



*Research article*

## **Multi-source online transfer algorithm based on source domain selection for EEG classification**

**Zizhuo Wu<sup>1</sup>, Qingshan She<sup>1,\*</sup>, Zhelong Hou<sup>1</sup>, Zhenyu Li<sup>1</sup>, Kun Tian<sup>2</sup> and Yuliang Ma<sup>1</sup>**

<sup>1</sup> School of Automation, Hangzhou Dianzi University, Hangzhou, Zhejiang 310018, China

<sup>2</sup> Zhejiang Kende Mechanical & Electrical Corporation

\* **Correspondence:** Email: [qsshe@hdu.edu.cn](mailto:qsshe@hdu.edu.cn).

**Abstract:** The non-stationary nature of electroencephalography (EEG) signals and individual variability makes it challenging to obtain EEG signals from users by utilizing brain-computer interface techniques. Most of the existing transfer learning methods are based on batch learning in offline mode, which cannot adapt well to the changes generated by EEG signals in the online situation. To address this problem, a multi-source online migrating EEG classification algorithm based on source domain selection is proposed in this paper. By utilizing a small number of labeled samples from the target domain, the source domain selection method selects the source domain data similar to the target data from multiple source domains. After training a classifier for each source domain, the proposed method adjusts the weight coefficients of each classifier according to the prediction results to avoid the negative transfer problem. This algorithm was applied to two publicly available motor imagery EEG datasets, namely, BCI Competition IV Dataset IIa and BNCI Horizon 2020 Dataset 2, and it achieved average accuracies of 79.29 and 70.86%, respectively, which are superior to those of several multi-source online transfer algorithms, confirming the effectiveness of the proposed algorithm.

**Keywords:** online transfer learning; brain-computer interface; motor imagery; source domain selection; data alignment

---

## 1. Introduction

As the most complex and important organ in the human body, the brain plays an indispensable role in regulating the body's motor functions and maintaining higher cognitive activities such as consciousness, language, memory, sensation and emotion. Therefore, studying the brain not only lets us have a better understanding of the brain, but it can also provide auxiliary methods for the treatment and diagnosis of brain disease. In recent years, many countries around the world have attached great importance to brain science research; researchers have started to study the brain and formed a unique brain science knowledge system. The brain-computer interface (BCI) [1,2] is a communication and control technology that directly translates the perceptual thinking generated by the brain into the corresponding action of external devices, which is of high scientific research value and is currently being widely applied in medical health, vehicle driving [3], daily life entertainment [4], etc. However, BCI systems still have problems that need to be solved urgently, such as a long user training time and limited online performance [5].

Currently, many advanced machine learning algorithms have been proposed and used in electroencephalography (EEG) classification, including Bayesian classifiers [6], support vector machines [7], linear discriminant analysis (LDA) [8], adaptive classifiers and deep learning classifiers [9]. In recent years, because transfer learning can take advantage of similarities between data, tasks or models, it has been utilized to analyze EEG signals with individual differences and those that are non-stationary, making it possible to apply models and the knowledge learned in the old domain to the new domain. In this way, not only can the classification performance of unlabeled data be improved by using labeled data, but the model training time can also be drastically reduced [10–12]. In transfer learning, a domain is defined as follows [13–14]: A domain  $\mathcal{D}$  consists of a feature space  $\mathcal{X}$  and its associated marginal probability distribution  $P(X)$ , i.e.,  $\mathcal{D} = \{\mathcal{X}, P(X)\}$ , where  $X \in \mathcal{X}$ . A source domain  $\mathcal{D}_s$  and a target domain  $\mathcal{D}_t$  are different if they have different feature spaces, i.e.,  $\mathcal{X}_s \neq \mathcal{X}_t$ , and/or different marginal probability distributions, i.e.,  $P_s(X) \neq P_t(X)$ .

However, most of the existing transfer learning methods train classifiers by batch learning in an offline manner, where all source and target data are pre-given [15,16], whereas this assumption may not be practical in real application scenarios where collecting enough data is very time-consuming. In addition, data are often transferred in a stream and cannot be collected in their entirety. Therefore, researchers have introduced online learning into the field of transfer learning and proposed an online transfer learning (OTL) framework. Unlike online learning, which merely considers the dynamic changes of data on a data domain, OTL also considers the changes in the data distributions in the source domain and target domain. OTL has been used in areas such as online feature selection [17] and graphical retrieval [18]. It combines the advantages of dynamically updating classification models for online learning with the ability to effectively exploit knowledge from source domains of transfer learning [19], aiming to apply online learning tasks to target domains by transferring knowledge from the source domain. Existing OTL approaches focus on how to use knowledge from the source domain for online learning in the target domain. Most of them use strategies based on integrated learning, i.e., directly combining source and target classifiers. Zhao et al. [20] proposed a single-source domain OTL algorithm in homogeneous and heterogeneous spaces, and it dynamically weights the combination of source and target classifiers to form the final classifier. Kang et al. [21] proposed an OTL algorithm for multi-class classification, utilizing a new loss function and updating method to extend the OTL of binary classification to multi-class tasks. Based on Zhao's work, Ge et al. [22] first proposed a multi-source online transfer framework. Zhou et al. [23] proposed an online transfer strategy, reusing the features extracted from a pre-trained model, and it can achieve automatic feature extraction and

achieve good classification accuracy. However, the OTL algorithm can only handle single-source domain transfer, which can easily lead to negative transfer [13] when facing multiple source domains. Negative transfer may happen when the source domain data and task contribute to the reduced performance of learning in the target domain.

Nevertheless, single-source transfer learning requires a high similarity between the source domain samples and the target domain samples during training. For example, in EEG signal classification, the class features of source subjects and target subjects are required to be as similar as possible. Since the knowledge that can be provided by a single source domain is limited, researchers tend to naturally think of applying transfer learning to the target domain by exploiting knowledge from multiple source domains. To utilize knowledge from multiple source domains, researchers have developed some boosting-based algorithms to adjust the weights of different domains or samples [24–26]. Eaton and DesJardins [24] proposed a method that uses AdaBoost to assign higher weights to source domains that are more similar to the target domain and adjust the weights of individual samples in each source domain. Yao and Doretto [25] proposed a multi-source transfer method that first integrates the knowledge from multiple source domains and then migrates the knowledge to the target domain; it achieved better performance on some benchmark datasets. Tan et al. [26] utilized different views from different source domains to assist the target domain task. Jiang et al. [27] proposed a general multi-source transfer framework to preserve independent information between different tasks. However, these studies assume that the data have already been known when dealing with target domain data, and they conducted transfer only via offline batch learning. Recently, after noticing the shortcomings of OTL, Wu et al. [17] proposed a HomOTLMS algorithm to train a classifier in each source domain and classify samples by weighting the classifiers in the source and target domains. Du et al. [28] proposed the HomOTL-ODDM algorithm to update the mapping matrix in an online manner, thus further reducing the differences between domains; the experimental results showed that the transfer effect can be significantly improved after considering the differences in data distributions. Li et al. [29] proposed an online EEG classification method based on instance transfer (OECIT). It can align the EEG data online, and combines with HomOTL, greatly reducing computational and memory costs.

To address the negative transfer problem in online single-source transfer and further reduce the individual differences between subjects, this paper presents a multi-source online transfer EEG classification algorithm based on source domain selection (SDS). Different from the HomOTLMS and HomOTL-ODDM algorithms, the proposed algorithm dynamically selects several suitable source domains and uses the multi-source online transfer algorithm to train a classifier in each selected source domain; it then combines the trained classifier with the target domain classifier for weighting, which not only reduces the training time but also achieves better classification results in the target domain. The algorithm was applied to two publicly available motor imagery EEG datasets and analyzed in comparison with the same type of multi-source online transfer algorithms to confirm the superiority of this algorithm.

## 2. Materials and methods

### 2.1. Euclidean alignment (EA)

Suppose a subject has  $n$  trials; we can have

$$\bar{\mathbf{E}} = \frac{1}{n} \sum_{i=1}^n \mathbf{X}_i \mathbf{X}_i^T \quad (1)$$

where  $\bar{E}$  is the Euclidean mean of all EEG trials from a subject and  $X_i \in \mathbb{R}^{c \times t}$  denotes the trial segments of EEG signals, where  $c$  is the number of EEG channels and  $t$  is the number of time samples. Then, we perform alignment by using

$$\tilde{X}_i = \bar{E}^{-1/2} X_i \quad (2)$$

## 2.2. Common spatial pattern (CSP) feature extraction

Given a spatiotemporal EEG signal  $X$  of one  $N$ -channel, where  $X$  is an  $N \times T$  matrix and  $T$  denotes the number of samples per channel, the normalized covariance matrix of the EEG signal is shown below.

$$C = \frac{XX^T}{\text{trace}(XX^T)} \quad (3)$$

The covariance matrices  $C_1$  and  $C_2$  for each category can be calculated from the sample means. The projection matrix of CSP is

$$W_{csp} = U^T P \quad (4)$$

where  $U$  is the orthogonal matrix and  $P$  is the whitening feature matrix.

After filtering with the projection matrix  $W_{csp}$ , the feature matrix is obtained:

$$Z_0 = W_{csp}^T X \quad (5)$$

The first  $m$  rows and the last  $m$  rows of the feature matrix are taken to construct the matrix  $Z = (z_1 \ z_2 \ \dots \ z_{2m}) \in R^{N \times 2m}$ . The feature vector  $F = (f_1 \ f_2 \ \dots \ f_{2m})^T \in R^{2m \times 1}$  can be obtained after normalization.

## 2.3. SDS

The SDS method can select the nearest source domain to reduce computational cost and potentially improve classification performance.

Supposing that there are  $Z$  different source domains, for the  $z$ -th source domain, the average feature vector for each source domain class  $m_{z,c}$  ( $c = 1, 2$ ) is calculated first. After obtaining a small number of target domain labels, the known label information is used to calculate the average feature vector for each target domain class  $m_{t,c}$ . Then, the distance between the two domains is expressed as

$$d(z, t) = \sum_{c=1}^2 \|m_{z,c} - m_{t,c}\| \quad (6)$$

The next step is to cluster these distances  $\{d(z, t)\}_{z=1, \dots, Z}$  to the  $k$  means. Assuming that the clusters are divided into  $(C_1, C_2, \dots, C_k)$ , the objective is to minimize the squared error  $E$ , which can

be expressed as

$$E = \sum_{i=1}^k \sum_{x \in C_i} \|d(z, t) - \mu_i\|_2^2 \quad (7)$$

where  $\mu_i$  is the mean vector of cluster  $C_i$ , which is also known as a centroid, with the following expression:

$$\mu_i = \frac{1}{|C_i|} \sum_{x \in C_i} d(z, t) \quad (8)$$

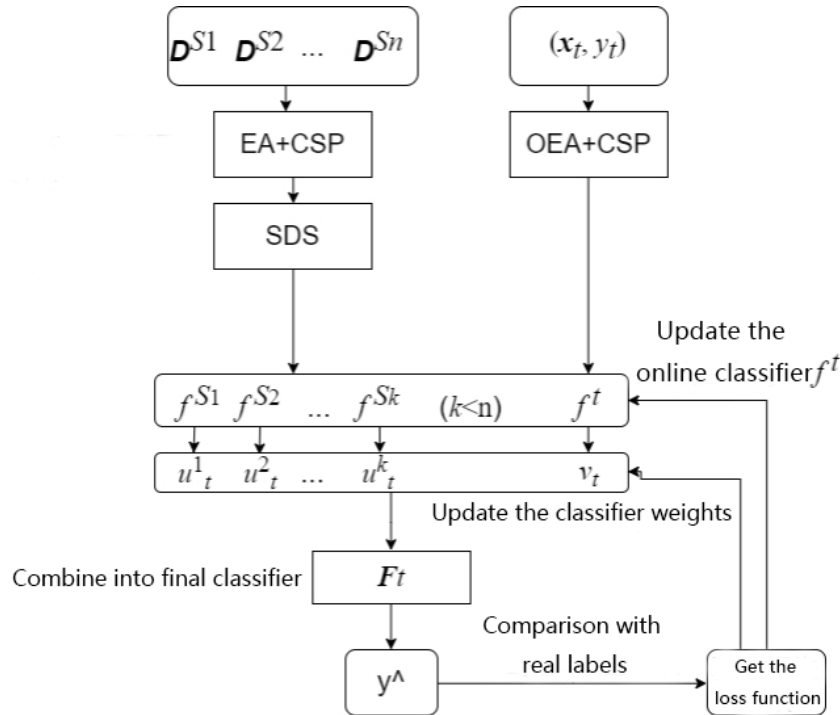
Finally, the cluster with the smallest center of mass, i.e., the source domain closest to the target domain, is selected. In this way, OTL is performed for only  $Z/l$  source domains where  $l$  is the number of clusters of the k-means cluster. The larger the value of  $l$ , the lower the computational cost. However, when  $l$  is too large, there may not be enough source domains for classification, resulting in unstable classification performance. Therefore, to minimize computational cost and improve classification performance,  $l$  was set to equal 2.

#### 2.4. Multi-source OTL

In this section, a multi-source OTL with SDS (MSOTL-SDS) algorithm is proposed. The algorithm flow is shown in Figure 1.

We set  $n$  source domains  $D^S = \{D^{S_1}, D^{S_2}, \dots, D^{S_n}\}$  and a certain number of labeled samples  $\{(x_t, y_t) | t = 1, 2, \dots, m\}$  from the target domain  $D^T$ . For the  $i$ -th source domain  $D^{S_i}$ ,  $x^{S_i} \times y$  represents the source domain data space,  $x^{S_i} \in R^{d_i}$  represents the feature space and  $y = \{-1, +1\}$  represents the label space.  $f^{S_i}$  represents the classifier learned on the second source domain.  $x \times y$  denotes the target domain data space, the feature space is  $x \in R^d$  and the target domain and source domain share the same label space  $y$ .

When the source domain data are extracted by domain alignment and CSP features, the source domains differing largely from the target domain samples are eliminated by SDS, which leaves only  $k$  ( $1 < k < n$ ) source domains; the classifier is trained by training each of these  $k$  source domains  $D^S = \{D^{S_1}, D^{S_2}, \dots, D^{S_k}\}$  and then assigning weights accordingly. In online mode, after domain alignment and feature extraction by online Euclidean space data alignment (OEA) and CSP, the samples  $(x_t, y_t)$  are sent to a classifier  $F_t(\cdot)$ , which is weighted by all classifiers, for label prediction. After obtaining the true label, the weight of the classifier and the online target domain classifier are updated by the loss function.



**Figure 1.** Flowchart for the multi-source OTL algorithm based on SDS.

Assume that the target domain classifier is given as follows:

$$f^T(x) = \sum_{i=1}^t \alpha_i y_i k(x_i, x) \quad (9)$$

where  $\alpha_i$  is the coefficient of the  $i$ -th target sample.

Set the weight vector of the source domain classifier as  $u_t = (u_t^1, u_t^2, \dots, u_t^n)^T$ , construct the target domain weight variable  $v$  and apply the Hedge algorithm to dynamically update the weights of the source and target domain classifiers as follows:

$$u_{t+1}^i = u_t^i \beta^{Z_t^i}, Z_t^i = I(\text{sign}(y_t f^{S_i}(x_t)) < 0), i = 1, 2, \dots, n \quad (10)$$

$$v_{t+1} = v_t \beta^{Z_t^v}, Z_t^v = I(\text{sign}(y_t f_t^T(x_t)) < 0), i = 1, 2, \dots, n \quad (11)$$

Finally, its class label is predicted by the following prediction function:

$$\hat{y}_t = \text{sign}\left(\sum_{i=1}^n p_t^i f^{S_i}(x_t) + p_t^v f_t^T(x_t)\right) \quad (12)$$

where  $p_t^i$  and  $p_t^v$  correspond to the weights of the classifiers in the source domain and target domain, respectively, and the calculation formula is given as follows:

$$p_t^i = \frac{u_t^i}{\sum_{j=1}^n u_t^j + v_t}, p_t^v = \frac{v_t}{\sum_{j=1}^n u_t^j + v_t} \quad (13)$$

**Algorithm 1 MSOTL-SDS algorithm**

**Input:** source domain classifier  $f^S = (f^{S_1}, f^{S_2}, \dots, f^{S_k})$ , initial trade-off parameter  $C$ , discount weight  $\beta \in (0,1)$

**Initialization:** online classifier  $f_1^T = 0$ , weight  $u_1 = \frac{1}{n+1}$ ,  $v_1 = \frac{1}{n+1}$

// SDS section

**for**  $s = 1, 2, \dots, k$  **do**

The distance  $d(s, t)$  from the  $s$ th source domain to the target domain is calculated by using Eq (6)

end if

The k-means clustering of  $\{d(S, t)\}_{s=1, \dots, k}$  is performed to select the source domain with a smaller center of mass to form  $S' = S_1, S_2, \dots, S_n$  ( $n < k$ )

// Multi-source online transfer

**for**  $t = 1, 2, \dots, m$  **do**

Receiving test specimens.  $x_t \in X$

Calculate the classifier weights by using Eq (10)

The predicted label is obtained by using Eq (12)

Receive real tags  $y_t = \{-1, +1\}$

The weights of each classifier are updated by Eqs (10) and (11)

Calculate the loss function  $l_t = [1 - y_t f_t^T x_t]_+$

**if**  $l_k > 0$  **then**

$f_{t+1}^T = f_t^T + \tau_t y_t x_t$ , where  $\tau_t = \min\{C, l_t / \|x_t\|^2\}$

**end if**

**end for**

**output:**  $f_t(x) = \text{sign}\left(\sum_{i=1}^n p_t^i f^{S_i}(x_t) + p_t^v f_t^T(x_t)\right)$

### 3. Experiments and result analysis

#### 3.1. Data description

Here, the algorithm was evaluated on two publicly available datasets of motor imagery EEG signals, where the first dataset was Dataset IIa from BCI Competition IV [30] and the second dataset was Dataset 2 from BNCI Horizon 2020 [31]. These two datasets have multiple subjects and are suitable for multi-source classification.

1) Dataset IIa of BCI Competition IV. It comprises 22-channel EEG signals obtained from two different sessions of nine healthy subjects, and the sampling rate was 250 Hz. Each subject was instructed to perform four motor imagery tasks, including movements of the left hand, right hand, feet and tongue. Each task had 72 trials in one session.

2) Dataset 2 (BNCI Horizon 2020 Dataset). It consists of EEG data from 14 healthy subjects, eight of whom were children. The data for each subject consists of two categories of motor imagery EEG of the subject's right hand and foot, with 50 samples in each category. Each experimental signal was recorded by using 15 electrodes, with electrode positions following the International 10–20 system; 15 channels of EEG signals were recorded and sampled at 512 Hz.

#### 3.2. Data preprocessing

The preprocessing part is introduced first. EEG signals from 2.5 to 5.5 s were selected for BCI Competition IV Dataset IIa, and, for Dataset 2 from BNCI Horizon 2020, the EEG signals ranged from 5.5 to 8.5 s. Bandpass filtering was performed by using a 5th-order Butterworth filter with the frequency ranging from 8 to 30 Hz. The samples in the target domain were randomly arranged 20 times and the online experiment was repeated 20 times; finally, the average results of

the 20 repetitions were recorded.

### 3.3. Comparison methods

To verify the effectiveness, the proposed MSOTL-SDS algorithm is compared against six state-of-the-art algorithms for EEG classification, grouped as follows:

i) OECIT-I [29]: After aligning the sample domains of the target domain using OEA, the classification was performed with an OTL algorithm.

ii) OECIT-II [29]: After aligning the sample domain of the target domain using OEA, the classification was performed with an OTL algorithm; a different weight update strategy was utilized instead of OECIT-I.

iii) HomOTLMS [17]: Migrate multiple-source domain knowledge in the OTL process by constructing the final classifier in a band-weighted classifier integration approach.

iv) HomOTL-ODDM [28]: A multi-source online transfer algorithm for the simultaneous reduction of marginal distribution and conditional distribution differences between domains via a linear feature transformation process.

v) EA-CSP-LDA [32]: EA was used to align the data from different domains; then, the source domain data were used to design the filters and an LDA classifier was employed for classification.

vi) CA-Joint distribution adaptation (JDA) [33]: The data were aligned before the JDA algorithm.

vii) Riemannian alignment-minimum distance to the Riemannian mean (RA-MDMR) [34]: It is a Riemannian space method that centers the covariance matrix relative to the reference covariance matrix.

## 4. Results and discussion

A comparison of the online classification results for the two datasets are given in Tables 1 and 2. The MSOTL-SDS algorithm achieved the highest average accuracies of 79.29 and 70.86% for the two datasets, respectively.

For BCI Competition IV Dataset Iia, we mainly selected 22-channel EEG data from the left and right hands. The classification accuracy of the MSOTL-SDS algorithm was higher than those of the other seven algorithms for more than half of the subjects. The HomOTLMS and HomOTL-ODDM algorithms had slightly lower accuracies, and HomOTL-ODDM was better than HomOTLMS. More importantly, except for Subject 3, the classification accuracies of the multi-source online transfer algorithms were higher than that of the single-source online transfer algorithm OECIT, indicating that the multi-source online transfer algorithm can effectively eliminate the effect of negative transfer when dealing with multiple source domains. The classification accuracy of the MSOTL-SDS algorithm was 2.4% higher than that of the best multi-source online transfer algorithm, HomOTL-ODDM, and 3.55% higher than the single-source online learning algorithm OECIT-II.

For Dataset 2, the MSOTL-SDS algorithm achieved the best classification accuracy on EEG data for subjects other than Subjects 4, 8, 10 and 11, while HomOTL-ODDM performed best on these four subjects, indicating that, for multi-source online classification, it is necessary to consider the conditional and marginal distributions of the samples. Similarly, the accuracies of the MSOTL-SDS algorithm were 1.85 and 3.99% higher than HomOTL-ODDM and OECIT-II, respectively.

Datasets 1 and 2 are both multi-source categorical datasets taken from healthy subjects, and they both have data on multiple subjects. The difference is that Dataset Iia is from the earlier BCI Competition IV 2008, and it has a lower sampling rate and a higher number of EEG channels; it also



contains more EEG feature information. Dataset 2 is from BNCI Horizon 2020, and it has a higher sampling rate but reduced number of channels; it is also more difficult, so all eight algorithms, including the MSOTL-SDS algorithm, achieved better performance on Dataset IIa.

**Table 1.** Comparison of online classification accuracy on BCI Competition IV Dataset IIa.

Subject	1	2	3	4	5	6	7	8	9	Avg
RA-MDMR	72.22	56.94	84.03	65.97	60.42	67.36	61.81	86.81	<b>82.64</b>	70.91
CA-JDA	66.75	47.45	65.52	59.34	54.55	54.27	54.27	73.01	65.18	60.04
EA-CSP-JDA	86.08	56.84	97.74	72.26	51.56	65.56	68.51	89.10	72.43	73.34
OECIT-I	88.91	56.31	97.56	73.57	58.44	67.01	71.04	92.88	75.87	75.73
OECIT-II	89.51	55.49	<b>97.88</b>	74.06	58.03	66.64	71.09	92.98	75.94	75.74
HomOTLMS	89.25	56.36	97.84	74.22	58.90	68.17	73.92	93.69	76.51	76.54
HomOTL-ODDM	<b>90.91</b>	57.31	97.54	75.57	58.84	68.01	74.04	<b>93.88</b>	75.87	76.89
<b>MSOTL-SDS</b>	90.88	<b>60.93</b>	97.80	<b>76.78</b>	<b>60.88</b>	<b>73.64</b>	<b>78.09</b>	94.69	79.94	<b>79.29</b>

**Table 2.** Comparison of online classification accuracy on Dataset 2.

Subject	OECIT-I	OECIT-II	HomOTLMS	HomOTL-ODDM	<b>MSOTL-SDS</b>
<b>1</b>	54.96	55.48	56.28	59.29	<b>62.31</b>
<b>2</b>	63.31	64.03	67.43	68.49	<b>71.77</b>
<b>3</b>	62.74	63.21	64.33	66.53	<b>72.24</b>
<b>4</b>	76.35	76.36	76.98	<b>78.43</b>	77.86
<b>5</b>	55.88	55.83	56.46	57.98	<b>60.61</b>
<b>6</b>	47.29	48.68	47.36	48.26	<b>49.67</b>
<b>7</b>	58.46	58.59	58.33	59.44	<b>60.79</b>
<b>8</b>	84.79	83.57	84.64	<b>85.68</b>	84.96
<b>9</b>	81.73	82.45	82.92	83.79	<b>85.93</b>
<b>10</b>	74.63	76.44	77.38	<b>78.29</b>	78.16
<b>11</b>	70.96	71.08	72.68	<b>73.46</b>	72.76
<b>12</b>	57.46	58.44	58.69	59.78	<b>62.21</b>
<b>13</b>	74.33	73.29	74.33	75.33	<b>78.38</b>
<b>14</b>	68.26	68.78	70.26	71.33	<b>74.35</b>
<b>Avg</b>	66.51	66.87	67.72	69.01	<b>70.86</b>

In Tables 1 and 2, it is noticeable that the MSOTL-SDS algorithm showed significant improvement in terms of accuracy for some subjects with poorly differentiated class features, such as Subjects 2, 6 and 7 in BCI Competition IV Dataset IIa and Subjects 1, 2, 3, 5, 9, 12 and 14 in Dataset 2 from BNCI Horizon 2020; the single-source OECIT and HomOTL-ODDM for each subject achieved average improvements of 5.85 and 3.69%, respectively. This indicates that MSOTL-SDS can effectively discover subjects with high similarity to the target subjects' characteristics by applying SDS when dealing with a multi-source online classification task, thus reducing individual differences and improving classification accuracy. Finally, as shown in Table 3, we compare the average online classification accuracy of some different transfer learning methods on the two datasets [35].

**Table 3.** Average online classification accuracies of different transfer learning approaches on Datasets IIa and 2.

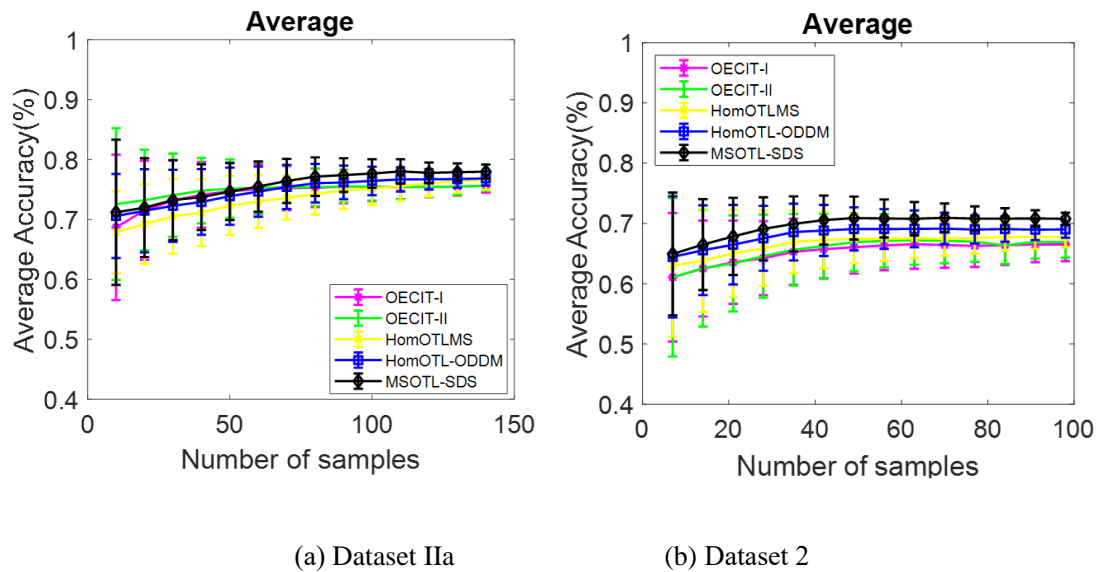
Algorithm	Dataset IIa Acc. (%)	Dataset 2 Acc. (%)
EA-RCSP-LDA [35]	73.72	64.09
EA-RCSP-OwAR [35]	74.78	65.71
OECIT-I [29]	75.73	66.51
OECIT-II [29]	75.74	66.87
HomOTLMS [33]	76.54	67.72
HomOTL-ODDM [34]	76.89	69.01
<b>MSOTL-SDS</b>	<b>79.29</b>	<b>70.86</b>

Table 4 gives the computing time consumption of the five algorithms on different datasets. Dataset 2 is less time-consuming because it has fewer EEG signal channels than Dataset IIa. OECIT handles the online single-source transfer task and thus takes the least amount of time. HomOTLMS and HomOTL-ODDM cost more time due to the increased complexity of the algorithm. MSOTL-SDS uses a multi-source online transfer framework, requiring less source domain data to train the classifier; so, it has an advantage in terms of time consumption. The reduction of computing time was 4.96 and 2.34 s on BCI Competition IV Dataset IIa and Dataset 2 from BNCI Horizon 2020, respectively, relative to HomOTLMS. Additionally, compared with HomOTL-ODDM, MSOTL-SDS was 36.56 and 31.28 s faster on the two datasets, respectively.

**Table 4.** Comparison of time cost using different algorithms.

	Dataset IIa		Dataset 2	
	Mean (s)	Std (s)	Mean (s)	Std (s)
EA-CSP-LDA	9.9918	0.9250	-	-
RA-MDRM	172.7034	27.7032	-	-
CA-JDA	68.0582	1.2880	-	-
OECIT-I	7.9428	0.8685	4.6735	0.7435
OECIT-II	7.8275	0.6577	4.7823	0.6421
HomOTLMS	40.6568	1.9549	20.8283	1.8543
HomOTL-ODDM	72.2592	4.9532	49.7724	4.7286
MSOTL-SDS	35.6982	1.9368	18.4931	1.9254

Figure 2(a),(b) show the average online classification accuracy curves with the number of samples for Datasets IIa and 2, respectively. The poor performance of the OECIT algorithm on a multi-source online classification task is mainly because it views multiple source subjects as a single individual during its training, which can cause negative transfer problems since there exist large individual differences among source subjects. For both datasets, MSOTL-SDS demonstrated the best performance in more than half of the subject accuracy variation plots. The HomOTLMS and HomOTL-ODDM algorithms present similar trends in most of the subject accuracy variation curves, which can be explained by their similar multi-source OTL framework.



**Figure 2.** Variation curves of mean online classification accuracy according to the number of samples for the two datasets.

**Table 5.** P-values of different algorithms according to paired t-test.

	MSOTL-SDS	
	Dataset IIa	Dataset 2
EA-CSP-LDA	0.0004	-
RA-MDRM	0.0081	-
CA-JDA	0.0001	-
OECIT-I	0.0010	0.0003
OECIT-II	0.0005	0.0004
HomOTLMS	0.0034	0.0021
HomOTL-ODDM	0.0018	0.0627

In order to compare the differences between the proposed MSOTL-SDS algorithm and the other four algorithms, a paired t-test was applied to the classification accuracy of the figure, and the significance level was set as  $\alpha = 0.05$ . The test results are shown in Table 5. The results show that MSOTL-SDS was significantly better than other algorithms on Dataset IIa, but there was no significant difference between MSOTL-SDS and HomOTL-ODDM on Dataset 2. This indicates that the multi-source online transfer algorithm is effective in eliminating the effect of negative transfer when dealing with multiple source domains. Moreover, for multi-source online classification, reducing the differences in conditional and edge distributions of samples is an effective method.

## 5. Conclusions

This methodology explores the negative transfer problem in online single-source transfer and tries to reduce the individual differences between subjects. We proposed a multi-source OTL method based on SDS, which was applied for the online classification study of motor imagery EEG signals. By utilizing some of the target domain sample labels in advance, the SDS method can select the source domain data that are similar to the target domain data while eliminating the source domain data that

are different from the target domain data. Then, by combining the weighting coefficients with the online classifier, each selected source domain is trained into a separate classifier, which is employed to predict sample labels. The proposed method was validated on two motor imagery EEG datasets and compared with other online transfer methods. The experimental results show that our MSOTL-SDS algorithm achieved the best performance, the best accuracy and the fastest calculation speed in the online scenarios. However, there are still some limitations of this study. The difference in conditional distribution between the source domain and the target domain was not fully considered. Hence, in future work, we can utilize some advanced transfer learning algorithms, such as balanced distribution adaptation [37] and manifold embedded distribution alignment [37], to align the conditional distribution between domains and improve the proposed method.

## Acknowledgments

This work was supported by Zhejiang Provincial Natural Science Foundation of China (Grant No. LZ22F010003) and the National Natural Science Foundation of China (Grant Nos. 61871427 and 62071161). The authors also gratefully acknowledge the research funding supported by the National Students' Innovation and Entrepreneurship Training Program (No. 202210336034).

## Conflict of interest

The authors declare that there is no conflict of interest.

## References

1. M. Shanechi, Brain-machine interfaces from motor to mood, *Nat. Neurosci.*, **22** (2019), 1554–1564. <https://doi.org/10.1038/s41593-019-0488-y>
2. B. J. Lance, S. E. Kerick, A. J. Ries, K. S. Oie, K. McDowell, Brain-computer interface technologies in the coming decades, in *Proceedings of the IEEE*, **100** (2012), 1585–1599. <https://doi.org/10.1109/JPROC.2012.2184830>
3. S. Aggarwal, N. Chugh, Review of machine learning techniques for EEG based brain computer interface, *Arch. Comput. Methods Eng.*, **29** (2022), 3001–3020. <https://doi.org/10.1007/s11831-021-09684-6>
4. D. Marshall, D. Coyle, S. Wilson, M. Callaghan, Games, gameplay, and BCI: The state of the art, *IEEE Trans. Comput. Intell. AI Games*, **5** (2013), 82–99, <https://doi.org/10.1109/TCIAIG.2013.2263555>
5. D. Wu, Y. Xu, B. Lu, Transfer learning for EEG-based brain-computer interfaces: A review of progress made since 2016, *IEEE Trans. Cognit. Dev. Syst.*, **14** (2022), 4–19, <https://doi.org/10.1109/TCDS.2020.3007453>
6. Y. Zhang, G. Zhou, J. Jin, Q. Zhao, X. Wang, A. Cichocki, Sparse bayesian classification of EEG for brain-computer interface, *IEEE Trans. Neural Networks Learn. Syst.*, **27** (2016), 2256–2267, <https://doi.org/10.1109/TNNLS.2015.2476656>
7. M. Krell, N. Wilshusen, A. Seeland, S. K. Kim, Classifier transfer with data selection strategies for online support vector machine classification with class imbalance, *J. Neural Eng.*, **14** (2017), 025003. <https://doi.org/10.1088/1741-2552/aa5166>.

8. R. Fu, Y. Tian, T. Bao, Z. Meng, P. Shi, Improvement motor imagery EEG classification based on regularized linear discriminant analysis, *J. Med. Syst.*, **43** (2019), 1–13. <https://doi.org/10.1007/s10916-019-1270-0>.
9. F. Fahimi, S. Dosen, K. Ang, N. Mrachacz-Kersting, C. Guan, Generative adversarial networks-based data augmentation for brain-computer interface, *IEEE Trans. Neural Networks Learn. Syst.*, **32** (2021), 4039–4051, <https://doi.org/10.1109/TNNLS.2020.3016666>.
10. V. Jayaram, M. Alamgir, Y. Altun, B. Scholkopf, M. Grosse-Wentrup, Transfer learning in brain-computer interfaces, *IEEE Comput. Intell. Mag.*, **11** (2016), 20–31. <https://doi.org/10.1109/MCI.2015.2501545>
11. H. He, D. Wu, Transfer learning for brain-computer interfaces: A Euclidean space data alignment approach, *IEEE Trans. Biomed. Eng.*, **67** (2021), 399–410. <https://doi.org/10.1109/TBME.2019.2913914>
12. L. Xu, M. Xu, Y. Ke, X. An, S. Liu, D. Ming, Cross-dataset variability problem in EEG decoding with deep learning, *Front. Hum. Neurosci.*, **14** (2020), 103–113. <https://doi.org/10.3389/fnhum.2020.00103>
13. S. J. Pan, Q. Yang, A survey on transfer learning, *IEEE Trans. Knowl. Data Eng.*, **22** (2010), 1345–1359. <https://doi.org/10.1109/TKDE.2009.191>
14. M. Long, J. Wang, G. Ding, S. J. Pan, P. S. Yu, Adaptation regularization: A general framework for transfer learning, *IEEE Trans. Knowl. Data Eng.*, **26** (2014), 1076–1089. <https://doi.org/10.1109/TKDE.2013.111>
15. X. Zhong, S. Guo, H. Shan, L. Gao, D. Xue, N. Zhao, Feature-based transfer learning based on distribution similarity, *IEEE Access*, **6** (2018), 35550–35557. <https://doi.org/10.1109/ACCESS.2018.2843773>
16. M. Jiang, W. Huang, Z. Huang, G. G. Yen, Integration of global and local metrics for domain adaptation learning via dimensionality reduction, *IEEE Trans. Cybern.*, **47** (2017), 38–51. <https://doi.org/10.1109/TCYB.2015.2502483>
17. Q. Wu, H. Wu, X. Zhou, M. Tan, Y. Xu, Y. Yan, et al., Online transfer learning with multiple homogeneous or heterogeneous sources, *IEEE Trans. Knowl. Data Eng.*, **29** (2017), 1494–1507. <https://doi.org/10.1109/TKDE.2017.2685597>
18. J. Wang, P. Zhao, S. C. H. Hoi, R. Jin, Online feature selection and its applications, *IEEE Trans. Knowl. Data Eng.*, **26** (2013), 698–710. <https://doi.org/10.1109/TKDE.2013.32>
19. P. Zhao, C. Steven, OTL: A framework of online transfer learning, in *Proceedings of the 27th International Conference on Machine Learning*, Haifa, Israel, (2010), 1231–1238.
20. P. Zhao, S. Hoi, J. Wang, B. Li, Online transfer learning, *Artif. Intell.*, **216** (2014), 76–102. <https://doi.org/10.1016/j.artint.2014.06.003>
21. Z. Kang, B. Yang, Z. Li, P. Wang, OTLAMC: An online transfer learning algorithm for multi-class classification, *Knowl.-Based Syst.*, **176** (2019), 133–146. <https://doi.org/10.1016/j.knosys.2019.03.024>
22. L. Ge, J. Gao, A. Zhang, Oms-tl: A framework of online multiple source transfer learning, in *Proceedings of ACM International Conference on Information & Knowledge Management*, ACM, (2013), 2423–2428. <https://doi.org/10.1145/2505515.2505603>.
23. H. Zhou, K. Wang, J. Tian, Online transfer learning for differential diagnosis of benign and malignant thyroid nodules with ultrasound images, *IEEE Trans. Biomed. Eng.*, **67** (2020), 27732780. <https://doi.org/10.1109/TBME.2020.2971065>

24. E. Eaton, M. DesJardins, Selective transfer between learning tasks using task-based boosting, in *Proceedings of the 25th AAAI Conference on Artificial Intelligence*, (2011), 337–342.
25. Y. Yao, G. Doretto, Boosting for transfer learning with multiple sources, in *Proceedings of IEEE Computer Vision and Pattern Recognition*, (2010), 1855–1862. <https://doi.org/10.1109/CVPR.2010.5539857>
26. B. Tan, E. Zhong, E. Xiang, Q. Yang, Multi-transfer: Transfer learning with multiple views and multiple sources, *Stat. Anal. Data Min.*, **7** (2014), 282–293. <https://doi.org/10.1002/sam.11226>
27. Y. Jiang, F. Chung, H. Ishibuchi, Z. Deng, S. Wang, Multitask TSK fuzzy system modeling by mining intertask common hidden structure, *IEEE Trans. Cybernet.*, **45** (2015), 534–547. <https://doi.org/10.1109/TCYB.2014.2330844>
28. Y. Du, Z. Tan, Q. Chen, Y. Zhang, C. Wang, Homogeneous online transfer learning with online distribution discrepancy minimization, in *Proceedings of the 24th European Conference on Artificial Intelligence*, (2020), 1–9. <https://doi.org/10.48550/arXiv.1912.13226>
29. Z. Li, Q. She, Y. Ma, J. Zhang, M. Sun, Online EEG classification method based on instance transfer, *Chin. J. Sens. Actuators*, **35** (2022), 1109–1116. <https://doi.org/10.3969/j.issn.1004-1699.2022.08.015>
30. C. Brunner, R. Leeb, G. R. Müller-Putz, A. Schlögl, G. Pfurtscheller, *BCI Competition 2008-Graz Data Set A*, Institute for Knowledge Discovery (Laboratory of Brain-Computer Interfaces), Graz University of Technology, **16** (2008), 1–6.
31. BNCI Horizon 2020, Data sets-BNCI Horizon 2020, <http://bnci-horizon-2020.eu/database/data-sets>.
32. H. He, D. Wu, Transfer learning for brain-computer interfaces: A Euclidean space data alignment approach, *IEEE Trans. Biomed. Eng.*, **67** (2020), 399–410. <https://doi.org/10.1109/TBME.2019.2913914>
33. W. Zhang, D. Wu, Manifold embedded knowledge transfer for brain-computer interfaces, *IEEE Trans. Neural Syst. Rehabil. Eng.*, **28** (2020), 1117–1127. <https://doi.org/10.1109/TNSRE.2020.2985996>
34. P. Zanini, M. Congedo, C. Jutten, S. Said, Y. Berthoumieu, Transfer learning: A Riemannian geometry framework with applications to brain-computer interfaces, *IEEE Trans. Biomed. Eng.*, **65** (2018), 1107–1116. <https://doi.org/10.1109/TBME.2017.2742541>
35. D. Wu, X. Jiang, R. Peng, Transfer learning for motor imagery based brain-computer interfaces: A tutorial, *Neural Networks*, **153** (2022), 235–253. <https://doi.org/10.1016/j.neunet.2022.06.008>
36. J. Wang, Y. Chen, S. Hao, W. Feng, Z. Shen, Balanced distribution adaptation for transfer learning, in *Proceedings of IEEE International Conference on Data Mining (ICDM)*, (2017), 1129–1134. <https://doi.org/10.1109/ICDM.2017.150>
37. J. Wang, W. Feng, Y. Chen, H. Yu, M. Huang, P. S. Yu, Visual domain adaptation with manifold embedded distribution alignment, in *Proceedings of the 26th ACM International Conference on Multimedia*, (2018), 402–410. <https://doi.org/10.1145/3240508.3240512>



AIMS Press

©2023 the Author(s), licensee AIMS Press. This is an open access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/4.0>)