Brain signal recognition system based on One-Against-One Multiclass Support Vector Machines: a comparison with Multiclass Neural Network

Winda Astuti1,4, Ari Fadli2, Sofyan Tan1 and Rini Akmeliawati3
1 Computer Engineering Department, Faculty of Engineering, Bina Nusantara University, Jakarta, Indonesia
2 Electrical Engineering Department, Universitas Jenderal Soedirman, Indonesia.
3 School of Mechanical Engineering, The University of Adelaide, Adelaide, South Australia, 5005 Australia
E-mail: wastuti@binus.edu

Abstract. The human mind recognition system is a complex state of mind which can be connected to trigger an external reaction based on internal stimuli. This paper investigates the possibility of recognizing brain signal using signal processing and intelligent technique. The system is differentiate four difference mind thinking, namely moving left hand, moving right hand, moving foot and moving tongue, respectively. The Linear Prediction Coding (LPC) is used to extract the signal, in which the extracted signal is later used as input to the Support Vector Machine (SVM) based identifier. The brain signal parameters are compared and classified to identify the human actions that are intended to be perform. The results of the computer simulation show that this technique produces better accuracy to that of the existing technique based on artificial neural network (ANN), and it requires less training time of one second.

1. Introduction
Recently, the development of neural technology has led to neurotechnology, this technology will connect the computer system and the activity of the human brain which is then called the Brain Computer Interface (BCI). BCI is able to understand the desires of its users by reading brain activity through various methods such as electroencephalography (EEG) [1]. The term "Electroencephalography" (EEG) is a way to measure the nerve activity of the human brain in the form of a recording of electrical current along the scalp produced from neurons in the human brain. [2].

There are three main aspects in BCI, namely data acquisition, noise acquisition and classification. In the aspect of data acquisition, the signal of the human brain is extracted in order to model the feature of the raw signal. There are some feature which applied to extract the brain signal, namely Continues Wavelet Transform, Fast Fourier Transform (FFT), Time Frequency Distribution and Autoregressive Distribution [3]. The release technique is then applied to eliminate noise that is present in the signals of the human brain [4]. Finally the classification algorithm is used to recognize patterns of brain signals. Classification is the most difficult task because each human-being does not produce the same brain signal. There are different types of human brain signal which led different features of
the signal, namely Beta (25Hz-40Hz), Alpha (8Hz-12Hz), Theta (4Hz-8Hz) and Delta (0.5Hz-4Hz), which have oscillated during wakefulness, Visual, deep meditation and deep sleep, respectively [5].

Furthermore, a classification system is the ability to recognize the brain signals properly, in order to translate the signal into useful information. The brain activity (EEG signals) are one of the biological systems in the human body that can be utilized [6]. The use of brain activity has been widely used including to communicate with other people or to control machines, and robots [5][7][8].

Based on the previous study, in this work, An automatic identification of brain activity pattern is developed. The system uses LPC for feature extraction techniques. One-Against-One Multi-Class Support Vector Machines and Multiclass Neural network algorithm as applics as recognition system of the brain activity patterns. The significance of this work is observed the better intelligent identification system suitable for brain recognition system. The effectiveness of the proposed system based on Multiclass SVM is evaluated experimentally and compared with the existing method, neural network to identify four difference brain activity.

2. Previous Research
An automatic brain signal activity recognition system has been developed to activate lights. By using Linear Predictive Coding (LPC) as a feature extraction method and Support Vector Machines as a method for recognition of brain signal activity. The accuracy of the training and the accuracy of the proposed system testing are 100% and 80%, respectively [9]. Other experiments that recognize brain activity in animals can be used to control the movement of computer cursors [10][11] or one-dimensional to three-dimensional movements of simple and elaborate robot arms [12][13]. In, Gao and Wang work record EEG signals in the case of right or left hand movements. By using artificial neural network classifiers they are able to classify the EEG signal with an accuracy of 80%. [14].

On other works, In spatial patterns of (event-related desynchronization) ERD on mu rhythms along the sensory-motor cortex, Gao and Wang, resulting classification accuracy for online and offline tests was 79.48% and 85.00%, respectively. [15]. In other studies on the EEG-based BCI system to control prosthetic hand prostheses based on left and right hand movement signals, an accuracy of 90% [16].

3. Proposed System
This paper compares the performance of two recognition engines for brain signal recognition systems. the first is SVM by using the Radial base kernel function. the second is artificial neural networks (ANN) with backpropagation learning algorithm. A block diagram of a conventional brain signal recognition system is shown in Figure 1. This system is trained to recognize a person’s specific thoughts through an electroencephalogram (EEG). Brain signals are digitized and some of them are done to create templates for thought patterns and are stored in memory.

![Figure 1. Block diagram of conventional speaker identification](image-url)

The overall model of the proposed system for SVM-based brain signal recognition system is depicted in Figure 2. A digitized brain signal passed through the processing stages from which Linear Prediction Coding (LPC) features are extracted and passed through the SVM learning algorithm for
both training and testing. In Figure 2, the speaker identification system is separated into two main functions.

1. Feature Extraction.
   Feature Extraction is used to perform feature extraction on brain signal vectors that will be used as input to the system. Some techniques that can be used are fundamental frequency and frequency formant, cepstral coefficient, fundamental frequency, and frequency formant. In this study, linear prediction coefficients are used to extract brain signals.

2. Pattern Matching.
   Pattern matching is a method used to measure the similarity between unknown feature vectors and feature vectors stored in system memory. The pattern matching algorithm compares the incoming brain with the reference model and the difference score is then referred to as distance. Distance is used to determine unknown patterns. There are many types of distance measurement, such as Exact congruence matching methods, and bottle neck matching[17]. The exact congruence matching method, applies of given two patterns A and B, the transformation g in G under which the image g(A) equals B, the distance matrix c, can be define as:

   \[c(A, B) = \begin{cases} 
   0 & \text{if } A = B \\
   1 & \text{otherwise}
   \end{cases}\] (1)

   Bottle neck matching, on the other hand, Let A and B be finite subsets of a space X with metric \(\rho\). Assume that A and B have the same cardinality. Let \(F(A,B)\) be the set of all bijections from A onto B. Then, the \(\rho\) based bottleneck distance \(b_\rho\) is defined as:

   \[b_\rho(A, B) = \min_{f \in F(A,B)} \max_{a \in A} \rho(f(a), a)\] (2)

   There are two types of models: stochastic models and template models. In stochastic models, the pattern matching is probabilistic result such as Hidden Markov Model (HMM), ANN, and SVM. For template models, the pattern matching is deterministic such as dynamic time warping (DTW) with K-nearest neighbors (KNN).

![Figure 2. Overall proposed speaker identification model](image)

3.1. Feature extraction of the brain signal
   The LPC method, which calculates a power spectrum of the signal, is plied in this work. It is used for formant analysis [18]. Linear Prediction Method is a parametric model that is calculated based on the theory of the least squares error. With this method, the speech signal is approximated as a linear combination of the parametric coefficients of the previous sample. In this technique, the LPC coefficient obtained illustrates the formant. [19]

   In this process, the modelling of the processed signal is applied. The autocorrelation analysis is applied to the windowing signal, this step is used to extract the important model properties of the brain
signal. Next step is LPC computation, this step converts auto coefficients into the LP coefficient. This problem can solved using recursive method. The most popular and well known of these recursive methods is the Levinson-Durbin algorithm (3). Finally, the cepstral coefficient provide better alternative to the LP coefficients for speech and speaker recognition. This step can be deriving either from LP analysis or Mel Filter-bank analysis [20].

In this process, the linear prediction analysis is applied to the windowed signal so as to extract the important harmonics and formant properties from the speech. LPA involves the computational of the autocorrelation function from the windowed speech; an autoregressive (AR) model is then fitted to the correlation function. An iterative procedure, termed Levinson algorithm is used to determine both the LPC and the AR order, from which the spectral estimates are computed. Taking the logarithm of the estimated spectral would provide the cepstral coefficient of the speech signal.

The main idea behind linear prediction is to extract the vocal tract parameters. Given a speech samples at time n, \( s(n) \) can be modeled as a linear combination of the past p speech samples, such that the equation (3):

\[
\hat{s}(a : n) = \sum_{k=1}^{p} s(n-k).a_p(k)
\]

where \( a_p = (a_p(1),a_p(2),...,a_p(p)) \) are unknown LPC coefficients use the equation (4), Summing the real and predicted samples we get the following signal :

\[
e(a : n) = s(n) + \hat{s}(n) = s(n) + \sum_{k=1}^{p} a_p(k).s(n-k)
\]

Apply to this signal the z - transform: \( R(z) = S(z)A(z) \). The filter is

\[
A(z)=1+\sum_{k=1}^{p} a_p(k)z^{-k}
\]

To find the vocal tract filter \( \theta (z) \) , we must first find the LPC coefficients \( p \) . By this aim, the following function is minimized

\[
\varepsilon_p(a) = \sum_{n=1}^{M} e(a : n)^2 \rightarrow \min
\]

where \( M \) is a number of frames. The solution of (6) can be obtained by using (7) as follows.

\[
\frac{\partial \varepsilon_p(a)}{\partial a_p(k)} = \frac{\partial}{\partial a_p(k)} \sum_{n=1}^{M} e(a : n)^2 = 2 \sum_{n=1}^{M} e(a : n) \frac{\partial}{\partial a_p(k)} [s(n) + \sum_{l=1}^{p} a_p(l)s(n-l)]
\]

\[
= 2 \sum_{n=1}^{M} e(a : n) s(n-l) = 0 \rightarrow k=1,2,...,p
\]

Thus, we get (8)
\[
\sum_{n=1}^{M} \left[ s(n) + \sum_{l=1}^{P} a_p(l) s(n-l) \right] s(n-k) = 0
\]  
(8)

Denote \( r_x(k) = \sum_{n=1}^{M} s(n).s(n-k) \). Consequently we can write (9) in the following form
\[
r_x(k) = \sum_{l=1}^{P} a_p(l).r_x(l-k) = 0 \quad \text{or} \quad -r_x(k) = \sum_{l=1}^{P} a_p(l).r_x(k-l)
\]  
\[\rightarrow k = 1, \ldots, p\]  
(9)

Consequently, can rewrite (9) as in the form of (10). Equation (8) is called the normal equation or the Yule-Walker equation. Using the expression (5) in the functional (6), we get the equation (10):
\[
q_p = \sum_{n=1}^{M} e(a;n)^2 = \sum_{n=1}^{M} e(a;n)e(a;n) = \sum_{n=1}^{M} \left[ s(n) + \sum_{k=1}^{P} a_p(k) s(n-k) \right]
\]  
\[= \sum_{n=1}^{M} e(a;n) s(n) + \sum_{k=1}^{P} a_p(k) \sum_{n=1}^{M} e(a;n)s(n-k)\]  
(10)

While \( \sum_{n=1}^{M} e(a:n).s(n-k) = 0 \) we can further rewrite (10) in the following form
\[
\epsilon_{p, \min} (a) = e_p(a) = \sum_{n=1}^{M} e(a:n)s(n)
\]  
\[= \sum_{n=1}^{M} \left[ s(n) + \sum_{k=1}^{P} a_p(k) s(n-k) \right] s(n)\]
\[= r_x(0) + \sum_{l=1}^{P} a_p(l) r_x(k)\]  
(11)

The coefficients \( a_p(k) \), which will minimize (11) are determined by using following Levinson-Durbin recursion[21].

1. \( a_0(0) = 1 \quad E_0 = r_x(0) \)
2. For \( j = 0,1, \ldots, p-1 \) calculated the following expressions:
   \[\gamma_{j} = r_{x}\left(j+1) + \frac{1}{a_{j}} a_{j}(i) r_{x}\left(j-i+1) \right.\]
   a. \( r_{j} = -\gamma_{j} / E_{j} \)
   b. \( i=1,2, \ldots, j \)
   c. \( a_{j+1}(i) = a_{j}(i) + \Gamma a_{j}(j-i+1) \)
   d. \( a_{j+1}(j+1) = \Gamma a_{j+1} \)
   e. \( E_{j+1} = E_{j} \left[1 - \left| \Gamma_{j+1} \right|^{2} \right] \)
3.2. Support Vector Machines based Identification

SVM is a relatively new learning machine technique, which is based on statistical learning theory. This method is mathematically simple and avoid over-fitting. The basic idea of SVM is the mapping of non-linearly training data into higher-dimensional feature space through the kernel function.

\[
\begin{align*}
3.2. \text{Support Vector Machines based Identification} \\
\text{SVM is a relatively new learning machine technique, which is based on statistical learning theory. This method is mathematically simple and avoid over-fitting. The basic idea of SVM is the mapping of non-linearly training data into higher-dimensional feature space through the kernel function.}
\end{align*}
\]

\[
\begin{align*}
\text{SVM with Linear separable data.}
\end{align*}
\]

There are three important aspects of SVM, namely discrimination (optimal) hyperplane, optimization via Lagrange multipliers and kernel function. The first two aspects are derived from linear separable form of classification (with no possibility of misclassification), whereas the third aspect deals with non-separable data classification, which may introduce any misclassification data. Fig. 3 shows linearly-separable data with no possibility of miss-classification data.

In this case the SVM training always seeks a global optimizes solution and avoids over-fitting, so it has ability to deal with a large number of feature. Thus, in the linear separable case, there exists a separating hyperplane whose function is:

\[
\begin{align*}
w \cdot x + b = 0 \quad w \in \mathbb{R}^N, b \in \mathbb{R}
\end{align*}
\]

For optimized linear division a hyperplane is constructed to separate the two classes. which implies

\[
\begin{align*}
y_i(w \cdot x + b = 0) \geq 1, i = 1, ..., N
\end{align*}
\]

By minimizing \(|w|\) subject to this constrain, the SVM approach tries to find a unique separating hyperplane. Here \(|w|\) is the Euclidean norm of \(w\), and the distance between the hyperplane and the nearest data points of each class is \(1/|w|\). By introducing Lagrange multiplier \(\alpha_i\), the SVM training procedure amounts to solving a convex quadratic problem (QP). The solution is a unique globally optimization result, which has following properties

\[
\begin{align*}
w = \sum \alpha_i y_i x_i
\end{align*}
\]

provided \(\alpha_i\) is not equal to zero, \(x_i\) is called the support vectors. When SVM is trained, a decision can be obtained by comparing each new example \(x\) with only the support vector \(\{x_i\}, i \in SV\);

\[
\begin{align*}
y = \text{sign}(\sum_{i \in SV} \alpha_i y_i (x \cdot x_i^T) + b)
\end{align*}
\]

In many practical situations, the non-linear separable case is often considered incorporation to linear separable case. Fig. 5 shows non-linear separable data with introduced miss-classification data[22]
In the case of a non-linear separable problem, SVM performs a non-linear mapping of the input vector $\mathbf{x}$ from the input space into a higher dimensional feature space, where the mapping data is determined by kernel function. Typical kernel functions often used in data classification are listed in Table 1.

$$K(\mathbf{x}, \mathbf{x}_j) = \begin{cases} 
\mathbf{x}^T \cdot \mathbf{x}_j, & \text{Linear} \\
(\mathbf{x}^T \cdot \mathbf{x}_j + 1)^d, & \text{Polynomial} \\
\exp(-\|\mathbf{x} - \mathbf{x}_j\|^2/2\sigma^2), & \text{Gaussian (RBF)} 
\end{cases}$$

The choice of the kernel function depends on the data. Different kernel function can be selected to obtain an optimal classification result. The choice of the degree in polynomial kernel and choice of $\sigma$ in Gaussian RBF kernel also depend on the data.

### 3.3. Multi-class Classification SVM

SVM was originally designed for binary classification. However, it can be extended to multi-class classification [23]. The extension of the SVM to solve the problem containing more than two classes is important. There are also some advantages applying multi-class classification as this would not only improve the error rate for the learning and testing process but it also produces faster adaptation and classification times. Many methods have been proposed [20] To solve multi-classification problems, a one-on-one procedure is used in this paper. This method is based on the problem of binary classification. The one-on-one multi-class classification builds $k (k-2) / 2$ classifiers in which each is trained on a combination of two classification data. For training data from the $i$-th and $j$-class, it is used to complete using binary classification. To test the data using a strategy if the results with binary classification decide that $\mathbf{x}$ is in that class, then select class $i$ increase by one. If not, class $j$ increases by one. Decisions are based on majority votes, this method is called 'Max win'. In case the two classes have the same vote, the decision will go to the smallest index.

In this proposed system, each of the extracted brain signal activity is combine one and others. The class is modify into binary classification, the number of the combination is based on $2^n$, where $n$ is the number of the brain activity to be classify. By ignore the same combination, the four class of extracted data have five modification. The identify brain signal actifity is based the max win of the testing data.
3.4. Artificial Neural Network (ANN) Approach

Artificial Neural networks algorithm is used to simulate same or all the characteristic of the biological neurons that form the structural constituents of the brain. Neural network can be trained to perform a particular function by updating the value of connections (weight) between elements [24]. Commonly neural networks are trained, so that particular input leads to a specific target output.

There are many forms of neural networks, they are; Multi-layer Perceptron (MLP), Radial Basis Function (RBF), and Learning vector quantizer (LVQ). The Multi-layer perceptron (MLP) network, as shown in Figure 5, architecture using a backpropagation learning algorithm is one of the most popular neural networks. The Multi-layer perceptron consists of three layers of neural networks; input layer, hidden layer and output layer [25]. The input layer is the non-functional layer that is responsible for sending input to all neurons in the hidden layer. The remaining layers are functional with weighted inputs and non-linear functions (activation). Each neuron in the output layer is directly related to class. Inputs are entered into MLP through input neurons and each output neuron contains posterior probabilities generated for a particular class. The input is then classified into the class whose neuron output accordingly has the highest score.

![Figure 5. MLP neural network](image)

Artificial Neural Network (ANN) based classifier is used in both text-dependent and text-independent speaker identification and speaker verification systems. The ANN is extremely efficient at learning complex mappings between inputs and output. The structure of the backpropagation neural network is shown in Figure 8. A back propagation is a supervised learning algorithm that calculates the change of weight in the network.

The MLP is trained using backpropagation algorithm [13]. The backpropagation algorithm is iterative in the nature, where the weights of the MLP are refined each iteration. Initially the weights of the MLP are set randomly in the range of -0.5 to 0.5. The algorithm performs operates in two phases, the forward and the backward phases. the forward phase passes all the training vectors and their corresponding labels are presented to the MLP and an overall error is found. The output of the $k^{th}$ output neuron, given the $n^{th}$ training vector ($y_n$), is

$$O_k(n) = f\left(\sum_{j} w_{kj} f\left(\sum_{i} w_{ij}y_n(i)\right)\right)$$  \hspace{1cm} (16)

where the summation $j$ is over all input neurons and summation $i$ is over all hidden neurons. Moreover, $wij$ is the weight associated with the connection of the $i^{th}$ neuron of given layer with the $j^{th}$ neuron of the proceeding layer. The activation function is a sigmoid-like function

$$f(x) = \frac{1}{1+e^{-x}}$$  \hspace{1cm} (17)
\[ E = \sum_n \sum_k \left( O_k(n) - d_k(n) \right)^2 \] (18)

where \( d_k(n) \) represent the desire output of the \( k \)th output neuron given the \( n \)th training vector and \( E \) is denote mean square error.

In the backward phase the weight are modified to minimize the error, starting at the output layer and following into to the hidden layer. The process of calculating \( E \) in the forward phase and refining the weights converge to a local minimum error.

In testing, each of the \( L \) feature vector \( Y = (y_1, y_2, \ldots, y_L) \) is given as input to the MLP. For the \( n \)th input feature vector \( (y_n) \), the corresponding output the neural network \( O_k(n) \) is generated, this is directly related to the posterior probability, \( P(\lambda|y_n) \). Considering that the priority of each speaker is same, and then the probability of sequence \( Y \) for a given model \( \lambda \) is

\[ P(Y|\lambda) = \prod P(\lambda|y_n) \] (19)

3.5. Multi-class Classification ANN

ANN was originally designed for binary classification. However, it can be extended to multi-class classification [7]. The extension of the ANN to solve the problem containing more than two classes is important. There are also some advantages applying multi-class classification as this would not only improve the error rate for the learning and testing process but it also produces faster adaptation and classification times [8]. Many methods have been proposed [25] to solve multi-classification problem, however one-against-one procedure is used in this paper. This method is based on binary classification problem.

One-against-one multi-class classification constructs \( k(k-2)/2 \) classifiers where each one is trained on the data from a combination of two classifications. For training data from the \( i \)th and \( j \)th classes, this is to solve using binary classification. For testing data the following strategy can be used; if the result by binary classification decided that \( x \) is in the class \( i \), then vote for class \( i \) is incremented by one. Otherwise, the \( j \)th class is increased by one. The decision is based on the largest vote, this method called ‘Max win’. In the case two classes have same vote, the decision will go to the smallest index.

4. Experimental Result

In this work, the data used are provided by the Laboratory of Brain-Computer Interfaces (BCI-Lab), Graz University of Technology, (Gert Pfurtscheller, Alois Schlögl) in Austria. The data consist of four different brain signals; left hand, right hand, foot and tongue [26].

In order to evaluate the effectiveness of the proposed brain signal recognition system, the brain signal data is applied. The database consists of four different signals, which are left hand, right hand, foot and tongue, which illustrated in Figure 6, Figure 7, Figure 8 and Figure 9, respectively.

In order to obtain the LPC coefficients, the 30 Cepstral coefficients are used. These filters are simulated by integrating the FFT spectrum of 20-ms Hamming-window speech segments in which the frame rate is 10-ms. Figure 10, Figure 11, Figure 12, and Figure 13 shown the 30 LPC extracted from the brain signal, respectively. The signal is then extracted by applied LPC method, which illustrated in, respectively.
4.1. Training and classification
In order to evaluate the effectiveness of the proposed system, four extracted brain signal activity are evaluated. Each of the signal is modeled based on LPC technique into 30 coefficient LPC parameters. The training data involve 45 brain signal per activity. Therefore, the training data involve 180 signals.
for all four brain signal activity. The other brain signal activity, 15 signals of each of the brain activity class are applied as testing data.

The SVMs-based multi-class classification is applied to perform classification process using one-against-one method. Here, RBF kernel function is used. Table 2 shows that during the training stage, SVM-based speaker identification system gives results with 100% classification rate. The SVM-based system identification can perfectly classify the people based on their mind. For testing the effectiveness of the proposed system, further experiment is carried out using another brain signal data. As shown in Table 2, the proposed system gives classification rate of 100%.

Based on the result shown in Table 2, it is clear that the proposed SVM-based brain signal recognition system can achieve classification rate better well known ANN-based brain signal recognition system. In addition, this good performance can be achieved in a shorter time as compared ANN-system.

The ANN based multilayer perceptron (MLP) network with the backpropagation learning algorithm has been used in this paper and it contained of a hidden layer. The numbers of neuron in the input, hidden and output layers are 3,4,5 respectively. As shown in Table 2, the ANN gives 75% classification in training stages. This means the ANN has not perfectly classifies all of the speakers. Furthermore, the results of analysis with another brain data testing stage produce 50% classification rate which is similar to the result obtained with ANN.

| Speaker Classification | SVMs | ANN | SVMs | ANN | SVMs | ANN |
|------------------------|------|-----|------|-----|------|-----|
| Right hand             | 100  | 100 | 100  | 0   | 1    | 2.53|
| Left hand              | 100  | 100 | 100  | 100 | 1    | 2.53|
| Foot                   | 100  | 100 | 100  | 0   | 1    | 2.53|
| Tongue                 | 100  | 100 | 0    | 100 | 1    | 2.53|
| Avr. Err               | 100  | 100 | 75   | 50  |

The testing classification rate results have shown that the performance of SVM based pattern matching technique has better performance compared to Neural Network in term of error correction. Moreover, the training time of SVM is less than that of neural network.

5. Conclusion
SVM-based brain recognition system has been proposed and discussed in this paper. This technique compared with the existing method based on ANN. The accuracy of the SVM based brain signal recognition was found better than the existing method based on ANN. It was also clear, from our computer simulation results, that the proposed method requires much less training time than the existing one. These results are four different thinking of the brain in the system considered. In the future, it is needed to clarify the performance and complexity with increasing number of human thinking to prove its usefulness in real system.

6. References
[1] D. Buvaneash and M. R. Stalin John, “Brain robot interface using artificial neural network,” IOP Conf. Ser. Mater. Sci. Eng., vol. 402, no. 1, 2018.
[2] E. Niedermeyer and F. H. L. da Silva, “Electroencephalography: Basic Principles, Clinical Applications, and Related Fields,” vol. Lippincott, pp. 24–28, 2005.
[3] A. S. Al-fahoum and A. A. Al-fraihat, “Methods of EEG Signal Features Extraction Using Linear Analysis in Frequency and Time-Frequency Domains,” ISRN Neurosci., vol. 2014, p. 7, 2014.

[4] and H. S. W. Lutzenberd, T. Elbert, N. Birbaumer, W.J. Ray, “The Scalp Distribution of The Fractal Dimension of the EEG and its Variations with Mental Task,” vol. 5, no. 1992, 2006.

[5] J. P. Donoghue, “Connecting cortex to machines: Recent advances in brain interfaces,” Nat. Neurosci., vol. 5, no. 11s, pp. 1085–1088, 2002.

[6] M. Zangeneh Soroush, K. Maghooli, S. K. Setarehdan, and A. Motie Nasrabadi, “A Review on EEG Signals Based Emotion Recognition,” Int. Clin. Neurosci. J., vol. 4, no. 4, pp. 118–129, 2017.

[7] J. E. Huggins et al., “Electrocorticogram as the basis for a direct brain interface: Opportunities for improved detection accuracy,” Int. IEEE/EMBS Conf. Neural Eng. NER, vol. 2003-January, no. April, pp. 587–590, 2003.

[8] J. R. Wolpaw, N. Birbaumer, D. J. Mcfarland, G. Pfurtscheller, and T. M. Vaughan, “BCIreview.pdf,” vol. 113, pp. 767–791, 2002.

[9] E. Science, “Intelligent Automatic Right-Left Sign Lamp Based on Brain Signal Recognition System INTELLIGENT AUTOMATIC RIGHT-LEFT SIGN LAMP BASED ON BRAIN SIGNAL RECOGNITION SYSTEM,” 2017.

[10] D. M. Taylor, S. I. H. Tillery, and A. B. Schwartz, “Direct Cortical Control of 3D Neuroprosthetic Devices,” vol. 296, no. June, pp. 1829–1832, 2002.

[11] D. J. Serruya MD, Hatsopoulos NG, Paninski L, Fellows MR, “Instant neural control of a movement signal. Nature 416:141 – 142,” p. 2002, 2002.

[12] M. A. L. Chapin, J. K., Markowitz, R. A., Moxon, K. A., and Nicolelis, “Direct real-time control of a robot arm using signals derived from neuronal population recordings in motor cortex. Nature Neuroscience 2, 664-670, (1999),” p. 1999, 1999.

[13] et al. Wessberg J, Stambaugh CR, Kralik JD, Beck PD, Laubach M, “Real-time prediction of hand trajectory by ensembles of cortical neurons in primates. Nature 408:361 – 365, 2000,” p. 2000, 2000.

[14] and M. P. G. Pfurtscheller, C. Neuper, D. Flotzinger, “EEGbased discrimination between imagination of right and left hand movement, Electroencephalography and Clinical Neurophysiology, vol. 103, pp. 642-651, 1997.,” vol. 103, p. 1997, 1997.

[15] and S. G. Y. Wang, B. Hong, X. Gao, “Implementation of a BrainComputer Interface Based on Three States of Motor Imagery, in 29th Annual International Conference of the IEEE Engineering in Medicine and Biology Society, EMBS2007, 2007, pp. 5059-5062,” p. 5062, 2007.

[16] and G. P. C. Guger, W. Harkam, C. Hertnaes, “Prosthetic Control by an EEG-based Brain-Computer Interface (BCI),in AAAE 5th European Conference for the Advancement of Assistive Technology, Düsseldorf, Germany, 1999,” p. 1999, 1999.

[17] U. Utrecht and R. Magnificus, Pattern matching using similarity measures, no. september. 2000.

[18] and S. N. M. R.V Pawar, P.P.Kajave, “Speaker Identification using Neural Networks,” vol. 7, no. August, p. I/165-I/168, 2002.

[19] M. Goyani, N. Dave, and N. M. Patel, “Performance analysis of lip synchronization using LPC, MFCC and PLP speech parameters,” Proc. - 2010 Int. Conf. Comput. Intell. Commun. Networks, CICN 2010, pp. 582–587, 2010.

[20] R. Rifkin and A. Kalutau, “In Defense of One-Vs-All Classification,” J. Mach. Learn. Res., vol. 5, pp. 101–141, 2004.

[21] K. R. A. Zade, C. Ardil, and S. S. Rustamov, “Investigation of Combined use of MFCC and LPC Features in Speech Recognition Systems,” World Acad. Sci. Eng. Technol., pp. 74–80, 2006.

[22] X. Huang, L. Shi, and J. a K. Suykens, “Support Vector Machine Classifier with Pinball Loss.,” IEEE Trans. Pattern Anal. Mach. Intell., no. 1, pp. 1–14, Sep. 2013.
[23] C.-W. Hsu and C.-J. Lin, “A comparison of methods for multiclass support vector machines,” IEEE Trans. Neural Netw., vol. 13, no. 2, pp. 415–425, Jan. 2002.
[24] JEFF SABATINI, “2017 Mercedes-Benz C-class Cabriolet,” Car and Driver, 2016. [Online]. Available: http://www.caranddriver.com/reviews/2017-mercedes-benz-c-class-cabriolet-first-drive-review. [Accessed: 07-Oct-1977].
[25] J. Uglov, L. Jakaite, V. Schetinin, and C. Maple, “Comparing Robustness of Pairwise and Multiclass Neural-Network Systems for Face Recognition NOISE PROBLEMS,” EURASIP J. Adv. Signal Process., vol. 2008, 2008.
[26] Berlin, Brain-Computer, and Interface, “BCI Competition III,” 2005. [Online]. Available: http://bnci.de/competition/iii/#download. [Accessed: 16-Jul-2019].