Article

Multi-Class Transfer Learning and Domain Selection for Cross-Subject EEG Classification

Rito Clifford Maswanganyi [1], Chungling Tu [1,*], Pius Adewale Owolawi [1] and Shengzhi Du [2,*]

[1] Department of Computer Systems Engineering, Tshwane University of Technology,
Pretoria 0002, South Africa
[2] Department of Electrical Engineering, Tshwane University of Technology, Pretoria 0002, South Africa
* Correspondence: duc@tut.ac.za (C.T.); dushengzhi@gmail.com (S.D.); Tel.: +27-012-382-9943 (C.T.)

Abstract: Transfer learning (TL) has been proven to be one of the most significant techniques for cross-subject classification in electroencephalogram (EEG)-based brain-computer interfaces (BCI). Hence, it is widely used to address the challenges of cross-session and cross-subject variability with more accurate intention prediction. In this case, TL utilizes knowledge (signal features) in the source domain(s) to improve the classification in the target domain. However, current existing transfer learning approaches on EEG-based BCI are mostly limited to two-class cross-subject classification problems, while multi-class problems are only implemented with a focus on within-subject classification due to the complexity of multi-class cross-subject classification problems. In this paper, we first extended the transfer learning approaches to a multi-class cross-subject scenario, then investigated the reason for transfer learning performance being poor in multi-class cross-subject classification. Secondly, we address the challenge of significant sessional and subject-to-subject variations originating from both known and unknown factors. It is discovered that such variations have a massive influence on the classification because of the negative transfer (NT) across domains. Based on this discovery, we propose a multi-class transfer learning approach based on multi-source manifold feature transfer learning (MMFT) framework and an enhanced version to minimize the effects of NT. The proposed multi-class transfer learning approach extends the existing MMFT to multi-class cases. Then enhanced multi-class MMFT firstly searches for domains with high transferability and selects only the best combination among source domains (SD), then utilize the best-selected combination of domains for transfer learning. Experimental results illustrate that the proposed multi-class MMFT can be employed in the cross-subject classification of both three-class and four-class problems. Experimental results also demonstrated that the enhanced multi-class MMFT could effectively minimize the effect of negative transfer and significantly increase the prediction rates across individual target domains (TD). The highest classification accuracy (CA) of 98% is obtained by the enhanced multi-class MMFT.

Keywords: brain-computer interface (BCI); electroencephalogram (EEG); transfer learning (TL); multi-source manifold feature transfer learning (MMFT); multi-class transfer learning; cross-session classification; cross-subject classification

1. Introduction

Brain-computer interface (BCI) aims to introduce a direct communication pathway between the human brain and external devices without any muscular stimulation. Hence, neural information can be interpreted as continuous EEG signals extracted from the brain and translated into control commands used to enable users with neurological disorders to interact with machines [1]. Consequently, EEG-based BCI has been widely utilized to translate EEG signals into control commands using signal pre-processing, feature extraction, and classification algorithms [2]. In recent years, enhancing classification performance has been one of the primary focuses of BCI. Therefore, numerous techniques, including experimental paradigms, have been developed to increase the signal-to-noise ratio of EEG signals, mainly

to improve classification performance. However, sessional and inter-subject variations in EEG dynamics, as a result of both known and unknown factors were demonstrated to pose a significant challenge to classification performance [3,4]. Such variations in EEG dynamics contribute to varying neural responses from the same stimulus across different subjects. Moreover, the same stimulus can result in varying neural responses from the same subject across different time frames, in turn resulting in poor classification performance [5,6]. In recent studies on BCI, it has been proven that Transfer learning and domain selection can effectively address the challenge of sessional and inter-subject variations [7]. Therefore, utilizing features from source domains to predict features in the target domain can effectively improve classification performance [8,9].

Zhang et al. [10] proposed a transfer learning framework based on a two-class problem to address the challenge of inter-subject and sessional variations in EEG dynamics. Manifold embedded knowledge transfer (MEKT) framework was proposed and achieved a highest mean classification accuracy of 79.86%. However, individual TDs demonstrated to be significantly affected by NT, which in turn affected the prediction rate across target domains. To address the challenge of NT across individual TDs, the domain transferability estimation (DTE) approach was further proposed for domain selection, with a highest accuracy of 82.14% observed after source selection. Li et al. [11] employed a forward floating-point search algorithm to address the challenge of negative transfer by selecting sources with high transferability for transfer mapping. The algorithm sequentially removed a single domain (subject) from the current subset of source domains (subjects) at each loop as long as the resulting subset is better than the previously evaluated one at that level.

Transfer learning has proven to yield a superior classification performance before source selection across individual TDs for a two-class cross-subject EEG classification. In a similar manner, a combination of both non-related and related sources can significantly deteriorate CA across different individual TDs. Hence, employing domain selection to address the challenge of NT by iterating through source domains to identify the most beneficial source domains and removing a single less-related domain at a time can still pose a negative impact on CA throughout each iteration, mainly because individual sources respond differently to different combinations of source domains. Moreover, the performance of TL on EEG-based BCI has not yet been evaluated and validated based on multi-class cross-subject EEG classification.

Kim et al. [12] further demonstrated the significance of TL on prediction rate for multi-class EEG classification through the implementation of a compact convolutional neural network (CNN) architecture for EEG-based BCI. However, the proposed deep learning model (EEGNet) was implemented based on a four-class problem but with a focus on within-subject classification, mainly to address the challenge of sessional variations [13]. Subsequently, a highest prediction rate of 91.34% was achieved for multi-class within-subject classification. Shahabi et al. [14] further illustrated the impact of TL on the prediction rate for a two-class cross-subject classification problem by evaluating the performance of five pre-trained CNN architectures (VGG16, Xception, DenseNet121, MobileNetV2, and InceptionResNetV2). The CNN models were employed to predict two classes for responder and non-responder, to Selective Serotonin Reuptake Inhibitors (SSRI) antidepressants in patients with Major Depressive Disorder (MDD). A highest CA of 95.74% was achieved when DenseNet121 was employed for a two-class cross-subject classification problem. Zhang et al. [15] proposed a deep CNN model based on a two-class problem, mainly to address the challenge of sessional and inter-subject variations. The network was first trained and evaluated on the same subject's data, then after the model was trained on a set of subjects and evaluated on a new target subject [13,16]. Subsequently, samples from the same session were used for training and validation for subject-specific classification, while for subject-independent classification, data from all subjects except the target subject were used to train, and the target subject was used to test the classifier. Subsequently, deep CNN obtained a highest classification accuracy of 84.19% for a two-class cross-subject classification or subject-independent.

Transfer learning is by far the best solution for cross-subject classification problems emerging as a result of significant variations in EEG dynamics. However, current existing studies have demonstrated that TL has been limited to or implemented based on a two-class problem. Furthermore, multi-class problems are only implemented with a focus on within-subject classification due to the complexity of multi-class cross-subject classification problems. Moreover, transfer learning suffers from the effect of NT. Therefore, domain selection has been introduced to address the challenge of NT originating from non-related domains. However, for more complex multi-class classification problems, removing a single non-related domain at a time when domain selection is employed by iterating through source domains to identify the most beneficial source domains can create a severe implication on classification performance, mainly because a single source domain can respond differently to different combinations of sources.

In this study, we first propose a multi-class MMFT framework to investigate the impact of three-class and four-class cross-subject classification problems on transfer learning performance. Secondly, we propose an enhanced multi-class MMFT framework to address the challenge of negative transfer originating from non-related sources, whereby enhanced multi-class MMFT firstly searches and selects only the optimal combination of source domains, then utilizes the best selected closely related sources to perform transfer learning. The main contributions of this study are as follows:

1. A multi-class MMFT approach is developed to investigate the impact of multi-class cross-session and cross-subject classification problems on transfer learning performance. The proposed multi-class MMFT enhanced performance of individual target domains for both three-class and four-class cross-session and cross-subject classification problems.
2. A comparative performance analysis between multi-class MMFT, manifold embedded knowledge transfer (MEKT), and a traditional BCI pipeline based on two classical machine learning algorithms (Linear discrimination analysis (LDA) and Regression tree (RegTree)) is carried out. The proposed approach outperforms all classification algorithms for three-class and four-class cross-session and cross-subject classification problems.
3. An enhanced multi-class MMFT framework is proposed for domain selection, mainly to minimize the impact of negative transfer in multi-source transfer mapping. The proposed enhanced multi-class MMFT improves classification performance by selecting only the optimal combination of closely related sources among source domains, then performing transfer learning on the optimal source domain combination, where the challenge of negative transfer is solved.
4. A comparative performance analysis shows that the source selection significantly improves the performance of the proposed enhanced multi-class MMFT for both three-class and four-class, cross-session, and cross-subject classification problems.

The rest of the paper is organized as follows. Section 2 describes details of the proposed multi-class transfer learning and its enhanced version. Section 3 illustrates the results of three-class and four-class experiments when cross-session and cross-subject classification problems are considered. Section 4 provides some insight and discussion on relevant issues. Finally, some conclusions are provided in Section 5.

2. Materials and Methods

2.1. Datasets

In this study, two EEG datasets consisting of MI and steady-state motion visual evoked potential (SSMVEP) samples are utilized. The MI dataset was acquired from a public database, while the SSMVEP dataset is our own recorded dataset. In this case, both datasets are utilized to validate the proposed method. These datasets are described in the following sections.

2.1.1. Dataset I (Our SSMVEP Dataset)

Our own SSMVEP dataset was recorded in a lab from five healthy experimental participants, whereby a g.tec EEG recording system consisting of sixteen EEG channels was utilized to acquire raw EEG signals from the brain [17]. Moreover, a 10–20 electrodes positioning system was utilized to arrange all EEG channels on the surface of the scalp, whereby raw EEG signals were recorded at a sampling rate of 250 Hz. When the experiment begins, a computer monitor is utilized to project a visual cue, with each participant seated facing the projecting screen. As such, four flickering objects were projected on a computer monitor, with each object flickering at a different frequency equivalent to four EEG classes [18]. In this case, all four objects were flickering in four different directions (left, right, up, and down). Each of the four objects took a turn to be projected on a monitor, while a beeping sound served as a notification to inform each subject to focus their attention on a projected object flickering at either of the four frequencies (29 Hz, 13.3 Hz, 17 Hz, and 21 Hz) [19,20]. Furthermore, each SSMVEP task was executed for 300 s equivalent to 75,000 samples, as depicted in Figure 1.
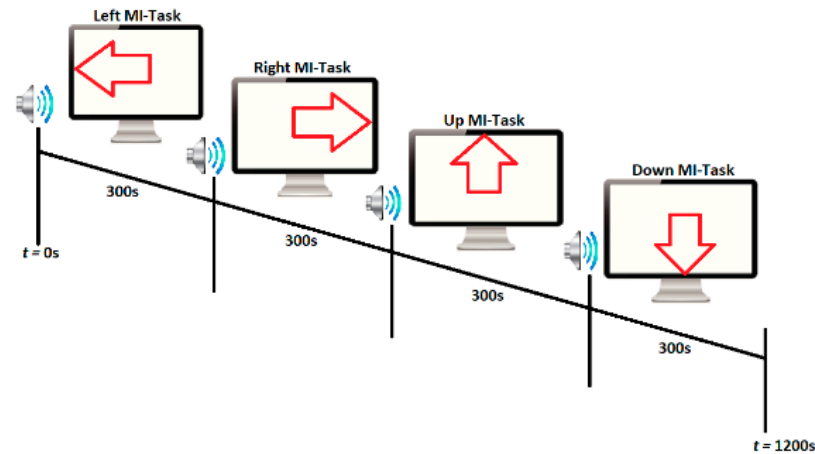

Figure 1. Our dataset experimental paradigm timing scheme.

2.1.2. Dataset II (BCI Competition IV-a Dataset)

This study also makes use of a publicly available database known as BCI competition IV-a to facilitate the investigation. The dataset consists of four MI classes (left, right, both feet and tongue) acquired from nine experiment participants or subjects [21]. In this case, twenty-two Ag/AgCl EEG channels were utilized to acquire raw EEG signals from the brain [22]. Moreover, a 10–20 electrodes positioning system was utilized to arrange EEG channels on the surface of the scalp [23]. Subsequently, when the experiment began, a fixation cross "+" was projected on a computer monitor to indicate the beginning of a trial denoted by $t = 0$ s, while participants were seated facing a computer monitor. An arrow pointing to four different directions was utilized as visual cues and projected on a computer monitor for $t = 1.25$ s, while subjects were requested to execute motor imagery tasks from $t = 3.25$ s to $6 = 6$ s, as depicted in Figure 2.

2.2. Data Pre-Processing

BCI competition IV dataset was firstly filtered using a 0.5 Hz to 100 Hz band-pass filter to eradicate the effect of non-physiological artifacts in the form of noise, and then a 50 Hz notch filter was applied to reduce the effect of line noise [23]. Moreover, MI signals were sampled at a frequency of 250 Hz. Our own SSMVEP dataset was recorded at a sampling frequency of 250 Hz. A 0.5 HZ to 60 Hz band-pass filter was then applied to reduce the effect noise. A notch filter at a cut-off frequency of 50 Hz was applied to reduce the effect of line noise. Moreover, a common average reference (CAR) was applied to eliminate the effect of noise originating from electrodes [17].
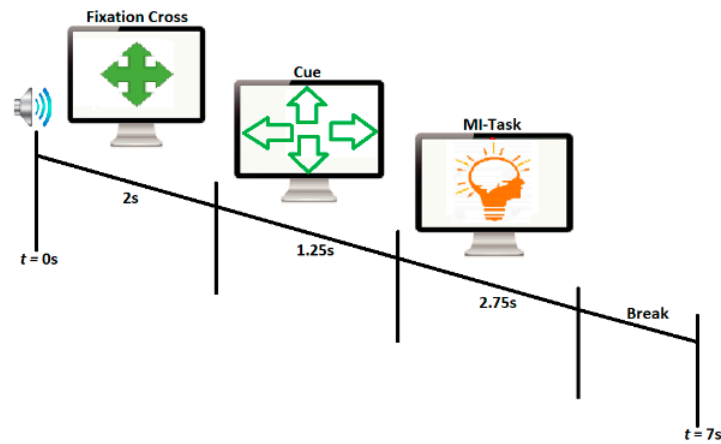
Figure 2. BCI Competition IV-a experimental paradigm timing scheme.

2.3. Methods

In this section, we present a novel framework of multi-class MMFT, which adapt the idea of a two-class cross-subject classification for MMFT in multi-class classification problems, using the three-class and four-class classification as examples. We first review the related MMFT framework for two-class cross-subject classification, then introduce the proposed multi-class MMFT for three-class and four-class cross-subject classification, which can be extended to even more classes. The proposed multi-class MMFT framework is further enhanced by applying source selection to identify the optimal combinations of domains, then utilizing the selected sources to perform transfer mapping.

2.3.1. Related Work

In transfer learning, features acquired in the source domain can be utilized to enhance the intention detection rate (IDR) in the target domain. Hence, it is highly significant to minimize variations between data distributions of the source and target domain [10]. The manifold feature transfer learning (MMFT) framework [24] is an effective method based on transfer learning. MMFT first performs distribution mean alignment (DMA), whereby all domains are similarly disseminated across the symmetric positive defined (SPD) manifold, using the objective function shown in (1).

$$J(A, B) = \min_{A,B} \left(\frac{\sigma_s^2 + \left(A^T M_s A - B^T M_t B \right)^2}{2\sigma_t^2} + \frac{\sigma_s^2 + \left(B^T M_t B - A^T M_s A \right)^2}{2\sigma_s^2} \right) \tag{1}$$

where A and B are linear transformations; M_s and M_t are the distribution means of source and target domains, respectively; σ_s^2 and σ_t^2 are the deviations of covariance matrices of source and target domains, respectively. The optimal solution of $A^T M_s A = B^T M_t B$ and two feasible solutions, as shown in (2) and (3), are obtained by minimizing the objective function in (1). In this way, the marginal probability distribution shift is minimized across source and target domains, i.e., aligning the distribution means.

$$A = M_s^{-1/2}, B = M_t^{-1/2} \tag{2}$$

A and B in (2) can be further used to align the distribution means to every domain's Riemannian center.

$$A = M_t^{1/2} M_s^{-1/2}, B = I \tag{3}$$

where I is the identity matrix with proper dimension. Using (3), the distribution mean of the target domain, M_t, is aligned with the distribution mean of the source domain M_s.

After domain distribution means are aligned, covariance matrices in both source and target domains are computed using (4) and (5).

$$\widetilde{P}_{s,i} = M_s^{-1/2} P_{s,i} M_s^{-1/2} \tag{4}$$

$$\widetilde{P}_{t,j} = M_s^{-1/2} P_{t,j} M_s^{-1/2} \tag{5}$$

where $P_{s,i}$ is the covariance matrix of the i-th trial EEG data in a source domain; $P_{t,j}$ is for the j-th trial EEG data in the target domain; $\widetilde{P}_{s,i}$ and $\widetilde{P}_{t,j}$ are the aligned covariance matrices of source and target domains, respectively.

The next phase of the MMFT framework is SPD manifold tangent space feature extraction using (6) and (7)

$$x_{s,i} = upper\left(\log\left(\widetilde{P}_{s,i}\right)\right) i = 1, \ldots, n_s \tag{6}$$

$$x_{t,j} = upper\left(\log\left(\widetilde{P}_{t,j}\right)\right) j = 1, \ldots, n_t \tag{7}$$

where n_s and n_t are the number of trials in a source domain and target domain, respectively; $x_{s,i}$ and $x_{t,j}$ represent tangent space features extracted from a source domain and target domain, respectively.

Equations (8) and (9) represent extracted feature matrices, whereby both source and target domain tangent space features are denoted by X_s and X_t, respectively [10,24].

$$X_s = [x_{s,1} \ldots x_{s,n_s}] \tag{8}$$

$$X_t = [x_{t,1} \ldots x_{t,n_t}] \tag{9}$$

Furthermore, Grassmann manifold feature learning (GFK) is implemented in the third phase of multi-class MMFT. Subsequently, throughout the conversion process on geodesics from the source to the target domain, GFK searches for average stable features [25]. The marginal probability distribution is reduced through GFK, which brings deviations of all covariance matrices closer [26].

Equation (10) represents either learned source or target domain Grassmann manifold features denoted by z, while tangent space features are denoted by x and geodesic flow is represented by Φ.

$$Z = g(X) = \Phi(t)^T X \tag{10}$$

The inner product of transformed features is obtained via (11).

$$\langle z_i, z_j \rangle = \int_0^1 \left(\Phi(t)^T x_i\right)^T \left(\Phi(t)^T x_j\right) dt = x_i^T G x_j \tag{11}$$

where Φ represents the geodesic flow; x and z are tangent space features and newly learned Grassmann manifold, respectively; $G = \Phi(t)\Phi(t)^T$ is the transformation from x to z. Using the transformation, one receives the Grassmann manifold features in (12).

$$z = g(x) = \sqrt{G} x \tag{12}$$

The MMFT classifier is implemented in the fourth phase of the MMFT framework. The structural risk (13) is minimized across source domains, and conditional alignment is summarized, mainly to train the prediction model [27].

$$\sum_{i=1}^n l(f(z_i), y_i) + \sigma\|f\|_K^2 = \sum_{i=1}^{n_s+n_t} E_{ii}(y_i - f(z_i))^2 + \sigma\|f\|_K^2 \tag{13}$$

where E is the diagonal label indicator matrix.

$$D_f(D_s, D_t) = \sum_{c=1}^{C} D_{f,(c)}(Q_s, Q_t) \tag{14}$$

Equation (14) signifies a conditional distribution discrepancy of a prediction model denoted by f. As such, the conditional distribution alignment is defined by $D_{f,(c)}(Q_s, Q_t)$.

$$f = arg \min_{f \in H_K} l(f(z_i), y_i) + \sigma \|f\|_K^2 + \lambda D_f(D_s, D_t) \tag{15}$$

Equation (15) represents a classifier denoted by f obtained by combining (13) and (14) with both D_s and D_t representing source and target domain samples, respectively [24].

$$f(z) = \sum_{i=1}^{n} \alpha_i K(z_i, z) \tag{16}$$

Equation (16) accepts an expansion when the representer theorem is implemented with coefficients vector denoted by α, while estimation of an original feature vector to Hilbert space through feature mapping generates a kernel denoted by K [24].

$$\sum_{i=1}^{n} l(f(z_i), y_i) + \sigma \|f\|_K^2 = \sum_{i=1}^{n_s n_t} E_{ii}(y_i - f(z_i))^2 + \sigma \|f\|_K^2 = \left\| \left(Y - \alpha^T K \right) E \right\|_F^2 + \sigma tr \left(\alpha^T K \alpha \right) \tag{17}$$

Equation (17) defines a structural risk minimization on the source domain. In this case, α represents the coefficient vector, while the frobenious norm is denoted by $\| \ \|_K^2$, and the kernel matrix is denoted by $K \in {}^{nxn}$. Both source and target domain label matrices are denoted by $Y = [y_1 \dots y_n]$, while the trace operation is represented by $tr()$ [24].

$$D_f(D_s, D_t) = tr \left(\alpha^T K M_c K \alpha \right) \tag{18}$$

To further transform the conditional distribution alignment, the representer theorem is combined with the kernel to produce (18), with the maximum mean discrepancy matrix denoted by M_c [24].

$$(M_c)_{i,j} = \begin{cases} \frac{1}{n_s^2} z_i, z_j \in D_{s,(c)} \\ \frac{1}{n_t^2} z_i, z_j \in D_{t,(c)} \\ -\frac{1}{n_s n_t} \begin{cases} z_i \in D_{s,(c)}, z_j \in D_{t,(c)} \\ z_i \in D_{t,(c)}, z_j \in D_{s,(c)} \end{cases} \\ 0 \ otherwise \end{cases} \tag{19}$$

Equation (19) is used to compute elements of a maximum mean discrepancy, with both source and target domain samples contained in class c denoted by $D_{s,(c)}$ and $D_{t,(c)}$.

$$f = arg \min_{f \in H_K} \left\| \left(Y - \alpha^T K \right) E \right\|_F^2 + \sigma tr \left(\alpha^T K \alpha \right) + \lambda tr \left(\alpha^T K M_c K \alpha \right) \tag{20}$$

Equation (20) representing the objective function denoted by f is obtained by combining (17) and (18) and used to reduce the loss function and conditional probability distribution shift.

$$\alpha = ((E + \lambda M_c)K + \sigma I)^{-1} E Y^T \tag{21}$$

A solution to the objective function was obtained using (21) through the setting of the derivative $\partial f / \partial \alpha = 0$. In this case, (18) is used to calculate the classifier once α is obtained.

$$f = \sum_{i=1}^{z} f_i \tag{22}$$

Quantified voting to transfer multi-source knowledge is implemented as the last phase of the proposed MMFT. In this case, to obtain labels for the target domain, a voting mechanism is utilized to integrate prediction information of various source domains. Subsequently, (22) is used to vote for the prediction results of individual classifiers. In this case, transfer of neural information across domains is dependent on the assessed classifier using a voting mechanism [24]. As such, each classifier is built using (18) whenever each of the source domains is transferred.

$$f = \arg\min_{f \epsilon H_K} \| \left(Y - \alpha^T K \right) W \|_F^2 + \sigma tr \left(\alpha^T K \alpha \right) + \lambda tr \left(\alpha^T K M_c K \alpha \right) \tag{23}$$

Furthermore, the weighted MMFT (w-MMFT) algorithm is also explored to address the challenge of class imbalance, which occur as a result of source domain label information, that tends to affect the structural risk function used to train the MMFT classifier [24,28]. In this instance, w-MMFT's objective function is defined by (23), with W representing the weight matrix.

$$\alpha = ((W + \lambda M_c)K + \sigma I)^{-1} E Y^T \tag{24}$$

Equation (24) defines the solution to the objective function denoted by (23), with M_c representing a maximum mean discrepancy matrix, while W represents the weight matrix [24]. The overall procedure of MMFT is summarized in Algorithm 1.

Algorithm 1 Multi-source Manifold Feature Transfer (MMFT)

Input: z source domains samples $\left\{ \left\{ X_{s,i}^{(j)}, y_{s,i}^{(j)} \right\}_{i=1}^{n_s} \right\}_{j-1}^{z}$; target domain samples $\{ X_{t,i} \}_{i=1}^{n_t}$;

regularization parameters λ, η; number of iterations N.
Output: Predict labels $\widetilde{y}_t$ for target domain
1: Calculate the covariance matrices for both z source domains and target domain, apply DMA,

get aligned covariance matrices $\left\{ \hat{P}_s^{(j)} \right\}_{j=1}^{z}$ and $\widetilde{P}_t$.

2: Get tangent space features $\widetilde{X}$ form SPD manifold.
3: Learn Grassmann manifold features $Z = \left[Z_t, Z_S^1, \dots, Z_S^z \right]$ for z source domains and target domain.
4: Pseudo labels for target domain $\hat{y}_t = []$.
5: **for** $n = 1, \dots, N$ **do**
6:      Multi-source classifier $f = []$.
7:      **for** $j = 1, \dots, z$ **do**
8:          Construct kernel K using features $\left[Z_t, Z_S^{(j)} \right]$.

9:          Calculate M_c using $\hat{y}_S^{(j)}$ and $\hat{y}_t$
10:          Compute α to obtain f_j trained by j-th source domain via the representer theorem.
11:          Quantified voting for target domain, $f = f + f_j$.
12:      **end for**
13:      Pseudo labels $\hat{y}_t = f(Z_t)$, update $\hat{y}_t$.
14: **end for**
15: **return** $\hat{y}_t$.

2.3.2. Multi-Class MMFT Framework

A novel multi-class MMFT is presented in this section to evaluate the effect of multi-class cross-subject classification on transfer learning performance. The proposed framework adapts the idea of a two-class cross-subject classification for MMFT in multi-class classification problems, using the three-class and four-class classification as examples.

Figure 3 gives a detailed overview of the proposed multi-class MMFT framework that performs transfer mapping, utilizing features acquired from multiple source domains to improve the prediction rate of individual target domains based on a three-class and a four-class problem. The proposed multi-class MMFT framework is composed of four modules that form part of transfer learning:

- Distribution means alignment (DMA). Multi-class MMFT firstly performs DMA for multi-class domains through rank of domain (ROD) utilizing KL divergence to evaluate similarities between source and target domain, mainly to align the distribution mean of each domain on (SPD) manifold.
- SPD manifold feature extraction. After domains are aligned, tangent space features are extracted from multi-class SDs and TD, respectively.
- Grassmann manifold feature learning. Once tangent space features have been extracted, the geodesic flow kernel is utilized to learn feature mapping.
- Classification utilizing MMFT classifier. After Grassmann manifold feature learning, the MMFT classifier is used to predict TD labels using knowledge from SDs. The overall procedure of the multi-class MMFT is summarized in Algorithm 2.

Algorithm 2 Multi-class MMFT

Input:z Multi-class source domains samples $\left\{ \left\{ X_{s,i}^{(j)}, y_{s,i}^{(j)} \right\}_{i=1}^{n_s} \right\}_{j-1}^{z}$; Multi-class target domain samples $\{X_{t,i}\}_{i=1}^{n_t}$; regularization parameters λ, η; number of iterations N.

Output: Predict multi-class labels $\widetilde{y}_t$ for target domain

1: Calculate the covariance matrices for both z source domains and target domain, apply DMA, get aligned covariance matrices $\left\{ \hat{P}_s^{(j)} \right\}_{j=1}^{z}$ and $\widetilde{P}_t$.

2: Get tangent space features $\widetilde{X}$ form SPD manifold.

3: Learn Grassmann manifold features $Z = \left[Z_t, Z_S^1, \ldots, Z_S^z \right]$ for z source domains and target domain.

4: Four-class pseudo labels for target domain $\hat{y}_t = []$.

5: **for** $n = 1, \ldots, N$ **do**

6:     Multi-source classifier $f = []$.

7:     **for** $j = 1, \ldots, z$ **do**

8:         Construct kernel K using features $\left[Z_t, Z_S^{(j)} \right]$.

9:         Calculate M_c using multi-class labels $\hat{y}_S^{(j)}$ and $\hat{y}_t$

10:         Compute α to obtain f_j trained by j-th source domain via the represeter theorem.

11:         Quantified voting for target domain, $f = f + f_j$.

12:     **end for**

13:     multi-class pseudo labels $\hat{y}_t = f(Z_t)$, update $\hat{y}_t$.

14: **end for**

15: return multi-class labels $\hat{y}_t$ for target domain.
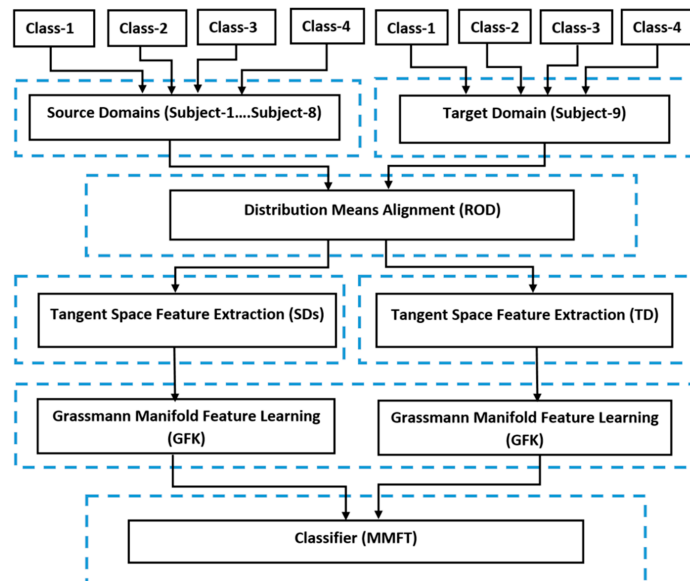

Figure 3. Proposed multi-Class MMFT framework.

2.3.3. Enhanced Multi-Class MMFT

The proposed enhanced multi-class MMFT approach consists of two phases, namely source selection and transfer mapping.

Figure 4 depicts the general framework of the proposed enhanced multi-class MMFT that firstly performs source selection aimed at identifying appropriate sources, then utilizes an optimal combination of sources to perform transfer mapping, mainly to minimize EEG variations across source domains and target domains. The proposed enhanced multi-class MMFT includes five phases:

- Domain selection through selection of optimal combination of closely related domains.
- Distribution means alignment through rank of domain.
- Tangent space feature extraction.
- Grassmann manifold feature learning.
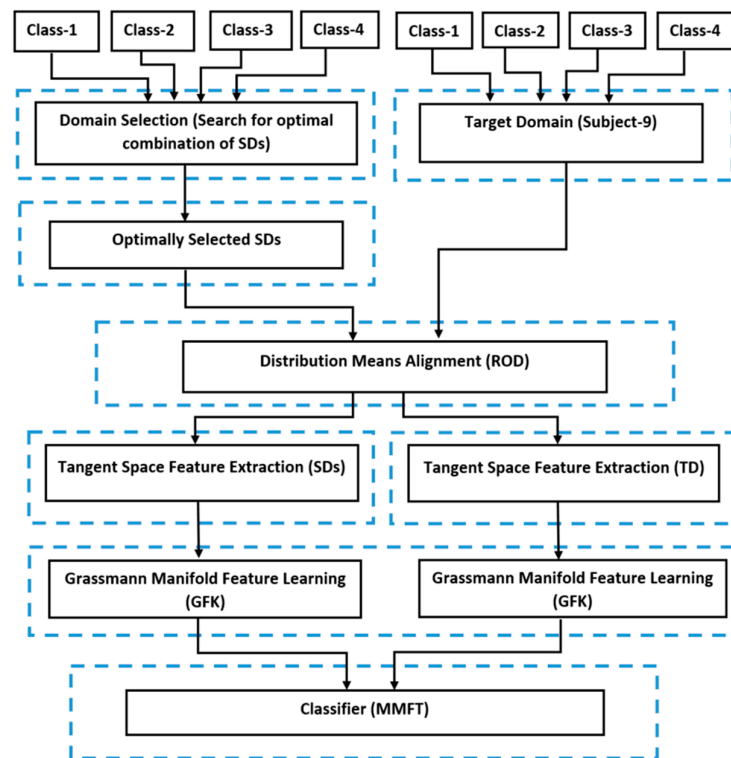- Classification utilizing MMFT classifier.

Figure 4. Proposed enhanced multi-Class MMFT integrating domain selection and transfer learning.

Domain Selection

In this case, source selection phase is firstly carried out to identify and select only closely related combinations of domains from multiple sources, then utilize the best-selected combination during transfer mapping to enhance the prediction rate of individual target domains. To achieve this goal, a MATLAB function known as nchoosek [29], denoted by $B = nchoosek(v, k)$, which is a representation of binomial coefficient, was implemented through our proposed multi-class MMFT, with the number of possible combinations shown in (25)

$$C(n,k) = \frac{n(n-1)\dots(n-k+1)}{k!} = \frac{n}{k}\frac{n-1}{k-1}\frac{n-2}{k-2}\cdots\frac{n-k+1}{1} \tag{25}$$

Subsequently, when the function is executed, it generates all possible combinations of closely related domains among multiple sources. In this instance, (4) is used to com-

pute binomial coefficients denoted by $C(n, k)$, with k representing the number of possible combinations from n. In this case, both k and n have to be a nonnegative integer [29].

Transfer Mapping

The main objective of the source selection phase is to discard non-related domains with low transferability but consider only closely related sources, mainly to reduce the effect of NT emerging from sessional and subject-to-subject variations [30]. After source selection, the optimal combination of domains is then utilized when transfer mapping is applied, whereby the selected sources are firstly aligned through DMA using ROD. Subsequently, covariance matrices are then generated once distribution means have been aligned, from which tangent space features for both source and target domain are extracted. Moreover, feature learning is carried out through GFK, making use of features acquired from the SPD manifold, whereby the MMFT classifier is applied after feature learning. In this case, to acquire the target domain's pseudo labels, prediction rates from multiple classifiers are then integrated through a voting mechanism.

The overall procedure of the proposed multi-class MMFT is summarized in Algorithm 3.

Algorithm 3 Enhanced Multi-class MMFT

Input: z source domains samples $\left\{ \left\{ X_{s,i}^{(j)}, y_{s,i}^{(j)} \right\}_{i=1}^{n_s} \right\}_{j-1}^{z}$; target domain samples $\{X_{t,i}\}_{i=1}^{n_t}$; regularization parameters λ, η; number of iterations N.
Output: Predict labels $\widetilde{y}_t$ for target domain
1: for $n = 1, \ldots, N$ do
2:     Compute binomial coefficients $C(n, k)$; get combination of sources closely related to the target domain.
3:         for $j = 1, \ldots, z$ do
4:             Get only the best combination from closely related sources.
5:         end for
6: end for
7: Calculate the covariance matrices for both z source domains and target domain, apply DMA, get aligned covariance matrices $\left\{ \hat{P}_s^{(j)} \right\}_{j=1}^{z}$ and $\widetilde{P}_t$.
8: Get tangent space features $\widetilde{X}$ form SPD manifold.
9: Learn Grassmann manifold features $Z = [Z_t, Z_S^1, \ldots, Z_S^z]$ for z source domains and target domain.
10: Pseudo labels for target domain $\hat{y}_t = []$.
11: for $n = 1, \ldots, N$ do
12:     Multi-source classifier $f = []$.
13:     for $j = 1, \ldots, z$ do
14:         Construct kernel K using features $[Z_t, Z_S^{(j)}]$.
15:         Calculate M_c using $\hat{y}_S^{(j)}$ and $\hat{y}_t$
16:         Compute α to obtain f_j trained by j-th source domain via the representer theorem.
17:         Quantified voting for target domain, $f = f + f_j$.
18:     end for
19:     Pseudo labels $\hat{y}_t = f(Z_t)$, update $\hat{y}_t$.
20: end for
21: return $\hat{y}_t$.

---

3. Results

3.1. Experiment Setup

In this study, the impact of multi-class cross-session and cross-subject classification on TL performance is investigated using our proposed multi-class MMFT approach. To facilitate the investigation, multi-class MMFT is firstly evaluated based on a three-class problem, then a four-class problem using both nine EEG sessions and subjects, respectively. The proposed multi-class MMFT receives multi-class (z) samples from different sources or domains (subjects/sessions), and class labels (y) for each domain as input parameters.

Multi-class MMFT iterate (N) through domains, and in each iteration, samples from each domain take a turn as a target domain (X_t), while the rest of the samples are source domains (X_s). In each iteration, the classifier utilizes source domain labels (Y_s) to predict labels (Y_t) for the corresponding target domain. The number of iterations in this instance is dependent on the number of sources (subjects/sessions), meaning less number of sources reduces while multiple sources increase computation time for multi-class MMFT. For our experiments, when multi-class MMFT is employed, eight sessions or subjects are assigned as source domains, while a single session or subject is assigned as a target domain [24].

For both three-class and four-class problems, multi-class MMFT is compared with both MEKT and two ML algorithms implemented through a traditional BCI pipeline. The same experiment setup utilized when multi-class MMFT is employed was utilized when MEKT is employed, and eight sessions or subjects are considered as source domains, while a single session or subject is considered as a target domain. To emulate the same TL experiment using LDA and RegTree classifiers [5,31], eight sessions or subjects are utilized for training, while a single session or subject is utilized to test both classifiers, respectively [17,32].

Multi-class MMFT is further enhanced to first perform domain selection and then TL, mainly to address the challenge of NT emanating from significant sessional and subject-to-subject variations [33]. When the enhanced multi-class MMFT is employed, the optimal combination among source domains is firstly selected from eight sessions or subjects, and then the best combination is selected as source domains to enhance the prediction rate in the target domain.

3.2. Three-Class Problems

3.2.1. Multi-Class MMFT for Three-Class Cross-Session Classification

In this section, we investigate the impact of a three-class cross-session classification problem on transfer learning performance. Multi-class MMFT framework is employed and compared with MEKT, and two ML algorithms are implemented through a traditional BCI pipeline. A three-class (Left, Right, and Up) EEG dataset consisting of nine sessions acquired from a single subject on different days is used for performance evaluation. Moreover, for this experiment, multi-class MMFT only performs transfer learning without domain selection. As such, when multi-class MMFT is employed, eight sessions are considered source domains, and a single session is a target domain; all nine sessions take a turn as a target domain.

Figure 5 shows the comparison results between the proposed method, MEKT, and ML algorithms (LDA and RegTree) [34]. From the comparison, one finds that both transfer learning algorithms achieved the best performance across individual target domains. Multi-class MMFT, in this instance, achieved the highest CA of 88.7% when Se6 is a TD and (Se1~Se5 and Se7~Se9) are SDs, while MEKT recorded a highest CA of 80% when Se3 is a TD and (Se1~Se2 and Se4~Se9) are SDs. Moreover, utilizing samples from sessions acquired on different days to train and test the classifiers demonstrated to be an obstacle for ML algorithms, mainly due to the challenge of overfitting, which in turn deteriorated the classification performance. Hence, an inferior CA of 48%, which is the best performance by LDA, is observed when Se7 is used to test and (Se1~Se6 and Se8~Se9) are used to train the classifier, while a highest CA of 42% was recorded when (Se1~Se8) are used to train, and Se9 is used to test RegTree [17,34,35].

Multi-class MMFT recorded superior classification performance, however, NT proved to pose a severe implication on classification performance across individual TDs. Hence, a significant decline in accuracy was observed when Se8 is the TD and (Se1~Se7 and Se9) are SDs. In this instance, a lowest CA of 22.7% was achieved when multi-class MMFT is employed. A drastic decrease in CA is also recorded when Se2 is the TD and (Se1 and Se3~Se9) are SDs. Subsequently, a lowest CA of 20% was achieved when MEKT is employed. The challenge of overfitting demonstrated to be an obstacle when ML algorithms are employed and contributed to significantly poor classification performance. Hence,

lowest accuracies of 34% and 26% were recorded for both LDA and RegTree when Se2 is used to test and (Se1 and Se3~Se9) are used to train both classifiers, respectively [17,36].
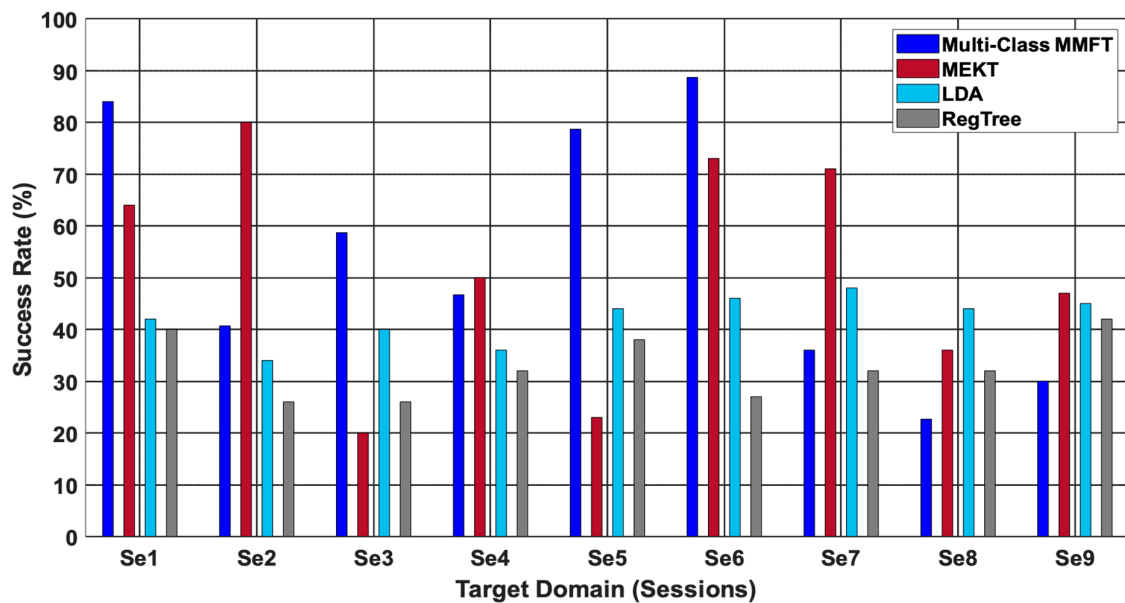

Figure 5. Results for three-class cross-session classification.

Multi-class MMFT in this experiment illustrated that individual sessions could yield a significant high prediction rate for a three-class cross-session classification. However, the existence of variations across sessions leads to the transfer of knowledge from non-related domains, which in turn affects the prediction rate across individual sessions, as illustrated in Se 4, 7–9 in Figure 5, where the multi-class MMFT performed poorly.

3.2.2. Enhanced Multi-Class MMFT for Three-Class Cross-Session Classification

To address the challenge of NT that constitutes low IDRs across individual sessions as a result of variations as observed in the last three sessions in Figure 5. We further enhanced the proposed multi-Class MMFT to first perform domain selection by identifying the optimal combination of domains among multiple sources. After use, the selected optimal combination during TL to improve the prediction rate in the target domain [37,38]. As such, the effectiveness of our proposed approach is evaluated based on classification performance after domain selection and compared with performance before domain selection.

Figure 6 depicts the performance evaluation results before source selection and after source selection when enhanced multi-class MMFT is employed. From the evaluation, one finds that non-related sources affect the prediction rate as a result of NT while selecting only closely related sources minimizes the impact of NT in turn, improves the prediction rate across individual TDs. Therefore, a highest CA of 88.7% was observed before source selection when Se6 is TD and (Se1~Se5 and Se7~Se9) are SDs. However, a CA of 94.7% was recorded after source selection when (Se1~Se2, Se5, Se7, and Se9) were optimally selected as the best combination for TL, as depicted in Table 1. Consequently, a 6% increase in CA was recorded across Se6 after source selection.

Furthermore, a 14% increase in CA is observed across Se1 when (Se4~Se7 and Se8) are optimally selected as the best combination of closely related source domains. Hence, a superior CA of 98%, which is the best performance across all individual TDs, was recorded after source selection, while a prediction rate of 84% was observed before source selection across Se1 when S2~Se9 are SDs.

Notably, a 56% increase in CA was observed across Se7 when (Se2, Se4, Se6, and Se8) were selected as the best combination of closely related sources. A combination of closely related sources proved to significantly enhance performance, with a CA of 92% recorded

after source selection. However, non-related sources were demonstrated to significantly contribute to poor classification performance, with a CA of 36% observed before source selection across Se7.
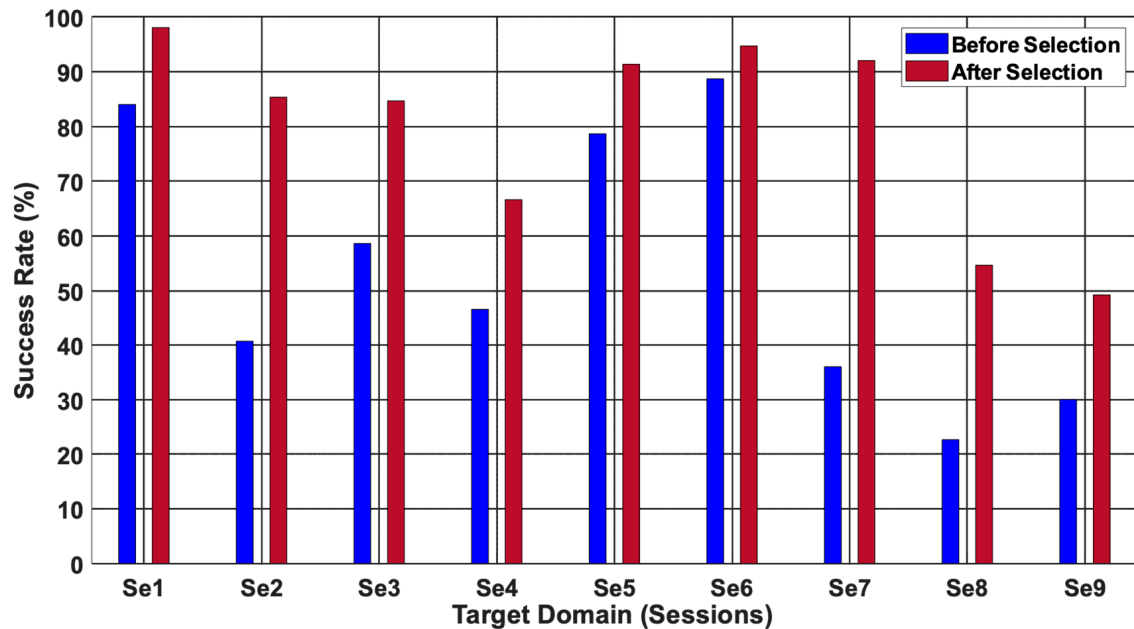

Figure 6. Multi-Class MMFT performance evaluation before and after domain selection based on a three-class cross-session classification.

Table 1. Selected source domain for a three-class cross-session classification.

Target Domain	Selected Source Domain(s)
1	4, 5, 6, 7, 8
2	3, 4, 6, 7, 8
3	5
4	5, 7, 8, 9
5	3, 6
6	1, 2, 5, 7, 9
7	2, 4, 6, 8
8	1, 2, 5
9	1, 4

From these results, one finds that different combinations of source domains can yield different prediction rates across individual domains, hence selecting only the optimal combination of source domains can minimize the impact of NT resulting from the transfer of knowledge from non-related domains, which will in turn, significantly enhance prediction rate across individual target domains.

3.2.3. Multi-Class MMFT for Three-Class Cross-Subject Classification

The impact of a three-class cross-subject classification problem on TL performance is further investigated using nine EEG subjects, mainly to further validate the classification performance of multi-class MMFT on a three-class cross-subject classification problem [39,40]. In a similar manner as the previous experiment, when multi-class MMFT is employed, eight subjects are considered as source domains, while a single subject is a target domain, and all nine subjects take a turn as a target domain.

From Figure 7, one finds that TL algorithms can significantly enhance the classification performance of individual TDs for a three-class cross-subject classification as compared to ML algorithms. Hence, emulating the same TL conditions for ML algorithms by utilizing samples from eight different subjects to train and samples from a single different subject to test the classifiers proved to have a negative effect on classification performance. Consequently, LDA achieved a highest CA of 48% when S5 is used to test and (S1~S4 and S6~S9) are used to train the classifier, while RegTree achieved a highest accuracy of 42% when S6 was used to test and (S1~S5 and S7~S9) were used to train the classifier [17,34,41]. However, multi-class MMFT demonstrated to effectively improve performance, with a highest CA of 75.3% recorded across S3 as a target domain and (S1~S2 and S4~S9) are source domains. An inferior CA was also recorded when MEKT was employed and compared with multi-class MMFT, with a highest CA of 70% observed across S4 as a TD and (S1~S3 and S5~S9) are SDs.
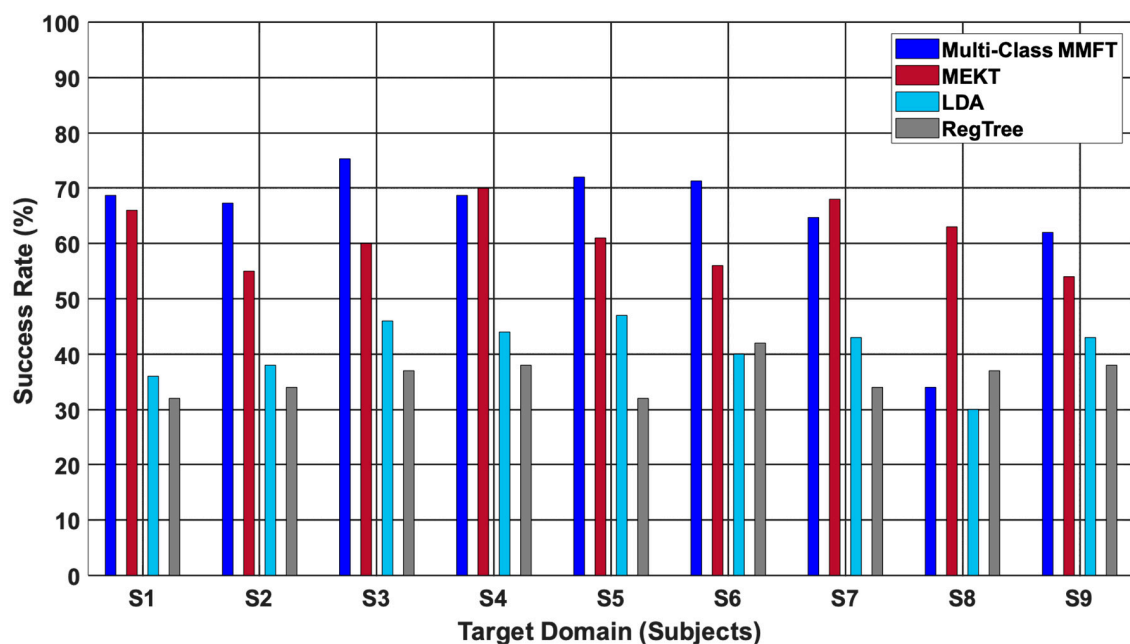

Figure 7. Results for three-class cross-subject classification.

Related source domains proved to be beneficial for multi-class MMFT. However, non-related sources constituted NT, which resulted in poor classification performance across individual TDs. The effect of NT can be attributed to a significant decline in accuracy which is observed across S8 as a TD and (S1~S7 and S9) are SDs, whereby S8 achieved an accuracy of 34% when multi-class MMFT is employed. Samples from non-related subjects were also demonstrated to be an obstacle for both ML classifiers. Hence, a significant decrease in accuracy is noted for ML classifiers, with LDA achieving a lowest accuracy of 30% when S8 was used to test and (S1~S7 and S9) were used to train the classifier, while RegTree achieved a lowest accuracy of 32% when S1 is used to test and (S2~S9) are used to train the classifier.

From these results, one finds that subject-to-subject variations can significantly affect the performance of TL for a three-class cross-subject classification. However, related domains illustrated that multi-class MMFT could significantly enhance classification performance, while non-related domains deteriorate the classification performance of individual target domains, as depicted in Figure 7.

3.2.4. Enhanced Multi-Class MMFT for Three-Class Cross-Subject Classification

In this section, multi-class MMFT is implemented based on a three-class cross-subject classification problem, mainly to address the challenge of inter-subject variations which sig-

nificantly constitute NT [42]. The proposed multi-class MMFT firstly discards non-related subjects by identifying the optimal combination of subjects among multiple source domains. After source selection, then utilize the best combination of closely related SDs to perform transfer learning. In a similar manner as the previous experiment based on EEG sessions, the effectiveness of our proposed approach is evaluated based on classification performance before domain selection and after domain selection using nine three-class subjects [32].

From the resulting Figure 8, a highest classification accuracy of 75.3% was observed before the selection of the optimal combination of source domains, when S3 is the target domain and (S1~S2 and S4~S9) are source domains. However, a 9.4% increase in CA was observed when only (S2 and S7~S8) were optimally selected, then utilized as source domains for TL, as depicted in Table 2. In this case, S3 achieved an accuracy of 84.7% after source selection, as illustrated in Figure 8.
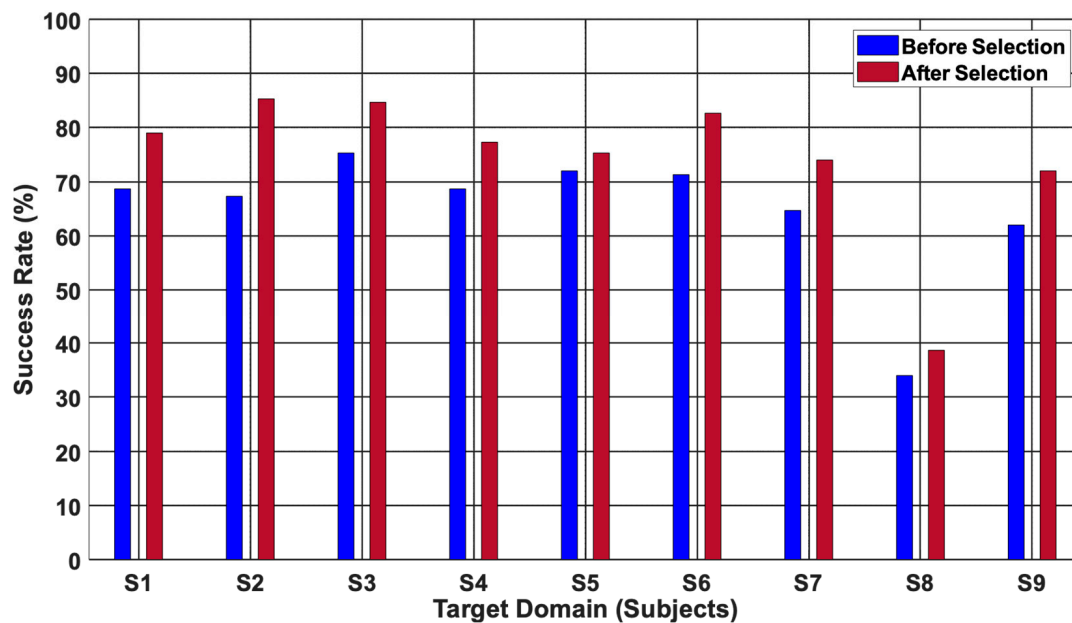

Figure 8. Multi-Class MMFT performance evaluation before and after domain selection based on a three-class cross-subject classification.

Table 2. Selected source domains for a three-class cross-subject classification.

Target Domain	Selected Source Domain(s)
1	3, 8, 9
2	3, 7, 8
3	2, 7, 8
4	5, 8
5	1, 7, 8
6	8, 9
7	2, 5, 9
8	1, 4, 5, 9
9	4, 8

Moreover, a highest classification accuracy of 67.3% was observed before source selection across S2 as a TD when (S1 and S3~S9) are SDs. Notably, an 18% increase in CA was recorded when (S3, S7~S8) were selected as the optimal combination of closely related source domains, with a highest CA of 85.3% observed across S2 after source selection.

Based on these results, it is worth noting that utilizing all subjects as source domains can increase the prediction rate across individual target domains. However, the same sources can deteriorate IDRs across different individual target domains due to the existence of non-related domains characterized by significant variations in EEG dynamics, which lead to NT. Hence, selecting only the best source domains with high transferability can significantly increase the prediction rate in the target domain when enhanced multi-class MMFT is employed.

3.3. Four-Class Problems

3.3.1. Multi-Class MMFT Four-Class Cross-Session Classification

The impact of multi-class cross-subject classification on transfer learning performance is further investigated based on a four-class cross-session classification problem [43]. To validate the classification performance, multi-class MMFT is compared with MEKT and ML algorithms. To facilitate the investigation, nine four-class EEG sessions are used. Subsequently, when multi-class MMFT is employed, eight sessions are assigned as source domains, while a single session is assigned as a TD, and all nine sessions take a turn as a target domain. Moreover, when a traditional BCI is implemented, both LDA and RegTree are trained using eight sessions, while a single session is used to predict the classifiers, respectively.

Consequently, a highest classification accuracy of 99% was recorded across Se7 as a TD and (Se1~Se6 and Se8~Se9) as SDs when MEKT is employed, while a CA of 70.5% was noted across Se4 as a target domain and (Se1~Se3 and Se5~Se9) as source domains when multi-class MMFT is employed. Compared with ML algorithms, an inferior CA of 43%, which is the best performance achieved by LDA, is noted when Se9 is used to test and (Se1~Se8) is used to train. In a similar manner, RegTree did not show traits of any improvement, with a highest accuracy of 40% achieved when Se5 was used to test and (Se1~Se4 and Se6~Se9) was used to train the classifiers.

Notably, a combination of (Se1~Se8) as source domains proved to contain non-related sources. Hence, an inferior classification performance of 24.5% is observed across Se9 when multi-class MMFT is employed. The complexity of a four-class problem and a combination of non-related sources proved to pose a significant challenge for MEKT, with poor performances recorded across Se2~Se6 as TDs. Furthermore, LDA achieved a lowest accuracy of 31% when Se4 was used to test and (Se1~Se3 and Se5~Se9) was used to train the classifier, while RegTree achieved a lowest accuracy of 27% when Se8 was used to test and (Se1~Se7 and Se9) is used to train the classifier.

From these results, one finds that sessional variations can significantly affect the performance of TL for a four-class cross-subject classification problem due to the complexity of a four-class problem. However, as sessions took turns as target domain, some sessions demonstrated to yield superior CA with different combinations of eight source domains. Hence, related domains with high transferability have proven to be more effective in enhancing the prediction rate of individual sessions, while non-related domains deteriorate the prediction rate, as illustrated in Figure 9.

3.3.2. Enhanced Multi-Class MMFT Four-Class Cross-Session Classification

Results from Figure 9 illustrated that the complexity of a four-class problem and non-related sources constituted NT. As a result, the classification performance in the target domain significantly deteriorated. However, closely related sources (sessions) have been demonstrated to enhance classification performance, while a combination of both non-related and related sources has been demonstrated to yield a significant variation in classification performance across individual target domains, as depicted in Figure 9. In this section, enhanced multi-class MMFT is employed to search and select only the optimal combination of closely related sources, then utilize the best-selected sources to enhance the performance of individual target domains for a four-class cross-session classification problem.
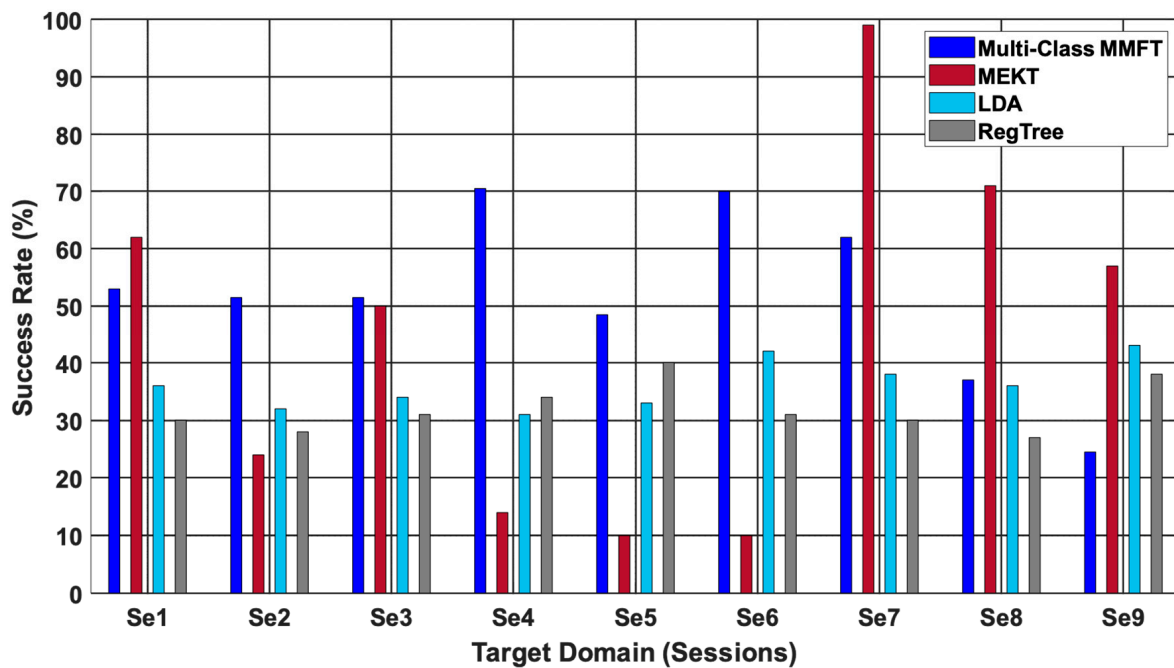
Figure 9. Results for four-class cross-session classification.

Table 3 depicts the best-selected combinations among all source domains as each session took a turn as the target domain. The selected sources vary across individual target domains since the proposed enhanced multi-class MMFT only searches for closely related sources for a specific target domain. As such, a superior classification performance of 70.5% was noted before domain selection across Se4 as a target domain, and the remaining eight sessions are SDs. However, a highest accuracy of 87% was recorded across Se4 after (Se7~Se8) were optimally selected as the best combination of closely related source domains. Hence, a 16.5% increase in accuracy after source selection can be observed across Se4 when enhanced multi-class MMFT is employed, as illustrated in Figure 10. Based on these results, it is worth noting that a combination of both non-related and related sources, including the complexity of a four-class problem, will yield significant variations in classification performances across individual TDs. Hence, classification performance before source selection is characterized by poor prediction rates. However, employing enhanced multi-class MMFT to select only the best sessions proved to effectively improve classification performance across target domains.

Table 3. Selected source domain for a four-class cross-session classification.

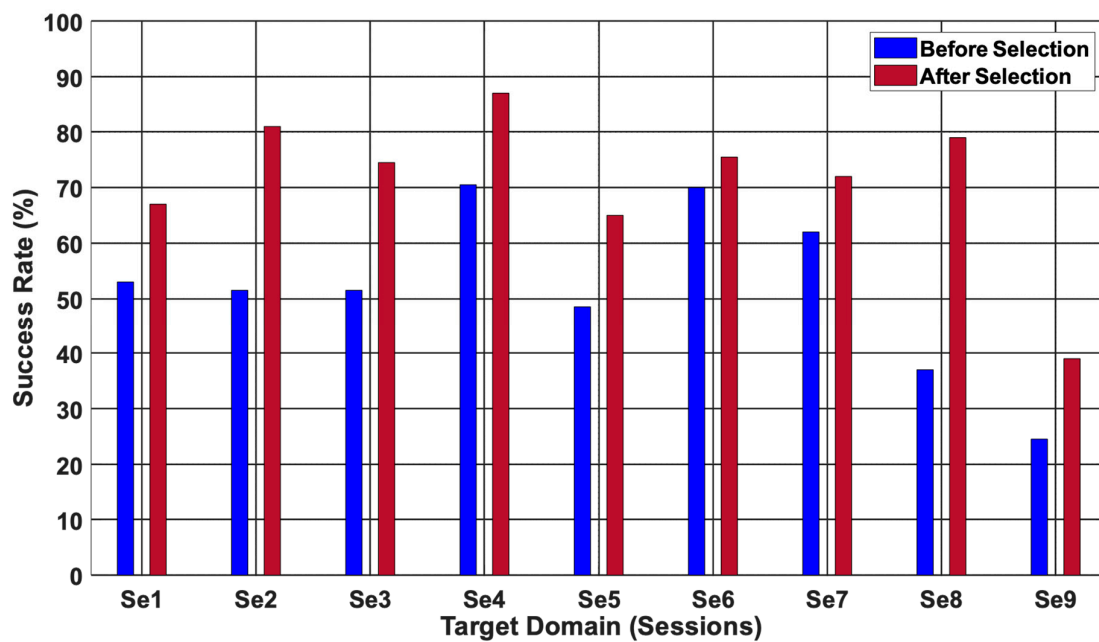| Target Domain	Selected Source Domain(s)
1	2, 3, 5, 6
2	1, 3, 5, 7
3	1, 2, 5, 9
4	7, 8
5	2, 3
6	3, 4, 8
7	3, 4, 6, 9
8	4, 6
9	3, 7

Figure 10. Multi-Class MMFT performance evaluation before and after domain selection based on a four-class cross-session classification.

3.3.3. Multi-Class MMFT for Four-Class Cross-Subject Classification

The impact of multi-class cross-subject classification on transfer learning performance is further investigated based on a four-class problem. A comparative performance analysis is carried out between the proposed multi-class MMFT, MEKT, and two traditional ML algorithms. However, to emulate the same transfer learning condition for a four-class cross-subject classification using ML, eight subjects are used to train, while a single subject is used to test the classifier. Despite the complexity of a four-class cross-subject classification problem, closely related subjects with similar EEG characteristics were demonstrated to effectively enhance the prediction rate in the target domain.

Figure 11 depicts a comparison of classification performance between our proposed multi-class MMFT with MEKT and two ML algorithms for a four-class cross-subject classification problem. The effect of closely related sources on a four-class cross-subject classification problem can be attributed to a high classification performance across individual TDs. Hence, the best performance of 59% is noted across S3 as a TD and (S1~S2 and S4~S9) as SDs when multi-class MMFT is employed. However, the complexity of a four-class cross-subject classification problem proved to be a challenge for TL algorithms. MEKT, on the other hand, did not show any traits of improvement, with the highest CA of 50% recorded across S1, while inferior CA accuracies of 38% and 31% were noted when both LDA and RegTree classifiers were employed, respectively.

Non-related subjects were further demonstrated to be an obstacle for TL algorithms for a four-class cross-subject classification problem. Hence, the effect of NT on classification performance in this instance was denoted by a significant deterioration in prediction rate across individual TDs. It can be observed that a combination of SDs (S1 ~ S8) proved to be dominated by non-related sources, which constituted a poor classification performance of 29.5% depicted across S9 when multi-class MMFT is employed, while MEKT achieved a CA of 38%. Moreover, utilizing samples from different subjects to train and test ML algorithms proved to significantly affect the performance of both LDA and RegTree, with highest accuracies of 41% and 38% recorded when both algorithms were employed, respectively.

Based on these results, it is worth noting that the complex nature of a four-class cross-subject classification problem, including significant subject-to-subject variations, can drastically deteriorate the performance of TL algorithms. Moreover, the presence of non-related sources proved to affect closely related domains with high transferability. Hence,

related sources in this instance demonstrated to have minimal impact on classification performance due to the complexity of a four-class cross-subject classification problem, as illustrated in Figure 11.
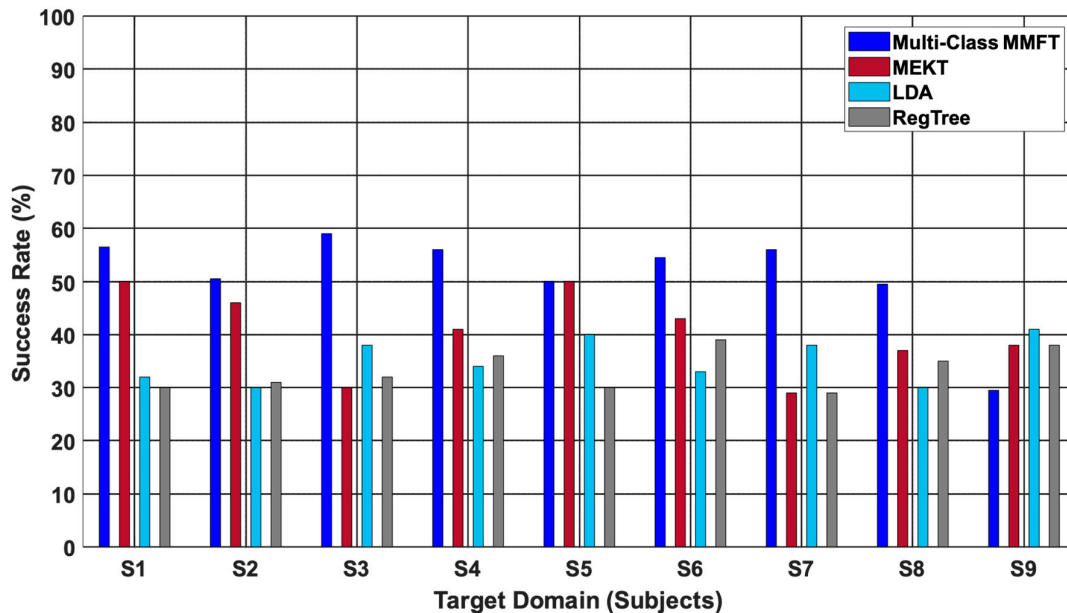

Figure 11. Results for four-class cross-subject classification.

3.3.4. Enhanced Multi-Class MMFT for Four-Class Cross-Subject Classification

Diverse mental states and complex neural connectivity across subjects can result in significant variations in EEG neural dynamics. Additionally, diversity in cognitive states over time constitutes variations in EEG neural dynamics from subject to subject. Moreover, subjects consisting of non-related EEG characteristics demonstrated to deteriorate the prediction rate of individual target domains as a result of NT. Subjects consisting of closely related EEG characteristics were demonstrated to enhance the classification performance of individual target domains. In this case, the enhanced multi-class MMFT was employed to select only the optimal combination of closely related subjects, then utilized the best-selected combinations of sources to enhance the classification performance of individual target domains for a four-class cross-subject classification problem.

Table 4 shows the optimal selected closely related subjects (source domains) as each subject took a turn as the target domain. Notably, optimally selecting (S3 and S7) as SDs proved to yield a superior classification performance, with a CA of 71.5% recorded across S6 as a TD after source selection, while an inferior performance of 54.5% was observed before domain selection across S6 when all eight subjects are SDs as illustrated in Figure 12. Consequently, a 17% increase in accuracy was observed after optimally selecting only closely related source domains.

Based on these results, it is worth noting that the complexity of a four-class problem, together with inter-subject variations in EEG dynamics, can significantly affect the performance of individual subjects for cross-subject EEG classification. However, selecting only the best combination of closely related subjects proved to effectively minimize the impact of NT, at the same time, enhance the performance of TL algorithms which is attributed to high prediction rate across individual target domains after source selection.

Table 5 depicts the mean classification performances of three-class and four-class problems for both sessions and subjects. From the comparison, one finds that multi-class MMFT achieved a superior mean classification performance of 64.9%, as compared to both MEKT and ML algorithms for a three-class problem. However, the impact of NT on classification performance was observed across both three and four-class problems when multi-class MMFT is employed. Hence, enhanced multi-class MMFT was proposed

to address the challenge of NT by selecting only the best combination of closely related source domains. Consequently, superior mean classification performance of 79.6% was recorded when enhanced multi-class MMFT was applied to three-class sessions. These results further support and validate the significance of selecting only the best combination of closely related sources as SDs for transfer mapping to minimize the impact of NT.

Table 4. Selected source domain for a four-class cross-subject classification.

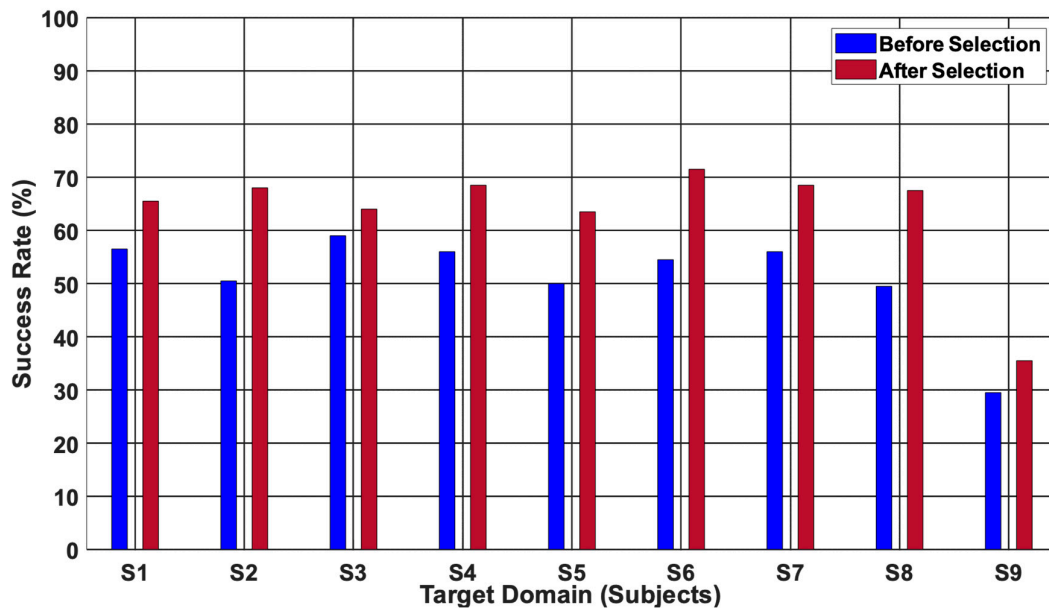| Target Domain	Selected Source Domain(s)
1	2, 7, 9
2	5, 8
3	6, 7, 8
4	1, 2
5	4
6	3, 7
7	3, 8
8	6, 7, 9
9	1, 3

Figure 12. Multi-Class MMFT performance evaluation before and after domain selection based on a four-class cross-subject classification.

Table 5. Mean (%) and standard deviation (in parenthesis) of both transfer learning and machine learning algorithms.

Methods	Three-Class (Sessions)	Three-Class (Subjects)	Four-Class (Sessions)	Four-Class (Subjects)
LDA	42.1 (4.65)	40.9 (5.6)	36.1 (4.23)	35.1 (4.23)
RegTree	32.77 (6.036)	36 (3.28)	32.3 (4.77)	32.9 (4.14)
MEKT	51.6 (22.02)	61.4 (5.79)	44.1 (31.4)	40.4 (7.73)
Multi-class MMFT	54 (24.6)	64.9 (12.23)	52 (14.84)	51.3 (8.81)
Enhanced Multi-class MMFT	79.6 (18.1)	74.3 (14.16)	71.1 (13.83)	63.6 (10.83)

Figure 13 depicts a comparison in classification performance between MI and SSMVEP datasets for both three-class and four-class problems. In this section, we evaluate the per-

formance of each dataset using the proposed multi-class MMFT, MEKT, and ML algorithms (LDA and RegTree) before source selection. A superior classification performance of 64.9% was recorded across a 3-class MI dataset when multi-class MMFT is employed, while a CA of 61.4% was noted when MEKT is applied on the same MI dataset. Moreover, the challenge of overfitting proved to be an obstacle for ML algorithms, with poor classification performance recorded when both LDA and RegTree are applied on the same 3-class MI dataset. The complexity of a four-class problem, including NT, demonstrated to have a significant impact on the classification performance of both 4-class datasets. Hence, a highest CA of 52% was recorded when multi-class MMFT is applied on a 4-Class SSMVEP dataset. MEKT, on the other hand, did not demonstrate any signs of improvements when applied on the same datasets. Moreover, using samples from different subjects to train and test the classifiers proved to give rise to the challenge of overfitting, in turn, results in poor classification performance when ML algorithms are applied to all datasets, as depicted in Figure 13.
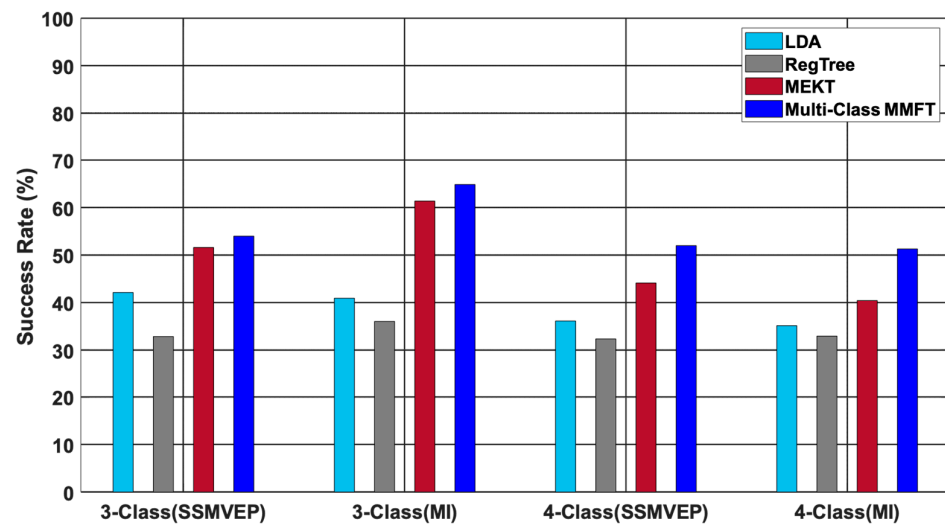

Figure 13. Classification performance for both SSMVEP and MI datasets.

3.4. Experiments on Computation Complexity

In this section, we empirically checked the computational complexity of four different approaches, including our proposed approach, which was implemented in Matlab 2020a on a laptop with i5-1035G1 CPU@1.00 GHz, 16 GB memory, running 64-bit Windows 11 Pro Edition.

In this section, we measure the computational complexity by analyzing the computational time across each method for both three-class and four-class problems.

Both sessions and subjects are considered for computation complexity analysis. Table 6 depicts the computational time for each method, and notably, a significantly low computational time was recorded when TL algorithms were employed as compared to both ML algorithms. Hence, a highest computational time of 5371.8 s is noted when a regression tree is employed. However, a lowest computational time of 6.63 s is noted for a three-class cross-session, while 8.47 s is observed for a three-class cross-subject classification when multi-class MMFT is employed. Moreover, a 10.17 s and 12.58 s computational time was observed for a four-class cross-session and cross-subject classification, respectively, when multi-class MMFT is employed. Domain selection demonstrated to have a significant impact on computation complexity. Hence, a highest computational time is noted for both four-class cross-session (6648.37) and cross-subject (6105.18 s) classification.

Table 6. Computation Complexity.

	Computation Time			
Methods	**Three-Class**	**Three-Class**	**Four-Class**	**Four-Class**
	(Sessions)	**(Subjects)**	**(Sessions)**	**(Subjects)**
LDA	231.71 s	229.7 s	266.96 s	233.19 s
RegTree	4439.61 s	4681.23 s	5371.8 s	5194.59 s
MEKT	7.564 s	8.59 s	10.36 s	13.95 s
Multi-class MMFT	6.63 s	8.47 s	10.17 s	12.58 s
Enhanced Multi-class MMFT	3314.09 s	3972.49 s	6648.37 s	6105.18 s

4. Discussion

Transfer learning is highly effective in terms of increasing classification performance across individual target domains utilizing features acquired from different source domains, and it has been demonstrated to yield superior CA for a two-class cross-subject classification problem [44], while for a multi-class problem, it has been demonstrated to be more effective for within-subject classification [9,45].

However, sessional diversity and subject-to-subject variations demonstrated that non-related domains pose a significant challenge across target domains as a result of NT, which in turn deteriorates the prediction rate across individual target domains [46,47].

In this paper, we proposed a multi-class MMFT approach to investigate the effect of multi-class cross-session and cross-subject classification on TL performance. The proposed approach was evaluated based on a three-class and a four-class problem and compared with MEKT and two classical ML algorithms under similar transfer learning conditions.

A comparative performance analysis demonstrated that the proposed multi-class MMFT could significantly enhance the classification performance of both three-class and four-class cross-session and cross-subject classification problems. Consequently, a superior classification performance of 88.7% was recorded across all individual TDs for three-class cross-session, while a highest prediction rate of 75.3% was recorded for three-class cross-subject classification problems when multi-class MMFT is employed. Moreover, compared to multi-class MMFT, the effect of NT and overfitting proved to pose severe implications on CA when both MEKT and ML algorithms are employed, respectively, as shown in Figure 5. Furthermore, a mean CA of 54% was noted when multi-class MMFT is employed, while a mean CA of 51.6% was observed when MEKT is employed for three-class cross-session classification. The effect of overfitting as a result of significant EEG variations was demonstrated to have an impact on low prediction rates when ML algorithms were employed. Hence, mean CAs of 42.1% and 32.7% were recorded for three-class cross-session classification when LDA and RegTree were employed, respectively, as shown in Table 5.

Moreover, utilizing samples from different sources to train and test ML algorithms further demonstrated to constitute overfitting, which resulted in lower accuracies obtained by the LDA and RegTree, with mean CAs of 40.9% and 36% recorded for three-class cross-subject classification, respectively. However, the proposed multi-class MMFT showed traits of improvements with the best mean classification performance of 64.8% achieved, while MEKT achieved a mean CA of 61.4% for three-class cross-subject classification, the highest accuracy as depicted in Table 5.

A similar decline in accuracy was observed for both four-class cross-session and four-class cross-subject classification problem when ML algorithms were employed [48]. The multi-class MMFT was demonstrated to be effective in minimizing the impact of inter-subject and sessional variations for both four-class cross-session and cross-subject classification problem, as depicted in Figures 9 and 11.

Although multi-class MMFT obtained satisfactory classification performance for both three-class and four-class cross-session and cross-subject classification problems, the impact

of NT and complexity of a four-class problem demonstrated to be an obstacle for the proposed multi-class MMFT across individual target domains, as depicted in Figures 9 and 11. To address this challenge, we further proposed an enhanced multi-class MMFT to first select only the optimal combination of closely related sources among all source domains, then utilize the best combination to perform TL.

The proposed enhanced multi-class MMFT was evaluated based on performance before and after the optimal source domain selection. It was noted that the enhanced multi-class MMFT is effective in minimizing the effect of NT, as depicted in Table 1 and Figure 6. A highest increase of 56% in CA was observed when Se7 is the TD and (Se2, Se4, Se6, Se8) are optimally selected as the best closely related sources. In this case, a CA of 36% was recorded before source selection, while a CA of 92% was recorded after source selection when enhanced multi-class MMFT is employed for a three-class problem.

Moreover, a combination of NT and the complexity of a four-class problem proved to negatively affect the performance of multi-class MMFT across individual target domains [49,50], as depicted in Figures 9 and 11. However, employing enhanced multi-class MMFT was demonstrated to effectively improve performance across individual target domains for both four-class cross-session and cross-subject classification. Notably, a CA of 37% was noted before source selection, while 79% was recorded after source selection when enhanced multi-class MMFT is employed for a four-class problem. A 42% increase in CA was observed after sources selection when Se8 is the TD and (Se4 and Se6) are optimally selected as the best combination.

Furthermore, the classification performance of individual subjects was demonstrated to be significantly affected by NT, which was prevented by the proposed methods, though the complexity of a four-class problem was demonstrated to be expensive in computation when multi-class MMFT is employed to perform cross-subject classification as depicted in Figures 10 and 12.

5. Conclusions

In this work, we proposed a multi-class transfer learning approach based on multi-source manifold feature transfer learning (MMFT) framework, then an enhanced version, using three-class and four-class classification problems considered as examples to validate the performance.

A comparative performance analysis was carried out to evaluate the impact of multi-class cross-session and cross-subject classification on transfer learning performance. The proposed multi-class MMFT was evaluated based on three-class and four-class problems and compared with MEKT and ML algorithms implemented based on a classical BCI pipeline under the same TL conditions.

Multi-class MMFT, in this instance, yielded superior classification performances across individual TDs for both three-class and four-class cross-session and cross-subject classification problems. However, a drastic deterioration in classification performance was recorded when ML algorithms (LDA and RegTree) were employed, with the poor performances attributed to the challenge of overfitting resulting from significant variations in EEG neural dynamics across individual subjects.

Despite the satisfactory performance recorded when multi-class MMFT is employed, non-related domains proved to pose a severe implication on prediction rate across individual target domains as a result of NT.

To address the challenge of NT, we further proposed an enhanced multi-class MMFT framework that first performed source selection, mainly to optimally select the combination of closely related sources, then utilizes the best-selected combination to perform transfer learning. The enhanced multi-class MMFT was evaluated based on performance before source selection, whereby all source domains were considered for TL and performance after source selection, whereby only the best combination of closely related sources wasw considered for TL. A decline in CA was observed across individual TDs before source selection as a result of non-related sources that contributed to the challenge of NT. However,

a significant increase in CA was recorded after source selection when enhanced multi-class MMFT was employed.

Although multi-class MMFT achieved satisfactory classification performance, the complexity of a four-class cross-subject classification problem was demonstrated to be expensive in computation for multi-class MMFT. Hence, future studies will consider more four-class cross-subject classification problems with a lighter computation load.

References

1. Xu, G.; Shen, X.; Chen, S.; Zong, Y.; Zhang, C.; Yue, H.; Liu, M.; Chena, F.; Che, W. A deep transfer convolutional neural network framework for EEG signal classification. *IEEE Access* **2019**, *7*, 112767–112776. [CrossRef]
2. Lin, Y.-P. Constructing a personalized cross-day EEG-based emotion-classification model using transfer learning. *IEEE J. Biomed. Health Inform.* **2019**, *24*, 1255–1264. [CrossRef] [PubMed]
3. Aldayel, M.S.; Ykhlef, M.; Al-Nafjan, A.N. Electroencephalogram-based preference prediction using deep transfer learning. *IEEE Access* **2020**, *8*, 176818–176829. [CrossRef]
4. Shajil, N.; Sasikala, M.; Arunnagiri, A.M. Deep learning classification of two-class motor imagery EEG signals using transfer learning. In Proceedings of the 2020 International Conference on e-Health and Bioengineering (EHB), Iasi, Romania, 29–30 October 2020; IEEE: Piscataway, NJ, USA, 2020; pp. 1–4.
5. Song, Z.; Deng, B.; Wang, J.; Yi, G.; Yue, W. Epileptic Seizure Detection Using Brain-Rhythmic Recurrence Biomarkers and ONASNet-Based Transfer Learning. *IEEE Trans. Neural Syst. Rehabil. Eng.* **2022**, *30*, 979–989. [CrossRef] [PubMed]
6. Wang, H.; Sun, Y.; Wang, F.; Cao, L.; Zhou, W.; Wang, Z.; Chen, S. Cross-Subject Assistance: Inter-and Intra-Subject Maximal Correlation for Enhancing the Performance of SSVEP-Based BCIs. *IEEE Trans. Neural Syst. Rehabil. Eng.* **2021**, *29*, 517–526. [CrossRef]
7. Lee, M.-H.; Kwon, O.-Y.; Kim, Y.-J.; Kim, H.-K.; Lee, Y.-E.; Williamson, J.; Fazli, S.; Lee, S.-W. EEG dataset and OpenBMI toolbox for three BCI paradigms: An investigation into BCI illiteracy. *GigaScience* **2019**, *8*, giz002. [CrossRef] [PubMed]
8. Bird, J.J.; Kobylarz, J.; Faria, D.R.; Ekart, A.; Ribeiro, E.P. Cross-domain MLP and CNN transfer learning for biological signal processing: EEG and EMG. *IEEE Access* **2020**, *8*, 54789–54801. [CrossRef]
9. Li, M.-A.; Xu, D.-Q. A Transfer Learning Method based on VGG-16 Convolutional Neural Network for MI Classification. In Proceedings of the 2021 33rd Chinese Control and Decision Conference (CCDC), Kunming, China, 22–24 May 2021; pp. 5430–5435.
10. Ju, C.; Gao, D.; Mane, R.; Tan, B.; Liu, Y.; Guan, C. Federated transfer learning for EEG signal classification. In Proceedings of the 2020 42nd Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC), Montréal, QC, Canada, 20–24 July 2020; pp. 3040–3045.
11. Zhang, W.; Wu, D. Manifold embedded knowledge transfer for brain-computer interfaces. *IEEE Trans. Neural Syst. Rehabil. Eng.* **2020**, *28*, 1117–1127. [CrossRef]
12. 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]
13. Kim, D.K.; Kim, Y.T.; Jung, H.R.; Kim, H.; Kim, D.J. Sequential Transfer Learning via Segment After Cue Enhances the Motor Imagery-based Brain-Computer Interface. In Proceedings of the 2021 9th International Winter Conference on Brain-Computer Interface (BCI), Gangwon, Republic of Korea, 22–24 February 2021; IEEE: Piscataway, NJ, USA, 2021; pp. 1–5.
14. Lee, D.-Y.; Jeong, J.-H.; Lee, B.-H.; Lee, S.-W. Motor Imagery Classification Using Inter-Task Transfer Learning via a Channel-Wise Variational Autoencoder-Based Convolutional Neural Network. *IEEE Trans. Neural Syst. Rehabil. Eng.* **2022**, *30*, 226–237. [CrossRef]
15. Shahabi, M.S.; Shalbaf, A.; Maghsoudi, A. Prediction of drug response in major depressive disorder using ensemble of transfer learning with convolutional neural network based on EEG. *Biocybern. Biomed. Eng.* **2021**, *41*, 946–959. [CrossRef]

16. Zhang, K.; Robinson, N.; Lee, S.-W.; Guan, C. Adaptive transfer learning for EEG motor imagery classification with deep Convolutional Neural Network. *Neural Netw.* **2021**, *136*, 1–10. [CrossRef] [PubMed]

17. Cao, J.; Hu, D.; Wang, Y.; Wang, J.; Lei, B. Epileptic classification with deep transfer learning based feature fusion algorithm. *IEEE Trans. Cogn. Dev. Syst.* **2021**, *14*, 684–695. [CrossRef]

18. Maswanganyi, R.C.; Tu, C.; Owolawi, P.A.; Du, S. Statistical Evaluation of Factors Influencing Inter-Session and Inter-Subject Variability in EEG-Based Brain Computer Interface. *IEEE Access* **2022**, *10*, 96821–96839. [CrossRef]

19. Gao, Z.; Yuan, T.; Zhou, X.; Ma, C.; Ma, K.; Hui, P. A deep learning method for improving the classification accuracy of SSMVEP-based BCI. *IEEE Trans. Circuits Syst. II Express Briefs* **2020**, *67*, 3447–3451. [CrossRef]

20. Maswanganyi, C.; Tu, C.; Owolawi, P.; Du, S. Factors influencing low intension detection rate in a non-invasive EEG-based brain computer interface system. *Indones. J. Electr. Eng. Comput. Sci.* **2020**, *20*, 167–175. [CrossRef]

21. Chu, Y.; Zhao, X.; Zou, Y.; Xu, W.; Zhao, Y. Robot-Assisted Rehabilitation System Based on SSVEP Brain-Computer Interface for Upper Extremity. In Proceedings of the 2018 IEEE International Conference on Robotics and Biomimetics (ROBIO), Kuala Lumpur, Malaysia, 12–15 December 2018; pp. 1098–1103.

22. Samanta, K.; Chatterjee, S.; Bose, R. Cross-subject motor imagery tasks EEG signal classification employing multiplex weighted visibility graph and deep feature extraction. *IEEE Sens. Lett.* **2019**, *4*, 1–4. [CrossRef]

23. Li, A.; Wang, Z.; Zhao, X.; Xu, T.; Zhou, T.; Hu, H. MDTL: A Novel and Model-Agnostic Transfer Learning Strategy for Cross-Subject Motor Imagery BCI. *IEEE Trans. Neural Syst. Rehabil. Eng.* **2023**, *31*, 1743–1753. [CrossRef]

24. Zhang, Y.; Li, H.; Dong, H.; Dai, Z.; Chen, X.; Li, Z. Transfer learning algorithm design for feature transfer problem in motor imagery brain-computer interface. *China Commun.* **2022**, *19*, 39–46. [CrossRef]

25. She, Q.; Cai, Y.; Du, S.; Chen, Y. Multi-source manifold feature transfer learning with domain selection for brain-computer interfaces. *Neurocomputing* **2022**, *514*, 313–327. [CrossRef]

26. Zhang, Z.; Wang, F.; Pang, Y.; Yan, G. Unsupervised Feature Transfer for Batch Process Based on Geodesic Flow Kernel. In Proceedings of the 2020 Chinese Control and Decision Conference (CCDC), Hefei, China, 22–24 August 2020; IEEE: Piscataway, NJ, USA, 2020; pp. 975–980.

27. Liu, X.; Zhan, Z.; Yuan, J. Domain Adaptation Algorithm based on Manifold Regularization. In Proceedings of the 2021 IEEE International Conference on Artificial Intelligence, Robotics, and Communication (ICAIRC), Fujian, China, 25–27 June 2021; IEEE: Piscataway, NJ, USA, 2021; pp. 30–33.

28. Meng, M.; Lan, M.; Yu, J.; Wu, J.; Liu, L. Dual-Level Adaptive and Discriminative Knowledge Transfer for Cross-Domain Recognition. *IEEE Trans. Multimed.* **2022**. [CrossRef]

29. Kuang, J.; Xu, G.; Tao, T.; Wu, Q. Class-imbalance adversarial transfer learning network for cross-domain fault diagnosis with imbalanced data. *IEEE Trans. Instrum. Meas.* **2021**, *71*, 1–11. [CrossRef]

30. Gu, X.; Cai, W.; Gao, M.; Jiang, Y.; Ning, X.; Qian, P. Multi-source domain transfer discriminative dictionary learning modeling for electroencephalogram-based emotion recognition. *IEEE Trans. Comput. Soc. Syst.* **2022**, *9*, 1604–1612. [CrossRef]

31. Dash, J.C.; Sarkar, D.; Antar, Y. Design of Series-fed Patch Array with Modified Binomial Coefficients for MIMO Radar Application. In Proceedings of the 2021 IEEE International Symposium on Antennas and Propagation and USNC-URSI Radio Science Meeting (APS/URSI), Singapore, 4–10 December 2021; IEEE: Piscataway, NJ, USA, 2021; pp. 1027–1028.

32. Li, J.; Qiu, S.; Shen, Y.-Y.; Liu, C.-L.; He, H. Multisource transfer learning for cross-subject EEG emotion recognition. *IEEE Trans. Cybern.* **2019**, *50*, 3281–3293. [CrossRef] [PubMed]

33. HALTAŞ, K.; ERGÜZEN, A.; Erdal, E. Classification methods in EEG based motor imagery BCI systems. In Proceedings of the 2019 3rd International Symposium on Multidisciplinary Studies and Innovative Technologies (ISMSIT), Ankara, Turkey, 11–13 October 2019; IEEE: Piscataway, NJ, USA, 2019; pp. 1–5.

34. Mathur, P.; Chakka, V.K. Graph signal processing based cross-subject mental task classification using multi-channel EEG signals. *IEEE Sens. J.* **2022**, *22*, 7971–7978. [CrossRef]

35. Zhang, W.; Deng, L.; Zhang, L.; Wu, D. A survey on negative transfer. *arXiv* **2020**, arXiv:2009.00909. [CrossRef]

36. Chen, Y.; Yang, R.; Huang, M.; Wang, Z.; Liu, X. Single-source to single-target cross-subject motor imagery classification based on multisubdomain adaptation network. *IEEE Trans. Neural Syst. Rehabil. Eng.* **2022**, *30*, 1992–2002. [CrossRef]

37. Cui, J.; Jin, X.; Hu, H.; Zhu, L.; Ozawa, K.; Pan, G.; Kong, W. Dynamic distribution alignment with dual-subspace mapping for cross-subject driver mental state detection. *IEEE Trans. Cogn. Dev. Syst.* **2021**, *14*, 1705–1716. [CrossRef]

38. Chen, C.; Li, Z.; Wan, F.; Xu, L.; Bezerianos, A.; Wang, H. Fusing frequency-domain features and brain connectivity features for cross-subject emotion recognition. *IEEE Trans. Instrum. Meas.* **2022**, *71*, 1–15. [CrossRef]

39. Demsy, O.; Achanccaray, D.; Hayashibe, M. Inter-Subject Transfer Learning Using Euclidean Alignment and Transfer Component Analysis for Motor Imagery-Based BCI. In Proceedings of the 2021 IEEE International Conference on Systems, Man, and Cybernetics (SMC), Melbourne, Australia, 17–20 October 2021; p. 3176.

40. Wei, X.; Ortega, P.; Faisal, A.A. Inter-subject deep transfer learning for motor imagery eeg decoding. In Proceedings of the 2021 10th International IEEE/EMBS Conference on Neural Engineering (NER), Virtual Event, Italy, 4–6 May 2021; pp. 21–24.

41. Chen, C.Y.; Wang, W.J.; Chen, C.C. Multiclass Classification of EEG Motor Imagery Signals Based on Transfer Learning. In Proceedings of the 2022 8th International Conference on Applied System Innovation (ICASI), Nantou, Taiwan, 22–23 April 2022; pp. 140–143.

42. Jiang, Z.; Chung, F.L.; Wang, S. Recognition of multiclass epileptic EEG signals based on knowledge and label space inductive transfer. *IEEE Trans. Neural Syst. Rehabil. Eng.* **2019**, *27*, 630–642. [CrossRef]

43. Lin, J.; Liang, L.; Han, X.; Yang, C.; Chen, X.; Gao, X. Cross-target transfer algorithm based on the volterra model of SSVEP-BCI. *Tsinghua Sci. Technol.* **2021**, *26*, 505–522. [CrossRef]

44. 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] [PubMed]

45. He, H.; Khoshelham, K.; Fraser, C. A multiclass TrAdaBoost transfer learning algorithm for the classification of mobile lidar data. *ISPRS J. Photogramm. Remote Sens.* **2020**, *166*, 118–127. [CrossRef]

46. Dai, M.; Wang, S.; Zheng, D.; Na, R.; Zhang, S. Domain transfer multiple kernel boosting for classification of EEG motor imagery signals. *IEEE Access* **2019**, *7*, 49951–49960. [CrossRef]

47. Jiang, Y.; Wu, D.; Deng, Z.; Qian, P.; Wang, J.; Wang, G.; Chung, F.L.; Choi, K.S.; Wang, S. Seizure classification from EEG signals using transfer learning, semi-supervised learning and TSK fuzzy system. *IEEE Trans. Neural Syst. Rehabil. Eng.* **2017**, *25*, 2270–2284. [CrossRef]

48. Saha, S.; Ahmed, K.I.; Mostafa, R.; Khandoker, A.H.; Hadjileontiadis, L. Enhanced inter-subject brain computer interface with associative sensorimotor oscillations. *Healthc. Technol. Lett.* **2017**, *4*, 39–43. [CrossRef]

49. Liu, Y.; Lan, Z.; Cui, J.; Sourina, O.; Müller-Wittig, W. Inter-subject transfer learning for EEG-based mental fatigue recognition. *Adv. Eng. Inform.* **2020**, *46*, 101157. [CrossRef]

50. Gaur, P.; Chowdhury, A.; McCreadie, K.; Pachori, R.B.; Wang, H. Logistic Regression with Tangent Space-Based Cross-Subject Learning for Enhancing Motor Imagery Classification. *IEEE Trans. Cogn. Dev. Syst.* **2021**, *14*, 1188–1197. [CrossRef]