J. Integr. Sci. Technol. 2024, 12(5), 814

Journal of Integrated
SCIENCE & TECHNOLOGY

Modelling EEG Signal using multivariate denoising and design of the optimized CSP filter for efficient feature extraction applicable in non-invasive BCI

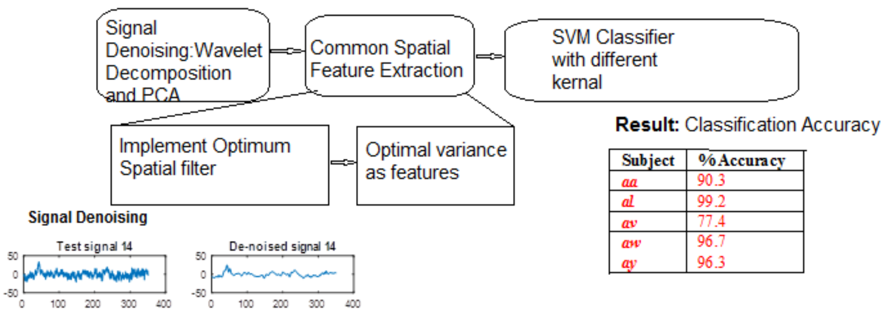
Vrushali G. Raut,[1*] Sandhya A. Shirsat,[1] Rohita P. Patil,[1] Omprakash S. Rajankar,[2] Supriya O. Rajankar[1]

[1]Electronics and Telecommunications Engineering, Sinhgad College of Engineering, Pune, India, [2]Electronics and Telecommunications Engineering, Dhole Patil College of Engineering, Pune, India

Received on: 12-Nov-2023, Accepted and Published on: 15-Feb-2024

Article

ABSTRACT



Electroencephalographic (EEG) signals corresponding to motor imagery (MI) are efficient input to the brain-computer interface (BCI), but are highly contaminated and have a spatial distribution of the activity related variation. This work proposes a filtering method having the combined advantage of wavelet transform and principal component analysis (PCA) for pre-processing of the signal. Transform domain helps to capture the main features of the signals using matching wavelet function while PCA reduces the feature dimension. Correlation structure of noise used here helps in cancelling interferences and to rebuild the signal. Optimized and subject-specific common spatial pattern (CSP) filter design is proposed for extracting the features. Empirical analysis of number of electrodes for building the CSP filter mask leads to selection of 21 electrodes from MI region gives the best performance. The method executes weighing of the electrodes and accordingly assigning the importance to the electrode while forming the filter. Filter induced optimized variance of the signals acts as the features for two-class support vector machine (SVM). Classification accuracy (CA) obtained for subject aa is 90.3%, and for subject al it is 99.2%. Subject aw having small training set gives accuracy to be 96.7% whereas for subject ay it is 96.3%.

Keywords: Multivariate; Common Spatial Pattern (CSP); Brain Computer Interface (BCI); Electroencephalography (EEG)

INTRODUCTION

Non-invasive Brain-computer Interface (BCI) uses the variation of electroencephalographic (EEG) signal for distinguishing the underlying activities. It is the rapidly developing technology as it allows users to command the external environment by using modulation of their brain wave due to their thoughts. It ultimately aims at providing environment control to the individuals suffering from motor loss due to accidents and has to depend on others.[1] Steady state visually evoked potential(SSVEP) creates the potential BCI but has dependency on the stimuli.[2] Movement planning or execution comes under Motor Imagery(MI), and it is one of the preferred choices for BCI as it provides variation in the signals corresponding to the movements.[3] The popularity of MI lies in the number of possible movements providing the corresponding variation in the signals.

Signals collected by EEG are weak and contaminated, but the non-invasive way of collecting the signals is appealing to the researchers.[4] These weak signals need multi-level signal processing for pre-processing the signals, feature extraction, and classification.[5] Volume conducted EEG signals are vulnerable to interference by neighboring EEG signals as well as other electrophysiological signals like electrocardiography (ECG) and electromyography (EMG). According to this logic, noise and artifacts must be eliminated from the signal through pre-processing. Effective pre-processing of EEG signals used for non-invasive BCI plays a significant role in improving final results mentioned in terms of classification accuracy (CA). Literature suggested the surface Laplacian (SL) method which estimates the signal in terms of radial current density at particular electrode sites thus reducing the interference from neighboring electrodes as well

*Corresponding Author: Dr. Vrushali G. Raut, E&TC Department, SCOE, India.
Tel: 9823227558 Email: vrushaliraut4@gmail.com

Cite as: *J. Integr. Sci. Technol., 2024, 12(5), 814.*
URN:NBN:sciencein.jist.2024.v12.814

as other sources.[6,7] SL is a useful method, but its estimate needs an array of the electrode for computation. Another method based on SL technique approximates the signal at the electrode by subtracting the average value of the adjacent channel from the channel of interest[8]. Independent Component Analysis (ICA) also suggested as the pre-processing tool for BCI, shows direct improvement in CA.[9] Improved signal to noise ratio (SNR) can be obtained using common average referencing (CAR) compared to the method of standard ear-reference. Deducting the average value of all the electrode of the montage from the one channel of interest gives reference free signal.[10] Small and large laplacian methods were used along with the referencing methods like CAR for pre-processing of the signal.[11]

In Motor Imagery (MI) based BCI μ (5-15Hz) and β (12-30Hz) are the prominent bands capturing event-related potential. Event-related desynchronization (ERD) of the μ rhythm and event-related synchronization (ERS) of β rhythm serves as a good source for feature extraction. These band limited characteristics of signal suggested frequency domain methods like Fourier transform (FT) for feature extraction. Na Lu et al. (2016) proposes Fourier transform and decomposition by wavelet packet for extracting the features and passes the features to the deep neural network for classification.[12] Literature suggested time-frequency methods such as short time Fourier transform (STFT) as well wavelet transform (WT) for feature extraction due to spanning of MI signals in time as well frequency. Jasmin Kevrica et al. (2017) proposes discrete mode of wavelet transform (DWT), wavelet packet decomposition(WPD) and empirical mode decomposition (EMD) for feature extraction by decomposing EEG signal, whereas classification uses k nearest neighbor algorithm.[13] Xiaofeng Xie et al. (2016) suggested symmetric positive-definite (SPD) covariance matrices for representing distinguishing information of EEG signals for feature extraction whereas preferred algorithm for classification is bilinear sub-manifold learning (BSML).[14] The other trend of literature emphasizes collecting distributed features from the signals on neighboring electrodes suggesting spatial filtering of the signal. Ioannis Xygonakis et al. (2018) offers Spatial features extraction using CSP filters they also suggested selected regions of interest (ROIs) for an individual with Ensemble model for classification.[15] Wei Wu et al. (2015) proposed probabilistic-CSP (P-CSP) for broad EEG spatiotemporal modeling framework. Computationally efficient fisher linear discriminant analysis (FLDA) used as a classifier and demonstrated by applying to classify three MI datasets.[16] Research must be carried out combining efficient method for pre-processing taking cognizance of time-frequency correlation of signal and feature extraction dealing with the spatial domain.

This work proposes rebuilding of the contamination-free EEG signals using the multivariate extension of the wavelet denoising algorithm. Adaptive thresholding of approximate coefficient band using PCA and detail coefficient band using univariate thresholding strategy based on noise covariance matrix helps in reproducing the denoised version of the signal keeping the MI related modulation intact. This study further used the pre-processed signal for implementing optimized CSP filter mask for distinguishing two populations of MI in the input EEG. This filter mask, when applied on the signal, optimizes the difference between the variance for different MI, making them distinguishable. Thus variance as the feature when passed to two class SVM classifier utilizing the non-linear kernel skill provides the best performance in terms of classification accuracy (CA).

METHODS

I. Pre-processing Method

The standard methods for signal denoising and filtering uses algorithms based on wavelet decomposition. On the other hand, PCA is a favorite statistical technique for the reduction of feature dimension in a new lower-dimensional subspace and helps to capture the main features of the signal. This work proposes a combination of wavelet decomposition and PCA for denoising of the signal.

The one-dimensional technique is generalized by the denoising process, which considers the noise's correlation structure. It first applies a basis modification, then applies a conventional one-dimensional soft thresholding [17][18]. By diagonalizing a reliable estimate of the noise covariance matrix provided by the Minimum Covariance Determinant (MCD) estimator based on the matrix of the finest details,[19] the basis change is achieved.

Procedure for multivariate denoising

Consider a signal matrix $X = n \times p$ with p signals from different electrodes and n number of samples where $n \gg p$.

1) Implement wavelet decomposition at the level of J on each column of X.
2) Estimate a noise covariance matrix $\sum_\epsilon$ by application of MCD estimator to detail coefficients at level D1. Resolve $\sum_\epsilon$ into eigenvalues (E) and eigenvector (V) i.e. $\sum_\epsilon = VEV'$.
3) Apply eigenvector matrix (V) obtained in step 2 on detail coefficients of remaining levels to give the change of basis as follows;
   $D_j V$ $(1 \le j \le J)$.
4) Apply p strategies of univariate thresholding on these $D_j V$ applying the threshold $T_i = \sqrt{2\lambda_i \log(n)}$ for the ith column of $D_j V$ where $\lambda_i (1 \le i \le p)$ are diagonal elements of eigenvalue matrix E and n number of samples.
5) The reconstructed final denoising matrix X1 used simplified detail matrices and approximation matrices by changing of basis using V' and taking inverse wavelet transform.
6) Apply PCA on final matrix X1 and select principal components using Kaiser's criteria, selecting the components corresponding to eigenvalues greater than the mean of all the eigenvalues.

II. Feature Extraction

After denoising suggested in section I, the signal without interference is available for feature extraction. It is the array of signals from selected electrodes with the number of samples and each array representing either LHM or RFM. The variations related to these motor movement are distributed spatially in this array, and they even have subject specific variations. Thus feature extraction method has to follow the approach of extracting subject specific patterns for discrimination of underlying movement from high-dimensional spatio-temporal signals[20]. This background suggested the need for common spatial pattern (CSP) approach proven in extracting discriminative spatial pattern. According to probability theory, joint variability of two random variables can be measured by the covariance of the signal. Simultaneous diagonalization of two such covariance matrices corresponding to two distinct movements is used to form an optimum spatial filter[21]. This mask, when applied on the signals under test, provides with the time series having an optimal variance for discriminating two tasks accommodated in the signals.

Matrix for each trial under every class of EEG data represented by $E = NxT$, where N denotes the number of electrodes and T as the sample. Normalized spatial covariance of that matrix can be obtained using equation 1

$$C = \frac{EE'}{trace(EE')} \tag{1}$$

This spatial covariance averaged on training trials for left-hand movement (LHM) provide us with $\bar{C}_l$ for all trails. The covariance over right foot movement (RFM) averaged over the training trials gives $\bar{C}_R$. C_c added the average of covariance $\bar{C}_l$ and $\bar{C}_R$ as in equation 2. Equation 3 computed eigenvalues and eigenvectors from C_c

$$C_c = \overline{C_l} + \overline{C_R} \tag{2}$$

$$C_c = U_c \lambda_c U_c' \tag{3}$$

Linear transformation called sphering or whitening transformation, transforms a signal matrix into a set of new variables whose covariance is the identity matrix, thus un-correlating the variables. When changing test data into an identity covariance matrix for modeling situations, this process is known as statistical whitening. The variance of the data along each dimension equals one when the data have identity covariance, which indicates that all dimensions are statistically independent. Statistical independence represents the joint-complex distribution of the data in a more straight forward manner.

Equation 4 computes the eigenvalues λ_c and eigenvectors U_c, which are used to create the whitening transform matrix P. The eigenvectors are organized in descending order based on the eigenvalues. The variance in the space that U_c spans is equalized by the whitening transformation.

$$P = \sqrt{\lambda^{-1}} U_c' \tag{4}$$

Whitening of C_c i.e. $PC_c P$ will result in unity eigenvalues. Applying this whitening separately on C_l and C_R as in equation 5 and 6 resulted in S_l and S_R respectively.

$$S_l = PC_l P' \tag{5}$$

$$S_R = PC_R P' \tag{6}$$

When S_l and S_R decomposed into eigenvalues and eigenvectors, it can be concluded that they share common eigenvectors and thus the addition of two diagonal matrices belonging to eigenvalues of them will result in identity matrix, i.e. if $S_l = B\lambda_l B$ and $S_R = B\lambda_R B$ then $I = \lambda_l + \lambda_R$, where I is the identity matrix. Since the sum of two corresponding eigenvalues is one, eigenvector with large eigenvalue for S_l will corresponds to small eigenvalue of S_R and vice-versa. This relation helps to conclude that eigenvector matrix B is useful for the classification. Whitening of the matrix B formed the projection matrix W as in equation 7 when W projected on EEG data E it gives Z as in equation 8. The matrix Z representing the decomposition of data matrix E acts as an effective source for features extraction. Further features can be constructed using equation 9, which is the log variance of Z.[22]

$$W = (B'P)' \tag{7}$$

$$Z = WE \tag{8}$$

$$F_p = \log\left[\frac{var(Z_p)}{\sum_{i=1}^{2m} \mathrm{var}(Z_i)}\right] \tag{9}$$

III. *Classification Methods*

Decomposing the signal by applying a spatial filter mask provides the features accommodating variations in variance according to underlying activity. These distributed features in feature space need the machine learning algorithm which will recognize the pattern and classify the feature according to the underlying task. This work uses support vector machine (SVM) a robust classifier working on the idea of building hyperplane to separate data according to their classes. While dealing with higher dimensional feature space, it is required to test for different kernel function. Kernel variations used for this work are quadratic, Gaussian, cubic along with linear kernels.[23,24,25]

MATERIAL AND METHODOLOGY

I. *Database Description and Experimental Setup*

Taken from BCI Competition III, the MI dataset utilized in this work is called dataset IVa. It is made up of signals that represent five subjects moving their left and right hands and feet, respectively. Signal processing in this dataset is hampered by the little amount of training data.[26]

Five healthy volunteers who were seated comfortably provided the data set for this analysis. For 3.5 seconds, they received visual cues telling them to move their left and right hands and feet. Target cues were spaced out by 1.75 to 2.25 seconds to give the individual a chance to unwind.

a. **Format of the Data**

According to the standard 10-20 system electrode montage, continuous signals from 118 EEG channels extract the signals. Markers and 280 cues indicating the time points are provided for each of the five subjects named aa, al, av, aw, ay. Out of available 280 trials for each subject aa and al are provided with training trails 168 and 224 respectively, whereas subject av, aw and ay are provided with small training sets of 84, 56 and 24 trials as shown in table 1. Time slot used for extracting the signal is 3.5s with the sampling frequency of 100Hz.

Table 1 Details of Data-Set

Subject	Training trails	Testing trails
aa	168	112
al	224	56
av	84	196
aw	56	224
ay	28	252

II. Methodology

Database utilized for this work has variability in terms of LHM and RFM and each trial as a matrix of 118x350, with 118 electrodes and 350 samples taken for the time slot of 3.5s with the sampling frequency of 100Hz. Figure 1 gives the flow of the procedure separating the training and testing sessions followed in this work. Subject wise change in training and testing trials are taken as specified in table 1. Out of available 118 electrodes selected 21 electrodes covering the MI region of the brain are used to provide the signal for processing.

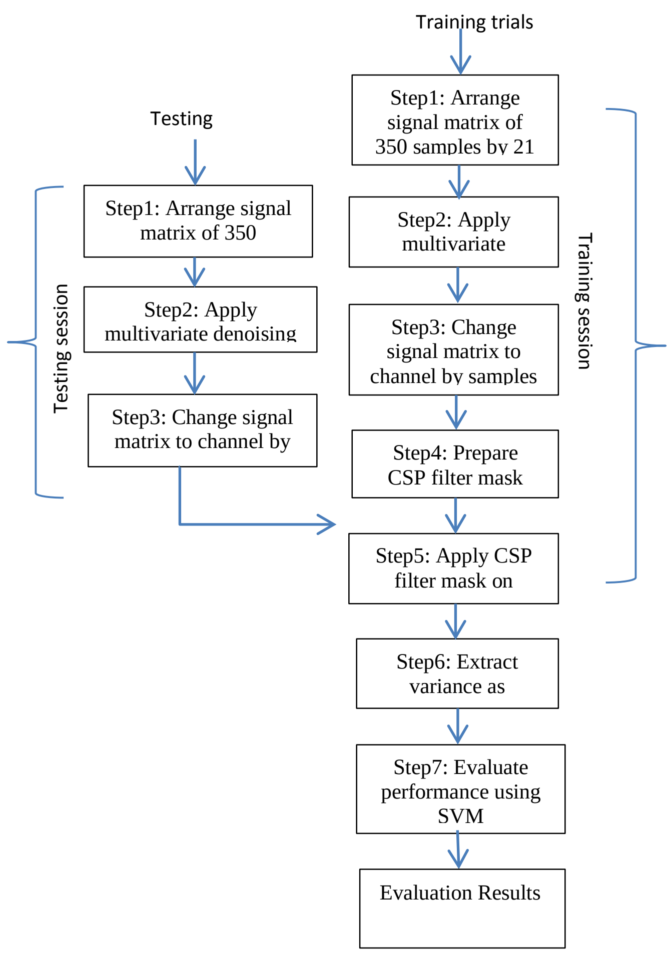


Figure 1. Flow of Analytical Methodology

Multivariate de-noising

Band energy compaction used in previous work for the selection of matching wavelet for EEG signal suggested *db10* and *bior6.8* for the wavelet decomposition of the signal required in the de-noising process.[27] Separation of μ (5-15Hz) and β (12-30Hz) band accommodating ERD and ERS is achieved up to decent extend as shown in figure 2 by using fifth levels of wavelet decomposition. The noise covariance matrix is computed from the detail coefficient (D_1) and decomposed into eigenvalues and eigenvectors.

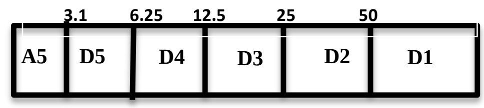


Figure 2 Wavelet Decomposition of Signal with Different Sub-bands and Frequency Ranges in Hertz

Change of basis for other detail coefficient levels D_1-D_5 is carried out by applying computed eigenvector on it to de-correlate p components of noise. Further threshold derived from eigenvalue λ_i of noise covariance matrix and number of samples is applied to the i[th] column of $D_j V$. Above thresholded detail band coefficients and approximate coefficients are used for reconstructing the signal using inverse wavelet transform. The principal components from the reconstructed signals are finally selected using Kaiser criteria to represent the signal. Figure 3 indicates the removal of the noise overriding four different signals from the database using *db10* for wavelet decomposition, and figure 4 displays noise removal using *bior6.8.* One can visually notice the removal of overriding high frequency noise from the signals, but CA obtained will be the proper criteria to judge the noise removal and maintaining task related variation.

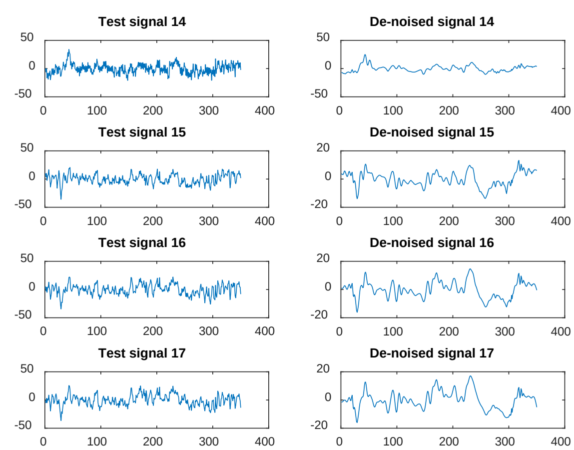


Figure 3. Comparison of Original and Denoised Signals using *db10* wavelet for decomposition

Common Spatial Filtering

Signals corresponding to 21 electrodes from motor imagery area of the brain are empirically selected to construct the spatial filter. Iteration for the different number of electrodes as well the combination of the electrode tested for the formation of CSP filter and the combination providing more CA is selected. Subject specific training trials mentioned in table 1 for selected 21 electrodes are used to construct the CSP filter.

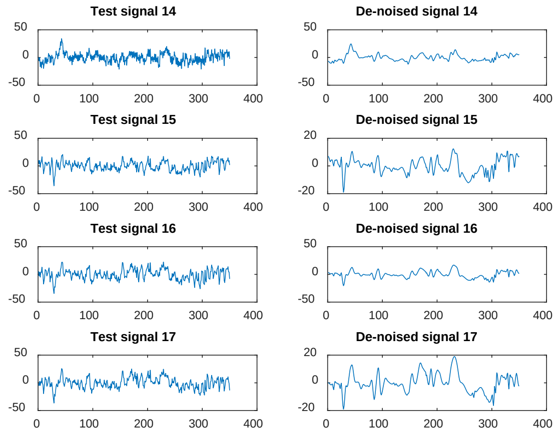


Figure 4. Comparison Original and Denoised Signals using *bior6.8* Wavelet for Decomposition

The designed filter mask, when applied on data, gives the matrix of 21x350 corresponding to particular movement and having an optimum variation for that movement. Out of 21 rows, eight are selected to calculate variance and passed as features to the classifier. Selected rows provide an optimum variance for the MI to differentiate.

Classification

The two-class support vector machine, with different kernel functions, is used for classification in this work. Training session utilizes the subject-specific trials mentioned in table 1. Five-fold cross-validation is done for the classifier using the available training trials by randomly partitioning the features into five subsets. Four subsets were used for training purpose whereas remaining one for classifier testing, and this process was repeated five times in such a manner that the classifier tested every subset. Classification accuracy(CA) is the most important evaluation parameter considered for BCI. It is given as

$$CA = \frac{\text{Number of Correct Classifications}}{\text{Total Number of Motor Imagery}} \times 100$$

RESULTS

Classification Accuracy (CA)

From BCI point of view as the environment has to be controlled by correctly guessing the underlying task, CA is the most important evaluation parameter, which represents the correct classifications. Results in table 2 & 3 displayed the difference as well improvement in task CA for the pre-processed signal. When analyzed for subject *aa* and *al* referring table 2, it can be concluded that denoising shows the visible increment in CA whereas decomposition using *bior6.8* and *db10* in denoising process boost the accuracy. Referring to table 3 for subject *av*, *aw*, and *ay*, it can be stated that the denoised version leads to more CA when *db10* is selected compared to *bior6.8*. Subject *aw*, *av*, and *ay* provides CA above 90% though having small training data. From the comparison of accuracy within the subjects (Figures 5 and 6), it can be concluded that subject al gives the highest accuracy for wavelet *bior 6.8*. as well as for *db10*.

Table 2. Percent classification accuracy for subject *aa*

Classification Accuracy for Subject aa			
	Pre-processed signal		
SVM kernel	**bior6.8**	**db10**	**Without pre-processing**
Quadratic	80.4	86.3	59.5
Cubic	**90.3**	78.6	57.1
Med. Gaussian	80.4	83.9	56.5
Linear	82.1	85.7	54.2
Cor. Gaussian	81	81	55.4

Table 3 Percent classification accuracy for subject *al*

Classification Accuracy for Subject *al*			
	Pre-processed signal		
SVM kernel	**bior6.8**	**db10**	**Without pre-processing**
Quadratic	91.5	92.4	61.8
Cubic	89.3	92	58.2
Med. Gaussian	92.4	92	63.9
Linear	**92.9**	92.4	60.4
Cor. Gaussian	95	99.2	62.1

Table 4 Percent classification accuracy for subject *av*

Classification Accuracy for Subject *av*			
	Pre-processed signal		
SVM kernel	**bior6.8**	**db10**	**Without pre-processing**
Quadratic	60.7	73.8	61.4
Cubic	53.6	71.4	53.2
Med. Gaussian	67.9	69	60.7
Linear	59.5	**77.4**	60
Cor. Gaussian	60.7	73.8	61.4

Table 5 Percent classification accuracy for subject *aw*

Classification Accuracy for Subject *aw*			
	Pre-processed signal		
SVM kernel	**bior6.8**	**db10**	**Without pre-processing**
Quadratic	77.5	89.3	76.1
Cubic	71.1	87.5	71.4
Med. Gaussian	77.9	92.9	76.1
Linear	80	**96.7**	76.8
Cor. Gaussian	77.5	89.3	76.1

Table 6 Percent classification accuracy for subject *ay*

Classification Accuracy for Subject *ay*			
	Pre-processed signal		
SVM kernel	**bior6.8**	**db10**	**Without pre-processing**
Quadratic	88.5	92.9	79.6
Cubic	88.5	92.1	72.1
Med. Gaussian	88.5	93.2	78.9
Linear	91	**96.3**	80
Cor. Gaussian	88.5	92.9	79.6

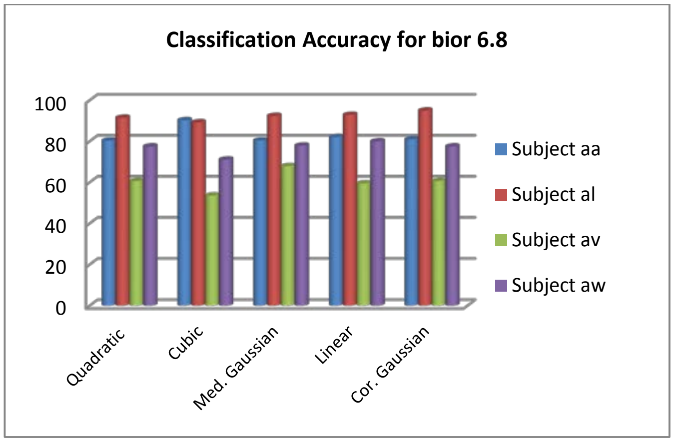


Figure 5. Comparison of Classification Accuracy for *bior6.8* wavelet for decomposition

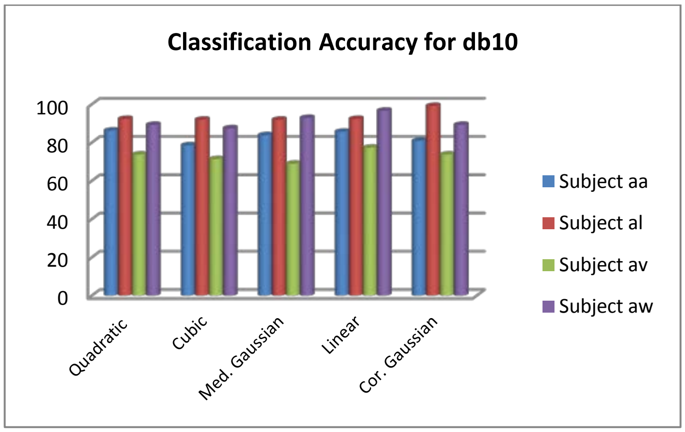


Figure 6. Comparison of Classification Accuracy for *db10* Wavelet for Decomposition

Discussion

In Brain-computer Interface Electroencephalographic (EEG) signals can be effectively used for classification of the MI task. However, the vital requirement is appropriate processing of the signals as the modulation due to MI is trivial to get hidden by the overriding noise. This paper emphasizes the concerns of noise removal by suggesting a multivariate denoising method. This work selected wavelet db10 and bior6.8 and five levels of decomposition for wavelet transform of the signals. The eigenvector obtained from noise covariance matrix estimated from detail coefficient band D1 carrying finest details of the signal is applied on all detail coefficient matrix to give change of basis. Soft thresholding using the eigenvalues obtained in previous step is the next strategy to be applied on the obtained coefficients. The thresholded detail coefficients and approximate coefficients are used to reconstruct the signal by inverse laplace transform. Further the Kaisers criteria is used to select principal components from the reconstructed signals. The efficiency of this filtered signal is proved when it responded in terms of improved CA by almost 30%. The significant contribution of this work is optimized Common Spatial Pattern filtering for feature extraction. Out of 118 electrodes, 21 area specific electrodes proposed in this work can form a capable filter. These electrodes are selected empirically for forming the filter. This work proficiently handles the challenge of small training set provided by this database. Results obtained by implementing the methodology suggested in this work gives comparable results with the state of art methods with the reduction in electrodes.

Conclusion

This work leads to the conclusion that the strong pre-processing method taking care of spatial spread of the signal is the key requirement of EEG based BCI. This paper suggested the efficient option which works on combination of wavelet transform and PCA. The optimized feature extraction strategy based on CSP filter design provided in this work further helps in effective classification of MI task. SVM with different kernel function is suggested as a strong classifier in this work which can be tuned according to the subject. Though the subject *al* having largest training data offers maximum CA of 99.2%, other subjects *aw* and *ay* also gives comparably good CA. CA as a performance parameter concluded as subject-specific, a subject can be trained to improve the MI modulation resulting in high accuracy no matters the number of trainings. Other important inference from the results is the dependency of accuracy on wavelet selection which is verified empirically for every subject.

Acknowledgments

Database utilized in this work is dataset IVa from BCI competition III and accordingly mentioned and cited in the paper. No funding has been utilized for this work.

Conflict of Interest

The authors declare that there is no academic or financial conflict of intetest for this work.

References and Notes

1. S.G. Mason, G.E. Birch. A brain-controlled switch for asynchronous control applications. *IEEE Trans. Biomed. Eng.* **2000**, 47 (10), 1297–1307.
2. X. Yuan, L. Zhang, Q. Sun, X. Lin, C. Li. A novel command generation method for SSVEP-based BCI by introducing SSVEP blocking response. *Comput. Biol. Med.* **2022**, 146, 105521.
3. G. Pfurtscheller, F.H. Lopes Da Silva. Event-related EEG/MEG synchronization and desynchronization: Basic principles. *Clinical Neurophysiology*. 1999, pp 1842–1857.
4. S. Borde, V. Ratnaparkhe. Optimization in channel selection for EEG signal analysis of Sleep Disorder subjects. *J. Integr. Sci. Technol.* **2023**, 11 (3), 527.
5. S.G. Mason, G.E. Birch. A general framework for brain - Computer interface design. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*. 2003, pp 70–85.
6. J. Kayser, C.E. Tenke. Issues and considerations for using the scalp surface Laplacian in EEG/ERP research: A tutorial review. *International Journal of Psychophysiology*. Elsevier B.V. 2015, pp 189–209.
7. J. Kayser, C.E. Tenke. Principal components analysis of Laplacian waveforms as a generic method for identifying ERP generator patterns: II. Adequacy of low-density estimates. *Clin. Neurophysiol.* **2006**, 117 (2), 369–380.
8. C. Carvalhaes, J.A. De Barros. The surface Laplacian technique in EEG: Theory and methods. *Int. J. Psychophysiol.* **2015**, 97 (3), 174–188.
9. I. Rejer, P. Górski. Independent component analysis for EEG data preprocessing - Algorithms comparison. In *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*; **2013**; Vol. 8104 LNCS, pp 108–119.

10. M. Cheng, W. Jia, X. Gao, S. Gao, F. Yang. Mu rhythm-based cursor control : an offline analysis. *Clin. Neurophysiol.* , *Elsevier* **2004**, 115, 745–751.
11. D.J. McFarland, L.M. McCane, S. V David, J.R. Wolpaw. Spatial filter selection for EEG-based communication. *Electroencephalogr. Clin. Neurophysiol.* **1997**, 103 (3), 386–394.
12. N. Lu, T. Li, X. Ren, H. Miao. A Deep Learning Scheme for Motor Imagery Classification based on Restricted Boltzmann Machines. *IEEE Trans. Neural Syst. Rehabil. Eng.* **2017**, 25 (6), 566–576.
13. J. Kevric, A. Subasi. Comparison of signal decomposition methods in classification of EEG signals for motor-imagery BCI system. *Biomed. Signal Process. Control* **2017**, 31, 398–406.
14. X. Xie, Z.L. Yu, H. Lu, Z. Gu, Y. Li. Motor Imagery Classification Based on Bilinear Sub-Manifold Learning of Symmetric Positive-Definite Matrices. *IEEE Trans. Neural Syst. Rehabil. Eng.* **2017**, 25 (6), 504–516.
15. I. Xygonakis, A. Athanasiou, N. Pandria, D. Kugiumtzis, P.D. Bamidis. Decoding Motor Imagery through Common Spatial Pattern Filters at the EEG Source Space. *Comput. Intell. Neurosci.* **2018**, 2018 (Mi).
16. W. Wu, Z. Chen, X. Gao, et al. Probabilistic common spatial patterns for multichannel EEG analysis. *IEEE Trans. Pattern Anal. Mach. Intell.* **2015**, 37 (3), 639–653.
17. D.L. Donoho. De-Noising by Soft-Thresholding. *IEEE Trans. Inf. Theory* **1995**, 41 (3), 613–627.
18. S.Y. Wang, X. Liu, J. Yianni, T.Z. Aziz, J.F. Stein. Extracting burst and tonic components from surface electromyograms in dystonia using adaptive wavelet shrinkage. *J. Neurosci. Methods* **2004**, 139 (2), 177–184.
19. M. Aminghafari, N. Cheze, J.M. Poggi. Multivariate denoising using wavelets and principal component analysis. *Comput. Stat. Data Anal.* **2006**, 50 (9), 2381–2398.
20. J. Müller-Gerking, G. Pfurtscheller, H. Flyvbjerg. Designing optimal spatial filters for single-trial EEG classification in a movement task. *Clin. Neurophysiol.* **1999**, 110 (5), 787–798.
21. K. Fukunaga. Introduction to Statistical Pattern Recognition, 2nd ed.; New York, Academic, 1990, **1990**.
22. H. Ramoser, J. Müller-Gerking, G. Pfurtscheller. Optimal spatial filtering of single trial EEG during imagined hand movement. *IEEE Trans. Rehabil. Eng.* **2000**, 8 (4), 441–446.
23. A.N. Ali. Simple Features Generation Method for SVM based Iris Classification. In *2013 IEEE International Conference on Control System, Computing and Engineering*; **2013**; pp 238–242.
24. K. Liao, R. Xiao, J. Gonzalez, L. Ding. Decoding individual finger movements from one hand using human EEG signals. *PLoS One* **2014**, 9 (1), 1–12.
25. B. Blankertz, G. Curio, K.-R. Müller. Classifying single trial EEG: Towards brain computer interfacing. *Adv. Neural Inf. Process. Syst. Conf.* **2002**, 1 (c), 157–164.
26. B. Blankertz, K.R. Müller, D.J. Krusienski, et al. The BCI competition III: Validating alternative approaches to actual BCI problems. In *IEEE Transactions on Neural Systems and Rehabilitation Engineering*; **2006**; Vol. 14, pp 153–159.
27. V.G. Raut, S.R. Ganorkar. Localized and Optimized Feature Extraction from a Sculpted Electroencephalographic Signal for Better Classification in Motor Imagery Based Brain Computer Interface. *J. Comput. Theor. Nanosci. Localized* **2018**, 15, 1–5.