International Journal of Engineering & Technology, 7 (4.44) (2018) 117-121

International Journal of Engineering & Technology

Website: www.sciencepubco.com/index.php/IJET

Research paper

Identifying Arabica Raw Coffee Bean Varieties through Feature Extraction GLCM and Circularity

M. Agung Nugroho [1]*, Maria Mediatrix Sebatubun [2], Sumiyatun [3]

[1, 2, 3] Department of Informatics Engineering, STMIK AKAKOM Yogyakarta, Indonesia
**Corresponding author E-mail: nugroho.agung.m@gmail.com*

Abstract

The morphology of raw coffee bean which has colour, texture, size and circularity features are used as standardization to calculate the price and the quality of the raw coffee bean. Meanwhile, coffee farmers have difficulties to distinguish the coffee varieties based on the features of raw coffee bean. Generally, the way of the distinguish the varieties of the coffee is through their own visual perception in form of the tree, leaves, and raw coffee bean. They find it difficult to distinguish the coffee varieties due to the similarity of the varieties forms and colours. This research proposes to solve the problem through the image processing method as the second opinion to help the coffee farmers in identifying the coffee varieties. The research is conducted in three steps processes: The first is pre-processing by cropping the image of raw coffee bean. The second is extracting the image feature of raw coffee bean with Gray Level Co-occurrence Matrices (GLCM) and circularity feature. The last is classifying the feature with multilayer perceptron. The results of the image processing method indicate that the accuracy is 90% with sensitivity is 90%, and 90% specific.

Keywords: *raw coffee bean, image processing, classification, GLCM*

1. Introduction

The morphology of raw coffee bean which has colour, texture, size and circularity features are used as standardization to calculate the price and the quality of the raw coffee bean. Meanwhile, coffee farmers have difficulties to distinguish the coffee varieties based on the features of raw coffee bean. Generally, the way of the distinguish the varieties of the coffee is through their own visual perception in form of the tree, leaves, and raw coffee bean. They find it difficult to distinguish the coffee varieties due to the similarity of the varieties forms and colours.

Image processing is a method to transform an image into digital form and perform some operations such as to obtain an enhanced image or to extract the useful information. It is a type of signal dispensation where the input is images, for example, a video frame or photo and output, can be image or characteristics which are associated with the image. Raw coffee beans image that has been taken would be processed using the technology of image processing. In image processing, feature extraction is a common technique used to take features or characteristics of an object. The method is used to take the features of the raw coffee bean so that the system can recognize the coffee varieties arabica based on the features obtained. This method will take the features of texture on the surface of coffee beans by using statistical calculations. Features the texture consists of a first-order texture feature and a second-order texture feature.

This study aims to help farmers and experts from coffee roastery to identify varieties based on raw coffee bean morphology. By utilizing image processing technology, this study will use a sample of 2 Arabica coffee varieties namely Sigararutang and Lini S 795. Both of these varieties will be selected using texture feature with with Gray Level Co-occurrence Matrices (GLCM) and circularity features. The results of the pattern recognition method of these two varieties can be a reference to distinguish between the two types of varieties.

2. Related Work

Some research has used image processing to recognize coffee bean morphology. One of the research has combining first order and second order texture features [1] that the results of identify can be more accurate. This method is used because based on the appearance of the object on the green coffee beans image that has diversities texture of coffee variety. After getting the features, the next step is classification to recognize arabica coffee varieties. Classification techniques consist of different methods. Based on previous research, the classification process can be done with the Support Vector Machine (SVM) method [2], [3], Naïve Bayes Classifier (NBC) [4], and others using MultiLayer Perceptron (MLP).

Green bean can indentify by using GLCM method. The segmentation process is the way to separate the background with the object in the image using contours algorithm, then the next step is feature extraction using GLCM with four angle directions. The features of GLCM that uses are contrast, dissimilarity, homogeneity, energy, entropy, correlation and variance. The last stage is the classification through SVM method [3].

The Research on the classification of coffee beans uses several methods. The first stage is pre-processing of the digital image with Multiscale Retinex with Color Restoration algorithm (MSRCR). The method is compared with Histogram Equalization and Contrast Limited Histogram Equalization (CLAHE). The feature extraction using Color GLCM method and classification using SVM. The result shows that the visual quality and accuracy is better if using pre-processing method MSRCR[2].

Research conducted to identify the variety of coffee in the geographic region based on colour. The ANN method is used as a later transformation model NBC to recognize coffee beans consisting of four types: whitish, cane green, green, and bluish-green. The ANN method only reaches an error level of 1.15% and NBC yielded 100% accuracy [4]. Other features proposes the system to recognize arabica coffee varieties using several stages of feature extraction by combining the first and second order texture feature features[1]. Stage classification is using ANN method and obtained the average accuracy of 80%.

3. Research Methodology

This research consists of several stages. The first process is collecting image data, then the image uses as an input image for the pre-processing stage to prepare the image before being extracted. Next process is feature extraction use GLCM method and circularity to obtain texture and shape features. After that, we conduct feature selection by using Correlation-based Feature Selection to find features that have a significant effect on the classification results. The last is the classification process by using the MultiLayer Perceptron. Data selection and classification were carried out with 10-fold cross-validation test method, which means the data will be divided into 10 parts randomly, then performed 10 experiments where each experiment uses 10 data as test data and the rest as training data.

Fig. 1: Research Method.

2.1. Collecting image and pre-processing

The data used in this research is the raw bean coffee image that taken using a camera. The research requires 60 images from two varieties. Green bean coffee is an Arabica coffee that taken from 2 varieties of coffee in Indonesia that is Sigararutang variety (Mandailing single origin) and Lini S 795 (Toraja single origin). The image is an RGB image which is taken from two sides of green coffee bean because it is difficult to distinguish with only one site of green coffee bean. The data used is the original image taken from a camera with the JPG extension. In order to take the green coffee beans object, the image will crop.

2.2. Features extraction

The results of feature extraction with GLCM can be used as input in the process of feature selection by using the CFS method. The result of the feature selection process is received features that significantly affect to the process of classification.

Gray Level Co-occurrence Matrices (GLCM) is one method that can be used to extract texture features. GLCM is one of the second-order texture analysis methods [5]. GLCM have the parameter which is the direction and the gap between the pixel of neighbouring reference and grey level in the images. Each pixel could have neighbouring pixels of the eight directions, i.e. 0°, 45°, 90°, 135°, 180°, 225°, 270°, or 315°. However, the selection of the angle was 0° will produce the same value GLCM valuable to the angle of 180°. The concept also applied to an angle of 45°, 90°, and 135° [5]. Therefore, the angles used were 0°, 45°, 90°, and 135°. Eight directions adjacent GLCM can be seen in Figure 3.

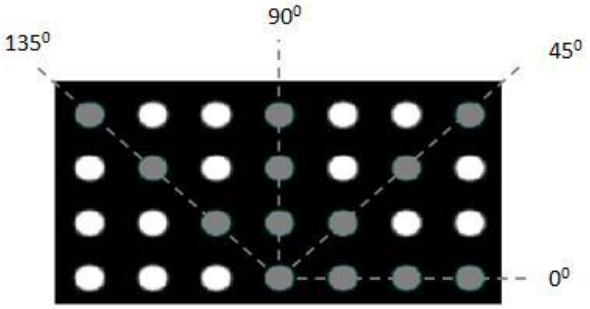

Fig. 2: Co-occurrence matrix directions for extracting texture features

Let f(x,y) is an image with size of N_x and N_y which has pixel with gray level probability (L level) and $\vec{r}$ is the offset vector. GLCM $\vec{r}$ (i,j) is defined as the number of pixels $\boldsymbol{j} \in \boldsymbol{1},\ldots,\boldsymbol{L}$, which happens to offset $\vec{\boldsymbol{r}}$ to pixels with $\boldsymbol{i} \in \boldsymbol{1},\ldots,\boldsymbol{L}$ that can be expressed by equation(1) as follows [6].

$$\begin{aligned} \boldsymbol{GLCM}_{\vec{r}}(\boldsymbol{i},\boldsymbol{j}) &= \#\{(\boldsymbol{x_1},\boldsymbol{y_1}),(\boldsymbol{x_2},\boldsymbol{y_2}) \\ &\quad \in (\boldsymbol{N_x},\boldsymbol{N_y})\boldsymbol{X}(\boldsymbol{N_x},\boldsymbol{N_y})|\boldsymbol{f}(\boldsymbol{x_1},\boldsymbol{y_1}) \\ &= \boldsymbol{j}^{\vec{r}} = \overrightarrow{(\boldsymbol{x_2}-\boldsymbol{x_1},\boldsymbol{y_2}-\boldsymbol{y_1})}\} \end{aligned} \quad (1)$$

In this case, # shows the number of set elements. Offset $\vec{\boldsymbol{r}}$ is the direction or distance.
To obtain GLCM features, only a few scales are used, namely angular second moment (ASM) or energy, contrast, correlation and homogeneity. ASM is a homogeneity measure of the image and gives the numbers of squared elements. It can be calculated using equation (2) below [6].

$$\boldsymbol{ASM} = \sum_{\boldsymbol{i}=\boldsymbol{1}}^{\boldsymbol{L}}\sum_{\boldsymbol{j}=\boldsymbol{1}}^{\boldsymbol{L}} \boldsymbol{GLCM}(\boldsymbol{i},\boldsymbol{j})^{\boldsymbol{2}} \quad (2)$$

In this case, L is the number of levels that have been used for computation.
Contrast (Cn) is the measure of variety existence of grey level pixel of image calculated by using equation (3) as follows [6].

$$\boldsymbol{Cn} = \sum_{\boldsymbol{n}=\boldsymbol{1}}^{\boldsymbol{L}} \boldsymbol{n}^{\boldsymbol{2}}\{\sum_{|\boldsymbol{i}-\boldsymbol{j}|=\boldsymbol{n}} \boldsymbol{GLCM}(\boldsymbol{i},\boldsymbol{j})\} \quad (3)$$

Correlation (Cr) is a measure of linear interdependence of grey level values or how big the relation between one pixel and the neighbour pixel in the image is, calculated by using equation (4) as follows [6].

$$\boldsymbol{Cr} = \frac{\sum_{\boldsymbol{i}=\boldsymbol{1}}^{\boldsymbol{L}}\sum_{\boldsymbol{j}=\boldsymbol{1}}^{\boldsymbol{L}}(\boldsymbol{i},\boldsymbol{j})(\boldsymbol{GLCM}(\boldsymbol{i},\boldsymbol{j}) - \boldsymbol{\mu_i'}\boldsymbol{\mu_j'}}{\boldsymbol{\sigma_i'}\boldsymbol{\sigma_j'}} \quad (4)$$

With :

$$\mu_i' = \sum_{i=1}^{L}\sum_{j=1}^{L} i * GLCM(i,j)$$

$$\mu_{ji}' = \sum_{i=1}^{L}\sum_{j=1}^{L} j * GLCM(i,j)$$

$$\sigma_j^2 = \sum_{i=1}^{L} \sum_{j=1}^{L} GLCM(i,j)(i-\mu_i')^2$$

$$\sigma_i^2 = \sum_{i=1}^{L} \sum_{j=1}^{L} GLCM(i,j)(i-\mu_i')^2$$

Homogeneity of inverse Different Moment (IDM) is a measure of proximity of elements distribution in GLCM which is calculated by using equation (5) below [6].

$$\boldsymbol{IDM} = \sum_{\boldsymbol{i=1}}^{\boldsymbol{L}} \sum_{\boldsymbol{j=1}}^{\boldsymbol{L}} \frac{\boldsymbol{GLCM(i,j)^2}}{\boldsymbol{1+(i-j)^2}} \quad (5)$$

Besides features of GLCM, shape feature (circularity) is also applied, and added as features of the selection process. Circularity is one of the important features that are commonly used in geometric features. Circularity is usually defined as follows [7]:

$$CI = 4\pi * \frac{area}{perimeter^2} \quad (6)$$

CI is an output that have value 1 if the object approaches a circle. Perimeter indicates the edge length of an object that can be calculated by the equation as follows [7]:

$$perimeter = \sum_{i=1}^{n} l_i \quad (7)$$

l_i is the pixels on the edge of the object obtained by using a four-direction chain code that can be defined as follows [6].

$$l_i = \overline{P_i P_{i-1}} = 1, \varepsilon_i = 0,1,2,3, \quad (8)$$

Perimeter consist of $P_0, P_1, \ldots, P_{n-1}, P_0$, and n present as total area that comes from a formula [8]:

$$area = \sum_{i=1}^{n} a_{ix}(y_{i-1}\frac{1}{2}a_{iy}), \quad (9)$$

with

$$y_i = \sum_{i=1}^{i} a_{jy} + y_0 \quad (10)$$

n is the number of chain codes, a_{ix} and a_{iy} present as component of x and y from the direction chain, y_i is coordinate at each, y_0 is coordinate for the starting point. At the first time, the algorithm calculates the perimeter and area by using chain code method and get the value of *CI*. In the implementation, a standard round shape is difficult to obtain so that the threshold value *s* that is set at the beginning. When *CI* is higher than s, the circularity can be detected based on the criteria of *s*.

2.3. Features selection

The feature selection use in this research is CFS algorithm. Correlation based Feature Selection (CFS) is algorithm that works by trying to find a feature subset with the purpose to reduce the dimensions of the dataset and improve classification accuracy [9]. This technique uses the Pearson correlation coefficient which is correlation metrics designed to look for features that have a high correlation with the class and features. The Pearson correlation coefficient is written using a formula [10]

$$M_S = \frac{k\overline{r_{cf}}}{\sqrt{k + k(k-1)\overline{r_{ff}}}} \quad (11)$$

M_S or "Merit" which indicates how feasible the S feature, *k* is a sum of feature. $\overline{r_{cf}}$ the average correlation between each feature with its class and $\overline{r_{ff}}$ is the paired correlation average between two features..

2.3. Classification

The result of features extraction would use as an input for the classification process using the Multi-Layer Perceptron method which aims to recognize varieties of coffee varieties. Multilayer perceptron is a development of the Perceptron Neural Network model which was developed in the early 1960s. Neural Networks have many layers that are limited to reducing time to solve existing problems. It takes several steps to run the ANN classification using the MLP architecture, which starts with data collection, then creates and configures the network. Then initialize the weights and biases. After the network can do training, data validation and use during classification. The output of the ANN can be computed by using equation (12) as follow [11]

$$O = f(IW_{io}) \quad (12)$$

W_{io} is weight matrix with the size i x o, *i* is the number of input nodes, *o* is the number of output nodes, *I* is the input vector and *O* output vector.

In general, the data are presented at the input layer, and then the network will conduct input process by multiplying the input and the weight layer. To make it easier to understand the way how the MLP works, algorithm [11] can be used as the following.

1. Conducting network initialization by randomly arranged all weights between -1 and +1.
2. Presenting the first training pattern on the existing network and store the output results.
3. Comparing the network output with existing output targets.
4. Fixing the error in a backward manner by :
   a. Fixing the output weight layer by using equation (13)

$$\omega_{h0} = \omega_{h0} + (\eta\delta_0 O_h) \quad (13)$$

ω_{ho} is the weight value of unit h which is hidden by the output of unit o, η is the training ratio, o_h is the output of unit h which is hidden, in which :

$$\delta_0 = O_0(1 - O_0)(t_0 - O_0) \quad (14)$$

O_0 is the node o of the output layer and t_0 is the output target for the node.

   b. Fixing the input weight by using equation (15)

$$\omega_{ih} = \omega_{ih} + (\eta\delta_h O_i) \quad (15)$$

ω_{ih} is the weight value of unit h which is hidden by the input of unit *i*, η is the training ratio, o_i is the input from node *i,* in which

$$\delta_h = O_h(1 - O_h)\sum_{o}(\delta_0\omega_{h0}) \quad (16)$$

5. Calculating error, by counting the average of the target value and the output vector. Calculation error was conducted by using equation (17) :

$$E = \frac{\sqrt{\sum_{n-1}^{p}(t_0 - o_0)^2}}{p} \quad (17)$$

where P = the number of units in the output layer.

6. Repeating steps No.2 for each pattern in the training dataset to complete one epoch.
7. Conducting exchange of training dataset randomly. This is to reduce the possibility of network affected by the order of the data.
8. Repeating step 2 to a number of epochs or until the error began to change.

4. Result and Discussion

This research proposes to use GLCM method because it is able to recognize features based texture. Besides texture features, shape features (circularity) are also used to identify the coffee varieties. The result of this extraction is a form of numbers that can be measured. Figure 1. The following is an example of green bean coffee image that taken from the front side and back side of Sigararutang variety and Lini S 795.

(a) Sigararutang (Mandailing Single Origin)

(b) Lini S 795 (Toraja Single Origin)

Fig. 3: Green bean coffee image which taken from camera

In order to focus the research only on coffee bean objects and speed up the computation process, the image is cropped manually. After cropping, the image colour converts from RGB to grayscale (figure 2).

(a) *grayscale* – Sigararutang

(b) *grayscale* – Lini S 795

Fig. 4: Grayscale image of green coffee bean

Based on the image obtained, it is difficult to identify the difference of shape between green bean coffee image each variety. To solve this problem, the next step is feature extraction using GLCM which will generate a number where one image has four features and each feature consists of four different directions so that one image has 16 GLCM features. This method can be used after a grayscale process whereas *circularity* needs to be done in pre-process stage first. Segmentation process used to analyze the shape of green bean and omit other object. The Segmentation process use Otsu method and result show in figure 3.

(a) Segmentation result of Sigararutang image

(b) Segmentation result of Lini S 795 image

Fig. 5: Segmentation result.

The Classification process use 16 feature of GLCM. The result show in table 1 with confusion matrix from GLCM feature classification

Tabel 1: Confusion Matrix GLCM Classification

		Target	
		Sigararutang	**Lini S 795**
Prediction	**Sigararutang**	22	8
	Lini S 795	7	23

Table 1 illustrate confusion matrix that shows the classification result with MLP and obtained a value of *True Positive* (TP)=22, *True Negative*(TN)=23, *False Negative* (FN)=8 and *False Positive* (FP)=7. By using 30 image of Sigararutang, MLP can identify 22 image of green bean coffee as Sigararutang and 8 image identify as Lini S 795. Meanwhile, by using 30 image of Lini S 795, MLP can identify 23 image of green bean coffee as Lini S 795 and 7 image identify as Sigararutang. Based on confusion m atrix, the result can show a value of accuration, sentitvy, and specification. The result shows that MLP can identify the variety of coffee with accuration 75%, sensivity 73.33% and specification 76,7%.

Based on the features used, there is the possibility that features has affect the classification results and there is also a feature that has absolutely no effect on the classification. Therefore, the feature selection process is performed using CFS to find the features that most significantly influence. The result of feature selection is shown in Figure 3.

Figure 3. The image shows that of the 16 features or attributes used, there is only 1 feature that have significant in the identification process of coffee varieties that is the contrast 0° feature. The result of this feature selection combined with circularity feature, and classification using contrast 0° and circularity feature. Table 2 shows confusion matrix using 2 features.

Tabel 2: Confusion Matrix using 2 features

		Target	
		Sigararutang	**Lini S 795**
Prediction	**Sigararutang**	27	3
	Lini S 795	3	27

Table 2 is a confusion matrix that describes the result of classification and obtained True Positive (TP) value = 27, True Negative (TN) = 27, False Negative (FN) = 3 and False Positive (FP) = 3. This means that from 30 images of coffee mandailing, MLP is able to recognize as image of Sigararutang variety as much as 27 image while 3 image is recognized as Lini S 795 variety. Furthermore, from 30 Lini S 795 variety image, MLP is able to recognize as image of Lini S as much as 27 images while 3 images are recognized as Sigararutang. Based on confusion matrix then can be calculated level of accuracy, sensitivity and spesifisitas. For the introduction of coffee varieties, obtained an accuracy of 90% with a sensitivity of 90% and specificity of 90%.

The GLCM method proposed in this research uses several features: contrast, energy, correlation and homogeneity with four different direction angles. These features are then used as input in the feature extraction process, then the results obtained are classified using MLP. After the classification process, 75% accuracy was obtained with a sensitivity of 73.33% and 76.7% specificity. The results are still low enough to be used as reference models in the classification process. therefore, a feature selection process is undertaken to look for features that affect the classification results significantly. After feature selection process using CFS, the contrast feature is found as the most influential feature. The classification is done again using the contrast feature and the results are shown in Table 3 below.

Tabel 3: Confusion Matrix Classification Contrast 0° Feature

		Target	
		Sigararutang	**Lini S 795**
Prediction	**Sigararutang**	22	8
	Lini S 795	12	18

Table 3 is a confusion matrix that describes the classification results using the contrast00 feature and obtained the value of True Positive (TP) = 22, True Negative (TN) = 18, False Negative (FN) = 8 and False Positive (FP) = 12. This means that from 30 images of mandailing coffee, MLP is able to recognize as image of Sigararutang variety as many as 22 images while 8 images are recognized as Lini S 795. Furthermore, from 30 images of Lini S 795, MLP is able to recognize as Lini S 795 of 12 images while 18 images are recognized as Sigararutang image. Based on the confusion matrix, 66.7% accuracy was obtained with a sensitivity of 73.3% and a specificity of 60%. The results are much lower when compared to using all the proposed GLCM features.Coffee beans can also be recognized by shape, so the circularity feature is also used for separate classification processes with GLCM features. The results are shown in Table 4 below.

Tabel 4: Confusion matrix classification by circularity

		Target	
		Sigararutang	**Lini S 795**
Prediction	**Sigararutang**	25	5
	Lini S 795	7	23

Table 4 is a confusion matrix that describes the classification result using the circularity feature and obtained True Positive (TP) value = 25, True Negative (TN) = 23, False Negative (FN) = 5 and False Positive (FP) = 7. This means that from 30 images of Sigararutang, MLP is able to recognize as image of Sigararutang as much as 25 image while 5 image is recognized as Lini S 795. Furthermore, from 30 images of Lini S 795, MLP is able to recognize as Lini S 795 of 23 images while 7 images are recognized as Sigararutang. Based on the confusion matrix, 80% accuracy was obtained with sensitivity of 83.3% and specificity of 76.7%. The result is much higher than the classification using the contrast feature. The accuracy can improve by combining texture feature in contrast and share feature (circularity). Both features are classified using MLP. The results are shown in Table 2.

5. Conclusion

In the traditional method, identification of coffee varieties use a farmer experiences. Sorting coffee varieties is subjective by observing colour attributes, sizes, shapes. As a result, the objectivity of coffee varieties and quality determination becomes inconsistent. Image processing might be useful as the second opinion to help this process. GLCM method can use to identify the features of coffee varieties that use by the farmer in the sortation process. In this research also use shape feature (circularity) to identify coffee bean varieties.

The results of this research combine contrast texture feature on GLCM and shape feature circularity. Both features are classified using MLP. By using 30 images of Sigararutang variety (Mandailing single origin coffee), MLP is able to recognize the 27 images as Sigararutang variety, while 3 images are recognized as Lini S 795 (Toraja single origin coffee). Furthermore, by using 30 Lini S 795 variety (Toraja single origin coffee), MLP is able to recognize 27 images while 3 images are recognized as Sigararutang variety. The result shows an accuracy of 90% with a sensitivity of 90% and specificity of 90%.

Acknowledgement

The research collaborates with black java coffee roastery and garasi roastery that provide the raw coffee bean and coffee expert. The research funded by STMIK AKAKOM Yogyakarta, Indonesia.

References

[1] Radi, Muhammad Rivai, and Mauridhi Hery Purnomo, "Combination of first and second order statistical features of bulk grain image for quality grade estimation of green coffee bean," *ARPN J. Eng. Appl. Sci.*, vol. 10, no. 18, Oct. 2015.

[2] R. G. Apaza, C. E. Portugal-Zambrano, J. C. Gutiérrez-Cáceres, and C. A. Beltrán-Castañón, "An approach for improve the recognition of defects in coffee beans using retinex algorithms," in *Computing Conference (CLEI), 2014 XL Latin American*, 2014, pp. 1–9.

[3] R. H. M. Condori, J. H. C. Humari, C. E. Portugal-Zambrano, J. C. Gutiérrez-Cáceres, and C. A. Beltrán-Castañón, "Automatic classification of physical defects in green coffee beans using CGLCM and SVM," in *Computing Conference (CLEI), 2014 XL Latin American*, 2014, pp. 1–9.

[4] E. M. de Oliveira, D. S. Leme, B. H. G. Barbosa, M. P. Rodarte, and R. G. F. A. Pereira, "A computer vision system for coffee beans classification based on computational intelligence techniques," *J. Food Eng.*, vol. 171, pp. 22–27, 2016.

[5] G. Liu, R. Wang, Y. Deng, R. Chen, Y. Shao, and Z. Yuan, "A new quality map for 2-D phase unwrapping based on gray level co-occurrence matrix," *IEEE Geosci. Remote Sens. Lett.*, vol. 11, no. 2, pp. 444–448, 2014.

[6] M. M. Sebatubun, C. Haryawan, and B. Windarta, "Classification of ground glass opacity lesion characteristic based on texture feature using lung CT image," *J. Exp. Theor. Artif. Intell.*, vol. 30, no. 2, pp. 203–215, 2018.

[7] Z. Fu and Y. Han, "A circle detection algorithm based on mathematical morphology and chain code," in *Computing, Measurement, Control and Sensor Network (CMCSN), 2012 International Conference on*, 2012, pp. 253–256.

[8] Z. Fu and Y. Han, "A Circle Detection Algorithm Based on Mathematical Morphology and Chain Code," in *2012 International Conference on Computing, Measurement, Control and Sensor Network*, 2012, pp. 253–256.

[9] H. Ghaderi and P. Kabiri, "Fourier transform and correlation-based feature selection for fault detection of automobile engines," in *Artificial Intelligence and Signal Processing (AISP), 2012 16th CSI International Symposium on*, 2012, pp. 514–519.

[10]R. Wald, T. M. Khoshgoftaar, and A. Napolitano, "Using correlation-based feature selection for a diverse collection of bioinformatics datasets," in *Bioinformatics and Bioengineering (BIBE), 2014 IEEE International Conference on*, 2014, pp. 156–162.

[11]L. Noriega, "Multilayer perceptron tutorial," *Sch. Comput. Staffs. Univ.*, 2005.