Classification of EEG-based motor imagery BCI by using ECOC

Jahangir Mobarezpour, Reza Khosrowabadi, Reza Ghaderi, Keivan Navi

Institute for Cognitive and Brain Science, Shahid Beheshti University GC, Evin Sq, Tehran, Iran

Abstract

Accuracy in identifying the subjects intentions for moving their different limbs from EEG signals is regarded as an important factor in the studies related to BCI. In fact, the complexity of motor-imagery and low amount of signal-to-noise ratio for EEG signal makes this identification as a difficult task. In order to overcome these complexities, many techniques such as various feature extraction methods, learning algorithms, and classifier schemes have been developed in this regard. However, conducting more research is necessary for improvement. The present study aimed to use an ensemble learning approach to improve the performance of MI-BCI systems. Therefore, filter bank common spatial pattern (FBCSP), as a well-known feature extraction method, was used to produce separable features from EEG signals. Accordingly, error correcting output codes (ECOC) was applied on several learning algorithms to classify four classes of motor imagery tasks. The proposed ECOC ensemble technique was tested on the data set 2a from BCI competition IV. Based on the results, the ECOC can lead to an improvement by using the naive Bayesian parzen window algorithm, compared to the winner algorithm of BCI competition IV, which is superior to other selected state of the art algorithms.

Keywords: Brain computer interface (BCI); Error Correcting Output Codes (ECOC); Electroencephalography (EEG); Motor imagery; Filter bank common spatial pattern (FBCSP)


1. Introduction

The present study aimed to focus on the classification algorithm used in brain computer interface (BCI) systems. First, BCI systems were created for disabled people to provide a non-muscular channel for controlling external devices with the brain electrical activities. For instance, it was used to control wheelchairs [1] or in neural prosthetics [2]. However, its applications in navigation for
virtual environment \cite{3,4} or assisting in performing highly-demanding mental tasks \cite{5,6,7,8,9} have been extended to healthy individuals during recent years.

In general, the brain activities could either be directly recorded from neural electrical activities by using electro-corticography (ECOG), magneto-encephalography (MEG), EEG or from blood de/oxygenation changes indirectly by using functional near infrared spectroscopy (fNIRS), and functional magnetic resonance imaging (fMRI). The ECOG is partially invasive while MEG, EEG, fNIRS, and fMRI are non-invasive methods and more safe in terms of surgery and infection. The EEG-based BCI is preferred to other existing methods since most of BCI applications require a noninvasive and portable system. Further, the EEG has been implemented in many other applications such as clinical diagnostics for epilepsy \cite{11,12,13}, dementia \cite{14,15}, dyslexia \cite{16}, autism spectrum disorder \cite{17,18}, biometric systems such as fraud detection \cite{19,20,21,22}, which can highlight its potential for the BCI studies.

The EEG signals are produced by the synchronous activity of clusters in neurons with similar spatial orientation, which could be recorded on the scalp. Typically, EEG is mixed with other non-cortical activities such as ECG, EMG, EOG, electrical power-line noises \cite{23}. Thus, separating only the brain activities seems difficult. The EEG-based BCI systems either use evoked potentials or the spontaneous activities. In addition, the evoked potentials are generated in response to a standard sensory stimulation such as steady state visually evoked potentials (SSVEP) \cite{24,25} by which brain can generate electrical activity at the same (or multiples of) frequency of the visual stimulation. This approach is regarded as accurate but requires external stimulation and long training time although it is not appropriate for a long-term use. However, spontaneous BCI systems are based on mental imaginations, without any training or external stimulation. Further, the imagination of different limb movement can activate very close areas in the brain, leading to the difficulty of the separation task. In order to overcome this problem, a large number of techniques were developed to extract separable features from EEG signals \cite{26,27,28}. Further, many learning algorithms with different schemas \cite{29,30,31} have been used to identify various mental tasks. By considering all the aforementioned issues, the present study aimed to use an ensemble learning approach to improve the performance of motor imagery (MI) BCI systems. Thus, filter bank common spatial pattern (FBCSP) was used to produce separable features from EEG signals. Accordingly, error correcting output codes (ECOC) was implemented with several learning algorithms to classify four classes of motor imagery tasks. Then, the proposed ECOC ensemble technique was tested on data set 2a from BCI competition IV (http://www.bbci.de/competition/iv). The data including four motor-imaginary class signals were recorded from nine subjects in two sessions during different days.

The concept of ECOC comes from information theory for the communication systems for the purpose of correcting noisy signals. Then, it was used in machine learning for an improvement in classification accuracy, face verification \cite{33,34,35}, text classification \cite{36,37}, handwritten recognition \cite{38}, and object recognition \cite{39}. Regarding the results of previous studies, the ECOC can reduce bias and variance of classification \cite{39,40,41}. Therefore, it is hypothesized that the proposed ECOC technique may help us improve the best results achieved in BCI competition IV.

The rest of paper is organized as follows. Section 2 explains the experimental protocol and signal analysis method. In this regard, the design of motor imagery tasks is introduced and accordingly feature extraction and classification are clarified. Section 3 presents the ECOC method and experimental results. Finally, conclusions and suggestions for further research are described in Section 4.
2. Materials and Method

In the present study, dataset 2a of the BCI competition IV was recruited. The dataset included EEG data recorded from nine subjects. A cue-based paradigm was applied for recording EEG data during four different motor imagery tasks including left hand (class 1), right hand (class 2), both feet (class 3), and the tongue (class 4) movement imagination. Data recording for each subject was performed in two sessions on different days. Each session included six runs separated by short breaks and each run contained 48 trials (12 for each of the four classes). Finally, 288 trials were yielded per each session.

The subjects were sitting in a comfortable armchair in front of a computer screen. In the beginning of each session, approximately 5-min EEG was recorded to estimate electrooculogram (EOG) artefact. The 5-min recording consisted of three blocks including (1) two minutes for eyes-open resting state by looking at a fixation cross on the screen, (2) one minute for eyes-closed resting state, and (3) one minute for eye movements. In the next procedure, motor imagery tasks were performed. At the beginning of each trial \((t = 0\, \text{s})\), a fixation cross appeared on the black screen. In addition, a short acoustic warning tone was presented. After two seconds \((t = 2\, \text{s})\), a cue in the form of an arrow pointing either to the left, right, down or up (corresponding to one of the four classes left hand, right hand, foot, and tongue) appeared and remained on the screen for 1.25 seconds. The cue prompted the subjects to perform the desired motor imagery task without providing any feedback. In the next procedure, the subjects were asked to conduct the motor imagery task until disappearing the fixation cross from the screen at \(t = 6\, \text{s}\). Finally, each trial was followed by a short break, where the screen was black again. Fig. 1 illustrates the experimental paradigm.

![Figure 1](image)

**Figure 1:** Experimental paradigm used for data collection in the BCI competition IV, dataset 2a

2.1. Data recording

First, twenty-two Ag/Ag-Cl electrodes (with inter-electrode distances of 3.5cm) were used to record the EEG signals (Fig. 2.A). Then, all signals were referenced to the left mastoid while the right mastoid was considered as the ground. The data were sampled with the frequency of 250 Hz and with a bandpass filter between 0.5 to 100 Hz. In addition, the sensitivity of the amplifier was set to 100 \(\mu\text{V}\) and an additional 50 Hz notch filter was used to suppress the electrical powerline noise. Further, three EOG electrodes were used with the same sampling frequency (Fig. 2.B). Furthermore, the sensitivity of the EOG amplifier was set to 1 mV and EOG signals were filtered similar to EEG signals. The EOG signals are only used for artifact rejection [13] and are not used for the classification.
purpose. Finally, a visual inspection was conducted on all data by an expert and accordingly the trials involving artifacts were remarked [27, 44].

Figure 2: Montage of electrodes used for data collection in the BCI completion IV, dataset 2a. (A) Placement of EEG electrodes according to the international 10-20 system, (B) Placement of EOG electrode

2.2. Feature extraction by filter bank common spatial pattern (FBCSP)

The EEG signal is usually contaminated with artifacts, which are either easily removed by an appropriate filtering (e.g., notch filtering at 50/60 Hz for the powerline noise) or suppressed by using source decomposition techniques such as independent component analysis (ICA) [45, 46]. However, some artifacts could not be removed by the above-mentioned methods such as the effect of volume conduction and skull tissue. The skull tissue acts as a low-pass filter [47], which causes localized cortical sources, which are originated by a cognitive/motor task, and dominated by widely spread sources. In addition to the low-pass filtering effect of the skull, the conductive surface area of each electrode acts as low pass spatial filter, which can eliminate the signal components with wavelengths shorter than the electrode diameter. The combined effect of the skull tissue and the conductive surface of electrodes enable us to sample the EEG signal spatially without aliasing. In order to avoid spatial aliasing completely, the electrode diameter should be selected either equal to the edge-to-edge distance between the electrode conductors or spread of gel layer. In addition, the approach of estimating local sources could be performed by common spatial patterns (CSP). The CSP method was first applied in BCI studies for binary classification purpose (left- and right-hand movements) (48-50). This algorithm applies spatial filters on a set of training data in order to maximize inter-class variance while minimizing intra-class variances [48, 50].

However, it is logical to consider both spectral and spatial characteristics of the EEG signals simultaneously since the present study aimed to extract the features from four different motor imagery tasks, in order to separate the features of four classes better [51, 52, 53]. Therefore, the filter bank
CSP (FBCSP), as a well-known feature extraction algorithm in BCI research, was implemented in the present study [54].

In the next procedure, the EEG signals were filtered to 9 bands of 4 Hz, from 4 to 40 Hz without overlap, and accordingly CSP method was separately used for each band. Finally, two features for each band and totally 18 features from all bands were calculated.

2.3. Classification

Now, the extracted features should be classified by using a learning algorithm. In this study, the Naive Bayesian Parzen Window (NBPW) method was selected as the main learning algorithm. The NBPW is a stable and non-iterative classifier, which is appropriate for combining with ECOC technique [55]. In addition, a good performance with FBCSP features was observed in the field of BCI [54]. Further, the results were compared with other well-known classifiers such as support vector machines (SVM) [31], linear discriminant analysis (LDA) [31], and K-nearest neighbour (KNN) [31].

The NBPW approach can use any feature vector \( f \), which classify the classes by equation (2.1),

\[
\hat{C} = \arg \max_{C_i} P(C_i | f)
\]

where \( C_i \) denotes the class labels and \( P(C_i | f) \) presents the probability \( f \) related to the class \( C_i \). The \( P(C_i | f) \) is determined by using the Bayes rule as described in the following equation.

\[
P(C_i | f) = \frac{P(f | C_i)P(C_i)}{P(f)} = \frac{P(f | C_i)P(C_i)}{\sum_i P(f | C_i)P(C_i)}
\]

(2.2)

The NBPW method assumes that the elements of the feature vectors \( f = [f_1, \ldots, f_E] \) are independent. Thus, \( P(f | Ci) \) can be calculated as follows;

\[
P(f | C_i) = \prod_{f=1}^{E} P(f_j | C_i)
\]

(2.3)

where \( P(f_j | C_i) \) is estimated by using Parzen windowing method [56] presented in the following equation.

\[
P(f_{j,i} | \hat{C}) = \frac{1}{n_{\hat{C}}} \sum_{k \in T_{\hat{C}}} \phi(f_{j,i} - f_{i,k}, w)
\]

where \( n_c \) denotes the number of data samples in class \( \hat{C} \), \( T_{\hat{C}} \) represents a part of the training data related to class \( \hat{C} \), \( f_k \) is considered as the feature value of the \( k \)th trial of \( T_{\hat{C}} \) from \( f_i \), and \( \phi \) indicates the smoothing kernel function with smoothing parameter of \( w \). The univariate Gaussian kernel was utilized according to the equation (2.4) as the kernel function.

\[
\phi(f, w) = \frac{1}{\sqrt{2\pi}} e^{-\left(\frac{f^2}{2w^2}\right)}
\]

(2.4)

However, it is worth noting that this assumption presents a limitation and may lead to a classification error.
2.4. Error correcting output codes (ECOC)

As it was already mentioned, the ECOC ensemble learning approach is applied to improve the classification accuracy of the MI-BCI system. The ECOC consists of encoding and decoding. Different designs have been used for encoding and decoding [57]. In the encoding phase, an \( M \times N \) dimensional matrix called “matrix code” is formed, where \( M \) denotes the number of classes and \( N \) represents the binary classifiers. Each row of matrix code is a code word for a class. The unlabelled data are classified after training binary classifiers with the assigned labels in the decoding phase. Typically, a class label is assigned to the data by calculating its closeness to the code-word defined in matrix code. The closest code word is defined as the target class.

The ECOC in classification requires both rows and columns of the matrix code in order to be separated well like when it is computed by the hamming distance [58]. The row separation helps the classes to be maximally far apart from each other while the column separation ensures that the learned bits are uncorrelated so that the errors in each bit can be independent from others. Therefore, an error may not proceed to other bits if an error occurs in a bit. In this regard, different types of coding have been introduced so far [39, 40] and the results of some studies indicated that BoseChaudhuriHocquenghem (BCH) codes could produce a better performance [40]. In the present study, the BCH-ECOC method was implemented using a software package introduced by Escalera et al [57].

A BCH code word has the length of \( (2^n - 2) \) for multiclass classification problem, where \( n \) presents the number of classes. The second half of the BCH code word is considered as complementary for the first half, which is used in unstable binary classifiers to reduce the variation of the results. However, the NBPW used in this study was considered as a stable classifier algorithm. Therefore, a reduction of \( (2^{n-1} - 1) \) occurs in the length of BCH code word. Thus, the code word length should be 7 by considering four classes of motor imagination. As shown in Table 1, the coding is independent from the classes and the distances between the classes are equal. In addition, Hamming method, defined in equation (2.5) was used for the decoding purpose.

\[
HD(x, y_j) = \frac{\sum_{j=1}^{n} 1 - \text{sign}(x^j \cdot y^j)}{2} \tag{2.5}
\]

where \( x \) denotes the output code word of the test data, and \( y_j \) presents the code word from matrix code corresponding to class \( C_i \). The minimum Hamming distance determines the target class.

Finally, the performance of the proposed MI-BCI system was compared with other algorithms by using the ECOC technique. Several classification schemas including Divide and conquer (DC), Pair wise (PW), and One versus the rest (OVR) [59, 60, 61] methods were implemented to adapt the FBCSP method for the present multi-class classification problem.

<table>
<thead>
<tr>
<th>code words</th>
<th>binary classifiers</th>
</tr>
</thead>
<tbody>
<tr>
<td>Class 1</td>
<td>1 1 1 0 0 1 0</td>
</tr>
<tr>
<td>Class 2</td>
<td>1 1 0 0 1 0 1</td>
</tr>
<tr>
<td>Class 3</td>
<td>0 0 1 0 1 1 1</td>
</tr>
<tr>
<td>Class 4</td>
<td>0 1 1 1 0 0 1</td>
</tr>
</tbody>
</table>
3. Results

Several classification schemas by using different learning algorithms such as ECOC, DC, PW and OVR were implemented in the present study. Then, inter-subject classification accuracy of the proposed EEG-based MI-BCI system was compared for discriminating four classes of motor imaginations (left hand, right hand, both feet, and tongue). As shown in Table 2 and Fig. 3, the ECOC method had better performance, compared to other methods. In addition, the ECOC method demonstrated higher classification accuracy, compared to other methods by using NBPW, LDA, KNN classifiers. Further, the MI-BCI system by using NBPW learning algorithm with ECOC technique could outperform other methods.

Table 2: Inter-subjects classification accuracy of the EEG-based MI-BCI system for discrimination of 4 classes of motor imaginations (mean ± std)

<table>
<thead>
<tr>
<th>Feature</th>
<th>Classifier scheme</th>
<th>NBPW</th>
<th>SVM</th>
<th>LDA</th>
<th>KNN</th>
</tr>
</thead>
<tbody>
<tr>
<td>FBCSP with 22 channels</td>
<td>ECOC</td>
<td>68.36 ± 13.47</td>
<td>64.93 ± 16.46</td>
<td>63.31 ± 14.35</td>
<td>65.35 ± 14.28</td>
</tr>
<tr>
<td></td>
<td>PW</td>
<td>64.93 ± 14.76</td>
<td>57.37 ± 15.65</td>
<td>61.42 ± 17.07</td>
<td>61.88 ± 16.56</td>
</tr>
<tr>
<td></td>
<td>OVR</td>
<td>66.05 ± 13.23</td>
<td>65.24 ± 16.87</td>
<td>58.33 ± 12.24</td>
<td>64.47 ± 14.32</td>
</tr>
<tr>
<td></td>
<td>DC</td>
<td>56.83 ± 10.67</td>
<td>57.37 ± 15.65</td>
<td>62.35 ± 13.41</td>
<td>55.40 ± 11.51</td>
</tr>
</tbody>
</table>

Abbreviations: NBPW Naive Bayesian parsons window, SVM Support vector machine, LDA Linear discriminant analysis, KNN K-nearest neighbors, ECOC Error correcting output code, PW Pairwise, OVR One versus rest, DC Divide and concur, FBCSP Filter bank common spatial pattern, MI-BCI Motor imagery brain computer interface.

Figure 3: Inter-subjects classification accuracy of the EEG-based MI-BCI system for discrimination of 4 classes of motor imaginations
Furthermore, the ECOC results indicated a slight improvement, compared to the winner algorithm of the BCI Competition IV (http://bbci.de/competition/iv/results) (Table 4).

Table 3: Intra-subject classification accuracy and kappa factor of the EEG-based MI-BCI system for discrimination of 4 classes of motor imagination using ECOC technique

<table>
<thead>
<tr>
<th>Subject</th>
<th>Classifier</th>
<th>NBPW</th>
<th>SVM</th>
<th>LDA</th>
<th>KNN</th>
<th>Winner algorithm of BCI competition IV</th>
</tr>
</thead>
<tbody>
<tr>
<td>Subject 1</td>
<td>NBPW</td>
<td>80.21 – 0.74</td>
<td>78.47 – 0.71</td>
<td>72.92 – 0.64</td>
<td>73.96 – 0.65</td>
<td>75.7 – 0.68</td>
</tr>
<tr>
<td>Subject 2</td>
<td>SVM</td>
<td>50.35 – 0.34</td>
<td>49.31 – 0.32</td>
<td>49.65 – 0.33</td>
<td>52.08 – 0.36</td>
<td>56.27 – 0.42</td>
</tr>
<tr>
<td>Subject 3</td>
<td>LDA</td>
<td>82.64 – 0.77</td>
<td>82.99 – 0.77</td>
<td>74.65 – 0.66</td>
<td>80.90 – 0.75</td>
<td>80.87 – 0.74</td>
</tr>
<tr>
<td>Subject 4</td>
<td>KNN</td>
<td>57.29 – 0.43</td>
<td>57.99 – 0.44</td>
<td>59.38 – 0.46</td>
<td>54.51 – 0.39</td>
<td>61.07 – 0.48</td>
</tr>
<tr>
<td>Subject 5</td>
<td></td>
<td>63.19 – 0.51</td>
<td>45.14 – 0.27</td>
<td>49.65 – 0.39</td>
<td>51.04 – 0.35</td>
<td>54.85 – 0.40</td>
</tr>
<tr>
<td>Subject 6</td>
<td></td>
<td>49.65 – 0.33</td>
<td>42.01 – 0.23</td>
<td>41.32 – 0.22</td>
<td>43.40 – 0.25</td>
<td>45.47 – 0.27</td>
</tr>
<tr>
<td>Subject 7</td>
<td></td>
<td>82.64 – 0.77</td>
<td>80.90 – 0.75</td>
<td>85.76 – 0.81</td>
<td>85.42 – 0.81</td>
<td>82.97 – 0.77</td>
</tr>
<tr>
<td>Subject 8</td>
<td></td>
<td>74.65 – 0.66</td>
<td>78.13 – 0.71</td>
<td>67.36 – 0.57</td>
<td>75.35 – 0.67</td>
<td>81.62 – 0.75</td>
</tr>
<tr>
<td>Subject 9</td>
<td></td>
<td>74.65 – 0.66</td>
<td>69.44 – 0.59</td>
<td>69.10 – 0.59</td>
<td>71.53 – 0.62</td>
<td>70.45 – 0.61</td>
</tr>
<tr>
<td>Average</td>
<td></td>
<td>68.36 ± 13.47</td>
<td>64.93 ± 16.46</td>
<td>63.31 ± 14.35</td>
<td>65.35 ± 14.28</td>
<td>67.70 ± 13.72</td>
</tr>
<tr>
<td>mean ± std</td>
<td></td>
<td>−0.58 ± 0.18</td>
<td>−0.53 ± 0.22</td>
<td>−0.52 ± 0.18</td>
<td>−0.54 ± 0.20</td>
<td>0.57 ± 0.18</td>
</tr>
</tbody>
</table>

Kappa factor=(Obtained classification accuracy – Chance classification)/(1- Chance classification )

As this table shows, the average of intra-subject classification accuracy by using the NBPW classifier with ECOC method is 68.36 ± 13.47 (kappa factor: 0.58 ± 0.18), which is slightly higher than the result of competition winner as 67.70 ± 13.72 (kappa factor: 0.57 ± 0.18).

4. Conclusion

Extracting useful information from EEG signals, along with their assignment to a unique mental task is considered as the objective of BCI research, which is not an easy task. The task is more difficult while working on the identification of motor imagination. The motor imagery tasks activate small part of cortex called primary motor region. Therefore, the information related to motor imagination of various limbs is highly overlapped, which is regarded as a big challenge for their identification. Therefore, extracting separable features and improving learning algorithms can help us to overcome this challenge since the ECOC technique has a potential to reduce bias and variance of classification [38, 40, 41]. In the present study, we hypothesized that using an ensemble learning algorithm called ECOC can improve the performance of the MI-BCI systems. Based on the results, a slight improvement occurred, compared to the results of the best algorithm used in BCI completion IV, which could confirm the hypothesis. It is hoped that the findings could be regarded as an asset for further improvement of the BCI systems.

Conflicts of interest:

Authors do not claim any conflict of interest.

References

Classification of EEG-based motor imagery BCI by using ECOC


Dietterich T. BG. Solving multiclass learning problems via error-correcting output codes. Journal of Artificial
