EEG Motor-Imagery BCI System Based on Maximum Overlap Discrete Wavelet Transform (MODWT) and cubic SVM

Samaa S. Abdulwahab *, Hussain K. Khleaf and Manal H. Jassim

Electrical Engineering Department, University of Technology, Iraq, Baghdad
Email: 316393@student.uotechnology.edu.iq

Abstract. Communication of the human brain with the surroundings became reality by using Brain-Computer Interface (BCI) based mechanism. Electroencephalography (EEG) being the non-invasive method has become popular for interaction with the brain. Traditionally, the devices were used for clinical applications to detect various brain diseases but with the advancement in technologies, companies like Emotiv, NeuSkY are coming up with low cost, easily portable EEG based consumer graded devices that can be used in various application domains like gaming, education etc as these devices are comfortable to wear also. This paper reviews the fields where the EEG has shown its impact and the way it has proved useful for individuals with severe motor disorder, rehabilitation and has become a means of communication to the real world. This paper investigates the use of Cubic SVM algorithm in the EEG classification. EEG feature extraction is implemented by maximum overlap discrete wavelet transform (MODWT) to reduce the dimensionality of data. The Sliding Window Technique is used to calculate the mean within each window samples. The feature vectors are loaded into the support vector machine (SVM) and optimize tree.

1. Introduction
In the last two decades, the research focus on Brain-Computer Interface (BCI) applications has been increased [1]. A BCI application is defined as the action of recording the Electroencephalography (EEG) activity of the brain. Then, the recorded data is filtered from noise, features are extracted and classified to perform a predefined action (such as opening or closing an artificial arm) [2]. Since the EEG data is large (1 sec of EEG data can have up to 256 time samples), feature extraction is needed, which is the first phase of the EEG classifier. Feature extraction is mandatory because most classifiers include matrix operations, and when a matrix is large in size, it is called an ill-conditioned matrix. The inverse of such matrices includes large numerical errors. Therefore, reducing the data size is needed. Usually, common EEG features are extracted in time or frequency bands [3]. The second step of the classifier includes two phases: training and recognition. In the training phase, EEG data sets of known classes are used to train the classification algorithm offline [4]. Then, the unclassified EEG data is fed to the classifier and a decision is made (determining which class the EEG samples belong to). The classification decision is then fed to the implemented hardware and the required action is done (moving an artificial arm or driving a wheelchair). Many feature reduction and classification techniques have been considered with EEG based Motor Imagery (MI) applications. Several feature extraction techniques have been considered in the literature [5], [6]:
1) time features.
2) spectrum energy features.
3) statistical features.

Time features can be extracted by many approaches, most commonly used are: Eigen Value Decomposition (EVD), Independent Component Analysis (ICA), Principal Component Analysis (PCA) and Linear Discriminant Analysis (LDA)[1], [7]. Because BCI applications requires real-time or semi real-time approaches, time features are the best candidate regarding computation time and complexity, compared to other feature extraction approaches. Many researchers investigated the use of different feature extraction and they showed promising results. [8], [9] discussed the small size problem of the feature matrix, for which the number of time features is largely higher than the number of channels. It reaches an accuracy of 84% for two class problem,[10] used Common Spatial Pattern (CSP) filtering along with LDA to reach a maximum accuracy of 99%. [11] showed that xDAWN algorithm accuracy outperforms both ICA and PCA. [7] and [8] demonstrated an accuracy comparison between different spectral feature extraction techniques, including Power Spectral Density (PSD), time-frequency energy distributions, periodogram, Spectrogram and Morlet Wavelet features. In this work, an offline EEG-MI classifier is built based on Support Vector Machine (SVM). It includes both steps mentioned above: feature extraction and classification. This work focuses on all implementation issues of the SVM classifier as well as the related theoretical background. The paper is organized as follows. In Section 2, the brain signal acquisition is described. In the section 3 the preprocessing of the signal is described. In the section 4 the feature extraction and classification is described. Then, in Section III the limitations of LDA are discussed with proposed solutions. The SVM classification algorithm is described in details. In Section IV, the setup procedure as well as the recording process is described. The algorithm parameters are varied and the corresponding accuracy are shown in figures.

![Block diagram of the BCI-Based motor imagery system](image)

Figure 1: the block diagram of the BCI-Based motor imagery system

2. Brain Signal Acquisition:
For BCI systems, the first layer is employed to gather brain signals with either invasive or noninvasive techniques. This is called electroencephalography (EEG) in the context of monitoring the electrical activity in the brain in relation to recording it in an experiment. Which transmit signals as electrical spikes. market or clinical grade instruments are being developed to detect these electrical signals. The first consumer system for measuring the brain activity is Emotiv Epoc has 14 electrodes, otherwise NeuroSky's Epoc measures brain activity. Collecting information from these devices wirelessly or through a Bluetooth link. These businesses sell devices at low prices.

**Figure 2:** EMOTIV EPOC headset

2.1 Database Description
The data used in this paper recorded by EMOTIV EPOC+. A data contains of four separate tasks with respect to one of either left-hand or right-hand imagination (class 1, class 2) respectively, and one per foot (class 3) and the last one for tongue picture involving the tongue (class 4). The recording sessions were held on two different days in two different days. Every session included six runs that were broken up by brief rest periods. Every run consisted of 48 trial (12 for each class). The figure 3 displays the timeline of the experiment being used for data collection.

**Figure 3:** Illustrative timeline of an acquisition trial

3. EEG Signal Pre-Processing:
The pre-processing after signal acquisition is to remove any noise, artifacts which are recorded while capturing the signals of the devices. Some of the unwanted signals include: Whenever electronic device is mounted some inference is there. Some muscular activity leads to EMG signals, Eye movement or blinking cause ocular artifact. The unwanted noises present in the EEG recording will lead to wrong conclusions and may bias the analysis of the EEG readings. So various filter are used to remove out the noises from the signals. In this paper used firstly, each channel was resampled to 128 Hz and filtered using a low pass filter (Chebyshev Type II Lowpass filter, cutoff: 40 Hz).
4. EEG Signal Extraction and Classification:

Since the EEG signal time samples show no obvious difference between different MI commands, it is not possible to use classification directly on time samples. Figure 5 depicts EEG time samples for two classes. It can be seen that there is no obvious classification pattern between the two classes. Therefore, feature extraction is needed, on which the classification accuracy relies on
4.1 Feature Extraction

Maximum Overlap Discrete Wavelet Transform (MODWT) The relative tolerances are independent of detail and approximation elements, but can change with both changes. In the experiment described in this paper, the wavelet method was employed to classify theta (θ), delta (delta, theta), and theta (theta) patterns of neural activity. Delta waves (deep meditation and deep sleep) are most commonly found within the range of 0-4 Hz. On the very low frequencies, theta waves (4 to 8 Hz) can be found in meditation, learning, and while slightly higher frequencies can be found in sleep, study, and memory. Alpha waves are in the 8 to 12 Hz range and are most prominent while the brain is at rest. Beta waves, located between 12-32 Hz, are produced when the brain is actively engaged in thought or in the outside of the external environment. Mu rhythm is typically located at 8-13 Hz, and is often correlated with the physical coordination of muscle groups in that both rhythmic and non-rhythmic ways. It was necessary to resample the raw EEG signal before wave decomposition in order to obtain 128 Hz. It was performed both on the discrete wavelet transform (DWT) as well as the full overlap wavelet transform (MODWT). For the analysis, a multistep process was required, in both cases a Haar basis has been employed. The Fig. 6 demonstrates a decimation of the EEG to four separate components, using a maximum wavelet transform (MODWT).

![Diagram showing EEG signal decomposed into 4 bands, with each significant sub-band highlighted]

Figure 6: EEG signal decomposed into 4 bands, with each significant sub-band highlighted
Figure 7: MODWT Graphic band (Alpha, Beta, Delta, Gamma, Theta) after decomposition
It uses statistical metrics computed from each sub-band after wavelet decomposition over a window of 25 samples. Mean and variance, as shown in Equations 1-2. Another measurement widely used in this type of experiment is the median, which is defined as the value of the central in a group of ordered data. The fourth function vectorization generator, which measures the uncertainty or complexity of a random signal, is defined as in Equation 3

\[
\overline{x} = \frac{1}{n}\sum_{i=1}^{n} x_i
\]

(1)

\[
\sigma^2 = \frac{1}{n}\sum_{i=1}^{n} (x_i - \overline{x})^2
\]

(2)

\[
H(x) = \sum_{i=1}^{n} x_i \log_2(x_i)
\]

(3)

4.2 Classification
Several kernels were implemented as part of the support vector machines (SVMs) for comparison to each data type (e.g., differentiate among different groupings of variables. Structural risk minimization consists of a statistical learning with a technique which iterates on the data structure to construct a model and a probability distribution. Though linear discrimination is not possible in the
The definition of the algorithm does help us discriminate among various features in the input space is realized as features are mapped into a higher dimensional space that does not have this separation mapping can be performed using a linear or nonlinear algorithm, depending on the purpose of the used kernel. Instead of training a decision trees, it first locates the class separators with the largest margin between instances. Then, it finds the class separators and selects the ones that divide the classes with the smallest error. In a variety of instances, the optimum hyperplane is expressed as a combination of a number of characteristics, known as the support vectors for the optimum. Here, we will use a straight-line, the standard, and the exponential functions for mappings, so that you can research the differences between them.

![Figure 8](image_url): confusion matrix obtained using MODWT and SVM with cubic kernel

| Classes       | True positive rate | False negative rate | Positive predictive value | False discovery rate |
|---------------|--------------------|---------------------|---------------------------|----------------------|
| Foot          | 94.8%              | 5.2%                | 99.1%                     | 0.9%                 |
| Left hand     | 96.8%              | 3.1%                | 94.8%                     | 5.2%                 |
| Right hand    | 97.8%              | 2.2%                | 99.1%                     | 0.9%                 |
| Tongue        | 99.1%              | 0.9%                | 95.8%                     | 4.2%                 |

Table 2: Comparison result with previous works of detection and classification of EEG-based Motor-Imagery.
5. Implementation and Results

The EEG dataset was collected using EMOTIV EPOC headset, which provides 14 channels. The subject was asked to sit at rest state for 10 to 20 minutes prior to recording. This allows the relaxation of the brain activities in general and the accuracy of recorded EEG data. The subject is then asked to put both of his hands rest on a table in front of him, while his eyes are opened. Then, a cue is shown to him to imagine the movement of his right hand (Close and Opening). While he closes and opens his right hand repeatedly, the EEG data is recorded for 5 seconds at a rate of 250 samples per second. Then repeat the operation for all 3 classes. Those trails are imported to MATLAB. Each trail consists of 640 samples. Trial recording is repeated until 120 trials are recorded and stored of all classes. These trials are named as Class Right training. The same procedure is repeated for Class Left training samples. MATLAB 2020a environment is used to record, store and implement the SVM algorithm. After the training data is acquired, the features representing each class are extracted and stored as matrices. The subject then performed either left- or right-hand simple movement (closing and opening of each hand) repeatedly for five seconds. The choice of which hand to move is random. During the subject session, the EEG signal is recorded at a sampling rate of 250 samples per second. This step is repeated for 120 trails with different random choices for either left- or right-hand tongue and foot of the subject. For each trial, the SVM algorithm then extracts the features from the EEG data set and those features are compared against the training features extracted during the training phase. According to the sign of table 3 the winning class result is recorded. The accuracy of the algorithm is calculated as:

\[
\text{Classification Rate} = \frac{\text{True Positives} + \text{True Negatives}}{\text{True Positives} + \text{False Positives} + \text{True Negatives} + \text{False Negatives}} = \frac{TP + TN}{TP + TN + FP + FN} \tag{4}
\]

\[
\text{Sensitivity} = \frac{\text{True Positives}}{\text{True Positives} + \text{False Negatives}} \tag{5}
\]

\[
TPR = \frac{TP}{P} = \frac{TP}{TP + FN} = 1 - FNR \tag{6}
\]

\[
\text{Specificity} = \frac{\text{True Negatives}}{\text{False Positives} + \text{True Negatives}} \tag{7}
\]

\[
\text{Positive predictive value (PPV)} = \frac{\sum \text{True positive}}{\sum \text{Predicted condition positive}} = \frac{TP}{TP + FP} = 1 - FDR \tag{8}
\]

\[
FNR = \frac{FN}{P} = \frac{FN}{FN + TP} = 1 - TPR \tag{9}
\]

\[
FDR = \frac{FP}{FP + TP} = 1 - PPV \tag{10}
\]

As it can be seen from Fig. 3, the best channel accuracy occurs at channel 3, which complies with F3 sensor. This sensor is the nearest to the MI region in the brain, related to EMOTIV EPOC headset.
(see Fig. 4). The regularization parameter (η) selection plays a role in increasing the accuracy of the inverse of the matrix. Several values were tested for η against accuracy, as shown in Fig. 5. The best accuracy occurs between [0.85, 0.9]. Fig. 6 shows the effect of windows size (in samples) on the classification accuracy. It can be seen that a window size of approximately 1.2 seconds gives the best accuracy. In another word, the mean value for every 150 samples represents certain EEG signal. On the other hand, increasing the window size beyond 175 samples will decrease the accuracy. The reason for this accuracy degradation is that larger window size will overlap with adjacent EEG samples, therefore, the unique time feature characteristics will be destroyed. As a result, an EEG recording of 1.5 seconds is sufficient for LDA classification. Therefore, the user does not have to repeat EEG emotion within a single trial.

Accuracy for SVM: 97.39  
Sensitivity for SVM: 99.03  
Specificity for SVM: 99.47  
True Negative Rate (TNR) for SVM: 99.47  
True Positive Rate (TPR) for SVM: 99.03  
Positive Predictive Value (PPV) for SVM: 99.47  
False Negative Rate (FNR) for SVM: 0.97  
False Discovery Rate (FDR) for SVM: 0.53

Table 3: comparison of different types of SVM classifier

| Classifier type          | Train accuracy | Test accuracy |
|--------------------------|----------------|---------------|
| Medium Gaussian SVM      | 68.30%         | 73.70%        |
| Coarse Gaussian SVM      | 51.60%         | 53.00%        |
| Linear SVM               | 52.30%         | 58.30%        |
| Quadratic SVM            | 77.10%         | 85.50%        |
| Cubic SVM                | **92.60%**     | **97.36%**    |

6. Conclusion
A study was done in this paper that demonstrated how an EEG can be used to classify patterns of electrical activity linked to motor imagery with BCI systems. Maximum Overlap Discrete Wavelet analysis worked as an excellent class separator using the defined features from the Maximum-Overlap Discrete classifier to find the specific subbands yielded excellent results. The support vector machine provided an accuracy of 98.81% on average, but the support vector algorithm was only correctable to 97.77% while operating on a kernel of nine points. Other movements of this research is presently in the robot control phase of trying to build upon involve simple ones.

References

[1] S. S. Abdulwahab, H. K. Khleaf, M. H. Jassim, and S. Abdulwahab, “A Systematic Review of Brain-Computer Interface Based EEG,” *Iraqi J. Electr. Electron. Eng.*, vol. 16, Dec. 2020, doi: 10.37917/ijeee.16.2.9.

[2] S. N. Abdulkader, A. Atia, and M. S. M. Mostafa, “Brain computer interfacing: Applications and challenges,” *Egypt. Informatics J.*, vol. 16, no. 2, pp. 213–230, Jul. 2015, doi: 10.1016/j.eij.2015.06.002.

[3] J. Preethi and C. Science, “A Survey on EEG Based Emotion Analysis using various Feature Extraction Techniques,” *Int. J. Sci. Eng. Technol. Res.*, vol. 3, no. 11, pp. 3113–3120, 2014.

[4] F. Lotte *et al*., “A review of classification algorithms for EEG-based brain-computer interfaces: A 10 year update,” *J. Neural Eng.*, vol. 15, no. 3, 2018, doi: 10.1088/1741-2552/aa821f.
[5] S. Snyder and X. A. Shen, “A Review of Brain Signal Processing Methods,” Int’l Conf. Biomed. Eng. Sci., vol. BIOENG’17, pp. 10–16, 2017.

[6] A. Al-Saegh, S. A. Dawwd, and J. M. Abdul-Jabbar, “Deep learning for motor imagery EEG-based classification: A review,” Biomedical Signal Processing and Control, vol. 63. Elsevier Ltd, Jan. 01, 2021, doi: 10.1016/j.bspc.2020.102172.

[7] A. Khosla, P. Khandnor, and T. Chand, “A comparative analysis of signal processing and classification methods for different applications based on EEG signals,” Biocybern. Biomed. Eng., vol. 40, no. 2, pp. 649–690, 2020, doi: 10.1016/j.bbe.2020.02.002.

[8] J. Meng, S. Zhang, A. Bekyo, J. Olsoe, B. Baxter, and B. He, “Noninvasive Electroencephalogram Based Control of a Robotic Arm for Reach and Grasp Tasks,” Sci. Rep., vol. 6, p. 38565, Dec. 2016, doi: 10.1038/srep38565.

[9] R. Vishwakarma, H. Khwaja, V. Samant, P. Gaude, M. Gambhir, and S. Aswale, “EEG Signals Analysis and Classification for BCI Systems: A Review,” Int. Conf. Emerg. Trends Inf. Technol. Eng. ic-ETITE 2020, pp. 1–6, 2020, doi: 10.1109/ic-ETITE47903.2020.066.

[10] B. Rivet, A. Souloumiac, V. Attina, and G. Gibert, “xDAWN Algorithm to Enhance Evoked Potentials: Application to Brain-Computer Interface,” IEEE Trans. Biomed. Eng., vol. 56, no. 8, pp. 2035–2043, 2009, doi: 10.1109/TBME.2009.2012869.

[11] N. Brodu, F. Lotte, and A. Lécuyer, “Comparative study of band-power extraction techniques for Motor Imagery classification,” in IEEE SSCI 2011 - Symposium Series on Computational Intelligence - CCMB 2011: 2011 IEEE Symposium on Computational Intelligence, Cognitive Algorithms, Mind, and Brain, 2011, pp. 95–100, doi: 10.1109/CCMB.2011.5952105.

[12] N. Lu and H. Miao, “A Deep Learning Scheme for Motor Imagery Classification based on Restricted Boltzmann Machines,” ieeexplore.ieee.org, doi: 10.1109/TNSRE.2016.2601240.

[13] G. Sagee, S. H.-2017 I. C. on, and undefined 2017, “EEG feature extraction and classification in multiclass multiuser motor imagery brain computer interface using Bayesian Network and ANN,” ieeexplore.ieee.org, Accessed: Mar. 28, 2021. [Online]. Available: https://ieeexplore.ieee.org/abstract/document/8342691/.

[14] S. Kumar, A. Sharma, K. Mamun, and T. Tsunoda, “A Deep Learning Approach for Motor Imagery EEG Signal Classification.” Accessed: Mar. 28, 2021. [Online]. Available: http://www.bbci.de/competition/iii/.

[15] H. Lee, Y. C.-2018 I. C. on, and undefined 2018, “A convolution neural networks scheme for classification of motor imagery EEG based on wavelet time-frequency image,” ieeexplore.ieee.org, Accessed: Mar. 28, 2021. [Online]. Available: https://ieeexplore.ieee.org/abstract/document/8343254/.

[16] K. S. Hong et al., “Hybrid brain-computer interface techniques for improved classification accuracy and increased number of commands: A review,” Front. Neurorobot., vol. 11, no. JUL, pp. 1–22, 2017, doi: 10.3389/fnbot.2017.00035.

[17] Y. Kim, N. Kwak, S. L.-2018 6th I. Conference, and undefined 2018, “Classification of motor imagery for Ear-EEG based brain-computer interface,” ieeexplore.ieee.org, Accessed: Mar. 28, 2021. [Online]. Available: https://ieeexplore.ieee.org/abstract/document/8311517/.

[18] S. Sakhati, C. Guan, and S. Yan, “Learning Temporal Information for Brain-Computer Interface Using Convolutional Neural Networks,” IEEE Trans. Neural Networks Learn. Syst., vol. 29, no. 11, pp. 5619–5629, Nov. 2018, doi: 10.1109/TNNLS.2018.2789927.

[19] W. Ko, J. Yoon, E. Kang, E. Jun, … J. C.-2018 6th I., and undefined 2018, “Deep recurrent spatio-temporal neural network for motor imagery based BCI,” ieeexplore.ieee.org, Accessed: Mar. 28, 2021. [Online]. Available: https://ieeexplore.ieee.org/abstract/document/8311535/.