The Identification of Beef and Pork Using Neural Network Based on Texture Features

DOI : 10.36909/jer.17983

Khoerul Anwar*, Sigit Setyowibowo**

**Department of Information Engineering, STMIK PPKIA Pradnya Paramita,, Malang, 65124, Indonesia*
***Department of Information Engineering, STMIK PPKIA Pradnya Paramita,, Malang, 65124, Indonesia*
**Corresponding Author: alqhoir@stimata.ac.id*

ABSTRACT

The actual problem that frequently happens related to meat sales at conventional markets is the manipulation of pork and beef. It can happen as both visual textures bear resemblances. Texture is a crucial part of an object. In image processing, textures can be used for classification, recognition or prediction of an image. This paper offers the Minimum Overlap Probability - Neural Network method for the identification of digital image features of pork and beef.. Minimum Overlap Probability was employed to select features of the strongest characteristics, whilst Neural Network is used for training and classification. Based on the test results, the strongest features are maximum probability, contrast, sum average, autocorrelation, and energy and entropy sum. Based on MOP-NN Model test result, the digital image identification of beef and pork has performance with an accuracy of 96% on 400 images of sample data.

Keywords: single feature, multiple features, minimum overlap probability, neural network, identification, beef and pork**.**

INTRODUCTION

The actual problem that frequently happens related to meat sales at conventional markets is the manipulation of pork and beef. This manipulation is repeated every year when the demand for beef increases sharply in a certain month. It can happen as both visual textures bear resemblances. The average consumer of beef cannot distinguish the identity of beef and pork. This fact becomes a challenge for researchers to conduct research on image processing or computer vision.

The identification of two objects having similar visual texture based on digital image in the field of visual computer is a heavy task. Beef and pork bear resemblances in their visual texture. Physically, pork and beef can be distinguished into five aspects, i.e. color, meat fibers, type of fat, aroma, and texture. The common method used for the identification is physical testing technique carried out in the laboratory, experts / specialists observations, and electronics nose. It is a difficult thing in differentiating based on digital image. This paper offers the identification method based on digital image feature. The challenge faced by the researcher for identifying digital image of pork and beef is to obtain the strongest feature as a key and model class for identification task of them both.

Until this decade, many methods have been presented to analyze image texture. In general, this method presents a technique for extracting repetition patterns that are presented with numbers so that they can be used for machine learning processes (Li, 2022). Features are unique identity belong to an image. Identifying the image that has a consistent pattern of texture, the pattern can be used as a key. However, if the image does not have a certain pattern, the analysis can use its statistical feature attachment. In this study, the object is the digital image of beef and pork from the acquisition with a smartphone camera. The image produced from the acquisition does not show unique physical characteristics such as meat fiber and chewiness. Meanwhile, only the color and texture of the image appear. For color, it's easy to change. Fresh meat has a different color from meat that is not fresh. So that color is not chosen as a key feature in this identification process. The choice of texture as a key feature is based on the reality that the texture does not change in fresh or unfresh meat. Therefore, the texture of the meat image is an important element for the identification process of beef and pork. The texture can be extracted to get a special feature by extracting its statistical features. Identification or classification that involves a large amount of data requires a feature or a combination of several features to get the best performance. On the basis of which it is possible to automatically conduct a more accurate classification of this image (Ilyasova et al., 2020). One of the statistical (Ibrahim et al., 2021) feature extraction methods that has quite a number of features is GLCM (Naveen et al., 2020). Features can be obtained by doing extraction of image texture. The relevant paper in connection with extraction technique includes GLCM (Shahare, 2020).

The key feature is an important part, especially in supporting the classification task(Sun & Bourennane, 2020). In the present paper, selecting a key feature in the identification case of beef and pork becomes an important task. In fact, a feature of two different objects for similar types of features with different values also becomes a separated problem. A set of features in a large number has probability to have partly or entirely irrelevant or redundant (Ortiz-jimenez et al., 2020)features. The irrelevant or redundant features do not represent as a unique feature. Ideally, a key feature is independent to one another. When the ideal condition is unachievable, then the best solution is to choose the features with minimum overlap. It has been conducted in feature selection by means of selecting features on redundancy minimum criteria for lymphoma classification process. The feature selection technique with minimum redundancy is also used by (Ilyasova et al., 2020) for text classification and texture features. The main function in the selection is selecting the minimum redundancy (Almayyan, 2022)of the two objects usable for forming each class. However, both features selection is used in the multi-label features in one single object. In contrast with the previous two researchers in this paper, the author conducted multi features selection for multi objects applied to the digital image identification of beef and pork. The key feature is considered relevant when having minimum overlap (Rana et al., 2021). The fundamental problem for overlap feature is that this set of features do not comprehensively represent the value of a feature as a target as it still contains the values of other features. The challenge encountered is how to choose one or several of those overlap features as candidates for the best features(Rana et al., 2021). The basic assumption is the smaller range of overlap value of a feature, the better the feature is chosen as the winner of the selection (key feature).

The key features affect the establishment of data classification or formation. In the identification task, the underlying problem is the availability of the model class. The capable classes are adaptive to new data, hence the system is able to generalize or predict new data with high accuracy. There are several classifiers usable for obtaining the model class, among others KNN (Yasiran et al., 2021), harris fragments model and classification KNN method (Lazrak et al., 2022), CNN (Ma & Zhang, 2021), GAT (Series, 2020), Deep CNN (Guo et al., 2020), SVM (Naveen et al., 2020)

In the present paper, the authors offer a solution for classification problems in getting a model class with the combination of Minimum Overlap Probability (MOP) (Anwar et al., 2016)and Neural Network (NN). MOP is a feature selection method to get the strongest feature and NN as the classifier. The NN employed is Back Propagation Neural Network (BPNN). The steps built to form the adaptive model class are obtained by conducting input variation to NN, and subsequently to select variation having the highest accuracy. There are two variations done. Firstly, input with a single feature, and secondly, input with many features. The features used are those selected from MOP selection process. NN output with the best accuracy is chosen as the model class. Such class is what used as a reference in the identification process of pork and beef.

This paper is organized as follows: Section Material and methodes consists of the materials and methods concerning GLCM, MOP-NN. Section results presents the experimental result about the feature extraction, feature selection and the classification with the computation of single feature, multiple features in BPNN classifier performance. Section conclusion contains the conclusions.

MATERIALS AND METHODS

The objects in this research were pork and beef. The types of beef selected were those from the types of loin, rump, round, and for pork were loin and ham. The meat variation was obtained from 10 cows and 10 pigs. Each of them was divided into a group of 7 for training data retrieval, and 3 for testing data retrieval. Each type of beef and pork was made into data samples of 10 slices measured 4 x 4 x 0.5 (cm). On each slice, acquisition was done 10 times. The tool used for data acquisition was mobile device digital camera (HP) Lenovo A706_ROW. The digital image was captured at a resolution of 5 MP pixel. The acquisition distance between the meat slice and the camera was 10 cm. The digital image format resulted from the acquisition of this tool was JPEG-format image. The acquisition was performed in open space under bright light condition. The total digital images resulted from acquisition is 2000; with details 1000 images of beef and 1000 images of pork.

In this paper, we build an identification model of pork and beef based on digital image texture features. The model built is shown in Figure. 1.

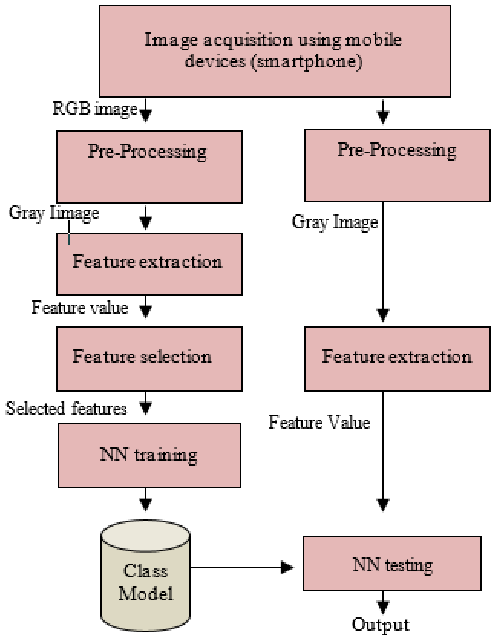

Figure 1. Identification model of beef and pork

Preprocessing

The identification of pork and beef from the acquisition result with mobile device is not easy to do, hence initial processing to improve the quality is frequently done. The acquisition result image is color image (RGB). At the preprocessing stage, the original image is resized into original image with dimension of 255x255 pixels. The change is intended to accelerate the computing process. The color image needs to be converted to grayscale since the extraction process is based on GLCM feature. RGB conversion to grayscale used the formula 0.29*R + 0.59*G + 0.11*B. The next stage is histogram equalization process. This process is to overcome the problem of image contras and brightness. The final stage of preprocessing is to filter the attached image noise and refine texture in the image. The process of image filter is done by using Sobel filter. The original image and the pre-processed image are shown in Figure. 2.

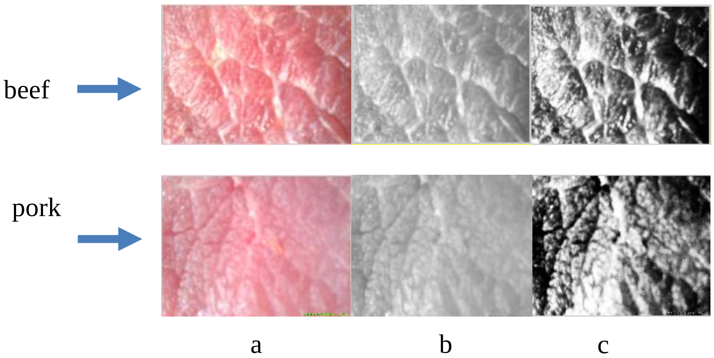

Figure 2. Preprocessing beef and pork imagery, a) RGB image, b) gray image, c) histogram equalization

Extraction

Features are obtained from texture extraction using statistical analysis. The texture feature is computed based on the statistical distribution of pixels intensity on relative position one another in image pixels representation in the matrix. The feature's value depends on the number of pixels or dots in any combination of second-order statistics. The second order is measured based on the relationship between the dual groups of pixels in the actual image. The feature extraction is based on gray level co-occurrence matrix (GLCM) method which is the second order statistics. The method has a variety of suitable features for the need or task of texture analysis of an image. In this section, the author extracted 20 types of GLCM features. The pixels adjacency analysis is calculated based on two pixels with a distance of one space in the direction 0^0. At this adjacency direction, the feature's value in digital images of pork and beef shows a higher value than in the other direction.

Feature Selection

The feature dataset which consists of 20 GLCM features from the beef and pork extraction process has the opportunity to have a relationship, namely: independent, overlap (intersection), superset/subset and equal. Therefore, a technique is needed to obtain the least overlapping features. The best features are ideally free from redundancy with each other (Rana et al., 2021). However, if the features of beef and pork do not have independent features, then different solutions should be considered.

In this article, the study focuses on the texture features of cow and pig images with GLMC feature values overlapping one another. To discuss this, the MOP technique is used (Anwar et al., 2016). MOP is used to get some of the best features with the lowest overlapping value. In the article, 20 GLCM features are presented from beef and pork processed to get the best fit. The selection technique uses a comparison of the minimum and maximum values of each GLCM feature value from beef and pork images. MOP is used to perform feature selection by calculating each feature and then displaying the results in descending order. The winning feature can be selected by applying a value threshold. In his article, the range of minimum and maximum values of each digital image feature of beef is denoted by Fs, while for pork it is denoted by Fb. Fs has a minimum point value range (x1) and the maximum value is (x2). While for Fb has a range of values at the maximum point (x4) and the value at the minimum point (x3). Based on this description, the area of Fs can be defined as follows:

$Fs = x_1 : x_2$

$$\Delta x_s = |x_2 - x_1| \quad (1)$$

And Fb can be defined as follows:

$Fb = x_3 : x_4$

$$\Delta x_b = |x_4 - x_3| \qquad (2)$$

Fs and Fb are considered to have overlapping values if they meet the following rules:

1. Fs (min) < Fb (min) and Fs (max) > Fb (min) and Fs (max) < Fb (max)
2. Fs (min) > Fb (min) and Fs (min)<> Fb (max) and Fs (max) >Fb (max)

(Anwar et al., 2016) write the overlapping formula between two features based on the range of values from Fs (X, x2) and Fb (X3, X4) is

$$Lo = |x_2 - x_3| \text{ or } |x_4 - x_1| \qquad (3)$$

Referring to equation (3), if there is a feature value in it, it can be indicated that it will produce problems in the identification process of beef and pork image identification. The problem arises when there is a duplication of the values of the two features of beef and pork. If Lo has a large value, it means that the duplication between two features has a large overlapping value and vice versa if Lo is a small value, the overlap of the two features is getting smaller. What needs to be understood is that the area is based on a range of feature values, so that the overlap area of each feature is not absolute in a certain number range because it is influenced by the stability of the minimum and maximum values of each feature. However, by using the area range based on the min-max feature, this method can obtain overlapping areas. The possibility of overlap or the chance of overlap is a major concern in the MOP method. Furthermore, to determine the probability of each feature overlap between Fs and Fb, it is calculated by comparing Lo with the overall feature area (FS + Fb). In this article, the calculation of the overlap area is referred to as the overlap probability (Anwar et al., 2016). The calculation is

$$\text{Prob overlap} = Lo \;/\; ((\Delta x_s + \Delta x_b) - \text{Lo}) \qquad (4)$$

The explanation of Equation (4), the smaller the Lo value, the less the chance of overlap and vice versa, the bigger the chance of the overlap.

MOP Algorithm

1. Calculating the min-max value of the digital image extraction feature of beef and pork.
2. Define the F_x feature area, digital image of beef and pork.
   a. If F_x has no slices (broken) then it is the selected feature
   b. if F_x is a subset or superset , then F_x is rejected
   c. if in the calculation of delta F_x the feature database is not included in the processes to a and b, then proceed to keno 3.
3. Calculating the value of ProbError
4. Computing the ProbError value
5. Ranking
6. Determine the threshold
7. Choose a feature with a ProbError less than the threshold

8. Selected features

After extraction and selection, features are used as NN input for classification in beef and pork classes. In this classification, we use Backpropagation Neural Network with multi-layer perception (MLP)

Backpropagation Neural Network (BPNN)

The MLP algorithm was first introduced in the thesis by Paul Werborn in 1974 in a general context. In 1980s, it developed rapidly and was used for a variety of specialized tasks. The working principle of MLP, each layer has a weight matrix ***W*** and bias vector ***b***, an input and output vectors. The separation between categories is taken by placing a hyperplane between the two categories.

In this paper, the BPNN design is comprised of an input layer, two hidden layers and an outer layer. At each layer, bias factor is added. Then for NN computation, activation function used is sigmoid [0 1] and learning method used is Levenberg Marquardt (LM). The number of node for each layer is designed as follows: in the input layer, the node number is adjusted with the selected feature (the number of feature selection winner). For the number of nodes in the hidden layer, it is determined by a formula adopted from (Chen et al.,2010), stating that the number of neurons in the hidden layer (n_H) is

$$n_H = \sqrt{n_i + n_o} + l \qquad (5)$$

n_H: the number of neurons in the input layer, n_i: number of neurons in the hidden layer, n_o: the number of neurons in the output layer, *l*: constants in integer (1,2,3 ... 10). BPNN outer layer architecture is designed with two neurons. The output layer with these two neurons is designed separately to get the reliable model class valid identification result. BPNN is classifier type with determined class. In this paper, the number is class is specified as many as two classes of prototypes or targets. Two classes of binary-type targets are [00] for pork class and [11] for beef class.

Figure. 3. NN outer layer contains two nodes, first as a node for Y_1 and Y_2 second as node for Y_2. Then, NN output in node at Y_1, Y_2 is recorded in vector C with element $Y_{1i}Y_{2i}$, $i = 1, 2, 3 \ldots$n is the index data. Such vector element value will be matched with the target. Y_1 and Y_2 data type as an output of the NN in numerical numbers, whilst target class is included binary-type. Steps taken to equal the data type is to transform NN output into binary. It is done by providing threshold value of 0.35 to the vector C,

$$F(i) = C[Y_{1i}Y_{2i}] = \begin{Bmatrix} i \geq 0.1 & = 1 \\ otherwise & = 0 \end{Bmatrix} \qquad (6)$$

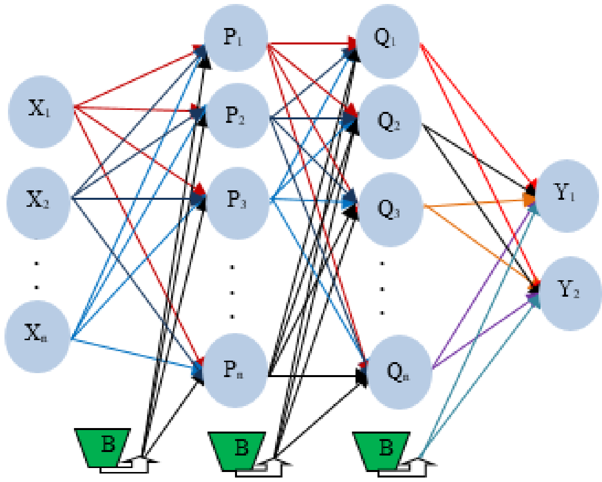

Figure 3. NN architecture

Therefore, when the output at Y_1 = = 0.0000012 and Y_2 = 0.0000089, refer to (6), each Y_1 becomes 0 and Y_2 becomes 0. With the same scheme as Y_1= 0.1237 and Y_2 = 0.4781 then each will be 1 (Y_1=1;Y_2=1). Another matter considered by the author is by the presence of two nodes, NN outputs have the possibility of 2^n; n is the number of nodes. Thereby, the NN outcome has four output variations, as shown in Table 2.

In Table 2, four possibilities happening at NN outputs are presented. Symbol X_1, X_n, respectively is the symbol for the digital image features of beef and pork, while x.xx is the input symbol of NN in the form of numerical data with two digits behind the comma. For symbols Y_1 and Y_2 are NN outputs symbol with binary data type. The output type is binary with respective target class is [00] for the pork class and [11] for beef class. Meanwhile, the output with binary order [01] or [10] is defined as the class "Non Pork or Beef". The definition is provided as the focus of information to be obtained on the task identification is to obtain the type of beef and pork.

Table 2. Prototiye class model

No	Input		Output		Class
	X_1	X_2	Y_1	Y_2	
1	x.xx	x.xx	0	0	Pork
2	x.xx	x.xx	0	1	Non pork or beef
3	x.xx	x.xx	1	0	Non pork or beef
4	x.xx	x.xx	1	1	Beef

Training

In the training phase, the network architecture is built with model I 7H 7H 2O, eppoch 1000, the learning rate, momentum = 0.91. In this training phase, there are two types of inputs on BPNN, namely the training model with one input feature and training model with multiple input features. In the first model, each selected feature is tested as input. In contrast with the first model, multiple features model is the combination of many selected features. Features used are selected features of the feature selection process.

Testing

In the testing phase, each BPNN model for a single feature and multiple features are tested with the real data, the data which are not from the training dataset. In addition, in order to determine the effect of total overlap data on the accuracy, the model is tested with four simulations of data amount in overlapped area, i.e. 0%, 25%, 50% and 100% of the total data.

RESULTS

Some of experiments results that have been done are presented in tables and pictures in graphical form. Table 3 indicates the tabulation results of GLCM extraction result feature value on the direction of adjacent pixel 0^0. The extraction is carried out on 800 textures of pork and beef digital images. From the extraction result, there are a number of 20 features extracted for each image of pork and beef. The analysis result on value for 20 features of these two shows value overlapping. As overlap feature sample is contras, contrast feature value of beef for minimum is 0.039 and maximum of 0.486, the contrast feature value of pork for minimum 0.003 and maximum of 0.224. Thereby, the contrast feature range of two meats also undergoes overlap. This condition causes difficulties to choose the strongest feature as a key identification. Nevertheless, the feature selection is still possible to do by selecting the minimum overlap.

Table 3. Minimum and maximumtabulation of pork and beef feature value

Fitures Name	Beef		Pork	
	min	max	min	max
Autocorrelation	1.469	3.458	1.287	2.220
Contrast	0.039	0.486	0.015	0.201
Correlation	0.716	0.935	0.818	0.966
Cluster Prominence	10.497	241.140	86.487	190.430

Cluster Shade	2.022	24.919	7.613	16.607
Difference entropy	0.164	0.756	0.042	0.490
Difference variance	0.039	0.486	0.015	0.201
Dissimilarity	0.039	0.343	0.008	0.177
IDN	0.963	0.996	0.981	0.999
IDMN	0.993	0.999	0.997	1.000
Energy	0.321	0.811	0.504	0.973
Entropy	0.445	1.777	0.099	1.189
Homogeneity	0.841	0.981	0.914	0.996
Maximum probability	0.538	0.898	0.695	0.986
Sum of squares: Variance	1.454	3.644	1.262	2.279
Sum average	2.214	3.244	2.076	2.660
Sum entropy	0.418	1.488	0.093	1.040
Sum variance	3.464	7.634	3.983	4.940
Information moc 1	0.668	0.226	0.643	0.335
Information moc 2	0.533	0.823	0.288	0.708

In Table 4, the feature selection method result is shown using MOP method with a overlap probability threshold value below 0.35 or five names of top features with the smallest overlap values. Based on the top five ranking for overlap feature value, the selected feature chosen are autocorrelation, information measure of correlation 2, sum of square variance, contrast and difference variance. Based on the overlap features value, there are two groups of twin feature values. First, is the information measure of correlation 2 feature and Sum of square variance with overlap probability value of 0.33. The second is the contrast and difference variance with overlap probability value of 0.4 overlap.

Table 4. The names of the features selected by mop

No	Features Name	Beef		Pork		Over Lap
		Min	max	Min	max	
1	Autocorrelation	1.469	3.667	1.287	2.220	0.32
2	Information.moc 2	0.533	0.823	0.288	0.708	0.33
3	Sum of square variance	1.454	3.781	1.262	2.279	0.33
4	Contrast	0.039	0.486	0.015	0.201	0.34
5	Difference variance	0.039	0.486	0.015	0.201	0.34

In Table 5, the result of BPNN training result is shown using a single feature as input. The obtained result shows this input model has a low accuracy. Each feature of selection winner when used as input to NN as training data cannot contribute maximally to the performance of NN. The test result for the single feature obtains the highest accuracy value at the feature information measure of correlation 1 (77.50%) and the lowest accuracy on autocorrelation feature (69.38%). These test results indicate that a single feature is not capable of supporting the NN performance maximally at the training process to form model class.

.

Table 5. BPNN performance based on one selected feature input

No	Features Name	Model Accuracy (%)
1	Autocorrelation	69.38
2	Information moc 1	77.50
3	Sum of square variance	69.39
4	Contrast	76.50
5	Difference variance	73.97

In Table 6, it is shown that the result of BPNN training with multiple features of the selected features. This features combination is performed by pairing the features sequentially. The combination of two features, tree features, four features and the last five features. The combinations 1,2 (autocorrelation, information measure of correlation 2). The combination of 1,2,3 (autocorrelation, information measure of correlation 2, sum of square variance). The combination 1,2,3,4 (autocorrelation, information measure of correlation 2, sum of square variance, contrast), and the combination 1,2,3,4,5 (autocorrelation, information measure of correlation 2, sum of square variance, contrast). The results found show the BPNN performance for each combination successfully identifies training data perfectly (100%). Meanwhile, the iterations required on each combination are 239, 202, 58 and 37, respectively. In combination 1,2,3,4 and 5 requires the fewest iterations (37 iterations) or most rapidly achieved and the largest iterations on a combination of *1,2* (239 iterations). This 100% success indicates that the BPNN have maximum performance when the input has multiple features type. Another indication of this result is that multiple features are suitable to be used as a special feature in the identification process.

Table 6. BPNN performance based on a combination of selected features

Features combination	Iteration	Model Accuracy (%)
1,2	239	100
1,2,3	202	100
1,2,3,4	58	100
1,2,3,4,5	37	100

To test the performance of the proposed method, 4 different types of data were tested. First, the system was tested with data that did not overlap (0%) between the values for each key feature of the digital image of beef and pork.

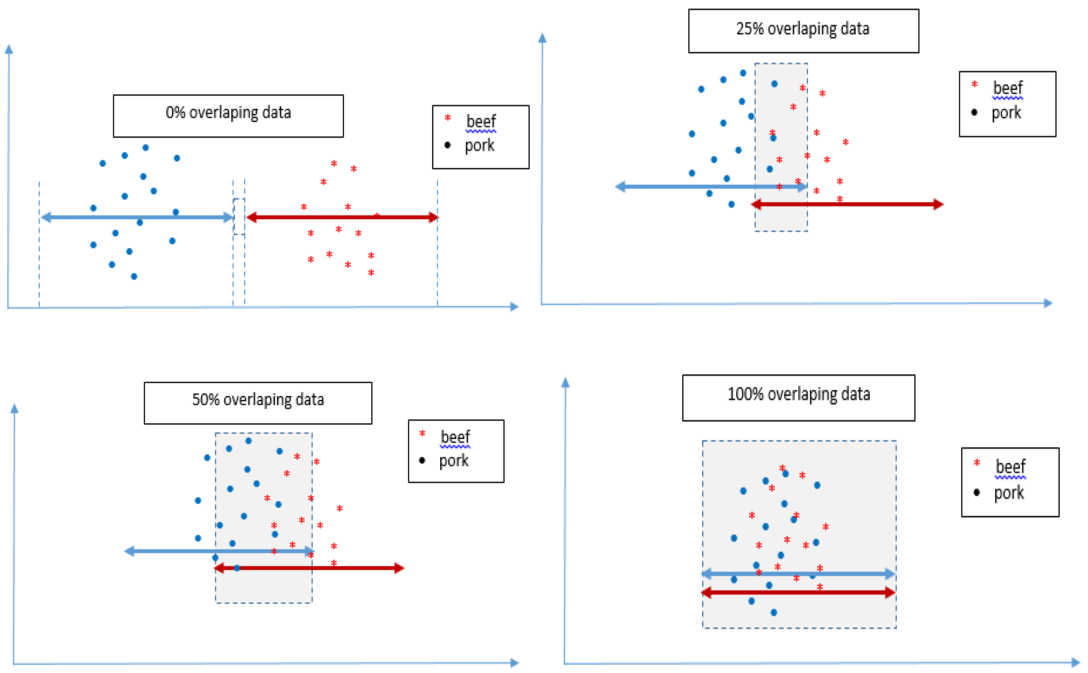

Figure. 4. Overlapping data illustration based on the amount of overlap data, a= 0%, b = 25%, c=50$, d=100%

Second, the system was tested with 25 percent of the test data being overlap data (25%). Third, the system is tested with data in the amount of 50 percent of the test data is overlap data (50%). Finally, the system is tested with data in the amount of 100 percent of the test data is overlap data (100%). An illustration of overlapping data is shown in Figure. 4. Meanwhile, the results of system performance in identification are shown in Table 4.

In Table 7, are shown in the test result of overlapped area is obtained by testing the untrained data. The number of overlap data is given, ranging from 0%, 25%, 50%, to 100% of the test data. The value is a percentage of data amount of feature value in the overlapped area. In this testing, the feature data entered in the overlapped area is the digital image feature of the pork only. The objective of this test is to determine the correlation of the data amount in the overlapped area with an error rate of identification. Testing is done on five features: maximum probability, contrast, energy, entropy and autocorrelation. Total value of beef feature is 200 and total value of pork feature is 200, thus total data are 400 features value.

Table 7. The identification results are based on each selected feature overlap range

Features Name	Overlap Data and Misidentification			
	0%	25%	50%	100%
Max probability	0	10	21	41
Contrast	11	15	18	27
Energy	0	7	13	26
Entropy	4	17	32	53
Autocorreation	4	16	29	34

The test results for non-overlapped data (0%) find misidentification by 11, 4, 4 sequentially in contrast feature, entropy and autocorrelation. It indicates that the error may still occur for testing data located outside the overlapped area. Meanwhile for the data with 100% overlap, error less than 60% is found. The result provides information that the model overlapped area is able to identify the objects correctly by 40%.

Figure. 5 shows a graph representing the correlation between the amount of data in overlapped area and the percentage of errors. The relationship between the amount of data in overlapped areas with the error showed that the higher the amount of test data is in the overlapped area, the higher the percentage of errors occurred. it shows that the percentage of error is affected by the large amount of testing data in the overlapped area is directly proportional to the amount of data in the overlapped area.

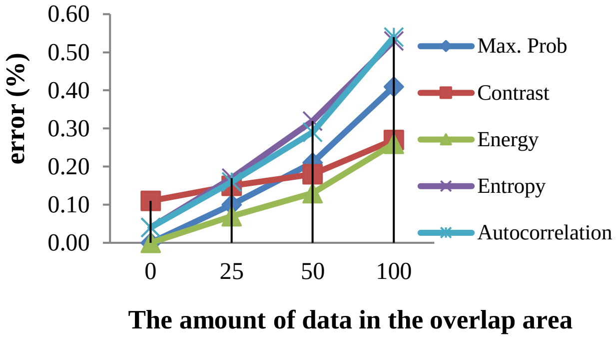

Figure 5. Graph of error percentage based on the amount of data in overlapped probability range area

In order to get the best feature as a feature that can be used in the identification task of beef and pork, the author tested several combinations of the amount of input features to NN. Testing was done to one feature and multiple features.

In Figure. 6, the graph of testing result of one feature with real data is shown. The test results obtain highest accuracy value of 73.46% on information measure feature of correlation 2 and the lowest accuracy of 52.04% on autocorrelation feature, and in overall, the average accuracy value is 59.89%.

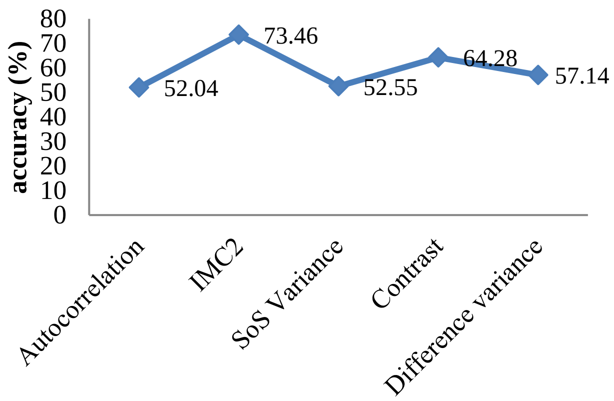

Figure 6. Graph of identification of accuracy percentage of each input feature in neural network

The exploration of misidentification is based on each feature, as indicated in Figure. 3, the obtained result is shown in Table 8. In single feature testing, the percentage of errors based on each feature on beef digital image obtains the highest score on contrast feature by 10% and the lowest autocorrelation feature of 2%. For pork digital image, the highest value in autocorrelation feature is 46% and the lowest being the measure of correlation.1 and Sum of square variance at 23%. The average beef misidentification (6%) and pork misidentification (29%).

Table 8. The number of identification errors of the feature in the classification based on one feature

Features Name	Beef	Pork
	False Positive	
Autocorrelation	2%	46%
Information moc 1	3%	23%
Sum of square variance	4%	23%
Contrast	10%	26%
Difference variance	14%	29%
Average	6%	29%

Figure. 7 shows a graph of test result for many features with real data, the highest accuracy obtained is 90% and the lowest accuracy is 72%.

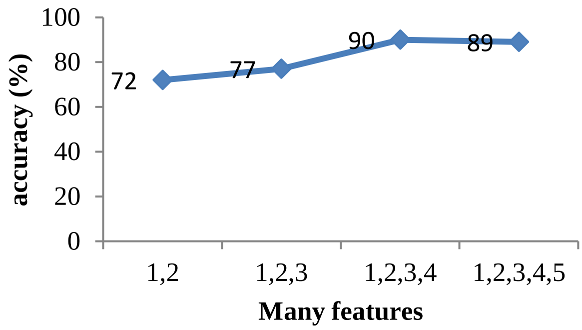

Figure 7. Graph of identification accuracy percentage based on each feature combination

The highest accuracy is obtained on a feature combinations 1,2,3 and 4 (autocorrelation, Information measure of correlation 2, Sum of square variance and contrast. The lowest accuracy is obtained at feature combinations 1 and 2 (autocorrelation, Information measure of correlation 2)

In Table 9, the test result for many features is obtained, the errors percentage based on each feature combination in two types of digital imagery. In beef, the highest error value is obtained at feature combinations 1,2,3 at 26% and the lowest at feature combinations 1,2; 1,2,3,4 and 1,2,3,4,5 at 20%. In the digital image of pork, the highest error value is obtained at the feature combinations 1.2 at 34% and the lowest feature combinations 1,2,3,4 at 0%. The average value of beef misidentification is 21% and pork misidentification is 13%.

Table 9. The amount of identification errors of the features in the classification based on multiple features

Features Name	Beef	Pork
	False Positive	
1,2	20%	34%
1,2,3	26%	16%
1,2,3,4	20%	0%
1,2,3,4,5	20%	2%
Average	21%	13%

The combination feature with highest accuracy value is 1,2,3,4 in this research is selected as the strongest characteristics and used as a key feature for the identification of beef and pork.

DISCUSSIONS

Based on test result shown in Table 4, the information obtained is that the overlap probability is not directly correlated with accuracy. That is, feature with smaller overlap probability does not necessarily have a higher accuracy than the feature with bigger probability overlap. As a sample, the autocorrelation feature probability value is 32% and information measure of correlation 2 feature is 33%. Apparently, autocorrelation accuracy is smaller than information measure of correlation 2.

The testing result on one feature and multiple features obtains result that multiple features have an average identification accuracy of (90.00%), higher than one feature (73.46%). Multiple features as an identity for identifying beef and pork contribute in raising the accuracy value of (16.54%).

The best combination of 1,2,3,4 have an average of 90% when used for the identification of beef, this combination is able to work with an accuracy of 80% and for the identification of pork model is able to work with 0% error or in other words, have an accuracy of 100%. Based on these results, it can be said that this combination is suitable for use as a key feature in the model built.

The testing results on the model built for the identification of beef and pork are based on digital image texture feature using real data obtain an accuracy of 90.00%, or in other words, the model has an error of 10.0%.

Based on Table 4, the number of test data in the overlapped area has an influence on the identification accuracy. The bigger the amount of testing data in overlapped area, the bigger the identification errors.

CONCLUSIONS

Based on MOP-NN selection, the best feature selected is autocorrelation, Information measure of correlation 2, Sum of square variance and contrast.

Identification based on one feature obtains an accuracy of 73.46%, whilst the identification based on multiple features obtains an accuracy of 90.00%. It shows that the use of many features as key identification contributes to accuracy increase of 16:54%.

The model offered by MOP-NN is able to be used as a solution to identification problems of pork and beef based on digital image texture features. It proves that the model offered has performance with an accuracy of 90.00% or error rate of 10.00% on 400 sample data images.

REFERENCES

Almayyan, W. (2022). *A Modified Maximum R Elevance M Inimum R Edundancy Feature Selection Method Based On T Abu S Earch For P Arkinson ' S Disease Mining*. *March 2020*. https://doi.org/10.5121/ijaia.2020.11201

Anwar, K., Harjoko, A., & Suharto, S. (2016). Feature Selection Based on Minimum Overlap Probability (MOP) in Identifying Beef and Pork. *International Journal of Advanced Computer Science and Applications*, *7*(3). https://doi.org/10.14569/ijacsa.2016.070345

Guo, H., Liu, J., Xiao, Z., & Xiao, L. (2020). Deep CNN-based hyperspectral image classification using discriminative multiple spatial- spectral feature fusion Deep CNN-based hyperspectral image classification using discriminative multiple spatial-spectral feature fusion. *Remote Sensing Letters*, *11*(9), 827–836. https://doi.org/10.1080/2150704X.2020.1779374

Ibrahim, L. A., Nasser, N., & Ali, M. (2021). Face Recognition Based on Statistical Texture Features. *Embedded Selforganising Systems*, *7*(1), 10–15. https://doi.org/10.14464/ess.v7i1.471

Ilyasova, N. Y., Shirokanev, A. S., & Demin, N. S. (2020). Segmentation of Lung Images Using Textural Features. *Journal of Physics: Conference Series*, *1438*(1). https://doi.org/10.1088/1742-6596/1438/1/012015

Lazrak, S., Semma, A., Ahmer El Kaab, N., El Youssfi El Kettani, M., & Mentagui, D. (2022). Writer identification using textural features. *ITM Web of Conferences*, *43*, 01027. https://doi.org/10.1051/itmconf/20224301027

Li, J. (2022). *Research on Image Texture Feature Extraction Based on Digital Twin*. *2022*.

Ma, Y., & Zhang, S. (2021). *Feature Selection Module for CNN Based Object Detector*. https://doi.org/10.1109/ACCESS.2021.3073565

Naveen, M., M., V., & Hemantha, G. (2020). Leaf Classification based on GLCM Texture and SVM. *International Journal of Computer Applications*, *177*(35), 18–21. https://doi.org/10.5120/ijca2020919846

Ortiz-jimenez, G., Modas, A., & Frossard, P. (2020). *Redundant features can hurt robustness to distribution shift Redundant features can hurt robustness to distribution shift*. *July*.

Rana, P., Dinh, T. N., & Ghosh, P. (2021). *Relevant and Non-Redundant Feature Selection for Cancer Classification and Subtype Detection*. *August*. https://doi.org/10.3390/cancers13174297

Series, C. (2020). *Image classification model based on GAT Image classification model based on GAT*. https://doi.org/10.1088/1742-6596/1570/1/012082

Shahare, N. (2020). *Analysis of Different Filtering Methods for Pre- processing and GLCM Feature Analysis of Different Filtering*

Methods for Pre- processing and GLCM Feature Extraction Using Wavelet in Mammogram Images. *June*. https://doi.org/10.9734/bpi/immr/v3

Sun, Q., & Bourennane, S. (2020). Hyperspectral image classification with unsupervised feature extraction. *Remote Sensing Letters*, *11*(5), 475–484. https://doi.org/10.1080/2150704X.2020.1731769

Yasiran, S. S., Salleh, S., Sarmin, N., Mahmud, R., & Halim, S. A. (2021). Curvelet Based Texture Features for Breast Cancer Classifications. *Journal of Physics: Conference Series*, *1988*(1). https://doi.org/10.1088/1742-6596/1988/1/012037