

Autoencoder-based transfer learning in brain–computer interface for rehabilitation robot

Chuanqi Tan¹, Fuchun Sun, Bin Fang, Tao Kong and Wenchang Zhang

Abstract

The brain–computer interface-based rehabilitation robot has quickly become a very important research area due to its natural interaction. One of the most important problems in brain–computer interface is that large-scale annotated electroencephalography data sets required by advanced classifiers are almost impossible to acquire because biological data acquisition is challenging and quality annotation is costly. Transfer learning relaxes the hypothesis that the training data must be independent and identically distributed with the test data. It can be considered a powerful tool for solving the problem of insufficient training data. There are two basic issues with transfer learning, under transfer and negative transfer. We proposed a novel brain–computer interface framework by using autoencoder-based transfer learning, which includes three main components: an autoencoder framework, a joint adversarial network, and a regularized manifold constraint. The autoencoder framework automatically encodes and reconstructs data from source and target domains and forces the neural network to learn to represent these domains reliably. The joint adversarial network aims to force the network to learn to encode more appropriately for the source domain and target domain simultaneously, thereby overcoming the problem of under transfer. The regularized manifold constraint aims to avoid the problem of negative transfer by avoiding geometric manifold structure in the target domain being destroyed by the source domain. Experiments show that the brain–computer interface framework proposed by us can achieve better results than state-of-the-art approaches in electroencephalography signal classification tasks. This is helpful in aiding our rehabilitation robot to understand the intention of patients and can help patients to carry out rehabilitation exercises effectively.

Keywords

Brain–computer interface, rehabilitation robot, transfer learning, autoencoder, adversarial network, regularized manifold constraints

Date received: 7 September 2018; accepted: 12 February 2019

Topic: Medical Robotics

Topic Editor: Henry Leung

Associate Editor: Antonio Fernandez-Caballero

Introduction

One of the main reasons for patients with physical disabilities is that their neural pathways are impeded, and medical knowledge suggests that the patient's neural pathways can be reactivated by repeat movements. Rehabilitation robot automatically assists the limb movements in rehabilitation exercises through mechanical devices, which can significantly improve the effectiveness of their rehabilitation

Department of Computer Science and Technology, Tsinghua University, Beijing, China

Corresponding author:

Fuchun Sun, State Key Laboratory of Intelligent Technology and Systems, Tsinghua National Laboratory for Information Science and Technology, Department of Computer Science and Technology, Tsinghua University, Beijing 100084, China.

Email: fcsun@mail.tsinghua.edu.cn



Creative Commons CC BY: This article is distributed under the terms of the Creative Commons Attribution 4.0 License

(<http://www.creativecommons.org/licenses/by/4.0/>) which permits any use, reproduction and distribution of the work without further permission provided the original work is attributed as specified on the SAGE and Open Access pages (<https://us.sagepub.com/en-us/nam/open-access-at-sage>).

exercises and improve their quality of life. In recent years, the brain-computer interface (BCI)-based rehabilitation robot has quickly become a very important research area due to its natural interaction. Compared with traditional rehabilitation robots, the BCI-based rehabilitation robot can read the patients' brain information and decode it into instructions for controlling external devices, such as robots, wheelchairs, and so on. It interacts more naturally with patients than other approaches. The electroencephalography (EEG) signal has become the input signal selected by most popular BCI systems because of its noninvasiveness and convenience.

Currently, it is difficult for the EEG-based BCI to be widely used. One of the most important problems is the insufficient training data. Because the cost of biosignal acquisition and labeling is extremely high, it is difficult to construct a large-scale training data set to train advanced classifiers, which makes it difficult to improve the classification accuracy of BCI. Insufficient training data is a serious problem in all domains related to bioinformatics. Large-scale annotated EEG data sets are almost impossible to acquire because biological data acquisition is challenging and quality annotation is costly. This difficulty has plagued BCI research for a long time. In contrast to biosignals, in other areas, such as natural images, there are already large annotated data sets due to the ease of sample collection and annotation. For example, ImageNet has more than 14 million annotated natural images that have been widely used in computer vision.

Transfer learning utilizes the samples from the source domain to train the model in the target domain, relaxing the hypothesis that training data must be independent and identically distributed (i.i.d.) with the test data, so that the samples in source and target domains in the transfer learning can obey different edge probability distributions or conditional distributions. This motivates us to use transfer learning to solve the problem of insufficient training data in bioinformatics.

We proposed an autoencoder-based transfer learning (ATL) framework that learns representations that are appropriate for both domains and effectively transfer this representation as knowledge to the target domain. The transfer learning framework consists of three main components: an autoencoder framework, a joint adversarial network, and a regularized manifold constraint. The autoencoder automatically encodes and reconstructs the data of the source and target domains and forces the neural network to learn a useful coding feature in the hidden layers. It tries to learn an identity function with the same input and output. Useful information is included in the hidden layers, which is a reliable representation of the input vector. In another way, the role of hidden layers can be identified as "dimension reduction" or "representation." This high-level representation can be transferred from the source domain to the target domain as hidden knowledge.

In the knowledge transfer process, how to improve the result of transfer learning is the main challenge. Under transfer and negative transfer are the two main basic problems in transfer learning that cannot be ignored. Under transfer means that the effect of transfer learning is not obvious or falls short of expectations, often because the knowledge transferred does not match the target domain and belongs to the specialized knowledge in the source domain. To solve the problem of under transfer, we used a joint adversarial training strategy. The joint adversarial training strategy uses an adversarial network to force the network to learn a representation approach that is suitable for the target domain through the data of the source domain by using a specific loss function. Negative transfer refers to the negative effect on the task in the target domain by transfer learning, usually because the difference between the target domain and the source domain is too large, and the data in the target domain is affected by the source domain data. To solve the problem of negative transfer, we used a regularized manifold constraint. The regularized manifold constraint overcomes the problem of negative transfer by avoiding the geometric manifold structure in the target domain being destroyed by the source domain.

Based on the ATL framework we proposed, we can easily apply it to EEG signal classification tasks. We represent the raw EEG signal as a new form, EEG optical flow. The EEG optical flow is a good description of the multimodal information of the raw EEG signal and unifies the original electrical signal format into a similar natural image format. This new representation has many advantages over the raw EEG signal, it can be easily used in our transfer learning framework, and the results can be improved by transfer learning from the natural image. Thus, the natural image is then treated as the source domain, and the EEG optical flow is considered as the target domain. We can apply our ATL framework for knowledge transfer and use the large-scale, well-labeled target data set in the natural image to improve the accuracy of EEG signal classification tasks.

Based on the EEG signal decoding algorithm we proposed, we developed a rehabilitation robot system using hybrid BCI. The rehabilitation robot system uses the electrooculogram (EOG) signal and the EEG signal as a hybrid input and decodes the users' commands by utilizing the characteristics of both the EOG signal and the EEG signal. We have also developed an interface software that will allow patients to choose the type of rehabilitation exercises and how they will use them. Based on the command decoded from the BCI, the intention of the user is executed by a UR-5 robot arm equipped with a Barrett three-finger hand. It will grasp the patients' upper limb, stimulating the patients to do the rehabilitation exercise.

The experimental results show that our ATL-based BCI framework can achieve good results in EEG signal classification tasks and outperform the state-of-the-art approaches. Our rehabilitation robot is able to understand

the users' instructions accurately and can assist the patient in rehabilitation exercises effectively.

The *contributions* of our approach are as follows:

- We proposed an ATL framework to transfer knowledge from source domain to target domain in an effective way. Within this framework, we utilized a joint adversarial training approach based on an adversarial network to overcome the problem of under transfer and use the regularized manifold constraint to overcome the negative-transfer problem.
- We converted the raw EEG signals to EEG optical flow, which was designed to preserve multimodal EEG information in a uniform representation similar to a natural image and obtain the ability to transfer knowledge from natural images.
- We have developed an advanced BCI-based rehabilitation robot system to help the patients to perform upper limb rehabilitation exercises by using the ATL framework we proposed.
- We experimented with a public data set and public subjects, and the results show that our approach has many advantages over traditional approaches.

The rest of this article is organized as follows. We reviewed related work in the second section, the ATL framework is described in the third section, and the BCI-based rehabilitation robot is described in the fourth section. Our experimental results are presented in the fifth section, and the conclusions and further plans are discussed in the last section.

Related work

Substantial work has been conducted to improve the EEG classification accuracy. The performance of this pattern-recognition-like system depends on both the selected features and the employed classification algorithms. Traditionally, a great variety of hand-designed features have been proposed such as band power¹ and power spectral density.² In recent years, common spatial pattern (CSP)³ has proven to be an excellent feature of EEG signals. Abundant related work has been reported, such as common spatio-spectral pattern (CSSP), wavelet CSP, and separable CSSPs. In addition to these single-domain approaches, techniques for extracting multimodal information from EEG signal⁴ and how to fuse the information⁵ have been the focus of many researchers.

As a new classification platform, deep learning has recently received increasing attention from researchers⁶ and has been successfully applied to many classification problems, such as image classification, video classification, and speech recognition. From hand-designed to data-driven features, deep learning has played an important role in various fields in which the artificial intelligence community has struggled for many years. Deep learning has

recently achieved great success in computer vision, which is partially attributed to ImageNet, a large-scale annotated image data set with more than 1.2 million 256×256 images categorized into 1000 object class categories. Many excellent networks have been presented by researchers, such as AlexNet, VGG16, VGG19, GoogLeNet, and ResNet, which outperformed other algorithms in the ImageNet Large-Scale Visual Recognition Competitions (ILSVRC) challenge. Bioinformatics has also benefited from deep learning. In recent years, many public reviews⁷ have discussed deep learning applications in bioinformatics research, for example, applying deep belief networks to the frequency components of EEG signals to classify left-hand and right-hand motor imagery (MI) skills. An et al.⁸ and Cecotti and Graser⁹ used a convolutional neural network (CNN) to decode P300 patterns. Soleymani et al.¹⁰ conducted an emotion detection and facial expressions study with both EEG signals and face images using a recurrent neural network (RNN).

Transfer learning¹¹ and deep transfer learning¹² enable the use of different domains, tasks, and distributions for training and testing. Traditional machine learning techniques attempt to learn each task from scratch, whereas transfer learning techniques attempt to transfer knowledge from previous tasks to the target task, which has a lower-quality training data. Jayaram et al.¹³ reviewed the current state-of-the-art transfer learning approaches in BCI. Tan et al.¹⁴ proposed a novel EEG representation that reduces the EEG classification problem to the image classification problem, which implicates the ability of transfer learning. Wei et al.¹⁵ found that EEG correlates of drowsiness were stable within individuals across sessions and used transfer learning to build a robust drowsiness detector. Wu et al.¹⁶ applied collaborative filtering to reduce the required quantity of training data by combining training data from an individual subject with additional training data from other, similar subjects. Hajinoroozi et al.¹⁷ transferred general features via a convolutional network across subjects and experiments. Zheng and Lu¹⁸ applied kernel principle analysis and transductive parameter transfer to identify the relationships between classifier parameter vectors across subjects. Lin and Jung¹⁹ evaluated the transferability between subjects by calculating distance and transferred knowledge in comparable feature spaces to improve accuracy. Tan et al.²⁰ design a deep transfer learning framework which is suitable for transferring knowledge by joint training, which contains an adversarial network and a special loss function.

For general transfer learning, Tzeng et al.²¹ proposed an approach that can simultaneously transfer knowledge across domains and tasks. Huang et al.²² divide the network into two parts, language-independent feature transform and language-relative classifier, trained on multi-language data sets with common feature transformation and share knowledge. Long et al.^{23–25} attempted the learning of transferable features by embedding task-specific layers in a reproducing

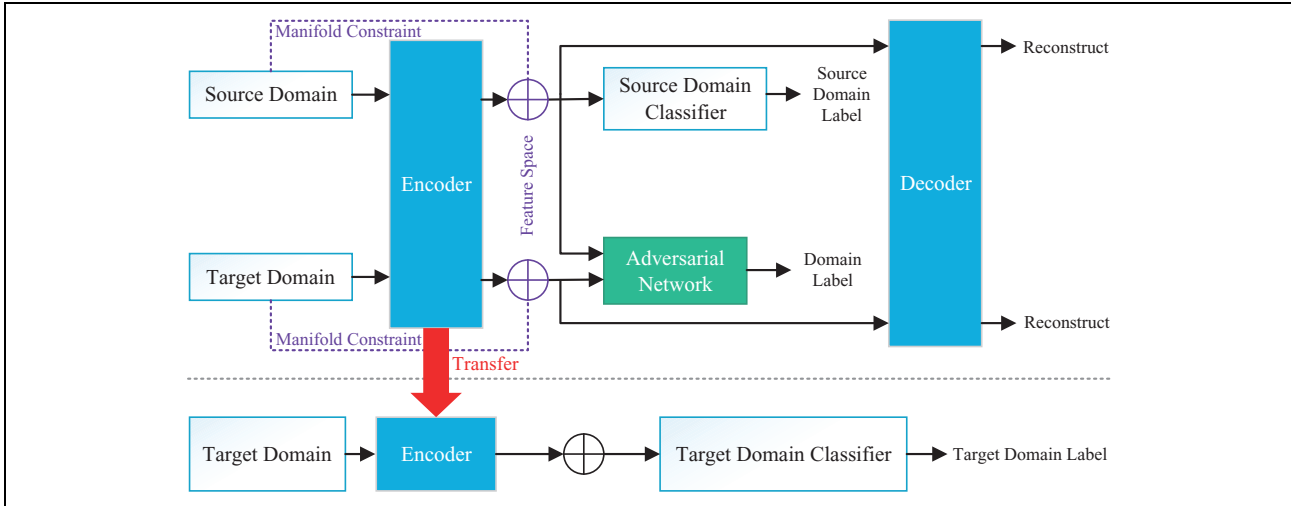


Figure 1. Architecture of our ATL framework. There are three main components in this framework: the autoencoder shown in blue, the adversarial network shown in green, and the manifold constraint shown in purple. First, the samples in the source and target domains are trained using an autoencoder (including adversarial network and manifold constraint). Second, the pretrained encoder in the previous step is transferred and treated as a feature extractor; the features are extracted and provided to the target domain classifier to obtain the final target domain label. ATL: autoencoder-based transfer learning.

kernel Hilbert space (RKHS) where the mean embeddings of different domain distributions can be explicitly matched. Bousmalis et al.²⁶ introduce similar loss; the similar part and difference part of the source domain and the target domain are trained separately. A domain classifier is established between the two domains to learn from the results generated by the gradient reversal layer. Tzeng et al.²⁷ proposed a new approach for unsupervised learning tasks, named adversarial discriminative domain adaptation, which combines discriminative model, untie weight sharing, and generative adversarial net (GAN) loss. It learns a discriminative representation using the label of the source domain, and then uses asymmetric mapping to map the data on the target domain to the same space through domain-adversarial loss.

Being able to measure the similarity between two domains is very important for transfer learning. Many previous works have proposed a list of distance measurement approach. One of the most widely used is the nonparametric statistics named maximum mean discrepancy (MMD) proposed by Gretton et al.,²⁸ which measures the difference between the mean function values on the two samples in kernel space. Long et al.²⁹ proposed an approach to solve the problem of variance drift in transfer learning by using MMD. Kim et al.³⁰ presented an algorithm to learning an online trajectory generator based on supervised imitation learning by using MMD measure. Tommasi et al.³¹ discussed the data set bias problem of CNN features in detail.

ATL framework

Sufficient and balanced data are the most important assumptions in deep learning, satisfying the need to

optimize numerous weight parameters in neural networks. Unfortunately, these assumptions are usually not satisfied in many domains, like bioinformatics and robotics; insufficient training data is a serious problem in these domains. In these domains, the collection of training data is complex and expensive. Therefore, it is extremely difficult to build a large-scale, high-quality annotated data set in these domains. But in some similar domains, it is easier to build large-scale data sets.

Transfer learning can be described as follows, consider in a classification task, X is the input space and $Y = \{0, 1, \dots, N - 1\}$ is the label of input data. In our ATL framework, we have two different distributions over $X \times Y$ with named source domain D_s and target domain D_t . Our transfer learning is then provided with a labeled source sample S drawn i.i.d. from D_s , and a labeled target sample T drawn i.i.d. from D_t . In addition, we have a unlabeled target test sample T_{test} draw i.i.d. from D_t .

$$S = \{(x_i, y_i)\}_{i=1}^n \sim (D_s)^n \quad (1)$$

$$T = \{(x_i, y_i)\}_{i=n+1}^N \sim (D_t)^{n'} \quad (2)$$

with $N = n + n'$ being the total number of train samples from source domain and target domain. In common, $n \gg n'$. The goal of our transfer learning approach is to obtain a feature extractor $\theta(x)$ that is suitable for target domain and a classifier $\vartheta(\theta)$ with a low target risk

$$R_{D_t}(\vartheta) = \Pr_{(x,y) \sim T_{\text{test}}} (\vartheta(\theta(x)) \neq y) \quad (3)$$

We proposed that the ATL framework can achieve better transfer results with low target risk. An overview of ATL is shown in Figure 1. ATL consists of *three main components*:

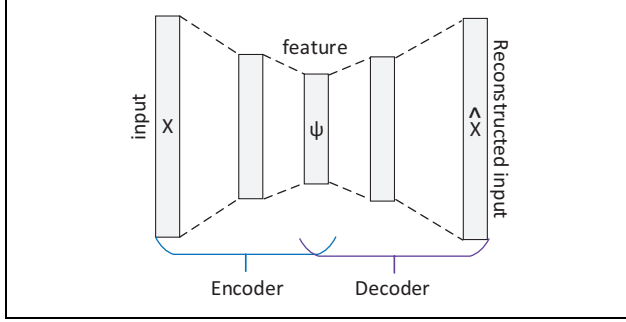


Figure 2. Schematic structure of autoencoder. X represents the original input, the high-level representation of the original input, and the reconstructed input.

- *ATL framework: knowledge is transferred from the source domain to the target domain by hiding knowledge in the features which are extracted by the “encoder”;*
- *adversarial network: utilized to overcome the under-transfer problem by identifying the domain of the feature source; and*
- *manifold constraints: utilized to overcome the negative-transfer problem by avoiding the geometric manifold structure in the target domain being destroyed by the source domain.*

Framework

Autoencoding is an unsupervised machine learning technique that uses neural networks to reconstruct input data, which means that the autoencoder tries to approximate an identity function, making the output as close as possible to the input. The autoencoder contains two main components, an encoder and a decoder. The role of the encoder is to find a compressed representation of the given data. The decoder is the mirror image of the encoder, trying to reconstruct the input as closely as possible. During the training step, the decoder forces the autoencoder to select the most informative feature within given limits, stored in the hidden layers. The closer the reconstructed input to the original input, the better the resulting representation. The architecture of the ATL framework is shown in Figure 2.

The autoencoder is a kind of feed-forward neural network. It is usually used as a tool for data dimensionality reduction or feature extraction. Unlike other feed-forward neural networks, other feed-forward neural network focuses on the output layer and error rate, while the autoencoder focuses on the features contained in the hidden layer. The hidden layer obtains a representation of the raw input, and this representation can be used as hidden knowledge for transfer between similar domains.

Inspired by this idea, we proposed the ATL framework that consists of *two steps*:

- step 1, autoencoder training step (shown in the upper part of Figure 1) and
- step 2, classification step (shown in the lower part of Figure 1).

In step 1, the autoencoder attempts to reconstruct the data in both the source and target domains, forcing the encoder to learn a high-level feature representation. The difference between reconstructed input \hat{X} and origin input X can be used to calculate a reconstruction loss $\mathcal{L}_{\text{reco}}$. In the process of transfer learning, there is a large amount of labeled data in the source domain. In order to utilize the label information in the source domain, we provide the features extracted by the encoder to the source domain classifier, which helps to train a better encoder according to the labels in the source domain. Here, the task loss $\mathcal{L}_{\text{task}}$ can be calculated by the source domain classifier. Thus, the loss function in step 1 can be defined as follows

$$\mathcal{L} = \mathcal{L}_{\text{task}} + \alpha \mathcal{L}_{\text{reco}} + \beta \mathcal{L}_{\text{adver}} + \gamma \mathfrak{R}(v) \quad (4)$$

where α , β , and γ are hyper parameters; $\mathcal{L}_{\text{task}}$ is the loss function of the source domain classifier; $\mathcal{L}_{\text{reco}}$ is the loss function to identify the reconstruction error; and $\mathcal{L}_{\text{adver}}$ and $\mathfrak{R}(v)$ will be introduced in the following sections.

The loss of task can be calculated using the following equation

$$\mathcal{L}_{\text{task}} = - \sum_k \mathbb{I}[y = k] \log p_k \quad (5)$$

where k is the number of categories, $\mathbb{I}(\cdot)$ is the one-hot encoding of the label and predicted label, and p_k is the softmax of the classifier activations.

Bousmalis et al.²⁶ demonstrated that the reconstruction loss $\mathcal{L}_{\text{reco}}$ can be calculated by using a scale-invariant mean squared error term which³² applied to both domains

$$\mathcal{L}_{\text{reco}} = \sum_k \mathcal{L}_{\text{si_mse}}(x_k^s, \hat{x}_k^s) + \sum_k \mathcal{L}_{\text{si_mse}}(x_k^t, \hat{x}_k^t) \quad (6)$$

$$\mathcal{L}_{\text{si_mse}}(x, \hat{x}) = \frac{1}{k} \|x - \hat{x}\|_2^2 - \frac{1}{k^2} ([x - \hat{x}] \cdot \mathbf{1}_k)^2 \quad (7)$$

where k is the number of data in input, $\mathbf{1}_k$ is a vector of ones of length k , and $\|\cdot\|_2$ is the squared L2-norm.

The valuable result of step 1 is the encoder, which will be transferred to step 2 (shown by the red arrow in Figure 1). The data in the target domain will be encoded by the encoder to obtain high-level features and fed to a special target domain classifier. Finally, the final target domain label will be obtained from the classifier.

Adversarial network

Many studies have demonstrated that the hidden layers in an autoencoder can extract the *high-level features* of inputs. However, the encoder trained by source domain does not fully match the target domain. Many previous works have

shown that in order to achieve a better transfer effect, the edge distribution of features from the source domain and target domain should be as similar as possible. Inspired by GANs, we apply an adversarial network to train a better general feature extractor. We use features extracted from the source and target domains as the inputs for the adversarial network and train it to identify their origins. If the adversarial network achieves worse performance, it indicates a small difference between the two types of features and better transferability, and vice versa. The pipeline to obtain a general feature extractor is shown in Figure 1 in green. Joint adversarial training and adaptive sample selection aim to learn a better representation for the source and target domains.

Suppose we have data in source domain $\{X_s, Y_s\}$ and data in target domain $\{X_t, Y_t\}$, where Y is data labels. The goal of joint adversarial training is to produce a general feature extractor with parameters θ_{extr} that can represent the source and target domains in a suitable way and can make correct classifications in the future. θ_{img} and θ_{adver} indicate parameters of the source classifier and the adversarial network, respectively.

To extract more transferable general features, we use a special loss function in the source classification training, which takes into account the performance of the adversarial network with loss function $\mathcal{L}_{\text{adver}}$ defined as follows

$$\mathcal{L}_{\text{adver}} = - \sum_d \frac{1}{D} \log p_d \quad (8)$$

$\mathcal{L}_{\text{adver}}$ is the cross entropy of the adversarial network and β in equation (4) is the hyper-parameter indicating how strongly $\mathcal{L}_{\text{adver}}$ influences the optimization.

The training process via loss function will minimize the performance of the adversarial network. Therefore, the edge distributions of the two types of features will be as similar as possible and will not be easily distinguished by the adversarial network.

When training the source domain classification network, the loss function attempts to reduce the performance of the adversarial network by optimizing equation (9). However, when training the adversarial network, the loss function attempts to improve the performance of the adversarial network by optimizing equation (10). These two goals stand in direct opposition to one another, and we overcome this problem by *iteratively optimizing* the following two goals while fixing the other parameters.

$$\underset{\theta_{\text{extr}}, \theta_{\text{task}}}{\operatorname{argmin}} \mathcal{L}_{\text{task}}(X_s, X_t, \theta_{\text{adver}}; \theta_{\text{extr}}, \theta_s) \quad (9)$$

$$\underset{\theta_{\text{extr}}, \theta_{\text{adver}}}{\operatorname{argmin}} \mathcal{L}_{\text{adver}}(X_s, X_t, \theta_s; \theta_{\text{extr}}, \theta_{\text{adver}}) \quad (10)$$

Joint adversarial training forces the transfer network to discover general features with greater transferability, which is important to obtain useful knowledge from the source

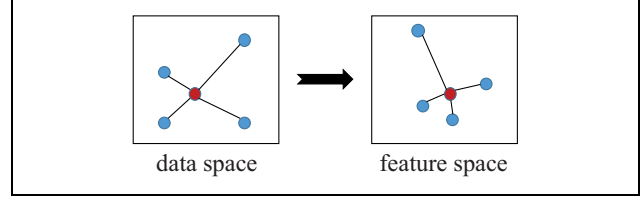


Figure 3. Geometric structures are modeled by p -nearest neighbors, and manifold constraints keep these p -nearest neighbor relationships in data space and feature space.

domain and transfer it to the target domain that can *effectively overcome under transfer*.

Manifold constraints

Joint adversarial training-based transfer learning forces the transfer network to discover a greater “encoder” which can extract features suitable to both source domain and target domain. It is important to obtain useful knowledge from the source domain and target domain. However, negative transfer is one of the biggest challenges for practical application of transfer learning. If the data distribution between the source domain and the target domain is sufficiently large, attempting to transfer knowledge will inevitably lead to inconsistency and thus to negative transfer.

We apply manifold constraints to overcome this problem to avoid the geometric manifold structure in the target domain being destroyed by the source domain. Due to the complete maintenance of the manifold structure in the target domain, the possibility of negative transfer is greatly reduced. Manifold constraints force the learning algorithm to transfer useful knowledge from the source domain and ignore the knowledge which may destroy the manifold structure of the target domain.

From a geometric perspective, the input data can be treated as sampling by probability distribution on a low-dimensional manifold, which is embedded in the high-dimensional surround space. Keeping the geometric structure allows the learning model to respect the domain’s own data distribution and virtually circumvent the negative transfer problem. According to the local invariant assumption, if samples x_i and x_j are close to each other on the inner geometric manifold of the data distribution, their embedded representations v_i and v_j should also be close to each other. According to the manifold theory, the geometric structure can be effectively modeled by the p -nearest neighbor graph on the data set. Manifold constraints can be regarded as keeping p -nearest neighbor relationships in data space and feature space as shown in Figure 3. Consider the sample graph G , which includes n vertices, each vertex represents a sample point in the domain, then the adjacency matrix of graph G can be defined as

$$(W)_{ij} = \begin{cases} \cos(x_{*i}, x_{*j}), x_{*i} \in N_p(x_{*j}) \vee x_{*j} \in N_p(x_{*i}) \\ 0, \text{others} \end{cases} \quad (11)$$

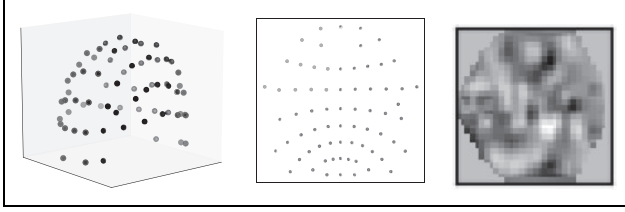


Figure 4. Generating EEG video by projecting and interpolating. EEG: electroencephalography.

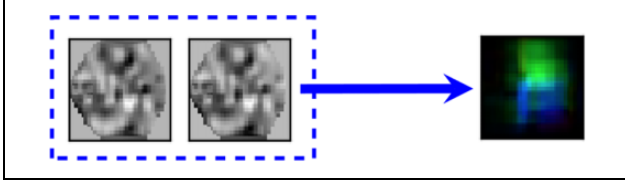


Figure 5. Visualization of the EEG optical flow frame. EEG: electroencephalography.

where $\cos(\cdot, \cdot)$ is the cosine similarity measure function, and $N_p(x)$ is p -nearest neighbors set of x .

Cai et al.³³ demonstrated that keeping the geometric structure can be reduced to the regularization of

$$\mathfrak{R}(v) = \frac{1}{2} \sum_{i,j=1}^n \zeta(v_{i*}, v_{j*})(W)_{ij} \quad (12)$$

where $\zeta(v_{i*}, v_{j*})$ is the loss function to measure the Euclidean distance of v_{i*} and v_{j*} .

We optimize loss function and the regularization term $\mathfrak{R}(v)$ simultaneously during the joint adversarial training; therefore, the manifold structure in the sample space and the feature space can be preserved. After adding $\mathfrak{R}(v)$ to the loss function, our algorithm preserves the manifold structure of the sample space and feature space, which can refine the knowledge structure, such as implicit semantics, and can effectively avoid negative transfer.

BCI-based rehabilitation robot

EEG representation

Traditional approaches do not fully exploit multimodal information. For example, they ignore the locations of the electrodes and the inherent information in the spatial dimension. Tan et al.¹⁴ introduce EEG optical flow as a new EEG presentation, which is designed to preserve multimodal EEG information. First, EEG video is converted from the raw EEG signal by Azimuthal equidistant projection and the Clough–Tocher interpolation algorithm. The processes are shown in Figure 4.

Second, EEG optical flow is extracted from the converted EEG video by the algorithm introduced by Farneback,³⁴ which can describe the obvious motion of objects in a vision scene. This process is shown in Figure 5.

Many benefits can be gained from using the EEG optical flow.

- *Uniform representation of multimodal information:* The spatial structure of the electrodes is preserved by the AEP, and the spectral information extracted via five stereotyped frequency filters and the temporal information are represented by the optical flow.
- *Suitable for CNNs:* Due to the inherent structure of CNNs, the EEG optical flow is more compatible with the image and video data structure. CNNs can discover the regional information in the EEG optical flow, which reflects the regional information of brain regions.

In addition, EEG optical flow obtains the *ability of transfer knowledge from natural images*. By reducing the EEG classification problem to a video classification problem, we gain the ability to transfer knowledge from computer vision, which has large-scale annotated data sets, such as ImageNet, and many excellent networks.

ATL-based BCI

The EEG optical flow representation approach gives us the ability to transfer knowledge from natural images. ImageNet is a very good candidate for the source domain because ImageNet is a large-scale, well-labeled data set with a data structure similar to EEG optical flow. We use ImageNet as the source domain and EEG optical flow as the target domain and apply them to the ATL we proposed in the previous section, as shown in Figure 6.

To further overcome the negative transfer and enhance the effect of transfer learning, we apply a *batch selection* to increase the weights of samples similar to the target domain and reduce the weights of the samples that are significantly different from the target domain. Therefore, objectively measuring the distance between the source and target domains is an important step.

In real analysis and probability theory, the two-sample problem theory provides a theoretical criterion for determining whether two probability distributions P and Q are the same with an infinite sample. Gretton et al.²⁸ provide a kernel-based test approach called MMD, which defines a more general class of statistic for the as yet unspecified function classes $\mathcal{F} \subset C(\mathcal{X})$, to measure the discrepancy between P and Q .

$$\text{MMD}[\mathcal{F}, \mathcal{X}_s, \mathcal{X}_t] = \left\| \sum_{x_i \in \mathcal{X}_s} \frac{\phi(x_i)}{|\mathcal{X}_s|} - \sum_{x_j \in \mathcal{X}_t} \frac{\phi(x_j)}{|\mathcal{X}_t|} \right\| \quad (13)$$

where \mathcal{F} be a class of functions $f: \mathcal{X} \rightarrow \mathbb{R}$ and $\phi(\cdot)$ is a nonlinear feature map of \mathcal{H} in RKHS, which satisfies $f(x) = \langle \phi(x), f \rangle$. Let \mathcal{F} be a unit ball in a universal RKHS \mathcal{H} ; then, $\text{MMD}[\mathcal{F}, P, Q] = 0$ if and only if $P = Q$.

Many previous works suggest that in deep neural networks (DNNs), the front layers can be considered to be a

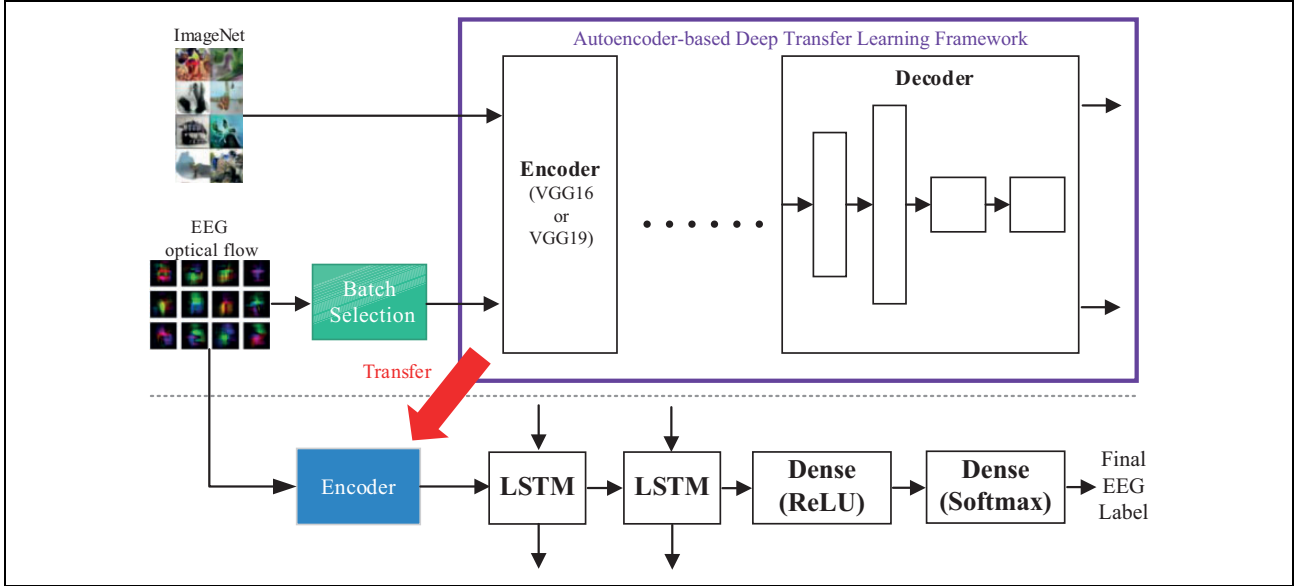


Figure 6. Pipeline of our ATL-based BCI. We use ImageNet as the source domain and EEG optical flow as the target domain and apply them to the ATL framework (in purple) we proposed before. After transfer encoder to step 2, the features extracted by the pretrained encoder are processed by an RNN network to obtain the final EEG label. ATL: autoencoder-based transfer learning; BCI: brain-computer interface; EEG: electroencephalography; RNN: recurrent neural network.

feature extractor and the back layers can be considered to be a classifier. Inspired by these results, we proposed deep representation discrepancy (DRD) based on MMD. We treat the front layers as a nonlinear feature extraction function $R(x)$ to replace the manually specified kernel function $\phi(\cdot)$ in MMD. The formal definition is as follows:

$$\text{DRD}[\mathcal{F}, \mathcal{X}_s, \mathcal{X}_t] = \left\| \sum_{x_i \in \mathcal{X}_s} \frac{R(x_i)}{|\mathcal{X}_s|} - \sum_{x_j \in \mathcal{X}_t} \frac{R(x_j)}{|\mathcal{X}_t|} \right\| \quad (14)$$

where $R(x) : \mathcal{X} \rightarrow \mathcal{H}$ be a nonlinear feature extraction function equal to the front layers of a neural network learned during training.

The DRD distance can measure the similarity between the features extracted by the DNN. This distance can help to select samples similar to the target domain and avoid samples that are very different from the target domain.

The theory of transfer learning tells us that the effect of transfer learning is positively related to the similarity of the probability distributions of the source and target domains. Therefore, we should select similar auxiliary samples and filter out the different auxiliary samples. Therefore, we apply the DRD distance we proposed previously for joint adversarial training to estimate the similarity between each category in source domain and the target domain and assign different weights to the relevant samples based on this value.

$$W'_{(i)} = \exp^{\text{DRD}(X_{of}, X_{(i)})} \quad (15)$$

$$W_{(i)} = \frac{W'_{(i)}}{\sum_{i=1}^n W'_{(i)}} \quad (16)$$

where n is the number of categories in source domain and $X_{(i)}$ is the samples in the i th category of source domain.

This weight value was designed to affect the batch selection process by increasing the weights of samples similar to the target domain and reducing the weights of samples that are substantially different from the target domain.

After the pretrained encoder has been successfully trained in ATL, we transfer this encoder as a feature extractor. Then, features are fed to a classification network with two RNN layers and two fully connected layers. We use long short-term memory to prevent vanishing gradient problems in the time dimension when training. Two fully connected layers are applied at the end of the classification network, with the last layer applying a softmax activation function to obtain the final EEG label. The pipeline of EEG classification network is shown in the bottom of Figure 6. If a fine-tuning strategy is used, the last one or two layers of the transfer network are updated simultaneously in the EEG classification.

We use the categorical cross-entropy loss function, defined in equation (17), to evaluate the loss of the RNN network.

$$\mathcal{L} = -\frac{1}{n} \sum_x [y \ln a + (1 - y) \ln(1 - a)] \quad (17)$$

where n is the number of samples, x is a sample in the data set, y is the expected output, and a is the actual output of the neural network.

We optimize the parameters of the classification network by applying the RMSprop algorithm as the optimizer. Suppose we randomly select m samples in a batch during

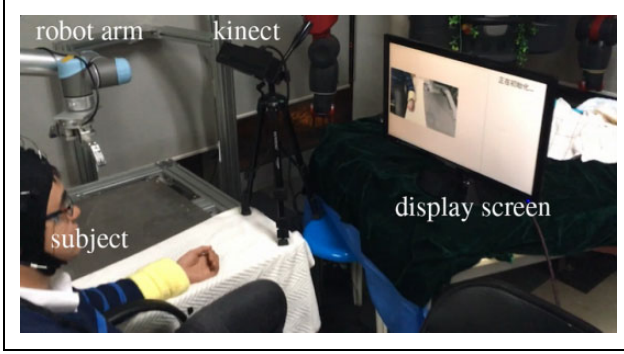


Figure 7. Overview of our rehabilitation robot system that contains a robot arm system, a vision system, a hybrid BCI system, a display system, and a subject. The subject interacts with the software displayed on the screen via the EEG signal, and after the BCI software decodes the commands, the robot arm system is controlled to assist the upper limb movement of the subject for rehabilitation. BCI: brain-computer interface; EEG: electroencephalography.

training, using $(x^{(i)}, y^{(i)})$ to identify the features and label of a sample, respectively. We update the network parameter θ using the following equation

$$\begin{aligned} g &\leftarrow \frac{1}{m} \nabla_{\theta} \sum_i \mathcal{L}(f(x^{(i)}; \theta), y^{(i)}) \\ r &\leftarrow \rho r + (1 - \rho)g \odot g \\ \theta &\leftarrow \theta - \frac{\epsilon}{\sqrt{\delta + r}} \odot g \end{aligned} \quad (18)$$

where $\rho = 0.9$ is the decay rate used to control the length of a move, $\epsilon = 0.001$ is the global learning rate, and $\delta = \frac{1}{10^6}$ is a constant in the equation.

Rehabilitation robot

According to the ATL-based BCI framework we described above, we developed a rehabilitation robot with a hybrid BCI. This rehabilitation robot consists of four main components: a robot arm system, a vision system, a hybrid BCI system, and a display system. An overview of the rehabilitation robot is shown in Figure 7. The rehabilitation robot grabs the arm of the subject and drives the upper limb of said subject to perform movement with the aim of achieving rehabilitation.

The hybrid BCI system uses the EOG, steady state visually evoked potential (SSVEP), and MI patterns as inputs to construct a convenient and fast interaction BCI.

- With EOG detection, it is easy to achieve a high success rate, but with fewer options. We use EOG to confirm or cancel the operation. We simply judge the threshold of the corresponding electrode to obtain the result of the EOG detection.
- SSVEP can achieve a high classification success rate with more options, but the interaction process is



Figure 8. UR-5 robot arm and Barrett three-finger hand used in our rehabilitation robot.

unnatural for the subject and easily makes the subject tired. We use SSVEP for menu selection by using canonical correlation analysis-based SSVEP frequency recognition algorithms that can capture the interrelationship between predictor and response variables.^{35,36}

- The MI interaction approach is natural for the subject and the burden is very light, so we use it for repeat rehabilitation actions. The MI classification uses the ATL-based BCI approach we proposed above.

The robot arm consists of a UR-5 robot arm (left in Figure 8) and a Barrett three-finger hand (right in Figure 8). When it receives the commands, the UR-5 robot arm will automatically move according to the vision system, then the Barrett three-finger hand will grab the upper limb of the subject and perform the rehabilitation exercise. The UR-5 robot arm can capture objects weighing 5 kg, which is sufficient to meet the needs of the upper limb rehabilitation exercise.

The vision system includes a Microsoft Kinect and an object detection algorithm. The subjects' upper limb was marked in yellow. Most object detection algorithms can achieve acceptable results in the simple task of detecting a specific color marker.

The display system displays the software we use to interact with the subject. The left side of the software shows the real-time content acquired by the Microsoft Kinect camera, and the object detected by the vision system will be identified by the red square. The right side of the software shows the menu items or tip information. In menu mode, each menu item has a box that flashes at different frequencies (such as 7 Hz, 9 Hz, 13 Hz, etc.). The subject makes selections with SSVEP by looking at the special flashing box. In addition, a dialog box will pop up in the software asking the subject to confirm or cancel the operation through EOG.

A typical rehabilitation exercise process is described as follows:

- The subject sits next to the robot arm, wearing the EEG electrode cap, and also has a yellow mark on the upper limb to identify the grasp point.
- The menu on the right side of the display screen shows the options for positive or passive rehabilitation mode.

The subject selects an option through SSVEP and then a confirmation dialog box pops up asking the subject to confirm with the EOG signal.

- The vision system detects the grasp point of the subjects' upper limb, which is identified by a red square on the screen, and the subject confirms with the EOG signal.
- The robot arm automatically grasps the upper limb of the subject according to the detected grasp point.
- In positive rehabilitation mode, the subject image is either bent arm or open arm, and the robot arm assists the subject to perform the two types of rehabilitation exercises. In the passive rehabilitation mode, the robot arm automatically stimulates the subjects' arm to perform a rehabilitation exercise according to the fixed pattern.
- The subject looks at the menu item on the right side, the BCI system infers the users intention to end the exercise according to the result of the SSVEP, and a dialog box pops up asking the subject to confirm using the EOG signal.
- The robot arm releases the upper limb of the subject and returns to the initial position, completing the rehabilitation exercise process.

Experiments

EEG classification

We apply our ATL-based BCI approach to a famous public data set in the BCI field called Open Music Imagery Information Retrieval (OpenMIIR), published by the Brain and Mind Institute at the University of Western Ontario.^{37,38} OpenMIIR is an excellent public data set of EEG recordings taken during music perception and imagination. The authors acquired this data set during an ongoing study involving 10 subjects listening to and imagining 12 short music fragments taken from well-known pieces (shown in Table 1). These music fragments were selected from different genres and musical styles and vary with respect to meter, tempo, and display of lyrics. These signals were recorded using 64 EEG electrodes at 512 Hz, and 240 trials were recorded per subject. The following parameters are used in our approach. We convert raw EEG signals into EEG videos with 13 frames and a resolution of 32×32 . These frames are resampled 50 times and converted to EEG optical flow with 12 frames. We employ VGG16 and VGG19 as the targets of encoder, which are the winners of past ILSVRC.³⁹

According to the approach described in the previous sections, we conduct classification experiments on the OpenMIIR data set. The OpenMIIR data set does not distinguish between training and test sets; therefore, we randomly selected 10% of the data set to use as the test data set. The experimental results show that our approach is superior to other current state-of-the-art approaches, with better classification accuracy. Figure 9 shows the 12-class

Table 1. Information about the tempo, meter, and length of the 12 classes short music fragments contained in the OpenMIIR.

ID	Name	Meter	Length	Tempo
1	Chim Chim Cheree (lyrics)	3/4	13.3 s	212 BPM
2	Take Me Out to the Ballgame (lyrics)	3/4	7.7 s	189 BPM
3	Jingle Bells (lyrics)	4/4	9.7 s	200 BPM
4	Mary Had a Little Lamb (lyrics)	4/4	11.6 s	160 BPM
11	Chim Chim Cheree	3/4	13.5 s	212 BPM
12	Take Me Out to the Ballgame	3/4	7.7 s	189 BPM
13	Jingle Bells	4/4	9.0 s	200 BPM
14	Mary Had a Little Lamb	4/4	12.2 s	160 BPM
21	Emperor Waltz	3/4	8.3 s	178 BPM
22	Hedwigs Theme (Harry Potter)	3/4	16.0 s	166 BPM
23	Imperial March (Star Wars Theme)	4/4	9.2 s	104 BPM
24	Eine Kleine Nachtmusik	4/4	6.9 s	140 BPM
	Mean		10.4 s	176 BPM

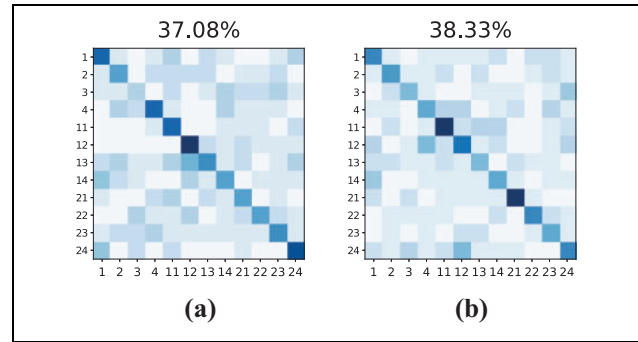


Figure 9. Confusion matrix of our approach with different encoders used in the experiments. (a) *OurVGG16*. (b) *OurVGG19*.

confusion matrix of the experimental results when different encoders are used in our ATL-based BCI framework.

Due to the design of our approach, it is possible to train a large-scale DNN using a limited EEG training data set by transferring knowledge from computer vision. As the baseline, we tested three recently proposed approaches: the support vector machine classifier described by Fan et al.,⁴⁰ the DNN described by Stober,³⁸ and the CNN described by Stober et al.⁴¹ Experiments on the OpenMIIR data set are conducted to compare the performance of our approach and that of the baseline approaches, and the results are shown in Table 2.

Rehabilitation experiment

We performed many experiments on the rehabilitation robot to verify the effectiveness of our ATL-based BCI approach. We selected two subjects for rehabilitation exercises. Both subjects were healthy men and performed rehabilitation exercises in a quiet and comfortable environment. One of the subjects was 30 years old and the other was 35 years old. Both subjects performed 10 positive rehabilitation processes for each BCI approach, and each of the processes included motor imaging 20 times. Each

Table 2. Classification accuracy on OpenMIIR data set and compared to the three baseline approaches.

	100%	50%	25%
SVC	23.10	16.69	9.83
DNN	27.22	20.83	12.47
CNN	27.80	19.20	8.55
OurVGG16	37.08	27.92	16.67
OurVGG19	38.33	30.00	16.67

SVC: support vector machine; DNN: deep neural network; CNN: convolutional neural network.

Table 3. The number of successful decoding and classification accuracy of limb movement images during positive rehabilitation exercise processes using our rehabilitation robot compared to the three baseline approaches.

	Open	Bend	Total (%)
SVC	151	155	76.50
DNN	159	162	80.25
CNN	166	161	81.75
OurVGG16	180	182	90.50
OurVGG19	184	181	91.25

SVC: support vector machine; DNN: deep neural network; CNN: convolutional neural network.

rehabilitation process interval lasted 10 min and each subject's experiment was spread over 5 days to ensure sufficient rest periods for the subject.

There are two types of movement in the positive rehabilitation exercises, called Open arm and Bend arm. We compared the effects of our approach with other baseline approaches, and the results are shown in Table 3 below.

Discussion

We can draw the following conclusions from the experimental results presented in the previous section:

- The experimental results shown in Figure 9 and Table 2 demonstrate that our proposed approach achieves accuracy that is clearly superior to the traditional approaches on public data set.
- Table 3 shows that our approach can achieve the best results in MI of rehabilitation exercise processes using our rehabilitation robot, thus helping patients with rehabilitation exercises effectively.
- VGG19 is a better choice than VGG16.

Conclusions

We propose a novel EEG signal classification approach based on an ATL-based framework in response to the serious problem of insufficient training data in EEG. It can be concluded that our approach is superior to other state-of-the-art approaches and provides a new perspective for

solving the problem of EEG classification. The ATL framework is a general-purpose transfer learning framework that can be used not only for BCI but also for other tasks, such as computer vision, natural language processing, and so on. Our approach can be viewed as a general bioelectrical signal classification framework that is suitable for other bioelectrical signals. In the future, we plan to develop an improved network based on state-of-the-art approaches in computer vision and other research fields. In addition, our rehabilitation robot should be tested on real patients and we plan to collaborate with hospitals in the future.

Acknowledgement

Thanks to the contributors of the open source software used in our system.


Declaration of conflicting interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: This work was jointly supported by National Natural Science Foundation of China under grants 91848206, 61621136008, and U1613212.

ORCID iD

Chuanqi Tan  <https://orcid.org/0000-0003-1197-4807>

References

1. Kaiser V, Kreiling A, Müller-Putz GR, et al. First steps toward a motor imagery based stroke BCI: new strategy to set up a classifier. *Front Neurosci* 2011; 5: 86.
2. Waldert S, Pistohl T, Braun C, et al. A review on directional information in neural signals for brain-machine interfaces. *J Physiol Paris* 2009; 103(3): 244–254.
3. Ramoser H, Müller-Gerking J, and Pfurtscheller G. Optimal spatial filtering of single trial EEG during imagined hand movement. *IEEE Trans Rehabil Eng* 2000; 8(4): 441–446.
4. Bashivan P, Rish I, Yeasin M, et al. Learning representations from EEG with deep recurrent-convolutional neural networks. *Comput Sci* 2015.
5. Tan C, Sun F, Zhang W, et al. Spatial and spectral features fusion for EEG classification during motor imagery in BCI. In: *2017 IEEE EMBS International Conference on Biomedical & Health Informatics (BHI)*, 16–19 February 2017, pp. 309–312. IEEE.
6. LeCun Y, Bengio Y, and Hinton G. Deep learning. *Nature* 2015; 521(7553): 436–444.
7. Mamoshina P, Vieira A, Putin E, et al. Applications of deep learning in biomedicine. *Mol Pharm* 2016; 13(5): 1445.
8. An X, Kuang D, Guo X, et al. A deep learning method for classification of EEG data based on motor imagery. In: *International Conference on Intelligent Computing*, 2014, pp. 203–210. Cham: Springer.

9. Cecotti H and Graser A. Convolutional neural networks for p300 detection with application to brain-computer interfaces. *IEEE Trans Pattern Anal Mach Int* 2011; 33(3): 433.
10. Soleymani M, Asghariesfeden S, Pantic M, et al. Continuous emotion detection using EEG signals and facial expressions. In: *IEEE International Conference on Multimedia and Expo*, 14–18 July 2014, pp. 1–6. IEEE.
11. Pan SJ and Yang Q. A survey on transfer learning. *IEEE Trans Knowl Data Eng* 2010; 22(10): 1345–1359.
12. Tan C, Sun F, Kong T, et al. A survey on deep transfer learning. In: Kůrková V, Manolopoulos Y, Hammer B, Iliadis L, and Maglogiannis I (eds) *International Conference on Artificial Neural Networks*, 2018, pp. 270–279. Cham: Springer.
13. Jayaram V, Alamgir M, Altun Y, et al. Transfer learning in brain-computer interfaces. *IEEE Comput Intell Magaz* 2016; 11(1): 20–31.
14. Tan C, Sun F, Zhang W, et al. Multimodal classification with deep convolutional-recurrent neural networks for electroencephalography. In: Liu D, Xie S, Li Y, Zhao D, and El-Alfy ES (eds) *International Conference on Neural Information Processing*, 2017, pp. 767–776. Cham: Springer.
15. Wei CS, Lin YP, Wang YT, et al. Selective transfer learning for EEG-based drowsiness detection. In: *2015 IEEE International Conference on Systems, Man, and Cybernetics (SMC)*, 9–12 October 2015, pp. 3229–3232. IEEE.
16. Wu D, Lance BJ, and Parsons TD. Collaborative filtering for brain-computer interaction using transfer learning and active class selection. *PloS One* 2013; 8(2): e56624.
17. Hajinoroozi M, Mao Z, Lin YP, et al. Deep transfer learning for cross-subject and cross-experiment prediction of image rapid serial visual presentation events from EEG data. In: *International Conference on Augmented Cognition*, 2017, pp. 45–55. Springer.
18. Zheng WL and Lu BL. Personalizing EEG-based affective models with transfer learning. In: *Proceedings of the Twenty-Fifth International Joint Conference on Artificial Intelligence*, New York, USA, 9–15 July 2016, pp. 2732–2738. AAAI Press.
19. Lin YP and Jung TP. Improving EEG-based emotion classification using conditional transfer learning. *Front Hum Neurosci* 2017; 11: 334.
20. Tan C, Sun F, and Zhang W. Deep transfer learning for EEG-based brain computer interface. In: *2018 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 15–20 April 2018, pp. 916–920. IEEE.
21. Tzeng E, Hoffman J, Darrell T, et al. Simultaneous deep transfer across domains and tasks. In: *Proceedings of the IEEE International Conference on Computer Vision*, 2015, pp. 4068–4076.
22. Huang JT, Li J, Yu D, et al. Cross-language knowledge transfer using multilingual deep neural network with shared hidden layers. In: *2013 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 26–31 May 2013, pp. 7304–7308. IEEE.
23. Long M, Wang J, Ding G, et al. Transfer feature learning with joint distribution adaptation. In: *Proceedings of the IEEE International Conference on Computer Vision*, 2013, pp. 2200–2207.
24. Long M, Cao Y, Wang J, et al. Learning transferable features with deep adaptation networks. In: *International Conference on Machine Learning*, pp. 97–105.
25. Long M, Wang J, and Jordan MI. Deep transfer learning with joint adaptation networks. arXiv preprint arXiv:160506636 2016.
26. Bousmalis K, Trigeorgis G, Silberman N, et al. Domain separation networks. In: *Advances in Neural Information Processing Systems*, 2016, pp. 343–351.
27. Tzeng E, Hoffman J, Saenko K, et al. Adversarial discriminative domain adaptation. In: *Computer Vision and Pattern Recognition (CVPR)*, 2017; 1: 4.
28. Gretton A, Borgwardt KM, Rasch M, et al. A kernel method for the two-sample-problem. In: *Advances in neural information processing systems*, 2007, pp. 513–520.
29. Long M, Wang J, Ding G, et al. Adaptation regularization: a general framework for transfer learning. *IEEE Trans Knowl Data Eng* 2014; 26(5): 1076–1089.
30. Kim B, Kim A, Dai H, et al. Generalizing over uncertain dynamics for online trajectory generation. In: Bicchì A and Burgard W (eds) *Robotics Research*, 2018, pp. 39–55. Cham: Springer.
31. Tommasi T, Patricia N, Caputo B, et al. A deeper look at dataset bias. In: Csurka G (eds) *Domain Adaptation in Computer Vision Applications*, 2017, pp. 37–55. Cham: Springer.
32. Eigen D, Puhrsch C, and Fergus R. Depth map prediction from a single image using a multi-scale deep network. In: *Advances in neural information processing systems*, pp. 2366–2374.
33. Cai D, He X, Han J, et al. Graph regularized nonnegative matrix factorization for data representation. *IEEE Trans Pattern Anal Mach Int* 2011; 33(8): 1548–1560.
34. Farneback G. Two-frame motion estimation based on polynomial expansion. In: *Scandinavian Conference on Image Analysis*, pp. 363–370.
35. Kalunga E, Djouani K, Hamam Y, et al. SSVEP enhancement based on canonical correlation analysis to improve BCI performances. In: *AFRICON*, 2013. 9–12 September 2013, pp. 1–5. IEEE.
36. Kumari M and Somani SB. Design and implementation of SSVEP based BCI using a canonical correlation analysis. *Int J Eng Sci* 2015; 6(4): 993–997.
37. Stober S, Sternin A, Owen AM, et al. Towards music imagery information retrieval: introducing the OpenMIIR dataset of EEG recordings from music perception and imagination. In: *ISMIR*, pp. 763–769.
38. Stober S. Learning discriminative features from electroencephalography recordings by encoding similarity constraints. In: *2017 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 5–9 March 2017, pp. 6175–6179. IEEE.
39. Guo Y, Liu Y, Oerlemans A, et al. Deep learning for visual understanding: a review. *Neurocomputing* 2016; 187: 27–48.
40. Fan RE, Chang KW, Hsieh CJ, et al. Liblinear: a library for large linear classification. *J Mach Learn Res* 2008; 9(Aug): 1871–1874.
41. Stober S, Sternin A, Owen AM, et al. Deep feature learning for EEG recordings. arXiv preprint arXiv:151104306 2015.