Banten Batik Classification with Backpropagation Neural Network

Rian Fahrizal[1]*' *Ridwan Prasetya* Parlindungan Siahaan[2], and *Romi* Wiryadinata[3]

[1,2,3] Department Electrical Engineering, Universitas Sultan Ageng Tirtayasa

Abstract. Batik Cloth is Indonesia heritage of the indigenous cultures. However, the lack of public awareness could made the other nations has claimed that Batik cloth as their heritage. Therefore, We will need serious attention to take the problems. The classification Banten pattern batik cloth had been made to protect the Batik pattern. The classification had been made with Artificial Neural Network with Lavender-Marquant methods. Gray-Level Co-Occurrence Matrix (GLCM) has been used to preprocessing the pattern from Gray scale images. The results showed that the Batik patterns had 97% accuracy and 2% error.

1 Introduction

One of the Indonesian heritage was Batik. It has been recognized as an intangible cultural heritage [1]. However, the lack of public awareness of national culture led to the recent years, batik was claimed and recognized as another nation's culture. Therefore, serious attention was needed to prevent this accident. There were many batik pattern in Indonesia, one of kind was Banten Batik. These batiks had the distinctive color, which is the soft gray color. It showed the character Banten Citizen, that showed the dream, ideas, willing, and the high tempered but the appearance was simple and calm. The Batik color showed the natural environment in Banten. The Batik Banten pattern were Datu Laya, Kapurban, Kawangsan, Mandalikan, Pamaranggen, Pancaniti, Pasepen, Pasulaman, Pejantren, Sabakingking, Srimanganti, and Surosowan. Batik Banten had 10 classification class.

Some methods had been used to clssified Batik Pattern. Batik classification's method were Support Vector Machine [2], Neural Network [3], and Decession Tree Classification [4]. The neural network had great advantages in parallel computation for classified the pattern. Neural network classification could recognized more than 10 class and fast learning convergence for classification. Another advantages were it could adaptive data driven self method that is able to regulate itself against the data that will be processed without requiring an explicit specification function or the form had been distributed, the apportion function of the general public which had accuracy in the classification process, a nonlinear model so that this model became flexible in modeling complex relationships, estimated the subsequent possibilities that provide the basis for classification rules and statistical analysis. Neural networks had been used to classified Batik Banten pattern because of their

---

* Corresponding author: rian.fahrzal@untirta.ac.id

advantages. The variety of batik pattern that vary in Indonesia made it difficult for the general public to identified batik patterns in each region in the archipelago. The same goes for the Banten batik motif which had approximately 10 batik motifs. This study tried to classify Banten batik motifs that aim to facilitated the public who do not knew the specifications of batik motifs in Indonesia, especially Banten regional batik motifs. Pattern feature extraction that is used as input to the system that is built is GLCM (Gray Level Co-occurrence Matrix) using a gray degree distribution (histogram). The final step is to classify using the artificial neural network method, that is, Backpropagation is used to classify Banten's batik motifs that are more accurate.

Banten Batik is one of the historic batik originating from Banten Province. In this paper, extensive information was needed in processing Banten Batik so that it matched the concept carried without changing the authenticity of Banten Batik. These were batik Banten pattern showed in Figure 1:

Fig. 1. Batik Banten

2 Research Method

Input image were batik images obtained from batik craftsmen using a Canon D550 camera. The banten batik taken was 300 images consisting of 200 training images and 100 test images. which consists of 10 classes patterns. The methods of artificial neural networks used in this research was backpropagation, in general the algorithm was referred to as back propagation because it can be described as follows, when the network is given an input pattern as a training pattern then the pattern goes to units in the hidden layer to be forwarded to the unit - output layer unit, then the output layer units gave a response called network output. If the output network is not the same as the expected output, the output spreads backward, in the hidden layer is forwarded to the unit at the input layer, therefore the training mechanism is called backpropagation.

Figure 2. explained the algorithm of an artificial neural network design, when the user starts to provide input training data, the training process will be done to get optimal weight on each layer. After optimal weight is obtained, then the optimal weight is stored for the testing process and information on the case of the classification of batik motifs. The results and tests will be compared with the existing targets to determine the system's ability to study the training data provided to the system.

2.1 Preprocessing Image

Batik's images taken from the camera and cropping it with different pixels size. Simplification had been made to resize the images. The resizing image became 640 x 640

pixels with RGB value. GLCM methods needed gray scale image. Resizing images converted to the gray images with *doubles floating point* values. Contrast Limited Adaptive Histogram Equalization (CLAHE) was implemented to enhancement gray images. This method was developed to prevent the over amplification of noise in picture.
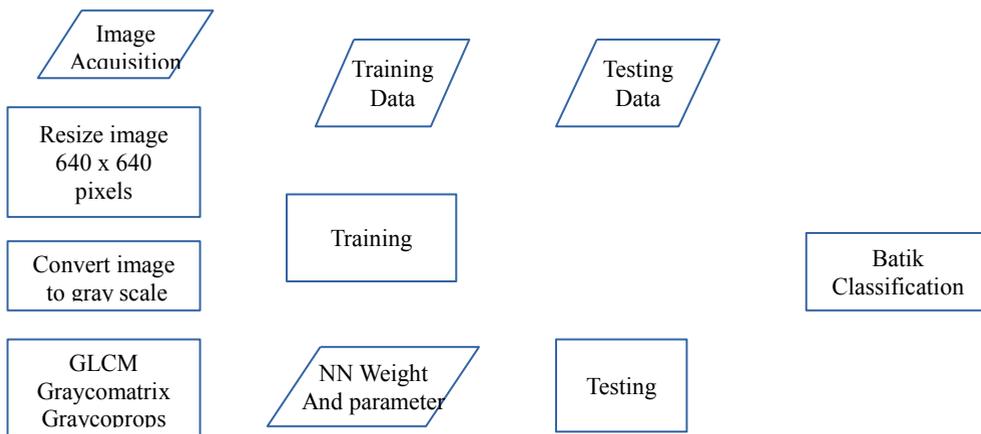
Fig. 2. Batik Banten Classification Algorithms

This method consist the following steps:
1. Collect all the inputs: Image, Number of regions in row and Column paths, Number of bins for the histograms used in image transform function (dynamic range), Clip limit for contrast limiting (normalized from 0 to 1)
2. Preprocess the inputs: Rectify real clip limit from the normalized value if necessary, glorify the image before splitting it into regions
3. Computing each contextual region (tile) thus made gray Level mappings: Thread a single image region, made a Histogram for this region using the specified number of bins, clip the histogram using clip limit, created mapping for this region
4. Computed gray level mapping in order to configure final CLAHE image: Theread cluster of four neighboring mapping Functions, compute image region partly overlapping each Of the mapping tiles, thread a single pixel, apply four Mappings to that pixel, and interpolate between the events to obtain the output pixel; render over the entire image

2.2 Feature Extraction

Feature extraction's method used was Gray Level Co-occurance Matrix (GLCM). GLCM method is a form of removing second order statistical texture features.

A GLCM is a matrix where the number of rows and columns is close to the number of gray levels, G, in the image. The maximum number for the level was 8 and minimum level was 1. The matrix point P (i, j | Δx, Δy) is the relational frequency with which two pixels, isolated by a pixel distance (Δx, Δy), pass within a given neighborhood, one with intensity 'i' and the other with intensity 'j'. The matrix point P (i, j | d, ɵ) included the second order statistical probability values for automatise between gray levels 'i' and 'j' at a specific replacing distance d and at a particular angle (ɵ). Using a large number of intensity levels G showed storing a lot of working data, i.e. a G × G matrix for each collection of (Δx, Δy) or (d, ɵ). Due to their large magnitude, the GLCM's are very reactive to the size of the texture instances on which they are asses. Thus, the number of gray levels is often attenuated. GLCM matrix formulation can be interpreted with the samples illustrated in fig 3. for four varied gray levels. Here one pixel offset is utilized (a reference pixel and its immediate

neighbor). If the window was large enough, using a larger offset is mathematical. The top left cell will be full with the number of times the combination 0,0 occurs, i.e. how many time within the image area a pixel with grey level 0 (neighbor pixel) falls to the right of another pixel with gray level 0(reference pixel).
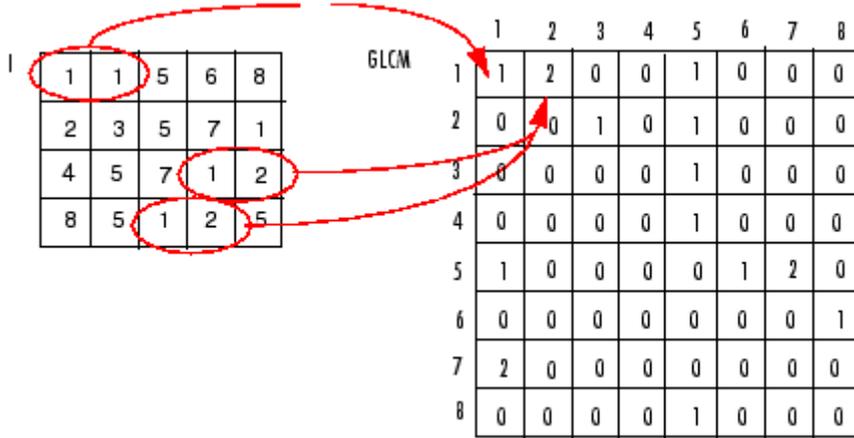

Fig. 3. GLCM Calculation

Four features extraction had been used to get GLCM values which were Contrast, Correlation, Energy, and Homogeneity.

2.2.1 Contrast

The definition of Contrast was the intensity contrast between a pixel and its neighbor over the whole picture. 'Sum of Square Variance' is the other name of Contrast. It defers the calculation of the intensity contrast linking pixel and its neighbor over the whole image. At constant image contrast value is 0.In contrast determine, weight accumulate exponentially(0,1,4,9) as continue from the diagonal.

$$\sum_{i,j} |i-j|^2 p(i,j)$$

(1)

Since (i-j) increases contrast proceeded to increase exponentially. When i and j are equal i.e. i-j=0.no contrast is there. When i and j are differ by 1,small contrast is there is 1. When i and j diverged by 2, the contrast is expanding and weight is 4

2.2.2 Correlation

It moved the computation of the correlation of a pixel and its neighbor over the whole picture means it figures out the linear state of gray levels on those of neighboring pixels. On behalf an absolutely positively or negative related image, the correlation value is 1 and -1.On behalf of unvarying image its value is NaN..Range=[-1,1] and the formula is

$$\sum_{i,j} \frac{(i-\mu i)(j-\mu j)p(i,j)}{\sigma_i \sigma_j}$$

(2)

2.2.3 Energy

Since energy is utilized for doing work, Thus system. It made use for the texture that computed orders in an image. It gave the sum of square elements in GLCM. It was fully different from entropy. When the window is good orderly, energy value is high .The square root of ASM(Angular Second Moment) texture character is used as Energy. Its range is[0 1].Since constant image its value is 1.The equation of energy is

$$\sum_{i,j} p(i,j)^2$$

(3)

2.2.4 Homogeneity

In brief term it is going by the name of HOM. It moved the value that computed the tightness of sampling of the elements in the GLCM to the GLCM diagonal. For diagonal GLCM its value is 1 and its range is [0,1].Opposite of contrast weight is homogeneity weight values, with weight decreases exponentially loose from the diagonal. The weight employed in contrast is (i-j)^2 and in homogeneity ,it is 1/1+(i-j)^2.The equation is

$$\sum_{i,j} \frac{p(i,j)}{1+|i-j|}$$

(4)

2.3 Backpropagation Neural Networks

The method of artificial neural network used in this study is backpropagation, in general, this algorithm is referred to as back propagation because it can be described as follows, when the network is given an input pattern as a training pattern then the pattern goes to units in the hidden layer to be forwarded to the unit and the output layer unit, then the output layer units give a response called network output, if the network output is not the same as the expected output, the output will spread backward, in the hidden layer forwarded to the unit at the input layer, therefore the training mechanism is called backpropagation or back propagation. Lavenberg-Marquardt had been used to classified Batik Banten. This method outperforms simple gradient descent and other conjugate gradient methods in a wide variety of problems. The algorithm provide a solution is called Nonlinear Least Squares Minimization. Figure 4 showed the input layer consisted 160 input, with 20 hidden layers, and 10 output layers, and 10 output layers.

Backpropagation is a supervised learning algorithm and is usually used by multiple layers of perceptrons to change the weights connected to the neurons in the hidden layer. The backpropagation algorithm uses error output to change its weight values in the backward direction. To get this error, the forward propagation stage (forward propagation) must be done first. When forward propagation, neurons are activated by using the sigmoid activation function. The working principle of the backpropagation algorithm is as follows, during the forward feed process, the neurons in the input layer will send information to each neuron in the hidden layer. This hidden layer neuron will do the computation process of the information received and produce output derived from its activation function. Then the output from the hidden layer will be sent again to the next layer, which in this case is the output layer. In this output, the same process takes place as in the hidden layer. However, the output from the activation output of this layer will be used as a response signal from Athe NN. This ANN output response signal will be compared with the desired target and calculated the error magnitude (the difference between the output produced with

the target to be achieved). This error was sent back to the previous layers and then used to correct and change the ANN weights based on the learning error given to the ANN.

2.3.1 Training

The training phase is a step in how an artificial neural network is practiced, that is by making a weighing change (the connection between layers that form a network through each unit), while solving new problems will be done if the training process is completed, the phase is the mapping phase or testing process. The training process uses 20 images of each batik, the image is an image.
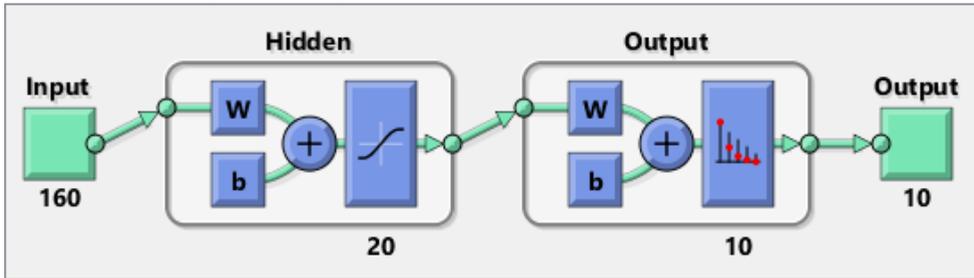

Fig. 4. Neural Networks Backpropagation

The training process by the backpropagation network required several parameters that needed to be set for learning with the trainlm function. The parameters that need to be set were:
1. Maximum epoch is the number of maximum or maximum epochs iterations that may be carried out during the training process.
2. Purpose performance is the target value of the performance function, or used to determine the Mean Square Error (MSE) value limit so that the iteration is stopped. The iteration will be stopped when the MSE value is smaller than the specified limit.
3. Learning Rate is the rate of learning. The greater the learning rate will have implications for the greater the learning step, if the learning rate is set too large, the algorithm will not become stable, and vice versa, if the learning rate is set too small, the algorithm will converge for a very long time.
4. The minimum gradient is the root of the least number of squares (input weight, layer weight, bias weight) that is the smallest allowed.

2.3.2 Testing

The testing process is carried out to determine the success rate of a network in recognizing a new input pattern, where the testing process compares the value of all outputs with the target. The testing processes used 100 images from 10 Batik patterns classification to check the system.Testing is done by entering images that have feature extraction using GLCM. The results of feature extraction are then entered into the neural network and classification is done to validate the results.

3 Result and Discussion

The original image of batik obtained from batik craftsmen in the city of Serang. then the cropping process was done manually to take normal Banten batik as many as 300 images. The obtained image data was divided into 200 images of Banten batik as training data

(training) consisting of normal and abnormal Banten batik while 100 images of Banten batik as testing data.

The results of the training are carried out like architectural designs that have been explained previously, the training process will stop when one of the requirements for termination of the network formation process is fulfilled. The conditions that determine the cessation of the formation of this backpropagation network include achieving maximum epoch or maximum iteration, the set MSE value has been reached, the minimum gradient value has been reached, and also the validation check that has been fulfilled. The training process that has been designed.
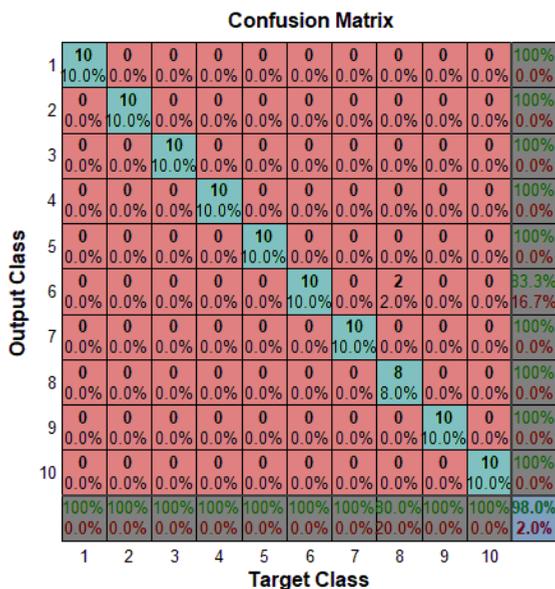

Fig. 5. Confusion Matrix

Figure 5. showed the result of the system. The system presented the 98% in accuracy value. Classification Class were :

1. Datu Laya,
2. Kapurban
3. Kawangsan
4. Mandalikan
5. Pamaranggen
6. Pancaniti
7. Pasepen
8. Pasulaman
9. Pejantren
10. Surosowan

Figure 5. showed the results of testing with the image of Banten batik, it can be seen that not all Banten batik can be recognized by the backpropagation network that has been built. Figure 5. showed that there are still targets that are not or close to the target that has been determined, therefore the network cannot recognize all Banten batik. System validation values for testing are:

$$accuracy = \frac{10+10+10+10+10+10+10+8+10+10}{10+10+10+10+10+10+10+10+10+10} x100\% = 98\%$$

$$error = \frac{2}{10+10+10+10+10+10+10+10+10+10} x100\% = 2\%$$

Batik Banten classification presented the accuracy value 98% and the error rate 2%. The system successfully classified Batik Banten class from 10 different pattern. The system had better value than [5] that had accuracy [3]97.67%,[6] with accuracy value 72%,[7] which had accuracy value 99,5%, but only 2 class,[8]images classification accuracy was 86%, [9] had accuracy value 99,96% but with 5 class had been used ,[4], [10]–[14] had accuracy < 92%,[15] had accuracy 80%, [16] had accuracy value 94,57%.

4 Conclusion

Batik Banten had special texture and shape of specific ornaments that could be used to classification the image of batik. The unique ornament of Batik Banten had size and shape that vary widely in 10 class. Training using Neural network with GLCM to feature extraction had high accuracy such as 98%, and low error as 2%. This method showed promising algorithm to classify Batik Banten, and could be used in face recognation.

References

1. J. Hidayat, "The Art and Sustainable Aspects of Natural Dyeing in KANAWIDA Hand Drawn Batik (Green Batik)," *IPTEK J. Proceeding Ser.*, (2014)

2. R. Azhar, D. Tuwohingide, D. Kamudi, Sarimuddin, and N. Suciati, "Batik Image Classification Using SIFT Feature Extraction, Bag of Features and Support Vector Machine," in *Procedia Computer Science*, (2015)

3. A. H. Rangkuti, "Content based batik image classification using wavelet transform and fuzzy neural network," *J. Comput. Sci.*, (2014).

4. A. H. Rangkuti, Z. E. Rasjid, and D. J. Santoso, "Batik Image Classification Using Treeval and Treefit as Decision Tree Function in Optimizing Content Based Batik Image Retrieval," in *Procedia Computer Science*, (2015)

5. R. Azhar, D. Tuwohingide, D. Kamudi, Sarimuddin, and N. Suciati, "Batik Image Classification Using SIFT Feature Extraction, Bag of Features and Support Vector Machine," *Procedia Comput. Sci.*, (2015)

6. A. Kitipong, W. Rueangsirasak, and R. Chaisricharoen, "Classification system for traditional textile: Case study of the batik," in *13th International Symposium on Communications and Information Technologies: Communication and Information Technology for New Life Style Beyond the Cloud, ISCIT 2013*, (2013)

7. F. Budiman, A. Suhendra, D. Agushinta, and A. Tarigan, "Wavelet decomposition levels analysis for Indonesia traditional batik classification," *J. Theor. Appl. Inf. Technol.* (2016)

8. T. Handhayani, "Batik Lasem images classification using voting feature intervals 5 and statistical features selection approach," *Proceeding - 2016 Int. Semin. Intell. Technol. Its Appl. ISITIA 2016 Recent Trends Intell. Comput. Technol. Sustain. Energy*, pp. 13–16, (2017)

9. I. Nurhaida, R. Manurung, and A. M. Arymurthy, "Performance comparison analysis features extraction methods for Batik recognition," in *2012 International Conference on Advanced Computer Science and Information Systems, ICACSIS 2012 - Proceedings*, (2012)

10. N. Suciati, A. Kridanto, M. F. Naufal, M. Machmud, and A. Y. Wicaksono, "Fast discrete curvelet transform and HSV color features for batik image clansificotlon," in *Proceedings of 2015 International Conference on Information and Communication Technology and Systems, ICTS 2015*, (2016)

11. H. Wijayanto, "Klasifikasi Batik Menggunakan Metode K-Nearest Neighbour Berdasarkan Gray Level Co-Occurrence Matrices (Glcm)," *Klasifikasi Batik Menggunakan Metod. K-Nearest Neighb. Berdasarkan Gray Lev. Co-Occurrence Matrices (Glcm)*, (2015)

12. C. S. K. Aditya, M. Hani'Ah, R. R. Bintana, and N. Suciati, "Batik classification using neural network with gray level co-occurence matrix and statistical color feature extraction," in *Proceedings of 2015 International Conference on Information and Communication Technology and Systems, ICTS 2015*, (2016)

13. F. U. Karimah and A. Harjoko, "Classification of batik kain besurek image using speed up robust features (SURF) and gray level co-occurrence matrix (GLCM)," in *Communications in Computer and Information Science*, (2017)

14. F. Budiman, A. Suhendra, D. Agushinta, and A. Tarigan, "Determination of SVM-RBF kernel space parameter to optimize accuracy value of Indonesian Batik images classification," *J. Comput. Sci.*, (2017)

15. A. A. Kasim, R. Wardoyo, and A. Harjoko, "The selection feature for batik motif classification with information gain value," in *Communications in Computer and Information Science*, (2017)

16. T. Handhayani, "Batik Lasem images classification using voting feature intervals 5 and statistical features selection approach," in *Proceeding - 2016 International Seminar on Intelligent Technology and Its Application, ISITIA 2016: Recent Trends in Intelligent Computational Technologies for Sustainable Energy*, (2017)