

Article

# A Domain Adaptation-Based Method for Classification of Motor Imagery EEG

Changsheng Li , Minyou Chen and Li Zhang \*

State Key Laboratory of Power Transmission Equipment & System Security and New Technology,  
School of Electrical Engineering, Chongqing University, Chongqing 400044, China; lichsh@cqu.edu.cn (C.L.);  
mchencqu@126.com (M.C.)

\* Correspondence: zldy02@cqu.edu.cn

**Abstract:** Non-stationarity of EEG signals lead to high variability across sessions, which results in low classification accuracy. To reduce the inter-session variability, an unsupervised domain adaptation method is proposed. Arithmetic mean and covariance are exploited to represent the data distribution. First, overall mean alignment is conducted between the source and target data. Then, the data in the target domain is labeled by a classifier trained with the source data. The per-class mean and covariance of the target data are estimated based on the predicted labels. Next, an alignment from the source domain to the target domain is performed according to the covariance of each class in the target domain. Finally, per-class mean adaptation is required after covariance alignment to remove the shift of data distribution caused by covariance alignment. Two public BCI competition datasets, namely the BCI competition III dataset IVa and the BCI competition IV dataset IIa were used to evaluate the proposed method. On both datasets, the proposed method effectively improved classification accuracy.

**Keywords:** electroencephalogram; motor imagery; mean and covariance; domain adaptation; linear discriminant analysis

**MSC:** 68U04



**Citation:** Li, C.; Chen, M.; Zhang, L. A Domain Adaptation-Based Method for Classification of Motor Imagery EEG. *Mathematics* **2022**, *10*, 1588. <https://doi.org/10.3390/math10091588>

Academic Editor: Luis Javier García Villalba

Received: 7 April 2022

Accepted: 5 May 2022

Published: 7 May 2022

**Publisher's Note:** MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Copyright:** © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

## 1. Introduction

A brain–computer interface (BCI) provides an alternate link to the external world for a subject using brain signals [1]. BCIs are especially useful for patients with impaired peripheral nerve or muscle functions to rebuild connection to the real world. For healthy people, BCIs can also provide a new control dimension, such as in games [2].

Motor imagery (MI)-based BCIs are the main type of BCIs. MI-BCIs are driven by neural signals modulated by users' voluntary movement imagination [3]. MI-related characteristic changes occur in some regions of the brain, especially the primary sensorimotor area and supplementary motor area [4]. These changes can be acquired through electroencephalograms (EEG). MI information in EEG can be captured through spatial filtering methods, such as common spatial pattern (CSP) [5], and then classified by machine learning methods to identify the intention of BCI users.

In machine learning theory, the data distributions of the training and test sets are supposed to be similar [6]. However, the non-stationarity of EEG leads to high variability across different recording sessions [7], which becomes a major obstacle to the accurate classification of EEG patterns. Studies indicate that the non-stationarity nature of EEG is induced by several factors, including physiological artifacts, state of subjects, and instrumental artifacts over different sessions [8]. To overcome this problem, transfer learning methods are proposed [9,10].

Transfer learning is widely investigated, including inductive transfer learning, transductive transfer learning, and unsupervised transfer learning [11]. Specifically, inductive

transfer learning is employed when the source and target tasks are different [12]. In this case, the source data is not labeled. Unsupervised transfer learning deals with the situation where the source and target tasks are similar, but labels of all data are unavailable. Transductive transfer learning deals with the situation in which the source data are labeled while the target data are unlabeled. Domain adaptation is related to transductive transfer learning, which is implemented based on the assumption that the source and target data are generated in the same task and distributed in different domains. Moreover, pre-trained models (with [13] or without [14] fine-tuning) are also introduced to search the shared feature patterns and reduce calibration with deep learning methods. Domain adaptation methods can easily learn the knowledge of labeled data in the source domain and transfer the knowledge to the target domain.

Recently, domain adaptation became a widely investigated approach to bridge the gap between the training set (i.e., the source domain) and the test set (i.e., the target domain) in the BCI field [15]. This technology aims at reducing the distribution shifting of the source and target domains. [16]. However, two issues arise, namely, the measurement of the discrepancy of different domains and the reduction of the discrepancy. Different measurement criteria are used to estimate the discrepancy between the source and target domains. Arithmetic mean, covariance, and correlation coefficients of data in source and target domains are frequently used as the statistical characteristics of distributions when performing alignment in data space [17–20]. Zheng et al. proposed a model based on transfer learning in which the mean and variance were used as the statistical characteristics shared across sessions [21]. Liang et al. aligned the tangent space mapping features according to the Riemannian center to handle the instability of channel covariance across sessions [22]. Azab et al. evaluated the distance between EEG data in the two domains using Kullback–Leibler (KL) divergence [12]. The spatial projection filters were then weighted by the divergence. In the regularized common spatial pattern (RCSP) method, optimization was used to obtain the spatial project matrix of the source and target EEG data simultaneously to minimize the distance between them [23]. The similarities between the two domains were also evaluated using other distance metrics, such as the Frobenius norm, Bhattacharyya distance, and cosine distance [24–26].

The single statistical characteristic that was investigated in previous studies may not estimate the data distribution properly, which may further influence the discrepancy minimization between the source and target domains. Furthermore, the estimation of the data distribution of the unlabeled data in the target domain is difficult. Mean and covariance are two important characteristics of data distribution. Due to the non-stationarity of EEG, they may be quite different in the source and target domains of BCI [7]. For this reason, a classifier trained with the source data may not perform well on the target data. Generally, the performance of classification depends on the similarity of the two domains [27]. Thus, transferring the data distribution of the source domain to the target domain may result in a more accurate classification.

To minimize the discrepancy between the source and target domains, an unsupervised method based on the alignment in Euclidean space is proposed in this paper. There are two major contributions. First, the discrepancy of distribution between the two domains is removed completely using this method. Second, the mean alignment and covariance alignment for each class can be realized simultaneously. In this method, mean and covariance are used to represent the center and dispersion of the data distribution in the two domains. The alignment contains two stages, namely the mean alignment (MA) and the per-class covariance and mean alignment (CMA). In the MA stage, considering that the mean and covariance of each class in the source and target domains may vary enormously and the data in the target domain remain unlabeled, the mean of the source data is aligned with the data mean in the target domain. The data in the target domain remain unchanged. After MA, a linear discriminant analysis (LDA) classifier is trained using the source data. Then, the LDA classifier is used to classify the target data. According to the predicted labels estimated by the classifier, the per-class basis covariance (CA) of the target data is

calculated. The covariance of the source data is then transformed to be the same as the covariance of target data for each class. After the covariance alignment, per-class basis MA is conducted again.

The remainder of this paper includes four sections. Section 2 introduces the details of the proposed method. In Section 3, the experimental datasets are introduced, and the experimental results are given. Section 4 presents the discussions, and Section 5 presents the conclusion.

## 2. Materials and Methods

A flowchart of the proposed method is shown in Figure 1. First, CSP is used for EEG feature extraction. The EEG features extracted from the training and test data are treated as the source and target data, respectively. Then, MA is performed in the source domain. After that, the LDA classifier is trained with the data in the source domain. Next, the trained classifier is used for the target data classification. CMA is then conducted in the two domains for each class. Finally, the classifier is retrained using the aligned source data.

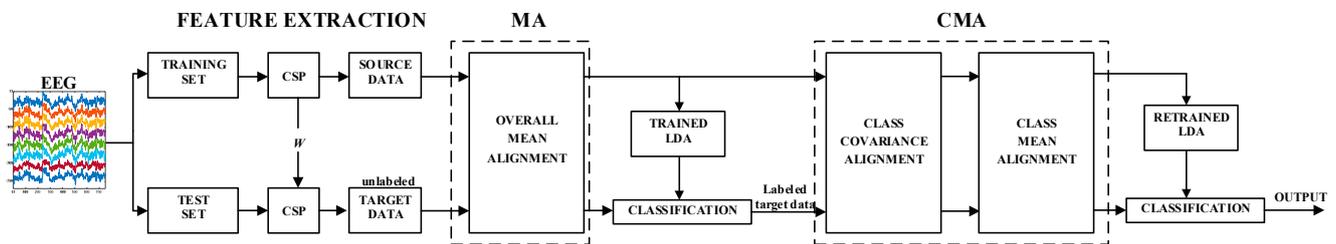


Figure 1. The flowchart of the CSP-MA-CMA method in this paper.

### 2.1. CSP for Feature Extraction

CSP was first proposed by Ramoser [28]. When given an EEG dataset that contains  $N$  trials in each class, the  $i$ -th trial of the  $l$ -th class ( $l = 1, 2$ ) is represented as  $E_i \in \mathbb{R}^{L \times S}$ , where  $L$  is the number of channels, and  $S$  is the number of sample points. The normalized covariance of the EEG per class is calculated by

$$R_l = \frac{1}{N} \sum_{i=1}^N \frac{E_i E_i^T}{\text{trace}(E_i E_i^T)} \tag{1}$$

where  $T$  is the transpose operator and  $\text{trace}(\cdot)$  calculates the sum of the diagonal elements of a matrix. A spatial filter can be found by solving an optimization problem given by [29].

$$\max_W J(W) = \frac{W^T R_1 W}{W^T R_2 W}, \text{ s.t. } \|W\|_2 = 1 \tag{2}$$

where  $\|*\|_2$  denotes the  $l$ -2 norm. After solving the above optimization problem through singular value decomposition, the spatial filter  $W$  can be obtained

$$R_1 W = \lambda R_2 W \tag{3}$$

The  $i$ -th trial  $E_i$  is then projected as

$$Z_i = W E_i \tag{4}$$

According to [28], the first and last  $n$  (usually,  $n = 3$ ) rows of  $\mathbf{Z}_i$  are selected to calculate the features. Since  $\mathbf{Z}_i$  and  $\mathbf{E}_i$  have the same number of rows, the CSP feature of the  $i$ -th trial can be represented as

$$f_i = \log \frac{\text{var}(\mathbf{Z}_{i,p})}{\sum_{p=1}^{2n} \text{var}(\mathbf{Z}_{i,p})} \tag{5}$$

where  $\text{var}(\cdot)$  denotes the covariance. Finally, six features can be obtained for each trial.

### 2.2. LDA for Classification

The aim of LDA is to find a projecting matrix to transform the data from different classes into a 2-dimensional space in which the inter-class scatter is maximized and the intra-class scatter is minimized [30]. First, the within-class scatter matrix  $\mathbf{S}_w$  is calculated by

$$\mathbf{S}_w = \mathbf{S}_1 + \mathbf{S}_2 \tag{6}$$

where  $\mathbf{S}_1$  and  $\mathbf{S}_2$  denote the class-scatter matrix of class 1 and class 2, respectively. They can be obtained by

$$\mathbf{S}_l = \sum_{i=1}^N (f_i - c_l)(f_i - c_l)^T, l = 1, 2 \tag{7}$$

where  $c_l$  denotes the per-class mean.

Then, the between class-scatter matrix  $\mathbf{S}_b$  is calculated by

$$\mathbf{S}_b = (c_1 - c_2)(c_1 - c_2)^T \tag{8}$$

Finally, after solving the following optimal problem, the linear projecting matrix  $\mathbf{M}$  can be obtained.

$$J(\mathbf{w}) = \frac{\mathbf{M}^T \mathbf{S}_b \mathbf{M}}{\mathbf{M}^T \mathbf{S}_w \mathbf{M}} \tag{9}$$

### 2.3. Overall Mean Alignment

The training and test features are treated as the source and target data, respectively. The mean of the data in the source domain and the mean of the data in the target domain are denoted as  $\mathbf{m}^s$  and  $\mathbf{m}^t$ , respectively, and can be utilized to measure the distance between the two domains.

$$\mathbf{d} = \mathbf{m}^s - \mathbf{m}^t \tag{10}$$

Then, the source data is aligned by

$$f_{i(MA)}^s = f_i^s - \mathbf{d} \tag{11}$$

where  $f_i^s$  is the  $i$ -th training feature calculated using CSP and  $f_{i(MA)}^s$  represents the feature after MA. Now the data in the source and target domains have the same center.

### 2.4. Per-Class Covariance and Mean Alignment

After overall MA, the data is aligned from the source domain to the target domain. The aligned data of the  $l$ -th class is denoted as  $\mathbf{F}_{l(MA)}^s$ .  $\mathbf{F}_{l(MA)}^s = [f_{i(MA)}^s]^T, I = 1, 2, \dots, N$ . The mean and covariance of  $\mathbf{F}_{l(MA)}^s$  are represented as  $\mathbf{m}_{l(MA)}^s$  and  $\mathbf{C}_{l(MA)}^s$ , respectively. The target data are labeled by LDA, which is trained with the source data after MA. The data of the  $l$ -th class in the target domain is denoted as  $\mathbf{F}_l^t$ , and its mean and covariance are represented as  $\mathbf{m}_l^t$  and  $\mathbf{C}_l^t$ , respectively. According to the label information, the per-class basis covariance alignment (CA) can be conducted using a transform matrix  $\mathbf{D}_l$ .

$$\mathbf{F}_{l(CA)}^s = \mathbf{D}_l \mathbf{F}_{l(MA)}^s \tag{12}$$

Let  $C_{l(CA)}^s$  and  $C_{l(MA)}^s$  represent the covariance of  $F_{l(CA)}^s$  and  $F_{l(MA)}^s$ , respectively.  $C_{l(CA)}^s$  can be represented using the method in [31].

$$C_{l(CA)}^s = D_l C_{l(MA)}^s D_l^T \tag{13}$$

The covariance of the source data is required to be the same as that of the target data,

$$C_{l(CA)}^s = C_l^t \tag{14}$$

So the transform matrix  $D_l$  can be calculated as

$$D_l = C_l^{t\frac{1}{2}} C_{l(MA)}^s^{-\frac{1}{2}} \tag{15}$$

After covariance alignment, the covariances of the data for the corresponding classes in the two domains are equal. However, it seems that the mean of the data in both domains is considerably different for each class [32]. So, per-class MA needs to be performed to achieve the mean alignment between the source and target data again. The per-class distance between the two domains is

$$d_l = m_{l(MA)}^s - m_{l(MA)}^t \tag{16}$$

The  $i$ -th data of the  $l$ -th class in the source domain, which is represented by the  $i$ -th row of  $F_{l(CA)}^s$ , namely  $f_{i,l(CA)}^s$ , is aligned to the target data.

$$f_{i,l(CMA)}^s = f_{i,l(CA)}^s - d_l \tag{17}$$

After MA and CMA, the per-class mean and covariance of the data in the two domains can be equal.

### 3. Results

Two datasets were used to evaluate the proposed method. The first dataset (dataset 1) is the BCI competition III dataset IVa [33]. Five subjects were guided to make the left hand, right hand, and right foot MI during EEG recording. For each subject, 280 EEG trials were obtained. The BCI competition IV dataset IIa [34] was used as the second dataset (dataset 2). Nine subjects participated in MI EEG recording. The recorded EEG data correspond to four classes of MIs (i.e., left hand, right hand, both feet, and tongue MIs). There are two sessions in dataset 2. Each session is comprised of 144 trials. The two sessions were recorded on two days. In this study, only two-class MI data (left hand vs. right hand) for each dataset were selected for classification. The EEG data were preprocessed by a band-pass filter (8 Hz to 30 Hz). In addition, classification accuracy was calculated based on five-fold cross validation.

The proposed CSP-MA-CMA method was compared with three competing methods, namely CSP, CSP-MA, and CSP-MA-CA. In all four methods, CSP and LDA were used for MI feature extraction and classification, respectively. In the CSP-MA method, after feature extraction by CSP, overall mean alignment was conducted. As for CSP-MA-CA, overall mean alignment and per-class covariance alignment were performed sequentially after CSP feature extraction. All the methods were tested with MATLAB 2016b on a PC with a 3.5 GHz processor and 8.0 GB RAM.

Tables 1 and 2 show the classification accuracies of the four methods on the two datasets. As shown in Table 1, compared with CSP, CSP-MA and CSP-MA-CMA improved the average accuracy by 1.3% and 2.3%, respectively, on dataset 1. As shown in Table 2, CSP-MA and CSP-MA-CMA improved the average accuracy by 3.1% and 3.3%, respectively, on dataset 2 compared to CSP. However, the classification performances of CSP-MA-CA on both datasets only reached the chance level. Moreover, an additional public EEG dataset with two-class data (i.e., left hand and right hand MIs) from 52 subjects is used to evaluate

the proposed method. For each subject, 200 trials were recorded during the experiment. More information on this dataset is provided in [35]. The average accuracies of CSP and CSP-MA-CMA (the proposed method) are 56.32% and 58.46%, respectively.

**Table 1.** Comparison of accuracies (%) using different alignment methods on dataset 1.

	AA	AL	AV	AW	AY	Mean
CSP	72.14	93.57	66.43	92.14	91.43	83.14
CSP-MA	75	96.43	68.57	89.29	92.86	84.43
CSP-MA-CA	50	50	49.29	50	50	49.86
CSP-MA-CMA	<b>75</b>	<b>96.43</b>	<b>70.71</b>	<b>92.14</b>	<b>92.86</b>	<b>85.43</b>

**Table 2.** Comparison of accuracies (%) using different alignment methods on dataset 2.

	A01	A02	A03	A04	A05	A06	A07	A08	A09	Mean
CSP	90.97	56.94	92.36	64.58	56.25	71.53	73.61	97.22	88.89	76.93
CSP-MA	<b>90.97</b>	58.33	99.31	<b>79.17</b>	58.33	71.53	77.08	97.22	88.19	80.02
CSP-MA-CA	50	50	50.69	50	50	50	50	50	50	50.08
CSP-MA-CMA	89.58	<b>59.03</b>	<b>99.31</b>	77.08	<b>59.03</b>	<b>72.22</b>	<b>79.86</b>	<b>97.22</b>	<b>89.58</b>	<b>80.32</b>

To test the statistically significant differences between different alignment methods, a paired *t*-test was used. There were no significant differences in classification accuracy between CSP and CSP-MA on dataset 1 ( $p > 0.1$ ) and dataset 2 ( $p = 0.09$ ), although CSP-MA increased the average accuracy. The performance of CSP-MA-CMA increased significantly compared with CSP on dataset 1 ( $p = 0.03$ ) and dataset 2 ( $p = 0.04$ ), as well as the additional dataset ( $p = 0.02$ ).

Datasets 1 and 2 are commonly used for MI EEG classification, whereas only a few methods not related to transfer learning are evaluated based on the above mentioned MI dataset from 52 subjects. Thus, the comparison between the proposed method and the methods in some previous studies was conducted only on datasets 1 and 2. As shown in Tables 3 and 4, the best average accuracy was achieved by the proposed method.

**Table 3.** Comparison of accuracies (%) of the proposed method and three existing methods on dataset 1.

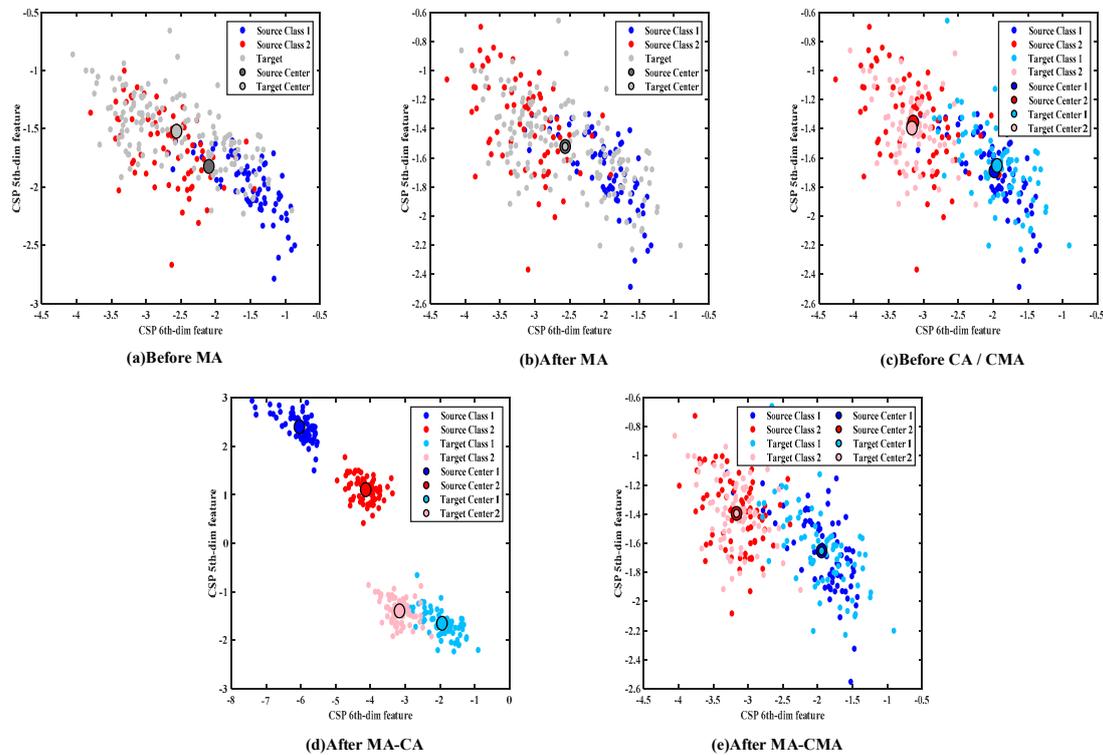
Methods	Year	AA	AL	AV	AW	AY	Mean
MSMV [36]	2021	<b>79.64</b>	94.64	75	78.57	<b>94.64</b>	84.51
p-LTCSP [37]	2020	77.68	<b>100</b>	71.94	<b>92.41</b>	74.21	83.25
MFCSP [38]	2021	77.68	<b>100</b>	<b>73.98</b>	84.82	88.1	84.91
Proposed		75	96.43	70.71	<b>92.41</b>	92.86	<b>85.43</b>

**Table 4.** Comparison of accuracies (%) of the proposed method and three existing methods on dataset 2.

Methods	Year	A01	A02	A03	A04	A05	A06	A07	A08	A09	Mean
SWCSP [39]	2021	86.11	<b>64.58</b>	95.83	64.58	68.06	68.75	<b>81.94</b>	<b>97.22</b>	90.97	79.78
CSCSP [40]	2019	88.89	63.89	95.14	69.44	<b>74.31</b>	65.97	72.92	92.36	88.19	79.01
DACSP [5]	2021	<b>91.67</b>	53.47	95.84	72.92	64.58	<b>73.61</b>	78.47	95.83	<b>92.37</b>	79.48
Proposed		89.58	59.03	<b>99.31</b>	<b>77.08</b>	59.03	72.22	79.86	<b>97.22</b>	89.58	<b>80.32</b>

The visualization of data distribution during the alignment process clearly shows how the data distribution was modified by the proposed CSP-MA-CMA method. The distribution of original features before and after different kinds of alignments in the two domains for subject 3 in dataset 2 is shown in Figure 2. To facilitate visualization, the last

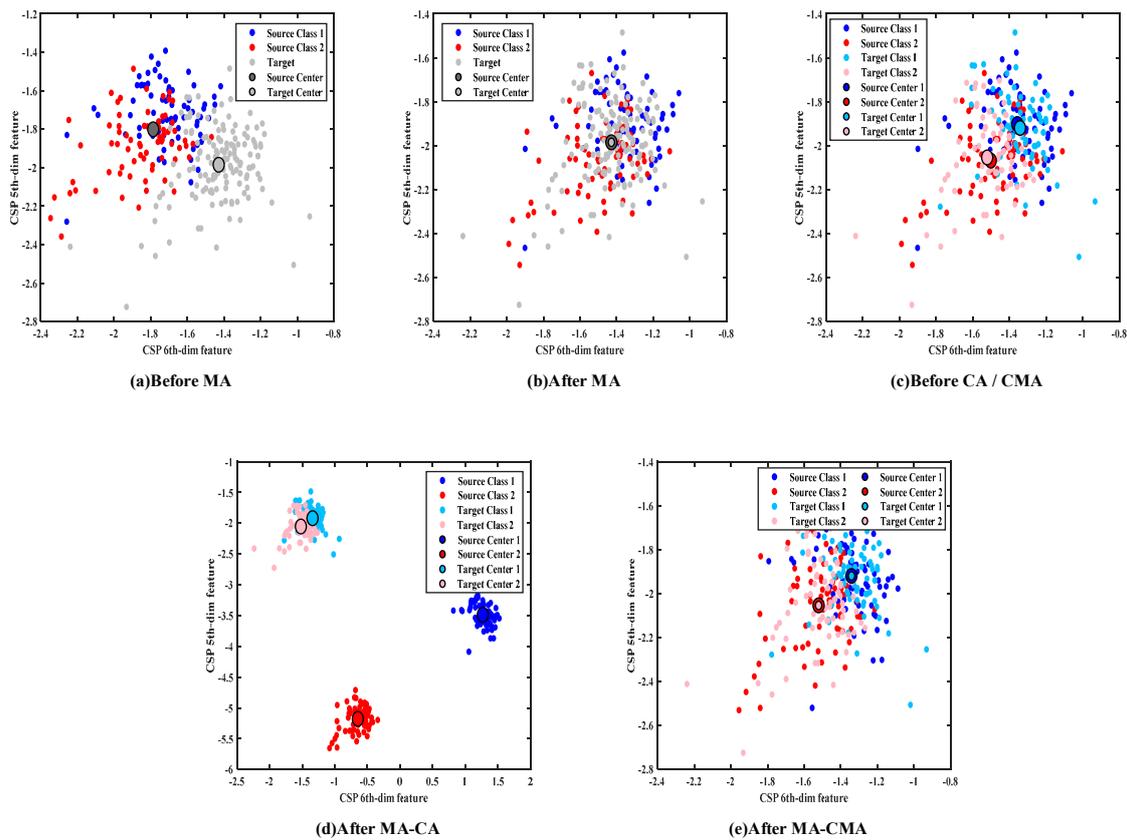
two dimensions of the EEG features, namely the sixth and fifth dimensions, were used for data distribution representation. The distribution before MA is shown in Figure 2a. The dots in dark blue are the samples of class 1, and those in dark red are the samples of class 2 in the source domain. The grey dots represent the unlabeled samples in the target domain. Two larger dots in dark grey and light grey show the overall centers in the two domains, respectively. They are different from each other before MA.



**Figure 2.** Visualization of data distribution during the alignment process for subject 3 on dataset 2. Data distribution (a) before MA, (b) after MA, (c) before CA/CMA, (d) after MA-CA, (e) after MA-CMA.

The data distribution after MA is shown in Figure 2b. The data in the source domain were aligned by MA to have the same center as the data in the target domain. The classifier was trained using the mean aligned data in the source domain and then used for data classification in the target domain. As shown in Figure 2c, the target data were divided into different classes by the pre-trained classifier, which are represented as class 1 in light blue and class 2 in light red, respectively. For each class, the mean and covariance of the data in the two domains were different. The data distribution after MA-CA and MA-CMA is shown in Figure 2d,e, respectively. In Figure 2d, although the per-class covariance in the two domains was forced to be equal after MA-CA, the difference between the means of the data in the two domains increased, which may be the reason for the low classification accuracy of the CSP-MA-CA method shown in Tables 1 and 2. Therefore, it is necessary to align the means of each class as the last step, which does not change the aligned class covariance in the two domains.

The data distribution of subject 5 in dataset 2 is shown in Figure 3. Similar results after each alignment step can be observed.



**Figure 3.** Visualization of data distribution during the alignment process for subject 5 on dataset 2. Data distribution (a) before MA, (b) after MA, (c) before CA/CMA, (d) after MA-CA, (e) after MA-CMA.

### 4. Discussion

The differences in the data distribution in the two domains before the MA step were evaluated, including the overall mean difference and the per-class covariance difference. The overall mean difference is indicated by the distance between the means of the data in the source and target domains. Per-class covariance difference is represented by the distance of class covariance in the two domains. The effect of MA was evaluated by the accuracy improvement expressed as the difference in accuracy between CSP (before MA) and CSP-MA (after MA). As shown in Table 5, the results are ranked from high to low according to the degree of improvement in accuracy on both datasets. It seems that after MA, a higher accuracy improvement can be achieved when the overall mean difference is larger for all subjects, except for the subjects A08 and A01 in dataset 1.

Table 6 show the average training and test time across all subjects required for CSP-MA-CMA and the three competing methods. Since there are no parameters for tuning in the whole processing, only a little time is required for data alignment. For CSP-MA-CMA, it took 1.038 s and 0.971 s for training on datasets 1 and 2, respectively. In the test phase, CSP-MA-CMA only cost 0.0022 s.

**Table 5.** The accuracy improvement after MA and the overall mean difference and per-class covariance difference in the two domains before MA.

		Improvement of Accuracies (%)	Overall Mean Difference	Per-Class Covariance Difference
Dataset 1	A04	12.5	0.66	0.13
	A03	6.94	0.64	0.17
	A07	6.25	0.55	0.07
	A05	2.78	0.45	0.06
	A02	2.08	0.41	0.17
	A06	0.69	0.40	0.22
	A09	0.69	0.14	0.47
	A08	0	0.26	0.09
Dataset 2	A01	−1.39	0.33	0.23
	AV	4.30	0.43	0.36
	AA	2.86	0.22	0.31
	AL	2.86	0.29	0.46
	AY	1.43	0.17	0.10
	AW	0	0.14	0.29

**Table 6.** Average time cost (s) using different alignment methods.

	Dataset 1		Dataset 2	
	Training Time (s)	Test Time (s)	Training Time (s)	Test Time (s)
CSP	0.4523	0.0016	0.4063	0.0016
CSP-MA	0.5754	0.0016	0.5111	0.0016
CSP-MA-CA	0.8981	0.0021	0.7841	0.0020
CSP-MA-CMA	1.038	0.0022	0.9710	0.0022

## 5. Conclusions

The classification of EEG is confronted with difficulty due to the high variability of EEG data recorded across different days. In this paper, CSP-MA-CMA was proposed to handle this problem. The proposed CSP-MA-CMA method was tested on two public MI datasets. CSP-MA-CMA significantly increased the performance compared with the competing method without alignment in feature space. Additionally, no parameters need to be determined in the proposed method.

**Author Contributions:** Conceptualization, M.C. and L.Z.; verification, C.L.; writing original draft, C.L.; review and editing, M.C. and L.Z. All authors have read and agreed to the published version of the manuscript.

**Funding:** This work was supported by the National Natural Science Foundation of China (Project No. 51977020).

**Institutional Review Board Statement:** Not applicable.

**Informed Consent Statement:** Not applicable.

**Data Availability Statement:** Data are available at <https://www.bbci.de/competition/> (accessed on 4 July 2021).

**Conflicts of Interest:** The authors declare no conflict of interest.

## References

1. Kubler, A.; Kotchoubey, B.; Kaiser, J.; Wolpaw, J.R.; Birbaumer, N. Brain-computer communication: Unlocking the locked in. *Psychol. Bull.* **2001**, *127*, 358–375. [[CrossRef](#)] [[PubMed](#)]
2. Nijholt, A. BCI for Games: A ‘State of the Art’ Survey. In Proceedings of the 7th International Conference on Entertainment Computing (ICEC 2008), Pittsburgh, PA, USA, 25–27 September 2008; pp. 225–228.

3. Wolpaw, J.R.; Birbaumer, N.; McFarland, D.J.; Pfurtscheller, G.; Vaughan, T.M. Brain-computer interfaces for communication and control. *Clin. Neurophysiol.* **2002**, *113*, 767–791. [[CrossRef](#)]
4. Pfurtscheller, G.; Neuper, C. Motor imagery activates primary sensorimotor area in humans. *Neurosci. Lett.* **1997**, *239*, 65–68. [[CrossRef](#)]
5. Wang, B.; Wong, C.M.; Kang, Z.; Liu, F.; Shui, C.; Wan, F.; Chen, C.L.P. Common Spatial Pattern Reformulated for Regularizations in Brain-Computer Interfaces. *IEEE Trans. Cybern.* **2021**, *51*, 5008–5020. [[CrossRef](#)] [[PubMed](#)]
6. Chai, X.; Wang, Q.; Zhao, Y.; Liu, X.; Bai, O.; Li, Y. Unsupervised domain adaptation techniques based on auto-encoder for non-stationary EEG-based emotion recognition. *Comput. Biol. Med.* **2016**, *79*, 205–214. [[CrossRef](#)]
7. Jayaram, V.; Alamgir, M.; Altun, Y.; Schoelkopf, B.; Grosse-Wentrup, M. Transfer Learning in Brain-Computer Interfaces. *IEEE Comput. Intell. Mag.* **2016**, *11*, 20–31. [[CrossRef](#)]
8. Bamdadian, A.; Guan, C.T.; Ang, K.K.; Xu, J.X. Improving session-to-session transfer performance of motor imagery-based BCI using Adaptive Extreme Learning Machine. In Proceedings of the 35th Annual International Conference of the IEEE Engineering in Medicine and Biology Society, Osaka, Japan, 3–7 July 2013; pp. 2188–2191.
9. Al-Saegh, A.; Dawwd, S.A.; Abdul-Jabbar, J.M. Deep learning for motor imagery EEG-based classification: A review. *Biomed. Signal Processing Control* **2021**, *63*, 102172. [[CrossRef](#)]
10. Huang, X.; Xu, Y.; Hua, J.; Yi, W.; Yin, H.; Hu, R.; Wang, S. A Review on Signal Processing Approaches to Reduce Calibration Time in EEG-Based Brain-Computer Interface. *Front. Neurosci.* **2021**, *15*, 1066. [[CrossRef](#)]
11. Pan, S.J.; Yang, Q. A Survey on Transfer Learning. *IEEE Trans. Knowl. Data Eng.* **2010**, *22*, 1345–1359. [[CrossRef](#)]
12. Azab, A.M.; Mihaylova, L.; Ang, K.K.; Arvaneh, M. Weighted Transfer Learning for Improving Motor Imagery-Based Brain-Computer Interface. *IEEE Trans. Neural Syst. Rehabil. Eng.* **2019**, *27*, 1352–1359. [[CrossRef](#)]
13. Zhang, D.; Yao, L.; Chen, K.; Wang, S. Ready for Use: Subject-Independent Movement Intention Recognition via a Convolutional Attention Model. In Proceedings of the 27th ACM International Conference on Information and Knowledge Management (CIKM), Torino, Italy, 22–26 October 2018; pp. 1763–1766.
14. Zhang, R.; Zong, Q.; Dou, L.; Zhao, X.; Tang, Y.; Li, Z. Hybrid deep neural network using transfer learning for EEG motor imagery decoding. *Biomed. Signal Processing Control* **2021**, *63*, 102144. [[CrossRef](#)]
15. Saha, S.; Baumert, M. Intra- and Inter-subject Variability in EEG-Based Sensorimotor Brain Computer Interface: A Review. *Front. Comput. Neurosci.* **2020**, *13*, 87. [[CrossRef](#)] [[PubMed](#)]
16. Fazli, S.; Daehne, S.; Samek, W.; Biessmann, F.; Mueller, K.-R. Learning From More Than One Data Source: Data Fusion Techniques for Sensorimotor Rhythm-Based Brain-Computer Interfaces. *Proc. IEEE* **2015**, *103*, 891–906. [[CrossRef](#)]
17. Abdi, L.; Hashemi, S. Unsupervised Domain Adaptation Based on Correlation Maximization. *IEEE Access* **2021**, *9*, 127054–127067. [[CrossRef](#)]
18. Li, P.; Ni, Z.; Zhu, X.; Song, J. Inter-class distribution alienation and inter-domain distribution alignment based on manifold embedding for domain adaptation. *J. Intell. Fuzzy Syst.* **2020**, *39*, 8149–8159. [[CrossRef](#)]
19. Zhang, W.; Zhang, X.; Lan, L.; Luo, Z. Maximum Mean and Covariance Discrepancy for Unsupervised Domain Adaptation. *Neural Processing Lett.* **2020**, *51*, 347–366. [[CrossRef](#)]
20. Lee, B.-H.; Jeong, J.-H.; Lee, S.-W. SessionNet: Feature Similarity-Based Weighted Ensemble Learning for Motor Imagery Classification. *IEEE Access* **2020**, *8*, 134524–134535. [[CrossRef](#)]
21. Zheng, M.; Yang, B.; Xie, Y. EEG classification across sessions and across subjects through transfer learning in motor imagery-based brain-machine interface system. *Med. Biol. Eng. Comput.* **2020**, *58*, 1515–1528. [[CrossRef](#)]
22. Liang, Y.; Ma, Y. A Cross-Session Feature Calibration Algorithm for Electroencephalogram-Based Motor Imagery Classification. *J. Med. Imaging Health Inform.* **2019**, *9*, 1534–1540. [[CrossRef](#)]
23. Cheng, M.; Lu, Z.; Wang, H. Regularized common spatial patterns with subject-to-subject transfer of EEG signals. *Cogn. Neurodyn.* **2017**, *11*, 173–181. [[CrossRef](#)]
24. Xu, Y.; Wei, Q.; Zhang, H.; Hu, R.; Liu, J.; Hua, J.; Guo, F. Transfer Learning Based on Regularized Common Spatial Patterns Using Cosine Similarities of Spatial Filters for Motor-Imagery BCI. *J. Circuits Syst. Comput.* **2019**, *28*, 1950123. [[CrossRef](#)]
25. Khalaf, A.; Akcakaya, M. A probabilistic approach for calibration time reduction in hybrid EEG-ftCD brain-computer interfaces. *Biomed. Eng. Online* **2020**, *19*, 295–314. [[CrossRef](#)] [[PubMed](#)]
26. Zheng, Q.; Zhu, F.; Qin, J.; Heng, P.-A. Multiclass support matrix machine for single trial EEG classification. *Neurocomputing* **2018**, *275*, 869–880. [[CrossRef](#)]
27. Lotte, F.; Bougrain, L.; Cichocki, A.; Clerc, M.; Congedo, M.; Rakotomamonjy, A.; Yger, F. A review of classification algorithms for EEG-based brain-computer interfaces: A 10 year update. *J. Neural Eng.* **2018**, *15*, 031005. [[CrossRef](#)]
28. Ramoser, H.; Muller-Gerking, J.; Pfurtscheller, G. Optimal spatial filtering of single trial EEG during imagined hand movement. *IEEE Trans. Rehabil. Eng.* **2000**, *8*, 441–446. [[CrossRef](#)]
29. Zhang, L.; Wen, D.; Li, C.; Zhu, R. Ensemble classifier based on optimized extreme learning machine for motor imagery classification. *J. Neural Eng.* **2020**, *17*, 026004. [[CrossRef](#)]
30. Tao, D.; Li, X.; Wu, X.; Maybank, S.J. Geometric Mean for Subspace Selection. *IEEE Trans. Pattern Anal. Mach. Intell.* **2009**, *31*, 260–274.
31. Li, Y.; Wei, Q.; Chen, Y.; Zhou, X. Transfer Learning Based on Hybrid Riemannian and Euclidean Space Data Alignment and Subject Selection in Brain-Computer Interfaces. *IEEE Access* **2021**, *9*, 6201–6212. [[CrossRef](#)]

32. Ma, L.; Crawford, M.M.; Zhu, L.; Liu, Y. Centroid and Covariance Alignment-Based Domain Adaptation for Unsupervised Classification of Remote Sensing Images. *IEEE Trans. Geosci. Remote Sens.* **2019**, *57*, 2305–2323. [[CrossRef](#)]
33. Dornhege, G.; Blankertz, B.; Curio, G.; Muller, K.R. Boosting bit rates in noninvasive EEG single-trial classifications by feature combination and multiclass paradigms. *IEEE Trans. Bio-Med. Eng.* **2004**, *51*, 993–1002. [[CrossRef](#)]
34. Fatourechhi, M.; Bashashati, A.; Ward, R.K.; Birch, G.E. EMG and EOG artifacts in brain computer interface systems: A survey. *Clin. Neurophysiol.* **2007**, *118*, 480–494. [[CrossRef](#)] [[PubMed](#)]
35. Cho, H.; Ahn, M.; Ahn, S.; Kwon, M.; Jun, S.C. EEG datasets for motor imagery brain-computer interface. *GigaScience* **2017**, *6*, 1–8. [[CrossRef](#)] [[PubMed](#)]
36. Padfield, N.; Ren, J.; Qing, C.; Murray, P.; Zhao, H.; Zheng, J. Multi-segment Majority Voting Decision Fusion for MI EEG Brain-Computer Interfacing. *Cogn. Comput.* **2021**, *13*, 1484–1495. [[CrossRef](#)]
37. Yu, Z.; Ma, T.; Fang, N.; Wang, H.; Li, Z.; Fan, H. Local temporal common spatial patterns modulated with phase locking value. *Biomed. Signal Processing Control* **2020**, *59*, 101882. [[CrossRef](#)]
38. Hou, Y.; Chen, T.; Lun, X.; Wang, F. A novel method for classification of multi-class motor imagery tasks based on feature fusion. *Neurosci. Res.* **2021**, *176*, 40–48. [[CrossRef](#)]
39. Gaur, P.; Gupta, H.; Chowdhury, A.; McCreddie, K.; Pachori, R.B.; Wang, H. A Sliding Window Common Spatial Pattern for Enhancing Motor Imagery Classification in EEG-BCI. *IEEE Trans. Instrum. Meas.* **2021**, *70*, 1–9. [[CrossRef](#)]
40. Raza, H.; Rathee, D.; Zhou, S.-M.; Cecotti, H.; Prasad, G. Covariate shift estimation based adaptive ensemble learning for handling non-stationarity in motor imagery related EEG-based brain-computer interface. *Neurocomputing* **2019**, *343*, 154–166. [[CrossRef](#)]