Mental Tasks EEG Signal Classification Using Support Vector Machine

Wahyu Caesarendra[1,a]**, Syahara U. Lekson**[2]** & Muhammad Agung**[3]

[1]*Faculty of Integrated Technologies, Universiti Brunei Darusslam, Brunei Darussalam*
[2]*Mechanical Engineering Departement, Faculty of Engineering, Diponegoro University, Indonesia*
[3]*Instrumentation Laboratory, Indonesian Institute of Science, Indonesia*

Corespondance   : [a]wahyu.caesarendra@ubd.edu.bn

ABSTRACT

This paper presents a result of electroencephalography (EEG) signal classification for mental tasks such as thinking forward, backward, left, and right. The EEG data in this study were recorded from Emotive device with 14 channels and 2 references. The aim of this study is to identify the most sensitive channels to the mental task classification. Prior to feature extraction, the EEG signal were decomposed using wavelet with three level decomposition. Eighteen features were extracted from the processed data. Principal component analysis (PCA) is then used to reduce 18 features into 3 principal components. The principal component were classified using support vector machine (SVM). The results show that the SVM classification accuracy of 75%.

Keywords: facility management, health facility, electroencephalography, principal component analysis, support vector machine

INTRODUCTION

Human brain generates an electrical signal namely electroencephalography (EEG) that related to the level of consciousness. Frequency of EEG signal ranges from 0.5 to 100 Hz. The frequency and amplitude of EEG signal are related to the brain activity, such as concentration level, cognitive ability, and relaxation. Based on the value of its frequency, EEG signal can be categorized into five classes, which are delta, alpha, beta, gamma, and theta.

Delta wave oscillates at 0-4 cycles per second with its amplitude reaches 10 mV; slowest compared to other waves. Delta wave is produced at dreamless sleeping state. Alpha wave oscillates at 8-13 times per second with the amplitude reaches 50 µV. Alpha wave is produced when one is relaxing or daydreaming. At 13-40 cycle per second, Beta wave oscillates faster than alpha wave. Beta wave is activated when one is conscious.

A number of researches have studied and analysed the application of EEG signal processing for the purpose of medical diagnosis, biomedical engineering, and automation. For example, the research by Deon Garrett reported about comparison of linear, nonlinear, and feature selection methods for EEG signal classification. The researchers in this paper measured EEG signal from six electrodes were placed at C3, C4, P3, P4, O1, and O2 as defined by the 10-20 system of electrode placement. The subjects were asked to do a mental task, such as relaxing, numbering, lettering, rotate a three-dimensional solid, and counting The study compared some classifiers with result LDA 4.8%, ANN 52.8%, and SVM 52.3% (Phinyomark, Limsakul & Phukpattaranont 2009). Other studies that focus on the application of EEG signal for brain computer interface (BCI) are presented in (Subas 2010; Garrett et al 2003).

Present study focused on the classification of the EEG signals using wavelet decomposition, feature extraction, feature reduction using PCA and classification using SVM.

The EEG signals were collected from 4 mental tasks experiment such as thinking forward, backward, righ and left.

EEG, as health detector facility, is vital diagnosis infrastructure. EEG must be well manage to be always in fuction and to be designed to improve its performance. Therefore a research, on the EEG facility development, needs to be conducted continously (Soemitro & Suprayitno 2018). By developing a proper EEG diagnosis equipment especially for the wheel chair application, it will optimize the facilty asset management.

Prior to the feature extraction, wavelet decomposition is employed to process the EEG signal. Wavelet decomposition is used for signal analysis prior the Fourier transforms to remove unwanted high frequency band. Wavelet is an effective method to obtain localized information (Tkach, Huang & Kuiken 2010). Wavelet decomposition represented the signal with a certain number of coefficients from wavelet transform.

METHOD

Pre-Processing

EEG is generally a non-stationary signal, the values of raw data EEG may vary over the time. This paper employed wavelet as a pre-processing step to decompose the EEG signal into a linear combination of time-scale units. The decomposition is basically produce two parts: (1) approximation coefficient (A) and detail coefficient (D). Approximation coefficient is a result from EEG signal that passes through the low-pass filter; while detail coefficient is result from EEG signal that passes through the high-pass filter.

Wavelet analysis begins by decomposing raw signal data into its component mutually orthogonal subspaces. The result of Wavelet decomposition with three level decomposition is shown in Figure.1.
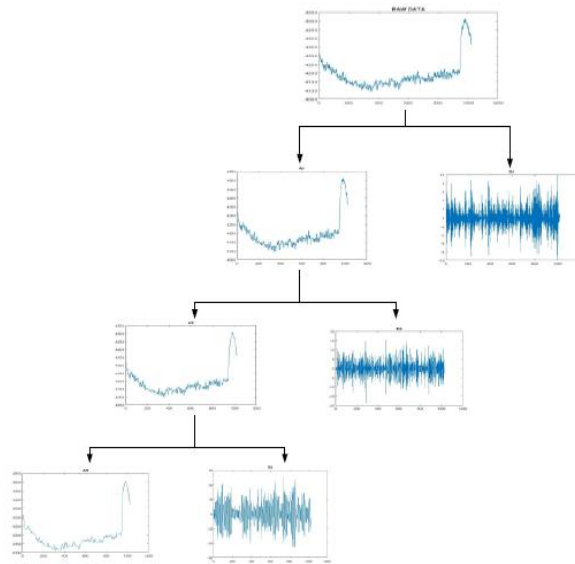

Figure1. Three level wavelet decomposition

Feature extraction

This study employed 18 feature extraction methods. The brief description is presented as follows :

- Average amplitude change (AAC): is nearly equivalent to waveform length (WL) feature, except that wavelength is averaged (Dubechies 1992).

$$AAC = \frac{1}{N}\sum_{i=1}^{N=1}\left|x_{i+1} - x_i\right| \tag{1}$$

- Difference absolute standard deviation value (DASDV): is a standard deviation of the wavelength (Tsai, Yeh & Lo 2008).

$$DASDV = \sqrt{\frac{1}{N-1}\sum_{i=1}^{N-1}\left(x_{i+1} + x_i\right)^2} \tag{2}$$

- Log detector (LOG): like the V-Order (V) feature, this feature estimates the brain contraction force (e.g. contraction force) (Kim et al 2011; Oskoei & Hu 2008).

$$LOG = e^{\frac{1}{N}\sum_{i=1}^{N}\log(|x_i|)} \tag{3}$$

- Mean absolute value (MAV): MAV is the average distance between each data value and the mean

$$MAV = \frac{1}{N}\sum_{i=1}^{N}\left|x_i\right| \tag{4}$$

- Modified mean absolute value type 1 (MAV1): MAV1 is an extension of MAV feature (Boostani & Moradi 2003; Zardoshti-Kermani et al 1995).

$$MAV1 = \frac{1}{N}\sum_{i=1}^{N}w_i\left|x_i\right|$$
$$w_i = \begin{cases} 1, & if\ 0.25N \le i \le 0.75N \\ 0.5, & otherwise \end{cases} \tag{5}$$

- Modified mean absolute value type 2 (MAV2): MAV2 is an expansion of MAV feature which is similar to the MAV1 (Boostani & Moradi 2003; Zardoshti-Kermani et al 1995)..

$$MAV2 = \frac{1}{N}\sum_{i=1}^{N}w_i\left|x_i\right|$$
$$w_i = \begin{cases} 1, & if\ 0.25N \le i \le 0.75N \\ 4i/N, & else\ if\ i < 0.25N \\ 4(i-N)/N, & otherwise \end{cases} \tag{6}$$

- The V-order (V): V is a non-linear detector that implicitly estimates brain contraction force. It is defined from a functional mathematical model of the EEG signal generation [5, 7].

$$V = \left(\frac{1}{N}\sum_{i=1}^{N}x_i^v\right)^{\frac{1}{v}} \tag{7}$$

- Hjorth 2 (mobility): The mobility parameter represents the mean frequency, or the proportion of standard deviation of the power spectrum (Rangayyan 2001).

$$Hjorth\_2 = \sqrt{\frac{\mathrm{var}\left(x\dfrac{dx}{dt}\right)}{\mathrm{var}(x)}} \tag{8}$$

- Hjorth 3 (complexity): The complexity parameter represents the change in frequency. The parameter compares the signal's similarity to a pure sine wave, where the value converges to 1 if the signal is more similar (Rangayyan 2001).

$$Hjorth\_3 = \frac{mobility\left(x\dfrac{dx}{dt}\right)}{mobility(x)} \tag{9}$$

- Skewness (Sk): Skewness measures the asymmetry of probability density function (pdf) of the signal. Similar to kurtosis, skewnes also has been used in vibration as a degradation feature of bearing condition (Park & Lee 1990).

$$Sk = \frac{\sum_{i=1}^{N}(x_i - m)^3}{(N-1)\sigma^3} \tag{10}$$

- Autoregressive (AR) coefficients: An approach for modelling the univariate time series (Tsai, Yeh & Lo 2008; Turnip 2016: Turnip, Soetraprawata & Kusumandari 2013).

$$y_t = a_1 y_{t-1} + a_2 y_{t-2} + \ldots + a_n y_{t-n} + \varepsilon_t = \sum_{i=1}^{n} a_i y_{t-1} + \varepsilon_t \tag{11}$$

- Correlation dimension: A larger correlation dimension corresponds to a larger degree of complexity and less-similarity. The correlation dimension is derived from the correlation integral presented in (Park & Lee 1990).

$$C(l) = \lim_{M \to \infty} \left[\frac{2}{M^2} \sum_{i=1}^{M-k} \sum_{j=i+k}^{M} \Theta\left(l\left|X_i - X_j\right|\right) \right] \tag{12}$$

where X_i, X_j are the position vectors on attractor of the phase-space vector, l is the distance under consideration, $\Theta(x)$ is the Heaviside step function, $\Theta(x) = 0$ if $X \leq 0$, $\Theta(x) = 1$ if $X > 0$, k is the summation offset, M is the number of reconstructed vectors from the original vibration signal, and $C(l)$ is the correlation dimension.

- Fractal dimension: Fractal dimension is used to measure the complexity of signals. Once the phase-space vector is obtained, the mean absolute length between the phase-space vector j^{th} and (j^{th}-1) of $\mathbf{X}$ can be defined as follows: (Park & Lee 1990).

$$
\begin{aligned}
&for\ i = 1:J \\
&\quad for\ j = 2:m \\
&\qquad L_1(i, j\text{-}1) = \left|\mathbf{X}(i, j) - \mathbf{X}(i, (j\text{-}1))\right| \\
&\qquad L_m(i, j\text{-}1) = mean\ (L_1) \\
&\quad end \\
&end
\end{aligned}
\tag{13}
$$

where J represents reconstruction delay and m is embedding dimension. Assuming that L_m is proportional to m^{-D}, then taking the logarithmic natural of $(1/m)$ and L_m. Hence, the fractal dimension, D can be computed using standard least-square fitting method.

- Approximate entropy: The approximate entropy computes the value of regularity in the signal. Smaller value indicates more regular behaviour and higher value of it indicates less regularity on the data set. Similar to three other nonlinear features described above, approximate entropy uses reconstruction vectors (1) as an input. The initial step of approximate entropy is measures the distance $d(X(i), X(j))$ between two vectors $X(i)$ and $X(j)$ which can be defined as the maximum difference in their respective corresponding elements: [13] Park, S.H. & Lee, S.P. (1990).

$$d(X(i), X(j)) = \max_{k=1,2,...,m} (| x(i+k-1) - x(j+k-1) |) \qquad (14)$$

where $i = 1, 2, ..., N-m+1$, $j = 1, 2, ..., N-m+1$, and N is the number of data points contained in the time series. For each vector $X(i)$, a measure that describes the similarity between the vector $X(i)$ and all other vectors $X(j)$, where $j \neq i$ can be constructed as

$$C_i^m(r) = \frac{1}{N-(m-1)} \sum_{j \neq i} \Theta\{ r - d[X(i), X(j)] \} \qquad (15)$$

- where $j = 1, 2, ..., N-m+1$ and the Heaviside step function, $\Theta(x)$ is similar to the symbol in correlation dimension where $\Theta(x) = 0$ if $X \leq 0$, $\Theta(x) = 1$ if $X > 0$. The symbol r in (16) denotes a predetermined tolerance value, defined as

$$r = k.std(\mathbf{Y}) \qquad (16)$$

- where $std(\mathbf{Y})$ means is the standard deviation of original time series or vibration signal, $\mathbf{Y}$ and k is a constant $(k > 0)$. By defining

$$\varphi^m(r) = \frac{1}{N-m+1} \sum_i \ln[C_i^m(r)], \ i = 1,2,...,N \quad m+1 \qquad (17)$$

- The approximate entropy value of time series can be calculated as

$$ApEn(m,r) = \lim_{N \to \infty} [\varphi^m(r) - \varphi^{m+1}(r)] \qquad (18)$$

Feature Reduction

Feature reduction is process to extract data using linier transformation. It is used to determine the best feature that influence of process classification.

In this paper, principal component analysis (PCA) is selected for reduced feature extraction method. In PCA, represent the d-dimensional data in a lower dimensional space.

Data from feature extraction computed by d-dimensional mean factor vector μ and $d \times d$ covariance matrix Σ. Eigenvalue and eigenvector are computed according to decreasing eigenvalue. Eigenvectors e_1 with eigenvalue λ_1, eigenvactor e_2 with λ_2, and so on. The largest k such eigenvector can be determine by looking at a spectrum of eigenvectors. the k eigenvector are columns that consist a $k \times k$ matrix. The feature extraction data represent: (Caesarendra et al 2013).

$$x' = A^t(x - \mu) \tag{19}$$

Let a set of centered data input vectors $\mathbf{x}_t$ ($t = 1, \ldots, l$ and $\sum \mathbf{x}_t = 0$), each of which is of m dimension $\mathbf{x}_t = [x_t(1), x_t(2), \ldots, x_t(m)]^T$ usually $m < l$, $\mathbf{s}_t$ linearly transforms each vector $\mathbf{x}_t$ as in (20)

$$\mathbf{s}_t = \mathbf{U}^T \cdot \mathbf{x}_t \tag{20}$$

Where $\mathbf{U}$ is the $m \times m$ orthogonal matrix whose ith column, $\mathbf{u}_i$ is the eigenvector of the sample covariance matrix $\mathbf{C}$. The $\mathbf{C}$ matrix can be calculated using (21)

$$\mathbf{C} = \frac{1}{l}\sum_{t=1}^{l} \mathbf{x}_t \cdot \mathbf{x}_t^T \tag{21}$$

The eigenvalue problem in PCA can be solved using equation (22)

$$\lambda_i \mathbf{u}_i = \mathbf{C} \cdot \mathbf{u}_i \quad i = 1, \ldots, m \tag{22}$$

where λ_i is one of the eigenvalues of $\mathbf{C}$. The components of $\mathbf{s}_t$ are then calculated as the orthogonal transformations of $\mathbf{x}_t$ based on the estimated $\mathbf{u}_i$

$$\mathbf{s}_t(i) = \mathbf{u}_i^T \mathbf{x}_t \quad i = 1, \ldots, m \tag{23}$$

The new extracted components are called principal components. The number of principal components in $\mathbf{s}_t$ can be reduced using only the first several eigenvectors sorted in descending order of the eigenvalues.

Feature Classification

Kernel function used to build linear boundaries through non-linear transformations or mapping to finding the best classes for decision plane. The SVM select the classes with maximal margin. The SVM is a supervised learning method. It is widely used for classification and regression. SVM applies the input vectors that are non-linearly mapped a very high dimension feature space.

The data input is given the matrix $\mathbf{x}$ where it is consist of element x_i ($i = 1, 2, \ldots, M$), M is the number of samples. It is assumed that there are two classes namely positive class and negative class. The two classes are denoted by $y_i = 1$ for positive class and $y_i = -1$ for negative class, respectively. For linearly data, it is possible to determine the hyper plane function of $f(\mathbf{x}) = 0$ splitting the given data as in (24).

$$f(\mathbf{x}) = \mathbf{w}^T \mathbf{x} + b = \sum_{i=1}^{M} w_i x_i + b = 0 \tag{24}$$

The M-dimensional vector $\mathbf{w}$ and scalar b are used to define the position of separating hyper plane. It is created by decision function of sign $f(\mathbf{x})$ to classify the input data either in positive or negative class. The constraint should be fulfilled by separating hyper plane that can be written in (25)

$$y_i f(\mathbf{x}_i) = y_i(\mathbf{w}^T \mathbf{x}_i + b) \geq 1 \quad \text{for } i = 1,2,\ldots M \tag{25}$$

The optimal separating hyper-plane is the maximum distance between the plane and the nearest data, i.e. the maximum margin created by separating hyper-plane. An example of the optimal hyper-plane of the two data sets can be seen in Figure 2.

A series data points for two different classes are presented in Fig. 2, black circle for positive class and white circle for negative class. The SVM tries to place a linear boundary between the two classes, and orients it in such way that the dash dotted line is maximized. Moreover, SVM tries to orientate the maximum of the distance between boundary and the nearest data point in each class. The boundary is located in the middle of margin between two points. Support vectors are the nearest data points used to define the margin. In Figure 2, support vectors are represented by square black circle and square white circle. In this linear system, the normal vector to the hyper plane is **w** and the perpendicular distance from the hyper plane to the origin is
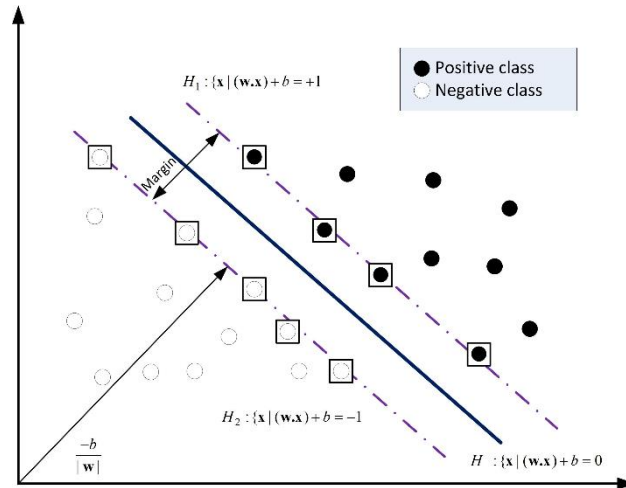
Figure.2 Classification of two linearly separable classes using SVM

The noise with slack variables ξ_i and the error penalty C, the optimal hyperplane separating the data can be calculated using (26) and (27)

$$\text{Minimize } \frac{1}{2}\|\mathbf{w}\|^2 + C\sum_{i=1}^{M}\xi_i \tag{26}$$

$$\text{Subject to } \begin{cases} y_i(\mathbf{w}^T\mathbf{x}_i + b) \geq 1 - \xi_i, i=1,...M \\ \xi_i \geq 0, i=1,...M \end{cases} \tag{27}$$

Where ξ_i is measuring the distance between the margin, the calculation can be simplified into the Lagrangian dual problem as in (28) using Kuhn-Tucker condition.

$$\text{Min } L(\mathbf{w},b,\boldsymbol{\alpha}) = \frac{1}{2}\|\mathbf{w}\|^2 - \sum_{i=1}^{M}\alpha_i y_i(\mathbf{w}.\mathbf{x}_i + b) + \sum_{i=1}^{M}\alpha_i \tag{28}$$

The task is to minimize (26) and (27) with respect to **w** and b. The saddle point at optimal point can be calculated using (29) and (30)

$$\text{Maximize } L(\alpha) = \sum_{i=1}^{M}\alpha_i - \frac{1}{2}\sum_{i,j=0}^{M}\alpha_i\alpha_j y_i\mathbf{x}_i.\mathbf{x}_j \tag{29}$$

$$\text{Subject to } \begin{array}{l} \alpha_i \geq 0, i=1,...,M \\ \sum_{i=1}^{M}\alpha_i y_i = 0 \end{array} \tag{30}$$

Solving the dual optimization problem, the coefficients αi is obtained which is required to express the **w** to solve (26) and (27). The non-linear decision function becomes (31)

$$f(\mathbf{x}) = sign\left(\sum_{i,j=1}^{M} \alpha_i y_i (\mathbf{x}_i \mathbf{x}_j) + b \right) \qquad (31)$$

It is possible to use SVM in non-linear classification with application of kernel functions. The data is mapped onto a high dimensional feature space using nonlinear vector function of $\mathbf{\Phi}(\mathbf{x}) = (\phi_1(\mathbf{x}),...,\phi_i(\mathbf{x}))$. The decision function can be calculated using (32)

$$f(\mathbf{x}) = sign\left(\sum_{i,j=1}^{M} \alpha_i . y_i (\mathbf{\Phi}^T(\mathbf{x}_i).\mathbf{\Phi}(\mathbf{x}_j)) + b \right) \qquad (32)$$

High dimensionality will cause over fitting and computational problem will occur due to large vectors. The problem can be solved by using kernel function $K(\mathbf{x}_i,\mathbf{x}_j) = \left(\mathbf{\Phi}^T(\mathbf{x}_i).\mathbf{\Phi}_j(\mathbf{x}_j) \right)$. The decision function will be

$$f(\mathbf{x}) = sign\left(\sum_{i,j=1}^{M} \alpha_i y_i K(\mathbf{x}_i,\mathbf{x}_j) + b \right) \qquad (33)$$

The SVM utilize different kernel functions such as linear, polynomial and Gaussian RBF. Kernel function defines the feature space, it is important to select the appropriate kernel function. Prior to classifying, the features were trained in SVM to define classes' category. A detail of SVM classification for PCA, ICA, and LDA is presented in (Turnip, Soetraprawata & Kusumandari 2013)..

MATERIALS

In this paper, Emotiv EEG device used to record EEG data. The device has 14 channels which are AF3, AF4, F7, F8, F3, F4, FC5, FC6, T7, T8, P7, P8, O1, O2 and 2 references such as gyro x, gyro y. Ten subjects were notified with the consent form and involved in the experiment. The location of the electrodes measured are shown in Figure.3. Each Subject were asked to do 4 mental tasks such as thinking forward, backward, right, and left. Each mental task activity was recorded for approximately 10 seconds. The experiment of each mental task activity and each subject were repeated five times. The sampling frequency of EEG signal acquistion is 128 Hz.
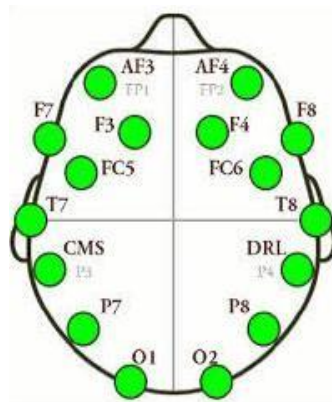

Figure.3 The electrode location of Emotiv device

RESULT AND DISCUSSION

Feature extraction

Features are play an important role in classification accuracy. The previous study investigated 18 features which were used in ANN classifier (Caesarendra et al 2015). The result shows that the better accuracy are channel F7 and F8 with accuracy of 80% and 85%, respectively.

Present study include three nonlinear features to   determine classes of EEG signal based four mental tasks. It has been known that EEG signal contain wide range frequency band and thus it's difficult to be analysed. The wavelet decomposition is used to extract the particular low frequency of EEG signal. The processed signal for feature extraction method were the third level wavelet decomposition (D3). Eighteen features were extracted from the D3 signal. The result of feature extraction for each mental tasks is shown in Figures 4-7. Number 1 to number 18 on x-axis represent the features that were explained in Section 2.2 and the value on y-axis represent the value of features. If we take one example feature extraction result i.e. AAC feature, the value of this features lay on negative except the backward mental task. Although there is a different between each mental task, the different still could not distinguish obviously. Therefore, the PCA method is need it.
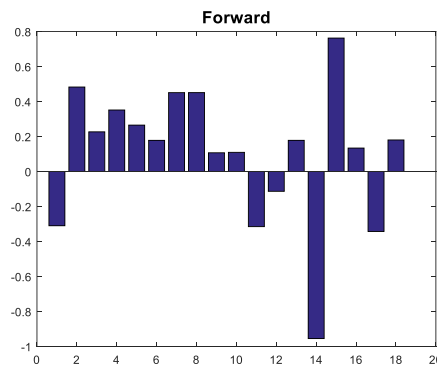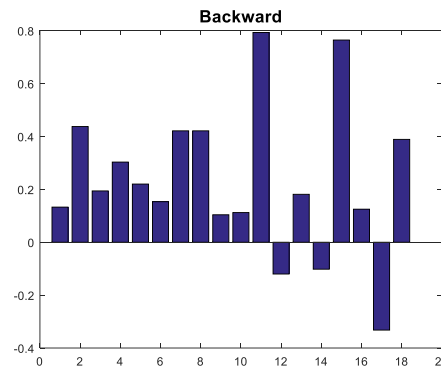

Figure.4 Extracted features of mental task: forward

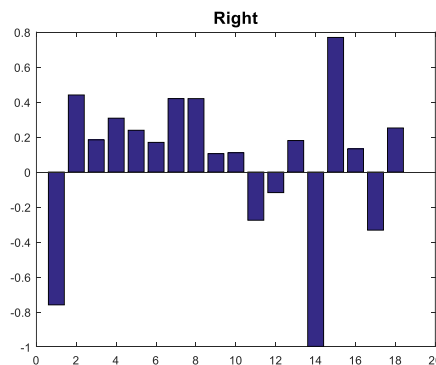**Figure.5** Extracted features of mental task: backward

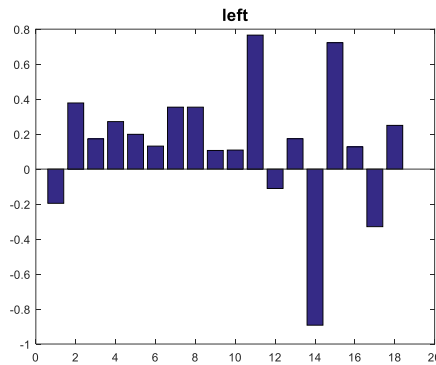**Figure.6** Extracted features of mental task: right

Figure.7 Extracted features of mental task: left

Feature extraction

In this paper, principal component analysis (PCA) is selected to reduce 18 features into 3 features namely principal components. This is necessary to calculate the optimum different between each mental task to build better classification model. From the three features, the combination of two features are studied. It is found that the combination of principal component 2 (PC2) and principal component 3 (PC3) shows that the features of 4 mental tasks can be distinguished. The PCA results for training and testing process are presented in Figure 8 and 9, respectively. It can be seen from the Figure 8 that one feature of mental task "backward" was appeared in the area of mental task "forward". In order to increase the classification accuracy, more data were used in the training process than testing process.
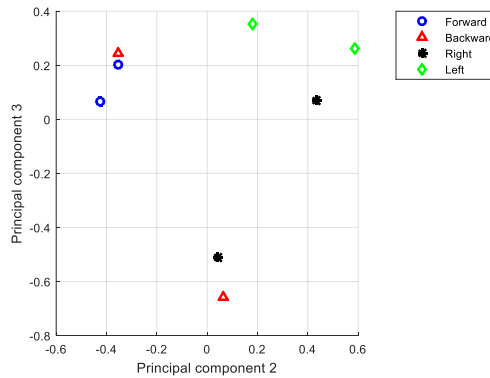

Figure.8 The results of PCA of channel F7 for training process: (a) 3D plot (b) PC2vs PC3 plot
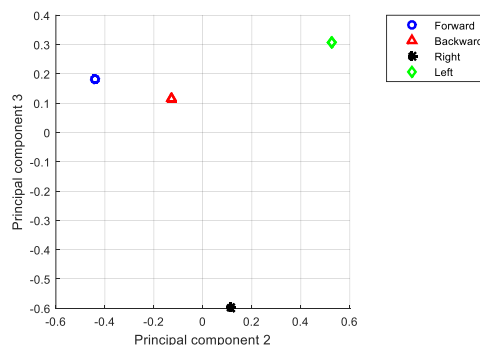

Figure.9 The results of PCA of channel F7 for testing process:

(a) 3D plot (b) PC2vs PC3 plot

Feature Clasification

The results of training and testing classification are presented in Figure 10. The plot combination of PC2 and PC3 was selected based on the distance between each feature. Self-minimum optimization SVM (SMO SVM) with kernel function is used to train and test these pairs. It can be seen from Figure 10(a) that there is overlapping between area of each mental task. For example, one feature of backward (red triangle) is lay on the forward area (blue circle). This overlapping phenomenon can reduce the classification accuracy.
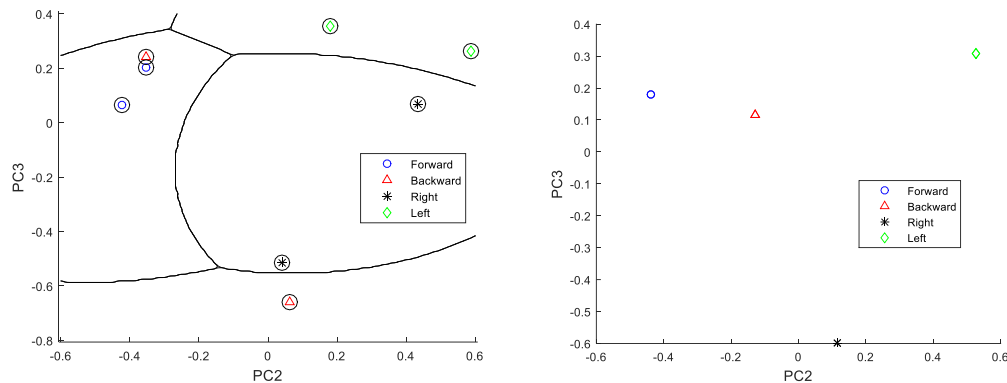

Figure.10 SVM classification for PCA: (a) Training; (b) Testing

CONCLUSION

SVM can determine decision from data testing based on data training. The accuracy of EEG signal for four mental tasks from SVM classifier is 75%. The result from SVM classifier known that EEG signal for thinking forward and backward have bad data because the data not match to decision.

From the methods to EEG classifying known that EEG signal are generally difficult to multi classifying. EEG signal should be pre-processed such as wavelet prior to classifying, is a practical and useful way to improve the accuracy. In this paper known that determination of feature calculation can be affect accuracy of classifying.

REFERENCES

Boostani, R. & Moradi, M.H. (2003). "Evaluation of the forearm EMG signal features for the control of a prosthetic hand". *Physiol. Meas., vol.24(2), pp. 309–319, March 2003*.

Caesarendra, W., Kosasih, B., Tieu, K. & Moodie, C.A.S. (2013). "An application of nonlinear feature extraction – A case study for low speed slewing bearing condition monitoring and prognosis". *Proc. of IEEE/ASME International Conf. On Advanved Intelligent Mechatronics, Wollongong, Australia, pp.1713-1718, 2013*.

Caesarendra, W., Ariyanto, M., Lekson, S.U. & Pasmanasari, E.D. (2015). "EEG based Pattern Recognition Method for Classification of Four Mental Tasking". *Proc. of IEEE International Confrence on Automation, Cognitive Science, Optics, Micro Electro Mechanical System, and Information Technology 2015*, Bandung, Indonesia.

Dubechies, I. (1992). "Ten lectures of wavelets". *Philadelphia: the Society for Industrial and Applied Mathematics 1992, pp. 7-10*.

Garrett, D., Peterson, D.A., Anderson, C.W. & Thaut, M.H. (2003). "Comparison of Linear, Nonlinear, and Feature Selection Methods for EEG Signal Classification", IEEE Transactions on Neural Systems and Rehabilitation Engineering, vol. 11, no. 2, pp 141-144 June 2003.

Kim, K.S., Choi, H.H., Moon, C.S. & Mun, C.W. (2011). "Comparison of k nearest neighbor, quadratic discriminant and linear discriminant analysis in classification of electromyogram signals based on the wrist-motion directions". *Current Applied Physics, vol.11(3), pp. 740– 745, November 2011.*

Oskoei, M.A. & Hu, H. (2008). "Support vector machine based classification scheme for myoelectric control applied to upper limb". *IEEE Trans..Biomed. Eng., vol. 55(8), pp. 1956–1965, August 2008.*

Park, S.H. & Lee, S.P. (1990). "EMG pattern recognition based on artificial intelligence techniques". *IEEE Trans. Rehabil. Eng., vol. 6, no.4, pp. 400– 405, 1998.*

Phinyomark, A., Limsakul, C. & Phukpattaranont, P. (2009). "A novel feature extraction for robust EMG pattern recognition". *J. Comput, vol. 1(1), pp.71–80, Desember 2009.*

Rangayyan, R.M. (2001). "Biomedical Signal Analysis : Case-Study Approach". *IEEE Press series on Biomedical Engineering, Wiley , New York, 2002.*

Soemitro, R.A.A. & Suprayitno, H. (2018). "Pemikiran Awal tentang Konsep Dasar Manajemen Aset Fasilitas". *Jurnal Manajemen Aset Infrastruktur & Fasilitas, Vol. 2, Sup. 1, Juni 2018, Hal. : 1-13.*

Subas, A. (2010). "EEG signal classification using PCA, ICA, LDA and support vector machines". *Expert Systems with Application, December 2010.*

Tkach, D., Huang, H. & Kuiken, T. (2010). "A Study of stability of time-domain features for electromyographic pattern recognition" *J. Neuroeng. Rehabil., vol. 7(21), 2010.* https://emotiv.com/forum/forum4/topic1232/messages/ Accessed on March 2, 2016.

Tsai, K.H., Yeh, C.Y. & Lo H.C. (2008). "A novel design and clinical evaluation of a wheelchair for stroke patients". *Int. J. Ind. Ergon., vol 38(3), pp. 264-71, April 2008.*

Turnip, A., Soetraprawata, D., Turnip, M., and Joelianto, E. (2016). "EEG-Based Brain-Controlled Wheelchair with Four Different Stimuli Frequencies". *Internetworking Indonesia Journal, vol.8, no. 1, 2016.*

Turnip, A., Soetraprawata, D., and Kusumandari, D. E. (2013). "A Comparison of EEG Processing Methods to Improve the Performance of BCI". *International Journal of Signal Processing Systems, vol. 1, No. 1, 2013.*

Zardoshti-Kermani, M., Wheeler, B.C., Badie, K. & Hashemi, R.M. (1995). "EMG feature evaluation for movement control of upper extren mity prostheses". *IEEE Trans. Rehabil. Eng.., vol. 3, pp. 324–333, Desember 1995.*