

Determination of SVM-RBF Kernel Space Parameter to Optimize Accuracy Value of Indonesian *Batik* Images Classification

¹Fikri Budiman, ²Adang Suhendra, ²Dewi Agushinta and ²Avinanta Tarigan

¹Department of Computer Science, University of Dian Nuswantoro, Semarang, Indonesia

²Department of Computer Science, University of Gunadarma, Depok, Indonesia

Corresponding Author:
Fikri Budiman
Department of Computer
Science, University of Dian
Nuswantoro, Semarang,
Indonesia
Email: fikri.budiman@dsn.dinus.ac.id

Abstract: Image retrieval using Support Vector Machine (SVM) classification very depends on kernel function and parameter. Kernel function used by dot product substitution from old dimension feature to new dimension depends on image dataset condition. In this research, parameter of Gaussian /Radial Basis Function (RBF) kernel function is optimized using multi class non-linear SVM method and implemented to training and test datasets of traditional Indonesian *batik* images. The *batik* images dataset is limited to four geometric motifs textures, which are *ceplok/ceplok*, *kawung*, *nitik* and *parang/lerang*. Discrete Wavelet Transform level 3 daubechies 2 is used to result feature dataset of traditional *batik* images dataset of four classes geometric motifs textures. The *batik* images are used for training and test dataset in SVM-RBF kernel parameter optimization to maximize accuracy value in non-linear multi-class classification. Cross Validation and Grid-search methods are used to analyze and evaluate SVM-RBF kernel parameter optimization. Confusion matrix measurement method is used to result accuracy value in every evaluation conducted in every combination of cost function/C and gamma/ γ as SVM-RBF kernel parameter. Maximum accuracy parameter value is $C = 2^7$ and $\gamma = 2^{-15}$ achieved by 10 times evaluation wit different test dataset for each evaluation. Maximum accuracy value is 0.77 to 0.86.

Keywords: Kernel Parameter, Gaussian/RBF, Support Vector Machine/SVM, Discrete Wavelet Transform/DWT, Indonesian Traditional *Batik*, Geometric Motifs

Introduction

This study aims to preserve traditional *batik* motifs texture culture came from Solo and Yogyakarta palaces, Indonesia. One of studies which can be developed is digital image retrieval based on texture of geometric motifs in traditional *batik* (Tirta, 2009), as seen in Fig. 1.

The dataset uses 80 images of traditional *batik* divided into four geometric motifs, which are *ceplok*, *ceplok*, *kawung*, *nitik* and *parang/lereng*. The feature dataset which can properly characterize images is needed to achieve high accuracy in classification accuracy optimization result. Discrete Wavelet Transform (DWT) is a good method to be implemented in *batik* images. It has multi resolution dimension with various scale transforms and can result features by differentiating image intensity in several sub-band dimension (Virmani *et al.*, 2013).

A study of DWT optimization using wavelet transform level 3 daubechies 2 to achieve proper feature in traditional *batik* motifs texture has been conducted by Budiman *et al.* (2016). The accuracy value optimization retrieval was achieved using binary class non linier SVM-RBF kernel classification method. Energy feature and deviation standard used in the classification with *non-linear binary support vector* method to recognize *batik keraton* and *batik pesisir* images had accuracy at 96.7%. Hence, DWT level 3 daubechies 2 was the feature extraction method implemented to result feature vector of every image in this study. Wavelet coefficient value in every sub-band resulted from DWT was used to define a feature vector for every image in form of energy value and Standard Deviation (STD) of every sub-band.

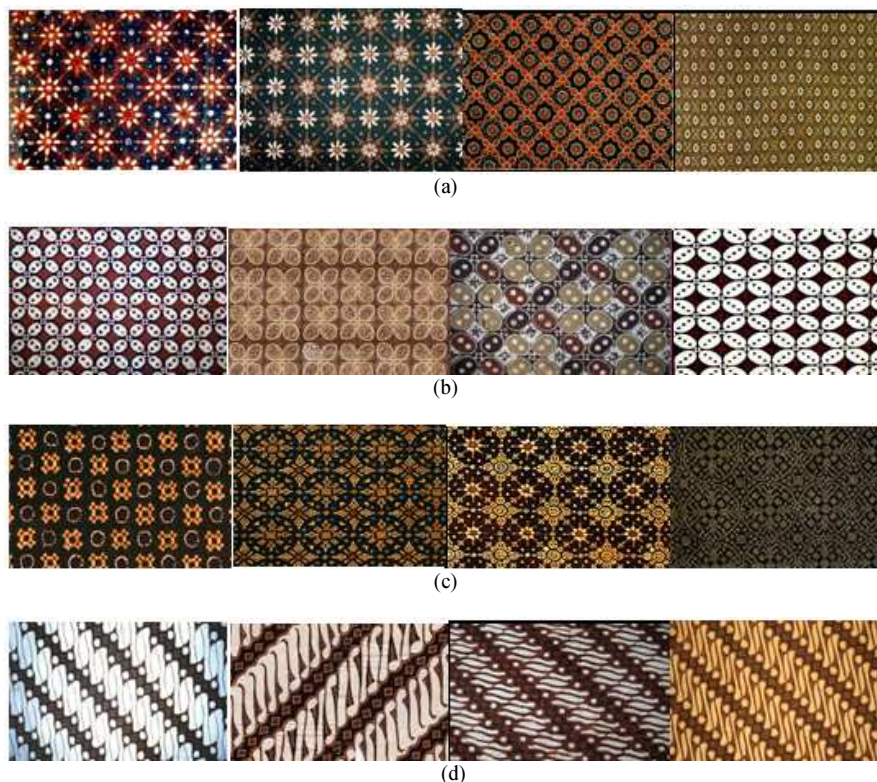


Fig. 1: Examples of traditional *batik* motifs (a) *ceplok* (b) *kawung* (c) *nitik* (d) *parang*

The performance of a combination of mean, energy and standard deviation as features has been compared and the use of energy and standard deviation is the best combination of image feature (Kokare *et al.*, 2003). Standard Deviation (STD) is used as a feature of every image since it can show a distribution of grey intensity; the lower STD value the more average distribution. The higher energy value as an image feature shows variation of texture elements (texel) with white color which is more dominant and the distribution of gray level/ the image structure becomes more irregular. The lower value shows variation of texel with black color which is more dominant and the distribution of gray level/the image structure becomes more regular.

Multi class SVM Classification with wavelet transform feature extraction integration has better tested than *Artificial Neural Network* (ANN) and empirical model to classify dataset of diffuse solar radiation (Shamshirband *et al.*, 2016). Classification using SVM method with default kernel parameter has higher classification than Minimum distance method and back propagation neural network in recognizing five *batik* motifs textures (Yuan *et al.*, 2014). In the texture classification of honey pollen images, SVM has better accuracy than classification methods of Multi-layer Perceptron, Minimum Distance Classifier and K-Nearest Neighbor (Fernandez-Delgado *et al.*, 2003). Classification optimization of computed tomography

medical images shows that SVM is better than *K-Nearest Neighbor* (Renukadevi and Thangaraj, 2013). SVM method is tested to two DNA micro array public datasets and compared to ANN method Ahmad (2013) and the result showed that SVM has success level than ANN.

Classification which uses SVM method for multi class and non-linear dataset needs to select the best kernel implementation to dataset. Kernel function optimization depends on dataset condition (Han and Michelin, 2006; Rosales-Perez *et al.*, 2013; Renukadevi and Thangaraj, 2013), so it needs to conduct kernel parameter value optimization to minimize classification error on test dataset with parameter estimation using cross validation dan grid search method (Hsu *et al.*, 2010). In previous studies, kernel function was recommended to be implemented is Radial Based Function (RBF). It has same performance with linear kernel in certain parameter (Rosales-Perez *et al.*, 2013; Renukadevi and Thangaraj, 2013). RBF is an ideal function since it has calculation that is efficient, simple and easy to adapt to its parameter optimization (Gaspar *et al.*, 2012; Shamshirband *et al.*, 2016).

A parameter estimation is needed since gaussian RBF kernel function is used for dot product mapping substitution from old dimension feature to new dimension depends on dataset condition (Han and Michelin, 2006; Rosales-Perez *et al.*, 2013; Renukadevi and Thangaraj, 2013).

A classification with maximum accurate value in recognizing *batik* images with various motifs, in which each *batik* image has a unique and non-linear feature motifs. The multi class SVM method needs experiments of the use of proper parameter range C and γ in the Gaussian RBF kernel, so it can define better new features (high dimensions) and result maximum classification function (hyperplane) for dataset of geometric motifs.

The best parameters combination for value C and γ has maximum accuracy from a classification result. The evaluation for optimizing parameters C and γ of non-linear multi-class SVM-RBF Gaussian kernel is used to achieve classification with small errors in dataset of traditional *batik* images with geometric motifs textures. The accuracy evaluation of classification was conducted. The value of the classification accuracy identified each use of parameters pairs of C and γ which were in the parameter scale. This was to optimize hyperplane in data training, so the classification can classify testing data in the proper class.

Materials and Method

This study which was a synthesis of several ideas of previous studies, aims to develop a method of determination the optimal value of space parameter in non-linear multi-class SVM-RBF kernel for classification of *batik* motifs with geometric ornament. First phase was to prepare Discrete Wavelet Transform (DWT) extraction feature optimization for classification of SVM-RBF kernel using standard parameter $C = 1$ and $\gamma = 0.5$. A comparison of wavelet types and level experiments was conducted to result the best vector feature for optimizing classification accuracy. Next phase was to optimize traditional *batik* motifs

recognition with geometric motifs. All phase conducted in this study is showed in Fig. 2. The first phase (step 1 to 4) has been conducted our previous research and has resulted DWT level 3 and type of db2 feature extraction method. This method was used in this research (step 5 to 7).

The experimental setup of the second phase consists of step 5 to 7 with dataset contains 80 images of traditional *batik* divided into four geometric motifs, which are *ceplok*, *ceplokan*, *kawung*, *nitik* and *parang/lereng*. The feature extraction was to find statistical feature value by conducting pre-process to define recognized feature classes and change every color in the images into grey scale mode. The optimization of the result of SVM classification with grid-search and cross validation process is to minimize over fitting and to achieve a combination of parameter value of RBF kernel (C and γ) in the space parameter which results maximum classification accuracy value. The classification evaluation is to classify by learning features served by the extraction result and to equate to the dataset features by measuring accuracy level using confusion matrix method. The evaluation was conducted in every combination of C and γ .

Evaluation process to find C and γ value is to result the best classification conducted using *one-against-all svm* (oaasvm) method. Accuracy calculation process which uses *confusion matrix* for every pair C and γ is showed in Fig. 3. Evaluation was conducted to achieve C and γ pair which has maximum accuracy value for classification result.

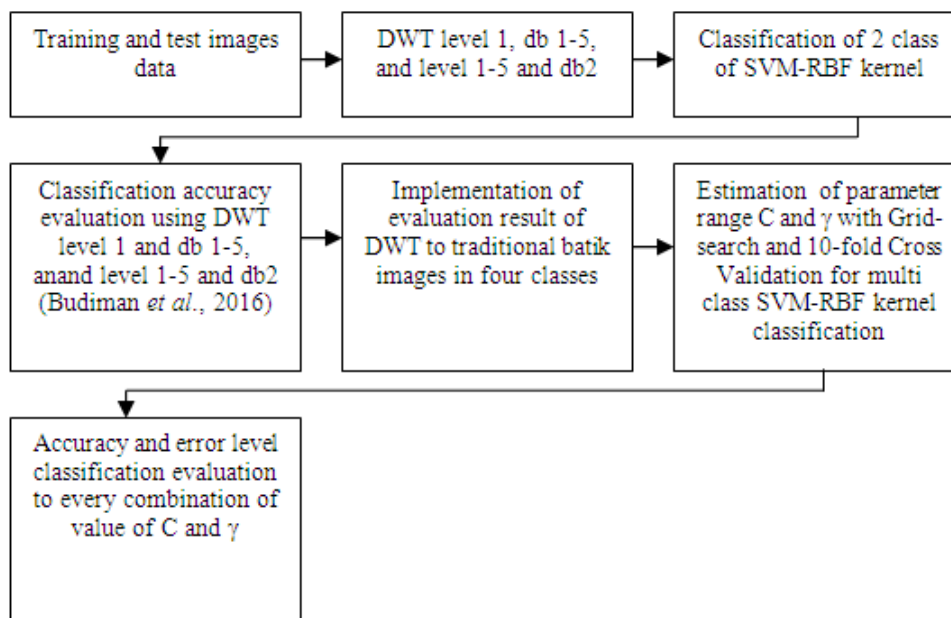


Fig. 2: Experiment phases

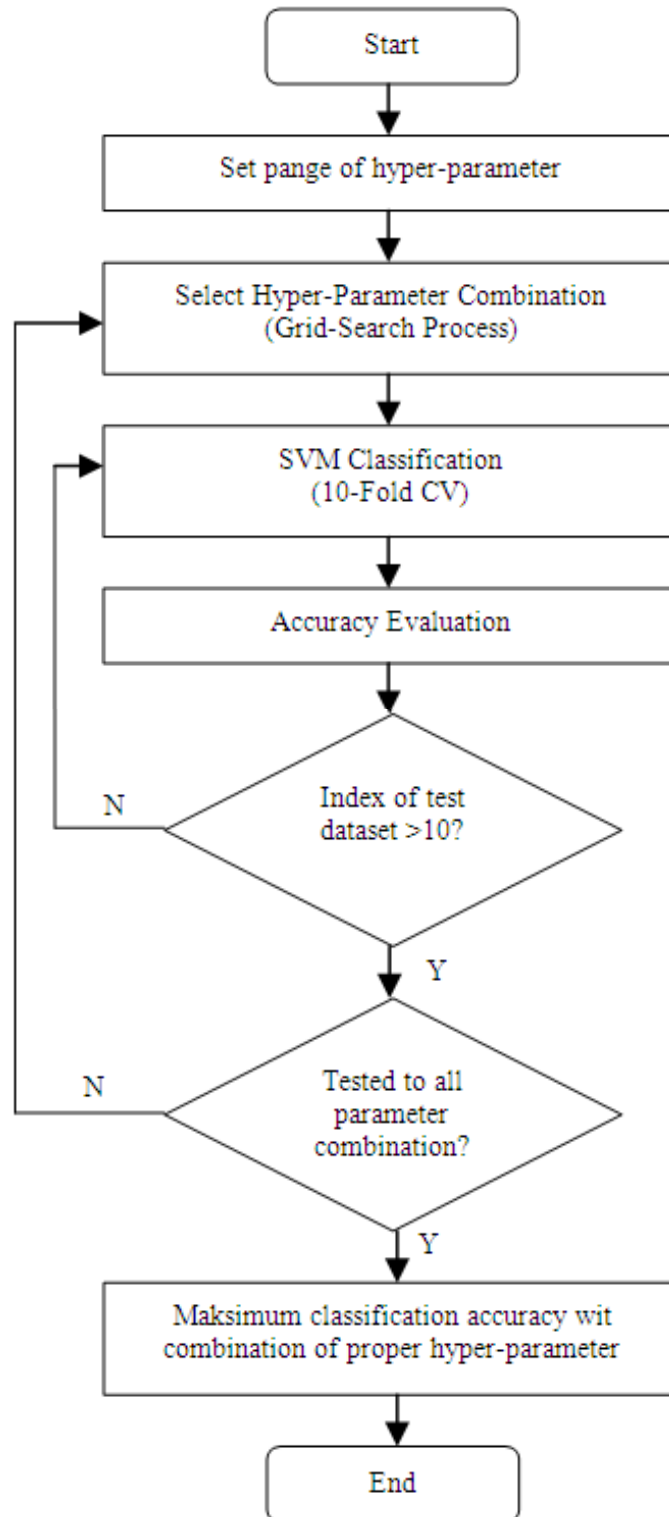


Fig. 3: Determination of SVM-RBF kernel parameter value optimization

In order to maximize classification pattern, non-linear SVM handles over fitting with soft margin conducted by replacing every dot product of the testing feature with

non-linear kernel function matrices (Boser *et al.*, 1992). This kernel function is defined with focus to find that the dot product of two data in feature space can be replaced by

kernel function with classification result (hyperplane) rather than not with knowing Φ mapping of what is used for every datum. Kernel function is used to map more than old training data to new training data which has feature space with higher dimension without defining mapping function of input space to new feature space (Hsu *et al.*, 2010):

$$\Phi = D^q \rightarrow D^r \quad (1)$$

x = training data, so $\{x_1, x_2, x_3, \dots, x_n\} \in D^q$, to new dimension feature $\{\Phi(x_1), y_1, \Phi(x_2), y_2, \dots, \Phi(x_n), y_n\} \in D^r$:

$$\Phi = D^2 \rightarrow D^3; (x_1, x_2) \rightarrow (x_1, x_2)^2 = (x_1^2, \sqrt{2x_1 \cdot x_2}, x_2^2) \quad (2)$$

Kernel function mapping, $K: D^2 \rightarrow D^3$:

$$\begin{aligned} K(x, x_j) &= (x, x_j)^2 \\ &= (x \cdot x_{j1} + x_2 \cdot x_{j2}) \\ &= (x_1^2 \cdot x_{j1}^2 + 2x_1 \cdot x_{j1} \cdot x_2 \cdot x_{j2} + x_2^2 \cdot x_{j2}^2) \\ &= (x_1^2, \sqrt{2x_1 \cdot x_2}, x_2^2)^T (x_{j1}^2, \sqrt{2x_{j1} \cdot x_{j2}}, x_{j2}^2) \\ &= \Phi(x)^T \cdot \Phi(x_j) \end{aligned} \quad (3)$$

The function is symmetric (*Mercer* theory):

$$K(x, xi) = \Phi(x)^T \cdot \Phi(xi) = \Phi(xi)^T \cdot \Phi(x) = K(xi, x) \quad (4)$$

Kernel trick to calculate kernel matrices/gram matrices $N \times N$ (N = total data dot) and every element value of matrices K (row, column) is used to replace dot product $X_i \cdot X_j$ in equation multi player Lagrange duality (Ld):

$$Ld = \sum_N^{i=1} \alpha_i - \frac{1}{2} \sum_{i,j} \alpha_i \alpha_j Y_i Y_j (X_i \cdot X_j) \quad (5)$$

SVM searches Lagrange parameter value of each feature ($\alpha_{1..n}$) for maximum Ld value and has condition $0 \leq \alpha_i \leq C$ and $\sum_{i=1..Ni} \alpha_i Y_i = 0$. SVM only uses chosen data dot (support vector) to build classification model. By knowing value α_i , so variable m (number of feature dot as support vector) can be achieved, which is feature data dot which has value $\alpha_i > 0$. Based on values α_i as support vector, optimum value w and b can be calculated by choosing one of values $\Phi(X_i)$ with positive class support vector for value x^+ and choosing one of values $\Phi(X_i)$ with negative class support vector for value x^- :

$$w = \sum_m^{i=1} \alpha_i Y_i \Phi(X_i) \quad (6)$$

$$b = \frac{1}{2} (w \cdot x^+ + w \cdot x^-) \quad (7)$$

Hyperplane:

$$f(x) = \sum_m^{i=1} \alpha_i Y_i K(X, X_i) + b = 0 \quad (8)$$

Margin:

$$f(x) = \sum_m^{i=1} \alpha_i Y_i K(X, X_i) + b = +1 \quad (9)$$

$$f(x) = \sum_m^{i=1} \alpha_i Y_i K(X, X_i) + b = -1 \quad (10)$$

To identify class of every data (X_d) in test dataset, kernel trick $K(X_d, X_i) = \Phi(x_d)^T \cdot \Phi(x_i)$ is used to replace dot product $X_d \cdot X_i$ on classification decision function:

$$f(x_d) = \text{sign} \left(\sum_m^{i=1} \alpha_i Y_i X_d \cdot X_i + b \right) \quad (11)$$

The classification error in linear SVM classification is reduced by margin optimization which maximizes hyperplane distance with support vector (Boser *et al.*, 1992), but features classification cannot be separated by linear. Kernel function is used as kernel trick in equation of SVM non-linear multi player Lagrange duality (Ld) to handle classification in SVM non-linear. There are several kernel function to replace $K(x, x') = \Phi(x)^T \cdot \Phi(x')$ used in SVM non linear literatures, which are *Polynomial of degree up to d*, *Gaussian RBF*, *Sigmoid*, *multi-quadratic inversion*, *Additive* (Hofmann, 2006).

The SVM non-linear training optimization very depends on the use of kernel function and its parameters. Gaussian RBF kernel is more recommended to result maximum non-linear classification for a new dataset (Hofmann, 2006; Hsu *et al.*, 2010; Rosales-Perez *et al.*, 2013; Renukadevi and Thangaraj, 2013), since it has same performance with linear kernel on parameter *cost* (C) and *gamma* (γ) with certain value in classification optimization. Gaussian RBF kernel is stated as:

$$K(x, x') = \exp \left(-\frac{\|x - x'\|^2}{2\sigma^2} \right) \quad (12)$$

$$\gamma = 1 / 2\sigma^2; \exp(x) = e^x$$

$$K(x, x') = \exp \left(-\gamma \|x - x'\|^2 \right) \quad (13)$$

$$\|x - x'\| x^T \cdot x + (x')^T \cdot x' - 2x^T \cdot x' \quad (14)$$

$$K(x, x') = \exp\left(-\gamma x^T \cdot x - \gamma (x')^T \cdot x' + \gamma x^T \cdot x'\right) \quad (15)$$

The hyper parameter estimation which is a constant parameter value for soft margin (C) and kernel parameter (γ) is needed to result a new test dataset classification with maximum SVM non-linear Gaussian RBF kernel (Rosales-Perez *et al.*, 2013; Renukadevi and Thangaraj, 2013). Parameter C and γ with proper value can keep ambiguous (measure of the error contribution) and variance (measure of the deviations) low when it is used for different training datasets with v-fold cross validation method and grid-search.

The high parameter C value and small γ value cause over fitting, while small value C and high value γ cause under fitting (Boser *et al.*, 1992). The smallest value γ omits more feature dots as support vector which are nearby hyperplane and increase maximum margin. The higher of gamma value/ γ increases support vector area and also flexibility from *decision boundary (hyperplane)*. The evaluation of training and testing in the use of geometric motifs dataset is needed to maximize result of SVM non-linear multi class classification with Gaussian RBF kernel. The evaluation is for success in estimation of *hyper parameter* RBF kernel. The grid search with v-fold cross validation is used in this study. Grid search is a model of hyper parameter value searching in certain interval range (Hsu *et al.*, 2010). The beginning range for this experiment is value $C = \{2^{-3}, 2^{-1}, \dots, 2^{15}, 2^{17}\}$ and $\gamma = \{2^{-17}, 2^{-15}, 2^{-13}, \dots, 2^1, 2^3\}$. The cross validation method divides dataset into v partition (v-fold) randomly and each partition has index number 1 to v. Common number of partition is 10 partition or 10-fold cross validation (Gaspar *et al.*, 2012; Virmani *et al.*, 2013; Syarif *et al.*, 2016). 10 times tests are conducted for 10 partitions by *leave-one-out* technique, in which one part is used alternately as test dataset and other parts (v-1) are used as training dataset. Every value in the interval is evaluated and the next tested value is addition of its parameter value exponentially. The next evaluation for smaller range between that best hyper parameter is conducted when maximum hyper parameter value has been defined.

The accuracy value measurement for every test dataset resulted from cross validation partition is conducted based on combination of parameter value used to classify. The measurement uses confusion matrix technique which is divided into positive data prediction class (positive right/BP; positive wrong/SP) and negative data (negative right/BN negative wrong/SN). Accuracy value is $(BP+BN) / (BP+BN+SP+SN)$.

Result and Discussion

The classification was conducted ten times using ten different training and test datasets. The classification was conducted for every C and γ parameter combination

evaluation (10-fold) to make sure there is no over fitting in evaluation with different test data. In every evaluation, training and test datasets are chosen randomly and equally for number of class 1, 2, 3 and 4 of vector file feature with size 80x21. A dataset which contains 80 traditional *batik* images was used in every training and test. The 10-fold method CV with leave-one-out process in feature vector images which was randomly selected resulted 8 feature vector for every test dataset used in every test.

The random selection were conducted using hold out process using percentages at 30% for test dataset and the rest were used for training dataset to increase number of test dataset. Process hold out in CV for vector file feature selected records for test data randomly. Class 1 (*ceplok* motifs) contained 1-20 records and 6 records were randomly selected; class 2 (*kawung* motifs) were 21-36 records and 4 record were randomly selected, class 3 (*nitik* motifs) were 37-55 records and 5 records were randomly selected; 7 record were selected in class 4 (*parang* motifs).

The experiment analysis for every combination of C and γ parameters value contained in grid-search range was used for ten times classifications, in which every training and test evaluation process used different training and test datasets resulted from CV. Ten times evaluations were conducted to avoid over fitting and under fitting that may result very different accuracy for different test datasets. The optimization of parameter C and γ was to achieve RBF kernel value which can result multi class non linear SVM classification with low bias (measure of the error contribution) and low variance (measure of deviation).

The best parameter value estimation used parameter value range known with grid-search method. The pre-evaluation of grid-search method was conducted using wide range of parameter C and γ . Narrower range of parameters was conducted when the best parameter was found to achieve parameter value that can result the best classification accuracy value. There is no rule for proper range in grid-search method. The wider parameter range was effective in resulting combination C and γ as the best parameter to significantly increase accuracy value in image retrieval. The proper parameter value of C and γ can maximize classification function (hyperplane) and adjust the balance of margin distance (+1 and -1) with proper hyperplane. There were four hyperplanes and margins (+1 and -1) for four geometric motifs classes in SVM classification using one-against-all, so there were four times training to result hyperplane and margin. The training is for non-linear binary class SVM-RBF classification in *ceplok* motifs (+1) with non-*ceplok* motifs (-1); *kawung* motifs (+1) with non-*kawung* motifs (-1); *nitik* motifs (+1) with non-*nitik* motifs (-1); and *parang* motifs (+1) with non-*parang* motifs (-1).

The phases conducted for experiments of every value pair of C and γ used one-against-all svm (oaasvm) method and accuracy calculation process used confusion matrix for every value pair of C and γ was described in algorithm 1. The evaluation of every pair value of C dan γ was conducted to achieve pair of C and γ which has maximum classification accuracy value.

Algorithm 1. Accuracy evaluation in every classification of value pair of C and γ
 For value i = 1:10

```

Load 'training_data (i)', 'test_data_class (i)';
model = oaasvm training(training_data (i),
test_data_class (i), 'rbf', C= n,  $\gamma$  = m);
svm training training_data (i) class 1 (class 1 and –
class 1);
svm training training_data (i) class 2 (class 2 and –
class 2);
svm training training_data (i) class 3 (class 3 and –
class 3);
svm training training_data (i) class 4 (class 4 and
class 4);
load 'test_data (i);
result (i) = 1-10;
end (i)

for value i = 1 : 10
load 'test_data_class (i);
confusion_matrix (test_data_class (i), result (i);
for nilai k = 1 : 22
if test_data_class (i, k) = hasil (i, k)
true = true +1;
end

false = 22 – true;
accuracy = true/22;
error = false/22;
result = [error accuracy]
list = [list, asil]
end
    
```

The parameter value range in the beginning of experiment is $C = \{2^{-3}, 2^{-1}, 2^1, 2^3, 2^5, \dots, 2^{13}, 2^{15}, 2^{17}\}$ and $\gamma = \{2^{-17}, 2^{-15}, 2^{-13}, \dots, 2^{-5}, 2^{-3}, 2^{-1}, 2^1, 2^3\}$. It is a standard parameter as binary classification with value C = 1 and $\gamma = 0.5$. Table 1 shows accuracy and error level in every evaluation of the beginning experiment which used 10 training datasets and 10 test datasets. The evaluation resulted the average accuracy = 0.354545455 and average error = 0.645454545. The use of parameter C = 1 and $\gamma = 0.5$ has low accuracy that is lower than 0.5, since it cannot result hyperplane and margin which can well recognize test images for class 1, 2 and 3. High bias in the use of parameter C = 1 and $\gamma = 0.5$ caused under fitting for class 1, 2, dan 3.

Table 1: Accuracy and error level of C = 1 and $\gamma = 0.5$

No	Accuracy	Error level
1	0.363636364	0.636363636
2	0.318181818	0.681818182
3	0.363636364	0.636363636
4	0.363636364	0.636363636
5	0.363636364	0.636363636
6	0.409090909	0.590909091
7	0.363636364	0.636363636
8	0.318181818	0.681818182
9	0.318181818	0.681818182
10	0.363636364	0.636363636

Table 2: Accuracy and error level of C = 2⁵ and $\gamma = 2^{-11}$

No	Accuracy	Error level
1	0.590909091	0.409090909
2	0.318181818	0.681818182
3	0.545454545	0.454545455
4	0.545454545	0.454545455
5	0.500000000	0.500000000
6	0.590909091	0.409090909
7	0.500000000	0.500000000
8	0.545454545	0.454545455
9	0.454545455	0.545454545
10	0.454545455	0.545454545

The next experiment used parameter $C = \{2^{-3}, 2^{-1}, 2^1, 2^2, 2^3\}$ and $\gamma = \{2^3, 2^1, 2^{-1}, 2^{-3}, 2^{-5}, 2^{-7}, 2^{-9}\}$ and there was no significant change. Motifs of class 1, 2 and 3 cannot be well recognized. These motifs started to be recognized in parameter combination of C = 25 dan $\gamma = 2^{-11}$ which has average accuracy = 0.504545455 and average error = 0.495454546, with accuracy value and error level as described in Table 2.

The one-against-all method in non-linear classification SVM with four class consists of four binary class non-linear SVM classification, so parameter value of C and γ must maximize classification result to form four hyperplanes and margin (+1 and -1) and parameter value of $C \leq 2^3$ and $\gamma > 2^{-9}$ cannot recognize motifs of class 1, 2 and 3. The parameter value $C \leq 2^3$ with calculation of maximum Lagrange multiplier equation value resulted more value of $\alpha_i > 0$ for more soft margin forming. The more feature vector as support vector at parameter value of $C \leq 2^3$ made soft margin maximizing margin by ignoring support vector which is nearby hyperplane and used support vector dot with maximum distance of hyperplane.

The parameter value of $C \leq 2^3$ which caused margin distance with maximum hyperplane causes high bias, since there was under fitting or it was not able to recognize class for feature vector (x_i) which was between margin limit and hyperplane ($-1 < x_i \leq 0$ and $0 < x_i < +1$). Value of $\gamma > 2^{-9}$ resulted higher value $-\gamma \|x - x_i\|^2$ in RBF kernel. This effected in the increasing of support vector area and hyperplane flexibility, so it caused margin forming linkages between support vectors which are close together. Margin

limit cannot recognize feature vector extensively and classification result tends to high bias due to under fitting.

The next experiment used parameter value of $C \geq 2^5$ and $\gamma \leq 2^{-11}$. The balance of parameter value of C and γ needed to be achieved for maximum accuracy value result of four binary class non-linear SVM classifications by optimizing hyperplane and margin of those classifications. The parameter optimization was conducted to achieve low bias and low variance in traditional *batik* images retrieval for four geometric motifs classes. The next experiments used parameter value of $C \geq 2^5$ and $\gamma \leq 2^{-11}$. The high value for C and lower value for γ caused over fitting that results high variance. The higher value of C which made maximum equation value calculation of lagrange multiplier was resulted for less value of $\alpha_i > 0$, so the number of support vector became less and be at nearest distance to hyperplane. The higher value of C made the number of support vector less and made margin distance and hyperplane became narrower. This caused over fitting, or error in classification class area placement to feature vector which were nearby hyperplane. The lower value of γ caused less support vector, so the forming margin had decision boundary which was similar to linear.

The maximum retrieval with low bias and low variance at parameter range defined in this experiment was achieved at parameter value of $C = 2^7$ and $\gamma = 2^{-15}$ with the number of feature vector properly recognized was 17 to 19 of 22 test dataset feature vector in every evaluation. The result of accuracy for ten times evaluation with parameter value of $C = 2^7$ and $\gamma = 2^{-15}$ and different test dataset in every evaluation, as shown in Table 3, set the average accuracy = 0.813636364 and average error = 0.186363636.

The next evaluation of value combination was conducted with parameter $C \geq 2^9$ and $\gamma \leq 2^{-17}$ (Table 4). The accuracy value in every evaluation relatively decreased and the measure of deviations (variance) increased that caused an increase in over fitting in the class definition of testing feature data.

Based on the result of classification with accuracy value as described in Table 4, the grid-search evaluation was conducted in lower range which were between $2^5 < C < 2^9$ and $2^{-17} < \gamma < 2^{-15}$, so the next evaluation with grid-search was conducted in range that were lower than $C = \{2^{6.5}, 2^{6.75}, 2^7, 2^{7.25}, 2^{7.5}, 2^{7.75}, 2^8\}$ and $\gamma = \{2^{-14.5}, 2^{-14.75}, 2^{-15}, 2^{-15.25}, 2^{-15.5}, 2^{-15.75}, 2^{-16}\}$.

There is no significant change in the experiment analysis of the use of the parameter combination with that range. The best result was still the parameter combination of $C = 2^7$ and $\gamma = 2^{-15}$. This experiment showed that parameter estimation to optimize the performance of SVM-RBF kernel classification was needed to achieve accuracy value in traditional *batik* image retrieval that has geometric motifs texture which as multi scale motifs and multi resolution color.

Table 3: Accuracy and error level of $C = 2^7$ dan $\gamma = 2^{-15}$

No	Accuracy	Error level
1	0.863636364	0.136363636
2	0.772727273	0.227272727
3	0.863636364	0.136363636
4	0.863636364	0.136363636
5	0.772727273	0.227272727
6	0.818181818	0.181818182
7	0.818181818	0.181818182
8	0.818181818	0.181818182
9	0.772727273	0.227272727
10	0.772727273	0.227272727

Table 4: Accuracy value and the use of parameter value $C = \{2^7, 2^9, 2^{11}, 2^{13}\}$ dan $\gamma = \{2^{-15}, 2^{-17}\}$

Test	The value of accuracy				
	$C = 2^7; \gamma = 2^{-15}$	$C = 2^7; \gamma = 2^{-17}$	$C = 2^9; \gamma = 2^{-15}$	$C = 2^{11}; \gamma = 2^{-15}$	$C = 2^{13}; \gamma = 2^{-15}$
1	0.864	0.818	0.773	0.773	0.818
2	0.773	0.682	0.818	0.773	0.773
3	0.864	0.818	0.773	0.682	0.682
4	0.864	0.773	0.773	0.682	0.682
5	0.773	0.682	0.773	0.773	0.682
6	0.818	0.773	0.818	0.773	0.818
7	0.818	0.682	0.682	0.682	0.636
8	0.864	0.773	0.864	0.864	0.864
9	0.773	0.773	0.682	0.682	0.636
10	0.773	0.682	0.773	0.773	0.773

The maximum parameter value and the smaller parameter value range defined in this experiment can be a reference to SVM-RBF kernel classification of digital image retrieval with geometric motifs textures.

Conclusion

The optimization of RBF kernel parameter to achieve maximum accuracy value in non-linear SVM classification method for traditional *batik* image retrieval that has four geometric motifs texture has been successfully conducted in this study. The image feature vector used as training and test datasets of SVM-RBF kernel classification can inform image feature properly. The feature vector was resulted from feature extraction of traditional *batik* images dataset with DWT level 3 and type of db2 which has been evaluated in the beginning phase. The feature vector can inform the image characteristic. The parameter optimization of C and γ conducted using grid-search and 10 fold Cross Validation with holdout is resulted by 10 training datasets and 10 test datasets from feature vector selected randomly. SVM-RBF kernel classification for every combination of parameter C and γ in grid-search range has been conducted using 10 training datasets and 10 test datasets. 10 times evaluation for every combination of C

and γ that used different test dataset for each evaluation achieved parameter combination of C and γ which result low bias and low variance.

Confusion matrix is used for accuracy value measurement conducted 10 times for every parameter combination of C and γ of RBF kernel and different test dataset for each evaluation. Every parameter combination of C and γ used in RBF kernel had 10 accuracy values to show that there was no significant different for error retrieval when a different dataset was used. The maximum accuracy value for parameter $C = 2^7$ and $\gamma = 2^{-15}$ are 0.77 to 0.86 with success image retrievals in every evaluation were 17 to 19 images of 22 images and had average accuracy = 0.813636364 and average error = 0.186363636. Based on the result in this research, a classification of images with geometric motifs texture using non-linear multi class SVM-RBF kernel are recommended to use grid-search range of $C = \{2^{6.5}, 2^{6.75}, 2^7, 2^{7.25}, 2^{7.5}, 2^{7.75}, 2^8\}$ and $\gamma = \{2^{-14.5}, 2^{-14.75}, 2^{-15}, 2^{-15.25}, 2^{-15.5}, 2^{-15.75}, 2^{-16}\}$.

In order to complete the *batik* motifs knowledge base, the range value of C and γ recommended in this research is still need to be evaluated with SVM-RBF kernel non-linear for more than four classes of *batik* motifs, as well as a study of feature extraction with feature combination using another methods and evaluation for process time and accuracy of classification.

Acknowledgement

The author would like to thank to Mr. Arry Maulana Syarif and Mr. Edi Sugiharto for supporting the codes writing and programming in MatLab.

Author's Contributions

Fikri Budiman: Dissertation author who wrote his research in this manuscript.

Adang Suhendra: Principal advisor of the first author's doctoral dissertation.

Dewi Agushinta: Advisor of the first author's doctoral dissertation.

Avinanta Tarigan: Advisor of the first author's doctoral dissertation.

Ethics

The authors ensure that there is no ethical issue that may rise after the publication of this manuscript.

References

Ahmad, M.S., 2013. Wavelet-based feature extraction for DNA microarray classification. *Artificial Intell. Rev.*, 39: 237-249. DOI: 10.1007/s10462-011-9269-z

- Boser, B.E., I. Guyon and V. Vapnik, 1992. A training algorithm for optimal margin classifiers. *Proceedings of the 5th Annual Workshop on Computational Learning Theory*, Jul. 27-29, ACM, New York, pp: 144-152. DOI: 10.1145/130385.130401
- Budiman, F., A. Suhendra, D. Agushinta and A. Tarigan, 2016. Wavelet decomposition levels analysis for indonesia traditional *batik* classification. *J. Theoretical Applied Inform. Technol.*, 92: 389-394.
- Fernandez-Delgado, M., P. Carrión, E. Cernadas, J.F. Gálvez and P. Sá-Otero, 2003. Improved classification of pollen texture images using SVM and MLP. *Proceedings of the International Conference on Visualization and Image Processing*, Sep. 8-10, ACTA Press, Benalmádena, Spain, pp: 1-1060.
- Gaspar, P., J. Carbonell and J.L. Oliveira, 2012. On the parameter optimization of support vector machines for binary classification. *J. Integrative Bioinformat.*, 9: 201-201.
- Han, J. and K. Michelin, 2006. *Data Mining: Concepts and Techniques*. 2nd Edn., Elsevier Inc, San Fransisco.
- Hofmann, M., 2006. *Support Vector Machines- Kernels and the Kernel Trick*. 1st Edn., Bamberg University.
- Hsu, C.W., C.C. Chang and C.J. Lin, 2010. *A practical guide to support vector classification*. Department of Computer Science, National Taiwan University, Taipei 106, Taiwan.
- Kokare, M., B.N. Chatterji and P.K. Biswas, 2003. Wavelet transform based texture feature for content based image retrieval. *Proceedings of the 9th National Conference on Communications, (NCC' 03)*, Chennai, India, pp: 443-447.
- Renukadevi, N.T. and P. Thangaraj, 2013. Performance evaluation of SVM - RBF kernel for medical image classification. *Global J. Comput. Sci. Technol. Graph. Vis.*, 13: 1-7.
- Rosales-Perez, A., H.J. Escalante, J.A. Gonzalez and C.A. Reyes-Garcia, 2013. Bias and variance optimization for SVMs model selection. *Proceedings of the 26th International Florida Artificial Intelligence Research Society Conference, (RSC'13)*, St. Pete Beach, Florida, USA, pp: 136-141.
- Shamshirband, S., K. Mohammadi, H. Khorasanizadeh, P.L. Yee and M. Lee *et al.*, 2016. Estimating the diffuse solar radiation using a coupled support vector machine-wavelet transform model. *Renewable Sustainable Energy Rev.*, 56: 428-435. DOI: 10.1016/j.rser.2015.11.055
- Syarif, I., A. Prugel-Bernett and G. Wills, 2016. SVM parameter optimization using grid search and genetic algorithm to improve classification performance. *Telkomnika J.*, 14: 1502-1509. DOI: 10.12928/TELKOMNIKA.v14i4.3956
- Tirta, I., 2009. *Batik Sebuah Lakon*. 1st Edn., Gaya Favorit Press, Jakarta, ISBN: 9789795154280, pp: 278.

Virmani, J., V. Kumar, N. Karla and N. Khandelwal, 2013. SVM-Based characterization of liver ultrasound images using wavelet packet texture descriptors. *J. Digital Imag.*, 26: 530-543.
DOI: 10.1007/s10278-012-9537-8

Yuan, Q.N., J. Lu, H. Huang and W. Pan, 2014. Research of *batik* image classification based on support vector machine. *Comput. Modell. New Technol.*, 18: 193-196.