Texture-Based Woven Image Classification using Fuzzy C-Means Algorithm

Soetrisno, Dwi Ratna Sulistyaningrum and Isi Bifawa'idati

***Abstract*—There are a lot of texture-based image data stored in the storage media Internet. Most of these data portray the cultural fabric texture results from a State. Because of the many variants of the existing texture, the data need to be easily accessible through the Internet. Moreover, the area of origin of weaving the surface is easily known. Therefore, it is necessary to develop a classification system based on woven image data. The texture of the image data stored in a database on the Internet can be grouped/clustered well, making it easy to access. This study examines a texture-based woven image classification using fuzzy c-means algorithm. This method combines extraction methods Gabor filter, fuzzy c-means algorithm and Euclid distance similarity measure. An experiment was done using the system as many as 60 woven images from Bali, NTT and Central Java areas, each taken as many as 25 images weaving. The test results stated that testing using the test images taken from the images in the database generates a 100% accuracy rate, and testing using test images taken from outside the database produces an accuracy rate of 94%.**

***Index Terms*—Image classification, feature extraction, Gabor filters, fuzzy C-Means algorithm.**

I. Introduction

All States have a culture that has continuously been developed and has always sought to preserve. Many countries build culture by producing textured fabrics and tried to introduce them through the Internet. Because of many texture-based images of weaving on the Internet, it's easier to find the area of origin. However, the system must store all these images in groups that show from where they originated. Therefore, it is necessary to develop a classification system to classify woven images based on image texture

Many researchers have researched and developed several methods of image classification. A research on image classification based on the texture has been done in [1]. This study classifies medical data by extracting the image features using wavelet and then performs interval training type-2 fuzzy logic system with fuzzy c-means clustering and supervised parameter tuning with genetic algorithm. The result is a classification of patients based on the type of cancer that affects the patient. Then, there is another research on Rock Lithological Classification Using Multi-Scale Gabor Feature Sub-images and Voting Rock Contour Information [2]. In the study, feature extraction using Gabor filter and further each sub-image is classified based class-stone carvings using support vector machine (SVM). Pusparini et.al. [3] have used fuzzy set theory to do bullet defect detection. The results of this study provide that the accuracy of the method used is 70.58%. Therefore, this study examines the texture-based woven image classification using the combination Gabor filter with the fuzzy c-means method.

The authors are with the Department of Mathematics, Institut Teknologi Sepuluh Nopember, Kampus ITS-Sukolilo, Surabaya 60111, Indonesia e-mail: soetrisno@matematika.its.ac.id, dwiratna@matematika.its.ac.id, isyi-faid@gmail.com.

Manuscript received July 2, 2021; accepted January 12, 2022.

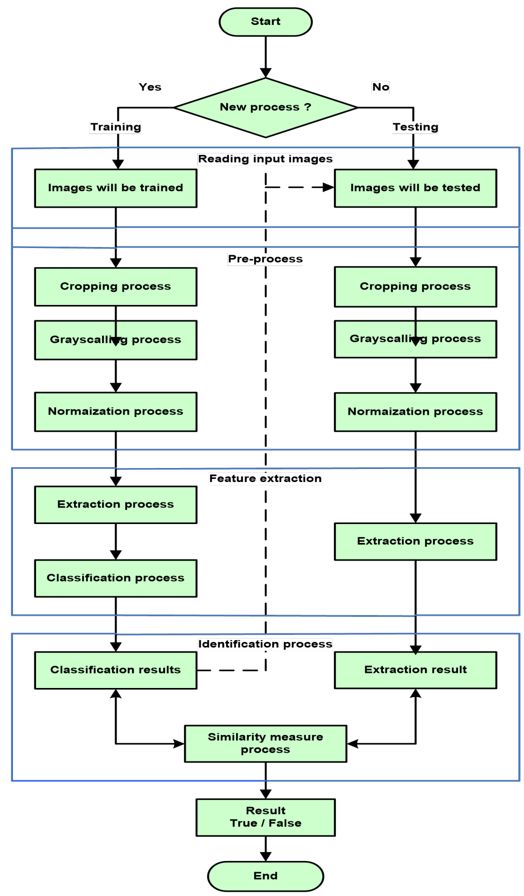

Fig. 1. The process flow diagram.

II. The Method of Classifications

What is meant by the texture-based image classification in this study was based on image grouping woven textures that are accessed online into groups where each group has a specific set of characteristics. The method developed has several stages, and a flowchart of this method can be seen in Fig. 1.

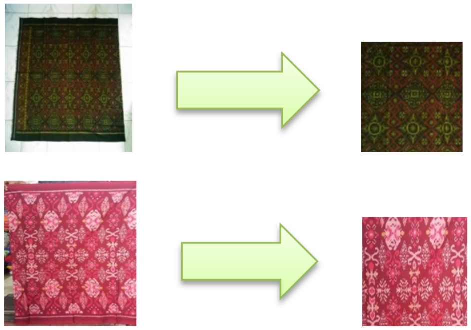

Fig. 2. Some results of the image cropping process.

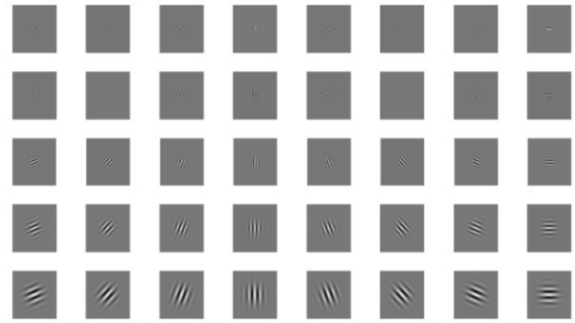

Fig. 3. Extraction result images of a trained image.

A. Pre-Processing

Texture-based image data on the Internet have different sizes, rectangular-shaped uprights, horizontals, rectangular and squares. No texture distribution meets space, and some do not meet the space, so that every image needs to have the best part with a regular pattern and meeting space that can represent the image. On average, the part with a pattern that meets the length and texture regularly is at the centre of the image, resulting in a pre-process each image is cut off (cropped) in the middle of the same size, which in this study been sized 300×300 pixels. Some results of the cropping process can be seen in Fig. 2.

The grey scaling process is then performed using (1), which is a process to change the RGB image (Red, Green, Blue) to grayscale images.

$$grayscale = \frac{R+G+B}{3} \tag{1}$$

and then each image is normalized using (2)

$$N(x,y) = \begin{cases} M_0 + \sqrt{\dfrac{V_0\{(I(x,y)-M)^2\}}{V}}, & I > M \\ M_0 - \sqrt{\dfrac{V_0\{(I(x,y)-M)^2\}}{V}}, & I < M \end{cases} \tag{2}$$

where $N(x,y)$ is the normalized grey level value, $I(x,y)$ is the intensity value at (x,y) pixel, M is the average of intensity value, M_0 is the average of intensity value wanted, V is the variant of pixel intensity value in the image, and V_0 is the desired variant of pixel intensity value.

B. Feature Characteristic Extraction using Gabor Filter

At this stage, the texture feature extraction from each image using the Gabor filter is stated by (3), as in [4].

$$G(x,y) = \frac{1}{2\pi\sigma_x\sigma_y}\exp\left\{-\frac{1}{2}\left[\frac{x^2+y^2}{\sigma_x^2+\sigma_y^2}\right]\right\} \exp\left[2j\pi(ux\cos\theta + uy\sin\theta)\right] \tag{3}$$

where σ_x and σ_y respectively a Gaussian distribution direction x and the direction y and here assumed σ_x and σ_y are the same, u_0 with σ_x and σ_y respectively a Gaussian distribution towards x is the central frequency of sinusoidal and orientation of sinusoidal [5], [6], [7]. Here we chose some satisfied frequency parameters; in this case, we decided on five frequency parameters, that is 4, 5.6, 8, 11.3, and 16. We also chose some satisfied orientation angle parameters; in this case, we selected eight orientation angle parameters, that is, $0°$, $23°$, $45°$, $68°$, $90°$, $113°$, $135°$, and $158°$. Furthermore, we count the magnitude of convolution results of the Gabor filter where the magnitude is a complex value of Gabor filter results that is a sum of the imaginer and the real filter. There were 40 magnitudes, where from the first magnitude, we chose the maximum value of its entry by using the Wo window procedure with the measure of 4×4. The maximum value of the entry in the Wo window is called feature points. So, each extraction result image gives 75×75 feature points, and the average value of those are projected to be 1 to 40. Thus, the feature extraction result is a characteristic vector (feature vector) with 40 components of the pre-processing result image. The example of the extraction result image is shown in Fig. 3.

C. Classification of Extraction Results using Fuzzy C-Means Algorithm

At this stage, classification was done by using fuzzy c-means algorithm. After the characteristic vectors of the texture feature, the data clustering process was completed. The data clustering process divides data elements into some clusters so that the elements in the same clusters are similar and the element in the different clusters are not similar. This process was done by using fuzzy c-means algorithm [8], [9] and has steps as below:

Step-1 Initialize parameters: A matrix, X, ordered $m \times n$ where m is the number of data clustered, and n is the number of feature vector components. The number of clusters will be built, $C = 3$. Weight, $w = 2$. The maximum number of iterations is 100—the criteria of stopping, $\varepsilon = 0.01$.

Step-2 Make an initial partition matrix, U (degree of membership in a cluster) is made random.

$$U = \begin{bmatrix} \mu_{11}(x_1) & \mu_{12}(x_2) & \dots & \mu_{1m}(x_m) \\ \mu_{21}(x_1) & \mu_{22}(x_2) & \dots & \mu_{2m}(x_m) \\ \vdots & \vdots & & \vdots \\ \mu_{C1}(x_1) & \mu_{C2}(x_2) & \dots & \mu_{Cm}(x_m) \end{bmatrix}$$

Step-3 Counting the matrix of cluster central, V, using (4)

$$V_{ij} = \frac{\sum_{k=1}^{m} (\mu_{ik})^2 x_{kj}}{\sum_{k=1}^{m} (\mu_{ik})^2} \tag{4}$$

Step-4 Repairing the degree of membership of each data in each cluster using (5)

$$\mu_{ik} = \left[\sum_{j=1}^{c} \left(\frac{d_{ik}}{d_{jk}} \right)^2 \right]^{-1}, \tag{5}$$

where

$$d_{ik} = d(x_k - v_i) = \left[\sum_{j=1}^{n} (x_{kj} - v_{ij}) \right]^{1/2}$$

Step-5 Evaluate the criteria of stopping of iteration. If $\Delta < \varepsilon$, then the iteration process stops.

$$\Delta = \| U^t - U^{t-1} \| \tag{6}$$

D. Counting the Similarity level of Images

This stage is the system testing stage. In this stage, the similarity level of two images is measured using Euclidean distance between the characteristic vectors of image input that are classified and the central vector of a cluster. The formula to count Euclidean distance is stated by equation (7).

$$d(X,Y) = \sqrt{\sum_{i=1}^{n} (x_i - y_i)^2} \tag{7}$$

Here, X is any vector pointed to the centre of an image cluster in a database, and Y is a feature vector of the input image in the testing process. If the feature vector of the input image, Y, has a minimum distance to the central vector of a cluster, C_i, then that input image is decided as a member of image cluster C_i.

III. Testing and Discussion

The classification system testing is done using software of MATLAB. The data used in the testing process is some 75 data woven images from Bali, NTT, and Central Java; each area took 25 image data. From these 75 woven image data, we take 60 image data for the testing process, where each area takes 20 image data. The testing process is divided into two categories. First, the testing process was carried out ten times; we selected 10 data were randomly from 60 training image data in each repetition. Second, the testing process was carried ten times. Each repetition uses 10 test data chosen randomly from 15 image data not used in the training process. At this testing process, the frequency parameters chose is several five different values, that is 4, 5.6, 8, 11.3, and 16; and we decided to have 8 orientation angle parameters, that is 0°, 23°, 45°, 68°, 90°, 113°, 135°, and 158°. To measures the level of accuracy, we used the formula is stated by (8).

$$accuracy = \frac{B + N + J}{numberOfTestingImages} \times 100 \tag{8}$$

TABLE I
Results of Testing

Test	Data testing (true)	Level of accuracy	Not data testing (true)	Level of accuracy
1	10	100%	9	90%
2	10	100%	9	90%
3	10	100%	10	100%
4	10	100%	9	90%
5	10	100%	10	100%
6	10	100%	10	100%
7	10	100%	9	90%
8	10	100%	9	90%
9	10	100%	10	100%
10	10	100%	9	90%
	Average	100%	Average	94%

where B represents the number of Bali images with accurate results, N represents the number of NTT images with accurate results, and J represents the number of Central Java images with accurate results.

Table I shows that a test using test data retrieved from the database generates an accuracy rate of 100% correctly identified. Still, a test using test data taken outside the database generates an accuracy rate of 94% accurately determined.

The leading cause is not identified correctly according to the image of the class include:

1) The image was acquired has a texture very similar to another image coming from different regions. For example, weaving Bali has a texture similar to the surface of weaving almost Central Java. So that extraction characteristics also become relatively identical.
2) The quality of the image acquired is not so good. Texture to the image is less visible, so that the system cannot find a robust characteristic feature of the image.

IV. Conclusions and Future Works

Based on the test results, we can be concluded that the texture-based woven image classification system using the fuzzy c-means algorithm produces a 100% accuracy rate in tests using test images part of the image database. The classification system of woven image-based texture using fuzzy c-means algorithm produces a 94% accuracy rate in tests using test images that are not a part of the image database. The identification error is due to the woven texture of a class almost resembles a woven texture of the other class and the image quality is woven tested is less, resulting in better image characteristics loom less assertive.

In this study, the data involved is still limited to image data looms from regions with variations in texture weaving. The training process is carried out at the level many still images uniform database of each class/region. Additionally, the system to access the data require user involvement. Because it so that the system ready to be used online still thought necessary to develop a system that can access image databases that are not uniform.

REFERENCES

[1] T. Nguyen, A. Khosravi, D. Creighton, and S. Nahavandi, "Medical data classification using interval type-2 fuzzy logic system and wavelets," *Applied Soft Computing*, vol. 30, pp. 812–822, 2015.

[2] C. Perez, J. Saravia, C. Navarro, D. Schulz, C. Aravena, and F. Galdames, "Rock lithological classification using multi-scale gabor features from sub-images, and voting with rock contour information," *International Journal of Mineral Processing*, vol. 144, pp. 56–64, 2015.

[3] E. Pusparini, B. Setiyono, and D. Sulistyaningrum, "Bullets defect detection based on digital image processing using line detection and fuzzy sets," *Global Journal of Pure and Applied Mathematics*, vol. 11, no. 4, pp. 2215–2222, 2015.

[4] R. Carmona, W.-L. Hwang, and B. Torresani, *Practical Time-Frequency Analysis: Gabor and wavelet transforms, with an implementation in S*. Academic Press, 1998.

[5] J. Raheja, S. Kumar, and A. Chaudhary, "Fabric defect detection based on glcm and gabor filter: A comparison," *Optik*, vol. 124, no. 23, pp. 6469–6474, 2013.

[6] D. Sulistyaningrum, B. Setiyono, J. Anita, and M. Muheimin, "Measurement of crack damage dimensions on asphalt road using gabor filter," in *Journal of Physics: Conference Series*, vol. 1752, no. 1, 2021, p. 012086.

[7] K.-L. Mak, P. Peng, and K.-F. Yiu, "Fabric defect detection using multi-level tuned-matched gabor filters," *Journal of Industrial & Management Optimization*, vol. 8, no. 2, p. 325, 2012.

[8] J. Bezdek, R. Ehrlich, and W. Full, "Fcm: The fuzzy c-means clustering algorithm," *Computers & geosciences*, vol. 10, no. 2-3, pp. 191–203, 1984.

[9] Z. Ji, Y. Xia, Q. Chen, Q. Sun, D. Xia, and D. Feng, "Fuzzy c-means clustering with weighted image patch for image segmentation," *Applied soft computing*, vol. 12, no. 6, pp. 1659–1667, 2012.