

Article

Mental Workload Classification Method Based on EEG Cross-Session Subspace Alignment

Hongquan Qu ¹, Mengyu Zhang ¹  and Liping Pang ^{2,*}

¹ School of Information Science and Technology, North China University of Technology, Beijing 100144, China; qhqphd@ncut.edu.cn (H.Q.); zhangmeng_yu@foxmail.com (M.Z.)

² School of Aeronautic Science and Engineering, Beihang University, Beijing 100191, China

* Correspondence: pangliping@buaa.edu.cn; Tel.: +86-010-8231-3186

Abstract: Electroencephalogram (EEG) signals are sensitive to the level of Mental Workload (MW). However, the random non-stationarity of EEG signals will lead to low accuracy and a poor generalization ability for cross-session MW classification. To solve this problem of the different marginal distribution of EEG signals in different time periods, an MW classification method based on EEG Cross-Session Subspace Alignment (CSSA) is presented to identify the level of MW induced in visual manipulation tasks. The Independent Component Analysis (ICA) method is used to obtain the Independent Components (ICs) of labeled and unlabeled EEG signals. The energy features of ICs are extracted as source domains and target domains, respectively. The marginal distributions of source subspace base vectors are aligned with the target subspace base vectors based on the linear mapping. The Kullback–Leibler (KL) divergences between the two domains are calculated to select approximately similar transformed base vectors of source subspace. The energy features in all selected vectors are trained to build a new classifier using the Support Vector Machine (SVM). Then it can realize MW classification using the cross-session EEG signals, and has good classification accuracy.

Keywords: Mental Workload; EEG; domain adaptation; KL divergence; subspace alignment

MSC: 68T01



Citation: Qu, H.; Zhang, M.; Pang, L. Mental Workload Classification Method Based on EEG Cross-Session Subspace Alignment. *Mathematics* **2022**, *10*, 1875. <https://doi.org/10.3390/math10111875>

Academic Editor: Konstantin Kozlov

Received: 11 April 2022

Accepted: 27 May 2022

Published: 30 May 2022

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



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

1. Introduction

The Mental Workload (MW) represents the amount of brain activity in a unit of time. It is closely related to the operator's professional knowledge, personality, task type, and physiological variables [1]. Excessive MW can lead to fatigue, distraction, and reduced efficiency, which leads to errors in analysis and decision-making tasks. However, too low MW also leads to a waste of human resources [2]. Hence, the MW level is a key design factor in human–machine collaboration tasks. It is of great significance to identify the level of MW and evaluate operator performance.

The existing methods of MW measurement include subjective measurement, performance measurement, and psychophysiological measurement [3]. Subjective measurement involves some subjective feeling investigations, such as the Subjective Workload Assessment Technique (SWAT), NASA Task Load Index (NASA-TLX), and Workload Profile (WP) [4]. Performance measurement is to measure the operating performance of the subject based on tasks [5]. The indicators of physiological measurement usually include electroencephalogram (EEG), electromyogram (EMG), electrocardiogram (ECG), and electrooculogram (EOG) [6,7].

EEG data can reflect the electrical activity of the brain [8,9]. It is possible to quantify brain responses to variations in task demand [10]. Many studies showed that some features of EEG signals were responsive to the MW change [11–14]. Most research divided EEG signals into five rhythms, namely Alpha, Beta, Theta, Delta, and Gamma [15,16]. Some studies revealed that the energy features of EEG signals can classify the MW level [17,18].

In recent years, many EEG classification methods were developed based on machine learning and deep learning, such as Logistic Regression (LR) [19], Linear Discriminant Analysis (LDA) [20], Support Vector Machine (SVM) [20,21], and Neural Network (NN) [22]. Qu et al., proposed an MW classification method based on EEG independent component features with an SVM classifier and compared the SVM classifier with others [23]. Pang et al., used a Stochastic Configuration Network (SCN) for MW classification [24]. The obtained results showed that most of these traditional machine learning methods could achieve good classification results when the trained and classified data were collected in a test period with a short time interval.

EEG signals are non-stationary [23,24] and usually change due to personnel fatigue, environmental change, electrode–skin impedance, and instrument change [25]. Mühl et al., mentioned the influence of environment and affective context on the classification results of MW [26]. Non-stationary EEG leads to a decrease in generalization capability and classification accuracy if the classifier only used the data from one previous session as its training set [27]. Therefore, most statistical classification models need to be retrained from scratch using newly collected data because of the distribution change of EEG data [28].

Domain adaptation methods have been utilized in some studies [29,30], including Transfer Component Analysis (TCA) [31], Subspace Alignment (SA) [32], and Geodesic Flow Kernel (GFK) [33]. Some studies applied domain adaptation to emotion classification using EEG data [34,35] and achieved high accuracy. Zheng et al., applied the TCA method to cross-subject emotion recognition [36]. Chai et al., used the Multi-Subject Subspace Alignment (MSSA) method for multi-subject emotion recognition based on the SA method [37]. Sciaraffa N. et al., highlighted the advantage of the use of calibration in cross-subject MW classification [38]. Zhang et al., used the kernel spectral regression and transfer learning techniques to improve the accuracy of cross-subject MW classification [39]. Liu et al., found a workload decoder using data from other subjects as an approach to improve workload classification performance [40]. Yin et al., used an adaptive deep learning model to deal with the cross-session EEG features [41].

Different from the above research, a mental workload classification method based on EEG Cross-Session Subspace Alignment (CSSA) is proposed in this paper. This method tries to solve the tricky problem of cross-session MW classification and improve its accuracy. In this proposed method, the SA method is used to obtain the domain-invariant feature space with different sessions, and the method not only uses a single source domain as a training set but tries to make full use of all data in different sessions and allows adding new data into the training data set. The trained cross-session MW classifier can improve classification accuracy because it builds a personalized MW classification model using the data from multiple sessions including both new and existing sessions.

The rest of this paper is organized as follows. The first section introduces the experiment about collecting EEG data and extracting EEG features. The second section introduces the principle of the presented CSSA method. The third section applies the CSSA method to classify MW and compares it with the existing methods. The fourth section involves the conclusion of this study.

2. Materials and Methods

2.1. Materials

2.1.1. Experiments

A 32-channel Neuroscan Neuamps (Neuroscan, Synamps2, Scan4.3, EI Paso, TX, USA) system was used to record all EEG signals. EEG raw data were digitized at a rate of 1024 samples per second. Band-pass filter settings were 0.1–200 Hz.

The EEG signals were collected from 13 subjects with an engineering knowledge background from Beihang University (aged between 22 and 25 years, with 2 females and 11 males). The Multi-Attribute Task Battery (MATB-II, National Aeronautics and Space Administration (NASA), Washington DC, MD, USA) was used as the platform for

the experiment. It is composed of 4 sub-tasks, including system monitoring, tracking, scheduling, and resource management.

The details of the four sub-tasks are shown in Table 1. The system monitoring task is presented in the upper left window of the display. The demands of monitoring gauges and warning lights were simulated here, and the demands of manual control were simulated by the tracking task. This task is shown in the upper middle window. The communication task presented pre-recorded auditory messages to the operator at selected intervals during the simulation, and the demands of fuel management were simulated by the resource management task. The four regions corresponded to visual and operational missions. During the experiment, two levels of MW were set; these were Low Mental Workload (LMW) and High Mental Workload (HMW).

Table 1. Frequency of occurrence of four tasks.

Task Title	Task Content	LMW	HMW
System Monitoring	Monitoring the scales of F1-F4 in Area 1 and responding with a mouse when the scales are not around the center.	1	24
Tracking	In Area 2, the target is kept in the grid center by a joystick in MANUAL mode and no action is required in AUTO mode.	1	24
Scheduling	Monitoring scheduling bar in Area 3 and responding to the activated communication with keyboard immediately.	1	24
Resource Management	Clicking the corresponding oil pump with the mouse when failure occurs.	1	24

The EEG data of 13 subjects were collected for 7 days. To avoid the learning effect, the participants were trained for 12 days before the experiment to ensure their familiarity with different tasks in the MATB-II task [42]. They were required to complete four sub-tasks of MATB-II under different MW levels. The order of experiments was designed with a Latin square design method to avoid the order effect. Three minutes of EEG data were collected as baseline data before the MW experiment on every subject.

2.1.2. Data Preprocess

Firstly, the average values of M1 and M2 were used as the reference for data processing [43]. M1 and M2 were located at the mastoid position behind the left and right ears, respectively. Then, the 30-channel EEG signals were filtered with a 1–30 Hz Finite Impulse Response (FIR) band-pass filter, where the filter order was 3300 [44].

Two selection criteria for subject EEG data were used in our study. To ensure the reliability of data analysis, the accuracy of the MATB-II task operation on the subject should be greater than 90%. To reduce the interference of the artifact signals, the artifacts of the subject EEG data can not exceed 40% [45]. Finally, the EEG data of 10 participants (sub1-sub10, aged between 22 and 25 years, with 2 females and 8 males) were selected for analysis.

For the 12-min EEG data, only the middle 10 min of EEG data from each MW experiment were used. They were divided into 300 data segments, where the length of a segment was 2 s. The segments were labeled according to the different MW levels induced by the MATB-II task. There were 600 data segments for two MW tasks.

2.2. Methods

2.2.1. Blind Source Separation of EEG Signals

The EEG signals are usually collected by non-invasive measurement technology through electrode sensors placed on the scalp. As shown in Figure 1, S_i ($i = 1, 2, \dots, n$) represents the signals generated by neural activities in the brain. Due to the mixed effect, the measured mixed signals, X_i ($i = 1, 2, \dots, m$), are different from the signals generated by neurons, S_i . This phenomenon will have a negative effect on feature extraction. This mixed effect depends on the distances from neuron sources to electrode sensors,

a_{ij} ($i = 1, 2, \dots, m; j = 1, 2, \dots, n$) [46]. The problem of separating mixed EEG signals is similar to the problem of Blind Source Separation (BSS), which is used to analyze mixing speech signals [47]. Independent Component Analysis (ICA), based on information independence, is used to separate the original signals from their mixtures [23].



Figure 1. Signal mixing process of the electroencephalogram (EEG).

The m -dimensional observation vectors can be expressed as a linear combination of n -dimensional ICs, s_i ($i = 1, 2, \dots, n$), as given in Equation (1) [48].

$$x_j = a_{j1}s_1 + a_{j2}s_2 + \dots + a_{jn}s_n, j = 1, 2, 3, \dots, m \quad (1)$$

This can be expressed as:

$$X = AS \quad (2)$$

where $S = [s_1, s_2, \dots, s_n]^T$ is a set of ICs and the number of ICs is 30, which is the same as the number of channels in the EEG signals. $X = [x_1, x_2, \dots, x_m]^T$ is a vector of the observed data and A is a mixing matrix.

2.2.2. Feature Extraction

The Power Spectral Densities (PSDs) at different frequencies are extracted as features of EEG signals in the frequency domain. Fourier transform in Equation (3) is performed on each data segment to obtain $F_s(n)$.

$$F_s(n) = \sum_{n=0}^{N-1} S e^{\frac{-2\pi j}{N} n} \quad (3)$$

where $F_s(n)$ is the frequency spectrum of EEG signals after the Fourier transform; S is the ICs of EEG signals and shows the Fourier transform of the signals; N is the number of signal samples; n is 0 to $N - 1$.

The corresponding PSDs of the EEG ICs are calculated with Equation (4).

$$p_s(n) = \frac{F_s(n)F_s^*(n)}{N} \quad (4)$$

where $p_s(n)$ is the PSD of the EEG ICs and $F_s^*(n)$ is the conjugate of $F_s(n)$.

Many studies have shown that the human mental state is related to four rhythms [15]. According to the frequency band distribution, EEG signals can be divided into δ (1~4 Hz), θ (4~8 Hz), α (8~14 Hz), and β (14~30 Hz) [23,24].

$$\begin{aligned}
 E_{S,\delta} &= \sum_{freq=1}^4 P_{s,freq} \\
 E_{S,\theta} &= \sum_{freq=4}^8 P_{s,freq} \\
 E_{S,\alpha} &= \sum_{freq=18}^{14} P_{s,freq} \\
 E_{S,\beta} &= \sum_{freq=14}^{30} P_{s,freq}
 \end{aligned} \tag{5}$$

where $P_{s,freq}$ refers to the power spectral density at a certain frequency and $E_{S,\delta}$, $E_{S,\theta}$, $E_{S,\alpha}$, and $E_{S,\beta}$ are the four energy features of the EEG ICs.

The absolute power spectrums are normalized to obtain the relative power spectrums.

$$E_{s,\delta}^* = \frac{E_{S,\delta}}{E_{all}}, \quad E_{s,\theta}^* = \frac{E_{S,\theta}}{E_{all}}, \quad E_{s,\alpha}^* = \frac{E_{S,\alpha}}{E_{all}}, \quad E_{s,\beta}^* = \frac{E_{S,\beta}}{E_{all}}, \quad E_{all} = \sum_{freq=1}^{30} P_{s,freq} \tag{6}$$

where E_{all} refers to the power spectral density at 1–30 Hz and $E_{s,\delta}^*$, $E_{s,\theta}^*$, $E_{s,\alpha}^*$, and $E_{s,\beta}^*$ are the four relative energy features of the EEG ICs.

2.2.3. Non-Stationarity of EEG Signals

EEG signals are sensitive to several influence factors, such as personnel fatigue, electrode placement, and the impedance value of the EEG cap, etc. All these lead to the non-stationarity characteristic of EEG signals. In order to explain this clearly, two groups of EEG signals are shown in Figure 2. These EEG signals were obtained from the same subject in MW experiments, but their time interval was 24 h. The three principal components of EEG ICs, PC1, PC2, and PC3 with 23.8% variance lost, are obtained with the Principal Components Analysis (PCA) method. In Figure 2a, the principal components are represented in four colors. The data marked in the colors red and blue are the first-day data, while the data marked in yellow and green are the second-day data. The LMW and HMW data are marked with “+” and “×”, separately. The top and left views of Figure 2a are shown in Figure 2b. Even for the same subject, the energy features of the EEG are separated after 24 h, that is their marginal distributions have changed, but the conditional distributions in each domain are almost the same.

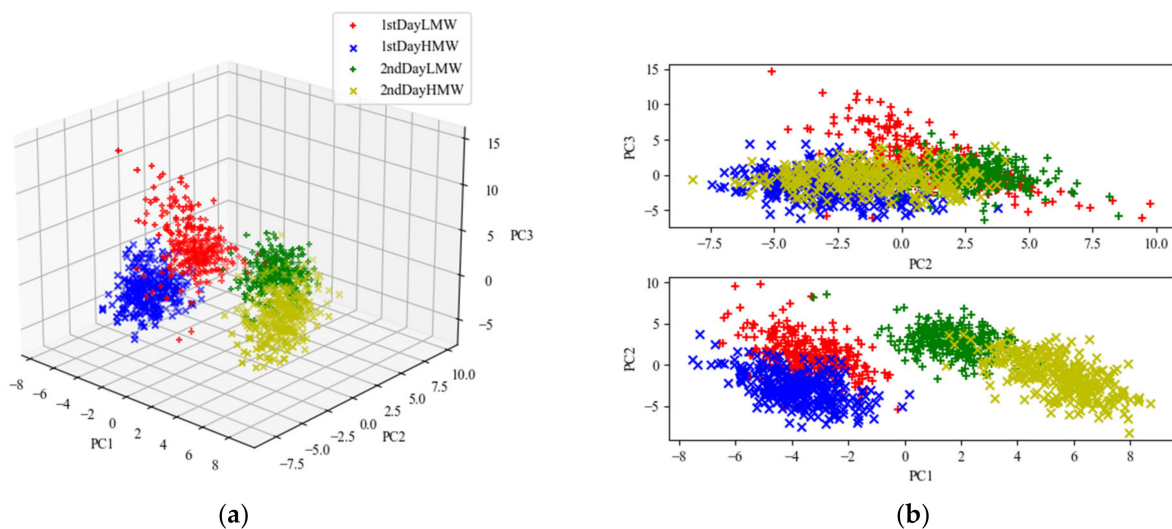


Figure 2. (a) Energy features of EEG signals in different sessions; (b) Top view and left view of the energy features.

The EEG data from the 1st day and the 2nd day can be used as the source and target domain data, respectively. The traditional machine learning method simply takes the source domain data as training samples to construct a classifier. Due to the distribution difference between the source domain and the target domain, the generalization ability of the traditional method is insufficient.

2.2.4. SA Method

As the middle 10 min of EEG data in each MW task were chosen and divided into 300 data segments, there were 600 segments and 120-dimensional energy features for two MW tasks in each segment. Let $X \in x$ be the recorded EEG of a sample (X, y) and $x = \mathbb{R}^{C \times D}$, where C is the number of channels and D is the number of samples of time series. In this paper, $C = 30$ and $D = 300$. $y \in Y$ represents the corresponding MW label. Let $P(X)$ be the marginal probability distribution of X and $P(Y|X)$ be the conditional probability distribution. The source domain is represented as $X_{Si} = (x_{Si,1}, y_{Si,1}), \dots, (x_{Si,j}, y_{Si,j}), \dots, (x_{Si,D_S}, y_{Si,D_S})$, $i = 1, 2, \dots, N$, $j = 1, 2, \dots, D_S$. Similarly, the target domain is represented as $X_T = (x_{T,1}, y_{T,1}), \dots, (x_{T,j}, y_{T,j}), \dots, (x_{T,D_T}, y_{T,D_T})$, $j = 1, 2, \dots, D_T$, where $(X_{Si}, X_T) \in X$ is the normalized relative power spectral density.

As shown in Figure 3, although the source domain data and target domain data are in D -dimensional space, their marginal distributions are different, while their conditional distributions are almost the same. Here, assume that $P(X_S) \neq P(X_T)$ but $P(Y_S|X_S) = P(Y_T|X_T)$, which means there are differences between different domains. The purpose of SA is to make $P(\phi(X_S)) = P(X_T)$ with the linear mapping ϕ .

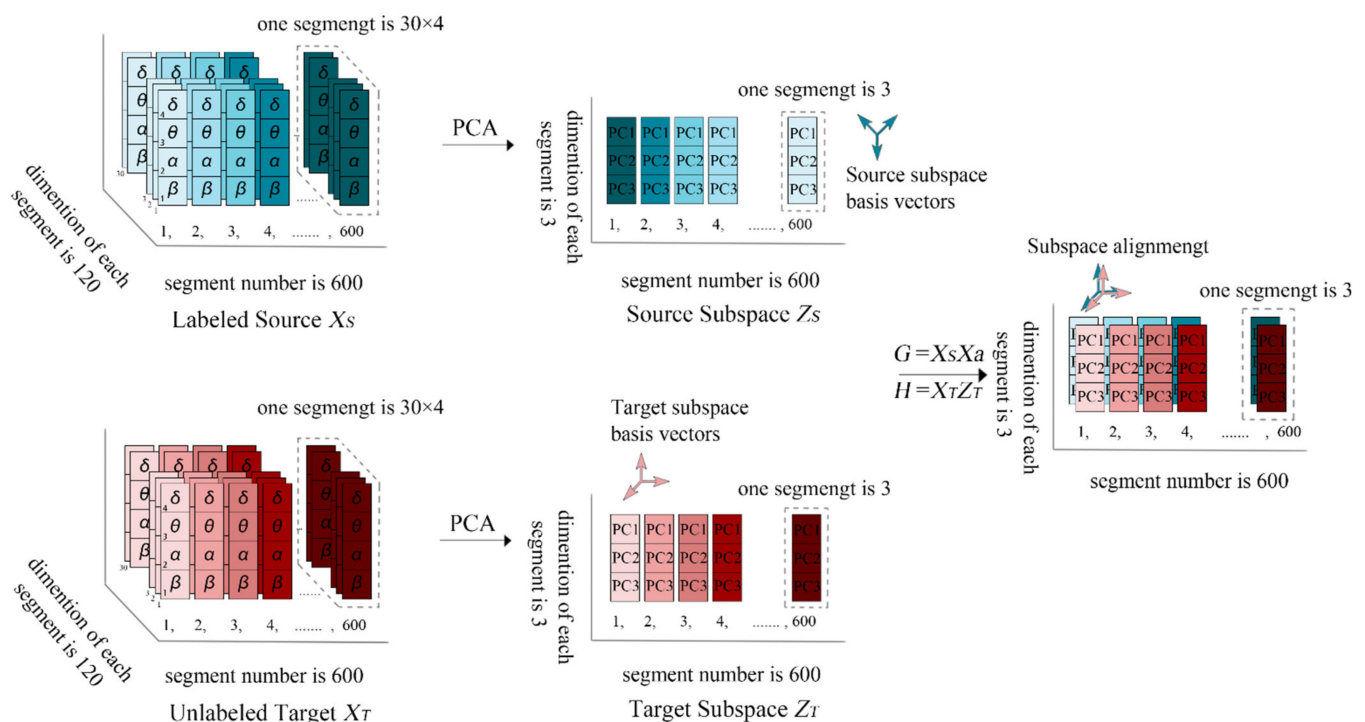


Figure 3. Subspace Alignment (SA) method.

According to the PCA method, the eigenvectors corresponding to the first l maximum eigenvalues can be found in each domain. Here, l represents the number of principal components and $l = 3$ in this paper. These eigenvectors are defined as the source domain subspace and target domain subspace, which are represented as Z_S and Z_T , respectively. Assume that there is a linear relationship between Z_S and Z_T , since Z_S and Z_T only depended on X_S and X_T , separately, the non-stationarity of EEG did not affect the method. Then we can find a linear mapping to align the source subspace with the target subspace.

We use the transformation matrix, S , from Z_S to Z_T to align the basis vectors. S transforms the source subspace into the target subspace by aligning the source and target basis vectors. S is achieved by minimizing the following Bregman matrix divergence:

$$F(S) = \min \|Z_S S - Z_T\|_F^2 \quad (7)$$

where $\|\cdot\|_F^2$ is the F-norm.

Due to the orthogonality of the F-norm, Equation (7) can be rewritten as:

$$\min \|Z_S^T Z_S S - Z_S^T Z_T\|_F^2 = \min \|S - Z_S^T Z_T\|_F^2 \quad (8)$$

The minimum value of Equation (8) is obtained when $S^* = Z_S^T Z_T$. For the coordinate system of the source domain, X_a can be written as:

$$Xa = Z_S Z_S^T Z_T \quad (9)$$

With Equation (10), the representation of the source domain and the target domain in the new coordinate system can be obtained after the SA method. Currently, the target domain and the source domain have the most similar marginal distribution. Because we mapped the source subspace to the target one, Z_T could be called the domain-invariant feature spaces of the source domain and target domain. G and H represent the domain-invariant features in the subspace.

$$\begin{aligned} G &= X_S X_a \\ H &= X_T Z_T \end{aligned} \quad (10)$$

2.2.5. KL Divergence

The KL divergence, also known as relative entropy, is an asymmetric measure of the difference between two probability distributions. When two distributions are the same, the KL divergence is zero. When the difference between the two distributions increases, the KL divergence also increases. The formula is given in Equation (11).

$$D_{KL}(p||q) = \sum_{i=1}^N p(x_i) \log\left(\frac{p(x_i)}{q(x_i)}\right) \quad (11)$$

where p and q represent different distributions and $p(x_i)$ and $q(x_i)$ represent the probability density functions of two distributions, respectively.

For the EEG data used in this paper, $p(x_i)$ and $q(x_i)$ represent the probability density function of the source domain data G and the transferred target domain data H . The purpose of SA is to make $P(\phi(X_S)) = P(X_T)$ through a linear mapping ϕ . By calculating the KL divergence of the source domain and the target domain, similar transformed source subspace base vectors can be selected.

2.2.6. MW Classification Method Based on CSSA

To improve the MW classification performance for different time sessions, an MW classification method based on the CSSA is proposed. The proposed method consists of two parts, as shown in Figure 4.

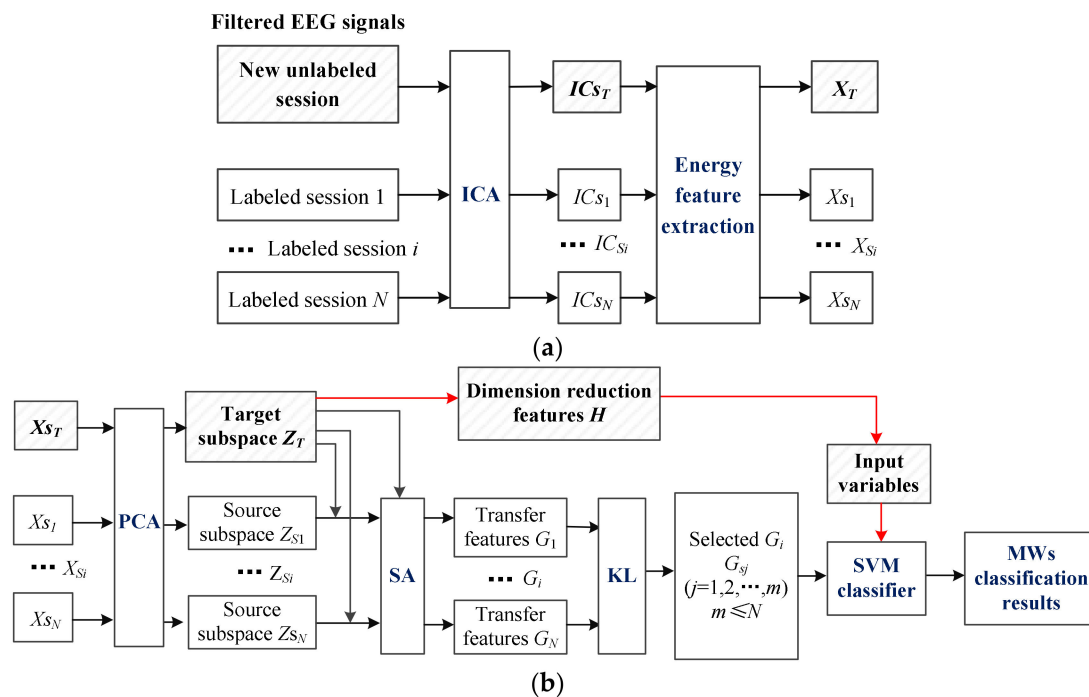


Figure 4. SA method; (a) Independent Components (ICs) of EEG and energy feature extraction; (b) Cross-Session Subspace Alignment (CSSA) method.

As shown in Figure 4a, the first step is to calculate the ICs of the EEG and extract their energy features. The filtered labeled EEG signals in N different time intervals or sessions form N labeled sessions, labeled session i , $i = 1, 2, \dots, N$. New filtered unlabeled EEG signals form a new unlabeled session. Their independent components, IC_{S_i} and IC_{S_T} , can be obtained using the ICA. Then, their energy features, X_{S_i} and X_{S_T} , can be extracted.

As shown in Figure 4b, the second part is the CSSA method. The source and target subspaces, Z_{S_i} and Z_{S_T} , are built using the PCA method. The transfer features of source subspaces, G_i , are extracted using the SA method. Meanwhile, the dimension-reduced features of the target subspace, H , are extracted by multiplying Z_T and X_T . The KL divergences of G_i marginal distributions are calculated in order to select approximately similar transformed source subspace base vectors, G_{S_j} , $j = 1, 2, \dots, m$, $m \leq N$. Then, the energy features in the selected vectors are trained to build a new classifier.

The SVM is a classical statistical classification algorithm. It is very effective in dealing with classification and regression problems [23]. In this paper, a linear SVM is finally used to classify the MWs in the target domains, H , and 5-fold cross-validation is used to reduce overfitting and ensure the robustness of the model. The grid search method is adopted to obtain the optimal model parameters. The penalty coefficient, C , indicates the tolerance of the error, and its search space is $[0.001, 0.01, 1, 10, 30]$.

3. Results

3.1. MW Classification Comparison with Different Methods

The MW classification accuracies of different methods are compared with those of different methods in this section.

In our method, the CSSA can adapt the energy features of ICs in a source subspace to its target subspace through linear mapping. In this way, the source domain subspace is aligned with the target domain subspace. Then the SVM classifier is used to obtain the MW classification results.

The CSSA method mentioned in Section 2.2.6 is named Method 1, and the compared methods are named Method 2 to Method 4, respectively. The SA method mentioned in Section 2.2.4 is named Method 2 [32]. The TCA method is named Method 3, as shown

in Figure 5. The PCA method is named Method 4 [31,49,50], as shown in Figure 6. For Method 3 and Method 4, the EEG signals are preprocessed in the same way, as shown in Figure 4a.

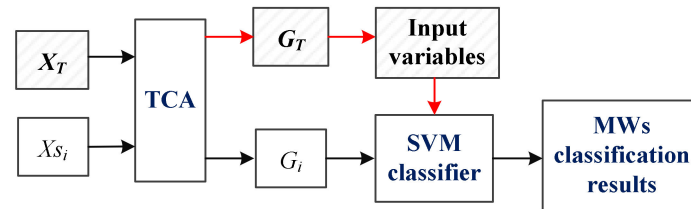


Figure 5. Implementation of Method 3.

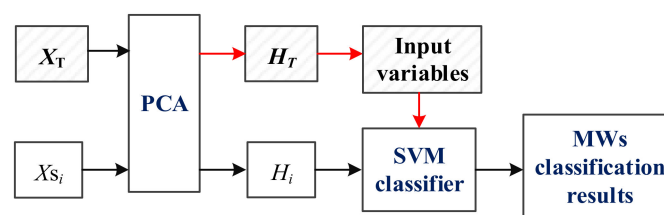


Figure 6. Implementation of Method 4.

Here, G_T and G_i represent the transfer of features from X_T and X_{S_i} , respectively.

Method 3 assumes that ϕ is the feature map induced by a universal kernel, and the distance of two domains can be obtained with Equation (12), where H is a universal RKHS [51], and D is the number of samples.

$$Dist(X_S, X_T) = \left\| \frac{1}{D_S} \sum_{i=1}^{D_S} \phi(x_{si}) - \frac{1}{D_T} \sum_{i=1}^{D_T} \phi(x_T) \right\|_{\mathcal{H}}^2 \quad (12)$$

Method 3 used the kernel learning:

$$\begin{aligned} \min W \operatorname{tr}(W^T W) + \mu \operatorname{tr}(W^T K L K W) \\ \text{s.t. } W^T K H K W = I \end{aligned} \quad (13)$$

where μ is a trade-off parameter and I is an identity matrix. K and L can be calculated as follows:

$$K = [X_S^T, X_T^T] \quad (14)$$

$$L_{ij} = \begin{cases} \frac{1}{D_S^2}, & xi, xj \in X_S \\ \frac{1}{D_T^2}, & xi, xj \in X_T \\ -\frac{1}{D_S D_T}, & \text{other} \end{cases}$$

The PCA method can extract the main feature components of data for the dimensionality reduction of high-dimensional data [52]. In Method 4, assuming that there is an n -dimensional data sample X with D samples, X can be decomposed into [53]:

$$X = U \Sigma V^T \quad (15)$$

where $U U^T = V V^T = 1$, and Σ is a diagonal matrix.

Divide Σ into r columns, named Σ_r . H can be calculated by Equation (16):

$$H = U \Sigma_r \quad (16)$$

Here, the calculation method used for X_T and X_{S_i} is the same as in Figure 4a. The PCA method is directly used to obtain X_T and X_{S_i} . H_T and H_i represent the reduced dimension features from X_T and X_{S_i} , respectively.

The EEG signals of the 10 subjects over seven days were obtained and preprocessed. Session 0 represents the EEG data obtained on the first day, and the data from this session were only used as training data. Session 1 to Session 6 represents the corresponding EEG data from the second day to the seventh day.

It should be noted that in Methods 2 to 4, the classified EEG data are the ones from the current day and the training data are the ones from the day before in this section. For example, if the EEG data from the fourth day were classified to determine their MW level, then the data from the third day would be used as the training data. In Method 1, the number of source domains corresponds to the days. For example, the fourth day has three source domains, in which the labeled EEG data are those from the first three days.

The classification accuracies of the four methods for the 10 subjects are compared in Figure 7. For the same time and data, the accuracy of Method 1 ranged from 82.0% to 88.2%, and the average classification accuracy was 84.7%. The accuracy of Method 2 ranged from 74.8% to 88.2%, and the average classification accuracy was 80.2%. The accuracy of Method 3 ranged from 66.0% to 87.1%, and the average accuracy was 74.0%. The accuracy of Method 4 ranged from 50.2% to 63.3%, and the average classification accuracy was 52.9%.

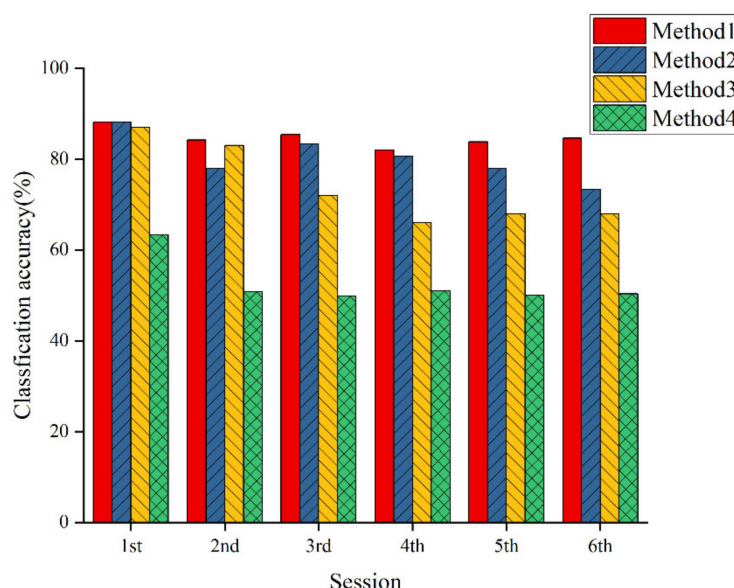


Figure 7. Classification accuracy comparison.

For Session 2, Method 2 faced the problem that its accuracy decreased due to the orthogonality of the source domain and target domain for Subject 3. Fortunately, this problem could be solved by Method 1 because the number of source domains gradually increased over the days.

Compared with Method 2, the accuracy of Method 1 was improved by 1.4% to 11.3%, for an average of 4.5%. Compared with Method 3, the accuracy of Method 1 was improved by 1.2% to 16.6%, for an average of 10.7%. Compared with Method 4, the accuracy of Method 1 was improved by 23% to 31.5%, for an average of 27.3%.

The average classification accuracies of Method 1 and Method 2 for 10 subjects were higher than those of Method 3. This is because Method 3 only used shared features. Method 1 and Method 2 both used relevant features. In addition, the stability of Method 1 was better than that of the others. Therefore, we chose to use Method 1 to classify the MWs across time.

The Friedman test based on algorithmic sorting was used to prove whether these methods perform equally. The data obtained in each session were considered as a data set,

D_i ($i = 1, 2, \dots, 6$). Each dataset contained data for Subject 1 to Subject 10. In comparing k algorithms on N data sets, r_i represents the average order value of the i^{th} algorithm, while τ_{χ^2} and τ_F were calculated from Equations (17) and (18), respectively.

$$\tau_{\chi^2} = \frac{12N}{k(k+1)} \left(\sum_{i=1}^k r_i^2 - \frac{k(k+1)^2}{4} \right) \quad (17)$$

$$\tau_F = \frac{(N-1)\tau_{\chi^2}}{N(k-1) - \tau_{\chi^2}} \quad (18)$$

When $\alpha = 0.05$, $K = 4$, and $N = 7$, the critical value of this test is 3.29. According to the above equations, $\tau_F = 17$. The assumption that all algorithms perform equally can therefore be rejected.

3.2. MW Classification Comparisons Using Different Source Domains

As CSSA only changes the mapping of the source domain, it can transform multiple source domains into the same subspace to minimize the difference between all source domains and the target domain. In this section, the impact of the multi-source domains on the classification performance will be discussed using only Method 1.

Six classification cases were used, and they contained 1 to 6 source domains. The number of source domains corresponded to the case number. For example, Case 3 had three source domains, in which the labeled EEG data were those obtained on the first three days.

The average classification accuracies of 10 subjects according to the number of source domains are shown in Figure 8. The results show that the classification accuracy was constantly improved as the number of source domains increased.

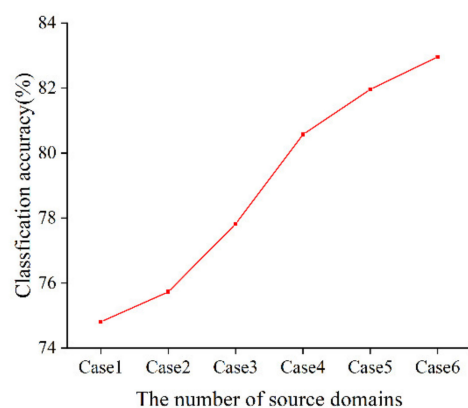


Figure 8. Impact of the number of source domains on the classification accuracy.

4. Discussion

In the traditional MW classification method, the classifier is trained only using the existing given EEG data, and then identifies the level of MW for unlabeled data. However, the distribution of EEG signals usually changes in different sessions due to the EEG signals being non-stationary. This may reduce the generalization ability of the classifier and lead to low classification accuracy. Hence, the above methods cannot be effective for cross-session classification when the distribution of EEG data changes greatly, which might lead to low classification accuracy.

The TCA method maps the source domain and target domain data into a common space, and the number of its source domains is fixed. The SA method maps the source domain to the target domain, but it is difficult to find the domain-invariant feature when there is an orthogonal relationship between the source domain and the target domain. The CSSA method proposed in this paper can handle this problem because it can make full use of existing and new data. With the increase in the number of source domains, the accuracy

of the classifier might be improved and become stable when the number of source domains is large enough.

5. Conclusions

In this study, we propose an MW classification method based on EEG cross-session subspace alignment. In this method, the source and target subspaces are obtained separately using the energy features extracted from the ICs of filtered EEG signals. The source subspace base vectors are aligned with the target subspace coordinate system base vector by linear mapping. The source domains are optimized by KL divergence. This study proposes a method to solve the problem of the marginal distribution of EEG signals being different in the same space due to their random non-stationary characteristics. The method may be able to improve cross-session generalization ability and classification accuracy.

Based on the analysis and discussion of the results, it can be concluded that:

- (1) In order to study the impact of a migration algorithm on classification performance, Method 1 to Method 4 were discussed. Compared with Method 2, Method 1 with CSSA could improve the MW classification accuracy from 80.2% to 84.7%. Compared with Method 3, Method 1 could improve the classification accuracy by 6%. Compared with Method 4, Method 1 could improve the classification accuracy from 52.9% to 80.2%.
- (2) The number of source domains also influences the classification performance. Datasets with 1 to 6 source domains were studied using Method 1. The classification accuracy of Method 1 could be improved from 74.8% for one source domain to 82.9% for six source domains. The results show that the classification accuracy can be improved by increasing the number of source domains.

Author Contributions: Conceptualization: H.Q.; methodology: H.Q.; software: M.Z.; validation: M.Z. and H.Q.; formal analysis: H.Q.; investigation: M.Z.; resources: L.P.; data curation: M.Z.; writing original draft preparation: M.Z.; writing—review and editing: M.Z. and L.P.; visualization: M.Z.; supervision: H.Q.; project administration: H.Q.; funding acquisition: L.P. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: The study was conducted in accordance with the Declaration of Helsinki, and approved by the Institutional Biological and Medical Ethics Committee, Beihang University (protocol code BM201900078, June 2019).

Informed Consent Statement: Informed consent was obtained from all subjects involved in the study.

Data Availability Statement: The raw data supporting the conclusions of this article are available on request from the corresponding author.

Acknowledgments: The authors are grateful to the subjects for their contributions to the experiment.

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

References

1. Hao, G.U.; Zhong, Y. Mental Workload Assessment Based on EEG and A Hybrid Ensemble Classifier. *Softw. Guide*. **2019**, *18*, 1–4.
2. Salomao, T.; Alberto, L. Operator Functional State Modelling and Adaptive Control of Automation in Human-Machine Systems. Ph.D. Thesis, University of Sheffield, Sheffield, UK, 2016.
3. Thea, R. Dual Frequency Head Maps: A New Method for Indexing Mental Workload Continuously during Execution of Cognitive Tasks. *Front. Physiol.* **2017**, *8*, 1019. [[CrossRef](#)]
4. Rubio, S.; Díaz, E.; Martín, J.; Puente, J.M. Evaluation of Subjective Mental Workload: A Comparison of SWAT, NASA-TLX, and Workload Profile Methods. *Appl. Psychol.* **2004**, *53*, 61–86. [[CrossRef](#)]
5. Ryu, K.; Myung, R. Evaluation of mental workload with a combined measure based on physiological indices during a dual task of tracking and mental arithmetic. *Int. J. Ind. Ergon.* **2005**, *35*, 991–1009. [[CrossRef](#)]
6. Reilly, R.B.; Lee, T.C. Electrograms (ECG, EEG, EMG, EOG). *Stud. Health Technol. Inform.* **2010**, *152*, 90–108. [[CrossRef](#)]

7. Aricò, P.; Reynal, M.; Flumeri, G.D.; Borghini, G.; Sciaraffa, N.; Imbert, J.P.; Hurter, C.; Terenzi, M.; Ferreira, A.; Pozzi, S.; et al. How Neurophysiological Measures Can be Used to Enhance the Evaluation of Remote Tower Solutions. *Front. Hum. Neurosci.* **2019**, *13*, 303. [\[CrossRef\]](#)
8. Zhang, J.; Cao, X.; Wang, X.; Pang, L.; Liang, J.; Zhang, L. Physiological responses to elevated carbon dioxide concentration and mental workload during performing MATB tasks. *Build. Environ.* **2021**, *195*, 107752. [\[CrossRef\]](#)
9. Acampora, G.; Trinchese, P.; Vitiello, A. A Dataset of EEG signals from a single-channel SSVEP-based Brain Computer Interface. *Data Brief* **2021**, *35*, 106826. [\[CrossRef\]](#)
10. Van, E.; Anne-Marie, B.; Zander, T.O. Editorial: Using neurophysiological signals that reflect cognitive or affective state. *Front. Neurosci.* **2015**, *9*, 193.
11. Sya, B.; Zhong, Y.; Ywa, B.; Zhang, W.; Wang, Y.; Zhang, J. Assessing cognitive mental workload via EEG signals and an ensemble deep learning classifier based on denoising autoencoders. *Comput. Biol. Med.* **2019**, *109*, 159–170.
12. Banville, H.; Parent, M.; Tremblay, S.; Falk, T.H. Toward Mental Workload Measurement Using Multimodal EEG–fNIRS Monitoring. *Neuroergonomics* **2018**, *2018*, 245–246. [\[CrossRef\]](#)
13. Abhang, P.A.; Gawali, B.W.; Mehrotra, S.C. Technological Basics of EEG Recording and Operation of Apparatus. In *Introduction to EEG- and Speech-Based Emotion Recognition*; Elsevier: Amsterdam, The Netherlands, 2016; pp. 19–50.
14. So, W.; Wong, S.; Mak, J.N.; Chan, R.H.M. An evaluation of mental workload with frontal EEG. *PLoS ONE* **2017**, *12*, e0174949. [\[CrossRef\]](#) [\[PubMed\]](#)
15. Nithin, G.; Sathidevi, P.S.; Ameer, P.M. Graph energy based centrality measures to detect epileptogenic focal invasive EEG electrodes. *Seizure* **2021**, *85*, 127–137.
16. Herrmann, C.S.; Demiralp, T. Human EEG gamma oscillations in neuropsychiatric disorders. *Clin. Neurophysiol.* **2005**, *116*, 2719–2733. [\[CrossRef\]](#) [\[PubMed\]](#)
17. Fasil, O.K.; Rajesh, R. Time-Domain Exponential Energy for Epileptic EEG Signal Classification. *Neurosci. Lett.* **2018**, *694*, 1–8.
18. Sharmila, A.; Mahalakshmi, P. Wavelet-based feature extraction for classification of epileptic seizure EEG signal. *J. Med. Eng. Technol.* **2017**, *41*, 670. [\[CrossRef\]](#)
19. Siuly, L.Y.; Wen, P. Modified CC-LR algorithm with three diverse feature sets for motor imagery tasks classification in EEG based brain–computer interface. *Comput. Methods Programs Biomed.* **2014**, *113*, 767–780. [\[CrossRef\]](#)
20. Subasi, A.; Gursoy, M.I. EEG signal classification using PCA, ICA, LDA and support vector machines. *Expert Syst. Appl.* **2010**, *37*, 8659–8666. [\[CrossRef\]](#)
21. Edla, D.R.; Ansari, M.F.; Chaudhary, N.; Dodia, S. Classification of Facial Expressions from EEG signals using Wavelet Packet Transform and SVM for Wheelchair Control Operations. *Procedia Comput. Sci.* **2018**, *132*, 1467–1476. [\[CrossRef\]](#)
22. Ieracitano, C.; Mammone, N.; Bramanti, A.; Hussain, A.; Morabito, F.C. A Convolutional Neural Network approach for classification of dementia stages based on 2D-spectral representation of EEG recordings. *Neurocomputing* **2019**, *323*, 96–107. [\[CrossRef\]](#)
23. Qu, H.; Shan, Y.; Liu, Y.; Pang, L.; Fan, Z.; Zhang, J.; Wanyan, X. Mental Workload Classification Method Based on EEG Independent Component Features. *Appl. Sci.* **2020**, *10*, 3036. [\[CrossRef\]](#)
24. Pang, L.; Guo, L.; Zhang, J.; Wanyan, X.; Qu, H.; Wang, X. Subject-specific mental workload classification using EEG and stochastic configuration network (SCN). *Biomed. Signal Process. Control* **2021**, *68*, 102711. [\[CrossRef\]](#)
25. Von Wegner, F.; Tagliazucchi, E.; Laufs, H. Information-theoretical analysis of resting state EEG microstate sequences—non-Markovianity, non-stationarity and periodicities. *Neuroimage* **2017**, *158*, 99. [\[CrossRef\]](#) [\[PubMed\]](#)
26. Bozhokin, S.V.; Suslova, I.B. Wavelet-based analysis of spectral rearrangements of EEG patterns and of non-stationary correlations. *Phys. A Stat. Mech. Appl.* **2015**, *421*, 151–160. [\[CrossRef\]](#)
27. Zanini, P.; Congedo, M.; Jutten, C.; Said, S.; Berthoumieu, Y. Transfer Learning: A Riemannian geometry framework with applications to Brain-Computer Interfaces. *IEEE Trans. Biomed. Eng.* **2017**, *65*, 1107–1116. [\[CrossRef\]](#) [\[PubMed\]](#)
28. Mühl, C.; Jeunet, C.; Lotte, F. EEG-based workload estimation across affective contexts. *Front Neurosci.* **2014**, *8*, 114. [\[CrossRef\]](#)
29. Fahimi, F.; Zhang, Z.; Goh, W.B.; Lee, T.S.; Ang, K.K.; Guan, C. Inter-subject transfer learning with an end-to-end deep convolutional neural network for EEG-based BCI. *J. Neural Eng.* **2019**, *16*, 2. [\[CrossRef\]](#)
30. Pan, S.J.; Yang, Q. A Survey on Transfer Learning. *IEEE Trans. Knowl. Data Eng.* **2010**, *22*, 1345–1359. [\[CrossRef\]](#)
31. Pan, S.J.; Tsang, I.W.; Kwok, J.T.; Yang, Q. Domain Adaptation via Transfer Component Analysis. *IEEE Trans. Neural Netw.* **2011**, *22*, 199–210. [\[CrossRef\]](#)
32. Fernando, B.; Habrard, A.; Sebban, M.; Tuytelaars, T. Unsupervised Visual Domain Adaptation Using Subspace Alignment. In *Proceedings of the IEEE International Conference on Computer Vision*, Sydney, NSW, Australia, 1–8 Dec. 2013; pp. 2960–2967.
33. Gong, B.; Yuan, S.; Fei, S.; Grauman, K. Geodesic flow kernel for unsupervised domain adaptation. In *Proceedings of the 2012 IEEE Conference on Computer Vision and Pattern Recognition*, Providence, RI, USA, 16–21 June 2012; pp. 2066–2073.
34. Yin, Z.; Wang, Y.; Liu, L.; Zhang, W.; Zhang, J. Cross-Subject EEG Feature Selection for Emotion Recognition Using Transfer Recursive Feature Elimination. *Front. Neurobotics* **2017**, *11*, 19. [\[CrossRef\]](#)
35. Li, J.; Qiu, S.; Du, C.; Wang, Y.; He, H. Domain Adaptation for EEG Emotion Recognition Based on Latent Representation Similarity. *IEEE Trans. Cognit. Dev. Syst.* **2019**, *12*, 344–353. [\[CrossRef\]](#)
36. Zheng, W.L.; Lu, B.L. Personalizing EEG-based affective models with transfer learning. In *Proceedings of the Twenty-Fifth International Joint Conference on Artificial Intelligence (IJCAI'16)*, New York, NY, USA, 9–15 July 2016; pp. 2732–2738.

37. Xin, C.; Wang, Q.; Zhao, Y.; Liu, X.; Liu, D.; Bai, O. Multi-subject subspace alignment for non-stationary EEG-based emotion recognition. *Technol. Health Care* **2018**, *26*, 327–335.
38. Sciaraffa, N.; Germano, D.; Giorgi, A.; Ronca, V.; Vozzi, A.; Borghini, G.; Di Flumeri, G.; Babiloni, F.; Arico, P. Mental Effort Estimation by Passive BCI: A Cross-Subject Analysis. In Proceedings of the 2021 43rd Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC), Guadalajara, Mexico, 31 October–4 November 2021.
39. Zhang, J.; Wang, Y.; Li, S. Cross-subject mental workload classification using kernel spectral regression and transfer learning techniques. *Cogn. Technol. Work* **2017**, *19*, 587–605. [[CrossRef](#)]
40. Liu, Y.; Hasan, A.; Shewokis, P.A. Multisubject “Learning” for Mental Workload Classification Using Concurrent EEG, fNIRS, and Physiological Measures. *Front. Hum. Neurosci.* **2017**, *11*, 389. [[CrossRef](#)] [[PubMed](#)]
41. Yin, Z.; Zhang, J. Cross-session classification of mental workload levels using EEG and an adaptive deep learning model. *Biomed. Signal Processing Control* **2017**, *33*, 30–47. [[CrossRef](#)]
42. Schwiedrzik, C.M.; Singer, W.; Treisman, M.A. Subjective and objective learning effects dissociate in space and in time. *Proc. Natl. Acad. Sci. USA* **2011**, *108*, 4506–4511. [[CrossRef](#)]
43. Li, D.; Li, F.; Qiang, L.; Wen, A.; Lai, Y.; Xu, P.; Yao, D. MATLAB Toolboxes for Reference Electrode Standardization Technique (REST) of Scalp EEG. *Front. Neurosci.* **2017**, *11*, 601.
44. Nitschke, J.B.; Miller, G.A.; Cook, E.W. Digital filtering in EEG/ERP analysis: Some technical and empirical comparisons. *Behav. Res. Methods Instrum. Comput.* **1998**, *30*, 54–67. [[CrossRef](#)]
45. Khatun, S.; Mahajan, R.; Morshed, B.I. Comparative analysis of wavelet based approaches for reliable removal of ocular artifacts from single channel EEG. In Proceedings of the 2015 IEEE International Conference on Electro/Information Technology (EIT), Dekalb, IL, USA, 21–23 May 2015; pp. 335–340. [[CrossRef](#)]
46. Bell, A.J.; Sejnowski, T.J. Blind separation and blind deconvolution: An information-theoretic approach. In Proceedings of the International Conference on Acoustics, Detroit, Michigan, 9–12 May 1995; IEEE: Piscataway, NJ, USA, 1995.
47. Cichocki, A.; Shishkin, S.; Musha, T.; Leonowicz, Z.; Asada, T.; Kurachie, T. EEG filtering based on blind source separation (BSS) for early detection of Alzheimer's disease. *Clin. Neurophysiol.* **2005**, *116*, 729–737. [[CrossRef](#)]
48. Hyvarinen, A.; Oja, E. Independent Component Analysis: Algorithms and Applications. *Neural Netw.* **2000**, *13*, 411–430. [[CrossRef](#)]
49. Zhang, Y.Q.; Zheng, W.L.; Lu, B.L. Transfer Components Between Subjects for EEG-based Driving Fatigue Detection. In Proceedings of the International Conference on Neural Information Processing, Istanbul, Turkey, 9–12 November 2015; Springer International Publishing: Berlin/Heidelberg, Germany, 2015.
50. Zheng, W.; Zhang, Y.; Zhu, J.; Lu, B. Transfer components between subjects for EEG-based emotion recognition. In Proceedings of the 2015 International Conference on Affective Computing and Intelligent Interaction (ACII), Xi'an, China, 21–24 September 2015; pp. 917–922. [[CrossRef](#)]
51. Steinwart, I. On the Influence of the Kernel on the Consistency of Support Vector Machines. *J. Mach. Learn. Res.* **2002**, *2*, 67–93.
52. Jolliffe, I.T. *Principal Component Analysis*; Springer: Berlin/Heidelberg, Germany, 2005.
53. Mirkin, B. *Principal Component Analysis and SVD*; Springer: London, UK, 2011.