



Article Comparison of Domain Selection Methods for Multi-Source Manifold Feature Transfer Learning in Electroencephalogram Classification

Rito Clifford Maswanganyi¹, Chungling Tu^{1,*}, Pius Adewale Owolawi¹ and Shengzhi Du^{2,*}

- ¹ Department of Computer Systems Engineering, Tshwane University of Technology,
- Pretoria 0002, South Africa; rcmaswa0809@gmail.com (R.C.M.); owolawipa@tut.ac.za (P.A.O.)
- ² Department of Electrical Engineering, Tshwane University of Technology, Pretoria 0002, South Africa

* Correspondence: duc@tut.ac.za (C.T.); dus@tut.ac.za (S.D.); Tel.: +27-012-382-9943 (C.T.)

Abstract: Transfer learning (TL) utilizes knowledge from the source domain (SD) to enhance the classification rate in the target domain (TD). It has been widely used to address the challenge of sessional and inter-subject variations in electroencephalogram (EEG)-based brain-computer interfaces (BCIs). However, utilizing knowledge from a combination of both related and non-related sources can significantly deteriorate the classification performance across individual target domains, resulting in a negative transfer (NT). Hence, NT becomes one of the most significant challenges for transfer learning algorithms. Notably, domain selection techniques have been developed to address the challenge of NT emerging from the transfer of knowledge from non-related sources. However, existing domain selection approaches iterate through domains and remove a single low-beneficial domain at a time, which can massively affect the classification performance in each iteration since SDs respond differently to other sources. In this paper, we compare domain selection techniques for a multi-source manifold feature transfer learning (MMFT) framework to address the challenge of NT and then evaluate the effect of beneficial and non-beneficial sources on TL performance. To evaluate the effect of low-beneficial and high beneficial sources on TL performance, some commonly used domain selection methods are compared, namely, domain transferability estimation (DTE), rank of domain (ROD), label similarity analysis, and enhanced multi-class MMFT (EMC-MMFT), using the same multi-class cross-session and cross-subject classification problems. The experimental results demonstrate the superiority of the EMC-MMFT algorithm in terms of minimizing the effect of NT. The highest classification accuracy (CA) of 100% is achieved when optimal combinations of high beneficial sources are selected for two-class SSMVEP problems, while the highest CAs of 98% and 87% are achieved for three- and four-class SSMVEP problems, respectively. The highest CA of 98% is achieved for two-class MI classification problems, while the highest CAs of 90% and 71.5% are obtained for three- and four-class MI problems, respectively.

Keywords: brain–computer interface (BCI); electroencephalogram (EEG); transfer learning (TL); negative transfer (NT); multi-source manifold feature transfer learning (MMFT); enhanced multi-class MMFT (EMC-MMFT)

1. Introduction

The brain–computer interface (BCI) is a communication system that interprets commands generated from the brain and utilized to control external devices, such as prosthetic limbs or motorized wheelchairs [1]. Therefore, the neural activities of the brain can be extracted in the form of EEG signals and transformed into control commands utilized to aid individuals suffering from severe motor disabilities in communicating with the surrounding environment without any muscular stimulation [2]. In this case, the efficient translation of EEG signals into control commands is significantly dependent on feature extraction and classification algorithms. Hence, in recent years, improving classification



Citation: Maswanganyi, R.C.; Tu, C.; Owolawi, P.A.; Du, S. Comparison of Domain Selection Methods for Multi-Source Manifold Feature Transfer Learning in Electroencephalogram Classification. *Appl. Sci.* 2024, *14*, 2326. https:// doi.org/10.3390/app14062326

Academic Editor: Qi-Huang Zheng

Received: 14 January 2024 Revised: 1 March 2024 Accepted: 7 March 2024 Published: 10 March 2024



Copyright: © 2024 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (https:// creativecommons.org/licenses/by/ 4.0/). performance has been a focus area for both online and offline EEG-based BCIs. As such, different methods have been implemented to enhanced classification performance.

However, sessional and inter-subject variations in EEG characteristics, emerging as a result of factors such as concentration or mental fatigue, can significantly contribute to poor classification performance in EEG-based BCIs [3,4]. Hence, transfer learning was introduced in this field to address the challenge of inter-subject and sessional variations with more accurate predictions. Moreover, recent studies have demonstrated that TL can successfully enhance performance for inter-subject EEG classification, making use of features acquired from SDs to improve the learning performance in the TD [5–7]. However, negative transfer emerging as a result of the transfer of knowledge from non-related source domains to a target domain still remains one of the most significant challenges when transfer learning is applied for both inter-session and inter-subject EEG classification [8]. As such, domain selection techniques have been developed to address the challenge of negative transfer. However, the currently existing domain selection algorithms iterate through domains, evaluating the transferability or similarities between the SDs and the TD, mainly to identify and remove a single less-related domain in each iteration. In this instance, removing a single non-related domain in each iteration can still pose a significant challenge in the next iteration, since different domains respond differently to various sources.

Li et al. [9] proposed a multi-source TL framework for emotion recognition using fifteen sessions acquired from a single subject. The proposed framework firstly performs source selection and then transfer mapping to minimize variations across the target and the selected source domains. Multi-source supervised-style transfer mapping (MM-S-STM) in this instance achieved the highest mean classification accuracy of 88.92%. Zhang et al. [10] employed domain transferability estimation (DTE) to address the challenge of negative transfer based on a two-class cross-subject classification problem. As such, DTE iterated through domains, removing a single less-beneficial domain in each iteration mainly to minimize the impact of NT on TL performance. Consequently, the highest classification accuracy of 82.15% was achieved when DTE was employed to perform domain selection. In this instance, the complexity of a two-class problem demonstrated to cause a minimal impact on transfer learning performance; however, the combination of both related and non-related source domains in each iteration when a single non-related domain is removed can affect the performance across individual target domains. Furthermore, She et al. [11] evaluated the transferability of domains utilizing labels to analyze similarities between domains. Hence, label similarity analysis multi-source manifold feature transfer (LSA-MMFT) was proposed to address the challenge of NT based on a two-class problem. The proposed framework iterated through domains, removing a single less-beneficial domain in each iteration using label similarities. Subsequently, LSA-MMFT achieved a stable CAs across individual TDs, with the highest mean CA of 76.85%. However, attempts to eradicate the effect of NT by iterating through domains and discarding a single lessbeneficial domain in each iteration from a pool of domains possessing significant individual differences was demonstrated to drastically deteriorate classification performance. As such, removing a single low-beneficial domain can still pose severe implications for classification performance in the next iteration, since each domain responds differently to another domain.

In this study, we investigate the impact of less-beneficial and highly beneficial source domains on transfer learning performance and also address the challenge of NT in multiclass inter-session and inter-subject classification scenarios. As such, four domain selection algorithms are introduced and compared mainly to evaluate the impact of both highly and less-beneficial source domains on classification performance across individual TDs. Moreover, the effect of multi-class TL on classification performance is further evaluated across individual TDs, mainly to address the challenge of negative transfer emerging from the transfer of features from non-related sources. Therefore, all four domain selection algorithms (EMC-MMFT, DTE-MMFT, LSA-MMFT, and ROD-MMFT) are firstly implemented based on a two-class problem and further extended to a three-class and four-class inter-session and inter-subject classification scenarios [12]. The main contributions of this study are listed below:

- 1. Investigate the impact of highly beneficial and less-beneficial sources contributing to negative transfer.
- 2. Evaluate the performance of domain selection methods based on two-class, three-class, and four-class cross-subject EEG classification problems.
- 3. A comparative performance analysis is carried out among EMC-MMFT, DTE-MMFT, LSA-MMFT, and ROD-MMFT. The experiment's results show that EMC-MMFT outperforms all the other domain selection algorithms for all the two-class, three-class, four-class, inter-session, and inter-subject classification problems.

The remainder of this paper is organized as follows: Section 2 introduces related work on domain selection algorithms. Section 3 introduces the materials and methods. Section 4 presents the experimental results. Section 5 discusses and analyzes the results. And, finally, Section 6 is the conclusion.

2. Related Work

In this section, we firstly review three related domain selection algorithms—namely, rank of domain (ROD) [10], domain transferability estimation (DTE) [13], and label similarity analysis (LSA) [11]—for multi-source manifold feature transfer (MMFT). In this case, all three algorithms are implemented on two-class cross-subject classification problems but can be extended to solve multi-class cross-subject classification problems.

2.1. Rank of Domain (ROD)

The primary focus of domain selection is to minimize EEG variations between the TD and SDs, mainly to minimize the effect of NT, which has a significant impact on classification performance across individual TDs [14,15]. ROD iterates through domains to identify less-beneficial sources, and, in each iteration, a single less-beneficial domain is discarded, while the remaining domains are considered SDs, and a single domain is considered TD when transfer mapping is applied. To achieve this goal, ROD makes use of *KL* divergence to calculate the adaptation level across both source and target domains; as such, similarly distributed domains are denoted by a lower *KL* divergence [13].

$$\mathcal{R}(S,T) = \frac{1}{d^*} \sum_{i}^{d^*} \theta_i [KL(S_i||T_i) + KL(T_i||S_i)]$$

$$\tag{1}$$

ROD makes use of three steps to rank domains that are closely related and computed using Equation (1). The first step identifies the best dimensionality for the subspaces, with the dimensionality being denoted by d^{*}. The second step calculates the symmetrized *KL* divergences across data distributions, after having estimated the data distributions of the two domains with two one-dimensional Gaussian functions. The third step measures the *KL* divergence-weighted average of the principal angles. In this case, the two Gaussian distributions are denoted by S_i and T_i , whereby the principal angle denoted by *i* is correlated with the principal vectors utilized to acquire the two Gaussian distributions from the projected data. A combination of sources consisting of low principal angles and *KL* divergences will results in a low value of $\mathcal{R}(S, T)$, which simply implies that the sources are closely related [13].

2.2. Domain Transferability Estimation (DTE)

DTE is an unsupervised source selection technique that iterates through domains to identify more-beneficial domains and, at the same time, discards less-beneficial domains by evaluating the transferability across multiple domains [10].

$$\mathbb{S}_i = \left\{ X_S^{(i)}, \mathbf{y}_S^{(i)} \right\}_i^{\mathcal{Z}} = 1$$
(2)

In this case, Equation (2) represents the number of labelled source domains depicted by \mathbb{S}_i , with the label vector being denoted by $\mathbf{y}_S^{(i)}$, while the feature matrix acquired from the

source domains is denoted by $X_S^{(i)}$. Moreover, each individual-source domain is depicted by *i*, while the overall number of domains is denoted by \mathcal{Z} [10].

$$DIS(\mathbb{S}_i) = \left\| S_b^{\mathbb{S}_i} \right\|_1 \tag{3}$$

The discriminability of individual-source domains is denoted by $DIS(\mathbb{S}_i)$ and computed utilizing Equation (3), with \mathbb{S}_i depicting the number of labelled source domains and the between-class scatter matrix being denoted by S_b .

$$DIF(\mathbb{S}_{i},\mathbb{T}) = \left(\left\| S_{\mathbf{b}}^{\mathbb{S}_{i},\mathbb{T}} \right\|_{1} \right)$$
(4)

Variations across the source domains and the target domain are computed using Equation (4) and represented by $DIF(\mathbb{S}_i, \mathbb{T})$, whereby S_b represents the between-class scatter matrix, \mathbb{S}_i represents the labelled source domains, and the unlabelled feature matrix is denoted by \mathbb{T} [10].

$$r(\mathbb{S}_i, \mathbb{T}) = \frac{DIS(\mathbb{S}_i)}{DIF(\mathbb{S}_i, \mathbb{T})}$$
(5)

The discriminability of individual-source domains depicted by Equation (3) and the differences between the source domains and the target domain depicted using Equation (4) are utilized to form Equation (5). In this instance, Equation (5) is utilized to compute the transferability of the source domains, denoted by $r(\mathbb{S}_i, \mathbb{T})$, with $DIF(\mathbb{S}_i, \mathbb{T})$ representing the variation between the source and target domains, while $DIS(\mathbb{S}_i)$ represents the discriminability of the *i*-th SD, \mathbb{T} represents the target domain, and the labelled source domains are denoted by \mathbb{S}_i [10].

2.3. Label Similarity Analysis MMFT (LSA-MMFT)

Identifying non-related domains to minimize the effect of NT is highly significant in enhancing classification performance across individual target domains. Hence, the LSA-MMFT algorithm first performs source selection by iterating through domains to identify the beneficial sources utilized to perform transfer mapping. To achieve this goal, LSA-MMFT utilizes both single-to-single (STS) and multi-source transfer (MST) to evaluate the similarities across domains [11]. Consequently, to effectively detect the transferability of individual-source domains, the labels acquired using the STS classifier are compared with the labels acquired through the MST classifier. As such, LSA-MMFT iterates through domains, evaluating the label similarities across the domains, and then discards a single less-related domain in each iteration based on said label similarities.

$$sim(y_{real}, y_{ps}) = \sum_{i=1}^{n_t} A_i$$
(6)

$$A_{i} = \begin{cases} y_{real,i} = y_{ps,i} \\ 0 \text{ otherwise} \end{cases}$$
(7)

The label similarity across domains is evaluated using Equation (6) and is denoted by $sim(y_{real}, y_{ps})$, with n_t representing the number of samples in the target domain [11]. The predicted single-source domain labels are denoted by y_{ps} , and the target domain labels are denoted by y_{real} , while A_i represent the overall number of samples, as denoted by Equation (7).

$$Accuracy = sim(y_{real}, y_{ps}) / n_t$$
(8)

A single-to-single (STS) transfer prediction rate in this instance is utilized to measure the transferability of the individual-source domain denoted by Equation (8). In this instance, the number of samples in the target domain is n_t , while y_{ps} represents the predicted single-SD labels and y_{real} the TD labels.

$$y_{ps}^{1} = y_{pm}^{1} = y_{real}^{1} y_{ps}^{3} = y_{pm}^{3} = y_{real}^{3}$$
(9)

$$sim(y_{real}, y_{ps}) = sim(y_{real}^1, y_{ps}^1) + sim(y_{real}^2, y_{ps}^2) + sim(y_{real}^3, y_{ps}^3)$$
(10)

To evaluate the transferability of source domains based on STS, Equations (9) and (10) are both utilized, whereby a higher prediction rate is interpreted as a higher transferability, with y_{ps} depicting the predicted single-SD labels and y_{real} representing the TD labels, while the predicted multi-source labels used to determine the transferability of the SDs are denoted by y_{pm} . [11].

$$sim(y_{ps}, y_{pm}) = sim(y_{ps}^1, y_{pm}^1) + sim(y_{ps}^2, y_{pm}^2) + sim(y_{ps}^3, y_{pm}^3)$$
(11)

Equations (9) and (11) are utilized to evaluate the transferability of source domains based on multi-source transfer (MST). In this case, sim() is utilized to quantify the label similarity between two types of predicted TD labels, with the predicted single-SD labels being depicted by y_{ps} and y_{pm} representing the predicted multi-source labels [11].

$$\mathcal{R}\left(\mathcal{D}_{s}^{(j)},\mathcal{D}_{t}\right) = sim\left(f_{s}\left(\mathcal{D}_{s}^{(j)},\mathcal{D}_{t}\right), f_{m}\left(\left\{\mathcal{D}_{s}^{(j)}\right\}_{j=1}^{z},\mathcal{D}_{t}\right)\right)$$
(12)

Furthermore, Equation (12) is utilized to compute the transferability across the source domain and target domain, with the transferability being denoted by \mathcal{R} . The source domain in this instance is denoted by $\mathcal{D}_s^{(j)}$, and the target domain is depicted by \mathcal{D}_t , while *j* represents the number of source domains [11]. The STS transfer classifier is denoted by f_s , while the multi-source transfer classifier is denoted by f_m .

2.4. Enhanced Multi-Class MMFT (EMC-MMFT)

A novel enhanced multi-class MMFT [12] is presented in this section to evaluate the effect of less-beneficial and highly beneficial multi-class sources on transfer learning performance and, at the same time, address the challenge of NT, which has a severe implications for the prediction rate across individual target domains [12]. The EMC-MMFT framework adapts the idea of the two-class inter-subject classification for MMFT in multiclass classification problems, using three-class and four-class classifications as examples. The novel EMC-MMFT framework, in this instance, first performs source selection to minimize individual differences across the domains and then utilizes the selected sources to perform transfer mapping. Hence, the EMC-MMFT approach iterates through the domains to select only beneficial sources with a high transferability from a pool of domains and then utilizes the selected combination of beneficial sources to enhance learning performance in the target domain [12].

The EMC-MMFT is, at first, implemented based on a two-class inter-subject scenario and then is further extended to adapt the idea of two-class inter-subject classification for MMFT to a multi-class classification problem [12]. In this case, the EMC-MMFT framework is extended to three-class and four-class inter-subject scenarios, mainly to investigate the effect of less-beneficial and highly beneficial source domains on transfer mapping performance. To facilitate our investigation, EMC-MMFT is compared with three domain selection algorithms (ROD, DTE, and LSA-MMFT) based on two-class, three-class, and four-class problems mainly to identify highly beneficial SDs for transfer learning and, at the same time, discard the less-beneficial SDs constituting the NT phenomenon that significantly deteriorate the classification performance. As such, EMC-MMFT consists of two components, namely, source selection to search, identify, and then discard lowbeneficial sources and the optimal selection of high beneficial sources for TL [12]. As such, EMC-MMFT utilizes binomial coefficients to evaluate the relationship between domains using the MATLAB function depicted as nchoosek [16].

$$C(n,k) = \frac{n(n-1)...(n-k+1)}{k!} = \frac{n}{k} \frac{n-1}{k-1} \frac{n-2}{k-2} ... \frac{n-k+1}{1}$$
(13)

Moreover, the MATLAB function utilized to compute the binomial coefficients is depicted as B = nchoosek(v, k), whereby the number of all possible combinations of closely related sources is acquired using Equation (13) [16]. Therefore, C(n,k) represents the binomial coefficients to be computed, with the number of possible combination being denoted by k, with n representing all the sources from which high beneficial domains are being selected [16,17]. Once source selection has been completed, the second component of EMC-MMFT utilizes the selected sources to execute transfer mapping as depicted in Figure 1. To achieve this goal, the second component utilizes four phases to perform transfer mapping. The second component first performs domain distribution mean alignment (DMA), whereby the distribution means of each domain on a symmetric positive-defined (SPD) manifold are aligned utilizing KL divergence to evaluate similarities between the domains [12,18]. Then, tangent space features are generated from both the source and target domains, respectively, once the distribution means have been aligned [19]. After tangent space feature extraction, Grassmann manifold feature learning is implemented to learn feature mapping using a geodesic flow kernel [20,21]. Finally, the MMFT classifier is utilized to predict the target domain labels utilizing knowledge from the source domains.



Figure 1. Enhanced multi-class MMFT framework based on a multi-class cross-session and crosssubject scenario.

3. Materials and Methods

3.1. Dataset Descriptions

In this study, our own recorded steady-state motion visual-evoked potential (SSMVEP) dataset is first utilized to facilitate the investigation, whereby nine sessions acquired from a single SSMVEP subject on different days are first considered for experimentation, followed by nine different SSMVEP subjects [22]. Secondly, our own recorded MI dataset consisting of nine subject is further utilized for investigation. Lastly, an MI dataset acquired from a publicly available database, known as BCI competition IV-a, is utilized to validate the classification performance when the EMC-MMFT framework is employed [23]. Moreover, all the datasets are also utilized for a comparative performance analysis when ROD, DTE, and LSA-MMFT are applied. A description of the experimental paradigms for each of the datasets is given in the following sections.

3.1.1. BCI Competition IV-A Dataset

The BCI competition IV-a database consisting of four-class MI datasets acquired from nine subjects was considered for experimentation. In this case, the four-class MI (left, right, both feet, and tongue) data represented EEG signals recorded from the brain using twenty-two Ag/AgCI electrodes [24,25]. All the EEG channels were placed on the surface of the scalp according to a 10–20 electrodes positioning system. Subsequently, the beginning of the experiment was depicted as t = 0 s, and a fixation cross "+" was displayed, while the participants were requested to seat and focus their attention on a computer monitor. Visual cues in the form of arrows pointing to four different directions were displayed on the monitor for t = 1.25 s [26]. The experiment's participants were required to perform four MI classes from t = 3.25 s to t = 6 s, as illustrated in Figure 2.



Figure 2. BCI competition IV-a experimental paradigm timing scheme.

3.1.2. Our Own Recorded SSMVEP Dataset

Furthermore, an SSMVEP dataset was recorded in a confined lab using nine healthy experiment participants (eight male subjects and one female subject, between the ages of 25 and 35); all the experiment's subjects were recruited and randomly selected to facilitate the investigation. A Simulink model was interfaced with a gtec EEG recording system via MATLAB to enable the acquisition of our own recorded SSMVEP dataset. In this case, the EEG recording system consisted of a gNautilus headset and a gNautilus base station that established wireless communication between the two components. The gNautilus headset consisted of 16 electrodes and a transmission module that made it possible to transmit raw EEG signals to the Simulink model, which were then saved to the MATLAB workspace for further processing. All sixteen electrodes were dedicated only to EEG signal acquisition and positioned on the surface of the scalp according to a 10–20 positioning system. Therefore, the EEG signals were recorded at a sampling rate of 250 Hz. At the beginning of the experiment, the subjects were seated facing a projecting monitor, at which point visual cues were displayed on said computer monitor. Four objects corresponding to four SSMVEP classes were sequentially displayed on the monitor [27]. The objects were flickering at different frequencies (29 Hz, 13.3 Hz, 17 Hz, and 21 Hz) and also moving in four different directions (left, right, up, and down), equivalent to the four SSMVEP classes [4,12]. Moreover, whenever an object was projected on the monitor, a beeping sound was used to notify the participants to focus their attention on the flickering object. Each SSMVEP class was recorded for 300 s, resulting in 7500 samples, as illustrated in Figure 3.



Figure 3. Our own SSMVEP dataset experimental paradigm timing scheme.

3.1.3. Our Own Recorded Motor Imagery (MI) Dataset

Our own motor imagery (MI) dataset was recorded in a confined lab using nine healthy experiment participants (eight male subjects and 1 female subject, between the ages of 25

and 35); all the experiment's subjects were recruited and randomly selected. In a similar manner, the same Simulink model and gtec EEG recording system consisting of sixteen electrodes were utilized to extract the MI signals of the brain from all the experimental subjects. Notably, all sixteen electrodes were dedicated to only EEG signal acquisition and positioned on the surface of the scalp according to a 10–20 positioning system, while EEG data in the form of MI signals were sampled at 250 Hz. Moreover, it is worth noting that all nine experimental subjects had no BCI training prior to the experiments. At the beginning of a trial, the subjects were seated facing a projecting computer screen and requested to perform four MI tasks for 300 s each. At the beginning of every 300 s window, an arrow corresponding to four MI tasks was displayed on the screen; the arrow was used to notify the subjects about which MI task to perform. A beeping sound was also utilized to notify the subjects to begin performing the MI task, as illustrated in Figure 4.



Figure 4. Our own MI dataset experimental paradigm timing scheme.

3.2. Data Pre-Processing

Both of our own online EEG datasets representing SSMVEP and MI signals were recorded at a sampling frequency of 250 Hz. Three filters were then applied on the datasets to eliminate the impact of non-physiological artifacts [28]. In this case, a 0.5 Hz to 60 Hz band-pass filter was first applied on the raw SSMVEP and MI datasets to remove the effect of noise, followed by a 50 Hz notch filter utilized to minimize the effect of line noise [29]. Lastly, to remove the effect of noise emerging from the electrodes positioned on the surface of the scalp, a common average reference (CAR) filter was applied.

The BCI competition IV-a dataset was pre-filtered using a band-pass filter at a cut-off frequency between 0.5 Hz and 100 Hz to reduce the impact of background noise, while the effect of line noise was reduced by applying a notch filter on the EEG signals at a cut-off frequency of 50 Hz [30,31].

3.3. Experimental Methods

In this section, the impact of NT on the prediction rate of TL across individual TDs is investigated. In this case, the EMC-MMFT algorithm was introduced and compared with three domain selection algorithms (ROD, DTE, and LSA-MMFT) mainly to determine the effect of NT on performance. The four algorithms were evaluated based on performance after the selection of highly beneficial domains and the rejection of less-beneficial domains. Moreover, the impact of multi-class inter-session and inter-subject classification on transfer mapping performance was further investigated. Therefore, all four domain selection algorithms were first implemented based on a two-class inter-session and inter-subject classification scenario and then further extended to adapt the idea of two-class intersession and inter-subject classification for MMFT to multi-class classification problems. Consequently, each domain selection algorithm was first extended to a three-class intersession and inter-subject classification scenario and further extended to a four-class intersession and inter-subject classification scenario. Notably, the domain selection algorithms were extended to multi-class classification scenarios mainly to evaluate the impact of a three-class problem on TL performance and also determine whether the complexity of a four-class problem contributed to the effect of NT. Therefore, the performance of all the domain selection algorithms was evaluated and compared under two-class, three-class, and four-class inter-session and inter-subject classification problems. In this case, when all four domain selection algorithms were employed, the class labels (y) for each domain and the multi-class (z) samples from the various sources or domains (sessions or subjects) were received as the input parameters. Each algorithm iterated (N) through the domains, while samples from each domain took turn as the TD (X_t), and the remaining samples were selected to be SDs (X_s) in each iteration. As such, the classifier made use of the source domains labels (Y_s) to predict the corresponding target domain labels (Y_t). However, each domain selection algorithm first performed source selection using different techniques to identify closely related highly beneficial domains, followed by transfer mapping using the identified highly beneficial domains as the SDs.

All four domain selection algorithms analysed in this study first perform source selection followed by transfer mapping to combat the effect of NT, which leads to poor classification performances across individual TDs. As such, when EMC-MMFT is applied, the overall number of available domains is first checked and then each domain takes turn as the TD. Each time a TD is assigned, EMC-MMFT iterates through the remaining domains mainly to discard all the less-beneficial domains at once and selects only the best combination of highly beneficial domains, which are then assigned to the roles of SDs when transfer mapping is employed [12]. However, to select the best combination of highly beneficial domains, EMC-MMFT takes advantage of a MATLAB function depicting binomial coefficients, denoted by B = nchoosek (v, k), which is utilized to generate a possible combination of closely related highly beneficial domains in each iteration. Once the highly beneficial SDs have been selected and the TD has been assigned, the distribution means are aligned and the tangent space features extracted from the SDs and the TD, respectively. After feature extraction, Grassmann manifold feature learning is applied, followed by the MMFT classifier, utilized to predict the TD.

Moreover, when LSA-MMFT is applied, the number of available domains is first determined, while each domain takes turn being the TD. Whenever a domain is selected to be the TD, LSA-MMFT iterates through the remaining domains, discarding a single less-beneficial domain in each iteration until two domains remain. In this case, to identify the less-beneficial domains in each iteration, LSA-MMFT utilizes a label similarity analysis based on the STS approach to evaluate the transferability between the SD and the TD. Moreover, once a single less-beneficial domain is removed, the remaining domains are assigned the roles of SDs in each iteration, and then LSA-MMFT performs the distribution of means alignment for both the SDs and the TD in each iteration. Tangent space features are then extracted from the SDs and the TD, respectively. The extracted features are then utilized for Grassmann manifold feature learning, followed by the MMFT classifier, utilized to predict the TD [11]. This process is repeated in each iteration until two domains remain; then, the best performance across each iteration is considered. Moreover, whenever a new domain is assigned the role of TD, the above-mentioned steps are repeated.

Furthermore, ROD first identifies the number of domains available when it is employed to perform source selection, and each domain takes a turn being the TD. Therefore, whenever a domain is selected to be the TD, ROD iterates through all the source domains, removing a single less-beneficial domain in each iteration until only two SDs remain. Notably, to identify the less-beneficial domains in each iteration, ROD ranks the similarities between the SDs and the TD; this is achieved by computing the symmetrized *KL* divergence-weighted average of the principal angles [32]. However, once a single lowbeneficial domain has been identified and removed, the remaining SDs, including the assigned TD, are utilized to perform the distribution of means alignment in each iteration. This is followed by the extracted features are then used for Grassmann manifold feature learning, followed by the MMFT classifier, utilized to predict the TD. This process is repeated in each iteration until two domains remain; then, the best performance across each iteration is considered. Moreover, whenever a new domain is assigned the role of TD, the above-mentioned steps are repeated. When DTE is employed, the transferability between the SD and the TD is evaluated by linking the distribution information with the between-class scatter matrix in the SD and the between-domain scatter matrix between the SD and the TD [10]. As such, DTE iterates through the nine domains, and each of the nine domains takes a turn being the TD. In this instance, a single less-beneficial domain is discarded in each iteration when the transferability between the SD and the TD is evaluated, while the remaining domains are assigned the roles of SDs, utilized to enhance the prediction rate in the TD. The same processing techniques applied when DTE and LSA-MMFT are employed to perform transfer mapping after source selection in each iteration are utilized when ROD is applied. In this case, the process of identifying and discarding a single less-beneficial domain is repeated in each iteration until two domains remain; then, the best performance across each iteration is considered. Notably, whenever a new domain is assigned the role of TD, the abovementioned process is repeated.

4. Results

4.1. Experiment Setup

In this study, each of the four algorithms were implemented under both inter-session and inter-subject classification scenarios. In this section, four domain selection algorithms (ROD, DTE, EMC-MMFT, and LSA-MMFT) utilized for experimentation including data processing were implemented on Matlab 2020a. As such, during our experiments, MATLAB 2020a was running on a Lenovo laptop with i5-1035G1 CPU@1.00GHz and a 16 GB memory, running a 64-bit Windows 11 Pro Edition.

4.1.1. Two-Class Inter-Session and Inter-Subject Classification Scenario

In this experiment, the EMC-MMFT, ROD, DTE, and LSA-MMFT frameworks are implemented based on a two-class inter-session and inter-subject classification scenario. As such, the performance of each of the four domain selection algorithms is evaluated and compared based on a two-class problem, using nine of our own recorded SSMVEP sessions, acquired from a single subject over different days, and our own dataset recorded using nine SSMVEP subjects, respectively. Moreover, the data of nine motor imagery (MI) subjects acquired from a publicly available EEG dataset and our own recorded (MI) dataset acquired from nine subjects are also utilized for experimentation. In this case, when EMC-MMFT is applied on both the data from the nine sessions and subjects, respectively, EMC-MMFT first iterates through the sessions (domains) and then assigns one session/subject the role of TD. However, each of the nine sessions/subjects takes a turn being the TD in each iteration, while the best combinations of highly beneficial SDs are selected from the remaining eight sessions/subjects, whereby the best selected combinations of SDs are utilized for TL [12].

Moreover, in our experiment, LSA-MMFT, ROD, and DTE are all further enhanced and modified for performance evaluation. In this case, when both ROD and the DTE algorithm are applied on the nine sessions/subjects, a single domain is selected to be the TD, and each of the nine sessions/subjects takes a turn being the TD. However, each time a TD is assigned, the algorithms iterate through the remaining eight domains, removing a single less-beneficial domain in each iteration, using mathematical similarity calculations to evaluate the similarities between the domains. The similarities between the domains are evaluated based on the single-to-single (STS) transfer method, then followed by multisource-to-single-target (MTS) transfer [10,13]. As such, whenever a less-beneficial domain is removed in each iteration, the other remaining domains are assigned the roles of SDs and are then utilized for TL.

Furthermore, when LSA-MMFT is employed on both nine sessions/subjects, respectively, each of the nine domains takes a turn being the TD. Therefore, each time a single session/subject is assigned the role of TD, LSA-MMFT iterates through eight SDs, removing a single less-beneficial domain in each iteration based on a label similarities analysis, mainly to identify highly beneficial domains using the STS approach [11]. Once a single less-beneficial domain has been removed, the remaining SDs and the assigned TD are utilized to perform TL based on the MTS approach, in each iteration.

4.1.2. Three-Class Inter-Session and Inter-Subject Classification Scenario

In the second experiment, the EMC-MMFT, ROD, DTE, and LSA-MMFT frameworks are further extended to adapt the idea of a two-class inter-session and inter-subject classification for MMFT to a three-class classification problem. Therefore, nine motor imagery (MI) subjects acquired from the BCI competition IV-a dataset and our own recorded (MI) dataset acquired from nine subjects, including nine of our own recorded SSMVEP sessions acquired from a single subject and nine of our own recorded SSMVEP subjects, are utilized to evaluate the performance of each of the four domain selection algorithms based on a three-class problem. The same experiment procedures utilized in the first experiment are utilized in experiment two. Notably, when the LSA-MMFT, ROD, and DTE algorithms are applied on nine sessions/subjects, a single domain is selected to be the TD, respectively, and each of the nine sessions/subjects takes a turn being the TD. In this case, each time a TD is assigned, the algorithms iterate through the remaining eight domains, removing a single less-beneficial domain in each iteration until a single highly beneficial SD remains. Moreover, when the EMC-MMFT framework is employed, source selection is first applied to select only a combination of highly beneficial SDs, followed by transfer mapping, using the assigned TD and the selected sources. In a similar manner, each session/subject takes a turn being the TD.

4.1.3. Four-Class Inter-Session and Inter-Subject Classification Scenario

The effect of NT, including the impact of less-beneficial and highly beneficial domains on transfer learning performance, is further evaluated in this section based on a more complex classification problem—in this instance, a four-class problem. Hence, the ROD, DTE, LSA-MMFT, and EMC-MMFT frameworks are further extended to adapt the idea of a two-class inter-session and inter-subject classification for MMFT to a four-class classification problem. To facilitate this investigation, nine four-class motor imagery (MI) subjects acquired from the BCI competition IV-a dataset and our own recorded (MI) dataset acquired from nine subjects, including our own recorded SSMVEP dataset acquired from nine sessions and subjects, respectively, are utilized to evaluate the performance of each of the four domain selection algorithms based on a four-class problem. The same experiment procedures utilized in the first and second experiments to select highly beneficial sources and reject less-beneficial sources are still utilized in this experiment when all four domain selection algorithms are employed.

4.2. Experimental Results

4.2.1. Domain Selection Based on a Two-Class Problem

In this section, the impact of NT on transfer learning performance across individual TDs is investigated using nine of our recorded SSMVEP sessions. As such, to facilitate the investigation, four state-of-the-art domain selection algorithms are utilized to evaluate the effect of less-beneficial and highly beneficial domains on transfer mapping performance, mainly to determine the impact of NT on classification performance, depicted as low CAs across individual TDs [33]. Therefore, the EMC-MMFT framework is introduced and compared with three domain selection algorithms (DTE, ROD, and LSA-MMFT) under two-class inter-session classification scenarios. In this case, when EMC-MMFT is employed, a single domain (subject/session) is selected to be the TD, while the best combinations of highly beneficial sources are selected from eight domains and assigned the roles of SDs. In a similar manner, when DTE, ROD, and LSA-MMFT are employed, a single domain is selected to be the TD, while the remaining domains, after the discarding of a single less-beneficial domain in each iteration, are assigned the roles of SDs when transfer mapping is employed after source selection.

Figure 5 presents the classification performance of the EMC-MMFT framework when compared with the three domain selection algorithms (DTE, ROD, and LSA-MMFT) under a two-class inter-session classification scenario. Performance evaluation in this instance is carried out to determine the effect of both less- and highly beneficial domains (which significantly contribute to NT) on the transfer learning prediction rate across individual target domains. From the evaluation, one finds that iterating through the domains and removing a single less-beneficial domain in each iteration proves to pose severe implications for performance across individual TDs, denoted by a low prediction rate across the TDs [34]. Consequently, all four domain selection algorithms yield remarkable classification performances for a less complex classification problem—in this case, a two-class problem. As such, the highest CA of 100% is observed when Se1 is selected to be the TD and a single less-beneficial SD is removed when LSA-MMFT is employed, while the highest CA of 100% is recorded when Se1 is assigned the role of TD and four less-beneficial SDs are removed when DTE is employed. Moreover, the highest CA of 100% is also observed when Se1 is selected to be the TD and three less-beneficial SDs are removed when ROD is employed, as depicted in Figure 5. Notably, a significant decline in CAs is recorded across Se4 and Se9 when DTE, ROD, and LSA-MMFT are employed, with a significant decline of 16% in the CA recorded across Se3 when LSA-MMFT is employed. However, the EMC-MMFT framework maintains superior CAs across all individual TDs, with the highest CA of 100% recorded across Se1, Se3, Se4, Se5, Se6, and Se7. Notably, an inferior CA of 66% is recorded across Se9, and Se1, Se3, and Se9 are assigned the roles of SDs when EMC-MMFT is employed, as depicted in Table 1. In this case, when EMC-MMFT is compared with ROD, DTE, and LSA-MMFT across Se9, EMC-MFFT still maintains a superior CA.



Figure 5. EMC-MMFT performance evaluation based on a two-class inter-session scenario (our own SSMVEP sessions).

9

1

6,7,8

5,7

TD	Two-Class (SDs)				Three-Class (SDs)				Four-Class (SDs)			
	X-Se	X-S1	X-S2	X-S3	X-Se	X-S1	X-S2	X-S3	X-Se	X-S1	X-S2	X-S3
1	2, 3, 4, 5, 8	2, 3, 5, 7, 8	2, 4	2, 3, 7	4, 5, 6, 7, 8	3, 8, 9	2, 3, 4, 7	5, 7, 8, 9	2, 3, 5, 6	2, 7, 9	2, 3, 4, 5	2, 3, 4, 5
2	1, 4, 6, 8, 9	3, 8	3, 5, 6, 7	1, 5, 6, 7	3, 4, 6, 7, 8	3, 7, 8	3,7	1,9	1, 3, 5, 7	5, 8	1, 3, 4, 5, 6	3, 5, 7
3	1, 5, 8	4, 5, 8, 9	2, 4, 7	2, 4, 5, 6, 8	5	2, 7, 8	4	1, 4, 8, 9	1, 2, 5, 9	6, 7, 8	2, 4	2,4
4	2, 3, 5, 7, 8	7,8	1, 3, 6, 9	5	5, 7, 8, 9	5, 8	2, 3	5,9	7,8	1, 2	3, 6	3
5	1, 4, 6, 7, 8	9	2, 9	3, 4, 9	3,6	1, 7, 8	2, 8	4	2,3	4	1, 6, 7, 8, 9	1, 6, 7, 8, 9
6	1, 2, 7, 8	2, 7, 8, 9	2, 4, 7, 9	1, 2, 4, 9	1, 2, 5, 7, 9	8, 9	7	1, 2, 3, 7	3, 4, 8	3,7	2, 4, 7, 8	4, 5, 7
7	1, 3, 9	2, 3, 4, 5	2, 4, 6, 9	1, 4, 6	2, 4, 6, 8	2, 5, 9	1,6	1, 2, 5	3, 4, 6, 9	3, 8	4, 5, 6	3, 5, 6
8	1, 2, 3, 5, 6	3, 5, 6	1, 3, 9	3, 5	1, 2, 5	2,7	5, 9	1, 4, 5	4,6	6, 7, 5	5,9	1, 4, 5, 9

4,8

4.6.7

3, 4, 5,

6,8

5,6

Table 1. Optimal selected combination of beneficial source domains using EMC-MMFT.

In this section, the EMC-MMFT framework is compared with the DTE, ROD, and LSA-MMFT algorithms under a two-class inter-subject classification scenario. As such, nine of our own recorded SSMVEP subjects are further utilized to investigate the impact of less-beneficial and highly beneficial domains on transfer mapping performance. Therefore, the highest CA across all the individual TDs is observed across S6 when LSA-MMFT is employed, with an accuracy of 86.25% recorded, while the highest CAs of 68.75% and 85% are recorded across S6 when DTE and ROD are employed, respectively. However, EMC-MMFT maintains superior classification performances across all the individual TDs when compared with ROD, DTE, and LSA-MMFT. Hence, a superior CA of 95% is recorded when S6 is selected to be the TD, and S2, S4, S7, and S9 are selected as the optimal combination of highly beneficial SDs using EMC-MMFT, as illustrated in Figure 6. Notably, a 40% increase in CA is observed across S1, with the highest CA of 85% being recorded when EMC-MMFT is applied, while a 45% CA is recorded when DTE is employed. In this case, selecting only the best combination of highly beneficial SDs is demonstrated to effectively increase CA in the TD. Hence, a 33.75% increase in CA is observed across S8 when ROD is compared with EMC-MMFT, whereby a CA of 80% is recorded when EMC-MMFT is employed and a CA of 46.25% when ROD is applied. Moreover, removing a single less-beneficial domain in each iteration is demonstrated to significantly affect the CA, with the lowest CA of 43.75% being observed across S5 when LSA-MMFT is employed. However, selecting only highly beneficial domains results in a 23.75% increase in CA, whereby an accuracy of 67.5% is recorded across S6 when EMC-MMFT is applied.

2, 3, 4,

5,6

1, 6, 7

3, 7, 8

The impact of less- and highly beneficial domains on transfer mapping performance is further investigated using nine of our own MI subjects. As such, the EMC-MMFT framework is further compared with the DTE, ROD, and LSA-MMFT algorithms under a two-class inter-subject classification scenario. Consequently, the highest CA of 85% is observed across S2 when DTE is employed, while the highest CAs of 82.5% and 78.75% are recorded across S2 when LSA-MMFT and ROD are applied, respectively. In this case, superior prediction rates are observed across all the individual TDs when EMC-MMFT is employed, whereby the highest CA of 92.5% is recorded across S9 when S3, S4, S5, S6, and S8 are selected as the optimal combination of highly beneficial SDs using EMC-MMFT, as illustrated in Figure 7. Notably, a significant increase of 35% in CA is observed across S9 when EMC-MMFT is compared with DTE, whereby a CA of 57.5% is recorded when DTE is employed, while EMC-MMFT demonstrates its superiority, with a CA of 92.5%

1, 2, 3, 4,

6,7

1, 2, 3, 4,

6,7

being recorded. Moreover, utilizing only the best selected highly beneficial SDs is proven to effectively increase the CA in the TD, with a 32.5% increase in CA observed across S4 when EMC-MMFT is compared with both DTE and LSA-MMFT. In this case, a CA of 73.75% is recorded when EMC-MMFT is employed, while a CA of 41.25% is recorded when DTE and LSA-MMFT are applied, respectively. The results further illustrate that removing a single less-beneficial domain in each iteration can significantly deteriorate the classification performance, as a result of each subject responding differently to another different subject. However, selecting only the best combination of closely related highly beneficial sources can significantly minimize the effect of NT and, in turn, increase performance in the TD.



Figure 6. EMC-MMFT performance evaluation based on a two-class inter-subject scenario (our own SSMVEP subjects).

In this experiment, the impact of NT on transfer mapping performance is further investigated using a BCI competition IV-a dataset consisting of nine MI subjects. Furthermore, to determine the impact of NT on performance, four domain selection algorithms are evaluated based on the selected beneficial domains for transfer mapping. As such, EMC-MMFT is further compared with DTE, ROD, and LSA-MMFT across individual TDs under a two-class inter-subject classification scenario. In this case, the EMC-MMFT framework maintains and achieves superior classification performances across all the individual TDs compared to the DTE, ROD, and LSA-MMFT algorithms. Consequently, the highest CAs of 97% is recorded across S3 when DTE and LSA-MMFT are employed, respectively, while a CA of 96% is observed when ROD is employed. However, a superior CA of 98% is recorded when S3 is assigned the role of TD, and S4, S5, S8, and S9 are selected as the highly beneficial SDs, as depicted in Table 1.

Moreover, removing a single less-beneficial domain in each iteration is demonstrated to pose severe implications for the classification performance, with a significant decline in CA being observed across S8 when DTE, ROD, and LSA-MMFT are employed. The highest CA of 53% is observed when ROD is employed, compared to both DTE and LSA-MMFT, while a 44% increase in CA is observed, with a superior CA of 97% being recorded across S8, when EMC-MMFT is employed, as depicted in Figure 8.



Figure 7. EMC-MMFT performance evaluation based on a two-class inter-subject scenario (our own MI subjects).



Figure 8. EMC-MMFT performance evaluation based on a two-class inter-subject scenario (BCI competition IV-a).

From these results one finds that the complexity of a two-class problem has a minimal impact on the classification performance of ROD, DTE, LSA-MMFT, and EMC-MMFT. However, the presence of less-beneficial domains when transfer mapping is applied can significantly deteriorate classification performance as a result of NT. Notably, removing an individual domain when iterating through domains to identify highly beneficial sources is demonstrated to pose severe implications for the prediction rate. Nonetheless, selecting

only the optimal combination of highly beneficial sources and discarding less-beneficial domains is proven to significantly improve the prediction rate across individual TDs when EMC-MMFT is employed.

4.2.2. Domain Selection Based on a Three-Class Problem

In this section, a three-class inter-session classification scenario is adopted to further investigate the impact of NT on the classification performance of TL. Therefore, to determine the impact of NT on performance, nine of our own recorded SSMVEP sessions are utilized to evaluate the effect of less- and highly beneficial domains on classification performance across individual TDs. Moreover, to facilitate this investigation, EMC-MMFT is compared with DTE, ROD, and LSA-MMFT under the same TL conditions. In this case, discarding a single less-beneficial SD in each iteration from a pool of domains consisting of significant individual differences is demonstrated to severely affect the CA. Hence, an inferior CA of 23.3% is observed across Se8 when ROD is employed. However, a 31.4% increase in CA is observed across Se8 when EMC-MMFT is employed, with the highest CA of 54.7% being recorded. Furthermore, inferior CAs of 36.66% and 36.67% are observed across Se9 and Se7 when both DTE and LSA-MMFT are applied, respectively. Notably, a 53.34% increase is observed across Se8, while a 55.33% increase in CA is observed across Se7 when EMC-MMFT is employed, with Cas of 90% and 92% being recorded across Se8 and Se7, respectively. From this comparison, it can be noted that selecting only a combination of highly beneficial SDs can minimize the effect of NT. Hence, the EMC-MMFT framework maintains superior classification performances across each individual TD compared to ROD, DTE, and LSA-MMFT. Moreover, EMC-MMFT further demonstrates its superiority when Se1 is assigned the role of TD and Se4, Se5, Se6, Se7, and Se8 are optimally selected as highly beneficial SDs, with the highest CA of 98% being recorded, as depicted in Figure 9.



Figure 9. EMC-MMFT performance evaluation based on a three-class cross-session scenario (our own SSMVEP sessions).

Based on these results, it is worth noting that the complexity of a three-class problem, including the effect of NT, can significantly deteriorate classification performance across individual TDs when DTE, ROD, and LSA-MMFT iterate through domains, removing a single less-beneficial domain at a time. However, the complexity of a three-class problem

and the effect of NT demonstrate to pose a minimal impact on classification performance when EMC-MMFT is employed. Hence, first selecting only relevant or highly beneficial sessions can significantly enhance the prediction rate when transfer mapping is applied.

In this section, the impacts of less- and highly beneficial domains on transfer mapping are investigated to determine the impact of NT on classification performance across each TD. As such, a three-class inter-subject classification scenario is adopted to evaluate and compare the four domain selection algorithms using nine of our own SSMVEP subjects. Therefore, the EMC-MMFT framework is compared with the DTE, ROD, and LSA-MMFT algorithm mainly to evaluate the effect of less- and highly beneficial domains on TL performance. Consequently, discarding a single less-beneficial SSMVEP subject in each iteration is demonstrated to have a minimal effect on reducing the impact of NT. Hence, the highest CA of 60.8% is recorded across S6 when LSA-MMFT is employed, while ROD and DTE achieve the highest CAs of 57.5% and 56.6% across S6, respectively. However, discarding all the less-beneficial domains and selecting only the highly beneficial domains when EMC-MMFT is applied are demonstrated to significantly minimize the effect of NT. Hence, a 25% increase in CA is observed across S6 when S7 is optimally selected as a highly beneficial SD, with a superior CA of 85.8% being recorded when EMC-MMFT is applied, as depicted in Figure 10. Moreover, a significant decline in CA is observed across S8 when LSA-MMFT is compared with ROD and DTE, with the lowest CA of 24.16% being recorded when LSA-MMFT is employed. Notably, a 50% increase in CA is observed across S8 when EMC-MMFT is applied and S5 and S9 are selected as highly beneficial SDs; in this instance, a CA of 74.16% is recorded.



Figure 10. EMC-MMFT performance evaluation based on a three-class inter-subject scenario (our own SSMVEP subjects).

In this section, a three-class inter-subject classification scenario is adopted to further evaluate the impact of less-beneficial and highly beneficial domains on transfer mapping, mainly to determine the impact of NT on performance. As such, nine of our own recorded MI subjects are considered for experimentation, whereby the EMC-MMFT framework is compared with the DTE, ROD, and LSA-MMFT algorithms. In this case, ROD outperforms both DTE and LSA-MMFT, with the highest CA of 63.3% being recorded across S9, while CAs of 58.3% and 61.66% are recorded across S9 when both DTE and LSA-MMFT are

employed, respectively. However, a 4.3% increase is observed across S9, with the highest CA of 67.6% being recorded when EMC-MMFT is applied. Notably, EMC-MMFT further demonstrates its superiority when optimal combinations of highly beneficial domains are selected, whereby superior classification performances are recorded across all the individual TDs. Hence, the highest CA of 74.1% is recorded across S5 when S1, S4, S8, and S9 are selected as the best combination of highly beneficial SDs, as illustrated in Figure 11. In this instance, an inferior CA of 28.3% is recorded across S5 when both DTE and LSA-MMFT are employed, respectively. However, a 45.8% increase in CA is observed across S5 when EMC-MMFT is applied. From the comparison results, it can be noted that selecting only the best combination of highly beneficial SDs can effectively minimize the existing challenge of NT and, in turn, increase the CA across the TD, while removing a single less-beneficial SD in each iteration can massively deteriorate the CA. Hence, a 33.75% increase in CA can be observed across S8 when EMC-MMFT is compared with ROD, whereby a CA of 80% is recorded when EMC-MMFT is applied, while a CA of 46.25% is recorded when ROD is applied. In a similar manner, it can also be noted that removing only a single less-beneficial domain in each iteration can significantly decrease the CA, whereby LSA-MMFT records the lowest CA of 26.66% across S3 when seven less-beneficial SDs are removed. However, a 29.2% increase in CA is noted across S3 when EMC-MMFT is employed, with the highest CA of 55.8% being recorded at this time.



Figure 11. EMC-MMFT performance evaluation based on a three-class inter-subject scenario (our own MI subjects).

In this experiment, the impact of NT on the transfer learning performance across TDs is further investigated, whereby four domain selection algorithms are employed to facilitate this investigation by evaluating the effect of less- and highly beneficial source domains on the transfer learning performance. As such, EMC-MMFT is compared with three domain selection algorithms (ROD, DTE, and LSA-MMFT), using a BCI competition IV-a dataset consisting of nine MI subjects. In this case, ROD outperforms both DTE and LSA-MMFT across S3, with a high CA of 80.66% being recorded when ROD is employed, while CAs of 76.66% and 74.66% are noted when both LSA-MMFT and DTE are employed, respectively. EMC-MMFT, on the other hand, outperforms ROD, with the highest CA of

84.66% being recorded across S3 when EMC-MMFT is employed. Moreover, discarding a single less-beneficial domain is proven to contribute to NT, with an inferior CA of 36% being recorded across S8 when ROD is employed. However, a 54% increase in CA is observed across S8, whereby a superior CA of 90% is recorded across S8 when EMC-MMFT is employed, as illustrated in Figure 12.



Figure 12. EMC-MMFT performance evaluation based on a three-class cross-subject scenario (BCI competition IV-a).

From these results, it is worth noting that the number of source domains utilized for transfer mapping has no impact on the classification performance. However, removing a single less-beneficial domain from a pool of SDs consisting of significant individual differences can drastically deteriorate the prediction rate across individual TDs when transfer mapping is applied. Hence, inferior classification performances are recorded when DTE, ROD, and LSA-MMFT are employed. Moreover, the complexity of a three-class problem including the presence of less-beneficial domains is proven to significantly affect the performance of each of the three algorithms in each iteration. Nonetheless, employing the EMC-MMFT framework is proven to enhance the prediction rate, while the complexity of a three-class problem is demonstrated to cause a minimal impact on classification performance when EMC-MMFT is employed to select only optimal combinations of highly beneficial domains for TL.

4.2.3. Domain Selection Based on a Four-Class Problem

A more complex multi-class inter-session classification scenario is explored to evaluate the effect less- and highly beneficial source domains on the TL prediction rate across individual TDs. A comparative performance analysis is carried out under a four-class inter-session classification problem to further investigate the impact of NT on classification performance. As such, nine of our own four-class SSMVEP sessions are utilized to facilitate this investigation, whereby EMC-MMFT is compared with the DTE, ROD, and LSA-MMFT algorithms mainly to determine the effect of source selection on TL performance. Therefore, selecting only highly beneficial SDs with a high transferability to minimize the effect of NT demonstrates to significantly enhance the prediction rate across the individual TDs for a four-class problem.

Hence, superior classification performances are observed across all the individual TDs when EMC-MMFT is applied compared to DTE, ROD, and LSA-MMFT. As such, a superior CA of 87% is recorded across Se4 when Se7 and Se8 are selected as the best combination of highly beneficial SDs when EMC-MMFT is applied. In this case, EMC-MMFT outperforms all three domain selection algorithms, with the highest CA of 74% being recorded across Se4 when ROD is applied, while both DTE and LSA-MMFT achieve the highest CAs of 68.5%, and 79% across Se4, respectively. Notably, a significant decline in CA is observed across Se9, with the lowest CA of 25% being recorded when LSA-MMFT is applied. However, a 49.5% increase in CA is noted across Se9, whereby a CA of 74.5% is recorded when EMC-MMFT is employed, as illustrated in Figure 13.



Figure 13. EMC-MMFT performance evaluation based on a four-class inter-session scenario (our own sessions).

Based on these results, it is worth noting that the complexity of a four-class problem can significantly contribute to a poor performance when DTE, ROD, and LSA-MMFT iterate through the domains to discard a single less-beneficial domain from a pool of domains consisting of significant individual differences. In this case, the complexity of a four-class problem is proven to massively affect the performances of all the domain selection algorithms tested. Hence, inferior classification performances are observed across all the individual TDs. However, the EMC-MMFT framework outperforms all the algorithms and yields remarkable performances compared to ROD, DTE, and LSA-MMFT when the best combinations of highly beneficial domains are selected for TL.

In this experiment, a four-class inter-subject classification scenario is adopted to further investigate the impact of NT on the classification performance of TL. Therefore, a comparative performance analysis is carried out to evaluate the effect of less- and highly beneficial domains on transfer mapping, utilizing nine of our own recorded SSMVEP sessions. As such, the EMC-MMFT framework is compared with DTE, ROD, and LSA-MMFT under a four-class inter-subject classification scenario, mainly to determine the impact of NT on performance. In this case, the complexity of a four-class problem including the effect of NT is proven to pose severe implications for the prediction rate when DTE,

ROD, and LSA-MMFT are applied. Such implications are the result of significant individual differences across domains, mainly because different sources respond differently to other domains. Therefore, iterating through subjects and discarding a single less-beneficial subject, while utilizing the remaining combination of both highly and less-beneficial subjects as SDs, can drastically deteriorate the classification performance.

Hence, inferior performances are observed across individual TDs when DTE, ROD, and LSA-MMFT are employed. Therefore, the highest CA of 38.7% is recorded across S7 when LSA-MMFT is applied, while CAs of 35.6% and 36.25% are recorded across S7 when ROD and DTE are applied, respectively. However, a superior CA of 68.75% is noted across S7 when EMC-MMFT is employed, while a 30.1% increase in CA is observed across S7 when EMC-MMFT is compared with LSA-MMFT. From the comparison results, it can be noted that, despite the complexity of a four-class problem, employing EMC-MMFT to select the best combination of highly beneficial SDs can significantly enhance the CA and, at the same time, minimize the effect of NT. Hence, superior classification performances are observed across all the individual TDs when EMC-MMFT is employed compared to the three domain selection algorithms tested, with a superior CA of 76.2% being recorded across S3 when EMC-MMFT is applied, as illustrated in Figure 14.



Figure 14. EMC-MMFT performance evaluation based on a four-class inter-subject scenario (our own SSMVEP subjects).

In this section, nine of our own recorded four-class MI subjects are further utilized to carry out a comparative performance analysis, mainly to evaluate the effect of lessbeneficial and highly beneficial domains on transfer mapping performance. In this case, to determine the impact of NT on transfer mapping, EMC-MMFT is compared with the DTE, ROD, and LSA-MMFT algorithms under a four-class inter-subject classification scenario. Consequently, the highest CA of 52.6% is observed across S1 when LSA-MMFT is employed and two less-beneficial SDs are removed, compared to both DTE and ROD, with the highest CAs of 51.87% and 52.5% being recorded across S2 and S1, respectively. Therefore, employing EMC-MMFT is demonstrated to yield superior classification performances across all the individual TDs, with a 25% increase in CA being observed when compared with all the three domain selection algorithms tested, whereby a superior CA of 77.6% is recorded across S4 when S3 is selected as a highly beneficial SD using EMC-MMFT, as illustrated in Figure 15. Notably, a significant increase of 50.1% in CA is observed across S4 when EMC-MMFT is applied, compared to ROD, which instead achieves a CA of 27.5% across S4. Furthermore, discarding a single less-beneficial domain in each iteration whilst including the complexity of a four-class problem is proven to significantly affect the prediction rate, with the lowest CA of 17.5% being recorded across S2 when ROD is applied; meanwhile, selecting the combination of highly beneficial domains is proven to minimize the effect of NT and the complexity of a four-class problem. Hence, a 37.1% increase in CA is observed across S2, while a CA of 54.6% is recorded when EMC-MMFT is applied.



Figure 15. EMC-MMFT performance evaluation based on a four-class inter-subject scenario (our own MI subjects).

A four-class inter-subject classification scenario is adopted when a comparative performance analysis is carried out to further investigate the impact of NT on the classification performance of TL. In this case, to examine the impact of NT on the performance of individual TDs, four domain selection algorithms are employed to evaluate the effect of lessand highly beneficial domains on TL. As such, EMC-MMFT is compared with DTE, ROD, and LSA-MMFT, whereby a BCI competition IV-a dataset consisting of nine MI subjects is utilized to facilitate this investigation. The complexity of a four-class problem including the presence of less-beneficial SDs in each iteration is demonstrated to significantly contribute to NT, denoted by poor classification performances across the individual TDs when the DTE, ROD, and LSA-MMFT algorithms are applied [35]. Notably, the lowest CA of 29% is observed across S9 when ROD is employed, while CAs of 33.5% and 31.5% are recorded across S9 when DTE and LSA-MMFT are employed, respectively. However, employing EMC-MMFT to select highly beneficial SDs for transfer mapping is demonstrated to minimize the effect of NT and, in turn, significantly enhance the CA. Hence, a 41% increase in CA is observed across S9 compared to ROD, whereby a CA of 70% is recorded when EMC-MMFT is employed, as depicted in Figure 16.



Figure 16. Enhanced multi-class MMFT performance evaluation based on a four-class cross-subject scenario (BCI competition IV-a).

Furthermore, it can be noted from the comparison that EMC-MMFT maintains a superior CAs across all TDs, with the highest CA of 71.5% being recorded across S6 when S3 and S7 are selected as being highly beneficial SDs. However, removing a single less-beneficial domain in each iteration results in a CA of 70.5% across S6 when ROD is applied, while CAs of 65.5% and 67% are recorded across S6 when DTE and LSA-MMFT are applied, respectively.

From these results, one finds that the complexity of a four-class problem poses severe implications for the classification performance regardless of iterating through the domains and removing a single less-beneficial domain in each iteration. However, selecting only highly beneficial sources is demonstrated to minimize the effect of NT and, at the same time, reduce the impact of the complexity of a four-class problem, thus improving performance across the individual TDs when EMC-MMFT is employed.

4.2.4. Optimal Combination of Selected Beneficial Sources

In this section, we evaluate and compare the classification performances of ROD, DTE, EMC-MMFT, and LSA-MMFT across individual TDs when low-beneficial sources are discarded and high beneficial sources are selected for transfer mapping under two-class, three-class, and four-class inter-session and inter-subject classification scenarios. In this case, ROD, DTE, and LSA-MMFT iterate through the domains and remove a single low-beneficial domain in each iteration until only two domains remain, respectively, mainly to minimize the effect of NT. However, whenever a session/subject is assigned the role of TD, a single low-beneficial domain is removed in each iteration, and the remaining sources are assigned the roles of the SDs utilized for TL [36]. EMC-MMFT first performs source selection, whereby less-beneficial domains are discarded, while the best combination of closely related high beneficial sources are selected as the SDs utilized when TL is performed.

Selected Combination of High Beneficial Source Domains Using EMC-MMFT

Table 1 depicts the optimal selected combination of closely related beneficial sources with a high transferability, with X-Se and X-S1 representing the SSMVEP inter-session and inter-subject, respectively. X-S2 represents the MI inter-subject (our own recorded MI dataset), and X-S3 also represents the MI inter-subject (BCI competition IV-a dataset), whereby the selected sources are utilized as SDs to enhance the prediction rate across individual TDs when TL is employed. In this case, the EMC-MMFT algorithm select only the best combination of closely related high beneficial sources as the SDs from a pool of SDs and then assign a single session/subject to be the TD while each of the nine sessions/subjects takes a turn being the TD, respectively. Notably, EMC-MMFT maintains superior classification performances across each of the nine TDs for all two-class intersubject and inter-session classification scenarios, with the highest CA being recorded when EMC-MMFT is employed and selects Se1, Se5, and Se8 as the optimal combination of highly beneficial SDs for the SSMVEP sessions, while the highest CA is observed when S3 and S8 are optimally selected as being high beneficial SDs for a two-class inter-subject classification problem using the MI subjects acquired from the BCI competition IV-a dataset. Moreover, selecting S2, S4, S7, and S9 and S3, S4, S5, S6, and S8 as high beneficial SDs yields superior CAs for both of the two-class inter-subject classification problems, i.e., when using our own recorded SSMVEP and the MI dataset, respectively.

A similar trend in classification performance is observed for a three-class problem, when EMC-MMFT is employed to select only the best combination of high beneficial sources as SDs. In this case, the highest CA is observed when Se4, Se5, Se6, Se7, and Se8 are optimally selected as the best combination of high beneficial SDs for three-class inter-session classification, while the highest CA is observed when S2 and S7 are selected as the optimal combination of SDs for inter-subject classification, utilizing the MI subjects acquired from the BCI competition IV-a dataset. In a similar manner, selecting S7 and S4 as high beneficial SDs yields superior CAs for both of the three-class inter-subject classification problems, using our own recorded SSMVEP and the MI dataset, respectively.

Furthermore, the EMC-MMFT framework yields superior CAs across the individual TDs for a more complex classification problem—in this instance, a four-class problem. Notably, the highest CA is recorded when Se7 and Se8 are selected as high beneficial SDs for four-class inter-session classification, while the highest CA is recorded when S3 and S7 are selected as the optimal combination of beneficial SDs for four-class inter-subject classification using the MI subjects acquired from the BCI competition IV-a dataset. Notably, superior CAs for both four-class inter-subject classification problems are recorded when S2, S4, S7, and S9 and S3, S4, S5, S6, and S8 are selected as high beneficial SDs, using our own recorded SSMVEP and the MI dataset, respectively. Hence, discarding low-beneficial sources and selecting only a combination of high-beneficial sources is proven to significantly enhance the prediction rate and yields superior classification performances when the EMC-MMFT algorithm is employed.

Number of Removed Low-Beneficial Source Domains Using ROD

Table 2 depicts the number of discarded low-beneficial source domains across each individual target domain when ROD is employed for two-class, three-class, and fourclass inter-session and inter-subject classification scenarios. In this case, the highest CA is recorded when the ROD algorithm assigns Se1 to be the TD and three less-beneficial SDs are removed, while the highest CA is observed when S3 is assigned the role of TD and two low-beneficial SDs are removed for the two-class inter-session scenario using the SSMVEP sessions and for the inter-subject classification problem using the MI subjects acquired from the BCI competition IV-a dataset, respectively. Moreover, for the two-class inter-subject classification problem using our own recorded SSMVEP subjects, the highest CA is observed when S2 is assigned the role of TD and two low-beneficial SDs are removed for two-class inter-subject classification problem using the MI subjects acquired from the BCI competition IV-a dataset, respectively. Moreover, for the two-class inter-subject classification problem using our own recorded SSMVEP subjects, the highest CA is observed when S2 is assigned the role of TD and two low-beneficial SDs are removed for two-class inter-subject classification, using our own recorded MI subjects.

TD		Two-Cla	ass (SDs)			Three-Cl	ass (SDs)			Four-Cla	ass (SDs)	
	X-Se	X-S1	X-S2	X-S3	X-Se	X-S1	X-S2	X-S3	X-Se	X-S1	X-S2	X-S3
1	33	3	3	4	5	4	7	5	2	5	7	3
2	6	6	3	2	1	3	3	2	1	5	3	6
3	3	2	6	6	1	5	5	5	3	6	1	5
4	2	6	6	3	3	5	5	6	2	6	1	2
5	1	6	5	3	1	2	7	6	1	6	4	2
6	4	5	1	3	2	5	6	5	3	6	2	7
7	1	2	1	3	6	2	1	7	3	2	1	2
8	1	4	3	3	1	1	7	7	4	4	7	1
9	6	2	4	1	6	6	1	6	7	2	4	7

Table 2. Optimal selected combination of beneficial source domains using ROD.

Furthermore, when the ROD algorithm is employed for a three-class problem, the highest CA is recorded when Se6 is selected to be the TD and two low-beneficial SDs are removed, while the highest CA is observed when S3 is assigned the role of TD and five low-beneficial SDs are removed for both of the inter-session classification problem using the SSMVEP sessions and the inter-subject classification scenario using MI subjects acquired from the BCI competition IV-a database, respectively. In a similar manner, removing six low-beneficial SDs and assigning S6 to be the TD yields the highest CA across all the TDs when ROD is employed, based on a three-class inter-subject classification problem utilizing SSMVEP subjects, while the highest CA is recorded when S9 is selected to be the TD and six less-beneficial SDs are removed for the three-class inter-subject classification problem using our own recorded MI subjects.

Employing ROD for source selection based on a more complex classification problem results in a significant decline in CA across the individual TDs. Hence, the highest CA is observed when Se4 is selected to be the TD and two low-beneficial SDs are removed for the inter-session classification scenario utilizing the SSMVEP sessions; meanwhile, removing five low-beneficial SDs and assigning S6 to be the TD yields a superior performance compared to the other individual TDs for the inter-subject classification scenario using MI subjects acquired from the BCI competition IV-a dataset. Furthermore, assigning S7 the role of TD and discarding six low-beneficial SDs yields the highest CA across all the TDs when ROD is applied on SSMVEP subjects for a four-class inter-subject classification scenario, while the highest CA is recorded when S9 is selected to be the TD and six lowbeneficial SDs are removed for a four-class inter-subject scenario using our own recorded MI subjects. Significant variations in the classification performances across individual TDs further validate that the number of domains causes a minimal impact on the classification performance; however, discarding a single low-beneficial source domain in each iteration demonstrates to significantly affect the prediction rate, resulting in massive variations in the CAs across individual TDs.

Number of Removed Low-Beneficial Source Domains Using DTE

Table 3 shows the number of removed low-beneficial SDs across each individual target domain when DTE is employed to perform source selection based on two-class, three-class, and four-class inter-session and inter-subject classification scenarios. As such, removing four low-beneficial SDs and assigning Se1 the role of TD yields a remarkable CA; however, a decline in CA is observed across individual TDs when DTE is employed, based on the two-class inter-session classification scenario, while a remarkable CA is observed when S3 is assigned the role of TD and three low-beneficial SDs are removed for the two-class inter-subject classification scenario using MI subjects acquired from the BCI competition IV-a database. Moreover, removing three low-beneficial SDs and selecting S6 as the TD

yields the highest CA across all the TDs when DTE is employed, based on a two-class inter-subject scenario utilizing SSMVEP subjects, while the highest CA is recorded when S2 is assigned the role of TD and four low-beneficial SDs are removed for the two-class inter-subject scenario using our own recorded MI subjects.

TD		Two-Cla	iss (SDs)		Three-Class (SDs)				Four-Class (SDs)			
	X-Se	X-S2	X-S3	X-S3	X-Se	X-S1	X-S2	X-S3	X-Se	X-S1	X-S2	X-S3
1	4	2	5	2	1	7	7	6	7	4	1	6
2	3	4	5	4	6	2	7	5	6	6	5	7
3	6	3	3	1	5	4	2	4	3	2	1	5
4	3	7	3	1	6	4	6	7	1	7	2	3
5	3	2	7	7	1	6	5	7	5	7	7	5
6	2	2	3	5	1	7	5	2	3	7	3	7
7	3	1	3	5	5	2	3	6	1	1	1	5
8	6	2	1	3	6	7	2	6	2	3	7	4
9	7	5	1	1	7	7	5	3	6	7	7	1

Table 3. Optimal selected combination of beneficial source domains using DTE.

Furthermore, when the DTE algorithm is employed for a three-class problem, the highest CA is recorded when Se1 is assigned the role of TD and a single less-beneficial SD is removed, while the highest CA is observed when S6 is assigned the role of TD and seven low-beneficial SDs are removed for both the inter-session scenario using the SSMVEP sessions and the inter-subject classification scenario using our own recorded MI subjects, respectively. In a similar manner, for a three-class inter-subject scenario using our own recorded SSMVEP subjects, the highest CA is observed when the DTE algorithm assigns S6 to be the TD and five low-beneficial SDs are removed, while the highest CA is observed when S9 is assigned the role of TD and three low-beneficial SDs are removed for the three-class inter-subject scenario using our own recorded SDS are removed for the three-class inter-subject scenario using our own the DTE algorithm assigns S6 to be the TD and five low-beneficial SDs are removed.

Furthermore, a significant decline in CA across the individual TDs is observed for a more complex four-class classification problem when DTE is employed. In this case, assigning Se2 the role of TD and removing six low-beneficial SDs in the scenario utilizing the SSMVEP sessions yields a remarkable CA, while removing seven low-beneficial SDs and assigning S6 the role of TD for the scenario using our own recorded MI subjects acquired from the BCI competition IV-a database also yields a remarkable CA for both the inter-session and inter-subject classification scenarios, respectively.

Notably, employing DTE based on a four-class inter-subject problem making use of SSMVEP subjects yields the highest CA across all the TDs when a single less-beneficial SD is removed and S7 is assigned the role of TD, while the highest CA is recorded when S2 is selected as the TD and seven less-beneficial SDs are removed for the four-class inter-subject scenario using our own recorded MI subjects. However, a decrease in the classification performances is observed across the individual TDs whenever a single less-beneficial SD is removed in each iteration, as a result of individual variations across the domains, since different domains respond differently to other domains.

Number of Removed Low-Beneficial Source Domains Using LSA-MMFT

Table 4 depicts the number of low-beneficial SDs discarded in each iteration whenever an individual target domain is assigned when the LSA-MMFT algorithm is employed to perform source selection based on two-class, three-class, and four-class inter-session and inter-subject classification scenarios. In this case, employing LSA-MMFT to select high beneficial domains by evaluating label similarities across the domains and discard a single low-beneficial domain in each iteration yields remarkable CAs. However, significant variations in the CAs across individual TDs is observed with LSA-MMFT performing poorly across several individual TDs. As such, assigning Se1 the role of TD and discarding a single low-beneficial SD yields a remarkable CA; however, significant variations in neural dynamics result in a decline in the CAs across several individual TDs when LSA-MMFT is employed, based on the two-class inter-session classification scenario using the SSMVEP sessions, while a remarkable CA is also observed when S3 is selected as the TD and six low-beneficial SDs are removed in the two-class inter-subject classification scenario using the MI subjects acquired from the BCI competition IV-a database.

TD		Two-Cla	iss (SDs)			Three-Cl	ass (SDs)			Four-Cla	ass (SDs)	
	X-Se	X-S1	X-S2	X-S3	X-Se	X-S1	X-S2	X-S3	X-Se	X-S1	X-S2	X-S3
1	1	7	7	6	6	6	2	4	1	6	7	2
2	6	4	1	7	2	3	1	3	1	4	2	6
3	6	6	1	7	6	7	1	7	4	5	3	4
4	3	4	1	1	5	5	1	7	5	5	2	4
5	1	1	3	1	6	4	7	7	2	6	7	4
6	2	1	7	1	1	7	7	5	3	7	6	2
7	4	4	7	2	7	4	7	4	7	5	6	3
8	6	6	7	7	5	5	2	7	7	2	3	3
9	1	1	7	7	7	6	5	1	3	6	2	7

Table 4. Optimal selected combination of beneficial source domains using LSA-MMFT.

Moreover, employing LSA-MMFT based on a two-class inter-subject scenario making use of SSMVEP subjects yields the highest CA across all the TDs when a seven low-beneficial SD is removed and S7 is assigned the role of TD, while the highest CA is recorded when S2 is assigned the role of TD and seven low-beneficial SDs are removed in the two-class inter-subject scenario using our own recorded MI subjects.

When the LSA-MMFT framework is employed, based on a three-class problem using SSMVEP sessions, the highest CA is recorded when Se6 is assigned the role of TD and a single low-beneficial SD is removed, while the highest prediction rate is recorded when S3 is assigned the role of TD and seven low-beneficial SDs are discarded in the scenario using our own MI subjects obtained from the BCI competition IV-a database. Selecting S6 as the TD and removing seven low-beneficial SDs yield the highest CA across all the TDs when LSA-MMFT is employed for the scenario using SSMVEP subjects, while the highest CA is recorded when S9 is assigned the role of TD and a single low-beneficial SD is removed in the three-class inter-subject scenario using our own recorded MI subjects.

Notably, the four-class classification problem demonstrates to pose severe implications for the classification performance when LSA-MMFT is employed. Hence, assigning Se4 the role of TD and removing five low-beneficial SDs yield a remarkable CA, while selecting S6 as the TD and removing seven low-beneficial SDs also yield remarkable CAs for both the inter-session and inter-subject classification scenarios using the BCI competition IV-a dataset, respectively. Furthermore, for the four-class inter-subject scenario using our own recorded SSMVEP subjects, the highest CA is observed when S7 is assigned the role of TD and six low-beneficial SDs are removed, while the highest CA is observed when S1 is selected as the TD and two low-beneficial SDs are removed in the scenario using our own recorded MI subjects. However, removing a single low-beneficial SD in each iteration from a pool of domains consisting of both low- and high beneficial domains results in a decrease in the CA across individual TDs, mainly as a result of individual variations across the domains.

Table 5 depicts the mean prediction rate for a comparative performance evaluation across all four domain selection algorithms under two-class, three-class, and four-class inter-

session classification scenarios, whereby X-Sessions represents the inter-session classification scenario using our own recorded SSMVEP sessions. As such, DTE outperforms both ROD and LSA-MMFT, with a mean CA of 85.1% being recorded when DTE is employed for a two-class problem. However, a 9.1% increase in CA is observed, with a superior mean CA of 94.2% being recorded when EMC-MMFT is applied in a two-class inter-session classification scenario. Moreover, for a three-class inter-session scenario, DTE yields a remarkable mean CA compared to both ROD and LSA-MMFT, with a mean CA of 68.9% being recorded when DTE is employed. However, a superior mean CA of 84.1% is recorded when EMC-MMFT is applied, whereby a 15.2% increase in CA is observed compared to DTE.

Table 5. Mean (%) and standard deviation (in parenthesis) for domain selection algorithms based on cross-session classification.

Mada a la	Two-Class	Three-Class	Four-Class		
Methods	X-Sessions	X-Sessions	X-Sessions		
LSA-MMFT	75.7 (11.26)	64.1 (14.24)	58.3 (6.79)		
DTE	85.1 (21.3)	68.9 (17.93)	61.2 (14.43)		
ROD	81.6 (20.5)	63.2 (22.73)	57 (12.88)		
EMC-MMFT	94.2 (11.26)	84.1 (14.24)	75.06 (6.79)		

Furthermore, performance evaluation has also been carried out for a more complex classification problem—in this instance, a four-class problem. Superior performances across all inter-session classification scenarios under Table 5 is denoted by bold texts. As such, DTE yields a superior performance compared to ROD and LSA-MMFT, with the highest mean CA of 61.2% being recorded when DTE is employed in a four-class inter-session scenario. However, a 13.85% increase in the mean CA is observed when EMC-MMFT is employed and compared with DTE, with a superior mean CA of 75.05% being recorded when EMC-MMFT is employed. In this case, the superior performances recorded when EMC-MMFT is employed can be attributed to the minimal effect of NT, a result of the best selected combination of closely related highly beneficial SDs, which, in turn, significant increases the CA across individual TDs when TL is applied. However, a decline in CA when ROD, DTE, and LSA-MMFT are employed can be attributed to removing only a single low-beneficial SD in each iteration, since the remaining low-beneficial SDs still pose severe implications for the CA in the next iteration as a result of each domain responding differently to another domain.

Table 6 displays the mean classification result for a comparative performance evaluation across all four domain selection algorithms under two-class, three-class, and four-class inter-subject classification scenarios. In this case, X-S1 represent an inter-subject classification scenario using MI subjects acquired from the BCI competition IV-a dataset, while X-S2 represent an inter-subject classification scenario using our own recorded SSMVEP subjects, and X-S3 represent an inter-subject classification scenario using our own recorded MI subjects. Moreover, Superior performances across all inter-session classification scenarios under Table 6 is denoted by bold texts. From the comparison, one finds that the complexity of a classification problem including the effect of low-beneficial sources has severe implications for the classification performance when a single low-beneficial domain is removed from a pool of domains in each iteration. Consequently, ROD outperforms both DTE and LSA-MMFT, with a mean CA of 85.1% being recorded when ROD is employed. However, an 8.5% increase is observed when EMC-MMFT is employed in X-S1, with a superior mean CA of 93.9% being recorded. Moreover, a 20.5% increase and a superior mean CA of 84.5% are observed when EMC-MMFT is applied and compared with DTE. In this instance, DTE outperforms both ROD and LSA-MMFT, with a mean CA of 64% being recorded when DTE is employed in X-S2 for a two-class problem. A superior mean CA of 78.4% is recorded when EMC-MMFT is employed in X-S3 for a two-class problem, and

a 10.8% increase is observed as compared to the performance of ROD, which achieves a mean CA of 67.6% and outperforms both DTE and LSA-MMFT.

Table 6. Mean (%) and standard deviation (in parenthesis) for domain selection algorithms based on cross-subject classification.

N (1 1		Two-Class			Three-Class		Four-Class			
Methods	X-S1	X-S2	X-S3	X-S1	X-S2	X-S3	X-S1	X-S2	X-S3	
LSA-MMFT	83.3 (6.09)	60.4 (15.5)	65.1 (17.5)	67.1 (5.9)	40.9 (13.2)	41.8 (13.8)	56.6 (2.7)	30.7 (7.9)	35.2 (11.1)	
DTE	83.7 (14.8)	64 (14.2)	60.3 (18.3)	67.1 (11.2)	42.3 (10.6)	44.9 (12.1)	56.9 (9.4)	32.2 (2.7)	36.5 (11.1)	
ROD	85.1 (14.9)	58.2 (8.5)	67.6 (10.5)	69.4 (13.2)	42.1 (7.5)	47.6 (6.9)	56.8 (11.6)	30.2 (5.2)	27.8 (10)	
EMC-MMFT	93.9 (6.09)	84.5 (8.7)	78.4 (12.4)	80 (5.9)	72.8 (10.7)	64.3 (7.6)	67.4 (2.7)	59.8 (12.6)	62.9 (10)	

Moreover, for the three-class inter-subject scenario, ROD yields a remarkable mean CA compared to both DTE and LSA-MMFT, with a mean CA of 69.4% being recorded when ROD is employed in X-S1. However, a superior mean CA of 80% is recorded when EMC-MMFT is applied, while a 10.6% increase in CA is observed compared to ROD's. For the three-class inter-subject classification scenario based on X-S2, a superior mean CA of 72.8% is recorded when EMC-MMFT is applied, while a 30.5% increase in CA is observed when it is compared to DTE, whose highest mean CA of 42.3% is recorded when it is applied and compared to both ROD and LSA-MMFT. A superior mean CA of 64.3% is achieved when EMC-MMFT is implemented based on X-S3 for a three-class problem, with a 16.7% increase in CA being observed when compared to ROD's, which achieves a mean CA of 47.6% and, at the same time, outperforms both DTE and LSA-MMFT.

Furthermore, in our study, performance evaluation has also been carried out for a more complex classification problem—in this instance, a four-class problem. DTE yields a superior performance compared to ROD and LSA-MMFT, with the highest mean CA of 56.9% being recorded when DTE is employed based on X-S1. However, a 10.5% increase in the mean CA is observed when EMC-MMFT is employed and compared with DTE, with a superior mean CA of 67.4% being recorded when EMC-MMFT is employed.

Moreover, a 27.6% increase in the mean CA is observed when EMC-MMFT is employed and compared with DTE, with a superior mean CA of 59.8% being recorded when EMC-MMFT is employed based on X-S2, while DTE achieves the highest mean CA of 59.8% compared to both ROD and LSA-MMFT. An inferior mean CA is observed when ROD, DTE, and LSA-MMFT are employed based on X-S3 for a four-class problem. However, a 26.5% increase in CA is observed when EMC-MMFT is applied and compared to DTE, which achieves a mean CA of 36.5% and outperforms both ROD and LSA-MMFT. The superior classification performances achieved in this instance are attributed to the selected optimal combination of high beneficial domains after low-beneficial domains have been discarded at once, when EMC-MMFT is employed. In this case, high beneficial SDs are mainly selected to minimize the effect of NT and utilized to enhance the prediction rate across individual TDs. Hence, superior classification performances are observed when EMC-MMFT is employed, while inferior performances are attributed to the remaining non-related low-beneficial domains when a single low-beneficial domain is removed at a time, when ROD, DTE, and LSA-MMFT are employed.

These results further validate the fact that removing a single low-beneficial domain in each iteration poses severe implications for the classification performance across individual TDs, while selecting only the best combination of closely related high beneficial sources, denoted by superior classification performances, can significantly enhance the prediction rate across individual TDs.

5. Discussion

The classification performance when TL is employed is dependent on the transferability of features in the source domains, utilized to enhance the prediction rate in the target domain [37]. However, negative transfer has demonstrated to pose severe implications for the classification performance across individual target domains when TL is employed [12]. As such, sessional and inter-subject variations have proven to significantly contribute to the effect of NT, emerging as a result of the transfer of features from low-beneficial sources in the target domain [38]. However, a combination of both low-beneficial and high beneficial domains can positively and negatively affect classification performance, mainly as a result of different sources responding differently to other individual sources. In this case, low-beneficial sources can drastically deteriorate the prediction rate, while related high beneficial sources improve the prediction rate across individual target domains [39]. Hence, domain selection algorithms have been implemented to identify high beneficial sources with a high transferability, mainly to minimize the effect of NT and enhance the prediction rate across individual TDs when TL is employed after source selection.

In this paper, we investigated the impact of low-beneficial and high beneficial domains on TL performance, at the same time addressing the challenge of NT, significantly contributing to poor classification performances across individual TDs. To facilitate the investigation, four novel domain selection algorithms were introduced and compared to evaluate the performance of each algorithm when beneficial sources ware selected for transfer mapping. In this case, all the domain selection techniques (EMC-MMFT, ROD, DTE, and LSA-MMFT) adapted a two-class inter-session and inter-subject scenario to a multi-class inter-subject and inter-session scenario. The performances of EMC-MMFT, DTE, ROD, and LSA-MMFT were evaluated and compared under three inter-session and inter-subject EEG classification problems (namely, two-class, three-class, and four-class problems), mainly to validate the performance of each algorithm across individual TDs [11,12].

Consequently, when the EMC-MMFT framework was employed, only a combination of high beneficial domains was selected for transfer mapping, while low-beneficial domains were discarded to minimize the effect of NT, which can drastically deteriorate the prediction rate across individual TDs. A comparative performance evaluation illustrated that EMC-MMFT could significantly improve the classification performance across individual TDs compared to DTE, ROD, and LSA-MMFT. Consequently, all four algorithms achieved the highest CA of 100% for the two-class inter-session classification scenario. However, a significant decline in CA across individual TDs was observed when LSA-MMFT, DTE, and ROD were applied, while EMC-MMFT maintained superior classification performances across all the individual TDs, as depicted in Figure 5. Similar classification performances were observed for the two-class inter-subject classification scenario using the BCI competition IV-a dataset, whereby the highest CAs of 96% and 97% were recorded when LSA-MMFT, DTE, and ROD were applied, respectively, while EMC-MMFT achieved a superior CA of 98%. In this case, EMC-MMFT outperformed all the other three algorithms and maintained superior performances across all the individual TDs, as illustrated in Figure 8. Moreover, EMC-MMFT outperformed all three algorithms for the two-class inter-subject classification scenarios using our own recorded SSMVEP and MI subjects, with superior CAs of 95% and 92.5% being recorded when optimal combinations of high beneficial SDs were selected, as illustrated in Figures 6 and 7. Notably, the highest mean CA of 85.1% was recorded when DTE was employed, compared to ROD and LSA-MMFT; however, a superior mean CA of 94.2% was observed when EMC-MMFT was employed, based on the two-class inter-session scenario, while a superior mean CA of 93.9% was achieved for the two-class inter-subject classification scenario using the BCI competition IV-a dataset. EMC-MMFT further outperformed the other three domain selection algorithms, with superior CAs of 84.5% and 78.4% being recorded for both of the inter-subject scenarios utilizing our own recorded SSMVEP and MI subjects, respectively as depicted in Table 5.

Iterating through the domains while removing a single less-beneficial domain in each iteration demonstrated to crucially cause a negative impact on the classification performance when DTE and ROD were employed under the three-class scenario. Hence, mean CAs of 68.9% and 63.2% were recorded for DTE and ROD, respectively. Iterating through the domains utilizing LSA-MMFT to evaluate label similarities to identify the transferability

between domains also demonstrated to be less efficient in enhancing performance, with a mean CA of 64.1% being recorded. Therefore, selecting only a combination of high beneficial domains when EMC-MMFT was employed yielded superior performances across the individual TDs that translated to a mean CA of 84.1% for the three-class inter-session classification scenario, while a superior mean CA of 80% was noted for the three-class inter-subject classification scenario utilizing MI subjects acquired from the BCI competition IV-a dataset. Moreover, EMC-MMFT achieved the highest mean CAs of 72.8% and 64.3% for the three-class inter-subject classification scenarios using our own recorded SSMVEP and MI subjects, respectively.

The complexity of the four-class inter-subject EEG classification problem including the existence of NT proved to significantly affect classification performance when DTE, ROD, and LSA-MMFT were employed. Hence, inferior mean CAs of 57%, 61.2%, and 58.27% were recorded when DTE, ROD, and LSA-MMFT were employed to address the effect of NT based on the four-class problem. Notably, selecting only the optimal combination of highly beneficial SDs yielded remarkable performances regardless of the complexity of the four-class problem, with the highest mean CA of 75.06% being recorded when EMC-MMFT was employed, based on the four-class inter-session scenario, while a superior mean CA of 67.4% was achieved for the four-class inter-subject classification scenario using the BCI competition IV-a dataset, as depicted in Tables 5 and 6, respectively. EMC-MMFT further demonstrated its superiority and outperformed the other three domain selection algorithms, with superior mean CAs of 59.8% and 62.9% being recorded for the four-class inter-subject classification scenarios using our own recorded SSMVEP and MI subjects.

These results further validate that selecting only the best combination of closely related high beneficial sources can significantly minimize the effect of NT and, in turn, enhance the prediction rate across individual TDs. However, iterating through the domains and removing a single low-beneficial source domain at a time can still pose severe implications for the CA in the next iteration, since different domains respond differently to other domains.

6. Conclusions

In this paper, the effect of NT on the classification performance of transfer learning across individual target domains was investigated, whereby four domain selection algorithms were introduced to evaluate the effect of low- and high beneficial source domains on TL performance under three inter-session and inter-subject classification scenarios (two-class, three-class, and four-class problems). In this case, the EMC-MMFT framework was compared with three domain selection algorithms (LSA-MMFT, ROD, and DTE), mainly to determine the impact of NT on classification performance across individual TDs. Hence, the performance of each domain selection algorithm was evaluated based on the selection of high beneficial domains and the rejection of low-beneficial domains to determine the effect of NT on the performance. Therefore, discarding a single low-beneficial domain in each iteration from a pool of domains consisting of significant individual differences demonstrated to pose severe implications for classification performance. Hence, inferior classification performances were recorded across individual TDs when ROD, DTE, and LSA-MMFT were employed under the two-class and three-class problems. Moreover, the complexity of the four-class problem including the presence of low-beneficial domains in each iteration when ROD, DTE, and LSA-MMFT were employed to discard only a single low-beneficial domain per iteration demonstrated to significantly deteriorate the prediction rate across individual TDs. However, discarding all the low-beneficial SDs at once and selecting only the best combination of high beneficial SDs demonstrated to minimize the effect of NT and, at the same time, significantly enhance the prediction rate. Hence, superior classification performances were recorded across all nine individual TDs when EMC-MMFT was employed under the two-class and three-class problems. Moreover, the complexity of the four-class problem was demonstrated to cause a minimal impact on the classification performance when EMC-MMFT was employed to select only the optimal combination of high beneficial SDs, with superior performances being recorded across each individual TD compared with the results obtained with the ROD, DTE, and LSA-MMFT algorithms. This proves the study's hypothesis and indicates that selecting only high beneficial domains can significantly minimize the effect of NT and, in turn, enhance the performance of TL algorithms.

Author Contributions: Conceptualization, R.C.M.; methodology, R.C.M.; software, R.C.M. and S.D.; formal analysis, R.C.M.; investigation, R.C.M.; resources, R.C.M. and S.D.; data curation, R.C.M.; writing—original draft preparation, R.C.M.; writing—review and editing, C.T. and S.D.; visualization, R.C.M.; supervision, C.T., P.A.O. and S.D.; project administration, C.T. and P.A.O.; funding acquisition, C.T. and P.A.O. All authors have read and agreed to the published version of the manuscript.

Funding: This research is supported by the National Research Foundation of South Africa (grant numbers 145975 and 138429).

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The datasets used/or analyzed during the current study are publicly available online and from the corresponding author upon reasonable request.

Conflicts of Interest: The authors declare no conflicts of interest.

References

- 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]
- 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] [PubMed]
- Saha, S.; Ahmed, K.I.U.; Mostafa, R.; Hadjileontiadis, L.; Khandoker, A. Evidence of Variabilities in EEG Dynamics During Motor Imagery-Based Multiclass Brain–Computer Interface. *IEEE Trans. Neural Syst. Rehabil. Eng.* 2017, 26, 371–382. [CrossRef]
- 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]
- Zhu, L.; Yang, J.; Ding, W.; Zhu, J.; Xu, P.; Ying, N.; Zhang, J. Multi-Source Fusion Domain Adaptation Using Resting-State Knowledge for Motor Imagery Classification Tasks. *IEEE Sens. J.* 2021, 21, 21772–21781. [CrossRef]
- 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]
- Wu, D.; Xu, Y.; Lu, B.-L. Transfer Learning for EEG-Based Brain–Computer Interfaces: A Review of Progress Made Since 2016. IEEE Trans. Cogn. Dev. Syst. 2020, 14, 4–19. [CrossRef]
- 8. Zhang, W.; Deng, L.; Zhang, L.; Wu, D. A survey on negative transfer. IEEE CAA J. Autom. Sin. 2022, 10, 305–329. [CrossRef]
- 9. 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]
- Zhang, W.; Wu, D. Manifold Embedded Knowledge Transfer for Brain-Computer Interfaces. *IEEE Trans. Neural Syst. Rehabil. Eng.* 2020, 28, 1117–1127. [CrossRef]
- 11. 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]
- 12. Maswanganyi, R.C.; Tu, C.; Owolawi, P.A.; Du, S. Multi-Class Transfer Learning and Domain Selection for Cross-Subject EEG Classification. *Appl. Sci.* 2023, *13*, 5205. [CrossRef]
- 13. Gong, B.; Shi, Y.; Sha, F.; Grauman, K. Geodesic flow kernel for unsupervised domain adaptation. In Proceedings of the 2012 IEEE Conference on Computer Vision and Pattern Recognition, IEEE, Providence, RI, USA, 16–21 June 2012; pp. 2066–2073.
- 14. Jeng, P.-Y.; Wei, C.-S.; Jung, T.-P.; Wang, L.-C. Low-Dimensional Subject Representation-Based Transfer Learning in EEG Decoding. *IEEE J. Biomed. Health Inform.* **2020**, *25*, 1915–1925. [CrossRef]
- Liang, Z.; Zheng, Z.; Chen, W.; Wang, J.; Zhang, J.; Chen, J.; Chen, Z. Manifold Trial Selection to Reduce Negative Transfer in Motor Imagery-based Brain–Computer Interface. In Proceedings of the 2021 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), Prague, Czech Republic, 27 September–1 October 2021; pp. 4144–4149.
- 16. 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*, 3501111. [CrossRef]
- 17. 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]
- 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]

- 19. Rodrigues, P.L.C.; Jutten, C.; Congedo, M. Riemannian Procrustes Analysis: Transfer Learning for Brain–Computer Interfaces. *IEEE Trans. Biomed. Eng.* **2018**, *66*, 2390–2401. [CrossRef] [PubMed]
- 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]
- 21. 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]
- 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]
- 23. Duan, T.; Chauhan, M.; Shaikh, M.A.; Chu, J.; Srihari, S. Ultra efficient transfer learning with meta update for cross subject EEG classification. *arXiv* 2020, arXiv:2003.06113.
- Gao, Y.; Liu, Y.; She, Q.; Zhang, J. Domain Adaptive Algorithm Based on Multi-Manifold Embedded Distributed Alignment for Brain-Computer Interfaces. *IEEE J. Biomed. Health Inform.* 2022, 27, 296–307. [CrossRef]
- 25. 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*, 7000104. [CrossRef]
- 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.
- 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]
- 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, 2508215. [CrossRef]
- 29. 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]
- 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]
- Kim, D.-K.; Kim, Y.-T.; Jung, H.-R.; Kim, H. Sequential Transfer Learning via Segment After Cue Enhances the Motor Imagerybased 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; pp. 1–5.
- Samek, W.; Kawanabe, M.; Muller, K.-R. Divergence-Based Framework for Common Spatial Patterns Algorithms. *IEEE Rev. Biomed. Eng.* 2013, 7, 50–72. [CrossRef] [PubMed]
- Wei, M.; Yang, R.; Huang, M. Motor imagery EEG signal classification based on deep transfer learning. In Proceedings of the 2021 IEEE 34th International Symposium on Computer-Based Medical Systems (CBMS), IEEE, Aveiro, Portugal, 7–9 June 2021; pp. 85–90.
- Shajil, N.; Sasikala, M.; Arunnagiri, A. Faisal, Inter-Subject Deep Transfer Learning for Motor Imagery EEG Decoding. In Proceedings of the 2021 10th International IEEE/EMBS Conference on Neural Engineering (NER), IEEE, Virtual, Italy, 4–6 May 2021; pp. 21–24.
- 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; pp. 1–4.
- 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] [PubMed]
- 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]
- Ma, Y.; Zhao, W.; Meng, M.; Zhang, Q.; She, Q.; Zhang, J. Cross-Subject Emotion Recognition Based on Domain Similarity of EEG Signal Transfer Learning. *IEEE Trans. Neural Syst. Rehabil. Eng.* 2023, 31, 936–943. [CrossRef] [PubMed]
- Jeon, E.; Ko, W.; Suk, H.-I. Domain Adaptation with Source Selection for Motor-Imagery based BCI. In Proceedings of the 2019 7th International Winter Conference on Brain-Computer Interface (BCI), IEEE, Gangwon, Republic of Korea, 18–20 February 2019; pp. 1–4.

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.