

# Non-linear Multiclass SVM Classification Optimization using Large Datasets of Geometric Motif Image

Fikri Budiman, Edi Sugiarto  
Department of Computer Science  
Dian Nuswantoro University  
Semarang, Indonesia

**Abstract**—Support Vector Machine (SVM) with Radial Basis Functions (RBF) kernel is one of the methods frequently applied to nonlinear multiclass image classification. To overcome some constraints in the form of a large number of image datasets divided into nonlinear multiclass, there three stages of SVM-RBF classification process carried out i.e. 1) Determining the algorithms of feature extraction and feature value dimensions used, 2) Determining the appropriate kernel and parameter values, and 3) Using correct multiclass method for the training and testing processes. The OaO, OaA, and DAGSVM multi-class methods were tested on a large dataset of batik motif images whose geometric motifs with a variety of patterns and colors in each class and containing similar patterns in the motifs between the classes. DAGSVM has the advantage in classification accuracy value, i.e. 91%, but it takes longer during the training and testing processes.

**Keywords**—Geometric motif; image classification; multiclass; non-linear; large dataset

## I. INTRODUCTION

Studies focusing on classification of image recognition have a high level of complexity if it has a large dataset with many different groups or multiple classes. The use of large datasets in question is data with large amounts (Lot of Data) with structured data. Moreover, the boundaries among classes cannot be separated by linear hyperplane (non-linear) due to their high level of image feature similarity. There are several supervised Machine Learning algorithms which can be used for image recognition classification, such as hierarchical-based decision trees, K-Nearest Neighborhood algorithms, partition-based K-means and minimum-distance, and networks based Artificial Neural Network (ANN) like perceptron algorithm, Backpropagation Neural Network (BNN), and Support Vector Machine (SVM). These classification algorithms are categorized as Shallow Learning type since they still require some application of feature extraction algorithms to produce feature image dataset. Feature extraction is a fundamental part in classification as feature dataset obtained from proper feature extraction can maximize the accuracy value of classification results [1-2].

This article elaborates a comparative study and evaluation of multi-class SVM methods which have been carried out. The advantages of classification using SVM are such as the ability to significantly gain good classification accuracy values for

image features with high data dimensions [3, 4]. However, it is still necessary to test the classification with several different kinds of feature dimensions to get the right dimensions for use [1-2]. The advantages of SVM than those of Artificial Neural Network (ANN) in classification are that SVM does not involve all vector dataset for training image features in forming hyperplane and margins as class separators, and only vector datasets for contributing image features (support vector) are used for hyperplane formation and margins. In addition, SVM determines the hyperplane by maximizing the distance among classes (maximum margin), so that it has high generalization for the testing dataset. Thus, it is better than ANN in which it searches for a hyperplane by principally minimizing its gradient and depending on the number of parameters used [5-8]. SVM classification works using the principle of Structural Risk Minimization (SRM) to enable it to produce good generalizations with hyperplane fields which can minimize the average error in managing training dataset [7].

Several previous studies showed that classification with SVM can significantly increase the accuracy value as in the classification to recognize the texture of honey pollen images in which the SVM accuracy results were better than Multi-layer Perceptron classification method, Minimum Distance Classifier, and K-Nearest Neighbor [8-9]. SVM classification was tested by two public databases of DNA micro array to classify tumors and non-tumors resulting in a classification accuracy value in which SVM was better than ANN [10]. SVM by using default kernel parameter was applied for non-linear multi-class classification with a dataset of five types of batik textures. The results showed that the classification accuracy value using SVM was better than by using Minimum Distance Method and Backpropagation Neural Network [6, 11].

SVM was initially introduced by Boser, Guyon, and Vapnik in 1992 and only used as binary classifier [12]. SVM has the convenience to maximize non-linear classification patterns since SVM can overcome over-fitting with soft margins by replacing each dot product of testing feature using a non-linear kernel function matrix [12]. The strategies carried out in developing non-linear multi-class SVM classification are still considered to have some weaknesses for large number of dataset samples. To solve such weaknesses of non-linear multi-class SVM classification on large-scale image recognition, there are three experimental stages which can be carried out to optimize the use of SVM method in non-linear multiclass

classification, so that it may maximize classification accuracy value. The first stage requires an experiment focusing on the use of qualified feature extraction by determining feature extraction algorithm and its proper parameters to gain qualified dimensions and feature values [1-2,13], the second is conducting an experiment to determine the best kernel and its parameter values which are fit with the conditions of the dataset used [4,5,7,14], The classification of each new dataset with SVM depends on the kernel function used and the parameters used. And the third requires an experiment on using the right method to handle the training and testing process in multi-class. The method commonly used for multi-class is by using combination approaches of several SVM binary or two classes [15-16].

In obtaining maximum classification accuracy value using SVM, it requires the best results from the three experiments by using the new image dataset. The author has conducted the first and the second experiments for 4 non-linear classes using one-against-all method for geometric pattern image dataset with a high level of feature similarity [1-2,5]. The dataset used Batik images with traditional motifs from Indonesia. The traditional Batik images have very diverse geometric decorations and high similarity of motifs as well as possess multi-scale patterns and multi-color resolution [17-18]. Such batik with traditional Indonesian motifs along with the method of creating this batik has been recognized by UNESCO on 2 October 2009 as a "Representative List of The Intangible Cultural Heritage of Humanity".

The third experiment was as important as the first and the second ones in increasing the maximum accuracy value. Consequently, this article used a study using multi-class SVM methods by combining several binary SVM approaches, such as: one-against-all, one-against-one, and directed acyclic graph SVM from previous studies. Moreover, to find out the ability of these methods in maximizing the value of accuracy and classification time for multi-class non-linear with new datasets, an experimental evaluation of these methods was required using a large dataset of images with geometric motifs in the form of traditional Indonesian Batik images consisting of 4 classes and 7 classes. The results of this experiment can be used as the basis to develop new research methods which are more efficient in training and testing time and have maximum accuracy values in handling non-linear multi-class in SVM classification with large datasets.

## II. RELATED WORK

The fundamental first step in images classification is the process of applying feature extraction method to generate feature values. Features are unique characteristics of texture features of each image so that they can be recognized through digital image processing [19]. The results of feature extraction are converted in the form of statistical feature features which is then used for classification in the recognition phase. Maximum accuracy value of image classification depends on whether the value of generated image features is good or not. Thus, the experiment on applying the right feature extraction method can be an interesting stage in the initial image recognition research [1-2,19-20]. Feature extraction methods are commonly divided into three, i.e. structural, statistical, and spectral [20].

Discrete Wavelet Transformation (DWT) which is a spectral method is a feature extraction method often used and applied to 2-dimensional images having multi-resolution spaces with varying scale transformations and can produce features based on differences in image intensity in several sub-band spaces [20-22]. DWT can perform the process of changing signal or wave data which is a combination of time and frequency into a series of wavelet coefficients that are easier to analyze, and can be used for feature extraction in images with geometric decorative motifs that have aperiodic or interrupted signals disconnected and noisy. The non-linear multi-class classification experiment in this study used the value of the image feature based on the author's previous research [1], i.e. using statistical features in the form of energy and standard deviation values obtained from wavelet coefficients by using DWT level 3 method with wavelet type of daubechies 2. For the introduction of batik motifs in the research by Rangkuti, Harjoko, and Putro [22], the use of wavelet daubechies 2 was also better than the haar, coiflets, and biorhogonal types.

Following the experiment on the proper feature extraction method to use, it is also significantly important to determine non-linear kernel function and its parameters which best apply to SVM classification of non-linear multi-class [5, 23]. SVM classification was initially only used to distinguish two classes which could be separated by a hyperplane in the form of a linear line [24]. In the real case, classification in general is more about separation among nonlinear classes, or it will be very difficult to be linearly separated. Furthermore, SVM can be developed into a nonlinear classifier by using a kernel trick. In maximizing non-linear SVM classification pattern with the kernel trick, over fitting is minimized using soft margin concept on the hyperplane by replacing each dot product of feature with a non-linear kernel function matrix to determine the support vector [24-26].

One of the functions of kernel is to solve non-linear problem used to determine the support vector by mapping from the initial training feature data to the new training feature data which has a feature space with higher dimensions without defining the function of input space to the new feature space. [27-28]. Optimization of non-linear SVM classification highly depends on the use of kernel function and its proper parameters. There are several functions of non-linear kernel to replace the mapping to new dimensional feature space, including Polynomial, Gaussian/Radial Basis Function (RBF), sigmoid, Multi Quadratic Inversion, and Additive [7].

The Gaussian RBF kernel is highly recommended to gain maximum non-linear classification results for a new dataset [4, 7, 14, 28] since it has the same performance as the linear kernel on the parameters cost (C) and gamma ( $\gamma$ ) / sigma ( $\sigma$ ) with a certain value in the optimization of classification. The best parameters combination for C and  $\gamma$  values will be obtained by a hyperplane with the right soft margin so that maximum accuracy of the classification results can be achieved [7, 28]. There is no range requirement for the estimated value of C and  $\gamma$  as test values for the Gaussian RBF kernel parameters. Determination of the RBF-SVM kernel parameters is to obtain a hyperplane and margin with minimal classification errors. The use of correct parameter values of C and  $\gamma$  will also result

in a low measure of the deviations (variance) and a low measure of error contribution (bias). High variance with too high C value and too low  $\gamma$  value can cause over fitting and high bias. On the contrary, if C value is too small and  $\gamma$  is too big, under fitting will occur. In each test, the combination of parameters C and  $\gamma$  can be classified several times with several different training and testing datasets. This is to ensure that there is no excessive over fitting on the tests with different testing data [4-5]. Gaussian RBF kernel can be elaborated as follows:

$$\text{Gaussian: } K(x_i, x_j) = \exp(-\|x_i, x_j\|^2 / (2\sigma^2)) \quad (1)$$

$$\gamma = 1/2\sigma^2 \quad (2)$$

$$K(x_i, x_j) = \exp(-\gamma \|x_i, x_j\|^2) \quad (3)$$

$K(x_i, x_j)$  is the value of each kernel matrix element, with  $x_i$  and  $x_j$  as data point pairs. The value of gamma ( $\gamma$ ) as a kernel parameter is used to determine the maximum result in optimization of classification, so it is necessary to estimate a parameter in the form of a constant parameter value for the kernel parameter ( $\gamma$ ).

In some previous studies, RBF kernel and its parameter determination were recommended to use to obtain maximum non-linear classification results for the new dataset [5,14,21,29-30]. Determining the value of this parameter is necessary since the function of Gaussian RBF kernel is to substitute dot product mapping from old dimensional features to the new ones depending on the conditions of the image dataset used. Evaluation to optimize the parameters for Gaussian RBF-SVM kernel needs to be done to get the classification with the smallest errors in the image dataset with geometric textures which have many variations and a high level of texture similarity. The parameter values applied to non-linear multi-class experiments using a large number of geometric motif image dataset in this article are from the results of the author's previous research [5]. This study used 4 classes of image datasets with geometric motifs and the maximum classification accuracy value was obtained with a combination of low bias and low variance RBF kernel parameters at the value of  $C = 2^7$  and  $\gamma = 2^{-15}$  [5].

In addition to having its own complexity in determining the right hyperplane, in multi-class non-linear SMV classification applied to large datasets with image features which have similarities among classes, it is also necessary to evaluate which multiclass method to be applied. The use of non-linear SVM multi-class training and testing methods can be used based on the use of two classes/binary classification, i.e. by combining several binary SVMs. This method is much easier and more practical to apply than by combining all datasets consisting of several classes into the form of optimization problems [16,31]. Some other commonly used methods are: one-against-all (OaA)/winner takes all (WTA), one-against-one (OaO) / Max Wins Voting (MWA), and Directed Acyclic Graph (DAG) [15]. Of all these methods, the superiority of each method in terms of accuracy value, testing and training time duration has not been clearly found out yet, it still depends on the size of the dataset and the number of classes [15,31]. In the methods above, several binary classifications are used in training and testing phases. Of all the other methods, OaA uses

the least binary classification as if there are n classes; then n binary classification models are needed. In OaA, there is an unbalanced distribution of the training dataset classes in determining the hyperplane for each binary classification. For example, there are 5 classes, class +1 is class a dataset; then class -1 is a combination of the remaining class datasets, i.e. class dataset b, c, d, and e. Such class imbalance might be problematic when applied to large datasets with many classes since large combined class datasets will cause a slow training phase. Besides, incorrect classification predictions may occur because the training process focuses more on large combined class datasets. The one-against-all method in previous research [5], uses an approach for multi-class SVM for 4 classes by combining kernel parameter functions that can adjust multiple hyperplane and margins in several classifications of two classes, the test stages carried out are as shown in Fig. 1.

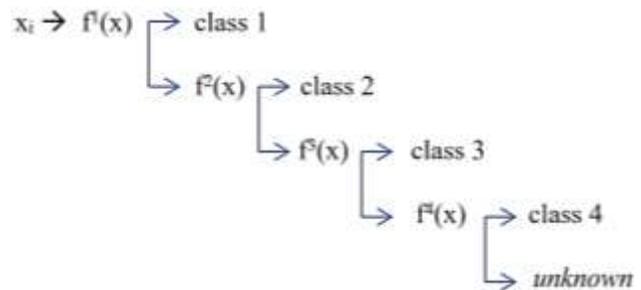


Fig. 1. Stages of Testing One-against-all 4 Classes.

In this study [5] a simplification of the process was carried out by using a combination of several binary SVMs with the one-against-all method (Fig. 1), which was compared using the one-against-one method. The one-against-one method for 4 classes is done with 6 SVM testing models as shown in Fig. 2. Each test classification model is carried out on data from two classes, namely for data from class i and class j.

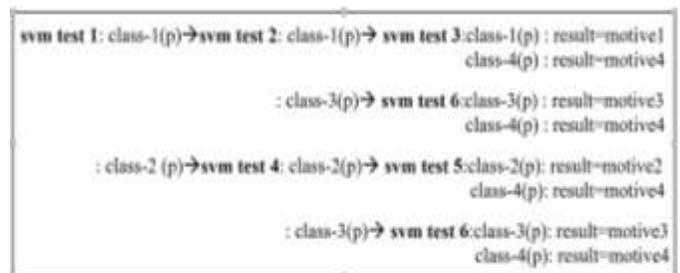


Fig. 2. Stages of Testing One-against-one 4 Classes, p=Positive.

In OaO and DAG there is no class merger, so in the training phase a binary classification model is needed as much as  $(n(n-1))/2$ , the value of 'n' is the number of classification class, thus the training phase will be slower as the number of classification classes increases [16]. The difference between OaO and DAG is in their test method, the binary classification model in OaO is first tested and the results are used for the voting method; while in DAG, binary directed acyclic graph method is used. The testing method in DAG is more efficient than OaO, because DAG produces a faster test time complexity for large datasets compared to using OaO, this is because OaO performs two stages in the testing phase [15, 32]. Although the

testing phase is slower, previous research experiments using OAO approach showed that the results of the accuracy value are better than the other multi-class approaches [33-34]. With the advantages and disadvantages of the OaA, OaO, and DAG methods, it is necessary to test the performance of these multiclass methods, the test is carried out in the case of large image datasets, the images used are images with geometric motifs that have high similarity between classes.

### III. PROPOSED METHOD

This section describes the proposed method to produce maximum accuracy values. In this case, the non-linear multiclass classification optimization method with SVM that uses Large Datasets of Geometric Motif Image can be explained in detail as follows:

1) The feature creation phase is used to optimize the classification accuracy results with feature extraction using Discrete Wavelet Transformation (DWT). Optimization of the use of DWT is done by comparing the use of decomposition level 1 to 5 and wavelet types db 1 to db 5. The features of each sub-band are energy and standard deviations values obtained from the wavelet coefficient values contained in each sub-band.

2) The optimization phase of the SVM classification results with the Grid-Search and Cross Validation processes, to minimize over fitting and obtain a combination of RBF kernel parameters values ( $C$  and  $\gamma$ ) in the space parameters that produces the maximum classification accuracy value.

3) The results of the first phase are in the form of optimization of the use of wavelet levels and types, and the results of the second phase are the optimal values of the  $C$  and  $\gamma$  parameters, used in experiments using DAG SVM, OAA SVM, and OAO SMV, in the classification of 4 classes and 7 classes, with variations in the number of datasets. 200 to 3000 batik image data.

The proposed classification optimization method is as shown in Fig. 3.

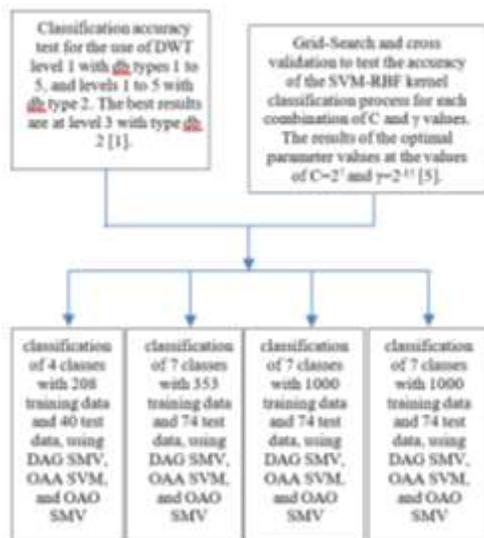


Fig. 3. Proposed Method.

### IV. RESULTS AND DISCUSSION

The implementations are conducted using a computing platform based on Intel Core i5 1.6 GHz CPU, 8 GB RAM, and Microsoft Windows 10 Professional 64-bit. The development environment is MATLAB programming language.

The datasets used were images of traditional Indonesian Batik which have several classes of decorative patterns with very diverse geometric motifs and high degree of motif patterns similarity among classes. Batik is an original cultural heritage from Indonesia in the form of the beauty of art works on cloth media and contains philosophical meaning of life in each depicted decorative motifs. Batik, which is a work of art on cloth, which is to decorate the surface of the textile by holding the dye. The process of making Batik artwork is by applying color retention popularly known as wax-resist dyeing process. This process of making Indonesian traditional Batik is recognized by UNESCO as a “Representative List of The Intangible Cultural Heritage of Humanity”. Each class of Batik motif patterns has a deep meaning of philosophy of life which reflects the noble elements of life. The dataset used in the experiment consisted of 4 classes and 7 classes of Indonesian traditional batik patterns. 4 classes consisted of motifs: Ceplok, Kawung, Nitik, and Parang. In addition, 7 classes consisted of motifs: Ceplok, Kawung, Nitik, Parang, Sidumukti, Lereng, and Slobog (Fig. 4). Total dataset for 4 classes consisted of 208 training data and 40 testing data. Tmeanwhile, 7 classes dataset were tested for 353 training data with 74 testing data, 1000 training data with 74 testing data, and 3000 training data with 74 testing data.



Fig. 4. Seven Classes of Batik Motif.

The level of prediction accuracy of the test dataset image recognition is measured by using the confusion matrix measurement technique which is divided into positive data prediction classes (true positive / TP; false positive / FP) and negative data (true negative / TN; false negative / FN) [22].

$$\text{Accuracy} = (\text{TP} + \text{TN}) / (\text{TP} + \text{TN} + \text{FP} + \text{FN}) \quad (4)$$

The correctness of the image class prediction from the classification results compared to the actual image class results in the TP, FP, TN, and FN values as shown in table 1.

TABLE I. CONFUSION MATRIX

Classification		Predicted Class	
		Class = yes	Class = no
Actual Class	Class = yes	TP	FN
	Class = no	FP	TN

In this non-linear multi-class classification experiment with this large dataset, the use of SVM classification was maximized through three experimental stages as described in the introduction section. At the first stage, based on the results of the author's previous experiment [1], it used a feature with a statistical characteristic of the energy value and the standard deviation of the Discrete Wavelet coefficients. The calculation of the energy value of each pattern component in the sub-band is with the parameter  $E_s$  (energy in a certain sub-band calculated by CA / CH / CV / CD),  $C_s$  (the matrix coefficient on the sub-band pattern calculated by CA / CH / CV / CD),  $c$  (vector coefficient values for all sub-bands), are as follows:

$$E = 100 * \sum_{x,y} (C_{s_{x,y}} \hat{^2}) / \sum_{n=1:m} (c_n \hat{^2}) \quad (5)$$

The standard deviation formula (Std) for each pattern component in the sub-band with the parameter  $C_s$  (the matrix coefficient on each component of the sub-band pattern calculated by CA / CH / CV / CD),  $r$  is the average value of  $C_s$ , and  $n$  is the amount of data  $C_s$ :

$$\text{Std} = (\sum_{x,y} (C_{s_{x,y}} - r)^2 / (n-1))^{1/2} \quad (6)$$

The level 3 wavelet decomposition is carried out on the sub-band approximation at level 2, as shown in Fig. 5.

Transform level 3 with wavelet type daubechies 2 which has been used in the previous experiment [1], produces 20 values in the feature vector for each image of the training and test data. Each sub-band has 2 (two) feature values, energy and standard deviation, so that the wavelet coefficient vector and the feature vector generated at each level 3 are:

$$[E_{CA3} \text{ Std}_{CA3} E_{CH3} \text{ Std}_{CH3} E_{CV3} \text{ Std}_{CV3} E_{CD3} \text{ Std}_{CD3} \\ E_{CH2} \text{ Std}_{CH2} E_{CV2} \text{ Std}_{CV2} E_{CD2} \text{ Std}_{CD2} \\ E_{CH1} \text{ Std}_{CH1} E_{CV1} \text{ Std}_{CV1} E_{CD1} \text{ Std}_{CD1}]$$

The results of two author's previously conducted experiments [1-2] showed that in addition to the application of the right feature extraction algorithm, the number of values in the feature vector used must also be precisely right, because any values would greatly affect the maximum accuracy value of the classification.

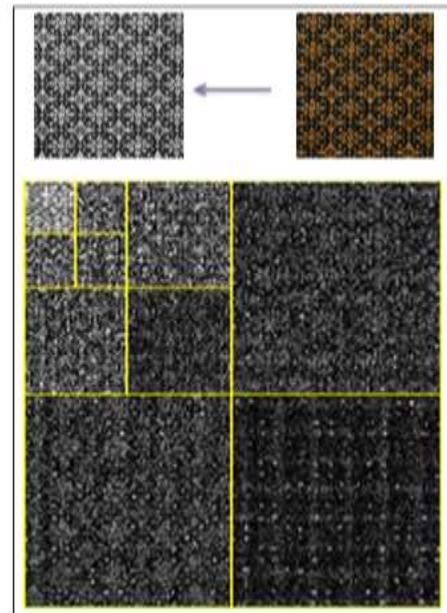


Fig. 5. Level 3 Wavelet Decomposition.

The second stage of the kernel and the parameter values used are the results of the previous author's experiment [5], the determination of the RBF kernel parameter values on the scale tested to obtain the maximum accuracy value in the non-linear multi-class SVM classification method has been successfully identified with the space parameter. In looking for the maximum value of the RBF kernel with low bias and low variance in the parameter range specified in the study, obtained at the optimal parameter value  $C=2^7$  and  $\gamma=2^{-15}$ . In gaining the maximum classification accuracy value, it is necessary to evaluate the use of the correct parameter value, since the optimal parameter value highly depends on the image datasets used. Testing the use of a combination of cost function/  $C$  and  $\gamma$  parameters on RBF kernel is better conducted in several times classification with different training and testing datasets. In each test the combination of parameters  $C$  and  $\gamma$  is carried out the k-fold Cross Validation (CV) method with 10 classifications using 10 training datasets and different tests (10-fold). In this study [5] the experiment using a small  $k$  value less than 10 resulted in more and more image class recognition errors, this is because the training dataset is getting less and less good at representing the hyperplane and margins for each class. Determining the best value for this parameter combination is important, to ensure that there are no over-fitting and under-fitting that could cause sharp differences in accuracy results when using different test datasets. From the test results in the previous study [5] it can be recommended to classify images that have geometric decorative motifs with a non-linear multi-class SVM-RBF kernel, using the Grid-search range  $C = \{2^{6.5}, 2^{6.75}, 2^7, 2^{7.25}, 2^{7.5}, 2^{7.75}, 2^8\}$  and  $\gamma = \{2^{-14.5}, 2^{-14.75}, 2^{-15}, 2^{-15.25}, 2^{-15.5}, 2^{-15.75}, 2^{-16}\}$ , these ranges are to get the optimal parameter combination value.

The third stage was conducted by comparing the methods of training and testing in OaA, OaO, and DAG multiclass using a large dataset of batik images. The use of these three methods was carried out in 4 stages to measure the accuracy value of the

classification and the time required for training and testing process (in seconds). The results of the first stage with 4 classes consisting of 208 training data and 40 test data (table 2) using DAG showed the highest accuracy value. However, it required longer training and testing time than those by using OaA and OaO. The results of the first stage and the second stage (table 3) are obviously the same which used 7 classes with 353 training data and 74 testing data. Furthermore, DAG method resulted in more superior accuracy value with longer training and testing time.

DAG has the best accuracy value than OaA and OaO in the third stage of the experiment (table 4) and in the fourth stage (table 5). However, this is not significant regarding the time required for the training process. There is a significant time leap when the training dataset was increased to 3000 images (fig. 6). By applying the three stages of the process using SVM for this nonlinear multiclass classification, DAG method is considered to be good for achieving the accuracy value of the classification. However, but there are still some constraint as the longer training time needed along with the increasing size of the dataset.

TABLE II. RESULTS OF 4 CLASSES EXPERIMENTS

Method	4 classes					
	Training data	Testing data	Correct classification	Accuracy (%)	Training Time	Testing time
DAG SVM	208	40	37	93	0,834	0,021
OAA SVM	208	40	35	88	0,274	0,016
OAO SVM	208	40	26	65	0,104	0,016

TABLE III. RESULTS OF 7 CLASSES EXPERIMENTS, 353 DATASETS

Method	7 classes					
	Training data	Testing data	Correct classification	Accuracy (%)	Training Time	Testing time
DAG SVM	353	74	62	84	1,863	0,038
OAA SVM	353	74	58	78	0,769	0,024
OAO SVM	353	74	39	53	0,165	0,017

TABLE IV. RESULTS OF 7 CLASSES EXPERIMENTS, 1000 DATASETS

Method	7 classes					
	Training data	Testing data	Correct classification	Accuracy (%)	Training Time	Testing time
DAG SVM	1000	74	67	91	3,755	0,058
OAA SVM	1000	74	63	85	1,439	0,029
OAO SVM	1000	74	42	57	0,477	0,020

TABLE V. RESULTS OF 7 CLASSES EXPERIMENTS, 3000 DATASET

Method	7 classes					
	Training data	Testing data	Correct classification	Accuracy (%)	Training Time	Testing time
DAG SVM	3000	74	67	91	13,726	0,276
OAA SVM	3000	74	63	85	4,550	0,033
OAO SVM	3000	74	42	57	1,544	0,020

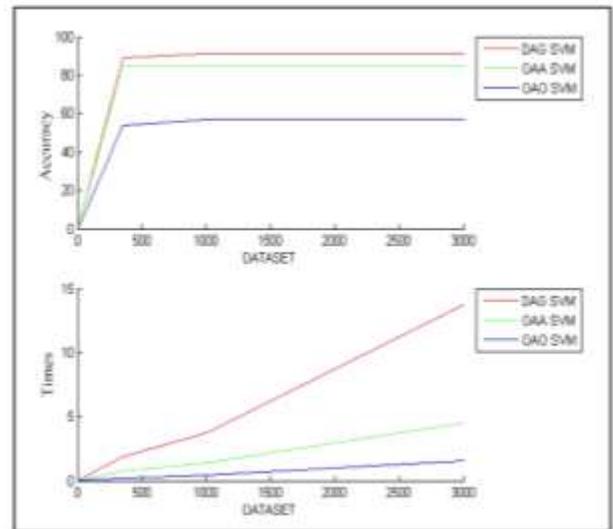


Fig. 6. Comparison of 7 Class Performances.

## V. CONCLUSION AND FUTURE WORK

These three stages of non-linear multiclass SVM-RBF classification are still able to produce good accuracy values, for the use of large datasets of traditional Indonesian batik images with highly varying geometric motifs. This good accuracy value depends on the image features, the kernel and its parameters, as well as on the methods used during the training and testing process.

The training and testing methods applied to 1000 and 3000 batik images as the dataset with 7 motif classes showed that DAG can produce a consistent accuracy value of 91%. Nevertheless, with the increase in the number of datasets from 1000 to 3000, it showed that the required training time was increasingly bigger. Thus the SVM classification with large dataset images has problems with the speed of the training process with the increasing number of datasets, thus further research is still needed in the form of developing a combination model of SVM-RBF with Deep Learning to minimize classification time on big data.

### REFERENCES

- [1] Budiman, F., Suhendra, A., Agushinta, D., and Tarigan, A., "Wavelet Decomposition Levels Analysis For Indonesia Traditional Batik Classification", Journal of Theoretical & Applied Information Technology, 92(2):389-394, 2016.
- [2] Budiman, F., Sugiarto, E., "Image Feature Extraction of Numbers and Letters Using Matrix Segmentation", Scientific Visualization, 12(1): 120-131, 2020.

- [3] Azhar, Ryfial, Tuwohingide, Desmin, Kamudi, Dasrit, Sarimuiddin, dan Suciati, Nanink, "Batik Image Classification Using SIFT Feature Extraction, Bag of Features and Support Vector Machine", *Procedia Computer Science*, Elsevier, 72(1) 24-30, 2015.
- [4] Renukadevi, N.T. dan Thangaraj, P., "Performance Evaluation of SVM-RBF Kernel for Medical Image Classification", *Global Journal of Computer Science and Technology Graphics & Vision*, Global Journal Inc, USA, 13(4):15-19, 2013.
- [5] Budiman, F., "SVM-RBF Parameters Testing Optimization Using Cross Validation and Grid Search to Improve Multiclass Classification", *Scientific Visualization*, 11(1), 80-90, 2019.
- [6] Yuan, Qing-Ni, Lu, Jian, Huang, Haisong, dan Pan, Weiji, 2014, "Research of Batik Image Classification Based On Support Vector Machine", *Computer Modelling & New Technologies* 18(12B) 193-196, 2014.
- [7] Hofmann, M., *Support Vector Machines- Kernels and the Kernel Trick*. 1st Edn., Bamberg University, 2006.
- [8] Ahmad, A., Hashim, U. K. M., Mohd, O., Abdullah, M. M., Sakidin, H., Rasib, A. W., & Sufahani, S. F., "Comparative analysis of support vector machine, maximum likelihood and neural network classification on multispectral remote sensing data", *International Journal of Advanced Computer Science and Applications (IJACSA)*, 9(9), 529-537, 2018.
- [9] Fernandez-Delgado, M., Carrion, P., Cernadas, E., Galvez, J. F., & S-Otero, P., "Improved classification of pollen texture images using SVM and MLP". In *3rd IASTED International Conference on Visualization, Imaging and Image Processing (VIIP2003) (Vol. 2)*, 2003.
- [10] Ahmad M. Sarhan, "Wavelet-based Feature Extraction For DNA Microarray Classification", *Springer Science+Business Media B.V. Artif Intell Rev*, 39:237-249. 2013.
- [11] Herulambang, W., Hamidah, M. N., & Setyatama, F., Comparison of SVM And BPNN Methods in The Classification of Batik Patterns Based on Color Histograms And Invariant Moments, In *2020 International Conference on Smart Technology and Applications (ICoSTA)*, pp. 1-4, IEEE, 2020.
- [12] Boser, B.E., Guyon, I.M., and Vapnik, V.N., "A training algorithm for optimal margin classifiers", In D. Haussler, editor, *5th Annual ACM Workshop on COLT*, pp. 144–152, Pittsburgh, PA. ACM Press, 1992.
- [13] Makri, E., Rotaru, D., Smart, N. P., and Vercauteren, F., "EPIC: Efficient private image classification (or: Learning from the masters)". In *Cryptographers' Track at the RSA Conference* (pp. 473-492). Springer, Cham, 2019.
- [14] Rosales-Perez, Alejandro, Escalante, Hugo Jair, Gonzales, Jesus A., dan Reyes-Garcia, Carlos A., "Bias and Variance Optimization for SVMs Model Selection", *Proceedings of the Twenty-Sixth International Florida Artificial Intelligence Research Society Conference*. pp. 136-141, 2013.
- [15] Hsu, C. W., & Lin, C. J., "A comparison of methods for multiclass support vector machines". *IEEE transactions on Neural Networks*, 13(2), 415-425, 2002.
- [16] Groenen, P. J. F., & van den Burg, G. J. J., Multiclass Support Vector Machines With Gensvm. In *Computer Data Analysis and Modeling: Stochastics and Data Science*, pp. 43-50, 2019.
- [17] Riyanto, Rahayu, Slamet, and Pamungkas, Wisnu, *Handbook of Indonesian Batik*, The Institute For Research and Development of Handicraft and Batik Industrie, Yogyakarta, 1997.
- [18] Tirta, I., *Batik Sebuah Lakon*, 1st Edn., Gaya Favorit Press, Jakarta, ISBN: 9789795154280, pp: 278. 2009.
- [19] Nixon, Mark S., dan Aguado, Alberto S., *Feature Extraction and Image Processing*, Academic Press is an imprint of Elsevier, 2008.
- [20] Selvarajah S., Kodituwakku, "Analysis and Comparison of Texture Features Content Based Image Retrieval", *International Journal of Latest Trends in Computing*, 2(1):108-113, 2011.
- [21] Virmani, Jitendra, Kumar, Vinod, Karla, Naveen, dan Khandelwal, Niranjana, "SVM-Based Characterization of Liver Ultrasound Images Using Wavelet Packet Texture Descriptors", *Journal of Digital Imaging* (2013) 26:530–543, Springer, 2013.
- [22] Rangkuti, Abdul Haris, "Content Based Batik Image Classification Using Wavelet Transform And Fuzzy Neural Network", *Journal of Computer Science* 10 (4): 604-613, Science Publications, 2014.
- [23] Gorunescu, Florin, *Data Mining – Concepts, Models, and Techniques*, Berlin : Springer-Verlag Berlin Heidelberg, 2011.
- [24] Wu, Xindong, Kumar, Vipin (ed.), *The top ten algorithms in data mining*. CRC press, Taylor & Francis Group, London, New York, 2009.
- [25] Boser, B., Guyan, I., dan Vapnik, V., 1992, "A Training Algorithm for Optimal Margin Classifiers", In *Proceedings of the Fifth Annual Workshop on Computational Learning Theory*, pp.144-152, New York:ACM Press.
- [26] Ali, Muhammad Asim, and Zain, Ahmed Siddiqui, "Automatic music genres classification using machine learning", *International Journal of Advanced Computer Science and Applications (IJACSA)*, 8(8) : 337-344, 2017.
- [27] Schölkopf, Bernhard, and Alexander J. Smola. "Learning with kernels: support vector machines, Regularization." *Optimization, and Beyond*. MIT press 1.2, 2002.
- [28] Hsu, Chih-Wei., Chang, Chih-Chung., dan Lin Chih-Jen., *A Practice Guide to Support Vector Classification*, Department of Computer Science, National Taiwan University, 2010.
- [29] Syarif, Iwan, Prugel-Bernnett, Adam, dan Wills, Gary, "SVM Parameter Optimization Using Grid Search and Genetic Algorithm to Improve Classification Performance", *Telkomnika Journal*, 14(4):1502-1509, 2016.
- [30] Gaspar, Paulo, Carbonell, Jaime, Oliveira, dan Jose Luis, "On the Parameter Optimization of Support Vector Machines for Binary Classification", *Journal of Integrative Bioinformatics*, 9(3):201, 2012
- [31] Duan, K. B., & Keerthi, S. S. , "Which is the best multiclass SVM method? An empirical study. In *International workshop on multiple classifier systems*", Springer, Berlin, Heidelberg, pp. 278-285, 2005.
- [32] Agarwal, N., Balasubramanian, V. N., & Jawahar, C. V., "Improving multiclass classification by deep networks using DAGSVM and Triplet Loss", *Pattern Recognition Letters*, 112, 184-190, 2018.
- [33] Zheng, H.B., Liao, R.J., Grzybowski, S dan Yang, L.J., "Fault diagnosis of power transformers using multi-class least square support vector machines classifier with particle swarm optimisation". *IET Elect. Power Appl.* Vol 5, Iss 9, pp. 691-696, 2011.
- [34] Liu, Yang, Jian-Wu Bi, and Zhi-Ping Fan, "A method for multi-class sentiment classification based on an improved one-vs-one (OVO) strategy and the support vector machine (SVM) algorithm", *Information Sciences* 394, 38-52, 2017.