Bulletin of Electrical Engineering and Informatics
Vol. 12, No. 6, December 2023, pp. 3584~3592
ISSN: 2302-9285, DOI: 10.11591/eei.v12i6.6079

Session to session transfer learning using regularized four parameters common spatial pattern method

Zaineb M. Alhakeem[1], Ramzy S. Ali[2]
[1]Department of Communication Engineering, Iraq University College, Basrah, Iraq
[2]Department of Electrical Engineering, University of Basrah, Basrah, Iraq

Article Info

Article history:

Received Sep 30, 2021
Revised Dec 30, 2022
Accepted Feb 23, 2023

Keywords:

Brain computer interface
Common spatial pattern
Electroencephalography
Motor imagery
Session to session learning

ABSTRACT

Brain computer interface (BCI) has many useful applications to help disabled people that have an active brain with difficulties in movements and speaking. One of these applications is the wheelchair, this device is operated always by just one user no sharing or borrowing the device. One-user applications need features extraction methods with high classification accuracy and small training datasets, the variability of the subjects' mood during the recorded sessions and the tiredness during the long sessions are serious problems that affect the classification accuracy in these applications. Transfer learning can solve the problem, by recording short and separated sessions for the same subject in different training times or days. The proposed method in this paper uses motor imagery (MI) signals from different recorded sessions by one user to build an acceptable size training dataset. To regularize different recording sessions, four tuning parameters that are independent from each other are generated using a loop, these parameters are used to find the ratios of the covariance matrices. The suggested method gives very good performance using a different number of training samples compared with six different common spatial patterns (CSP) methods using only two channels.

Corresponding Author:

Zaineb M. Alhakeem
Department of Communication Engineering, Iraq University College
Basrah, Iraq
Email: zaineb.alhakeem@iuc.edu.iq

1. INTRODUCTION

Brain activities of any person could be recorded using two types of methods non invasive or invasive. The non-invasive methods use channeles attached to the scalp of the user. Different channels of the recorded signal can be used to extract features by using different methods such as power of the signals directly or combined with other methods [1]. Spatial filters are one of theses features extraction methods. Ramoser *et al.* [2] used common spatial patterns (CSP) to extract motor imagery (MI) features. Conventional CSP is a user or session dependening method, where there is no information about other users or sessions that added to the training-dataset. Wang *et al.* [3] and Cao *et al.* [4] used the classic CSP in a hybrid wheelchair control system each one with different number of commands.

Černý and Šťastný [5] applied CSP on finger movement instead of hand movement. Sun *et al.* [6] used CSP and support vector machine (SVM) in online experiment. Classic CSP gives poor classification accuracy when the training data set is small. Pan *et al.* [7] used the same principles of CSP but with the frequency domain signals extracted from only two channels named as common frequency pattern (CFP). Modifications and improvements are made to increase the classification accuracy. Xygonakis *et al.* [8] extracted the spatial filter from more than one region named as regions of interest. Park and Chung [9] divided the head area into sub-areas each one centered by a channel, they found the best region to find the spatial filter of CSP.

Lv and Liu [10] selected the best channels to be used to find the spatial filter of CSP using binary particle swarm optimization (BPSO). Wang *et al.* [11] proposed a modification in the features extraction method of CSP and used it in frequency domain instead of the time domain to decrease the time consumed in extracting features. Zhang *et al.* [12] used wavelet analysis to enhance the signal while Gouy-Pailler *et al.* [13] used independent component analysis (ICA) before applying CSP. Mousavi *et al.* [14] extracted the frequency of interest using wavelet transfer analysis then used the signals to extract the spatial filter of CSP, while Robinson *et al.* [15] used only the useful components after clearing all the unneeded ones to use the reconstructed signal as input signal to CSP filter. Ang *et al.* [16] divided the frequency band into multiple frequency banks each one used to find the spatial filter and features then applied to different classifiers their method named as filter bank CSP (FBCSP), while Novi *et al.* [17] applied the sub-bands to the same classifier this method named as sub band CSP (SBCSP).

Kumar and Sharma [18] used the meta-heuristic algorithms to find the best frequency range and band-pass filter order that will give the best accuracy of classification. Ge *et al.* [19] used fusion features time domain and CSP features. Song and Yoon [20] introduced adaptive CSP by measuring the similarity between the training data and the new data to improve the spatial filter. Higashi and Tanaka [21] optimized the parameters and window size that used in calculating CSP filter. To increase the separability a complexity index is used by Li *et al.* [22] in calculating CSP spatial filter. To increase the robustness of the features Samek *et al.* [23] combined CSP with Tikhonov regularization method to produce stationary measure using CSP method.

To improve the accuracy of the classification and to overcome the lack of training-dataset problem, some researchers proposed a regularized CSP (RCSP) method. The regularization is done in two ways either on the objective functions or on the covariance matrices [24]. The regularization of covariance matricies is done by adding more information from other subjects or sessions beside the target ones. Lu *et al.* [25] proposed the RCSP with generic learning (RCSPGL). RCSPGL uses data from a subjects' population to increase the training dataset size. Lu *et al.* [26] used aggregation to choose the shrinking parameters of RCSPGL. Shin *et al.* [27] used RCSP method to classify the human moods. Yuksel and Olmez [28] emphased the channels that are close to the imagery area using spatially regularizing CSP method.

Li *et al.* [29] improved RCSP method by using statistical dependency method to extract features instead of variances. Lotte and Guan [24] proposed four different features extraction algorithms one of them is the RCSP with selected subject (SSRCSP). SSRCSP uses data from selected subjects or session, not from all of the population. Lotte and Guan [30] regularized the objective function using Laplacian penalty that produce smooth spatial filter. Park *et al.* [31] used the same method as FBCSP but with RCSP instead of CSP. Li and Wang [32] introduced two smoothing algorithms, the first used Gaussian prior and the other used ridge penalty function. Both are used to RCSP method in the objective function.

Cheng *et al.* [33] developed subject to subject transfer learning by RCSP, by finding the weights of the most similar subject to calculate the spatial filter. Alhakeem and Ali [34] used a combination of both RCSP and ICA to improve the accuracy of session to session transfer learning of MI data. Xu *et al.* [35] used cosine similarities between the source and the target subject, they proposed iteration CSP method to find the most suitable subject to fine the spatial filter. Kang *et al.* [36] developed two methods to find the composite CSP (CCSP) to increase the accuracy percentage of the classifiers. An adaptive session to session extreme learning machine is proposed by Bamdadian *et al.* [37] to improve the classification accuracy. Cho *et al.* [38] studied the effect of the background noise, they found that removing the background noises while working will reduce the needed training time.

Some researchers proposed algorithms based on CSP to classify multiclass rather than the previously mentioned ones which they are binary class only. Wu *et al.* [39] used multiple binary class CSP with one versus the rest algorithm to expand the binary class CSP to multiclass one. Grosse-Wentrup and Buss [40] used joint approximate diagonalization (JAD) with CSP to obtain multiclass CSP. Sun proposed a multiclass CSP by dividing the frequency range in to sub-ranges then find the spatial filter of multiple binary problems for each sub-range [41]. In one subject brain computer interface (BCI) applications such as wheelchair control, there is different recording sessions could be used as a regularizing dataset to increase the training trials numbers instead of using one day recordings. The proposed method uses four tuning parameters that represented the ratios of the covariance matrices that are used in spatial filters extraction.

2. RESEARCH METHOD AND DATASET

2.1. Four-parameter regularized common spatial pattern

There are many types of CSP methods starting from the traditional [2], this type depend only on the training set of the target no information of others or the old state of training. Two methods are proposed [34] named as CCSP, this method used two different ways to calculate the weights that refer to the importance of the covariance matrix. Other methods are proposed to regularize the covariance matrices each in its own way. RCSPGL [25] in this method a population of training data from other users are used and two shrinking parameters that will refer to the ratios of the covariance matrices. While the SSRCSP method that proposed [24] it works by selecting the most fitted subject for training even if there is a big group of subjects or enough data for training. This process is done by discarding the subjects that give useless information and keep the useful ones.

In the proposed method the covariance matrices are determined just like the traditional CSP method. As the following, the normalized covariance matrix is determined from the data matrix E, which is here the data of different channels over time, for each class i and trial j:

$$C_{ij} = \frac{E_{ij}E_{ij}^T}{trace\left(E_{ij}E_{ij}^T\right)} \tag{1}$$

For N training trials the average covariance matrix is of class i:

$$CN_i = \frac{1}{N}\sum_{j=1}^{N} C_{ij} \tag{2}$$

where N is the training trials number, I is 1, or 2 depending on the choosen class (1 or 2), j is 1, …, N (number of trails), trace: is the summation of the diagonal elements in square matrix.
The covariance matrices are as:

$$\widetilde{CN}_i = a_1 CN_{i(target)} + a_2 CN_{i(training)} \tag{3}$$

$$\widehat{CN}_i = a_3\widetilde{CN}_i + a_4 I \tag{4}$$

where $a_1, a_{2,}, a_3, and\ a_4$ are the tuning parameters, all of them belong to the period [0, 1] with a condition which is all parameters could not be zero at the same time. This condition is made to prevent the matrices from vanishing by multiplying them by a zero parameter.

The first phase has two parameters $a_1\ and\ a_2$ as shown in (3), they will define the ratio of the covariance matrix whether it go forward the training covariance matrix or to the target one. These parameters are not complemented and independent of each other so shrinking one matrix will not lead to enlarge the other and decrease its effect on the classification accuracy. The second phase in (4) has two parameters also these will give us two matrices one is a ratio of the identity matrix and the other is a ratio of the first phase resultant matrix, the parameters are not the complement of each other as we mentioned before and not dependent on the first two parameters. This condition will give us four matrices that each one has a portion of the target user, the other users and the identity matrix all that will give better extraction of different features to be classified later. To prevent the matrices from vanishing no two parameters in the same phase should be zero at the same time. After calculating the covariance matrices, the composite covariance matrix for two classes defined as (5):

$$sum = \sum_{i=1}^{2} \widehat{CN}_i \tag{5}$$

Then sum is factorized as (6):

$$sum = U\lambda U^T \tag{6}$$

The whitening transformation is determined as (7):

$$P = \frac{1}{\sqrt{\lambda}} U^T \tag{7}$$

The whitened sum_i is as (8):

$$\widehat{sum}_i = P\ \widehat{CN}_i P^T \tag{8}$$

The factorization of is as (9):

$$\widehat{sum}_i = B\ \lambda_i B^T \tag{9}$$

The full projection matrix is as (10):

$$\widetilde{W} = B^T \mathrm{P} \tag{10}$$

Select the first and last Q columns to produce the projection matrix W:

$$f = WE \tag{11}$$

The features are:

$$\tilde{f} = log \frac{var(f)}{\sum var(f)} \tag{12}$$

2.2. Support vector machine classifier [42]

SVM is one of the most popular classifiers. In this classifier there is a separation hyper-plane separate the members of two classes, Figure 1 shows the details of the SVM classifier. In linearly separable cases SVM maximize the distance between margins. This is formulated in (13):

$$y_i(x_i - w + b) - 1 \geq 0, \text{ for all } i \tag{13}$$

where x_i is the features and y_i is the classes of them. Linear Kernel SVM classifier is used in this work, this Kernel is suitable for the number of features and number of training trails that used.

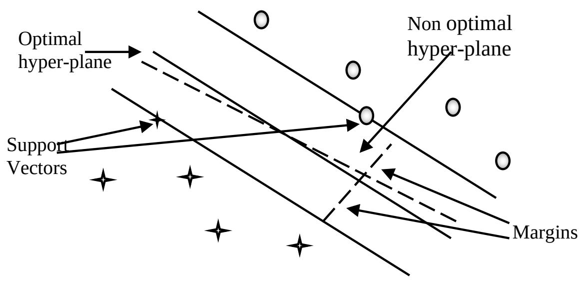


Figure 1. SVM classifier (separable case)

2.3. Datasets

2.3.1. Dataset_I

This dataset is provided by Dr Cichocki's Lab [43] which is recorded for MI signal, multi-user, and multiple-sessions for one user. This dataset is recorded by using two recording devices which they are (Neuroscan and g.tec). The sampling frequencies of them are 256 and 250 Hz respectively. The data is recorded using different number of channels which they are five (C3, C43, Cp3, Cp4, and Cz), six (C3, Cp3, Cz, Cpz, C4, and Fp4) and fourteen channels (C1, C2, C3, C4, C5, C6, Cz, Cpz, Cp1, Cp2, Cp3, Fp1, Fp2, and Fp3). Three classes are left-hand, right-hand, and both feet are used to record data for eight different healthy subjects mentioned in the dataset as (a, b, c, d, e, and h). In this work, only two classes are used imagining the movement of left and right hands, if the data recorded for three classes, we use left-hand and right-hand classes only. Our method focused only on signals from two channels C3 and C4 which they are close to the motor cortex area in the brain, therefore, the signals from the other channels are ignored.

2.3.2. Dataset_II

This dataset is recorded by the one of authors with three volunteers, who they are healthy people. Each one labelled by the first letter of the his/her name (A, H, and Z) all of them have no mental disease and record these signals for the first time they all have healthy or corrected eyesight. The same recording protocol [43] is used also. Two classes are recorded which they are the left-hand class and right-hand class. The experiment take place in an ordinary room environment using OPENBCI recording device, 16-channels which they are (C3, C4, Cz, O1, O2, Oz, P3, P4, Pz, T7, T8, F3, F4, Fp1, Fp2, and Fpz) and 125 Hz as a sampling frequency. Multiple-sessions are recorded by each user, only C3 and C4 channels are used in the experiments and all other channels are dropped from the experiment.

2.4. Evaluation method

The accuracy is the most suitable method to evaluate the performance of balanced datasets. The balance dataset means number of classes are equal for each one. The standard deviation is used to find the divergence of the results from their average:

$$Accuracy = \frac{number\ of\ true\ classified\ samples}{total\ umber\ of\ samples} \quad (14)$$

$$Standard\ Deviation = \sqrt{\frac{\sum_{i=1}^{N}(x_i - \bar{x})^2}{N-1}} \quad (15)$$

where N is the number of samples, xi is the tested sample and $\bar{x}$ is the average of the samples.

3. RESULTS AND DISCUSSION

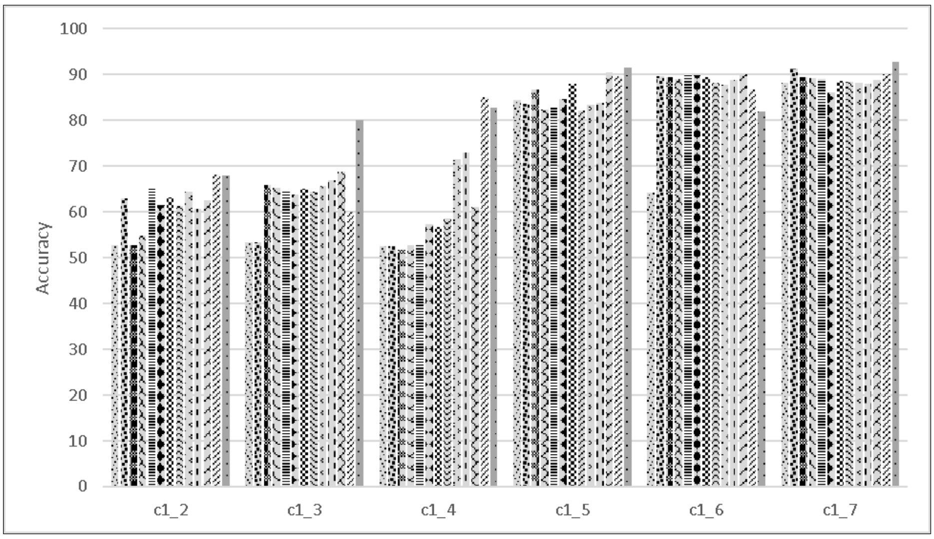
FPRCSP method is used to extract the features from both datasets-I and dataset-II that previously discussed in subsection 2.3. As this work studies the effects of adding new recorded sessions and the number of them on the accuracy of classification. Linear Kernel SVM is used as a classifier, because of the small features vector, no need to complex Kernel of SVM classifier. The average accuracy is taken for ten runs to overcome any randomness in initial values, the standard deviation of the runs is taken too. The data set silted in to training set and testing set each one is different from the other. As a data preprocessing step, the signals are filtered using band-pass filter in the range of alpha and beta activity frequency range of the brain, which is 7-30 Hz. All the tuning parameters that used in FPRCSP which they are ($a_1, a_{2,}, a_3, and\ a_4$) generated by using simple loops within the period [0,1] with step =0.1.

To show the results of changing the sizes of training-set and the effect of adding new session on the accuracy, subject-c/dataset-I is used in this experiment. Seven sessions are recorded each in different day are used to show the effect of adding new sessions to the training-set. The sessions are divided as the following the new session is always the target one and the old are the generic (e.g., session c1-4 means that the first three sessions (1, 2, and 3) are the generic training-data and the fourth one (session 4) is the target session). Figure 2 shows that when the sessions are few (from two to four sessions) the number of trails will not improve the classification accuracy unless it is relatively high (40 to 50 trails). This means that the subject did not train well to generate a powerful signal to be classified using low number of trails. After the fifth day of recording the accuracy starting to improve to reach more than 80% using whatever the number of trials that used. Now what is the effect one the training-set size? The answer is that, the training set size equals to number of trails (N)×number of sessions (i.e., if there are five sessions each one of N=50 the training-set size is 50×5=250 trails) when the data increases the old data will be unuseful and could be discarded to keep the training-set updated and as small as possible.

To compare the performance of the proposed method, the results of FPRCSP are compared with seven other methods which are (CSP, CCSP1, CCSP2, RCSPDL, RCSPGL, SSRCSP, and ICA_RCSP). The results of this comparison are shown in Tables 1-3, the comparison is done with other methods using different number of trails. Some results appeared as NaN which means that there is a zero division is some cases FPRCSP is not suffer from dividing by zero problem because there is no parameter in the denominator could be zero at all. From Tables 1 and 2 it is obvious that FPRCSP method give better results after adding the data of four days and above and the results of the classification using different number of trails per session which they are in the tables (12 and 20 trails), besides that the average classification accuracy is the best amount the other methods.

In Table 3 the number of trails is 50 from each session, because the data starting to be enough to extract good features from only two sessions and above, the accuracy increased from about 80 to 92%. Fifty trails from each session are not very big number of trails that will make the user tried to reduce the power of the recorded signals and will not make the dataset very bit to be unseparable using SVM classifier. The overall number of trails will be is the seven sessions are used is 50×7=350 trail only giving 92% classification accuracy.

The results of subject c/dataset I and subject z/dataset II are shown in Figure 3, the data are recorded during different 5 sessions for only two classes. The comparison is done among different CSP methods using number of training trails (N=40) from each session. FPRCSP method still has the best result among all other methods that compared with. Table 4 shows the results of subject A/dataset II the comparison is done for (N=20) trials per session. Both subjects A and subject Z have recorded the signals for the first time, both have no experience of motor imagination. Figure 4 shows the result of testing FPRCSP method using subject-to-subject transfer learning using (N=50) trails per subject each one is the target subject and the other are the rest populations.



*Each bar in this figure represents the results of training data that collected from the labelled sessions and the last one is the target

Figure 2. A comparison according to the classification accuracy of FPRCSP method using different training datasets' sizes of subject_c/dataset_I

Table 1. A comparison according to classification accuracy of subject_c/dataset_I when using 12 trails from each session

Session \ Method	CSP method	CCSP1 method	CCSP2 method	RCSPDL method	RCSPGL method	SSRCSP method	ICA_RCSP method	FRCSP method
Sessions c1_2	54.347	**71.014**	**71.014**	53.623	**71.014**	**71.014**	67.173	61.594
Sessions c1_3	46.296	NaN	NaN	**65.740**	NaN	NaN	55.555	63.888
Sessions c1_4	20.370	50	50.925	50	50.925	NaN	**62.129**	57.407
Sessions c1_5	58.333	75.694	76.388	67.361	81.944	75.694	82.152	**84.722**
Sessions c1_6	85.227	NAN	NaN	88.636	NaN	NaN	83.181	**89.772**
Sessions c1_7	88.888	**91.666**	62.037	63.888	**91.666**	NaN	78.703	86.111
Average accuracy	58.910	72.093	65.091	64.875	73.8878	73.354	71.482	**73.916**
Standard deviation	23.298	14.877	9.6528	12.3853	15.136	2.3399	10.5085	13.177

Table 2. A comparison according to classification accuracy of subject_c/dataset_I when using 20 trails from each session

Session \ Method	CSP method	CCSP1 method	CCSP2 method	RCSPDL method	RCSPGL method	SSRCSP method	ICA_RCSP method	FRCSP method
Sessions c1_2	51.538	52.307	52.307	53.076	52.307	52.307	**65.384**	60.7692
Sessions c1_3	42	NAN	NaN	44	NaN	NaN	50	**67**
Sessions c1_4	50	50	50	50	50	NaN	57	**73**
Sessions c1_5	31.617	74.264	77.205	69.852	77.205	77.205	74.558	**83.8235**
Sessions c1_6	87.50	NaN	NaN	90	NaN	NaN	76.25	**88.75**
Sessions c1_7	82	83	63	83	**92**	NaN	85.10	88
Average accuracy	57.4426	64.893	60.628	64.988	67.878	64.756	68.048	**76.8904**
Standard deviation	20.4212	14.105	10.7544	17.2274	17.5423	12.449	11.9395	10.6847

Table 3. A comparison according to classification accuracy of subject_c/dataset_I when using 50 trails from each session

Session \ Method	CSP method	CCSP1 method	CCSP2 method	RCSPDL method	RCSPGL method	SSRCSP method	ICA_RCSP method	FRCSP method
c1_2	67	77	77	**79**	77	77	54.30	68
c1_3	48.571	55.714	55.714	54.285	55.714	58.571	17.428	**80**
c1_4	81.428	**82.857**	**82.857**	**82.857**	**82.857**	**82.857**	40	**82.857**
c1_5	83.962	**92.452**	**92.452**	80.188	**92.452**	**92.452**	49.528	91.509
c1_6	86	**90**	**90**	**90**	**90**	NaN	62	82
c1_7	91.428	**95.714**	94.285	**95.714**	**95.714**	94.285	49.428	92.857
Average	76.3984	82.2897	82.0516	80.3409	82.2897	81.033	45.447	**82.870**
Standard deviation	14.5102	13.4009	13.1710	13.0170	13.4009	12.8853	14.139	8.21068

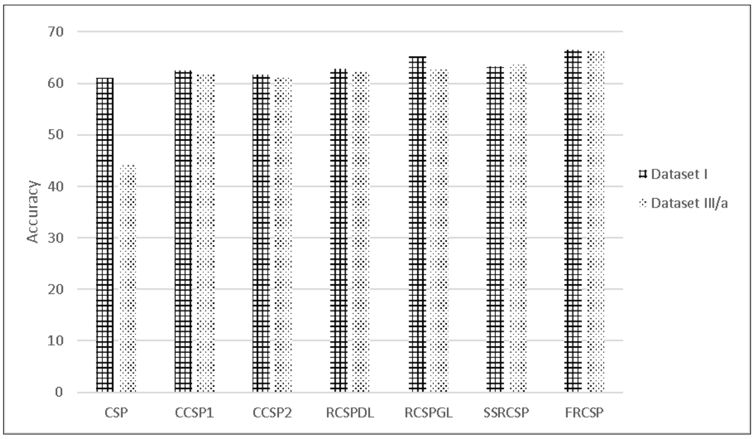


Figure 3. A comparison according to classification accuracy of subject_c/dataset_I and subject when using 40 trails from each session

Table 4. A comparison among different CSP methods using subject_a/dataset_II with N=20

Session \ Method	CSP method	CCSP1 method	CCSP2 method	RCSPDL method	RCSPGL method	SSRCSP method	FPRCSP method
Sessions A1_2	43.333	50	50	53.333	53.333	50	**56.667**
Sessions A1_3	46.666	**50**	**50**	**50**	**50**	**50**	**50**
Accuracy average	45	50	50	51.6665	51.6665	50	**53.335**
Standard deviation	1.666	0	0	1.6665	1.6665	0	3.335

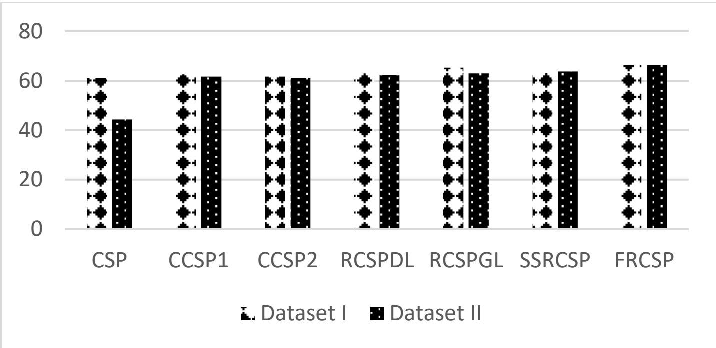


Figure 4. A comparison among different CSP methods using the data of 6 subjects from dataset_I and 3 subjects from dataset_II with N=50 according to accuracy

4. CONCLUSION

Transfer learning such as (session-to-session) is very important when a single user BCI applications, when there is no enough data to train the model, such applications are used by only one person so no need to train the model using other users' data such as wheelchairs control. FPRCSP method based on regularizing the subjects' data on different sessions, because the mode and attention the subject differ from time to time and from day to day. To achieve good performance according to the accuracy of classification electroencephalography (EEG) signals over multi-sessions for single user or multi-user in different sessions, tuning parameters are used to take ratios from different covariance matrices, four of them are used in this work. FPRCSP is applied to both subject-to-subject and session-to-session transfer learning.

Although the results of (session-to-session) transfer learning is better than (subject-to-subject) transfer learning according to the classification accuracy, the results of inter-subject are acceptable also as it is shown previously. Some methods such as (CSP, CCSP1, CCSP2, RCSPDL, RCSPGL, SSRCSP, and ICA_RCSP) are special cases of FPRCSP so no need to take time to choose one of these methods in feature extraction phase instead using FPRCSP will do the work in most cases. Adding a fresh recording session will increase the accuracy of the classification. For future work a method should be found to exactly choose the tuning parameters.

REFERENCES

[1] Z. M. Alhakeem and R. S. Ali, “Fast channel selection method using crow search algorithm,” in *Proceedings of the International Conference on Information and Communication Technology*, Apr. 2019, pp. 210–214, doi: 10.1145/3321289.3321309.

[2] H. Ramoser, J. Muller-Gerking, and G. Pfurtscheller, “Optimal spatial filtering of single trial EEG during imagined hand movement,” *IEEE Transactions on Rehabilitation Engineering*, vol. 8, no. 4, pp. 441–446, Dec. 2000, doi: 10.1109/86.895946.

[3] H. Wang, Y. Li, J. Long, T. Yu, and Z. Gu, “An asynchronous wheelchair control by hybrid EEG–EOG brain–computer interface,” *Cognitive Neurodynamics*, vol. 8, no. 5, pp. 399–409, Oct. 2014, doi: 10.1007/s11571-014-9296-y.

[4] L. Cao, J. Li, H. Ji, and C. Jiang, “A hybrid brain computer interface system based on the neurophysiological protocol and brain-actuated switch for wheelchair control,” *Journal of Neuroscience Methods*, vol. 229, pp. 33–43, May 2014, doi: 10.1016/j.jneumeth.2014.03.011.

[5] V. Černý and J. Šťastný, “Application of common spatial patterns on classification of right hand finger movements from EEG signal,” *Elektrorevue*, vol. 5, no. 3, pp. 31–35, 2014.

[6] H. Sun, Y. Xiang, Y. Sun, H. Zhu, and J. Zeng, “On-line EEG classification for brain-computer interface based on CSP and SVM,” in *2010 3rd International Congress on Image and Signal Processing*, Oct. 2010, pp. 4105–4108, doi: 10.1109/CISP.2010.5648081.

[7] Y. Pan, C. Guan, J. Yu, K. K. Ang, and T. E. Chan, “Common frequency pattern for music preference identification using frontal EEG,” in *2013 6th International IEEE/EMBS Conference on Neural Engineering (NER)*, Nov. 2013, pp. 505–508, doi: 10.1109/NER.2013.6695982.

[8] I. Xygonakis, A. Athanasiou, N. Pandria, D. Kugiumtzis, and P. D. Bamidis, “Decoding motor imagery through common spatial pattern filters at the EEG source space,” *Computational Intelligence and Neuroscience*, vol. 2018, pp. 1–10, Aug. 2018, doi: 10.1155/2018/7957408.

[9] Y. Park and W. Chung, “BCI classification using locally generated CSP features,” in *2018 6th International Conference on Brain-Computer Interface (BCI)*, Jan. 2018, pp. 1–4, doi: 10.1109/IWW-BCI.2018.8311492.

[10] J. Lv and M. Liu, “Common spatial pattern and particle swarm optimization for channel selection in BCI,” in *2008 3rd International Conference on Innovative Computing Information and Control*, 2008, pp. 457–457, doi: 10.1109/ICICIC.2008.196.

[11] J. Wang, Z. Feng, and N. Lu, “Feature extraction by common spatial pattern in frequency domain for motor imagery tasks classification,” in *2017 29th Chinese Control And Decision Conference (CCDC)*, May 2017, pp. 5883–5888, doi: 10.1109/CCDC.2017.7978220.

[12] L. Zhang, G. Liu, and Y. Wu, “Wavelet and common spatial pattern for EEG signal feature extraction and classification,” in *2010 International Conference on Computer, Mechatronics, Control and Electronic Engineering*, Aug. 2010, pp. 243–246, doi: 10.1109/CMCE.2010.5609989.

[13] C. Gouy-Pailler, M. Congedo, C. Brunner, C. Jutten, and G. Pfurtscheller, “Multi-class independent common spatial patterns: exploiting energy variations of brain sources,” in *BCI 2008-4th International Brain-Computer Interface Workshop and Training Course*, 2008, pp. 20–25.

[14] E. A. Mousavi, J. J. Maller, P. B. Fitzgerald, and B. J. Lithgow, “Wavelet common spatial pattern in asynchronous offline brain computer interfaces,” *Biomedical Signal Processing and Control*, vol. 6, no. 2, pp. 121–128, Apr. 2011, doi: 10.1016/j.bspc.2010.08.003.

[15] N. Robinson, A. P. Vinod, C. Guan, K. K. Ang, and T. K. Peng, “A wavelet-CSP method to classify hand movement directions in EEG based BCI system,” in *2011 8th International Conference on Information, Communications & Signal Processing*, Dec. 2011, pp. 1–5, doi: 10.1109/ICICS.2011.6174210.

[16] K. K. Ang, Z. Y. Chin, H. Zhang, and C. Guan, “Filter bank common spatial pattern (FBCSP) in brain-computer interface,” in *2008 IEEE International Joint Conference on Neural Networks (IEEE World Congress on Computational Intelligence)*, Jun. 2008, pp. 2390–2397, doi: 10.1109/IJCNN.2008.4634130.

[17] Q. Novi, C. Guan, T. H. Dat, and P. Xue, “Sub-band common spatial pattern (SBCSP) for brain-computer interface,” in *2007 3rd International IEEE/EMBS Conference on Neural Engineering*, May 2007, pp. 204–207, doi: 10.1109/CNE.2007.369647.

[18] S. Kumar and A. Sharma, “A new parameter tuning approach for enhanced motor imagery EEG signal classification,” *Medical & Biological Engineering & Computing*, vol. 56, no. 10, pp. 1861–1874, Oct. 2018, doi: 10.1007/s11517-018-1821-4.

[19] S. S. Ge, X. Li, and Y. Pan, “Fusion of CSP-derived features and time-domain parameters for mind robotics control,” in *2011 8th Asian Control Conference (ASCC)*, 2011, pp. 79–84.

[20] X. Song and S.-C. Yoon, “Improving brain–computer interface classification using adaptive common spatial patterns,” *Computers in Biology and Medicine*, vol. 61, pp. 150–160, Jun. 2015, doi: 10.1016/j.compbiomed.2015.03.023.

[21] H. Higashi and T. Tanaka, “Common spatio-time-frequency patterns for motor imagery-based brain machine interfaces,” *Computational Intelligence and Neuroscience*, vol. 2013, pp. 1–12, 2013, doi: 10.1155/2013/537218.

[22] L. Li, G. Xu, J. Xie, and M. Li, “Classification of single-trial motor imagery EEG by complexity regularization,” *Neural Computing and Applications*, vol. 31, no. 6, pp. 1959–1965, Jun. 2019, doi: 10.1007/s00521-017-3174-6.

[23] W. Samek, C. Vidaurre, K.-R. Müller, and M. Kawanabe, “Stationary common spatial patterns for brain–computer interfacing,” *Journal of Neural Engineering*, vol. 9, no. 2, pp. 1–14, Apr. 2012, doi: 10.1088/1741-2560/9/2/026013.

[24] F. Lotte and C. Guan, “Regularizing common spatial patterns to improve BCI designs: unified theory and new algorithms,” *IEEE Transactions on Biomedical Engineering*, vol. 58, no. 2, pp. 355–362, Feb. 2011, doi: 10.1109/TBME.2010.2082539.

[25] H. Lu, K. N. Plataniotis, and A. N. Venetsanopoulos, “Regularized common spatial patterns with generic learning for EEG signal classification,” in *2009 Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, Sep. 2009, pp. 6599–6602, doi: 10.1109/IEMBS.2009.5332554.

[26] H. Lu, H.-L. Eng, C. Guan, K. N. Plataniotis, and A. N. Venetsanopoulos, “Regularized common spatial pattern with aggregation for EEG classification in small-sample setting,” *IEEE Transactions on Biomedical Engineering*, vol. 57, no. 12, pp. 2936–2946, Dec. 2010, doi: 10.1109/TBME.2010.2082540.

[27] S. Shin, S.-J. Jang, D. Lee, U. Park, and J.-H. Kim, “Brainwave-based mood classification using regularized common spatial pattern filter,” *KSII Transactions on Internet and Information Systems (TIIS)*, vol. 10, no. 2, pp. 807–824, Feb. 2016, doi: 10.3837/tiis.2016.02.020.

[28] A. Yuksel and T. Olmez, “Task related and spatially regularized common spatial patterns for brain computer interfaces,” in *Proceedings IWBBIO 2014*, 2014, pp. 42–53.

[29] Y. Li, Y. Sun, F. Taya, H. Yu, N. Thakor, and A. Bezerianos, “Single trial EEG classification of lower-limb movements using improved regularized common spatial pattern,” in *2015 7th International IEEE/EMBS Conference on Neural Engineering (NER)*, Apr. 2015, pp. 1056–1059, doi: 10.1109/NER.2015.7146809.

[30] F. Lotte and C. Guan, “Spatially regularized common spatial patterns for EEG classification,” in *2010 20th International*

Conference on Pattern Recognition, Aug. 2010, pp. 3712–3715, doi: 10.1109/ICPR.2010.904.
[31] S.-H. Park, D. Lee, and S.-G. Lee, "Filter bank regularized common spatial pattern ensemble for small sample motor imagery classification," *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 26, no. 2, pp. 498–505, Feb. 2018, doi: 10.1109/TNSRE.2017.2757519.
[32] X. Li and H. Wang, "Smooth spatial filter for common spatial patterns," in *Neural Information Processing: 20th International Conference, ICONIP 2013*, 2013, pp. 315–322, doi: 10.1007/978-3-642-42042-9_40.
[33] M. Cheng, Z. Lu, and H. Wang, "Regularized common spatial patterns with subject-to-subject transfer of EEG signals," *Cognitive Neurodynamics*, vol. 11, no. 2, pp. 173–181, Apr. 2017, doi: 10.1007/s11571-016-9417-x.
[34] Z. M. Alhakeem and R. S. Ali, "Session to session transfer learning method using independent component analysis with regularized common spatial patterns for EEG-MI signals," *Iraqi Journal for Electrical and Electronic Engineering*, vol. 15, no. 1, pp. 13–27, Jun. 2019, doi: 10.37917/ijeee.15.1.2.
[35] Y. Xu *et al.*, "Transfer learning based on regularized common spatial patterns using cosine similarities of spatial filters for motor-imagery BCI," *Journal of Circuits, Systems and Computers*, vol. 28, no. 7, pp. 1–19, Jun. 2019, doi: 10.1142/S0218126619501238.
[36] H. Kang, Y. Nam, and S. Choi, "Composite common spatial pattern for subject-to-subject transfer," *IEEE Signal Processing Letters*, vol. 16, no. 8, pp. 683–686, Aug. 2009, doi: 10.1109/LSP.2009.2022557.
[37] A. Bamdadian, C. Guan, K. K. Ang, and J. Xu, "Improving session-to-session transfer performance of motor imagery-based BCI using adaptive extreme learning machine," in *2013 35th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, Jul. 2013, pp. 2188–2191, doi: 10.1109/EMBC.2013.6609969.
[38] H. Cho, M. Ahn, K. Kim, and S. C. Jun, "Increasing session-to-session transfer in a brain–computer interface with on-site background noise acquisition," *Journal of Neural Engineering*, vol. 12, no. 6, pp. 1–15, Dec. 2015, doi: 10.1088/1741-2560/12/6/066009.
[39] W. Wu, X. Gao, and S. Gao, "One-versus-the-rest (OVR) algorithm: an extension of common spatial patterns (CSP) algorithm to multi-class case," in *2005 IEEE Engineering in Medicine and Biology 27th Annual Conference*, 2005, pp. 2387–2390, doi: 10.1109/IEMBS.2005.1616947.
[40] M. Grosse-Wentrup and M. Buss, "Multiclass common spatial patterns and information theoretic feature extraction," *IEEE Transactions on Biomedical Engineering*, vol. 55, no. 8, pp. 1991–2000, Aug. 2008, doi: 10.1109/TBME.2008.921154.
[41] H. Sun, L. Bi, and B. Chen, "Partitioned common spatial pattern method for single trial EEG signal classification in brain-computer interface system," *Automatika*, vol. 57, no. 1, pp. 66–75, Jan. 2016, doi: 10.7305/automatika.2016.07.1078.
[42] C. J. C. Burges, "A tutorial on support vector machines for pattern recognition," *Data Mining and Knowledge Discovery*, vol. 2, no. 2, pp. 121–167, 1998, doi: 10.1023/A:1009715923555.
[43] Z. M. Alhakeem, H. Hakim, and O. A. Hassan, "Motor Imagery Electroencephalogram Datasets, Where and How to Find Them: A Review," *Iraq university college journal of engineering and applied sciences (IUCJEAS)*, vol. 1, no. 2, 2021.

BIOGRAPHIES OF AUTHORS

Zaineb M. Alhakeem was born in Basrah, Iraq in 1983. She received the B.Sc. and M.Sc. degrees in Computer Engineering from the University of Basrah, in 2005 and 2008 respectively. She is also had her Ph.D. degree within the Department of Electrical Engineering, College of Engineering in 2020, University of Basrah, Basrah, Iraq. Her teaching interests covering wide areas of modules across the Department of Computer Engineering and Communication Engineering, Iraq University College, include neural networks, distributed systems, control systems, modelling and simulation, microprocessor and microcontrollers and programming languages. Her research interests include biomedical engineering, digital signal processing, artificial intelligence, and renewable energy. She is a Member in the IEEE. She can be contacted at email: zainebalhakeem@buog.edu.iq.

Ramzy S. Ali received the B.Sc. and M.Sc. degrees in Electrical Engineering and Control and Computers Engineering from the University of Basrah, Basrah, Iraq in 1985 and 1989 respectively. He also received his Ph.D. degree in the Field of Control and Systems from the Saint Petersburg State Polytechnic University, Russia in 2003. He is currently a Professor at the University of Basrah. His teaching interests covering wide areas of modules across the Department of Electrical Engineering, University of Basrah, include intelligent control systems, robust control systems, microprocessor and microcontrollers, and industrial automation. He currently serves as a co-editor of the Basrah Journal for Engineering Sciences. His research interests include intelligent control of robotics, computational intelligence, chaos and nonlinear dynamics, renewable electrical energy systems, and PLC applications in industrial and engineering education. He is a Senior Member of the IEEE. He can be contacted at email: rsalwaily@ieee.org.