HAL
open science

Knowledge Transfer for Reducing Calibration Time in Brain-Computer Interfacing

Sami Dalhoumi, Gérard Dray, Jacky Montmain

To cite this version:

Sami Dalhoumi, Gérard Dray, Jacky Montmain. Knowledge Transfer for Reducing Calibration Time in Brain-Computer Interfacing. 26th International Conference on Tools with Artificial Intelligence (ICTAI 2014), Nov 2014, Limassol, Cyprus. 10.1109/ICTAI.2014.100 . hal-01933884

HAL Id: hal-01933884

https://hal.science/hal-01933884v1

Submitted on 21 Jun 2021

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Knowledge Transfer for Reducing Calibration Time in Brain-Computer Interfacing

Sami DALHOUMI, Gérard DRAY, Jacky MONTMAIN
Laboratoire d'Informatique et d'Ingénierie de Production (LGI2P), Ecole des Mines d'Alès
Parc Scientifique G. Besse, 30035 Nîmes, France.
e-mail: {sami.dalhoumi, gerard.dray, jacky.montmain}@mines-ales.fr

***Abstract*— Reducing calibration time while maintaining good classification accuracy has been one of the most challenging problems in electroencephalography (EEG)-based brain-computer interfaces (BCIs) research during the last years. Most of machine learning approaches that have been attempted to address this issue are based on knowledge transfer between different BCIs users. Assuming that there is a common underlying data generating process, they try to learn a subject-independent classification model from multiple users in order to classify data of future users. In this paper, we propose a novel approach that allows inter-subjects classification of EEG signals without relying on the strong assumptions considered in previous work. It consists of learning a prediction model of a new BCI user through an ensemble of classifiers where base classifiers are trained on data from other users separately and weighted according to the performance of the ensemble on few labeled data of the new user. Evaluation on real EEG data showed that our approach allows achieving good classification accuracy when the size of calibration set is small.**

Key words : Brain-Computer Interfaces (BCIs); Electroencephalography (EEG) signals classification; transfer learning; ensemble methods.

I. INTRODUCTION

Brain-computer interfaces (BCIs) are communication and control technologies that enable their users to interact with external environment without using peripheral nervous and muscular system, by directly monitoring electrical or hemodynamic activity of the brain [1]. The first aim of developing such technologies is to offer new communication means for people suffering from severe neuromuscular disorders.

Initial BCI systems were based on intensive user training, which may last several weeks, in order to learn the voluntary control of specific brain signals. With the incorporation of machine learning techniques in the learning process, second generation of BCIs relied on the adaptation of the system to the specificities of the user's brain activity patterns [2]. Although the machine learning approach reduced considerably the duration of user training phase, it still require a time consuming calibration phase (Fig. 1). During this phase, users are instructed to perform repeatedly predefined cognitive tasks (e.g. imagination of movement of left hand vs. imagination of movement of right hand) during specified time periods in order to collect enough labeled data necessary for features extraction and classifier training.

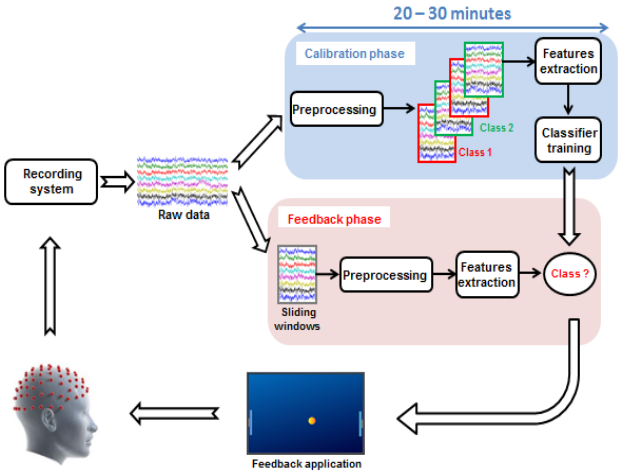

Figure 1. Pattern classification in brain-computer interfacing [3]

Because this learning scheme is a limitation to the use of BCI technology in realistic interaction settings, reducing calibration time while maintaining good classification accuracy has become one of the major challenges in BCI research during the last years. Many machine learning approaches have been attempted to address this issue [2], [4-11], [13-14]. Among them, approaches based on transfer learning have been shown to be the most promising ones [2], [6-11], [13-14]. They consist of incorporating data recorded from other users and/or other sessions in the learning process of the current user. The key issue in such approaches is how to manage the high inter-subjects and inter-sessions variability of brain activity patterns in order to extract useful information and avoid *negative transfer*. To do so, most of existing approaches rely on the assumption that there is a common underlying brain activity pattern which they try to model in order to learn a subject-independent classifier. Although this assumption can be effective for able-bodied users, it may be very strong for disabled users as their brain activity patterns are much more variable [14].

In this paper, we propose a novel approach for inter-subjects classification in EEG-based BCIs. It aims to transfer knowledge between BCI users without relying on the strong assumptions considered in related work. It consists of an ensemble of classifiers, where each classifier is trained on spatially filtered EEG data of one user. Base classifiers are weighted according to the performance of the ensemble on a

small amount of labeled data of a new user projected in different subspaces using the previously learned spatial filters. Negative transfer may occur when a classifier trained using only labeled data of new user outperforms the classifier learned from other users. This is likely to happen for experienced BCI users as they can rapidly control their brain activity. To avoid such situation, spatial filters and corresponding classifier of the new user are learned using calibration data. The final model is chosen depending on the leave one trial out cross-validation result on this small amount of labeled data.

The rest of the paper is organized as follows. In the next section, we give a comprehensive review on inter-subjects and inter-sessions classification in EEG-based BCIs. In section 3, we describe different steps of our approach. In section 4, we evaluate and discuss its performance using a real EEG data set. Finally, section 5 concludes the paper and gives directions for future work.

II. Related work

Several subject-transfer and session-transfer approaches have been attempted to address the issue of reducing calibration time in BCIs technology. *Krauledat et al* [6] used a clustering approach in order to identify spatial filters that have good generalization abilities across sessions of the same user. *Fazli et al* [2] used an ensemble strategy in which a large number of redundant features are extracted in order to train several classifiers. The use of leave one subject out cross-validation and l_1 regularization in the classifiers aggregation step allows finding the brain activity patterns that are common across all BCI users and predictive for future users. Another ensemble learning framework was proposed by *Tu and Sun* [7]. It aims to choose the best subset of spatial filters that allow positive knowledge transfer between subjects. For each subject, a robust filter bank and an adaptive filter bank are learned using quadratic regression with l_1 regularization. The robust filter bank allows extracting characteristics shared between users. Whereas, the adaptive filter bank contains characteristics specific to target user. More recently, *Liyanage et al* [8] proposed a dynamically weighted ensemble method that allows knowledge transfer from one session to another session of the same user. Base classifiers are learned using clustered data of the first session, and Bayesian optimal weighted majority voting is used to classify data of the target session. This approach is limited to learning from a single session. *Lu et al* [9] proposed a subject transfer framework for P300-based BCIs. In their framework, a subject-independent classification model is trained offline from a pool of subjects to capture common brain activity patterns. For a new user, a subject-specific classification model is learned using few labeled data. A confidence score is then calculated to find which classifier is likely to predict the rest of trials of the same user. *Lotte and Guan* [10] proposed an algorithm to regularize parameters of Common Spatial Patterns (CSP) filters and Linear Discriminant Analysis (LDA) classifier learned from few labeled data of current BCI user based on data from other users. It uses a sequential method for choosing the subset of subjects included in regularization. *Almagir et al* [11] applied the framework of multi-task learning [12] for inter-subjects EEG classification. It consists of simultaneously learning several related tasks and a shared feature representation via a common regularization term. The shared characteristics will be used as prior information in order to accelerate the classifier training process for a new user. Feature space transformation has been attempted by *Heger et al* [13]. A linear function was applied to the logarithmic variance features, estimated using little adaptation data of the current user, in order to transform these subject-specific features into subject-independent features. The parameters of the transformation were estimated using a pool of data of other BCI users. Recently, *Samek et al* [14] highlighted that most of the previously presented methods rely on assumptions that are very strong for disabled BCI users because of high inter-subjects and inter-sessions variability of their brain activity patterns. They proposed to transfer information about non stationarity instead of transferring discriminative information.

III. Methods

Instead of learning a feature representation and a classifier that allow modeling brain activity patterns that are common across multiple BCI users, our approach aims to learn multiple feature representations and classification models that provide flexibility of adaptation to specificities of brain activity patterns of each user. In this section, we start with a short background on features extraction in motor imagery based BCIs then we describe different steps of our approach. We use Common Spatial Patterns (CSP) algorithm for spatial filtering, but our approach applies to other features extraction techniques.

A. Features extraction using CSP algorithm

EEG signals are known to have a poor spatial resolution which makes the separation of signal of interest from other signals very difficult. In order to perform classification of multi-channels EEG signals related to different cognitive tasks, data should be projected into a low-dimensional space that spans most of brain signals fluctuations related to tasks of interest.

CSP algorithm is the most successful spatial filtering technique in EEG-based BCIs [3]. It performs a supervised decomposition of multi-channels EEG signals using a parameter matrix W, that projects data into a space in which signals related to different brain states are well discriminated. Each row w of the projection matrix is called spatial filter. For two-class classification, spatial filters are computed by solving the following generalized eigenvalue problem,

$$\text{maximize}_{\mathrm{w}} \frac{\mathrm{w}(\mathrm{C}_1-\mathrm{C}_2)\mathrm{w}^{\mathrm{T}}}{\mathrm{w}(\mathrm{C}_1+\mathrm{C}_2)\mathrm{w}^{\mathrm{T}}} \tag{1}$$

C_1 and C_2 are the spatial covariance matrices of the EEG signals from each class calculated using labeled trials $\{(T_1,y_1),\ldots,(T_P,y_P)\}$ recorded during calibration phase of current BCI user, where $T_p \in R^{t\times n}$ $p=1\ldots P$, is the multi-

channel EEG measurement of the pth trial (n the number of channels and t the number of samples per trial) and y_p the corresponding class label.

Given a new EEG measurement T_{P+1}, spatial filtering is performed as follows,

$$E_{P+1}=W_m T_{P+1}^T \quad (2)$$

where $W_m \in R^{2m\times n}$ is the matrix representing the first m and the last m rows of W that maximize the variance of one class and minimize the variance of the second class and vice versa.

The logarithmic variance feature vector is then calculated using the following formula,

$$x_{P+1}=\log\left(\frac{\mathrm{diag}\left(E_{P+1}E_{P+1}^T\right)}{\mathrm{trace}\left(E_{P+1}E_{P+1}^T\right)}\right) \quad (3)$$

where *diag* returns diagonal elements of the square matrix and *trace* returns the sum of these elements.

CSP algorithm is based on estimation of class covariance matrices which requires a lot of labeled data and consequently a long calibration time. Using spatial filters of other BCI users may solve this problem. However, a proper choice of spatial filters based on their performance in extracting relevant features for new users is necessary to perform positive transfer.

B. *Learning CSP filters and corresponding classifiers from other users*

We propose to learn spatial filters and corresponding classifier of current BCI user (from now called target user) using an ensemble strategy in which base classifiers are trained separately using the spatially filtered EEG signals of other BCI users that performed the same cognitive tasks (from now called source users) and weighted according to the performance of the ensemble in classifying few labeled trials of target user. We used Bayesian model averaging [15], which is a data dependent classifiers combination technique, in order to find the optimal weights for base classification models.

Let $h^1,\ldots,h^K$ be the K classification models trained on EEG data of source users and $L_t=\{(T_1,y_1),\ldots,(T_P,y_P)\}$ the few labeled trials recorded during calibration phase of target user. Let W_m^k be the matrix representing the first m and the last m spatial filters trained on EEG signals of user k, $k=1\ldots K$, and $L_t^k=\{(x_1^k,y_1),\ldots,(x_P^k,y_P)\}$ the set of labeled feature vectors calculated using L_t and W_m^k as shown in equations (2) and (3).

. Following the Bayesian model averaging framework, our ensemble-based approach can be formalized as follows: the probability of having class label y_{P+1} given a new multi-channels EEG measurement T_{P+1} recorded during feedback phase of target user can be expressed as,

$$p\left(y_{P+1}/T_{P+1}\right)=\sum_{k=1}^{K} p\left(y_{P+1}/x_{P+1}^k,h^k\right)p\left(h^k/L_t^k\right) \quad (4)$$

where x_{P+1}^k is the logarithmic variance feature vector of T_{P+1} filtered using the matrix $W_m^k, k=1\ldots K$.

In equation (4), $p\left(y_{p+1}/x_{p+1}^k,h^k\right), k=1\ldots K$, are the predictions made by each base classifier and $p\left(h^k/L_t^k\right)$, $k=1\ldots K$, are classifiers priors. The first term allows learning from each source user and the second term allows adapting the ensemble to the brain activity patterns of target user.

Given a convex loss function l (e.g., mean squared error, hinge loss, logistic loss, etc.), classifiers priors $p\left(h^k/L_t^k\right)$, $k=1\ldots K$, can be estimated empirically by finding the optimal weight vector w^* satisfying,

$$w^*=\mathrm{argmin}\sum_{p=1}^{P} l\left(\sum_{k=1}^{K} w_k h^k\left(x_p^k\right), y_p\right) \quad (5)$$

$$s.t.\sum_{k=1}^{K} w_k=1, w_k\geq 0, k=1\ldots K$$

Then,

$$\hat{p}\left(h^k/L_t^k\right)=w_k^*,\ k=1\ldots K \quad (6)$$

The learned ensemble classifier that will be used to predict class labels of EEG measurements recorded during feedback phase of target user is then,

$$h^*=\sum_{k=1}^{K} w_k^* h^k \quad (7)$$

C. *Avoiding negative transfer*

Negative transfer may occur when the classification model learned using only calibration data of current user outperforms the classification model learned from source users. This is more likely to happen for experienced BCI users because they can rapidly control their brain activity and consequently EEG signals related to different brain states are easily classified using few training data. In order to avoid this situation, projection matrix W_m^{K+1} and corresponding classifier h^{K+1} of target user are learned using labeled trials recorded during calibration phase. Then, classification performance of the weighted ensemble is compared to this single classifier by performing leave one out cross-validation *(LOOCV)* on the calibration set. The prediction model that has the lowest empirical loss is used to classify the rest of trials of current session.

The final prediction model is learned as shown in algorithm 1.

Given
- K classification models learned using EEG signals of source users: $h^1, \dots, h^K$.
- K projection matrices corresponding to the first m and the last m spatial filters of each source user : $W_m^1, \dots, W_m^K$.
- A small calibration set containing P trials of target user: $L_t = \{(T_1, y_1), \dots, (T_P, y_P)\}$.
- A convex loss function l.
- A classification performance measure *perf.*

for $k = 1{:}K$
  Calculate the set of labeled feature vectors L_t^k using L_t and W_m^k as shown in equations (2) and (3);
end for

Train an ensemble of classifiers $\sum_{k=1}^{K} w_k^* h^k$ using $l, h^1, \dots, h^K$ and $L_t^1, \dots, L_t^K$ as shown in equations (5) and (7);

Learn the projection matrix W_m^{K+1} and the set of labeled feature vectors L_t^{K+1} of target user using L_t as shown in equations (1), (2) and (3);

Train a single classification model h^{K+1} using L_t^{K+1};

Calculate *perf*$(\sum_{k=1}^{K} w_k^* h_S^k, \{L_t^1, \dots, L_t^K\})$ using leave one trial out cross-validation;

Calculate *perf*(h^{K+1}, L_t^{K+1}) using leave one trial out cross-validation;

if *perf*$(\sum_{k=1}^{K} w_k^* h_S^k, \{L_t^1, \dots, L_t^K\}) > $ *perf*(h^{K+1}, L_t^{K+1})
  $\mathrm{h}^* = \sum_{\mathrm{k}=1}^{\mathrm{K}} \mathrm{w}_\mathrm{k}^* \times h^k$;
else
  $\mathrm{h}^* = h^{K+1}$;
end if

ALGORITHM 1. LEARNING SPATIAL FILTERS AND CLASSIFIER FROM OTHER USERS.

IV. EVALUATION

Our approach was evaluated on the publicly available data set 2A in BCI competition IV, provided by the Graz group [16]. We used Linear Discriminant Analysis (LDA) for classification as it achieved the best performance compared to other classifiers. Mean squared error (MSE) was used as a loss function for calculating the empirical error of the ensemble on calibration data.

A. EEG data set

The data set consists of EEG signals recorded using 22 Ag/AgCl electrodes from 9 subjects. Subjects were asked to perform four different motor imagery tasks: left hand, right hand, both feet and tongue movement imagination. For each subject, a training and a testing set were collected. Both sets comprise 72 trials of duration 7 s from each class. At the beginning of a trial, a fixation point appeared on a computer screen. After two seconds, a cue appeared informing the subject which imagery task to perform. The subjects were asked to perform the task until the cue disappeared.

EEG measurements were band-pass filtered using a 5th order Butterworth filter in the frequency band 8-30 Hz as movements imagination is known to suppress idle rhythms in this frequency band contra-laterally [17]. Logarithmic variance features were extracted from the time segment 3-5 s after the beginning of each trial. For spatial filtering, we used the three most discriminative CSP filters *(m = 3)*.

B. Classification paradigms

In order to assess the classification performance of our approach, we compared it to two state of the art classification paradigms commonly used in previous work [7], [11], [13]:

- **CALIB**: spatial filters and classifiers are trained using calibration data of current user. This is the traditional machine learning approach in brain-computer interfacing.
- **POOL**: data from all users are pooled together in order to learn spatial filters and corresponding classifier. This allows assessing whether data from different users have the same distribution or not.
- **TL**: our ensemble-based transfer learning approach described in section 3.

C. Results

Fig 2. illustrates classification performance of the approaches CALIB, POOL and TL averaged across all users. Evaluation was performed offline using leave one subject out cross-validation. In each iteration, training sets of eight subjects are used to learn spatial filters and classifiers in POOL and TL approaches. For the ninth subject, the number of trials per class from his training set was gradually increased and in each time spatial filters and corresponding classifier were trained for the CALIB approach and classifiers weights were adjusted for the TL approach. Comparison was made using the test set of the target subject. Performing two-class classification allowed us to evaluate our method using multiple data sets.

Results illustrate that our approach outperformed the two other approaches when the amount of labeled data is small. When only few trials were available, CSP filters and classifier were poorly estimated for the approach CALIB which dramatically deteriorated classification performance. As more data are included in the training phase, CALIB performance increased rapidly. Regarding the approach POOL, classification performance does not depend on the size of calibration set as it is not included in the training phase. The poor classification accuracy of this approach reflects the high inter-subjects variability of EEG signals. Our approach significantly increased classification performance when the size of calibration set is small (up to 11% for right hand vs. both feet movement imagination). Its performance converged to the CALIB approach performance as the number of trials increased. This shows that when there is enough labeled trials for the current user, learning from other users is not useful *(negative transfer)*.

Table I. illustrates results of paired t-test over all subjects and all numbers of trials per class. This test of significance proves that performance achieved by our approach was statistically higher than performance achieved by the other approaches ($p < 0.05$ compared to CALIB and $p < 0.001$ compared to POOL).

TABLE I. P-VALUES OF PAIRED T-TEST OVER ALL SUBJECTS AND ALL NUMBERS OF TRIALS PER CLASS

	Left vs. right	***Left vs. feet***	***Left vs. tongue***	***Right vs. feet***	***Right vs. tongue***	***Feet vs. tongue***
TL vs. CALIB	*0.012*	*0.039*	*0.002*	1.5×10^{-4}	*0.043*	*0.010*
TL vs. POOL	3.6×10^{-7}	1.0×10^{-12}	2.1×10^{-18}	2.1×10^{-11}	1.2×10^{-18}	1.2×10^{-28}

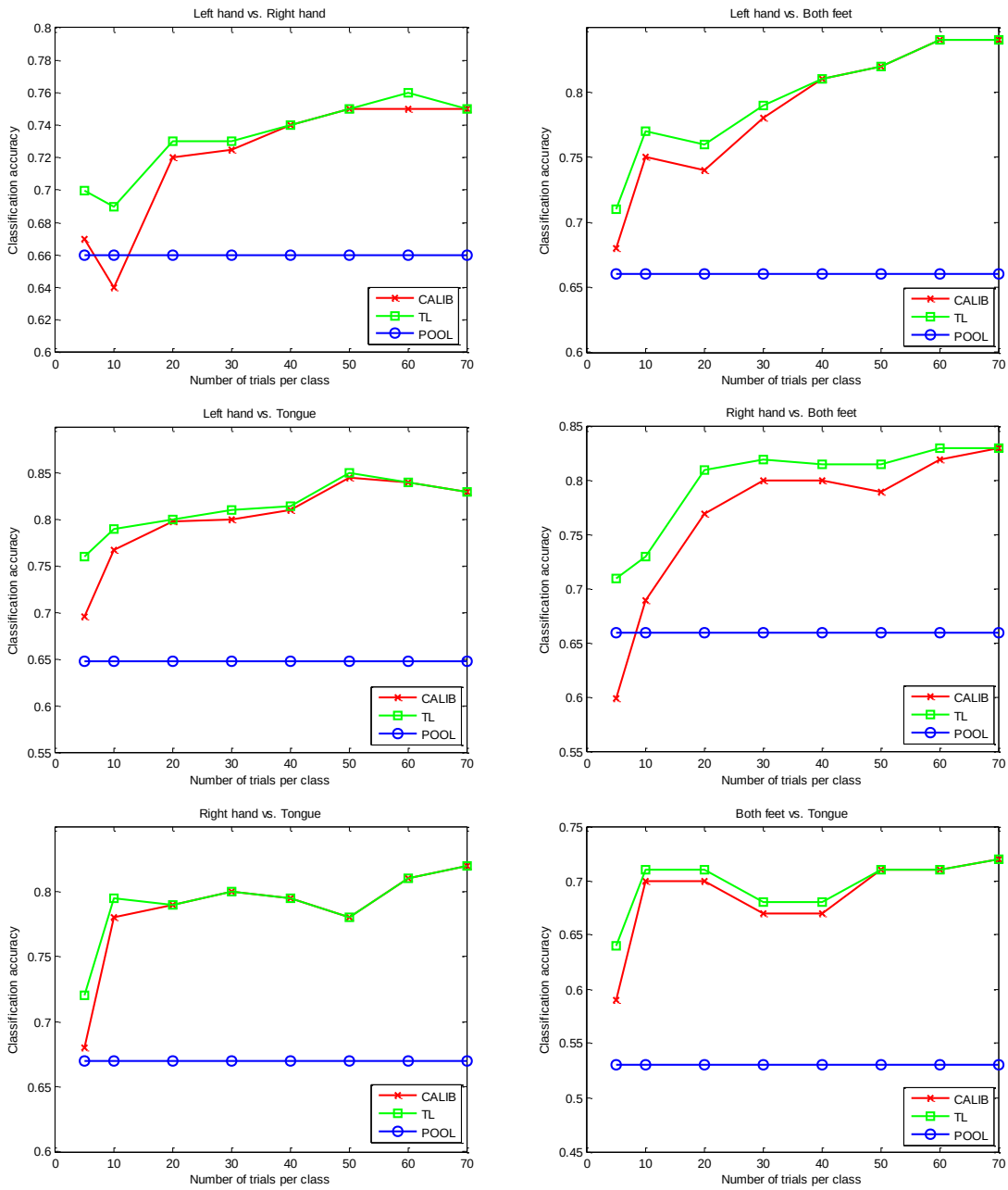

Figure 2. Classification performance of the learning paradigms CALIB, TL and POOL

V. CONCLUSION

In this paper, we presented a novel approach for reducing calibration time in EEG-based BCIs. Unlike most of existing approaches, our approach allows transferring knowledge between BCI users without relying on the strong assumption of common underlying brain activity pattern, which may not hold for disabled users. It is based on ensemble classification which allows modeling different brain activity patterns of different users simultaneously and provide flexibility of adjusting the weights of these models in order to adapt to new user. Empirical results showed that our approach significantly outperformed standard BCI designs when the amount of labeled data is small.

In future work, we will assess the performance of our approach in classifying EEG data recorded from disabled subjects. We will also investigate EEG signals non stationarity within the same session as brain activity patterns may change between calibration phase and feedback phase [8].

REFERENCES

[1] D.J. McFarland, J.R. Wolpaw, "Brain-computer interfaces for communication and control", Communications of the ACM, vol. 54, issue 5, May 2011, pp. 60-66, doi : 10.1145/1941487.1941506.

[2] S. Fazli, C. Grozea, M. Danoczy, F. Popescu, B. Blankertz, K.R. Muller, "Subject independent EEG-based BCI decoding", Proc. Neural Information Processing Systems (NIPS 23), 2009, pp. 513-521.

[3] B. Blankertz, R. Tomioka, S. Lemm, M. Kawanabe, K.R. Muller, "Optimizing Spatial filters for Robust EEG Single-Trial Analysis", IEEE Signal Processing Magazine, vol. 25, 2008, pp. 41-56, doi : 10.1109/MSP.2008.4408441.

[4] J. Blumberg, J. Rickert, S. Waldert, A. Schulze-Bonhage, A. Aeresten, C. Mehring, "Adaptive Classification for Brain Computer Interfaces", Proc. International Conference of the IEEE Engineering in Medicine & Biology Society (EMBS 29), 2007.

[5] Y. Li, C. Guan, H. Li, Z. Chin, "A self-training semi-supervised SVM algorithm and its application in an EEG-based brain computer interface speller system", Pattern Recognition Letters, vol. 29, issue 9, July 2008, pp. 1285-1294, doi : 10.1016/j.patrec.2008.01.030

[6] M. Krauledat, M. Tangermann, B. Blankertz, K.R. Muller, "Towards Zero Training for Brain-Computer Interfacing", PLoS ONE 3(8): e2967. doi:10.1371/journal.pone.0002967.

[7] W. Tu, S. Sun, "A subject transfer framework for EEG classification", Neurocomputing, vol. 82, April 2012, pp. 109-116, doi : 10.1016/j.neucom.2011.10.024.

[8] S.R. Liyanage, C. Guan, H. Zhan, K.K. Ang, J. Xu, T.H. Lee, "Dynamically weighted ensemble classification for non-stationary EEG processing", Journal of Neural Engineering, June 2013, 10(3):036007. doi: 10.1088/1741-2560/10/3/036007.

[9] S. Lu, C. Guan, H. Zhang, "Unsupervised Brain Computer Interface Based on Intersubject Information and Online Adaptation", IEEE Transactions on Neural Systems and Reabilitation Engineering, vol. 17, February 2009, pp. 135-145, doi : 10.1109/TNSRE.2009.2015197.

[10] F. Lotte, C. Guan, "Learning from other subjects helps reducing brain-computer interface calibration time", Proc. International Conference on Audio Speech and Signal Processing (ICASSP), IEEE Press, March 2010, pp. 614-617, doi : 10.1109/ICASSP.2010.5495183.

[11] M. Almagir, M. Grosse-Wentrup, Y. Altun, "Multitask Learning for Brain-Computer Interfaces", Proc. International Conference on Artificial Intelligence and Statistics (AISTATS 13), 2010.

[12] A. Argyriou, T. Evgeniou, M. Pontil, "Convex multi-task feature learning", Machine Learning, vol. 73, December 2008, pp. 243-272, doi : 10.1007/s10994-007-5040-8.

[13] D. Heger, F. Putze, C. Herff, T. Schultz, "Subject-To-Subject Transfer for CSP based BCIs: Feature Space Transformation and Decision-Level Fusion", Proc. International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC 13), 2013:5614-7. doi: 10.1109/EMBC.2013.6610823.

[14] W. Samek, F.C. Meinecke, K.R. Muller, "Transferring Subspaces Between Subjects in Brain-Computer Interfacing", IEEE Transactions on Biomedical Engineering, vol. 60, August 2013, pp. 2289-2298, doi : 10.1109/TBME.2013.2253608.

[15] J.A. Hoeting, D. Madigan, A.E. Raftery, C.T. Volinsky, "Bayesian Model Averaging: A Tutorial" Statistical Science, vol. 14, no. 4, 1999, pp. 382-417.

[16] B. Blankertz, "BCI Competition IV website : http://www.bbci.de/competition/iv/".

[17] G. Pfurtscheller, C. Brunner, A. Schlogl, F.H. Lopes da Silva, "Mu rhythm (de)synchronization and EEG single-trial classification of different motor imagery tasks", NeuroImage, vol. 31, 2006, pp. 153-159, doi : 10.1016/j.neuroimage.2005.12.003.