Cross-Domain Feature Learning Methods for Motor Imagery EEG Signals: Postprint

Authors: Wei Hongyu, Chen Lifei, Luo Tianjian

Date: 2022-04-07T00:00:00+00:00

Abstract

Motor imagery EEG signals are costly to acquire and exhibit significant individual variability, making the construction of EEG signal pattern recognition models across subjects a typical few-shot cross-domain learning task. To address this challenge, we propose a cross-domain feature learning method for motor imagery EEG signals. The method first selects an optimal metric approach to align covariances and extract common spatial pattern features; second, based on these features, domain adaptation techniques are employed to learn optimal cross-domain features for the target domain. To verify the feasibility and effectiveness of the proposed method, classical models are utilized to identify cross-domain features, and comparative experiments are conducted on two public datasets. Experimental results demonstrate that the cross-domain features learned through the proposed method significantly outperform those learned by existing methods in motor imagery pattern recognition. Furthermore, a detailed comparison of various parameter settings, performance, and efficiency of the cross-domain feature learning method is presented.

Full Text

Cross-Domain Feature Learning Method for Motor Imagery EEG Signals

Wei Hongyu[1], **Chen Lifei**[1,2,3]**†**, **Luo Tianjian**[1,2] [1]College of Computer & Cyber Security, [2]Digital Fujian Environmental Monitoring Internet of Things Laboratory, [3]Fujian Applied Mathematics Center, Fujian Normal University, Fuzhou 350108, China

Abstract

Motor imagery EEG signals are costly to acquire and exhibit significant individual variability, making cross-subject EEG pattern recognition a typical small-sample cross-domain learning task. To address this challenge, this paper proposes a cross-domain feature learning method for motor imagery EEG signals. The method first selects an optimal metric to align covariance matrices and extract Common Spatial Pattern (CSP) features. Subsequently, domain adaptation techniques are employed to learn optimal cross-domain features for the target domain. To validate the feasibility and effectiveness of the proposed approach, classical models are used to recognize the learned cross-domain features, with comparative experiments conducted on two public datasets. Experimental results demonstrate that the cross-domain features learned by our method significantly outperform those obtained by existing methods in motor imagery pattern recognition. Furthermore, the paper provides a detailed comparison of parameter settings, performance, and efficiency of the cross-domain feature learning method.

Keywords: motor imagery; EEG signal; cross-domain feature learning; domain adaptation; covariance alignment

0 Introduction

Brain-Computer Interface (BCI) technology bypasses the muscular system to establish a direct communication pathway between the brain and external devices, enabling direct information exchange between the nervous system and external equipment. BCI technology consists of three main steps: acquiring brain signals, analyzing and decoding them, and outputting control commands to external devices based on the decoding results. Scalp electroencephalography (EEG) is one of the most commonly used signal types in BCIs, recorded non-invasively to capture electrophysiological phenomena of brain activity. EEG-based BCIs primarily include three paradigms: Motor Imagery (MI), Steady-State Visual Evoked Potentials (SSVEP), and Event-Related Potentials (ERP). Among these, MI represents an active BCI mode, making EEG-based MI-BCI the most frequently used paradigm and the focus of this study.

In MI-BCI, users imagine movements of body parts according to cues, which modulates sensorimotor rhythms in corresponding cortical regions and produces characteristic patterns in EEG signals. Different imagined movements generate distinct spatially localized patterns, enabling the identification of various motor imagery rhythms for controlling external devices. Early MI-BCI applications in medicine helped severely paralyzed patients control wheelchairs without muscular involvement. More recently, MI-BCI has been used for rehabilitation training in stroke patients to improve motor function, and its entertainment applications have expanded to healthy populations. However, EEG signals record weak scalp potentials that are susceptible to noise. Common Spatial Patterns (CSP) is a widely used spatial filtering algorithm for feature extraction that effectively extracts optimal spatial distribution components for each class from multi-channel EEG signals. For pattern recognition, Linear Discriminant Analysis (LDA) or

Support Vector Machines (SVM) are typically applied to CSP features.

Due to the high cost of EEG acquisition, individual subjects usually provide limited samples, and significant inter-subject variability in EEG signal distributions prevents direct use of cross-subject data for building high-performance pattern recognition models. Consequently, researchers have employed transfer learning to reduce individual differences in EEG feature distributions and construct robust models across subjects. Domain adaptation is the most common transfer learning task, with classical methods including Transfer Component Analysis (TCA) by Pan et al., Joint Distribution Adaptation (JDA) by Long et al., Transfer Joint Matching (TJM) which considers sparse feature selection, and Balanced Distribution Adaptation (BDA) by Wang et al. These methods aim to reduce feature distribution discrepancies between source and target domains to enable joint classifier training, and have been extended to CSP features in MI-BCI to improve recognition performance.

In MI-BCI transfer learning, researchers have combined Kernel CSP (KCSP) with Transfer Kernel Learning (TKL) to propose Transfer Kernel Common Spatial Patterns (TKCSP) for cross-subject EEG processing. Jayaram et al. proposed a multi-task learning framework for cross-subject transfer learning without target domain labels, demonstrating effectiveness for a muscular dystrophy patient. Azab et al. further improved this approach using Kullback-Leibler divergence for weighted transfer learning in MI classification. In ensemble transfer learning, Hossain et al. proposed an ensemble method for cross-subject transfer in multi-class MI classification.

For data alignment, Zanini et al. proposed Riemannian Alignment (RA) and validated it with an improved MDRM classifier. Yair et al. introduced a domain adaptation method using parallel transport on the manifold of SPD matrices, showing promising results for cross-subject MI classification, sleep stage classification, and mental arithmetic recognition. He and Wu improved upon RA with Euclidean Alignment (EA), which aligns EEG signal covariance matrices to the identity matrix in Euclidean space, offering greater convenience and efficiency. Additionally, Rodrigues et al. proposed Riemannian Procrustes Analysis (RPA) for aligning cross-subject EEG signals, demonstrating effectiveness across MI, ERP, and SSVEP paradigms.

In manifold transfer learning, Zhang and Wu proposed Manifold Embedded Knowledge Transfer (MEKT) based on Riemannian manifold covariance alignment in tangent space. Singh et al. developed a sample selection approach that estimates sample covariance matrices for MDRM classifiers from limited target domain samples. Deep learning-based MI-BCI transfer learning has also shown promising results, with Wu et al. proposing a parallel multiscale filter CNN that enhances classification performance in transfer scenarios.

Motor imagery EEG signals exhibit non-stationary characteristics and noise sensitivity, with high acquisition costs and substantial distribution differences across individuals, making cross-subject model construction a typical small-sample cross-domain learning task. Existing research has focused on sample covariance alignment, manifold feature adaptation, or model transfer. To learn stable and reliable cross-domain features across subjects, this paper proposes a cross-domain feature learning method for motor imagery EEG signals to improve cross-subject pattern recognition performance. The method first selects an optimal covariance matrix distance metric to align sample sets to the identity matrix across domains, then extracts CSP features. Subsequently, four domain adaptation methods (TCA, JDA, BDA, and TJM) are introduced to learn optimal cross-domain features for the target domain. We validate the feasibility and effectiveness of the proposed method on two public datasets, with detailed comparative analysis of parameter settings, pattern recognition performance, and efficiency.

Existing cross-domain feature learning methods for MI-BCI address only sample alignment or CSP feature adaptation independently, resulting in insufficient cross-domain learning capability and suboptimal recognition performance. This paper simultaneously addresses both sample-level and CSP feature-level adaptation, selecting optimal sample alignment metrics and CSP feature domain adaptation methods for individual domains, and validates pattern recognition performance and efficiency on two public EEG datasets.

---

1 Optimal Covariance Matrix Sample Alignment

In cross-subject EEG signal sample alignment, this paper employs Euclidean Alignment (EA). Since this method requires aligning samples to a covariance center, we extend the covariance center calculation to different distance metrics and select the optimal metric for each individual.

1.1 Covariance Matrix Center Alignment Method

The covariance matrix center serves as a metric for multidimensional time series distributions. Let X_i denote the i-th trial obtained from a subject performing motor imagery tasks, with each subject having n EEG samples. The EEG signal sample covariance matrix center R for an individual domain is:

$$\langle MATH_0 \rangle$$

In the classic EA algorithm, Euclidean distance is used as the metric. Subsequently, each EEG sample is aligned according to the covariance center: $\langle MATH_1 \rangle$. This center is expressed using algorithmic averaging. After n alignments, the average covariance matrix of all subjects' EEG samples equals the identity matrix, making covariance matrix distributions across subjects more similar.

1.2 Optimal Covariance Matrix Center Distance Metric

Since different tasks and subjects affect EEG signal sample distributions, flexible selection of appropriate covariance center metrics is necessary for each individual domain. Beyond Euclidean distance, this paper employs seven additional mean metrics:

(1) **Riemannian Mean**:

$$\langle MATH_2 \rangle$$

where δ represents the distance between two covariance matrices in Riemannian space. Gradient descent is used for iterative mean computation in Riemannian space.

(2) **Log-Euclidean Mean**: $\langle MATH_3 \rangle$

(3) **Log-Det Mean**: $\langle MATH_4 \rangle$

(4) **Wasserstein Mean**:

$$\langle MATH_5 \rangle$$

Based on the Wasserstein metric, initialize iteration count and update until termination.

(5) **AJD Log-Mean**:

$$\langle MATH_6 \rangle$$

where matrix B is based on Pham' s algorithm for approximate joint diagonalization, and A is the inverse of B.

(6) **Harmonic Mean**: $\langle MATH_7 \rangle$

(7) **KL Divergence Mean**: $\langle MATH_8 \rangle$

In practice, this mean represents the geometric mean between arithmetic and harmonic means. In actual experiments, we adopt a trial-and-error approach to select the optimal metric based on each individual domain' s EEG sample distribution, compute the covariance matrix center using arithmetic averaging, and align different individual domains' EEG signal sample sets using Equation (1).

---

2 Cross-Domain CSP Feature Learning for EEG Signals

In transfer learning, domain adaptation is a common kernel-based approach that minimizes the distance between source and target domain feature distributions, most frequently measured using Maximum Mean Discrepancy (MMD). In cross-domain CSP feature learning for EEG signals, our objective is to minimize the MMD between multiple source domains and the target domain CSP features.

2.1 Transfer Component Analysis

Transfer Component Analysis (TCA) is a marginal distribution adaptation method that minimizes the difference between source and target domain marginal distributions. Assuming the marginal distributions are represented by $\langle MATH_9 \rangle$ and $\langle MATH_{10} \rangle$, TCA aims to minimize the MMD distance. The optimization objective is:

$$\langle MATH_{11} \rangle$$

where A is the projection matrix for MMD minimization, H is the centering matrix, and $\langle MATH_{12} \rangle$ represents the MMD distance. Using Lagrange multipliers, the optimal solution for A is obtained by selecting eigenvectors corresponding to the top m eigenvalues. Finally, source and target domain features are transformed through mapping $\langle MATH_{13} \rangle$.

2.2 Joint Distribution Adaptation

Joint Distribution Adaptation (JDA) minimizes the joint probability distribution distance between source and target domains. Since target domain labels are unavailable, $\langle MATH_{14} \rangle$ cannot be directly computed. JDA iteratively trains a simple classifier on source domain features and labels to predict pseudo-labels $\hat{y}_t$ on target domain features. The JDA optimization objective is:

$$\langle MATH_{15} \rangle$$

where M is the center matrix, I is the identity matrix, and MMD represents the distance computed from source domain samples with labels and target domain samples with pseudo-labels. The projection matrix A is obtained similarly to transform source and target domain features.

2.3 Balanced Distribution Adaptation

JDA treats marginal and conditional distributions as equally important, which may not hold across different BCI tasks and populations. To address this limitation, Balanced Distribution Adaptation (BDA) dynamically adjusts the importance between marginal and conditional distributions. The BDA optimization objective is:

$$\langle MATH_{16} \rangle$$

where μ is estimated using A-distance. Notably, BDA recalculates μ after each iteration using pseudo-labels and A-distance. The obtained projection matrix transforms source and target domain features.

2.4 Transfer Joint Matching

Both JDA and BDA use all source domain features to compute the projection matrix. However, only some source features are beneficial for transfer, while others may cause negative transfer. Therefore, Transfer Joint Matching (TJM) adds sparse regularization $\langle MATH_{17} \rangle$ to the source domain projection matrix while maintaining ℓ_2 norm constraints on the target domain matrix. This selects more representative source features during optimization, improving transfer performance. The TJM formulation extends TCA as:

$$\langle MATH_{18} \rangle$$

where λ is the regularization coefficient. Due to the non-convex objective, iterative solving is required. The sparse mapping matrix constructed from A transforms features while ensuring sparsity to extract the most representative features.

---

3 CSP Feature Extraction and Classification

3.1 Common Spatial Patterns

Common Spatial Patterns (CSP) extracts discriminative features from multi-channel EEG signals for binary classification. Let X_1 and X_2 represent EEG signals from two different motor imagery tasks. First, compute the average covariance matrix $\langle MATH_{19} \rangle$ from aligned EEG trials. Then, find the optimal spatial filter w that maximizes variance for Class 1 while minimizing variance for Class 2, with the optimization objective:

$$\langle MATH_{20} \rangle$$

where λ and w are generalized eigenvalues and eigenvectors, respectively. EEG signals are spatially filtered using $Z = w^T X$ to obtain feature vectors. The first and last m rows of Z are selected as separated spatial filter features. For EEG signal X_i, the final CSP-extracted feature is:

$$\langle MATH_{21} \rangle$$

3.2 Linear Discriminant Analysis

Linear Discriminant Analysis (LDA) is a commonly used classifier in MI-BCI that classifies extracted CSP features to complete motor imagery pattern recognition. Given feature set $D = \{(x_i, y_i)\}_{i=1}^n$ where $y_i \in \{0, 1\}$ represents class labels, LDA computes within-class scatter matrix S_w and between-class scatter matrix S_b, ensuring S_b is large while S_w is small. The optimization objective is:

$$\langle MATH_{22} \rangle$$

The optimal projection vector is solved using Lagrange multipliers: $w^* = S_w^{-1}(\mu_0 - \mu_1)$. In practical motor imagery pattern recognition, learned cross-domain features are projected onto the optimal direction and classified using LDA.

3.3 Cross-Domain Feature Learning Method Flowchart

The proposed MI-BCI EEG signal cross-domain feature learning method is illustrated in Figure 1 [Figure 1: see original paper]. The procedure consists of: 1. Selecting the optimal covariance distance metric using trial-and-error based on each individual domain' s EEG sample distribution, then aligning different individual domains' EEG signals to the identity matrix to increase distribution similarity. 2. Extracting CSP features from aligned EEG trials, designating one subject' s CSP features as the target domain and remaining subjects as source domains, then applying domain adaptation to minimize MMD between CSP features. The optimal domain adaptation method is selected from TCA, JDA, TJM, and BDA to learn the most suitable cross-domain features for the target domain. 3. Using an LDA classifier for cross-domain feature classification to complete motor imagery EEG signal pattern recognition.

---

4 Experiments and Analysis

4.1 Experimental Setup

To validate the feasibility and effectiveness of the proposed method, we selected two public MI-BCI datasets from BCI Competition IV [36,37]: MI-1 and MI-2a. The experimental paradigm for both datasets is shown in Figure 2 [Figure 2: see original paper]. Subjects sat comfortably facing a screen. At $t = 0$s, a cross appeared with a brief tone; at $t = 2$s, an arrow cue appeared directing the motor imagery task; at $t = 6$s, the imagery ended and the arrow disappeared, followed by a short rest before the next trial.

Dataset MI-1: Contains EEG data from 7 healthy subjects performing motor imagery, with 59 channels sampled at 100 Hz. Experiments extracted left- and right-hand imagery tasks (Class 1 and Class 2), with 100 trials per class per subject.

Dataset MI-2a: Contains EEG data from 9 healthy subjects, with 22 channels sampled at 250 Hz. Experiments extracted left- and right-hand imagery tasks (Class 1 and Class 2), with 72 trials per class per subject.

Preprocessing was performed in MATLAB, using samples from 0.5s to 3.5s after the imagery cue. This yielded 300 time points per sample for MI-1 and 750 for MI-2a.

Experiments ran on an Intel Core i7-9700 CPU with MATLAB R2017b and Python 3.7.4. Covariance alignment used code from https://github.com/hehe91/EA, and domain adaptation methods from https://github.com/jindongwang/transferlearning. For JDA, BDA, and TJM, iterations were set to 10, with other parameters tuned for optimal performance. The best result across all iterations was reported.

4.2 Covariance Matrix Center Alignment Results

To evaluate alignment effects of different mean metrics, we recorded both the average best-performing metric per dataset and the best metric per subject. Tables 1 and 2 present results for MI-1 and MI-2a, respectively. Each row shows classification accuracy for a subject across 8 mean metrics; each column shows accuracy across subjects for a given metric. The last row shows average results.

Key findings: - For MI-1, Euclidean mean performed best while KL divergence mean performed worst. KL divergence mean is sensitive to poor individual results (subjects 1 and 4 performed poorly), degrading the average. - For MI-2a, KL divergence mean performed best while harmonic mean performed worst. Subject 2 achieved only 50.69% accuracy with harmonic mean, the worst among all metrics. KL divergence mean' s worst-case accuracy was 54.86%, making it most robust.

4.3 Cross-Domain Feature Learning Results

For aligned EEG samples, CSP features were extracted and classified using LDA. Due to limited samples, leave-one-subject-out cross-validation was used: each subject served as target domain once while remaining subjects formed the source domain. Average classification accuracy across all subjects was reported.

4.3.1 Experimental Results Building on covariance alignment results, we tested domain adaptation performance with and without alignment. Table 3 and Table 4 compare four domain adaptation methods (TCA, BDA, JDA, TJM) with and without Euclidean alignment on both datasets.

Key observations: 1. Cross-domain feature learning (with alignment) consistently outperformed domain adaptation alone, demonstrating the crucial role of covariance alignment in reducing individual distribution differences. 2. For MI-1, TJM performed best; for MI-2a, JDA was optimal under Euclidean alignment, indicating method suitability depends on dataset characteristics. 3. Compared to domain adaptation without alignment, all four methods improved with alignment: EA+TJM improved by 14.35% on MI-1, and EA+JDA improved by 3.09% on MI-2a.

To explore optimal per-subject methods, Tables 5 and 6 present individual best results combining optimal alignment metrics and domain adaptation methods.

Key findings: 1. Optimal metric alignment combined with optimal domain adaptation further improved performance by 2.43% and 3.40% on average for MI-1 and MI-2a, respectively. 2. Metric selection significantly impacts final accuracy, and investing time to find optimal configurations substantially improves performance. 3. The proposed method enhances individual classification performance by flexibly selecting appropriate methods based on data distribution differences.

Figure 3 [Figure 3: see original paper] compares four scenarios per subject: (1) non-transfer (CSP+LDA), (2) domain adaptation without alignment (CSP+DA+LDA), (3) optimal alignment without adaptation (EA+CSP+LDA), and (4) the proposed optimal cross-domain method (EA+CSP+DA+LDA). The proposed method achieved highest accuracy for all subjects, with subject 4 in MI-1 showing a 36.5% improvement over non-transfer methods.

Figure 4 [Figure 4: see original paper] analyzes the effect of parameter dim on four domain adaptation methods under alignment. Small dim values limit transferable CSP features, reducing accuracy. Performance initially increases then decreases with dim, peaking at $dim = 4$.

Figure 5 [Figure 5: see original paper] shows results using RBF kernel. For MI-1, Euclidean mean with TJM performed best; for MI-2a, log-Euclidean mean with JDA was superior. This confirms that Euclidean+TJM suits MI-1 distribution while log-Euclidean+JDA suits MI-2a.

Figure 6 [Figure 6: see original paper] compares iteration efficiency. TCA is unsupervised and non-iterative, while BDA, JDA, and TJM require iterations. Performance improves with iterations, but most cases converge within 5 iterations, making the method suitable for online BCI applications.

4.3.2 Feature Visualization To visually assess cross-domain learning effects, we applied t-SNE to project features into 2D space. Figures 7 [Figure 7: see original paper] and 8 [Figure 8: see original paper] show results for subject 4 in MI-1 and subject 8 in MI-2a, respectively, with red and blue dots representing two classes.

Observations: - After sample alignment alone, classes remain highly overlapped. Domain adaptation after alignment reduces inter-class overlap and improves separability. - For MI-2a, pre-aligned data already shows good separation, but alignment further reduces overlap and cross-domain learning enhances class distinction. - Covariance alignment reduces marginal distribution differences across subjects. TCA further reduces marginal distribution differences within subjects, while JDA and BDA reduce both marginal and conditional distribution differences, improving accuracy. TJM ensures feature sparsity, making selected features more discriminative.

4.4 Time Complexity Analysis

Figure 9 [Figure 9: see original paper] presents runtime comparisons. Figure 9(a) shows average runtime for different mean metrics. Euclidean mean is fastest (2-10× speedup), while AJD-based log-Euclidean mean is slowest. MI-1, with larger data volume, requires more time.

Figure 9(b)(c) shows domain adaptation runtime under Euclidean alignment. With 5 iterations, BDA, JDA, and TJM approach optimal solutions. TCA is fastest, while BDA, JDA, and TJM have similar runtime (<1.5s), sufficient for online BCI applications.

Current BCI systems (e.g., exoskeleton rehabilitation) require 200-300 trials (4s each) for new subjects to achieve stable recognition. Our method needs only <100 trials from new subjects, leveraging existing data for improved performance. The time cost for optimal configuration is acceptable.

---

5 Conclusion

This paper proposes a cross-domain feature learning method for motor imagery EEG signals to address high acquisition costs and large inter-subject distribution differences. The method selects optimal covariance distance metrics for sample alignment, extracts CSP features, applies appropriate domain adaptation to learn optimal cross-domain features from multiple source domains, and uses LDA for classification. Experimental results demonstrate significant accuracy improvements. Cross-domain feature learning substantially enhances sample quantity and quality in BCI applications, improving device control performance.

A current limitation is the offline computation of covariance centers and CSP features. Future work will focus on iterative covariance computation methods and incremental cross-domain feature learning for online MI-BCI applications.

References

[1] Lance B J, Kerick S E, Ries A J, et al. Brain-computer interface technologies in the coming decades [J]. Proceedings of The IEEE, 2012, 100(1): 1585-1599.

[2] Wolpaw J R, Birbaumer N, Mcfarland D J, et al. Brain-computer Interfaces for Communication and Control [J]. Clinical Neurophysiology, 2002, 113(6): 767-791.

[3] Zhou Peng. Research on Brain-Computer Interface Based on Motor Imagery [D]. Tianjin: Tianjin University, 2007.

[4] Pfurtscheller G, Müeller-Putz G R, et al. Rehabilitation with Brain-Computer Interface Systems [J]. COMPUTER, 2008, 41(10): 58-65.

[5] Hashimoto Y, Kakui T, Ushiba J, et al. Development of Rehabilitation System with Brain-Computer Interface for Subacute Stroke Patients [C]// Proc of the IEEE International Conference on Systems Man and Cybernetics Conference Proceedings. Miyazaki: IEEE Press, 2018: 51-56.

[6] Kasashima-Shindo Y, Fujiwara T, Ushiba J, et al. Brain-Computer Interface Training Combined with Transcranial Direct Current Stimulation in Patients with Chronic Severe Hemiparesis: Proof of Concept Study [J]. Journal of Rehabilitation Medicine, 2015, 47(4): 318-324.

[7] Blankertz B, Tomioka R, Lemm S, et al. Optimizing Spatial Filters for Robust EEG Single-Trial Analysis [J]. IEEE Signal Processing Magazine, 2008, 25(1): 41-56.

[8] He He, Wu Dongrui. Transfer Learning Enhanced Common Spatial Pattern Filtering for Brain Computer Interfaces (BCIS): Overview and A New Approach [C]// Proc of the 24th International Conference on Neural Information Processing. Guangzhou: ICONIP Press, 2017: 811-821.

[9] Zhuang Fuzhen, Luo Ping, He Qing, et al. Survey on Transfer Learning Research [J]. Journal of Software, 2015, 26(1): 26-39.

[10] Pan S J, Tsang I W, Kwok J T, et al. Domain Adaptation via Transfer Component Analysis [C]// Proc of the 21st International Joint Conference on Artificial Intelligence. Pasadena: IJCAI Press, 2009: 1187-1192.

[11] Long Mingsheng, Wang Jianmin, Ding Guiguang, et al. Transfer Feature Learning with Joint Distribution Adaptation [C]// Proc of the IEEE International Conference on Computer Vision. Sydney: IEEE Press, 2013: 2200-2207.

[12] Long Mingsheng, Wang Jianmin, Ding Guiguang, et al. Transfer Joint Matching for Unsupervised Domain Adaptation [C]// Proc of the 27th IEEE Conference on Computer Vision and Pattern Recognition. Columbus: IEEE Press, 2014: 1410-1417.

[13] Wang Jindong, Chen Yiqiang, Hao Shuji, et al. Balanced Distribution Adaptation for Transfer Learning [C]// Proc of the 17th IEEE International Conference on Data Mining. New Orleans: IEEE Press, 2017: 1129-1134.

[14] Li Jingjing, Meng Lichao, Zhang Ke, et al. A Review of Domain Adaptation Research [J]. Computer Engineering, 2021, 47(06): 1-13.

[15] Tian Lei, Tang Yongqiang, Zhang Wensheng. Two-Stage Domain Adaptive Learning [J]. Pattern Recognition and Artificial Intelligence, 2019, 32(09): 773-784.

[16] Dai Mengxi, Zheng Dezhi, Liu Shucong, et al. Transfer Kernel Common Spatial Patterns for Motor Imagery Brain-Computer Interface Classification [J/OL]. Computational and Mathematical Methods in Medicine, 2018, 2018(9): Article ID 9871603.

[17] Albalawi H, Song Xiaomu. A Study of Kernel CSP-based Motor Imagery Brain Computer Interface Classification [C]// Proc of the IEEE Signal Processing in Medicine and Biology Symposium. New York: IEEE Press, 2012: 1-4.

[18] Long Mingsheng, Wang Jianmin, Sun Jiaguang, et al. Domain Invariant Transfer Kernel Learning [J]. IEEE Trans on Knowledge and Data Engineering, 2015, 27(6): 1519-1532.

[19] Jayaram V, Alamgir M, Altun Y, et al. Transfer Learning in Brain-Computer Interfaces [J]. IEEE Computational Intelligence Magazine, 2015, 11(1): 20-31.

[20] Azab A M, Mihaylova L, Ang K K, et al. Weighted Transfer Learning for Improving Motor Imagery-Based Brain-Computer Interface [J]. IEEE Trans on Neural Systems and Rehabilitation Engineering, 2019, 27(7): 1480-1490.

[21] Hossain I, Khosravi A, Hettiarachchi I, et al. Multiclass Informative Instance Transfer Learning Framework for Motor Imagery-Based Brain-Computer Interface [J]. Computational Intelligence and Neuroscience, 2018, 2018(9): Article ID 6323414.

[22] Zanini P, Congedo M, Jutten C, et al. Transfer Learning: A Riemannian Geometry Framework with Applications to Brain-Computer Interfaces [J]. IEEE Trans Biomed Eng, 2017, 65(5): 1107-1116.

[23] Barachant A, Bonnet S, Congedo M, et al. Multiclass brain-computer interface classification by Riemannian geometry [J]. IEEE Trans on Biomedical Engineering, 2012, 59(4): 920-928.

[24] Yair O, Dietrich F, Talmon R, et al. Domain Adaptation with Optimal Transport on the Manifold of SPD matrices [J]. IEEE Trans on Signal Processing, 2019, 67(7): 1797-1811.

[25] He He, Wu Dongrui. Transfer Learning for Brain-Computer Interfaces: A Euclidean Space Data Alignment Approach [J]. IEEE Trans on Biomedical Engineering, 2020, 67(2): 399-410.

[26] Rodrigues P L C, Jutten C, Congedo M. Riemannian Procrustes Analysis: Transfer Learning for Brain-Computer Interfaces [J]. IEEE Trans on Biomedical Engineering, 2018, 66(8): 2390-2401.

[27] Zhang Wen, Wu Dongrui. Manifold Embedded Knowledge Transfer for Brain-Computer Interfaces [J]. IEEE Trans on Neural Systems and Rehabilitation Engineering, 2020, 28(5): 1117-1127.

[28] Singh A, Lal S, Guesgen H W. Small Sample Motor Imagery Classification Using Regularized Riemannian Features [J]. IEEE Access, 2019, 7(1): 46858-46869.

[29] Wu Dongrui, Xu Yifan, Lu Bao-liang. Transfer Learning for EEG-Based Brain-Computer Interfaces: A Review of Progress Made Since 2016 [J]. IEEE

Trans on Cognitive and Developmental Systems, 2020, 2020(16): Article ID 9134411.

[30] Wu Hao, Niu Yi, Li Fu, et al. A Parallel Multiscale Filter Bank Convolutional Neural Networks for Motor Imagery EEG Classification [J]. Frontiers in Neuroscience, 2019, 13(1): 1-9.

[31] Wang Peitao, Lu Jun, Zhang Bin, et al. A Review on Transfer Learning for Brain-Computer Interface Classification [C]// Proc of the 5th International Conference on Information Science and Technology. Changsha: ICIST Press, 2015: 315-322.

[32] Ren Jiangtao, Liang Zhou, Hu Shaofeng. Multiple Kernel Learning Improved by MMD [C]// Proc of the 6th International Conference on Advanced Data Mining and Applications. Chongqing: ADMA Press, 2010: 63-74.

[33] Zhao Peng, Wu Guoqin, Liu Huiting, et al. Transfer Learning Algorithm Based on Feature Joint Probability Distribution and Instance [J]. Pattern Recognition and Artificial Intelligence, 2016, 29(08): 717-724.

[34] Mishuhina V, Jiang Xudong. Feature Weighting and Regularization of Common Spatial Patterns in EEG-Based Motor Imagery BCI [J]. IEEE Signal Processing Letters, 2018, 25(6): 783-787.

[35] Subasi A, Gursoy M I. EEG signal classification using PCA, ICA, LDA and support vector machines [J]. Expert Systems with Applications, 2010, 37(12): 8659-8666.

[36] Tangermann M, Muller K R, Aertsen A, et al. Review of the BCI Competition IV [J]. Frontiers of Neuroscience, 2012, 6(1).

[37] Ang K K, Chin Z Y, Wang C C, et al. Filter Bank Common Spatial Pattern Algorithm on BCI Competition IV Datasets 2a and 2b [J]. Frontiers of Neuroscience, 2012, 6(1).

Note: Figure translations are in progress. See original paper for figures.