Personalized Sleep Staging Leveraging Source Free Unsupervised Domain Adaptation

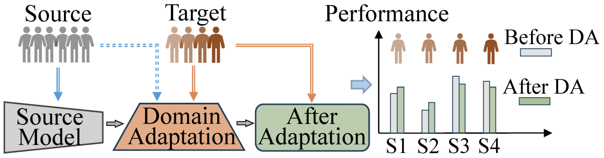
Yangxuan Zhou[1,2], **Sha Zhao**[1,2*], **Jiquan Wang**[1,2], **Haiteng Jiang**[3,4,1],
Shijian Li[1,2], **Benyan Luo**[5], **Tao Li**[3,4,1], **Gang Pan**[1,2,4]
[1]State Key Laboratory of Brain-machine Intelligence, Zhejiang University, Hangzhou, China
[2]College of Computer Science and Technology, Zhejiang University, Hangzhou, China
[3]Department of Neurobiology, Affiliated Mental Health Center & Hangzhou Seventh People's Hospital, Zhejiang University School of Medicine, Hangzhou, China
[4]MOE Frontier Science Center for Brain Science and Brain-machine Integration, Zhejiang University, Hangzhou, China
[5]The First Affiliated Hospital, College of Medicine, Zhejiang University, Hangzhou, China
{zyangxuan, szhao, wangjiquan, h.jiang, shijianli, luobenyan, litaozjusc, gpan}@zju.edu.cn

Abstract

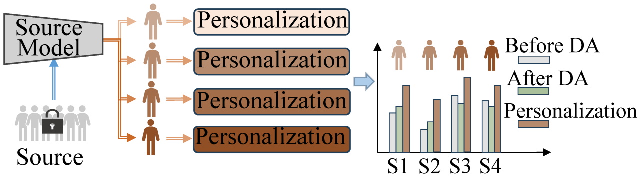
Sleep staging is important for monitoring sleep quality and diagnosing sleep-related disorders. Recently, numerous deep learning-based models have been proposed for automatic sleep staging using polysomnography recordings. Most of them are trained and tested on the same labeled datasets which results in poor generalization to unseen target domains. However, they regard the subjects in the target domains as a whole and overlook the individual discrepancies, which limits the model's generalization ability to new patients (i.e., unseen subjects) and plug-and-play applicability in clinics. To address this, we propose a novel Source-Free Unsupervised Individual Domain Adaptation (SF-UIDA) framework for sleep staging, leveraging sequential cross-view contrasting and pseudo-label based fine-tuning. It is actually a two-step subject-specific adaptation scheme, which enables the source model to effectively adapt to newly appeared unlabeled individual without access to the source data. It meets the practical needs in real-world scenarios, where the personalized customization can be plug-and-play applied to new ones. Our framework is applied to three classic sleep staging models and evaluated on three public sleep datasets, achieving the state-of-the-art performance.

Introduction

Sleep plays a crucial role in people's lives and has a significant impact on their overall well-being (Humphreys, Sharps, and Campbell 2005). Sleep staging is important for monitoring sleep quality and serves as a valuable tool to help diagnose sleep disorders(Wang et al. 2023), which refers to classify sleep periods into different stages. Recently, Polysomnography (PSG) has been widely used for sleep staging in clinics, which records various physiological signals by the sensors attached to different parts of one body, such as electroencephalography (EEG), electrooculography (EOG), and electromyography (EMG). The PSG recordings are usually divided into consecutive epochs of 30s. Experts manually identify each epoch into five distinct sleep stages, namely, W, N1, N2, N3, REM, according to the American Academy of Sleep Medicine (AASM) (Berry et al. 2012). Obviously, such process is subjective and time-consuming.



(a) Traditional Domain Adaptation



(b) Personalized Customization

Figure 1: (a) Traditional Domain Adaptation for a group of target (new) individuals and (b) Source-free Unsupervised Personalized Customization for each target (new) individual.

Numerous deep learning-based models have been proposed for automatic sleep staging (Tsinalis et al. 2016; Mousavi, Afghah, and Acharya 2019; Zhou et al. 2024; Wang et al. 2024c) in recent years. Such models are usually trained and tested on the same labeled source data, automatically classifying different stages. They have achieved good performance for sleep staging in the valid set, however, their performance is not so satisfying on unknown samples. Most of them overlook the **individual discrepancies**, such as physiological structures (Matsushima, Minami, and Takadama 2012), physical characteristics (i.e., Electrodermal Response) and sleep habits. In clinical practice, the new patients (unknown subjects) are probably different from the samples used for training the models offline. This limits the model's generalization ability to unknown subjects and dramatically degrades the performance, especially when they are directly tested on the new patients in clinics. Therefore, we desperately need to make the model adapt to new subjects. Unsupervised Domain Adaptation (UDA) is a suitable method which can transfer knowledge

---

[*]Corresponding authors
Copyright © 2025, Association for the Advancement of Artificial Intelligence (www.aaai.org). All rights reserved.

learned from a labeled dataset (i.e., source domain) to an unlabeled dataset (i.e., target domain). Traditional UDA approaches commonly apply a joint training strategy, which relies on not only the target but also the source data to mitigate domain shift through feature alignment. By doing so, the performance of some target subjects is improved after adaptation (e.g., S1, S2 subjects in Fig. 1 (a)). However, this adapting way will lead to the following drawbacks. First, the target domain is treated as a whole (i.e., a number of subjects), which requires waiting until a batch of target domain samples are available before conducting the adaptation. This is impractical in real life, where the arrival of each new individual is entirely random. **We need a plug-and-play adaptation rather than waiting for all the target individuals to arrive to conduct adaptation**. Second, they overlook the individual discrepancies, resulting in a failure to adapt to certain special ones whose distribution deviates significantly from the overall distribution (e.g., S3, S4 subjects in Fig. 1 (a)). **A personalized sleep staging model is needed for each new patient**. Third, using source data for joint-training is time-consuming and can lead to data privacy leakage. **A source-free adaptation is needed in clinical practice.**

To meet such practical needs, in this paper, we propose an **S**ource-**F**ree **U**nsupervised **I**ndividual **D**omain **A**daptation framework for automatic sleep staging, named SF-UIDA. First, we introduce **the concept of individual domains**, where the SF-UIDA framework treats each target subject as a distinct target domain. Meanwhile, the proposed SF-UIDA contains a two-step **subject-specific strategy** that considers individual discrepancies to mitigate their impact on the source model. It can be applied to the source pretrained models and enable them to adapt to each individual in a personalized manner, without the need to wait for all the test individuals to arrive. Moreover, SF-UIDA also adopts a source-free UDA strategy, which is a more practical setup that does not require accessing the source data, thereby lowering the time cost and protecting the data privacy. As shown in Fig. 1 (b), our framework enables the source model to rapidly adapt to each new target in a personalized and plug-and-play fashion without accessing the source data.

Our SF-UIDA is evaluated on three public datasets. The process of customization is efficient in time and acceptable in clinics, only taking a short amount of time to transform the source model into a personalized model. Besides, it is worth pointing that our SF-UIDA framework can be easily implemented without any modification to the source model structure, achieving plug-and-play application in practice. Our contributions are as follows:

- We devise a novel Source-Free Unsupervised Individual Domain Adaptation framework for automatic sleep staging, named SF-UIDA. It meets the practical needs for a plug-and-play personalized customization that can be applied to each newly appeared individual without accessing the source data.
- We propose a two-step subject-specific alignment strategy to mitigate the impact of individual discrepancies. It effectively transforms the source model into a personalized model for each new individual within a short adaptation time.
- Our SF-UIDA achieves the best generalization performance across three public sleep staging datasets, compared to other methods.

Related Work

Automatic Sleep Staging

There have been many models proposed for automatic sleep staging in recent years. Some studies commonly employ convolutional networks to extract local sleep features. For example, U-time (Perslev et al. 2019, 2021) is a fully CNN network based on the U-net architecture that can excellently model sleep-related features. SalientSleepNet (Jia et al. 2021) is also a fully CNN network based on U^2-net which can capture multi modal sleep feature. Considering the advantage of capturing long-term temporal information, there are also some studies utilizing recurrent neural networks (Mikolov et al. 2010) or transformer encoders (Vaswani et al. 2017) for sleep staging. Phan et al. (2022) proposed SleepTransformer, a transformer-based sequence-to-sequence model that improves the interpretability of the sleep-staging task. Although these models obtain good performance for sleep staging, they have not taken individual discrepancies into account, leading to dramatically degraded performance when applied to target domains (i.e., unknown subjects). In this paper, we propose a new training strategy to address individual discrepancies for sleep staging, so as to improve the model generalization ability in practice.

Source-free Unsupervised Domain Adaptation

To address the challenge of model generalization on unseen data, some studies have employed UDA methods to facilitate knowledge transfer between the source and target domains. Existing UDA studies (Tang et al. 2022; Wang et al. 2024a) have reduced domain shift and extracted domain-invariant features through distance-based alignment. Fan et al. (2022) utilized statistical alignment to mitigate domain shift across several sleep datasets. These UDA methods effectively enhance the generalization ability of the source model for sleep staging. However, they rely on source data for joint training, which is impractical and time-consuming for each new subject in real-life scenarios. In contrast, source-free UDA (Liang, Hu, and Feng 2020) offers a more practical solution by eliminating the need for source data during adaptation. Contrastive learning (CL) (Oord, Li, and Vinyals 2018; Chen et al. 2020; Chen and He 2021) is commonly used in source-free UDA, focusing on mining intrinsic representation features within the data. Chen et al. (2022) applied a self-supervised CL method to facilitate target feature learning and achieve test-time adaptation, while Wang et al. (2022) utilized weight-averaged and augmentation-averaged predictions to generate pseudo-labels for adaptation. These source-free approaches enable rapid adaptation of the source model to target domains. However, they often overlook individual discrepancies and treat the entire target domain, encompassing multiple subjects, as a single distribution for adaptation. This can hinder accurate predictions for individuals whose distributions significantly deviate from the over-

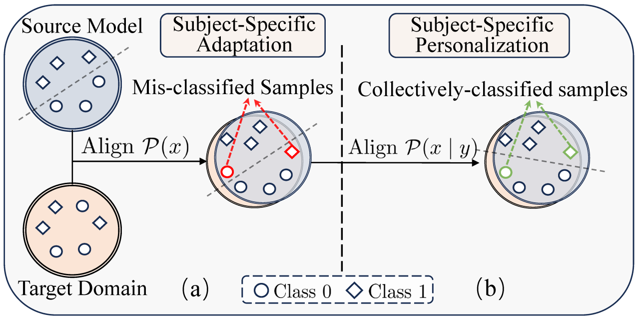


Figure 2: Illustration of the two-step alignment strategy in the SF-UIDA Framework: aligning the marginal probability distribution $\mathcal{P}_{\mathcal{T}}(x)$ and the class conditional probability distribution $\mathcal{P}_{\mathcal{T}}(x \mid y)$ for individual target domains

all distribution. To address this, we propose the SF-UIDA framework, which accounts for individual discrepancies and enables personalized customization for each individual.

Methodology

Problem Formulation

In this work, we try to address the issue of individual discrepancy for the task of automatic sleep staging by employing an UDA-based approach. Here, we introduce the concept of individual target domain, which consists of the recordings from only one subject. Formally, given a labeled source domain $\mathcal{D}_{\mathcal{S}}=\{\mathcal{X}_{\mathcal{S}}^{i}, \mathcal{Y}_{\mathcal{S}}^{i}\}_{i=1}^{\mathcal{N}_{\mathcal{S}}}$ with $\mathcal{N}_{\mathcal{S}}$ subjects and an unlabeled individual target domain $\mathcal{D}_{\mathcal{T}}=\{\mathcal{X}_{\mathcal{T}}^{j}\}_{j=1}^{\mathcal{N}_{\mathcal{T}}}$ with $\mathcal{N}_{\mathcal{T}} = 1$. We denote the distributions of different domains as $\mathcal{P}_{\mathcal{S}}(x, y)$ and $\mathcal{P}_{\mathcal{T}}(x, y)$ respectively, where $\mathcal{P}_{\mathcal{S}}(x, y) \neq \mathcal{P}_{\mathcal{T}}(x, y)$. We employ the sleep sequence $\mathcal{X}_{\mathcal{S}}=(x_1^{\mathcal{S}}, x_2^{\mathcal{S}}, x_3^{\mathcal{S}}, ..., x_L^{\mathcal{S}})$ of length L and its corresponding label $\mathcal{Y}_{\mathcal{S}}=(y_1^{\mathcal{S}}, y_2^{\mathcal{S}}, y_3^{\mathcal{S}}, ..., y_L^{\mathcal{S}})$ from $\mathcal{D}_{\mathcal{S}}$ as inputs, where $x_i^{\mathcal{S}}$, $y_i^{\mathcal{S}}$ denotes the data and label of $i\text{-}th$ epoch in the sequence $\mathcal{X}_{\mathcal{S}}$. Our main purpose is to accurately predict the label $\mathcal{Y}_{\mathcal{T}}=(y_1^{\mathcal{T}}, y_2^{\mathcal{T}}, y_3^{\mathcal{T}}, ..., y_L^{\mathcal{T}})$ of unlabeled individual target domain sequential sample $\mathcal{X}_{\mathcal{T}}=(x_1^{\mathcal{T}}, x_2^{\mathcal{T}}, x_3^{\mathcal{T}}, ..., x_L^{\mathcal{T}})$.

Overview

In this work, we first pretrain the source model from a labeled source domain $\mathcal{D}_{\mathcal{S}}$. The probability distribution of the source domain can be described as follows.

$$\mathcal{P}_{\mathcal{S}}(x, y) = \mathcal{P}_{\mathcal{S}}(x)\mathcal{P}_{\mathcal{S}}(y \mid x) = \mathcal{P}_{\mathcal{S}}(y)\mathcal{P}_{\mathcal{S}}(x \mid y) \tag{1}$$

Considering the individual discrepancies among each subject, our SF-UIDA framework contains a two-step alignment: **subject-specific adaptation** and **subject-specific personalization**, to align each of the individual target distribution $\mathcal{P}_{\mathcal{T}}$ with the source domain distribution $\mathcal{P}_{\mathcal{S}}$, customizing the sleep staging model for each individual. Specifically, we use the unlabeled data from each individual target domain and the generated pseudo-labels to make the source model's probability distribution $\mathcal{P}_{\mathcal{S}}(x, y)$ align with the individual target domain's probability distribution $\mathcal{P}_{\mathcal{T}}(x)$ and class-conditional probability distribution $\mathcal{P}_{\mathcal{T}}(x \mid y)$ illustrated in Fig. 2. The whole process can be totally divided into three training stages: **Source Model Pretraining**: We employ several classical sleep stage classification (SSC) models as the pretraining models to learn the general sleep features from the source domain. **Subject-Specific Adaptation**: We propose a subject-specific adaptation by proposing a sequential cross-view prediction task on individual target domain. It is used to capture subject-specific sleep representations and align with the individual target domain's probability distribution $\mathcal{P}_{\mathcal{T}}(x)$, mitigating the impact of individual discrepancies on the pretrained source model. **Subject-Specific Personalization**: We employ a teacher model based pseudo-labeling strategy for fine-tuning, so as to learn the fine-grained distribution of different classes in individual target domains. It enables the source model further align the class-conditional probability distribution $\mathcal{P}_{\mathcal{T}}(x \mid y)$, thus achieving personalized customization. **Notably, our model customization is absolutely source-free, using only the unlabeled target data for personalization.**

Source Model Pretraining

We first extract general characteristics from PSG recordings by pretraining a model on the source domain. Specifically, we employ three classical lightweight sleep staging models as pretrained models. Each model has a specialized extractor for sleep features and a temporal encoder to capture temporal information from sleep sequences. The parameters of the source model are then transferred to the subsequent process.

Subject Specific Adaptation

Due to the independence of individuals in the test set and inter-individual differences, each individual target domain exhibits a distinct distribution, often significantly deviating from the source domain distribution. Therefore, the pretrained source model struggles to generalize to individual target domains. **In this step, our objective is to reduce domain shift and mitigate the impact of individual discrepancies without access to the source data.** Inspired by the CPC and TS-TCC (Oord, Li, and Vinyals 2018; Eldele et al. 2021) algorithms, we propose a subject-specific adaptation scheme for unsupervised domain adaptation, aligning the probability distribution $\mathcal{P}_{\mathcal{T}}(x)$ through a complex sequential cross-view prediction task. This approach enables us to model specific representations for each individual.

Sequential Cross View Contrasting According to the AASM(Berry et al. 2012), the transition patterns of sleep stages between neighboring epochs are crucial for accurate sleep staging. **These transitions occur not only in the forward direction (e.g., W→N1→N2→N3) but also in reverse (e.g., REM←N3←N2).** Motivated by these patterns, we propose a novel **S**equential **C**ross-view **C**ontrasting (**SCC**) module to model the bidirectional transition relationships within subject-specific sequences, as illustrated in Fig. 3. To generate a new augmented view, reversing the original sequence effectively contrasts the temporal relationships

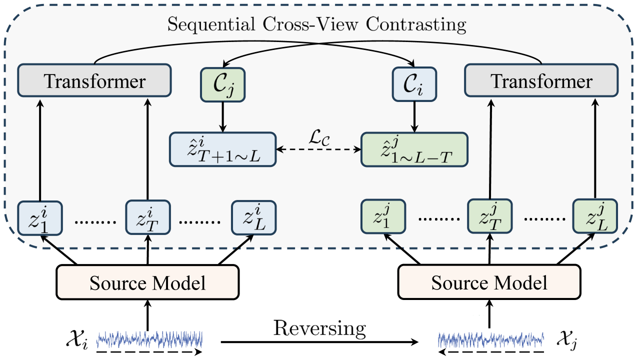


Figure 3: The architecture of the proposed SCC module.

between different views. Formally, given an input sleep sequence $\mathcal{X}_i$=$(x_1, x_2, ..., x_{L-1}, x_L)$, its augmented view $\mathcal{X}_j$=$(x_L, x_{L-1}, ..., x_2, x_1)$ is obtained by reversal. After feeding $\mathcal{X}_i$ and $\mathcal{X}_j$ into the Feature Extractor and Feature Encoder, we obtain their corresponding latent representations $\mathcal{Z}_i$=$(z^i_1, z^i_2, ..., z^i_{L-1}, z^i_L)$ and $\mathcal{Z}_j$=$(z^j_L, z^j_{L-1}, ..., z^j_2, z^j_1)$, respectively. For a given time step T $(1 < T < L)$, we utilize a transformer as an autoregressive model to encode $\mathcal{Z}_{i:t\leq T}$, $\mathcal{Z}_{j:L-T\leq t \leq L}$ into a contextual vector: $\mathcal{C}_i$ and $\mathcal{C}_j$. We then establish a sequential cross-view task, using linear layers to predict the future $L-T$ sleep timesteps from z^i_{T+1} to z^i_L in sequence $\mathcal{Z}_i$ by leveraging contextual vector $\mathcal{C}_j$, such that $\hat{z}^i_{T+k} = f_{T+k}(\mathcal{C}_j)$, where $\hat{z}^i_{T+k}$ denotes the predicted timesteps for z^i_{T+k} and f_{T+k} is the corresponding predicting linear layer. Similarly, we use contextual vector $\mathcal{C}_i$ to predict the past $L-T$ sleep timesteps from z^j_1 to z^j_{L-T} in sequence $\mathcal{Z}_j$. The cross-view predicted timesteps can be formulated as $\hat{\mathcal{Z}}^i_{T+1\sim L} = (f_{T+1}(\mathcal{C}_j), f_{T+2}(\mathcal{C}_j), ..., f_L(\mathcal{C}_j)) = (\hat{z}^i_{T+1}, \hat{z}^i_{T+2}, ..., \hat{z}^i_L)$ and $\hat{\mathcal{Z}}^j_{1\sim L-T} = (f_1(\mathcal{C}_i), f_2(\mathcal{C}_i), ..., f_{L-T}(\mathcal{C}_i)) = (\hat{z}^j_1, \hat{z}^j_2, ..., \hat{z}^j_{L-T})$, respectively. We then apply the Maximum Mean Discrepancy (MMD) loss to minimize the distance between the cross-view predicted timesteps $\hat{\mathcal{Z}}_i$ and $\hat{\mathcal{Z}}_j$ as follows:

$$\mathcal{L}_\mathcal{C} = -\frac{1}{\mathcal{K}} \sum_{k=1}^{\mathcal{K}} \mathcal{D}_{\mathcal{MMD}}(\hat{z}^i_{T+k}, \hat{z}^j_k) \tag{2}$$

where $\mathcal{K}$ denotes the $L-T$. Notably, for a specific predicted timestep $\hat{z}^j_k$, we minimize its MMD distance between the corresponding predicted timestep $\hat{z}^i_{T+k}$ rather than $\hat{z}^i_k$. This is because after reversing, the position corresponding to z^i_{T+k} aligns perfectly with z^j_k.

Algorithm 1: Source-Free Unsupervised Individual Domain Adaptation algorithm

```
Input: $\mathcal{X}_\mathcal{S}$, $\mathcal{Y}_\mathcal{S}$, $\mathcal{X}_\mathcal{T}$
Output: $\mathcal{F}_\theta$
Source Model Pretraining:
Pretrain the source model by $\mathcal{X}_\mathcal{S}$, $\mathcal{Y}_\mathcal{S}$ and return $\mathcal{F}_\theta$.
Subject-Specific Adaptation:
Generate the augmented view $\mathcal{X}_j$ for each input
  target sleep sequence $\mathcal{X}_i$.
for $i = 1$ to $n$ do
    Compute latent representations $\mathcal{Z}_i$, $\mathcal{Z}_j$.
    Compute contextual vectors $\mathcal{C}_i$, $\mathcal{C}_j$.
    Optimize $\mathcal{F}_\theta$ by minimizing Eq.2.
end
return retrained model $\mathcal{F}_\theta$.
Subject-Specific Personalization:
Initialize the teacher model $\mathcal{F}_{\theta_T}$ and transfer the
  weights $W_\theta$ to $W_{\theta_T}$.
for $i = 1$ to $n$ do
    Generate confident pseudo sequence labels $\hat{\mathcal{Y}}_\mathcal{T}$
      by $\mathcal{F}_{\theta_T}$ by Eq. 4, Eq. 5.
    Optimize $\mathcal{F}_\theta$ by minimizing Eq.6.
    Update $W_{\theta_T}$ by Eq.3
end
return $\mathcal{F}_\theta$.
```

Subject Specific Personalization

The subject-specific adaptation aligns the marginal distributions between the source domain and the individual target domain. Due to the influence of individual differences, there still exist class-conditional distribution discrepancies between each individual target domain and the source domain, which may lead to erroneous alignment of different classes across the source and target domains shown in Fig.2 (a). Conventional solutions typically rely on fine-tuning with labeled data to address this issue (Eldele et al. 2023). However, we are unable to employ supervised methods for class-conditional distribution alignment as the target labels are unavailable. Inspired by Tarvainen and Valpola (2017) and Ragab et al. (2022), we employ a teacher model based on pseudo-label generation approach to tackle this problem shown in Fig.2(b). Notably, we have introduced **sequence confidence** for each individual, producing robust pseudo sequence labels. We solely preserve the confident ones for further fine-tuning, which enable us to better align the class-conditional distribution with the source domain and model the personalized representations. Specifically, we migrate the model parameters $\mathcal{W}_\theta$ to the teacher model $\mathcal{F}_{\theta_T}$ using Exponential Moving Average (EMA). The updates to the teacher model parameters are as follows:

$$\mathcal{W}_{\theta_T} = \alpha \mathcal{W}_{\theta_T} + (1-\alpha)\mathcal{W}_\theta \tag{3}$$

where $\mathcal{W}_{\theta_T}$ denotes the parameters of the teacher model $\mathcal{F}_{\theta_T}$ and α is a hyper-parameter employed to regulate the update rate of the teacher model parameters. For each sleep sequence $\mathcal{X}_\mathcal{T}$=$(x_1, x_2, x_3, ..., x_L)$, we can obtain the corresponding predicted sequence probabilities $\mathcal{Y}_\mathcal{T}$=$(y_1, y_2, y_3, ..., y_L)$ by the teacher model. Given a sleep sequence $\mathcal{X}_\mathcal{T}$ with the length L, we retain it for subsequent fine-tuning iff there are $\mathcal{N}_c$ or more epochs in the sequence and the prediction probabilities of each epoch not less than the confidence threshold ξ. It can be formalized as follows:

$$\mathcal{Y}_\mathcal{T} = softmax(\mathcal{F}_\theta(\mathcal{X}_\mathcal{T})) \tag{4}$$

Dataset	Subjects (all)	Subjects (we choose)	CV (fold)	Sampling (Hz)	Scoring (standard)	EOG Channels (we choose)	EEG Channels (we choose)
ISRUC	100	98	10	200	AASM	E1, E2	F3, F4, C3, C4, O1, O2
HMC	151	145	10	256	AASM	E1, E2	F4, C4, O2, C3
SleepEDF-153	78	78	10	100	R&K	EOG horizontal	Fpz-Cz, Pz-Oz

Table 1: A brief description about three public sleep staging datasets.

$$\hat{\mathcal{Y}_{\mathcal{T}}} = \mathbb{1}_s([\sum_{i=1}^{L} \mathbb{1}_e(max(y_i^{\mathcal{T}}) > \xi)] \geq \mathcal{N}_c) \cdot \mathcal{Y}_{\mathcal{T}} \quad (5)$$

where $\mathbb{1}_e$ is the confident epoch indicator function, evaluating to 1 iff $max(y_t^k) > \xi$ and $\mathbb{1}_s$ is the confident sequence indicator function, evaluating to 1 iff there are $\mathcal{N}_c$ or more confident epochs $y_i^{\mathcal{T}}$ in the $\mathcal{Y}_{\mathcal{T}}$. In most existing studies, the confidence threshold ξ is set to be greater than 0.9. However, based on our confident sequence setting, we focus more on the overall confidence of the sequence rather than that of a single epoch. So we need a higher tolerance for the threshold ξ and we set the it equal to 0.8 and the $\mathcal{N}_c$ is set to 15. To ensure alignment of the class-conditional distribution, we employ the confident sequence labels to finetune the model, by using a cross-entropy loss:

$$\mathcal{L}_{ce} = -\mathbb{E}_{\mathcal{X}_{\mathcal{T}} \sim \mathcal{P}_{\mathcal{T}}}[\sum_{k=1}^{K} \mathcal{Y}_{\mathcal{T}}^k log(\hat{\mathcal{Y}_{\mathcal{T}}^k})] \quad (6)$$

During the whole process, only the unlabeled target individual's data is needed for personalized customization without access to the source data. The algorithm of SF-UIDA framework is illustrated in Algorithm 1.

Experiments

Datasets

As shown in Table 1, we evaluated our approach on three publicly available datasets: Sleep-EDF (Kemp et al. 2000), ISRUC (Khalighi et al. 2016), and HMC (Alvarez-Estevez and Rijsman 2021). For each dataset, we utilized both EEG and EOG channels as input. **ISRUC**: This public database consists of three sub-groups. We specifically selected subgroup 1, which includes all-night polysomnography (PSG) recordings from 100 adults, totaling 86,400 samples. Subjects 8 and 40 were excluded due to missing channels. **SleepEDF-153**: A public Physionet database comprising 78 healthy subjects aged 25 to 101, containing 188,760 samples. All subjects' recordings were used for evaluation. **HMC**: This public dataset includes recordings from 151 subjects at the Haaglanden Medisch Centrum (The Netherlands), consisting of 129,440 samples. Subjects 14, 32, 33, 64, 112, and 135 were excluded due to missing channels. All sleep recordings were bandpass filtered (0.3 Hz–35 Hz) and resampled to 100 Hz.

Settings

Baseline Models We need to select baseline models to evaluate our SF-UIDA framework. Considering the efficiency requirement in clinical practice, the fine-tuning for each individual cannot be time-consuming. Therefore, **we selected three lightweight sleep staging models from existing studies**, each of which is comprised of the feature extractor and the temporal encoder: **DeepSleepNet** (Supratak et al. 2017): a classical CNN-BiLSTM model for extracting sleep features and learning transition rules. **TinySleepNet** (Supratak and Guo 2020): a more lightweight model based on the DeepSleepNet. **RecSleepNet** (Nie, Tu, and Xu 2021): a CNN-LSTM model based on feature representation reconstruction. Here, we do not choose some other sleep staging models, which are also classical and perform well, such as UTime (Perslev et al. 2019), SalientSleepNet (Jia et al. 2021), CareSleepNet(Wang et al. 2024b) or SleepTransformer (Phan et al. 2022), because their network structures are complex and not suitable for our problem in clinics.

Implementation Based on the publicly available source code, we re-implemented the three baseline models using pytorch. The experimental settings are as follows: **Source Model Pretraining**: The pretraining epoch is set to 100. The learning rate is set to 1e-4. **Subject-Specific Adaptation**: The training epoch of this stage is set to 5. The learning rate is set to 1e-7. The time step T is set to 17. **Subject-Specific Personalization**: The fine-tuning epoch is set to 10. The learning rate is set to 1e-7. The momentum α is set to 0.996. We use the Adam optimizer to train the model, the β is set to [0.5,0.99], the weight decay is set to 3e-4, the size of mini-batch is set to 32. The model is trained on a single machine equipped with an Intel Core i9 10900K CPU and eight NVIDIA RTX 3080 GPUs. The source code is publicly available[1].

Performance Measurement We employ 10-fold cross-validation (CV) to evaluate the performance of our approach across three different datasets. Different from the conventional settings in previous studies, which only included training and validation sets, **we divided the dataset into the training, validation, and test sets and the ratio is 8:1:1**. The test set is regarded as unknown subjects, **where there are no repetitive individuals in the test set, ensuring that each individual appears only once in the test set throughout the 10-fold CV experiment**. In each fold, we employ the training and validation sets to pretrain the source model. Subsequently, **the source model goes though personalization customization on each individual in the test set.** Finally, we compute the average metrics for each individual in the test set. We employ Accuracy (ACC) and Macro-F1 score (MF1) as evaluation metrics.

---

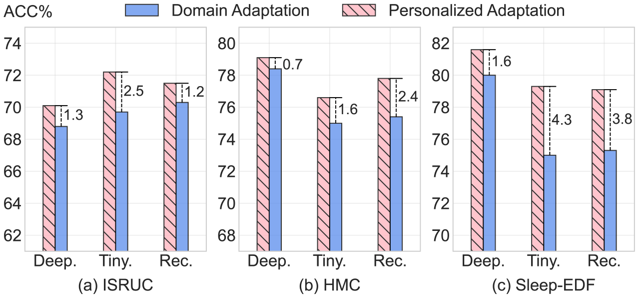
[1] https://github.com/xiaobaben/SF-UIDA

	ISRUC						HMC						SleepEDF						Average	
	Deep.		Tiny.		Rec.		Deep.		Tiny.		Rec.		Deep.		Tiny.		Rec.			
	ACC	MF1	ACC	MF1	ACC	MF1	ACC	MF1	ACC	MF1	ACC	MF1	ACC	MF1	ACC	MF1	ACC	MF1	ACC	MF1
Source Only	59.7	49.7	66.9	57.2	64.1	57.3	71.8	66.0	72.2	64.5	74.0	66.8	51.8	35.5	61.4	46.7	73.6	61.2	66.2	56.1
CPC	67.3	62.1	69.3	62.0	69.0	63.1	78.1	74.0	74.9	68.3	76.4	70.2	78.5	69.2	77.8	65.2	78.0	66.1	74.4	66.7
SimSiam	68.4	63.5	69.2	62.6	69.4	62.8	78.2	73.8	75.7	69.4	76.6	70.9	80.6	71.0	77.9	64.8	78.1	66.2	74.9	67.2
Adacontrast	68.6	63.7	69.7	63.2	69.6	62.9	78.1	73.7	75.4	69.2	76.8	71.1	80.5	70.9	77.6	64.5	77.9	65.7	74.9	67.2
CoTTA	67.8	62.8	70.2	62.9	69.3	62.8	77.7	73.5	75.0	68.3	76.4	70.2	80.2	70.7	77.9	65.4	78.0	66.1	74.7	67.0
C-SFDA	68.6	63.7	70.0	63.5	70.4	63.4	78.1	73.7	75.5	69.4	76.8	71.1	80.3	70.8	77.6	64.4	77.9	65.6	75.0	67.3
Ours	**70.1**	**64.7**	**72.2**	**65.1**	**71.5**	**64.6**	**79.1**	**75.0**	**76.6**	**70.4**	**77.8**	**72.3**	**81.6**	**72.2**	**79.3**	**66.5**	**79.1**	**67.4**	**76.4**	**68.7**

Table 2: Performance comparison with existing source-free UDA methods. Notably, Deep., Tiny., and Rec. refer to DeepSleepNet, TinySleepNet, and RecSleepNet, respectively.

Result Analysis

Compared with Other Existing Methods We compare our method with other classical source-free UDA methods in sleep staging, to further investigate the generalization ability of SF-UIDA. **Source only** : A method to directly test the individual target domain using source model. **CPC** (Oord, Li, and Vinyals 2018): A classical contrastive learning approach to learn representation of time-seires data by predicting the future timesteps. **SimSiam** (Chen and He 2021): An efficient contrastive learning approach, which focuses on representation learning using stop-gradient strategy and symmetrized loss. **Adacontrast** (Chen et al. 2022): A test-time adaptation method using contrastive learning to facilitate target feature learning. **CoTTA** (Wang et al. 2022): A test-time adaptation method , which can effectively adapt off-the-shelf source pretrained models to target domains. **C-SFDA** (Karim et al. 2023): A curriculum learning aided self-training framework for SFDA is designed to adapt efficiently and reliably to target domains. We implement these SFDA methods within our framework, and the performance comparison is shown in Table 2. Our method outperforms existing approaches, underscoring the effectiveness of our personalized customization strategy in enhancing overall performance. Compared to the Source Only method, our approach significantly enables the source model to adapt to individual target domains, achieving model customization and improved performance. Among the compared methods, it is worth noting that methods leveraging contrastive learning(e.g. SimSiam, Adacontrast, C-SFDA) exhibit better performance compared to other approaches (e.g. CPC, CoTTA).



(a) ISRUC (b) HMC (c) Sleep-EDF

Figure 4: Comparison with traditional domain adaptation paradigm (i.e., non-personalized domain adaptation).

Compared with Non-Personalized Domain Adaptation The traditional domain adaptation (DA) paradigm considers a batch of subjects as the target domain for subsequent adaptation which is impractical in real life. In contrast, our individual domain adaptation based method allows for plug-and-play application on each new subject without waiting. To evaluate the effectiveness of our individual DA setting compared to traditional DA paradigm, we conducted a comparative study. Notably, we maintained the consistent partitioning of the 10-fold cross validation and the SF-UIDA framework. The only difference is: **for the traditional DA paradigm, in each fold we use the data of all target individuals (i.e., test set) to perform adaptation, rather than conducting model customization for each individual separately.** We evaluated the performance of the fine-tuned model on each test individual shown in Fig. 4. Compared to the traditional domain adaptation paradigm, our personalized adaptation based paradigm can achieve comprehensive superiority in performance on all baseline models across the three datasets. It proves that our method can not only meet the practical needs in real-world scenarios, where the personalized customization can be plug-and-play applied to new individuals, but also better achieve individual performance improvement after adaptation.

Ablation Study To investigate the importance of our proposed two-step alignment strategy, we conducted this ablation experiments. The model variants are defined as follows:

- **SO**: which means source only that we directly use the source model for testing.
- **SO+SSA**: only the **S**ubject-**S**pecific **A**daptation (SSA) stage is preserved in the SF-UIDA framework.
- **SO+SSP**: only the **S**ubject-**S**pecific **P**ersonalization (SSP) stage is preserved in the SF-UIDA framework.
- **SO+SSA+SSP**: we employed the full two-step alignment process of the SF-UIDA framework.

As shown in Table 3, our ablation experiments convincingly demonstrate the efficacy of the proposed two-step alignment

	ISRUC						HMC						SleepEDF						Average	
	Deep.		Tiny.		Rec.		Deep.		Tiny.		Rec.		Deep.		Tiny.		Rec.			
	ACC	MF1	ACC	MF1	ACC	MF1	ACC	MF1	ACC	MF1	ACC	MF1	ACC	MF1	ACC	MF1	ACC	MF1	ACC	MF1
SO	59.7	49.7	66.9	57.2	64.1	57.3	71.8	66.0	72.2	64.5	74.0	66.8	51.8	35.5	61.4	46.7	73.6	61.2	66.2	56.1
SO+SSA	69.0	63.8	71.3	63.4	70.3	63.5	77.5	73.1	75.5	69.0	76.6	70.7	79.4	69.8	76.1	62.8	77.2	64.9	74.8	66.8
SO+SSP	68.0	62.4	70.9	64.7	68.0	62.4	78.4	74.3	74.5	70.3	77.7	71.9	79.5	69.9	78.9	66.4	79.1	67.4	75.0	67.7
SO+SSA+SSP	**70.1**	**64.7**	**72.2**	**65.1**	**71.5**	**64.6**	**79.1**	**75.0**	**76.6**	**70.4**	**77.8**	**72.3**	**81.6**	**72.2**	**79.3**	**66.5**	**79.1**	**67.4**	**76.4**	**68.7**

Table 3: Ablation experiment overview.

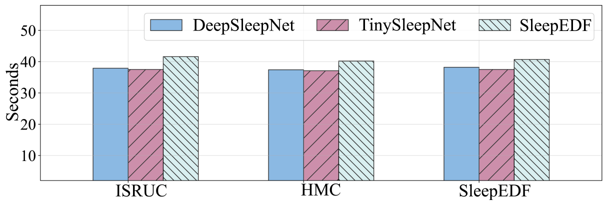
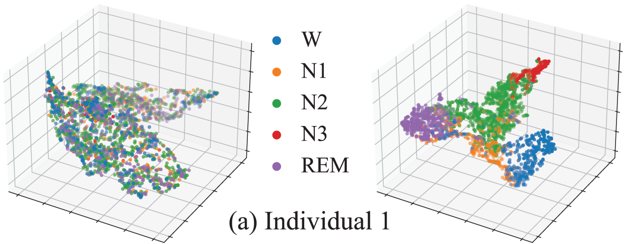


Figure 5: Average time cost per individual (seconds).

methodology for unsupervised individual domain adaptation. Compared to the source-only method, both the SSA and SSP modules yield improved performance, highlighting the effectiveness of our alignment strategy during model customization. Notably, the SSP module outperforms the SSA module slightly, suggesting that aligning the class-conditional distributions of the individual target domain is more impactful than aligning the marginal distribution. This is reasonable given the imbalance in sleep data classes, as evidenced by a greater performance gap on the MF1 metric (67.7% vs. 66.8%) compared to the ACC metric (75.0% vs. 74.8%). The MF1 metric more accurately reflects the classification performance for each class. By integrating both SSA and SSP alignment modules, our SF-UIDA framework achieves superior performance compared to single-alignment approaches.

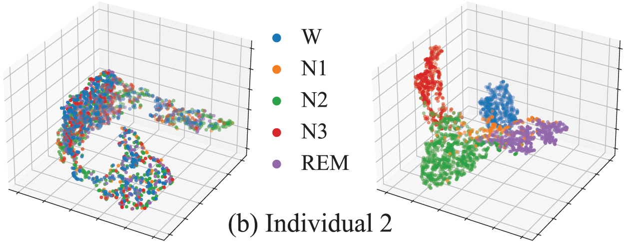
Computational Complexity To evaluate the computational complexity of our proposed SF-UIDA framework, we calculate the time cost per individual across three datasets shown in Fig. 5. Our method is capable of completing personalized customization for an unknown individual within an average of 40 seconds. When compared to the several hours' duration of one person's sleep records, this time cost is acceptable. Moreover, considering the unseen subject commonly appears one by one in practice, our method is applicable to enable the source model continuously adapt to new subjects and achieve plug-and-play personalized customization for each individual.

Feature Visualization To demonstrate the effectiveness of our method, we selected two individuals from the Sleep-EDF dataset to visualize intermediate features using the t-SNE method (Van der Maaten and Hinton 2008). Figures 6 (a-1) and (a-2) illustrate the feature distribution for the first individual, showing the distribution before and after personalization. Similarly, Figures 6 (b-1) and (b-2) depict the distribution for the second individual. As shown in Figure 6 (a-2), samples from the same sleep stage are well-clustered after personalization, compared to the distribution in Figure 6 (a-1). A similar trend is observed for the second individual, demonstrating that our method effectively achieves personalized customization for target subjects.



(a) Individual 1

(a-1) Direct Transfer (a-2) Personalization



(b) Individual 2

(b-1) Direct Transfer (b-2) Personalization

Figure 6: Sleep features visualization.

Conclusion

In this paper, we present a novel Source-Free Unsupervised Individual Domain Adaptation (SF-UIDA) framework for automatic sleep staging, employing a two-step subject-specific alignment scheme for adaptation. Our framework facilitates plug-and-play personalization for each new individual without requiring access to source data, meeting the practical needs in clinics. Experimental results across three public datasets demonstrate that the SF-UIDA framework effectively transforms a source model into a personalized one within a short adaptation period, highlighting its practical significance. Our future work will aim to extend the applicability of our method to a broader range of EEG tasks.

Acknowledgments

This work was supported by STI 2030 Major Projects (2021ZD0200400), the Key Program of the Natural Science Foundation of Zhejiang Province, China (No. LZ24F020004), the Natural Science Foundation of China (No. 61925603) and the Fundamental Research Funds for the Central Universities (No. 61925603). The corresponding author is Dr. Sha Zhao.

References

Alvarez-Estevez, D.; and Rijsman, R. M. 2021. Inter-database validation of a deep learning approach for automatic sleep scoring. *PLOS ONE*, 16(8): 1–27.

Berry, R. B.; Budhiraja, R.; Gottlieb, D. J.; Gozal, D.; Iber, C.; Kapur, V. K.; Marcus, C. L.; Mehra, R.; Parthasarathy, S.; Quan, S. F.; et al. 2012. Rules for scoring respiratory events in sleep: update of the 2007 AASM manual for the scoring of sleep and associated events: deliberations of the sleep apnea definitions task force of the American Academy of Sleep Medicine. *Journal of clinical sleep medicine*, 8(5): 597–619.

Chen, D.; Wang, D.; Darrell, T.; and Ebrahimi, S. 2022. Contrastive test-time adaptation. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 295–305.

Chen, T.; Kornblith, S.; Norouzi, M.; and Hinton, G. 2020. A simple framework for contrastive learning of visual representations. In *International conference on machine learning*, 1597–1607. PMLR.

Chen, X.; and He, K. 2021. Exploring simple siamese representation learning. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, 15750–15758.

Eldele, E.; Ragab, M.; Chen, Z.; Wu, M.; Kwoh, C.-K.; and Li, X. 2023. Self-Supervised Learning for Label-Efficient Sleep Stage Classification: A Comprehensive Evaluation. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, 31: 1333–1342.

Eldele, E.; Ragab, M.; Chen, Z.; Wu, M.; Kwoh, C. K.; Li, X.; and Guan, C. 2021. Time-series representation learning via temporal and contextual contrasting. *arXiv preprint arXiv:2106.14112*.

Fan, J.; Zhu, H.; Jiang, X.; Meng, L.; Chen, C.; Fu, C.; Yu, H.; Dai, C.; and Chen, W. 2022. Unsupervised domain adaptation by statistics alignment for deep sleep staging networks. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, 30: 205–216.

Humphreys, J.; Sharps, P. W.; and Campbell, J. C. 2005. What we know and what we still need to learn. *Journal of interpersonal violence*, 20(2): 182–187.

Jia, Z.; Lin, Y.; Wang, J.; Wang, X.; Xie, P.; and Zhang, Y. 2021. SalientSleepNet: Multimodal salient wave detection network for sleep staging. *arXiv preprint arXiv:2105.13864*.

Karim, N.; Mithun, N. C.; Rajvanshi, A.; Chiu, H.-p.; Samarasekera, S.; and Rahnavard, N. 2023. C-sfda: A curriculum learning aided self-training framework for efficient source free domain adaptation. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 24120–24131.

Kemp, B.; Zwinderman, A.; Tuk, B.; Kamphuisen, H.; and Oberye, J. 2000. Analysis of a sleep-dependent neuronal feedback loop: the slow-wave microcontinuity of the EEG. *IEEE Transactions on Biomedical Engineering*, 47(9).

Khalighi, S.; Sousa, T.; Santos, J. M.; and Nunes, U. 2016. ISRUC-Sleep: A comprehensive public dataset for sleep researchers. *Computer Methods and Programs in Biomedicine*, 124: 180–192.

Liang, J.; Hu, D.; and Feng, J. 2020. Do we really need to access the source data? source hypothesis transfer for unsupervised domain adaptation. In *International conference on machine learning*, 6028–6039. PMLR.

Matsushima, H.; Minami, S.; and Takadama, K. 2012. Age-Based Sleep Stage Estimation by Evolutionary Algorithm. In *2012 AAAI Spring Symposium Series*.

Mikolov, T.; Karafiát, M.; Burget, L.; Cernockỳ, J.; and Khudanpur, S. 2010. Recurrent neural network based language model. In *Interspeech*, volume 2, 1045–1048. Makuhari.

Mousavi, S.; Afghah, F.; and Acharya, U. R. 2019. SleepEEGNet: Automated sleep stage scoring with sequence to sequence deep learning approach. *PloS one*, 14(5): e0216456.

Nie, H.; Tu, S.; and Xu, L. 2021. Recsleepnet: An automatic sleep staging model based on feature reconstruction. In *2021 IEEE International Conference on Bioinformatics and Biomedicine (BIBM)*, 1458–1461. IEEE.

Oord, A. v. d.; Li, Y.; and Vinyals, O. 2018. Representation learning with contrastive predictive coding. *arXiv preprint arXiv:1807.03748*.

Perslev, M.; Darkner, S.; Kempfner, L.; Nikolic, M.; Jennum, P. J.; and Igel, C. 2021. U-Sleep: resilient high-frequency sleep staging. *NPJ digital medicine*, 4(1): 72.

Perslev, M.; Jensen, M.; Darkner, S.; Jennum, P. J.; and Igel, C. 2019. U-time: A fully convolutional network for time series segmentation applied to sleep staging. *Advances in Neural Information Processing Systems*, 32.

Phan, H.; Mikkelsen, K.; Chen, O. Y.; Koch, P.; Mertins, A.; and Vos, M. D. 2022. SleepTransformer: Automatic Sleep Staging With Interpretability and Uncertainty Quantification. *IEEE Transactions on Biomedical Engineering*, 69(8): 2456–2467.

Ragab, M.; Eldele, E.; Chen, Z.; Wu, M.; Kwoh, C.-K.; and Li, X. 2022. Self-Supervised Autoregressive Domain Adaptation for Time Series Data. *IEEE Transactions on Neural Networks and Learning Systems*, 1–11.

Supratak, A.; Dong, H.; Wu, C.; and Guo, Y. 2017. DeepSleepNet: A Model for Automatic Sleep Stage Scoring Based on Raw Single-Channel EEG. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, 25(11): 1998–2008.

Supratak, A.; and Guo, Y. 2020. TinySleepNet: An efficient deep learning model for sleep stage scoring based on raw single-channel EEG. In *2020 42nd Annual International*

Conference of the IEEE Engineering in Medicine & Biology Society (EMBC), 641–644. IEEE.

Tang, M.; Zhang, Z.; He, Z.; Li, W.; Mou, X.; Du, L.; Wang, P.; Zhao, Z.; Chen, X.; Li, X.; et al. 2022. Deep adaptation network for subject-specific sleep stage classification based on a single-lead ECG. *Biomedical Signal Processing and Control*, 75: 103548.

Tarvainen, A.; and Valpola, H. 2017. Mean teachers are better role models: Weight-averaged consistency targets improve semi-supervised deep learning results. *Advances in neural information processing systems*, 30.

Tsinalis, O.; Matthews, P. M.; Guo, Y.; and Zafeiriou, S. 2016. Automatic sleep stage scoring with single-channel EEG using convolutional neural networks. *arXiv preprint arXiv:1610.01683*.

Van der Maaten, L.; and Hinton, G. 2008. Visualizing data using t-SNE. *Journal of machine learning research*, 9(11).

Vaswani, A.; Shazeer, N.; Parmar, N.; Uszkoreit, J.; Jones, L.; Gomez, A. N.; Kaiser, Ł.; and Polosukhin, I. 2017. Attention is all you need. *Advances in neural information processing systems*, 30.

Wang, J.; Zhao, S.; Jiang, H.; Li, S.; Li, T.; and Pan, G. 2024a. Generalizable sleep staging via multi-level domain alignment. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 38, 265–273.

Wang, J.; Zhao, S.; Jiang, H.; Zhou, Y.; Yu, Z.; Li, T.; Li, S.; and Pan, G. 2024b. CareSleepNet: A Hybrid Deep Learning Network for Automatic Sleep Staging. *IEEE Journal of Biomedical and Health Informatics*, 28(12): 7392–7405.

Wang, J.; Zhao, S.; Luo, Z.; Zhou, Y.; Jiang, H.; Li, S.; Li, T.; and Pan, G. 2024c. CBraMod: A Criss-Cross Brain Foundation Model for EEG Decoding. arXiv:2412.07236.

Wang, J.; Zhao, S.; Zhou, Y.; Jiang, H.; Yu, Z.; Li, T.; Li, S.; and Pan, G. 2023. Narcolepsy Diagnosis With Sleep Stage Features Using PSG Recordings. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, 31: 3619–3629.

Wang, Q.; Fink, O.; Van Gool, L.; and Dai, D. 2022. Continual test-time domain adaptation. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 7201–7211.

Zhou, Y.; Zhao, S.; Wang, J.; Jiang, H.; Yu, Z.; Li, S.; Li, T.; and Pan, G. 2024. Simplifying Multimodal With Single EOG Modality for Automatic Sleep Staging. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, 32: 1668–1678.