The 2018 International Conference on Information Technology,
Engineering, Science, and its Applications

Available in the SSRN eLibrary: http://ssrn.com/link/ITES-2018.html

Batik Image Retrieval System Using Self Organizing Map

Natalia[*], *Janson* Hendryli, and *Dyah Erny* Herwindiati

Faculty of Information Technology, Tarumanagara University, Indonesia

Abstract. This paper describes an image retrieval system for batik motifs from Yogyakarta, Solo, and Lasem region using self organizing map method. The system receives user-uploaded batik image and returns 10 images of batik from Yogyakarta, Solo, and Lasem from the database with the closest distance that is computed using self organizing map. The experiments cluster the batik images into 13 clusters based on the 16 features extracted using gray level co-occurrence matrices (GLCM) to represent the texture of batik. In addition, color moment features are also considered as feature extractor to represent the color values of batik images. The average precision and recall for the experiments with both GLCM and color moment features are 0.5 and 0.082, respectively. Meanwhile, using only GLCM gives the result of 0.377 average precision and 0.060 average recall.

1 Introduction

Batik is an Indonesian cultural heritage with high artistic value. It has already been recognized by United Nations Educational, Scientific, and Culture Organization (UNESCO) as one of the world cultural heritage. The word batik itself is a technique in dyeing cloths to make specific patterns using "*canting*". The dyeing is conducted using a stranglehold material that was applied into a fabric which holds back the dyes from getting in. So the batik cloths are fabrics that have a lot of motives and patterns that are made with the "*canting*" [2].

The development of batik motifs has already been going for centuries resulting in a lot of batik patterns in Indonesia. The variation of batik patterns shows each region's identity and culture, such as batik Yogyakarta, batik Aceh, batik Lampung, batik Cirebon, batik Solo, batik Madura, batik Sidoarjo, batik Bali, batik Papua, and so much more.

A lot of unique batik motifs and variations becomes the attraction of batik in the heart of Indonesian people. Nowadays, batik is still popular with Indonesians shown by the creation of clothing, bags, or shoes made with batik clothes. Yet the lack of knowledge of the huge variety of batik motifs can be an obstacle for understanding and identifying the value and uniqueness of batik.

With the advancement of technology, we can easily search for a particular batik name from the web and learn a lot of its history, characteristics, and other valuable information. In spite of that, frequently, the batik name is unknown and the user want to get the information based on the batik image.

[*] Corresponding author: natalia.lia96@yahoo.co.id

Fig. 1. Example of the batik motifs from Yogyakarta, Solo, and Lasem

This paper describes an image retrieval system for searching a batik pattern based on user-uploaded image. The batik image retrieval system measures the similarity level between the batik image and the images of Yogyakarta, Solo, and Lasem batik motifs in the database. The system then returns similar images of batik and its information from the database. The application is expected to be used not only as a learning material in educating the public about batik patterns, but also to help those who want to discover particular batik patterns from images. For the experiments, we limit the batik patterns only for batik from Yogyakarta, Solo, and Lasem region.

The method for the batik image retrieval are the self organizing map (SOM) algorithm. Meanwhile, we employ two kinds of feature extractors: the gray level co-occurrence matrix (GLCM) and color moments. The SOM algorithm is grouping similar images of batik into clusters, and our system selects 10 images with the shortest distance in the cluster.

2 Theoretical Framework

2.1 Gray level co-occurrence matrix

The gray level co-occurrence matrix (GLCM) is a method which analyzes image's pixels to discover the grayness level [5]. In image processing literatures, the GLCM is one of the widely used method as feature extractor. It can also be seen as a method to retrieve the second order statistical value by computing the probability of neighborhood relationships between two pixels. The GLCM works by constructing a co-occurrence matrix, meaning common events or the number of events of one pixel value level be neighbors with other pixel value level in particular distance and angle orientation. The distance is declared in pixels and the orientation is declared in degree, which is in 0˚, 45˚, 90˚, and 135˚. This research employs the angular second moment, contrast, correlation, and homogenity features from the GLCM to describe the texture of batik images.

The angular second moment is the sum of squares of the image with similarity in colors and pixels. This can be computed as in Eq. (1) where f_1 is the angular second moment feature and $p(i,j)$ is the matrix element.

$$f_1 = \sum_{i=0}^{G-1} \sum_{j=0}^{G-1} \{p(i,j)\}^2 \tag{1}$$

The contrast is used to calculate the difference of the graylevel in the image. Visually, a large value of contrast is a measure of the variation of the greyness in an image. The contrast can be computed as in Eq. (2) where f_2 is the contrast feature.

$$f_2 = \sum_{n=0}^{G-1} n^2 \{\sum_{i=1}^{G} \sum_{j=1}^{G} p(i,j)\}, |i-j| = n \tag{2}$$

The third GLCM feature used in this research is correlation, which denotes the linear dependence of an image so as to inform the existence of a linear structure in the image. Eq. (3) computes the correlation, where f_3 is the correlation, μ_x and μ_y are the mean of p_x and p_y respectively, σ_x and σ_y are the standard deviation of p_x and p_y respectively.

$$f_3 = \sum_{i=0}^{G-1} \sum_{j=0}^{G-1} \frac{\{i \cdot j\} \cdot p(i,j) - (\mu_x \cdot \mu_y)}{\sigma_x \sigma_y} \tag{3}$$

The last feature is homogenity or inverse difference moment (IDM) which is used to measure the homogenity of the graylevel variation in the image. A homogenous image will have a large IDM value. The computation of homogenity feature can be seen in Eq. (4).

$$f_4 = \sum_{i=0}^{G-1} \sum_{j=0}^{G-1} \frac{1}{1+(i-j)^2} p(i,j) \tag{4}$$

2.2 Color moments

For the color moment features, the mean, standard deviation, and skewness of the hue, saturation, and value of the images are computed [1]. Those three features are representative enough to describe the distribution of colors in a digital image. Eq. (5), (6), and (7) shows the formula to compute the mean, standard deviation, and skewness, respectively.

$$E_i = \frac{1}{N} \sum_{j=1}^{N} p(i,j) \tag{5}$$

$$\sigma_i = \sqrt{\left(\frac{1}{N} \sum_{j=1}^{N} (p(i,j) - E_i)^2\right)} \tag{6}$$

$$S_i = \sqrt[3]{\left(\frac{1}{N} \sum_{j=1}^{N} (p(i,j) - E_i)^3\right)} \tag{7}$$

2.3 Self organizing maps

The self organizing map (SOM) algorithm or Kohonen's network was first introduced by Teuvo Kohonen in 1982. The SOM is one of the popular clustering methods that uses neural network approach and Euclidean distance [3]. It can recognize the distribution and topology of the input vector.

The SOM algorithm is as follows [4]:

1. Initialize the random weights w_{ij}.
2. If the termination conditions are not fulfilled, perform steps 3-8.
3. For each vector input x do step 4-6.
4. For each index j, calculate the distance between the input vectors with their corresponding weights using Eq. 8 where $D(j)$ is the Euclidean distance, w_{ij} is the weight node, and x_i denotes the input vector.

$$D(j) = \sum_i (w_{ij} - x_i)^2 \tag{8}$$

5. Find the winning unit index j with the smallest distance $D(j)$.
6. Calculate new values of w_{ij} where j is the winning unit index using Eq. 9. Note that α is the learning rate.

$$w_{ij}^{\text{new}} = w_{ij}^{\text{old}} + \alpha(x_i - w_{ij}^{\text{old}}) \tag{9}$$

7. Reduce the neighboring distance.
8. Return to step 2.

The training process stops when the termination conditions are fulfilled. The condition can be either when the learning rate approaches 0 or the nearest neighbor distance is 0. Other termination conditions such as convergence rate or maximum number of iteration can also be employed.

3 Experiments

3.1 Data

The images of batik are collected from various internet sources. The dataset consists of 1010 images from 13 classes of batik motifs, such as Kawung, Parang, Sidoluhur, Sidomukti, Slobog, Truntum, Aseman, Bledak, Gunung Ringgit, Krecak, Latohan, Naga, and Sekar Jagad. Fig. 2 shows the number of the batik images for each classes or motifs in the dataset. Furthermore, the data are splitted into two sets: training and testing. We use 90% of the data for training the SOM model and use the remaining 10% for model evaluation.

3.2 Evaluation Metrics

The experiments will be evaluated using precision and recall at the K-rank. The evaluations aim to measure the quality of a particular rank that is known to be relevant or irrelevant to the document. Eq. (10) and (11) calculate the precision and recall, respectively, where RET is the retrieved document, REL is the relevant document, and $|X|$ denotes the cardinality or the number of documents in set X.

$$P@K = \frac{|RET \cap REL|}{RET} \tag{10}$$

$$R@K = \frac{|RET \cap REL|}{REL} \tag{11}$$

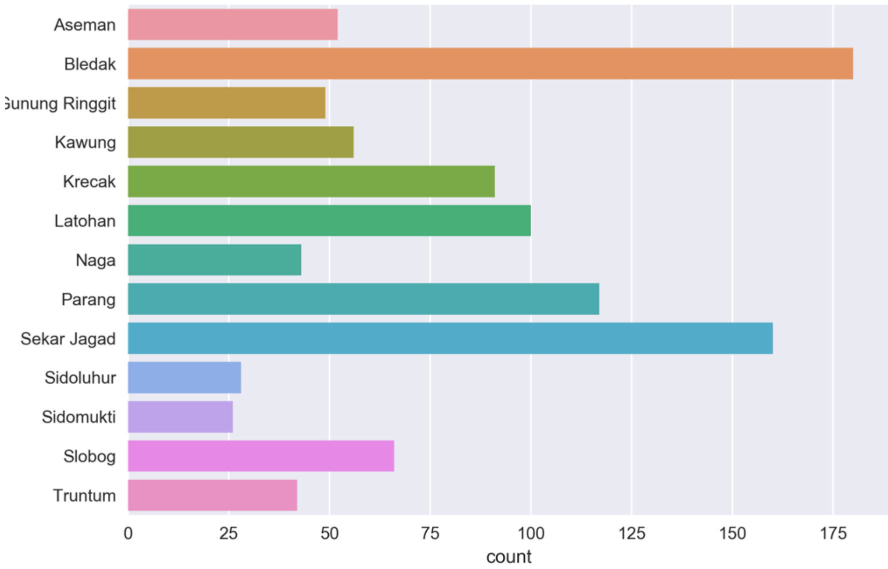

Fig. 2. The number of batik images per motifs in the dataset

3.3 Results

To understand the performance of the clustering method, we first evaluate the SOM algorithm and analyze the clusters formed in the training process. For this cluster evaluation, the GLCM and color moment features are employed. The feature values are used as the input vector to calculate the distances among the batik images.

Fig. 3-7 show the clusters formed after the SOM training. It can be seen that the members of the clusters demonstrate a lot of similarities in the textures and colors. As can be seen from the figures, the clusters are mostly formed around the images with similar color characteristics, e.g. reddish color in the first cluster. We can also argue that the texture characteristics are not really being considered by the SOM algorithm. And it should also be noted that there is one cluster with only single image as its members. The fourth cluster only consists of one Naga batik image.

Furthermore, the performance of the image retrieval system will be discussed below. There are two experiment scenarios, which are the retrieval system using GLCM and color moments as image features and using GLCM only. Table 1 shows the precision and recall from both experiments. From Table 1, we can see that the average precision and recall, though not necessarily high enough, of the experiments with both GLCM and color moments as features are better than when only GLCM is used.

From the GLCM + color moments experiment, the Aseman and Sidoluhur motif have the lowest precision, while the Latohan has the lowest recall. In contrast, for the GLCM-only experiment, Gunung Ringgit and Naga have the lowest precision. Surprisingly, the Gunung Ringgit also has the lowest recall in the last experiment. Meanwhile, Bledak and Sidomukti has the best precision and recall, respectively, in both experiments.

Fig. 3. Example of batik images in the first cluster. In total, the cluster has 15 images as its members. It can be seen that this cluster focuses on batik images with reddish color.

Fig. 4. Example of batik images in the second cluster. This cluster, which consists of 6 images, focuses on yellowish or brownish images with diagonal textures.

Table 1. The precision and recall from the evaluation of the image retrieval system using GLCM and color moments

No	Batik Motifs	GLCM + Color Moments		GLCM	
		Precision	**Recall**	**Precision**	**Recall**
1	Aseman	0.2	0.043	0.3	0.064
2	Bledak	1.0	0.062	0.8	0.049
3	Gunung Ringgit	0.5	0.114	0.1	0.023
4	Kawung	0.4	0.08	0.3	0.06
5	Krecak	0.7	0.085	0.5	0.061
6	Latohan	0.3	0.033	0.4	0.044
7	Naga	0.3	0.077	0.1	0.026
8	Parang	0.9	0.086	0.6	0.057
9	Sekar Jagad	0.8	0.056	0.6	0.042
10	Sidoluhur	0.2	0.08	0.2	0.08
11	Sidomukti	0.4	0.174	0.3	0.13
12	Slobog	0.4	0.068	0.4	0.068
13	Truntum	0.4	0.105	0.3	0.079
Average		**0.5**	**0.082**	**0.377**	**0.060**

Fig. 5. The third cluster consists of reddish images with flower motifs.

Fig. 8 shows the example of the batik image retrieval results for Bledak motif. As we can see, the system returns 10 images of Bledak with 100% precision.

Fig. 6. The member of the fourth cluster consists of only one image, which is the Naga motif from Lasem.

Fig. 7. The fifth cluster consists of dominantly white batik motifs.

4 Conclusion

This paper has demonstrated the self organizing map algorithm (SOM) for batik image retrieval using gray level co-occurrence matrix (GLCM) and color moments as feature extractors. The batik motifs that are being used in this research are from Yogyakarta, Solo, and Lasem region in Indonesia. From the experiments, we found that the performance of the model is still inconsistent. Particularly for Bledak motif, the retrieval system returns the best result (best precision), while for other motifs, the results are not good enough. From analyzing the cluster members formed after SOM training, it can be found that the color characteristics are too dominant in the clustering process. Although when we study the performance of GLCM and color moments as feature extractors, compared to when we only use only GLCM, the retrieval results are still better for the former.

For our future works, it will be beneficial to collect more comprehensive batik image dataset, so that we can train a better model. It is also interesting to study and analyze deeper about the effects of color features in batik image retrieval system.

(a)

(b)

Fig. 8. The results of the image retrieval. (a) the user-input image of Bledak motif; (b) top-6 result from the image retrieval system returns 6 batik images with Bledak motif.

References

[1] D. E. Herwindiati, S. M. Isa, R. Sagara, "The new notion distance of content based image retrieval," *Journal of Indonesian Mathematical Society*, vol. 16, no. 1, 2010.

[2] F. Hindayani, Mengenal dan membuat batik. Jakarta: Buana Cipta Pustaka, 2009.

[3] J. Ruiz-del-Solar, P. Navarrete, "Eigenspace-based face recognition: A comparative study of different approaches," *IEEE Transactions on Systems, Man, and Cybernetics – Part C: Applications and Reviews*, vol. 35, no. 3, 2005.

[4] L. Fausett, Fundamentals of neural networks architectures, algorithms, and applications. Englewood Cliffs: Prentice-Hall International Editions, 1994.

[5] X. Zhang, F. Zhang, "Image features extraction of tobacco leaves," in *Proc. of Congress on Image and Signal Processing*, Sanya, Hainan, China, May 27-30, 2008.

The 2018 International Conference on Information Technology,
Engineering, Science, and its Applications

Available in the SSRN eLibrary: http://ssrn.com/link/ITES-2018.html

Classification of Batik Motifs Using Convolutional Neural Networks

Jonathan Tristanto, *Janson* Hendryli, and *Dyah Erny* Herwindiati

Faculty of Information Technology, Tarumanagara University, Jakarta, Indonesia

Abstract. Batik is Indonesia's traditional cultural heritage which is a cloth-making technique using *canting* and *malam*. Each region of Indonesia usually produces their own batik motifs. This paper proposes the convolutional neural networks to recognize the batik motifs from images. The dataset consists of batik motifs from Lasem, Solo, and Yogyakarta region. In total, there are 13 classes of batik motif in the dataset. Experimental results show that the model performance is not sufficient enough to accurately detect most of the batik motifs.

1 Preliminary

Batik, which is designated as one of the UNESCO's Masterpieces of the Oral and Intangible Heritage of Humanity on 2009 [1], is a motif on a cloth made with a special technique using *canting*. The batik cloth is very popular in Indonesia with many regions produce their own motifs. For example, the Aseman motif in Fig. 1 is a batik motif from the Lasem region in Central Java, Indonesia.

The large number of batik motifs can be daunting for some people, especially the younger generations who aren't exposed too much to the traditional batik making industry. Usually, the images retrieval system such as Google Images can be used as a tool for searching batik motifs. The user will upload certain batik images and the web will return numerous similar images. However, the results are usually unsatisfactory since the system only search for similar looking images, without considering the batik motifs itself.

Recent advancements of deep learning shows the increase popularity of convolutional neural networks (CNN) for various tasks such as image [2] and video [3] classification, speech recognition [4], and others. The way the CNN works is quite similar to the multi-layered perceptron. The difference is the addition of the convolutional layers for extracting important features of an image before feeding it up to the fully-connected layers. This paper proposed the CNN model to automatically classify the batik motif from images.

2 Convolutional Neural Networks

Convolutional neural networks or CNN (Fig. 2) is a type of neural networks which utilizes the convolutional filters in the layers to extract local features from the input [5]. The CNN model has been used since the late 1980s [6] in the application of handwritten zip code recognition. In 2012, [7] used CNN to outperform other models in the ImageNet challenge

Corresponding author: jotris99@gmail.com

and revived the interests in the model. Since then, numerous researchers have improved the model by introducing various techniques, such as transfer learning [8], dropout regularization [9], and deep architectures like VGG19 [10] and ResNet [11].

Fig. 1 Batik with Aseman motif which comes from Lasem region in Indonesia

The convolutional layer is the first step in the CNN architectures which learns local patterns such as edges, textures, and so on from an input image. The convolution is a mathematical operation involving matrix multiplications of the input and a kernel matrix resulting in a feature map. The output of the convolutional layers can further be fed into the pooling layers. The role of the pooling layers are to downsample the feature maps. The usual choice is the max pooling layer which outputs the max value of each channel.

After a few convolutional and pooling layers, the networks are usually followed by the fully connected layer.

Each layers of convolution and fully connected are usually employing the rectified linear unit (ReLU) activation function which allows faster and effective training of the deep networks [12]. The last fully connected layer is the output layer which uses the softmax activation function for multiclass classification problem or sigmoid function for the binary class classification.

For the training process, a number of optimization algorithms can be chosen, such as the stochastic gradient descent, Adam, RMSProp, and so on. Furthermore, to avoid overfitting the training set, the dropout regularization can also be employed.

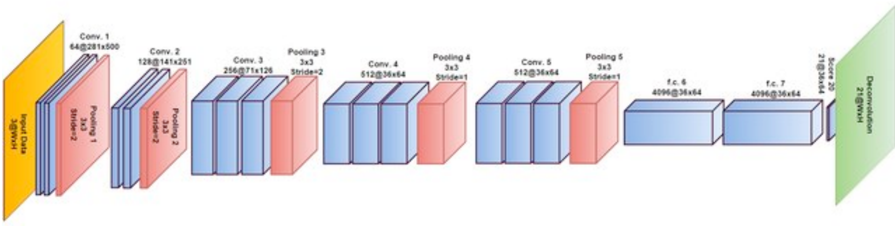

Fig. 2 Illustration of a convolutional neural networks model [13]

3 Experimental Settings

The batik image dataset for the experiments are mostly collected from the Google Images. There were 967 images of batik from the Lasem, Yogyakarta, and Solo region which consist of 13 motifs. Those 13 motifs are Aseman, Bledak, Gunung Ringgit, Krecak, Kawung, Latohan, Naga, Parang, Sido Mukti, Sido Luhur, Sekar Jagad, Slobog, and Tuntum. The Lasem batik images are collected by [14]. The dataset are divided with a proportion of 70% and 30% for training and test data respectively. Moreover, the training also utilizes data augmentation, such as rotation, shear mapping, zoom, horizontal flip, and translation. All the images are resized to 160 x 160 pixels.

The CNN model is implemented in Python and Keras library. The convolutional neural networks architecture consists of four convolutional layers followed by three fully connected layers. Each convolutional layers but the last consist of 32 filters of size 3 x 3. The number of filters in the last convolutional layer is 64. The ReLU activation function is used in the convolutional layers. After each convolutional layers, we also employ maxpooling layers of size 2 x 2. The first fully connected layer consists of 64 units with ReLU activation function. We also utilize the dropout regularization in this layer. The next fully connected layer has 32 units (ReLU) and the last layer's number of units is exactly the same as the number of our output classes. Also, the softmax activation function is used in the last layer. And finally we train the model using RMSProp optimizers to minimize the categorical cross entropy loss in 2000 epochs with early stopping and batch size of 32.

4 Results and Discussions

With the aforementioned experimental settings, we obtain the result shown in Fig. 3. The loss from training and validation/test sets are converging around the 1750th epoch. Meanwhile, the accuracy of the model is 56%. From the experiments, we find that Kawung are the motifs with the best precision of 91.67%. In contrast, the Naga and Sido Mukti can't be detected correctly most of the time. Furthermore, the precision of other motifs are also not very high, which are around 40-60%.

We argue that the trained model cannot identify the motifs accurately due to the lack of our data training, particulary in the number of images and their variations. For example, Fig. 4 shows the batik images of Aseman which cannot be correctly classified in the testing phase. Though the batik patterns are clearly Aseman (by human eye inspection), the model may not be able to differentiate it since the color of the batiks are different than the images in the training set. Thus, our dataset can be said to be lack of variations in color.

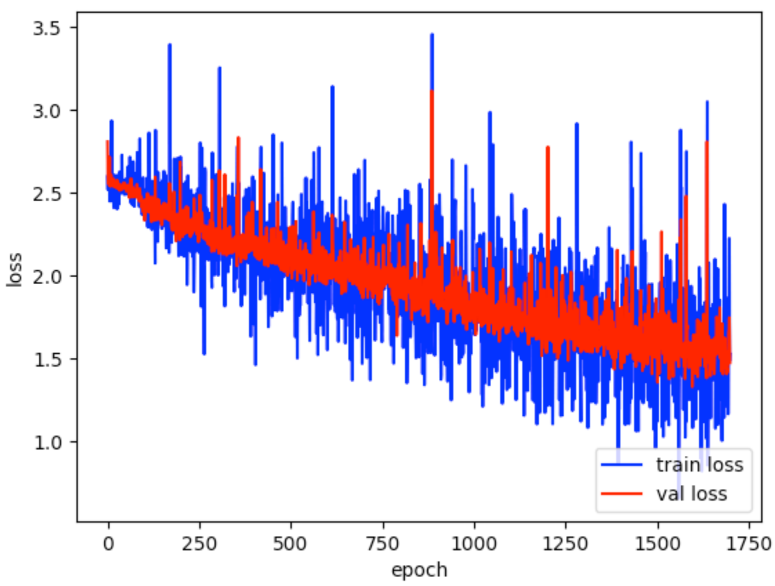

Fig. 3 The visualization of the training and validation loss on each epochs

Fig. 4 The examples of Aseman batik motif from the test set

5 Conclusions

In this paper, we explore the use of convolutional neural network (CNN) to classify batik motif images from Lasem, Yogyakarta, and Solo region in Indonesia. The experiment shows that our CNN model cannot reliably classify the batik motif with training accuracy of 56%. We found that the model performs well when classifying the Kawung motif but fail to accurately classify the Naga and Sido Mukti motif. For the future works, it is imperative to collect more comprehensive dataset of batik images with many variations to get better result in automatic classification of batik motif.

We gratefully acknowledge the support of NVIDIA Corporation with the donation of the Titan Xp GPU used for this research.

References

1. R. Smend, D. Harper, Batik Traditional Textiles of Indonesia, 9 (2011)
2. A. Krizhevsky, I. Sutskever, G. E. Hinton. Imagenet classification with deep convolutional neural networks, in Advances in neural information processing systems, pp. 1097-1105 (2012)
3. A. Karpathy, G. Toderici, S. Shetty, T. Leung, R. Sukthankar, and Li Fei-Fei, Large-scale video classification with convolutional neural networks, in Proceedings of the IEEE conference on Computer Vision and Pattern Recognition, pp. 1725-1732 (2014)
4. O. Abdel-Hamid, A. Mohamed, H. Jiang, and G. Penn, Applying convolutional neural networks concepts to hybrid NN-HMM model for speech recognition, in Acoustics, Speech and Signal Processing (ICASSP), pp. 4277-4280 (2012)
5. I. Goodfellow, Y. Bengio, A. Courville, Deep learning, MIT Press (2016)
6. Y. LeCun, B. Boser, J. S. Denker, D. Henderson, R. E. Howard, W. Hubbard, L. D. Jackel, Backpropagation applied to handwritten zip code recognition, J. Neural Computation, vol. 1 no. 4, pp. 541-551 (1989)
7. A. Krizhevsky, I. Sutskever, G. E. Hinton, Imagenet classification with deep convolutional neural networks, in Advances in neural information processing systems, pp. 1097-1105 (2012)
8. S. J. Pan, Q. Yang, A survey on transfer learning, IEEE Transactions on knowledge and data engineering, vol. 22 no. 10, pp. 1345-1359 (2010)

9. G. E. Dahl, T. N. Sainath, G. E. Hinton, Improving deep neural networks for LVCSR using rectified linear units and dropout, in Acoustics, Speech and Signal Processing (ICASSP), pp. 8609-8613, (2013)
10. K. Simonyan, A. Zisserman, Very Deep Convolutional Networks for Large-Scale Image Recognition, arXiv preprint, arXiv:1409.1556 (2014)
11. K. He, X. Zhang, S. Ren, J. Sun, Deep residual learning for image recognition, in Proceedings of the IEEE conference on computer vision and pattern recognition, pp. 770-778 (2016)
12. X. Glorot, A. Bordes, Y. Bengio, Deep sparse rectifier neural networks, AISTATS (2011)
13. F. Saleh, M. S. Aliakbarian, M. Salzmann, L. Petersson, S. Gould, J. M. Alvarez, Built-in foreground/background prior for weakly-supervised semantic segmentation, in European Conference on Computer Vision, pp. 413-432 (2016)
14. T. Handhayani, J. Hendryli, L. Hiryanto, Comparison of shallow and deep learning models for classification of Lasem batik patterns, in Proceedings of International Conference on Informatics and Computational Sciences, pp. 11-16 (2017)