

Adaptive Error Detection Method for P300-based Spelling Using Riemannian Geometry

Attaullah Sahito, M. Abdul Rahman, Jamil Ahmed
Department of Computer Science
Sukkur Institute of Business Administration
Airport road Sukkur, Pakistan

Abstract—Brain-Computer Interface (BCI) systems have become one of the valuable research area of ML (Machine Learning) and AI based techniques have brought significant change in traditional diagnostic systems of medical diagnosis. Specially; Electroencephalogram (EEG), which is measured electrical activity of the brain and ionic current in neurons is result of these activities. A brain-computer interface (BCI) system uses these EEG signals to facilitate humans in different ways. P300 signal is one of the most important and vastly studied EEG phenomenon that has been studied in Brain Computer Interface domain. For instance, P300 signal can be used in BCI to translate the subject's intention from mere thoughts using brain waves into actual commands, which can eventually be used to control different electro mechanical devices and artificial human body parts. Since low Signal-to-Noise-Ratio (SNR) in P300 is one of the major challenge because concurrently ongoing heterogeneous activities and artifacts of brain creates lots of challenges for doctors to understand the human intentions. In order to address above stated challenge this research proposes a system so called Adaptive Error Detection method for P300-Based Spelling using Riemannian Geometry, the system comprises of three main steps, in first step raw signal is cleaned by preprocessing. In second step most relevant features are extracted using xDAWN spatial filtering along with covariance matrices for handling high dimensional data and in final step elastic net classification algorithm is applied after converting from Riemannian manifold to Euclidean space using tangent space mapping. Results obtained by proposed method are comparable to state-of-the-art methods, as they decrease time drastically; as results suggest six times decrease in time and perform better during the inter-session and inter-subject variability.

Keywords—Brain Computer Interface; EEG; P300; Riemannian geometry; xDAWN; Covariances; Tangent Space; Elastic net

I. INTRODUCTION

In recent ML (Machine Learning) techniques have been vastly used to solve the various medical classification problems such as Liver, Heart, Neurological disorders and others. Particularly; human intuition interpretation using P-300 signals from EEG (Electroencephalogram) have become one of the significant problems, since BCI (Brain Computer Interface) establishes point of communication between doctors and patients where the subject is found with complete paralyzed situation. For example, locked-in syndrome paralysis is a condition in which patient is found awake and fully aware but no expression could be made by patients to communicate with doctors. The patients with spinal cord injury leading to ALS (Amyotrophic Lateral Sclerosis) so called neuro-degenerative disease in which loss of voluntary control muscles produces

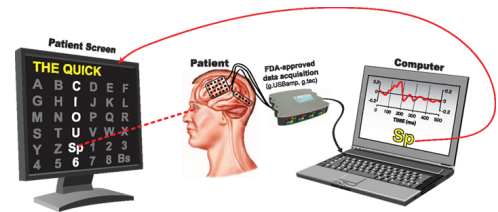


Fig. 1. P300 Speller

human paralysis and patient become unable to communicate. In these conditions Brain Computer Interface can help in controlling of external devices or as a reliable tool for communication. Electrical activity of neurons in brain can be recorded using Electroencephalogram (EEG). This EEG data is of great use: neurologists can use for diagnosing neurological disorders. BCI experts can use for means of communication. Nowadays BCI has become one of assistive tool for human beings, it can control wheel chairs, video games, entertainment, spellers, and other assistive technologies [1]. In order to address above stated challenge this paper proposes a system called Adaptive Error Detection method for P300-Based Spelling using Riemannian Geometry, the system comprises of three main steps, in first step raw signal is cleaned by preprocessing. In second step most relevant features are extracted using xDAWN spatial filtering along with covariance matrices for handling high dimensional data and in final step elastic net classification algorithm is applied after converting from Riemannian manifold to Euclidean space using tangent space mapping. Since BCI is vastly used for controlling external devices and as a reliable tool for communication. In these circumstances as stated above; BCI can be used to establish a reliable communication channel directly from patients brain signals to the computer [2], but low Signal-to-Noise Ratio (SNR) in P300 is one of the major challenges because concurrently ongoing heterogeneous activities and artifacts of brain create lots of difficulties for doctors to understand the human intentions.

P300 signal is one of the vastly explored non-invasive Brain Computer Interface scenario. Response of P300 signal can be considered as an oddball scenario, where low probability of desired events is mixed with high probability of undesired events. An ERP or event-related-potential can be described as recorded P300 response during the course of decision making when subject reacts to different kinds of stimuli. P300 [3] speller has 6X6 matrix of characters containing alpha-numeric characters as shown here in Figure. 1.

Different kinds of stimuli can be explored such as auditory,

visual etc., however P300 speller uses visual stimuli [4] only. In P300 Speller scenario a cap containing electrodes, is placed on scalp of subject. These electrodes record brain waves data produced against a visual stimuli of letter displayed on screen. Raw EEG signal is inherently very noisy; eventually, it has very low SNR. Therefore, it is very hard to differentiate true P300 response from recorded EEG signal. Low SNR is due to different ongoing electrical activities of neurons in brain while performing other physiological tasks simultaneously. This noise poses a challenge of extracting true ERP and interpreting brain activity correctly. Problem is to detect errors during spelling task made by P300 speller, given subject's brain waves. Machine Learning (ML) algorithms play crucial role in formulating Brain Computer Interface systems. For example, ML algorithms can help in extracting influential and most relevant features from data, because recorded signals are inherently noisy due to many reasons. Feature extraction can be helpful in removing artifacts, noise and other undesired factors from recorded data. Classification algorithm is used to set learning parameters on training data and employed on test data. In this scenario, goal of classifier is to predict based on training data; whether P300 speller has recognized Target letter correctly or not according to subject's intention. In this paper, we are proposing a ML pipeline for classification of EEG data from P300 speller having robust and reliable feature extraction method and classification algorithm, which minimizes misinterpreted classified commands against user's actual desired intentions and provide reliable means of communication. For this objective, BCI system must possess two key properties:

- 1) Better classification technique for P300 Speller which enhances across-subject generalization and across-session generalization, having optimal accuracy.
- 2) Fast convergence to optimal parameters requiring minimal calibration and training data.

This paper is organized in five sections. Section one is used to describe the introduction of this paper, section two describes the related works, section three defines the materials and methods, section four represents the results and in section five conclusion and discussion is discussed.

A. Related Work

Classification of P300 speller is of one the active research area(s) of ML (Machine Learning) and many approaches have been proposed to solve the classification problem. Our approach deals as productive modelling for P300 classification in machine learning domain. Some of the related works have been presented as below.

For P300 classification, comparison of different classifier algorithms is presented in [6], between different classifiers such as stepwise linear discriminant analysis (SWLDA), support vector machines (SVMs) and LDA (Linear Discriminant Analysis). Results suggest that SWLDA performs better than other classifiers. Although SWLDA works as simple LDA, but at start SWLDA has no feature in discrimination function; it starts adding features one by one using their statistical significance from p-value. In this paper only 16 electrodes were used. This paper proposes a scalable system, which can incorporate as many as 56 electrodes.

A system [7] was proposed to classify ERP data from P300 experiment with optimal accuracy. Spatial as well as spatio-temporal features are used along with Linear Discriminant Analysis with shrinkage as classifier and proposed an analytical method for estimating regularization parameter for LDA. Then comparison between different classifiers such as SWLDA [6] and simple LDA is presented. Results suggest their model performs better than other two. They used 55 channels but for only 7-time sample points after occurrence of an event. Epoch window size of 7-time sample points is very small; it is likely that important temporal information for whole trial can be missed. This paper proposes a scalable system, which can incorporate as many as 260 time points, maintaining full spatial as well as temporal information.

Comparison of different feature selection and classification methods is presented by [8] using five different datasets from Brain Computer Interface paradigm. For model building; spectral, spatial and spatio-temporal filtering as feature selection was applied and for classification four methods LDA, SWLDA, rLDA (regularized LDA) and rLR (regularized logistic regression) were used. Results suggest that regularized classifiers are better in terms of performance. This suggested approach requires a lot of manual parameter tuning and results are not generalizable across sessions and across users.

System based on Riemannian geometry was firstly introduced by Congedo et al. [9] for classification of BCI systems. Data from all three modalities of BCI Motion Imagery (MI), Steady State Visually Evoked Potential (SSVEP) and ERP from 22 electrodes. They proposed a classifier known as MDM (Minimum Distance to Mean) classifier, which works on principle of Riemannian geometry. This classifier works by calculating covariance mean matrices for each class in training data as a representative and assigns the label to test data by calculating their distances from mean covariance matrices of classes. For comparison they used Common Spatial Patterns+ [10] + Linear Discriminant Analysis [11]. Results suggest that, Riemannian geometry based classifier Minimum Distance to Mean (MDM) works better and requires small amount of data for training and also works better in across-subject and across-session generalization. As robust and efficient classification algorithms like Neural network, SVM, etc. work in euclidean space; so cannot be applied to covariance matrices belonging to riemannian manifold directly, this is main limitation of this approach.

Feature selection technique known as xDAWN based on spatial filtering was introduced by [12]. Main motivation behind this technique was to discriminate and enhance P300 evoked response. This technique works on assumption that P300 response related to an ERP is least frequent and target response occurs simultaneously, has very little spatial subspace span in the recorded signal. This algorithm is based on QR factorization of matrices [13]. For comparison, 3 subjects data having 29 electrodes were used. Results comparison was done between different feature extraction techniques such as xDAWN spatial filtering, ICA and PCA, and classification using Bayesian Linear Discriminant Analysis (BLDA). Results suggest that xDAWN spatial filtering technique has brought significant improvement on accuracy as compared with ICA and PCA. However, if xDAWN components are increased, then performance slightly decreases and computational complexity

increases, this is main limitation of algorithm.

As discussed earlier famous classification techniques works in euclidean space and hence cannot be applied directly in Riemannian Manifolds. To overcome this limitation [14],[15] introduced Tangent Space mapping (TS). Tangent Space bridges euclidean space and Riemannian manifolds. Tangent space mapping projects covariance matrices belonging to Riemann manifold into Euclidean space vectors. By using this mapping, one can use classical and efficient classifiers such as LDA, SVM etc. on covariance matrices directly, instead of using MDM [9]. Results suggest that, significant improvement can be achieved; without need of spatial filtering of electrodes as they are exploited in their native space. Although, from BCI; Motor Imagery (MI) was considered in these both papers [14],[15], though this technique can be extended for classification of P300 signal.

All above stated approaches have performed better, but some techniques work well on one data set while perform poorly on other data set. However, [5],[9], [16] proposed a system that uses Riemannian based covariance matrices as features along with other additional meta features for instance word length, feedback detail, etc. to enhance the prediction performance. These newly introduced techniques has brought significant improvement for classification, but also brought adverse effect on computational complexity of the classifier. Hence, P300 classification system direly needs a novel approach which can minimize computational complexity along with consideration of optimal features towards time accuracy trade-off.

II. MATERIAL AND METHODS

Proposed approach is based on modelling machine learning pipeline, which deals with the classification problem of the P300 speller; whether an event is erroneous or not. Our proposed methodology as graphically shown in Figure. 2, is divided into four interconnected steps where each step is assigned several inputs to interact with each other. In first step; raw data is cleaned using simple preprocessing techniques, in second step BCI related features are extracted from cleaned data using xDAWN spatial filters along with covariance matrices and tangent space mapping. In third step classification model is constructed by using elastic net and in fourth step the classification accuracy have been measured. These steps are discussed in detail below.

A. Preprocessing

In preprocessing step, EOG(ElectroOculoGram) channel is removed which gives information about the noise caused by blinking of subjects eye. 5th order butter-worth filter is employed to bandpass rest EEG channels between 1 to 40 Hz. Butter-worth filter is maximally flat magnitude filter, which have response in its passband. For each trial each epoch window is set to 1.3 seconds after the occurrence of feedback event. After preprocessing of raw signals, features are extracted using below defined ML pipeline.

B. Feature Extraction

Figure. 2 gives an overview of all preprocessing steps, feature extraction techniques and model building steps. Our

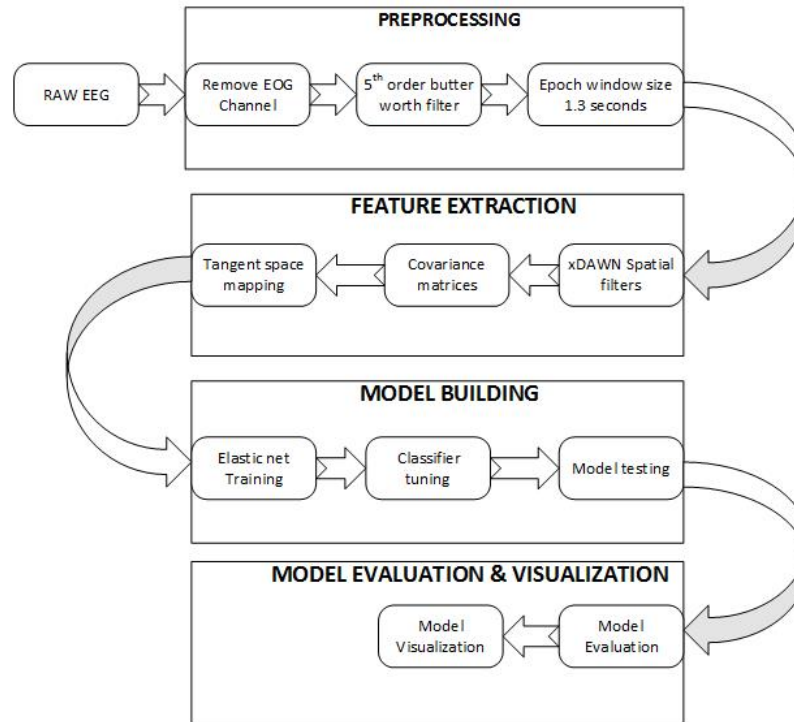


Fig. 2. Adaptive error detection system for P300 speller workflow

feature extraction step comprises of three interconnected sub steps Xdown spatial filters, covariances matrices and tangent space mapping. All these are discussed in detail as under.

1) xDAWN: As discussed earlier, brain waves recorded through EEG are noisy, this recorded EEG signal has information about true P300 signal as well as other ongoing background activities of the brain such as muscular artifacts, eye blinking etc. Therefore, acquired EEG data has very much noise and this low SNR makes classification task challenging. Different techniques have been investigated in literature to improve EEG signal, spatial filtering one of them. For instance, independent component analysis (ICA) was used [17] to enhance SNR. Limitation of such techniques are that, they are not designed specifically for Brain Computer Interface. To overcome this limitation [12] an unsupervised spatial filtering technique known as xDAWN was introduced. This technique works on assumption: P300 related signal is less frequent and has smaller synchronous subspace in the recorded EEG signal. These synchronous responses for each channel are calculated and then spatial filters are calculated using these responses, in order to make true P300 signals stand out from other artifacts and the noise.

Let us $X \in R^{N_t \times N_s}$ denotes actual recorded raw EEG signals. In which each entry $(i, j)^{th}$ is $x_j(i)$, which represents data from j^{th} channel at instant of time i . Here N_s represents index of channel and N_t represents time index of a trail. Let us assume $a_j(t)$ denote the Event Related Potential signal from j^{th} electrode channel at time course of index t for that trail, whereas $A \in R^{N_e \times N_s}$ denotes Event Related Potential signal, where $(i, j)^{th}$ value corresponds to $a_j(i)$. Here N_s represents index of channel and N_e is index of time point for Event Related Potential which here corresponds to 260 which is given

by single trial duration of 1.3 seconds. Recorded signal of EEG can be formulated as combination of Event Related Potential A and Noise N:

$$X = DA + N \quad (1)$$

Where D represents Toeplitz matrix $D \in R^{N_e \times N_s}$. Solution of (1) is given by QR factorization [13].

2) *Covariances*: As discussed above Signal to noise Ratio (SNR) can be enhanced by spatial filtering for better design of Brain Computer Interface systems, but it requires substantial amount of data for training purpose, therefore for BCI scenario more repeated trials from the subject are needed, this makes spatial filtering solely unfeasible choice for on-line scenario. For BCI systems, covariance matrices for feature extraction were introduced by [15], further [16] improved covariance matrices by including temporal information as well. This all has been possible due to recent theoretical advancements in field of Information geometry [18]. The information geometry is a field of mathematics, which takes probability distributions as points of a Riemannian manifold (Manifold has resemblance to homomorphic euclidean space near each point.). Nowadays it is widely applied in various fields. For example, processing of radar signals [19], diffusion tensor [20] and digital image processing [21]. Processing the covariance matrices in their native manifold has added benefit. Sample covariance matrix (SCM) also known as prototyped ERP response matrix is calculated for each class by taking average of all epochs belonging to same class. Let $P_1 \in R^{s \times t}$ be the sample covariance matrix (SCM) for class 1.

$$P_1 = \frac{1}{I} \sum_{i \in I} X_i \quad (2)$$

Where I is the indexes of i^{th} trail. For each trial x_i , modified trail \tilde{x}_i is given by combining it with SCM:

$$\tilde{x}_i = \begin{bmatrix} P_1 \\ X_i \end{bmatrix}$$

Final covariance matrices are computed using SCM:

$$\sum_{\tilde{i}} i = \frac{1}{N-1} \tilde{X}_i \tilde{X}_i^T \quad (3)$$

The resultant covariance matrix of each epoch is concatenated with xDAWN spatial filters and passed to next step. One can directly only apply Minimum Distance to Mean (MDM) classifier on covariance matrices [9], [16] because they belong to Riemannian manifold of symmetric positive definite matrices (SPD) [22].

3) *Tangent Space*: As discussed earlier [9,16] that we cannot apply classical classification algorithms to covariance matrices without modification, because they are from riemannian manifold. Although [15] used support vector machine (SVM) by modifying kernel, but this is not an obvious choice. Tangent space mapping helps us in using robust and efficient classifiers for instance svm, elastic net etc. easily by converting riemannian covariance matrices to euclidean space. Tangent vectors for euclidean space are calculated using tangent space mapping on covariance matrices belonging to riemannian manifold. Process of tangent space mapping is geometrically shown in Figure. 3.

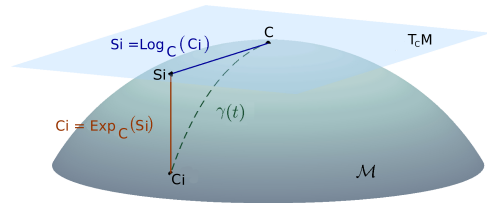


Fig. 3. Tangent space TcM corresponds to manifold M at point C. Si tangent vectors are estimated using Logarithmic map $\text{Log}_C(\cdot)$. The Exponential-map $\text{Exp}_C(\cdot)$ is used for invers mapping [15].

Logarithmic mapping is used to calculate tangent vectors S_i of tangent space TcM, for each point C (which here is actually covariance matrices) of manifold M as shown in Figure. 3. Tangent vector space TcM for that point in euclidean space, which is locally homomorphic to manifold can be calculated using:

$$S_i = \text{Log}_C(C_i) = C^{\frac{1}{2}} \log(C^{-\frac{1}{2}} C_i C^{-\frac{1}{2}}) C^{\frac{1}{2}} \quad (4)$$

C. Model building

Different classifier models were used, but elastic net performed better than others. Elastic net learning algorithm is widely used to predict in different classification problems. It is one of the linear regularized regression algorithm, which addresses shortcomings of lasso as well as ridge regression by introducing l1 and l2 penalty [23] and also handles numerical data very well. After feature extraction, elastic net is trained on training set comprises of 5440 trails, after tangent space mapping while $p = 2211$ number of predictors, class labels of training data is represented by $Y = (y_1, y_2, y_3, \dots, y_n)^T$ and $X = [X_1] \dots [X_p]$ be model matrix where $x_j = (x_{1j}, \dots, x_{nj})^T$. Let λ_1 and λ_2 be the fixed non-negative integers, then objective function for elastic net algorithm in terms of minimization can be defined as:

$$L(\lambda_1, \lambda_2, \beta) = |Y - X\beta|^2 + \lambda_2 |\beta|^2 + \lambda_1 |\beta|_1 \quad (5)$$

For training and testing purpose Elastic net is used. For tuning of hyper parameters of elastic net were tuned by cross validation. Prediction is done using conventional linear regression equation as:

$$\hat{y} = \beta_0 + x_1 \beta_1 + \dots + x_p \beta_p \quad (6)$$

D. Model evaluation

In performance evaluation, area under the ROC (Receiver Operating Characteristic) curve (AUC) is used, as it has already been used widely for binary classification tasks [24]. AUC is calculated using:

$$AUC = \frac{1}{N_+ + N_-} \sum_{i=1}^{N_+} \sum_{j=1}^{N_-} I(x_i > y_j) \quad (7)$$

Where N_+ denotes frequency of instances belonging to positive class and N_- denotes frequency of instances belonging to negative class. While x_1, \dots, x_{N_+} are probability scores predicted by model for N_+ positives and y_1, \dots, y_{N_-} denotes probability scores predicted by model for N_- negative class.

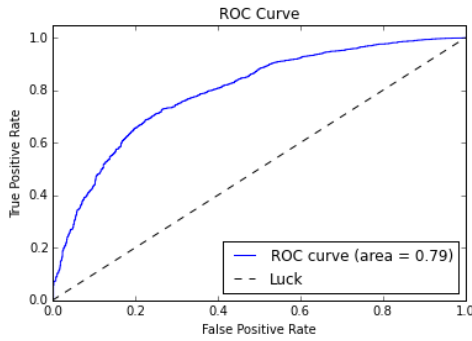


Fig. 4. ROC Curve estimated for P300 classification using proposed system

On other hand values of x_i and y_j are normalized between (0, 1).

For model building and evaluation Data set is used from an experiment [25]. In this experiment, brain activity was recorded with total 56 passive EEG electrode channels. Data was collected from 26 subjects (16 training subjects and 10 test subjects) during 5 different spelling sessions. All subjects participated in five sessions, only 12 five-letter words were spelt in first four sessions by subject and in last session 20 five-letter words were spelt. For each letter, trail is lasted for 1.3 seconds. All subjects in these five sessions constitute total training 5440 trials and 3400 test trials.

III. RESULTS

In figure. 4 presents the results of fully automated process after applying proposed ML pipeline using ROC curve having 0.79 AUC (Area under curve). While performance for different users across different sessions is shown in [Table 1] in terms of AUC. Figure. 5 compares the AUC obtained by proposed method with state-of-the-art and award winning method by Alexandre [5], which shows almost equivalent AUC obtained by both methods, this is also supported by figure. 7; where samples of size 200, 500, 800 and 1200 were used from test data and AUC is compared for both models. Another comparison is performed by considering time as a variable. Figure. 6 shows comparison between time taken by proposed approach and Alexandres [5] approach, which shows huge difference in time taken by proposed approach and Alexandres approach; on average proposed approach takes 37 seconds while Alexandres [5] approach took 224 seconds. This result is also supported by results presented in figure. 8; where time taken by both approaches for samples of size 200, 500, 800 and 1200 from test data is compared.

IV. CONCLUSION

Due to low Signal to Noise Ratio and high dimensional nature of EEG data, classification of P300 signal is one of the challenging problem. Although various approaches have been reported in the literature for classification of P300 signals based on feature extraction methods but most of them uses spatial filters [11] and exploits riemannian geometry [9] but classification of P300 data requires significant method to preprocess. This paper proposes a machine learning pipeline from preprocessing to construction of classification model for

TABLE I. OVERALL PERFORMANCE OF PROPOSED METHODOLOGY

Session→	1	2	3	4	5
Subject 1	0.79	0.80	0.81	0.66	0.56
Subject 3	0.84	0.88	0.69	0.75	0.67
Subject 4	0.80	0.94	0.91	0.90	0.80
Subject 5	0.61	0.70	0.65	0.52	0.65
Subject 8	0.91	0.90	0.90	0.86	0.75
Subject 9	0.63	0.75	0.71	0.87	0.75
Subject 10	0.90	0.96	0.90	0.93	0.96
Subject 15	0.94	0.79	0.76	0.99	0.82
Subject 19	0.64	0.58	0.63	0.65	0.69
Subject 25	0.83	0.81	0.64	0.72	0.65

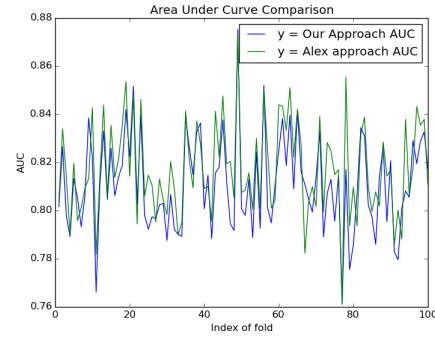


Fig. 5. Comparing Area Under Curve with Alexandre Model

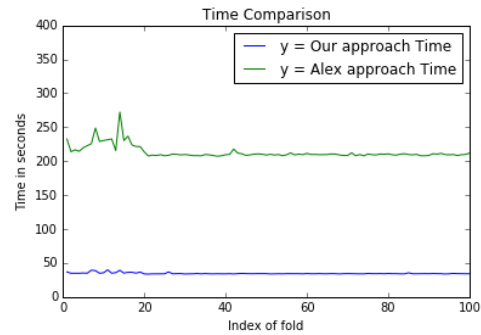


Fig. 6. Comparing time taken (seconds) with Alexandre Model

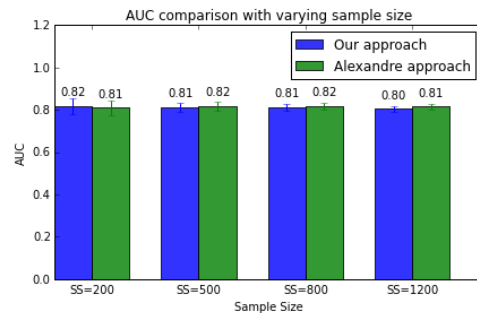


Fig. 7. AUC comparison with varying sample size

P300 data, by using tangent space mapping from riemannian geometry while same kind of approaches were also applied in Motion Imagery (another modality of Brain Computer Interface). System proposed in this study is comprised of three steps. In first step, raw EEG signals from brain are

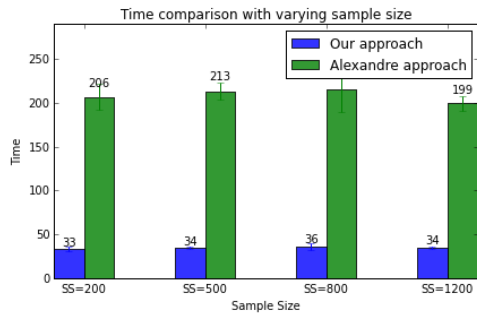


Fig. 8. Time comparison with varying sample size

preprocessed by removing irrelevant electrodes and filtering through 5th order butter-worth filter and discarding signals are bandpass 1-40 Hz. In second step, feature extraction is performed by estimating five xDAWN spatial filters and concatenating with covariances matrices and then tangent space mapping is applied to intermediate results for projecting xDAWN covariances from riemannian manifold into homologous euclidean space. In final step, classification using Elastic net is performed; whose parameters are tuned by cross validation. Based on presented results, it can be concluded that the system classifies the P300 signals efficiently and achieves better AUC while low computational complexity for most of the cases. Also there is drastic decrease in time taken to build a model while maintaining higher AUC while supplying low amount of training data. Proposed model shows better performance during inter-session and inter-subject variability and accomplishes comparable AUC 0.79 in just 37 seconds as compared against state-of-the-art approach, having AUC 0.80 in 224 seconds.

This study can be extended in different ways. As this paper only linear classifiers were used, but other complex classifiers such as neural networks can be used to improve accuracy. Furthermore, we used all electrodes for classification. However, one would like to minimize and select only most relevant electrodes for classification.

REFERENCES

- [1] J d R Millán, Rüdiger Rupp, Gernot R Müller-Putz, Roderick Murray-Smith, Claudio Giugliemma, Michael Tangermann, Carmen Vidaurre, Febo Cincotti, Andrea Kübler, Robert Leeb, et al. Combining brain-computer interfaces and assistive technologies: state-of-the-art and challenges. *Frontiers in neuroscience*, 4, 2010.
- [2] Thorsten O Zander, Christian Kothe, Sebastian Welke, and Matthias Rötting. Utilizing secondary input from passive brain-computer interfaces for enhancing human-machine interaction. In *Foundations of Augmented Cognition. Neuroergonomics and Operational Neuroscience*, pages 759–771. Springer, 2009.
- [3] Lawrence Ashley Farwell and Emanuel Donchin. Talking off the top of your head: toward a mental prosthesis utilizing event-related brain potentials. *Electroencephalography and clinical Neurophysiology*, 70(6): 510–523, 1988.
- [4] A Furdea, S Halder, DJ Krusienski, D Bross, F Nijboer, N Birbaumer, and A Kübler. An auditory oddball (p300) spelling system for brain-computer interfaces. *Psychophysiology*, 46(3):617–625, 2009.
- [5] Alexandre Barachant, Rafał Cycon, and Cedric Gouy-Pailler. P300-speller: Géométrie riemannienne pour la détection multi-sujets de potentiels d’erreur. In *GRETSI 2015*, 2015.

- [6] Dean J Krusienski, Eric W Sellers, François Cabestaing, Sabri Bayoudh, Dennis J McFarland, Theresa M Vaughan, and Jonathan R Wolpaw. A comparison of classification techniques for the p300 speller. *Journal of neural engineering*, 3(4):299, 2006.
- [7] Benjamin Blankertz, Steven Lemm, Matthias Treder, Stefan Haufe, and Klaus-Robert Müller. Single-trial analysis and classification of erp components a tutorial. *NeuroImage*, 56(2):814–825, 2011.
- [8] Jason Farquhar and N Jeremy Hill. Interactions between pre-processing and classification methods for event-related-potential classification. *Neuroinformatics*, 11(2):175–192, 2013.
- [9] Marco Congedo, Alexandre Barachant, and Anton Andreev. A new generation of brain-computer interface based on riemannian geometry. *arXiv preprint arXiv:1310.8115*, 2013.
- [10] Moritz Grosse-Wentrup and Martin Buss. Multiclass common spatial patterns and information theoretic feature extraction. *Biomedical Engineering, IEEE Transactions on*, 55(8):1991–2000, 2008.
- [11] Fabien Lotte and Cuntai Guan. Regularizing common spatial patterns to improve bci designs: unified theory and new algorithms. *Biomedical Engineering, IEEE Transactions on*, 58(2):355–362, 2011.
- [12] Bertrand Rivet, Antoine Souloumiac, Virginie Attina, and Guillaume Gibert. xdown algorithm to enhance evoked potentials: application to brain-computer interface. *Biomedical Engineering, IEEE Transactions on*, 56(8):2035–2043, 2009.
- [13] Gene H Golub and Charles F Van Loan. *Matrix computations*, volume 3. JHU Press, 2012.
- [14] Alexandre Barachant, Stéphane Bonnet, Marco Congedo, and Christian Jutten. Multiclass brain-computer interface classification by riemannian geometry. *Biomedical Engineering, IEEE Transactions on*, 59(4):920–928, 2012.
- [15] Alexandre Barachant, Stéphane Bonnet, Marco Congedo, and Christian Jutten. Classification of covariance matrices using a riemannian-based kernel for BCI applications. *Neurocomputing*, 112:172–178, 2013.
- [16] Alexandre Barachant and Marco Congedo. A plug&play p300 bci using information geometry. *arXiv preprint arXiv:1409.0107*, 2014.
- [17] Neng Xu, Xiaorong Gao, Bo Hong, Xiaobo Miao, Shangkai Gao, and Fusheng Yang. Bci competition 2003-data set iib: enhancing p 300 wave detection using ica-based subspace projections for bci applications. *IEEE transactions on biomedical engineering*, 51(6):1067–1072, 2004.
- [18] Shun-ichi Amari and Hiroshi Nagaoka. *Methods of information geometry*, volume 191 of translations of mathematical monographs. American Mathematical Society, page 13, 2000.
- [19] F Barbaresco. Innovative tools for radar signal processing based on cartans geometry of spd matrices & information geometry. In *Radar Conference, 2008. RADAR’08*. IEEE, pages 1–6. IEEE, 2008.
- [20] P Thomas Fletcher and Sarang Joshi. Principal geodesic analysis on symmetric spaces: Statistics of diffusion tensors. In *Computer Vision and Mathematical Methods in Medical and Biomedical Image Analysis*, pages 87–98. Springer, 2004.
- [21] Oncel Tuzel, Fatih Porikli, and Peter Meer. Pedestrian detection via classification on riemannian manifolds. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 30(10):1713–1727, 2008.
- [22] Maher Moakher. A differential geometric approach to the geometric mean of symmetric positive-definite matrices. *SIAM Journal on Matrix Analysis and Applications*, 26(3):735–747, 2005.
- [23] Hui Zou and Trevor Hastie. Regularization and variable selection via the elastic net. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 67(2):301–320, 2005.
- [24] Shaomin Wu and Peter Flach. A scored auc metric for classifier evaluation and selection. In *Second Workshop on ROC Analysis in ML*, Bonn, Germany, 2005.
- [25] Perrin Margaux, Maby Emmanuel, Daligault Sébastien, Bertrand Olivier, and Mattout Jérémie. Objective and subjective evaluation of online error correction during p300-based spelling. *Advances in Human-Computer Interaction*, 2012:4, 2012.