized other texture-based feature extraction technique such as Local Binary Pattern (LBP) [10], [11] and Invariant Moment [12]. Beside the texture-based feature, the other technique that also used is color feature [10], [11], [13], [14] e.g., RGB and HSV, which quite simple but can give high performance when well configured. Interestingly, color features, even much simpler than GLCM, can achieve as good as GLCM performance, even better in some works.

The use of color-based features is motivated by the significant visual difference between beef and pork, where the red color in beef slices is much darker compared to pork slices as provided in dataset section. Unfortunately, several works that used color features did not have optimum configuration which leads to unoptimized classification performance. In the previous works, the color feature was only represented by single mean value of each color channel, e.g., red, green, and blue (RGB) or hue, saturation, and value (HSV). Besides, most of the previous works that is mentioned above used small dataset that leads to performance bias. Therefore, this research aims to utilize and to improve the color-based feature to classify the image of beef and pork slice using larger dataset than previous works. The improvement of the color-based feature in this work is conducting by dividing the color histogram in every color channel (RGB and HSV) into three equals area that represent composition of the dark, fair, and bright pixels in each color channel and then summarizing the value by adding up the color probability in every area. This work also applied the Cross Validation technique in evaluating the system to obtain the comprehensive performance evaluation of the proposed system.

METHODS

A. *Dataset*

Different to previous works that was mostly conducted using small number of image dataset (less than 150 images), this research used primary dataset with total of 744 images containing 312 images of beef slices and 432 images of pork slices with size of 500 x 400 pixels each. To obtain the dataset, every meat slice which was bought from the local market was captured using smartphone camera that was positioned about 15 – 20 cm from meat surface. There are total of 124 raw images consisting of 52 and 72 raw images of beef slices pork slices respectively. As described in Figure 1, each raw image that is captured has a size of 4000 x 1800 pixels, then it is resized into 1000 x 450 pixels to reduce the computational cost. To obtain a lot of images in proportional size, each raw image is sampled into 500 x 400 pixels which results in 6 smaller images as illustrated in Figure 1. This dataset preparation step results in 312 sample images (52 raw images x 6 samples) of beef slices and 432 sample images (72 raw images x 6 samples) of pork slices. Figure 1 shows several images used in this work that were obtained from the sampling process.

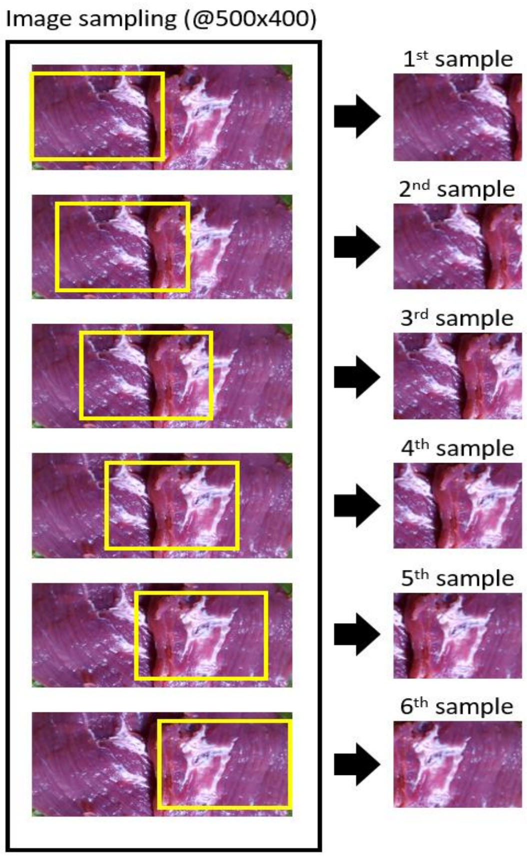
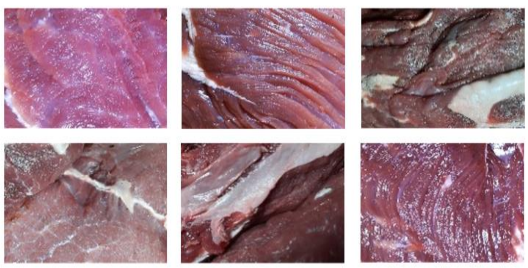


Figure 1. Process of obtaining sample images for dataset from raw image



(a)

(b)

Figure 2. Sample of beef (a) and pork (b) images used in the research

From each image, a set of features was extracted to obtain the pattern. Based on images in Figure 2, the main difference between beef and pork slices is color, where beef slices look brighter, and their color is rather darker than pork slices. Hence this research focused on extracting the color feature from the images to see how good color feature is in differentiating the two kinds of meats. The extracted feature is obtained from color histogram which resulted in nine features for each image.

B. *Feature Extraction*

The feature extraction is performed by constructing the color histogram from RGB and HSV channel, hence the input image should be split into two different color channels. After the channel separation, three histograms with 256 lengths on each channel, i.e., red, green, and blue histogram in RGB channel, and hue, value, and saturation in HSV channel were constructed. Then, the 256 features from each histogram are summarized into three features to reduce computational cost and eliminate unnecessary information. Figure 3 shows the main workflow of feature extraction process.

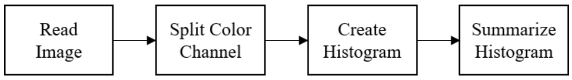


Figure 3. workflow in feature extraction

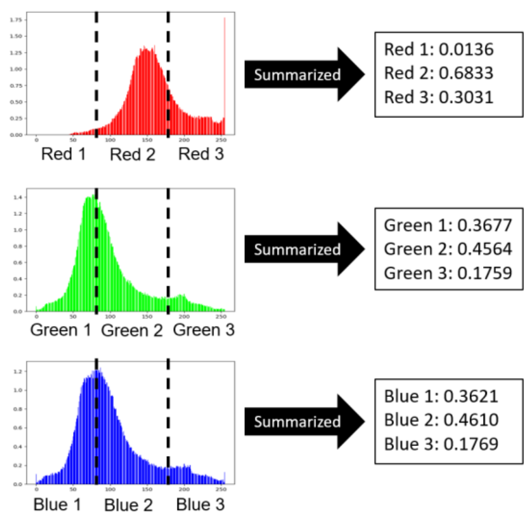


Figure 4. Illustration of histogram division to extract the color feature.

As illustrated in Figure 4, the three histograms in each channel are then divided into three equal areas to summarize the histogram by counting the average value of color frequency in each area. Basically, the histogram can be summarized into four, five, or more areas, but in this research, these three areas are considered to represent the weak, medium, and strong value of color. A similar procedure is also applied to images in HSV channel.

Each color channel (RGB and HSV) forms a separated dataset which was individually used to build a classification model. There are nine features/ attributes in each dataset that represent the value of ‘Red 1’ through ‘Blue 3’ in RGB channel and the value of ‘Hue 1’ through ‘Value 3’ in HSV channel. The use of the two channels in this research aims to clearly evaluate which color channel can distinguish between beef and pork precisely.

To expound on the method of segmenting the histogram into three categories: "weak," "medium," and "strong," a more detailed explanation is furnished herein. The process of histogram division entails an assessment of color distribution within the image. To be precise, the histogram is partitioned into three equally sized sections predicated on color intensity or frequency: "weak" for colors with lower intensity, "medium" for intermediate intensity, and "strong" for colors with higher intensity. This division is achieved by calculating thresholds that demarcate these categories, ensuring an equitable distribution of color data. For example, the "weak" category may encompass colors with intensity values falling below a specific threshold, while the "strong" category comprises colors with intensity values surpassing another threshold. This method facilitates a comprehensive representation of color features and guarantees the inclusion of various shades and intensities in the classification process.

C. *Pseudo Nearest Neighbor Rule (PNNR)*

The simplicity and the robustness of k-Nearest Neighbors (kNN) algorithm is the basic principle in PNNR which is an effective classification algorithm that was first introduced in [15], Additionally, it classifies the unlabeled data by using the distance weighted local learning in each class. PNNR was developed to solve the weakness of kNN by considering the local mean distance of every neighbor from each existing class, so that the decision is not based on the number of nearest neighbors but based on the distance of several k nearest neighbors as illustrated in Figure 5.

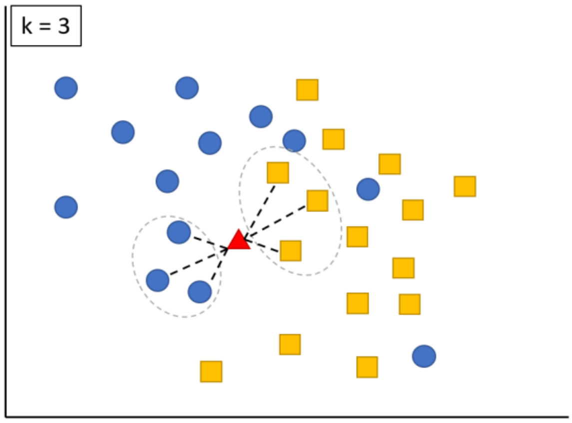


Figure 5. PNNR classifies unlabeled data (red triangle) based on the local mean distance from k nearest neighbors in every class

The local mean distance from each class is calculated using (1).

$$y_j = \sum_{i=1}^{k} w_i \times d_{ij} \qquad (1)$$

where $i = 1, \ldots, k$ and j is index of class, w is defined as the weight of each nearest neighbors which is defined in (2), and d is distance metric between unlabeled data (observed data) to k nearest neighbors in every class. In this research, Euclidean distance is used to calculate the value of d.

$$w_i = \frac{1}{i};\; i = 1, \ldots, k \qquad (2)$$

At the end, by using equation (3), the class that has minimum local mean distance is assigned to the unlabeled data.

$$x_{PNN} = \arg\min \{y_j\} \qquad (3)$$

As mentioned in [15], if more than one class has identical value of local mean distance, then the decision can be made by randomly choosing the class.

RESULT AND DISCUSSION

A. *Model Evaluation using RGB Channel*

The first experiment was conducted using color features from RGB channel. To build the classification model, the dataset which contains 744 records was separated using a 5-fold Cross Validation mechanism. The PNNR neighborhood parameter is set to several values: 3, 6, 9, 12, and 15 to identify the optimum one. The result of this first experiment, as given in Table 1, has quite a high accuracy of 86.70% on average, but this is still not a satisfying result.

Table 1. Experimental result using color features from RGB channel.

Neighborhood Parameter (*k*)	Accuracy
3	87.43%
6	87.43%
9	86.75%
12	86.35%
15	85.54%
Average Accuracy	86.70%

B. *Model Evaluation using HSV Channel*

The second experiment was conducted using color features from HSV channel and similar parameter to the first experiment. Surprisingly, color features in HSV channel give higher accuracy up to 93.78% for almost all *k* values as given in Table 2.

Table 2. Experimental result using color features from HSV channel.

Neighborhood Parameter (*k*)	Accuracy
3	93.78%
6	93.78%
9	93.78%
12	93.38%
15	93.78%
Average Accuracy	93.70%

Classification model using HSV color features has better accuracy than RGB channel with improvement up to 6%. HSV channel is better in representing the characteristic of color from beef and pork slices. HSV successfully separated inter-class data between beef and pork which means that features value between the two classes are significantly different. This can be analyzed by calculating the average value for each feature in RGB and HSV channel as given in Table 3 and Table 4.

Table 3. Average value of RGB features between 'Beef' and 'Pork' data.

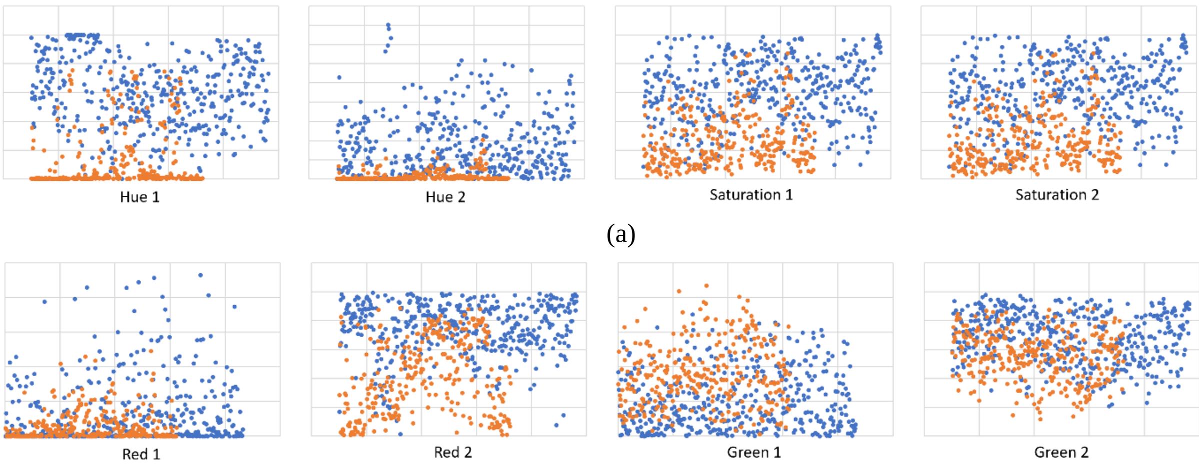
RGB Features	Average Value		Difference
	Beef	Pork	
Red 1	0.058	0.315	0.256
Red 2	0.629	0.736	0.107
Red 3	0.489	0.126	0.363
Green 1	0.453	0.416	0.036
Green 2	0.613	0.649	0.036
Green 3	0.111	0.111	0
Blue 1	0.198	0.418	0.220
Blue 2	0.719	0.641	0.078
Blue 3	0.260	0.118	0.142
Average Difference			0.138

Table 4. Average value of HSV features between 'Beef' and 'Pork' data.

HSV Features	Average Value		Difference
	Beef	Pork	
Hue 1	0.143	0.741	0.598
Hue 2	0.768	0.279	0.490
Hue 3	0.265	0.157	0.108
Saturation 1	0.314	0.712	0.398
Saturation 2	0.799	0.389	0.411
Saturation 3	0.063	0.076	0.013
Value 1	0.061	0.312	0.251
Value 2	0.625	0.715	0.091
Value 3	0.491	0.150	0.342
Average Difference			0.3

As seen in Table 3, each feature in RGB value between 'Beef' and 'Pork' data has a small difference average value. This indicates that inter-class data are not well separated because the value of data between two classes tends to be similar. In contrast, as seen in Table 4, HSV value has quite a large difference between 'Beef' and 'Pork' average feature value which indicates that HSV feature separate inter-class data effectively.

The scatter plot displayed in Figure 6 gives visual explanation where HSV channel has better inter-class separation compared to RGB channel. Well separated inter-class data gives very good performance to local based classification approach such as PNNR or other kNN variation.



(b)

Figure 6. Sample of scatter plot between beef (orange) and pork (blue) features in HSV (a) and RGB channel (b)

The use of color-based feature in this work is effective to classify the images of beef and pork slices, especially when using HSV channel instead of RGB channel. This is proven by the system performance that achieves up to 93.78% of accuracy as given in Table 5. The performance of color-based features in this work also outperforms the previous works in and [3], [5], [9], [10], [11] that use texture-based feature and the combination of color and texture-based features. The performance comparison between this study and previous works provided in Table 5 shows that the proposed system achieved the highest accuracy using only HSV color features with a simple classifier, i.e., Pseudo Nearest Neighbors Rule (PNNR).

Table 5. Performance comparison between proposed system and previous works

Work	Number of Class	Number images	Feature Extractor	Classifier Algorithm	Accuracy
This work	2	744	HSV histogram	Pseudo Nearest Neighbors	93.78%
[3]	2	340	GLCM, HSV	LVQ	76.25%
[5]	3	120	GLCM, HSV	LVQ3	91.67%
[9]	2	80	RGB, HSI, statistical features	Neural Networks	93.75%
[10]	2	100	HSV, LBP, PCA	Probabilistic Neural Networks	91.67%
[11]	3	30	HSV, LBP, statistical features	SVM	90.00%

CONCLUSION

This work has successfully proposed a system to classify the images of beef and pork slices using improved color-based features and Pseudo Nearest Neighbors Rule (PNNR). Instead of using complex feature extraction such as GLCM, invariant moment, or Local Binary Pattern, color features have significant contribution in classifying beef and pork images. Based on the experiment, HSV is much better than RGB in representing color features from beef and pork images. The proposed system successfully achieved the highest accuracy of 93.78% using HSV color features. For the future work, the combination of proposed color-based features in this work with texture-based features may give higher accuracy.

In conclusion, this study aimed to address a crucial issue in halal food recognition, particularly the distinction between beef and pork slices. We proposed and implemented the Pseudo-Nearest Neighbor Rule (PNNR) as a computer-based system for classifying meat slice images based on color features. Our research highlighted the significance of color as a distinguishing factor between beef and pork, especially when human observation may fall short, such as in cases involving mixed or processed meats. The performance evaluation of PNNR, utilizing a larger dataset than previous studies, yielded impressive results, achieving up to 93.78% accuracy when utilizing color features from the HSV channel. For future enhancements, we recommend exploring the integration of color-based features with texture-based features and real-time implementation in mobile applications to make halal food choices more accessible to consumers. In summary, this research not only provides a practical solution to the halal food recognition challenge but also underscores the role of computer-based systems in assisting consumers, particularly Muslims, in making informed halal dietary decisions, potentially contributing to the well-being of Muslim communities and their adherence to religious dietary.

REFERENCES

[1] C. Reviews, “Impact of Pork Consumption on Human Health,” vol. 7, no. 03, pp. 2442–2447, 2020.

[2] S. A. Laga and R. Sarno, “Optimal sample temperature of electronic nose for detecting beef and pork mixture,” *2019 International Conference on Information and Communications Technology, ICOIACT 2019*, pp. 398–402, 2019, doi: 10.1109/ICOIACT46704.2019.8938492.

[3] R. Sarno, K. Triyana, S. I. Sabilla, D. R. Wijaya, D. Sunaryono, and C. Fatichah, “Detecting Pork Adulteration in Beef for Halal Authentication using an Optimized Electronic Nose System,” *IEEE Access*, vol. 8, pp. 25–30, 2020, doi: 10.1109/ACCESS.2020.3043394.

[4] I. Tazi, N. Laila Isnaini, and A. Ainur, “Principal Component Analysis (PCA) Method for Classification of Beef and Pork Aroma Based on Electronic Nose,” *Indonesian Journal of Halal Research*, vol. 1, no. 1, pp. 5–8, 2019, doi: 10.15575/ijhar.4155.

[5] D. R. Wijaya, R. Sarno, and A. F. Daiva, “Electronic nose for classifying beef and pork using Naïve Bayes,” *Proceedings - 2017 International Seminar on Sensor, Instrumentation, Measurement and Metrology: Innovation for the Advancement and Competitiveness of the Nation, ISSIMM 2017*, vol. 2017-Janua, pp. 104–108, 2017, doi: 10.1109/ISSIMM.2017.8124272.

[6] U. Sudibyo, D. P. Kusumaningrum, E. H. Rachmawanto, and C. A. Sari, “OPTIMASI ALGORITMA LEARNING VECTOR QUANTIZATION (LVQ) DALAM PENGKLASIFIKASIAN CITRA DAGING SAPI DAN DAGING BABI BERBASIS GLCM DAN HSV,” *Jurnal SIMETRIS*, vol. 9, no. 1, 2018.

[7] J. Jasril and S. Sanjaya, “Learning Vector Quantization 3 (LVQ3) and Spatial Fuzzy C-Means (SFCM) for Beef and Pork Image Classification,” *Indonesian Journal of Artificial Intelligence and Data Mining*, vol. 1, no. 2, p. 60, 2018, doi: 10.24014/ijaidm.v1i2.5024.

[8] E. I. Sela and Sutarman, “Extracting the potential features of digital panoramic radiograph images by combining radio morphometry index, texture analysis, and morphological features,” *Journal of Computer Science*, vol. 14, no. 2, pp. 144–152, 2017, doi: 10.3844/jcssp.2018.144.152.

[9] E. I. Sela and R. Pulungan, “Osteoporosis identification based on the validated trabecular area on digital dental radiographic images,” *Procedia Comput Sci*, vol. 157, pp. 282–289, 2019, doi: 10.1016/j.procs.2019.08.168.

[10] S. Ayu Aisah, A. Hanifa Setyaningrum, L. Kesuma Wardhani, and R. Bahaweres, “Identifying Pork Raw-Meat Based on Color and Texture Extraction Using Support Vector Machine,” *2020 8th International Conference on Cyber and IT Service Management, CITSM 2020*, no. c, 2020, doi: 10.1109/CITSM50537.2020.9268892.

[11] L. Handayani *et al.*, “Comparison of target Probabilistic Neural network (PNN) classification for beef and pork,” *J Theor Appl Inf Technol*, vol. 95, no. 12, pp. 2753–2760, 2017.

[12] O. D. Nurhayati and I. P. Hastuti, “Metode Moment Invariant Geometrik untuk Menganalisis Jenis Daging Babi dan Daging Sapi,” *Jurnal Sistem Informasi Bisnis*, vol. 8, no. 2, p. 67, 2018, doi: 10.21456/vol8iss2pp67-72.

[13] J. Chairunnisa *et al.*, “Pork and Beef Features Extractions,” *Proceedings - 2018 International Seminar on Application for Technology of Information and Communication: Creative Technology for Human Life, iSemantic 2018*, pp. 295–298, 2018, doi: 10.1109/ISEMANTIC.2018.8549765.

[14] K. Adi, S. Pujiyanto, O. D. Nurhayati, and A. Pamungkas, “Beef Quality Identification using Color Analysis and K-Nearest Neighbor Classification,” in *2015 4th International Conference on Instrumentation, Communications, Information Technology, and Biomedical Engineering (ICICI-BME),* Bandung, 2015, pp. 180–184.

[15] Y. Zeng, Y. Yang, and L. Zhao, “Pseudo nearest neighbor rule for pattern classification,” *Expert Syst Appl*, vol. 36, no. 2 PART 2, pp. 3587–3595, 2009, doi: 10.1016/j.eswa.2008.02.003.