

# Cognitive-Affective Emotion Classification: Comparing Features Extraction Algorithm Classified by Multi-class Support Vector Machine

Nova Eka Diana\*, Ahmad Sabiq

Faculty of Information Technology, YARSI University, Jakarta, Indonesia.

\* Corresponding author. Tel.: +6281230973641; email: nova.diana@yarsi.ac.id

Manuscript submitted August 14, 2015; accepted December 20, 2015.

doi: 10.17706/ijcce.2016.5.5.350-357

---

**Abstract:** Emotional quotient (EQ) is one of the main factors determining the outcome of a learning process. A cognitive-affective states that usually appear during a learning process are bored, confuse, and excited/enthusiastic. Emotion state can be detected by identifying human facial expressions. Here, Principal Component analysis (PCA) and Gabor features extract salient information from facial expression database. Each feature space obtained from these methods is then classified using multi-class Support Vector Machine (SVM) with two cross-validation methods, Holdout and 10-fold cross validation. Experiment results show that classification process using Gabor features and 10-fold cross validation of multi-class SVM give the best accuracy rate.

**Key words:** Cognitive-affective emotions, features extraction, multiclass classification, cross-validation.

---

## 1. Introduction

In academic learning, study outcome is not only determined by Intelligent Quotient (IQ). Another factor such as Emotional Quotient (EQ) also has a significant role in deciding the output of each student. IQ only participates about 20% for the success of learning process, and 80% is affected by other parameters such as EQ. Emotional Quotient (EQ) is a competence to motivate own self, control negative emotion, redeem frustration, empathize and work together in a group of people [1].

P. Ekman divided human emotion into six primary groups, which are fear, anger, happiness, sadness, disgust, and surprise [2]. The relevance of these feelings toward a learning process is still being questioned by many researchers. Hence, they tried to find another alternative term of emotions that affecting the output of learning process. William Damon classified emotion into two categories, positive and negative emotion. Negative emotion may motivate willingness to study by giving a punishment when the student fails to achieve the goal. Otherwise, positive emotion can increase students' empathy towards people and process in a learning environment [3]. Instead of using basic emotions to measure the output of learning process, several researchers suggested using a set of cognitive-affective states as emotions that usually arose during a learning session. Those affective states were boredom, confusion, delight, engaged concentration, and surprise that could be identified based on the human facial expression [4]. Paul Ekman has proposed 46 Action Unit (AU) which express facial features movements as a form of emotion representation [5].

Computer vision techniques have been widely used to process human facial images, either for detection

or recognition purpose. Two primary processes in face recognition areas are feature extraction and recognition or classification. In the past years, many researchers had used Principal Component Analysis (PCA) or known as Karhunen-Loeve method for face recognition purpose [6]-[8]. The main idea of this algorithm is representing the significant variations in facial images in a lower dimensionality size. Hence, it can reduce the cost of computing extraction process both of memory and time consumption. Another popular algorithm for features extraction is Gabor features. Many other researchers also employed Gabor filter to extract significant features from facial image database [9]-[12]. The main reason of Gabor features popularity is its insensitivity towards pose variations and lighting condition. Hence, it can keep useful features as much as possible [13], [14].

Support Vector Machine (SVM) and its variant, Multi-class SVM, have been widely used to classify data into a respected group based on fitted parameters condition. As opposed to the initial purpose of SVM, which is only processing binary classification, Multi-class SVM tries to map data into  $n$  multiple of classes, with  $n > 2$ . Many approaches have been proposed to compute multiclass classification effectively. Those methods are "one-against-all", "one-against-one", "directed acyclic graph (DAG)", and ECOC (Error Corrected Output Coding)" [15]-[17].

The purpose of this research is to classify human emotion based on facial expressions image. We focus only on three cognitive-affective emotions which affecting the outcome of a learning process: bored, confuse, and excited or enthusiastic. Here, we compare the correct rate of PCA and Gabor features extraction methods classified with multi-class SVM.

## 2. Features Extraction

### 2.1. Principle Component Analysis (PCA)

Principal Component Analysis (PCA) which also called as Karhunen-Loeve expansion has been widely used to create features representation of relevant information in data, such as images database. The goal of PCA is to reduce the dimensionality of image matrix representation while keeping as much as possible useful features and variations present in the original database. Given  $A$ , matrix representation of all pictures in the database with the dimensionality of  $N$ , PCA will reduce its dimensionality to  $K$  where  $K \ll N$ .

Assume that facial images database consist of  $n$  numbers of data,  $\{x_i\}$  where  $i = 1, 2, \dots, n$ . Each image has a dimensionality of  $N$  (width  $\times$  height of the image). Equation (1) denotes the projection from facial space into features space of dimension  $K$ , where  $K < N$ . The primary purpose of PCA is to maximize the variance of  $\{y_i\}$  as described in equation (2). Let  $S_x$  is the covariance matrix of  $\{x_i\}$ , and then by using Lagrangian multiplier we can get the eigenvector as defined in equation (3).

$$y = Ax, A = [U_1^T, U_2^T, \dots, U_K^T] \quad (1)$$

$$A^* = \arg \max_A (S_y)^T \quad (2)$$

$$S_y = \frac{1}{n} \sum_{i=1}^N (y_i - \mu)(y_i - \mu)^T$$

$$\mu = \frac{1}{n} \sum_{i=1}^N X_i$$

$$X_i = \sum_{k=1}^N (x_i^T U_k) U_k \quad (3)$$

$$S_x U_k = \lambda_k U_k$$

where,  $U_k$ : The eigenvector of  $S_x$  corresponding to the  $k$ -largest eigenvalues.

## 2.2. Gabor Features

Gabor filter has characteristics that less sensitive towards pose and orientation variations. Equation (4) defines Gabor filters for features extractions.

$$\psi_{\mu,\nu} = \frac{\|k_{\mu,\nu}\|^2}{\sigma^2} e^{(-\|k_{\mu,\nu}\|^2 \|z\|^2 / 2\sigma^2)} \left[ e^{ik_{\mu,\nu}^z} - e^{-\frac{\sigma^2}{2}} \right] \quad (4)$$

$$\begin{aligned} k_{\mu,\nu} &= k_v e^{i\varphi u} \\ k_v &= \frac{k_{\max}}{f^v}, \varphi u = \pi\mu/8 \end{aligned} \quad (5)$$

where,

$\mu$ : Orientation

$\nu$ : Scale of Gabor filters

$z$ : Pixel value at  $(x, y)$  position

$k_{\max}$ : Maximum frequency

$f$ : Spacing factor between kernels in the frequency domain

Gabor features perform convolution by employing Gabor filter to extract features information from the database as described in equation (6).

$$F_{\mu,\nu}(z) = I(z) \times \psi_{\mu,\nu}(z) \quad (6)$$

## 3. Multi-class Support Vector Machine

Multi-class SVM is employed to tackle multiclass classification problems. Here, each point in the data training could belong to more than one different groups of class. The main purpose of multiclass classification is to construct a function that will correctly predict the label to which a data point should belong. In this research, we use "one-against-all" or "one-versus-rest" method to perform multiclass classification of facial expression emotion. Given  $k$  numbers of a different class, then it will construct  $k$  SVM models. Each  $i$ -th SVM model is trained using all facial images in the  $i$ -th class as correct labels, and the rest of pictures as negative labels.

Given  $m$  numbers of training data, then the  $i$ -th SVM model will solve the problem as stated in (7). The function will map training data  $x_i$  to a higher dimensional space. If it cannot be linearly separated, then penalty parameter  $C$  will be given. The purpose of this computation is to maximize the distance between two classes of data. Hence,  $x$  belongs to one class with the largest value of decision function as depicted in (8).

$$\begin{aligned} \min_{w^i, b^i, \xi^i} \quad & \frac{1}{2} (w^i)^T w^i + C \sum_{j=1}^l \xi_j^i \\ & (w^i)^T \phi(x_j) + b^i \geq 1 - \xi_j^i, \text{ if } y_j = i. \\ & (w^i)^T \phi(x_j) + b^i \leq -1 + \xi_j^i, \text{ if } y_j \neq i. \\ & \xi_j^i \geq 0, j = 1, \dots, m \end{aligned} \quad (7)$$

$$\text{class of } x \equiv \arg \max_{i=1, \dots, k} \left( (w^i)^T \phi(x) + b^i \right) \quad (8)$$

## 4. Results and Discussions

This research has a purpose of classifying student cognitive-affective state based on still images of facial expressions. Based on the observation of classroom activities and literature mentioned, three dominants emotions that arise during learning process are *Bored*, *Confuse*, and *Excited/Enthusiastic*.

#### 4.1. Emotions Database

Human emotions can be perceived by interpreting facial expression or mimics. A total of 408 facial image expressions of Indonesian student were collected to build a database that represents three kinds of emotions: Bored, Confuse, and Excited/Enthusiastic. Each facial expression of 300x300 pixels size was captured from frontal view direction with the non-uniform lighting condition.

Fig. 1 illustrates a sample of facial expression images in the database. Several different expressions can portray one particular emotion state. The first two columns (a) depict different kinds of facial mimics that identify a bored emotion state. The next two columns (b) shows expression for a confuse emotion. Last columns (c) express mimics for excited/enthusiastic feeling.

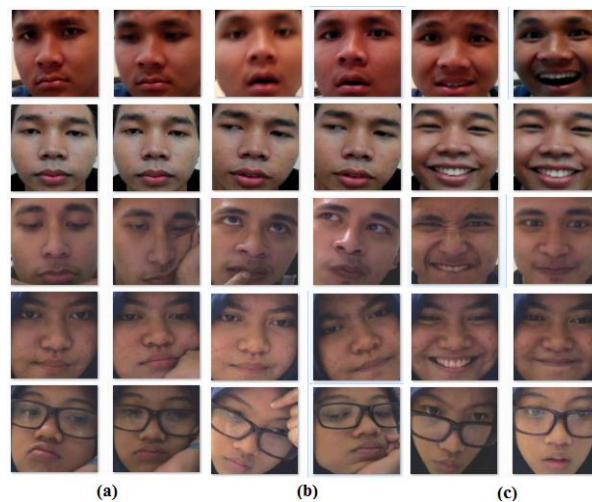


Fig. 1. Emotion expressions of (a) bored (b) confuse and (c) excited/enthusiastic.

#### 4.2. Application Architecture

Fig. 2 shows the application architecture to identify an emotional state of facial expressions data. First, image pre-processing is executed to maximize the quality of data by eliminating noises or unimportant information in pictures. Next, PCA and Gabor features algorithm is employed to extract salient features of facial expressions. Then, multi-class SVM classified the testing data based on the training data that selected from the database. Classification model will map the testing data to an appropriate cluster and generate the correct rate value that represents the accuracy of multi-class SVM model.

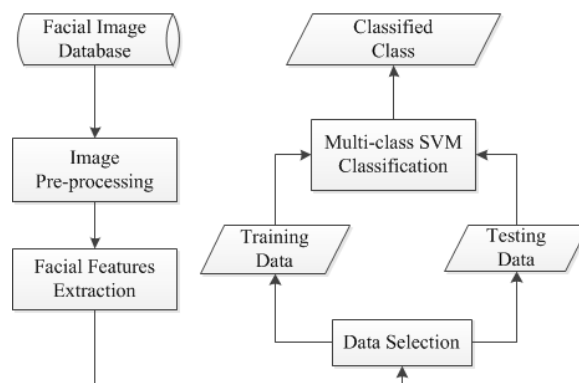


Fig. 2. Emotion classification architecture.

### 4.3. Discussions

Two algorithms, Principal Component Analysis (PCA) and Gabor Features, were employed to generate two set of facial features extracted from a single database. Each features space was then classified using Multi-class Support Vector Machine (multi-class SVM). Holdout and 10-fold cross validation were employed to estimate the performance of multi-class SVM model. The performance of the methods was justified by comparing the correct rate of facial emotion classification. Table 1 shows four types of method combination that were conducted to find the best approach to classifying facial emotion. The fitness rate for each procedure was measured by executing five times of experiments for each method. Three correct labels defined for multi-class emotion classification are bored, confuse, and excited.

Table 1. Method Combination

Method Combination	Extraction/Classification Algorithm
PCA-Holdout	PCA / Multi-class SVM with holdout cross validation
PCA-10fold	PCA / 10-fold cross-validation of Multi-class SVM
Gabor-Holdout	Gabor features / Multi-class SVM with holdout cross validation
Gabor-10fold	Gabor features / 10-fold cross-validation of Multi-class SVM

Table 2 depicts the comparison of correct rate classification for each method mentioned above. All methods give a very high value of correct rate in executing emotion classification, which is above 70%. Among all approaches, Gabor features and 10-fold cross-validation of multi-class SVM can classify emotions accurately with a high correct rate value about 94.8%. Otherwise, the classification rate for features space extracted by PCA gives the lowest correct rate. Overall, multi-class SVM classification combined with 10-fold cross-validation can better classify facial images into the respected affective-cognitive emotion.

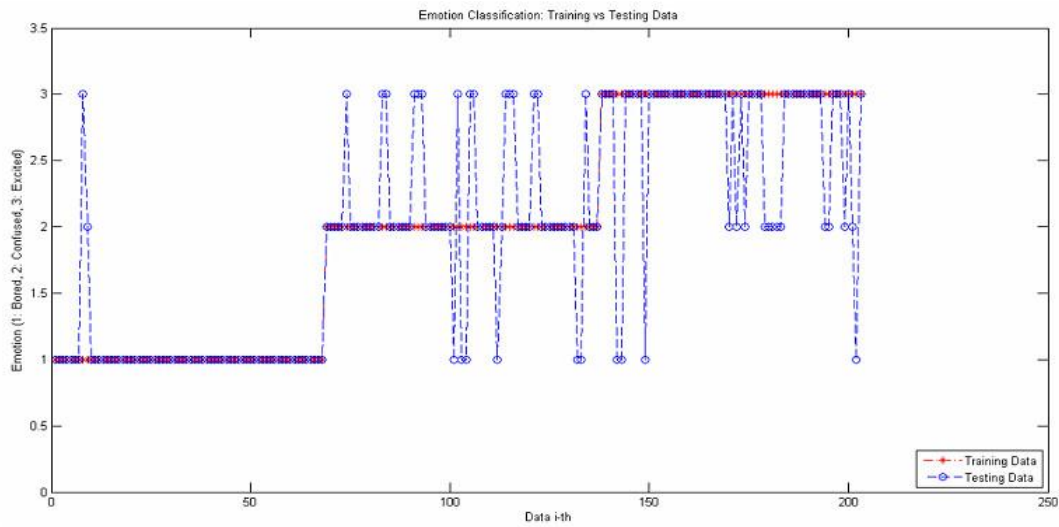
Table 2. Emotion Classification Comparison

Experiment	Correct rate			
	PCA-Holdout	PCA-10fold	Gabor-Holdout	Gabor-10fold
1	0.8079	0.8834	0.8784	0.9505
2	0.8325	0.9012	0.9057	0.9381
3	0.8030	0.9212	0.9007	0.9505
4	0.8276	0.9329	0.9181	0.9381
5	0.8473	0.9329	0.9256	0.9628
Average	0.8237	0.91432	0.9057	0.9480

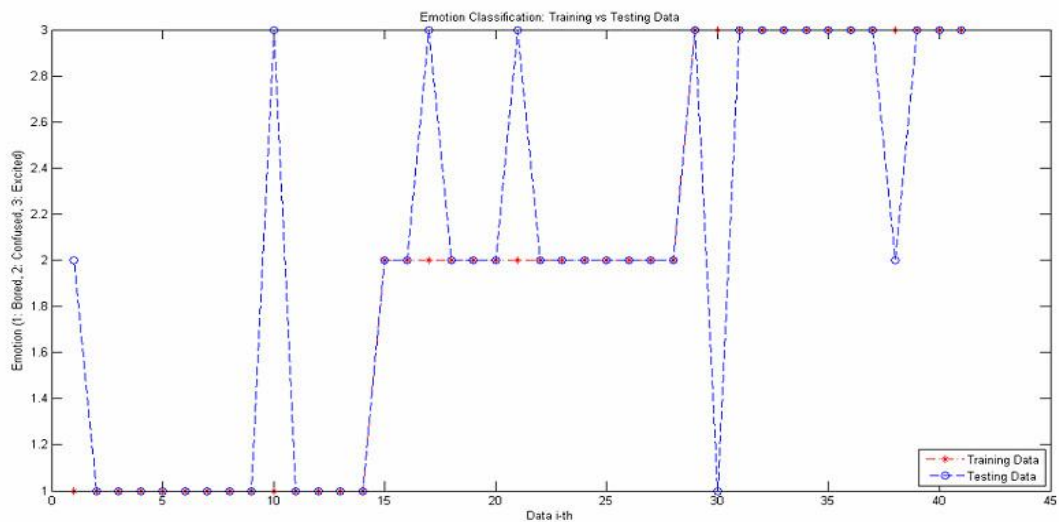
Fig. 3 illustrates the classification result for three affective-cognitive emotions: bored, confuse, and excited/enthusiastic using methods as mentioned earlier (PCA-Holdout, PCA-10fold, Gabor-Holdout, and Gabor-10fold). Comparative analysis of experiments results is justified based on two criteria, features extraction quality and correct rate classification. Gabor filter outperforms PCA algorithm in extracting the best features of emotional expressions. PCA is very sensitive to changes in lighting condition, pose variations, and facial expressions. It yields some valuable information might be eliminated during an extraction process. Otherwise, Gabor filters is insensitive to pose variations and orientations so it can keep much necessary information in image facial expression.

In term of classification result, 10-fold cross-validation combined with multi-class SVM give the best accuracy rate than Holdout method. Holdout method uses brute force approach that employed an approximator function for predicting the classification outcome of the testing data. This method introduces

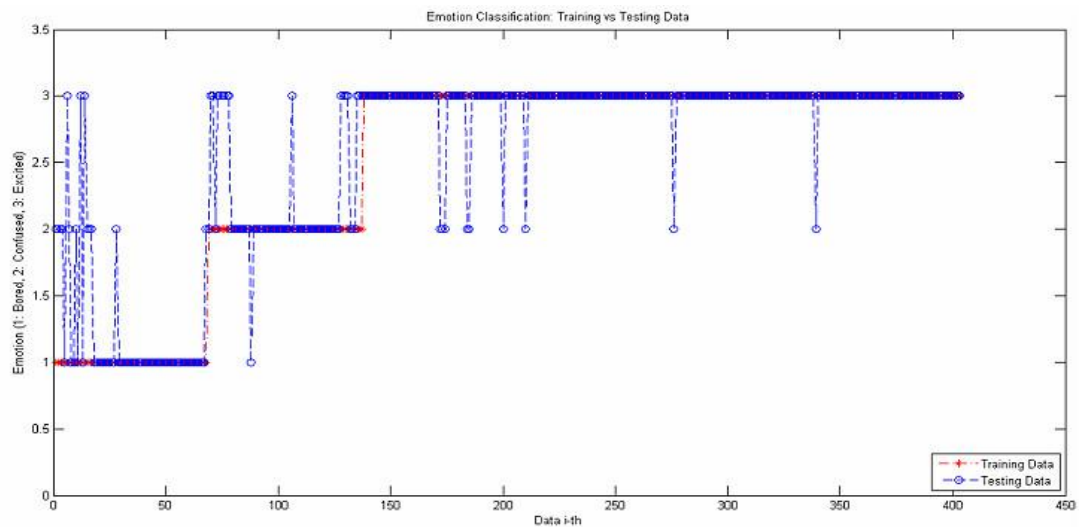
a high variance of the evaluation result. On the contrary, 10-fold cross validation is a holdout method that is repeated for ten times. The variance value of the estimating result will decrease as  $k$ -th value increases. Small variance value of the estimating result will improve the accuracy of Multi-class SVM model.



(a)



(b)



(c)



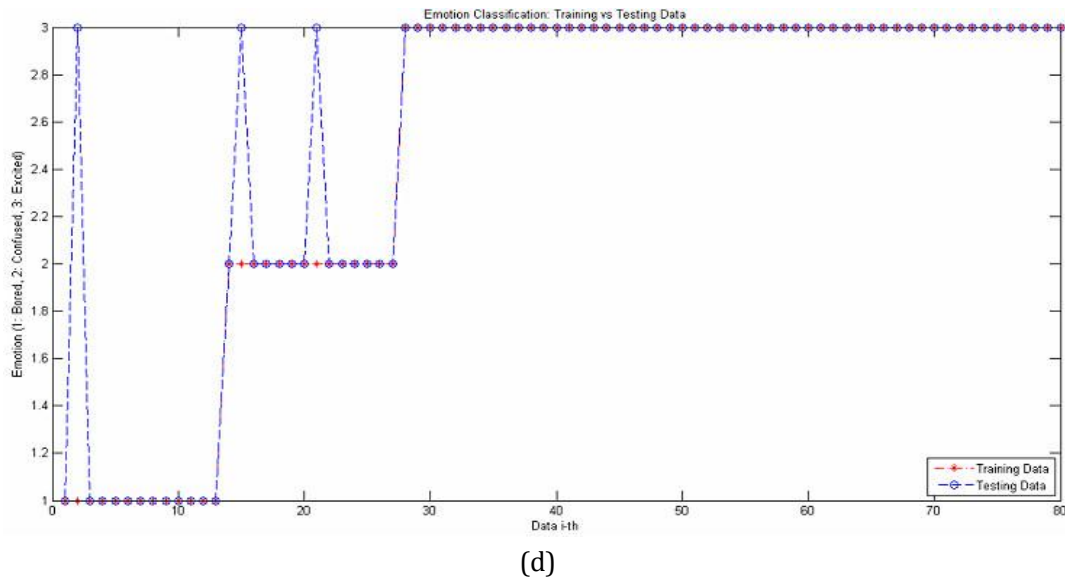


Fig. 3. Correct rate classification using (a) PCA-holdout method (b) PCA-10fold method (c) Gabor-holdout method and (d) Gabor-10fold method.

## 5. Conclusion

This research has a purpose of classifying cognitive-affective emotion states based on facial expression data. PCA and Gabor Filters are utilized to extract features space of facial expression database. Holdout and 10-fold cross validation were chosen to estimate the best parameters for Multi-class SVM. Experiments results suggest that 10-fold cross-validation of Multi-class SVM give a higher accuracy rate compared to Holdout cross validation. 10-fold cross validation gives the best result because it decreases the variance of the estimating result as iteration number increased. However, 10-fold cross validation has a drawback as it takes a longer time to classify compared to Holdout method.

## Acknowledgment

We would like to thank Ministry of Higher Education (DIKTI) Indonesia for giving a funding grant for this research.

## References

- [1] Goleman, D. (2011). The brain and emotional intelligence: New insights. *More than Sound*.
- [2] Ekman, P. (1999). *Handbook of Cognition and Emotion: Basic Emotions*. John Wiley and Sons.
- [3] Damon, W., & Lerner, R. (2006). *Handbook of Child Psychology* (6th ed.). New York: John Wiley and Sons.
- [4] Baker, R. S. J. D., D'Mello, S. K., Rodrigo, M. M. T., & Graesser, A. C. (2010). Better to be frustrated than bored: the incidence, persistence, and impact of learners' cognitive-affective states during interactions with three different computer-based learning environments. *International Journal of Human-Computer Studies*, 68, 223-241.
- [5] Ekman, P., & Friesen, W. V. (1978). *The Facial Action Coding System: A Technique for the Measurement of Facial Movement*. Palo Alto, CA: Consulting Psychologists Press.
- [6] Kumar, D., & Rajni. (2014). An efficient method of PCA based face recognition using Simulink. *International Journal of Advances in Computer Science and Technology*, 3(5), 364-368.
- [7] Dashore, G., & Raj, V. C. (2012). An efficient method for face recognition using principal component analysis (PCA). *International Journal of Advanced Technology and Engineering Research*, 2(2), 23-29.

- [8] Pankaj, D. S., & Wilsy, M. (2013). Comparison of PCA, LDA, and gabor features for face recognition using fuzzy neural network. *Proceedings of 2nd International Conference on Advances in Computing and Information Technology* (pp. 413-422).
- [9] Barbu, T. (2010). Gabor filter-based face recognition technique. *Romanian Academy Series A*, 11(3), 277-283.
- [10] Javed, M. (2014). Face recognition using principle gabor filter. *International Journal of Advanced Research in Computer Science and Software Engineering*, 4(2), 176-180.
- [11] Yang, M., Zhang, L., Shiu, S. C. K., & Zhang, D. (2013). Gabor feature based robust representation and classification for face recognition with Gabor occlusion dictionary. *Pattern Recognition*, 46(7), 1865-1878.
- [12] Samad, R., & Sawada, H. (2011). Edge-based facial features extraction using gabor wavelet and convolution filters. *Proceedings of MVA2011 IAPR Conference on Machine Vision Applications* (pp. 1-4).
- [13] Wang, J., & Chen, J. (2010). Face recognition based on fusion of Gabor and 2DPCA features. *International Signal Processing and Communication Systems (ISPACS)* (pp. 1-4).
- [14] Kumar, C. M., Thiyagarajan, R., Natarajan, S. P., & Arulselvi, S. (2011). Gabor features and LDA based face recognition with ANN classifier. *International Conference Emerging Trends in Electrical and Computer Technology* (pp. 831-836).
- [15] Desai, C., Deva R., & Charless, C. F. (2011). Discriminative models for multi-class object layout. *International Journal of Computer Vision*, 95(1), 1-12.
- [16] Gryllias, K. C., & Antoniadis, I. A. (2012). A support vector machine approach based on physical model training for rolling element bearing fault detection in industrial environments. *Engineering Applications of Artificial Intelligence*, 25(2), 326-344.
- [17] Zhi-Xia, Y., Yuan-Hai, S., & Xiang-Sun, Z. (2013). Multiple birth support vector machine for multi-class classification. *Neural Computing and Applications*, 22(1), 153-161.



**Nova Eka Diana** was born in Lamongan, East Java Province, Indonesia, in 1985. She received the B.S. degree from computer science, University of Indonesia, Indonesia, in 2007 and the M.S. degree from the Korea Institute of Science and Technology (KIST), South Korea, in 2012, in HCI & Robotic Department. She is currently a faculty member and researcher in the Faculty of Information Technology, University YARSI, Indonesia. Her research interests include affective computing, HCI, biomedical engineering and simulation.



**Ahmad Sabiq** was born in Tegal, Central Java, Indonesia, in 1983. He graduated his M.S. degree of computer science, from University of Indonesia, in 2012. Currently, he is a faculty member in the Faculty of Information Technology, University YARSI, Indonesia. His research interests include affective computing, mobile computing and mobile ad-hoc network.