SSVEP-DAN: Cross-Domain Data Alignment for SSVEP-Based Brain–Computer Interfaces

Sung-Yu Chen, *Graduate Student Member, IEEE*, Chi-Min Chang, Kuan-Jung Chiang, and Chun-Shu Wei, *Member, IEEE*

***Abstract*— Steady-state visual-evoked potential (SSVEP)-based brain-computer interfaces (BCIs) offer a non-invasive means of communication through high-speed speller systems. However, their efficiency is highly dependent on individual training data acquired during time-consuming calibration sessions. To address the challenge of data insufficiency in SSVEP-based BCIs, we introduce SSVEP-DAN, the first dedicated neural network model designed to align SSVEP data across different domains, encompassing various sessions, subjects, or devices. Our experimental results demonstrate the ability of SSVEP-DAN to transform existing source SSVEP data into supplementary calibration data. This results in a significant improvement in SSVEP decoding accuracy while reducing the calibration time. We envision SSVEP-DAN playing a crucial role in future applications of high-performance SSVEP-based BCIs. The source code for this work is available at: https://github.com/CECNL/SSVEP-DAN.**

***Index Terms*— Electroencephalogram (EEG), brain–computer interface (BCI), steady-state visual-evoked potentials (SSVEPs), domain adaptation, data alignment.**

I. Introduction

STEADY-STATE Visual Evoked Potential (SSVEP) is a type of electroencephalogram (EEG) signal observed in the occipital region when an individual focuses their attention on visual stimuli flickering at specific frequencies [1], [2]. The SSVEP signal is recognized for its stability [3], [4] and has become a reliable control signal for non-invasive brain-computer interfaces (BCIs) [5], [6] in various practical applications, such as spelling [7], [8], gaming [9], [10], and device control [11], [12].

To accurately detect and analyze users' SSVEPs for distinguishing corresponding stimuli, the development of efficient decoding algorithms has become significantly important. Canonical correlation analysis (CCA) [13], [14] has been proposed as a training-free technique used to assess the relationships between multichannel SSVEP and the reference sinusoidal signals corresponding to each stimulation frequency. Although the performance of training-free methods is limited by individual differences, training-based SSVEP detection leverages individual calibration data to improve the performance. The most widely used training-based SSVEP detection algorithm is task-related component analysis (TRCA) [15], [16], aiming to separate task-related from non-task-related information by maximizing SSVEP data reproducibility within each trial. As the gold standard approach to online SSVEP-based BCI [17], [18], [19], the success of TRCA has inspired the development of novel neural network-based SSVEP detection models such as ConvCA [20] and bi-SiamCA [21]. However, the calibration process for collecting individual data is often time-consuming and laborious, leading to significant visual fatigue in subjects [22]. Furthermore, due to substantial variability between subjects, simply concatenating training data from a larger pool of participants can potentially lead to a decrease in decoding performance [23]. This technical challenge manifests itself as a domain adaptation problem, a subcategory of transfer learning that aims to transfer knowledge from a source domain to improve the performance of a model on a target domain [24]. In this study, we develop a domain adaptation technique to mitigate inter-domain disparities and enhance SSVEP decoding algorithm performance, particularly in scenarios with limited calibration data. Figure 1 illustrates the proposed framework, where the target domain simulates a new user of the SSVEP-based BCI, and the source domain represents subjects providing existing SSVEP data.

Recently, various studies have explored the application of domain adaptation techniques to streamline calibration efforts or enhance detection performance in SSVEP-based BCIs [23], [25], [26], [27], [28]. Many of these methods aim to align data from different participants into a shared subspace to mitigate inter-subject distribution discrepancies. However, they often

Manuscript received 30 November 2023; revised 10 April 2024 and 30 April 2024; accepted 18 May 2024. Date of publication 23 May 2024; date of current version 29 May 2024. This work was supported in part by the National Science and Technology Council (NSTC) under Contract 109-2222-E-009-006-MY3, Contract 110-2221-E-A49-130-MY2, Contract 110-2314-B-037-061, Contract 112-2321-B-A49-012, and Contract 112-2222-E-A49-008-MY2; and in part by the Higher Education Sprout Project of National Yang Ming Chiao Tung University and Ministry of Education. *(Corresponding author: Chun-Shu Wei.)*

Sung-Yu Chen and Chi-Min Chang are with the Department of Computer Science, National Yang Ming Chiao Tung University (NYCU), Hsinchu 30010, Taiwan.

Kuan-Jung Chiang is with Arctop Inc., La Jolla, CA 92093 USA.

Chun-Shu Wei is with the Department of Computer Science, the Institute of Education, and the Institute of Biomedical Engineering, National Yang Ming Chiao Tung University (NYCU), Hsinchu 30010, Taiwan (e-mail: wei@nycu.edu.tw).

Digital Object Identifier 10.1109/TNSRE.2024.3404432

TABLE I
COMPARISON OF EXISTING DOMAIN ADAPTATION METHODS FOR SSVEP DECODING. SUBSPACE ALIGNMENT APPROACHES IMPOSE CONSTRAINTS ON SUBSEQUENT DECODING FEATURE TYPES, WHEREAS DATA ALIGNMENT METHODS GENERATE TIME-SERIES DATA THAT CAN BE WIDELY UTILIZED FOR SSVEP DECODING

Method	Approach of domain adaptation	Stimulus-specific	Output	Constraint on SSVEP decoding	Domain transferred		
					Subjects	Sessions	Devices
RPA [25]	Subspace alignment	No	SPD matrix	Yes	Yes	Yes	No
SLR [26]	Subspace alignment	Yes	Spatial patterns	Yes	Yes	Yes	Yes
ALPHA [27]	Subspace alignment	Yes	Spatial patterns	Yes	Yes	Yes	Yes
TSA [28]	Subspace alignment	No	Tangent vectors	Yes	Yes	Yes	Yes
LST [23]	Data alignment	Yes	Time series	No	Yes	Yes	Yes

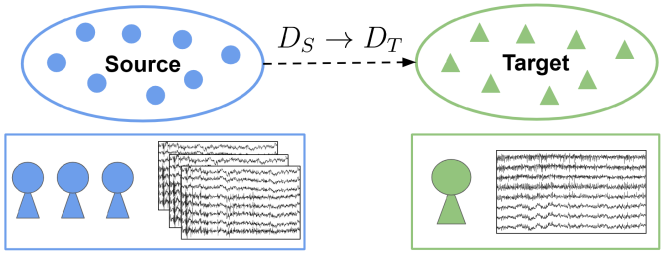


Fig. 1. An illustration of domain adaptation for boosting the calibration of SSVEP-based BCI. Transferring existing data from the source subjects (D_S) to the target subject (D_T) provides additional calibration for a target user and therefore reduce the required amount of data from individual calibration.

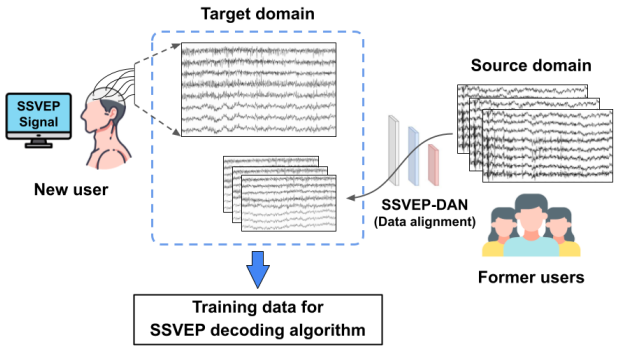


Fig. 2. The proposed SSVEP data alignment framework based on SSVEP-DAN. The SSVEP-DAN learns the transformation between the source and target domain through minimizing the difference between the transformed and target SSVEP data. Then, the transformed SSVEP data, together with the target SSVEP data, serve as the calibration data for the training-based SSVEP detection.

output non-SSVEP signals, such as covariance matrices [25], tangent vectors [28], or spatial patterns [26], [27], posing challenges for integration with conventional or deep-learning-based SSVEP detection algorithms like TRCA or Conv-CA.

On the contrary, data alignment techniques adapt samples from the source and target domains to generate time-series SSVEP signals as output. A notable method in this regard is least-square transformation (LST) [23], which transforms SSVEP waveforms from source subjects into additional calibration data for target subjects seamlessly integrated with subsequent SSVEP detection algorithms. However, emerging evidence suggests SSVEP signals possess non-linear characteristics [29], [30], [31], yet LST has limited capacity to accommodate non-linear transformations and noise tolerance. Moreover, the stimulus-dependent training process of LST restricts the data availability during model fitting, potentially leading to redundant or invalid transformation models.

To address the current domain adaptation challenges in SSVEP detection, we present SSVEP-DAN, a neural network-based method. SSVEP-DAN provides non-linear mapping for transforming source SSVEP signals into target domain data, enabling robust transformations through stimulus-independent training. The transformed SSVEP signals serve as supplementary calibration data for the target subject, compatible with any training-based SSVEP detection algorithm. Our novel architecture, depicted in Figure 2, incorporates innovative training approaches, including stimulus-independent training and pre-training techniques, to overcome data scarcity challenges. We evaluate the effectiveness of SSVEP-DAN alongside the standard TRCA framework [15], [16], [32] in various cross-domain adaptations that reflect practical SSVEP-based BCI scenarios.

II. RELATED WORK

Domain adaptation techniques aim to adapt the trained model from the source domain to the target domain by leveraging the available data from the target domain while utilizing the knowledge learned from the source domain. The goal is to improve the model's performance and generalization capabilities across different users without the need for extensive re-training or user-specific calibration. This section provides the background of domain adaptation for SSVEP-based BCI and reviews relevant studies on existing domain adaptation approaches as summarized in Table I.

Domain adaptation methods in SSVEP-based BCIs are typically categorized into two groups: 1) subspace alignment, which involves alignment between domains of feature spaces or subspaces, and 2) data alignment, which performs alignment on the SSVEP signals between domains based on their transformation relationships [33], [34], [35].

A. Subspace Alignment

Subspace alignment methods align source and target domains of feature subspace to a common subspace where the the discrepancy between the two domains reduces. Riemannian Procrustes Analysis (RPA) [25] achieves this by applying simple geometric transformations (translation, scaling, and rotation) to symmetric positive definite matrices (SPD), aligning the source and target domains to the same subspace. Although this method can be applied across subjects and sessions, its practicality is limited due to the output being SPD

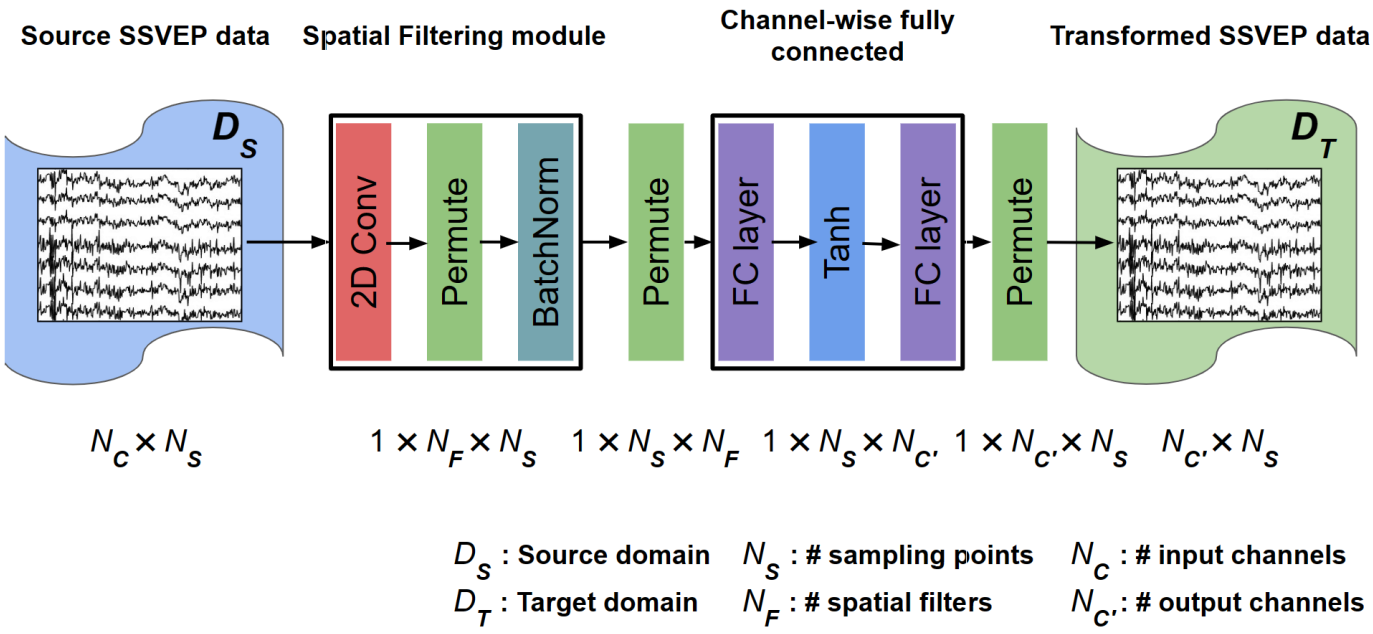


Fig. 3. The architecture of the proposed SSVEP-DAN.

matrices. Shared Latent Response (SLR) [26] uses common spatial filtering methods, including CCA and TRCA, to extract features from the training data and then uses least squares regression to obtain new spatial filters that project test data onto the same subspace as the training data. This approach is applicable to cross-subject, cross-session, and cross-device scenarios, with input data being common time series data, providing more flexibility in practical applications. ALign and Pool for EEG Headset domain Adaptation (ALPHA) [27] aligns spatial patterns through orthogonal transformations and aligns the covariance between different distributions using linear transformations, mitigating variations in spatial patterns and covariance. This method further improves upon obtaining new spatial filters and achieves better performance than SLR. Tangent Space Alignment (TSA) [28] shares similarities with RPA as it aligns different domains to the same subspace through translation, scaling, and rotation, but operates within the tangent space. The tangent space being Euclidean allows for faster decoding, and rotation can be achieved with a singular value decomposition (SVD), making it computationally efficient compared to RPA. The tangent vector, as the output of TSA, has limited compatibility to most classification methods for SSVEP detection that requires time-series signals and thus its practicality is restricted. Similar issues are also found in applying subspace alignment methods that identify target stimuli by computing correlation coefficients between spatial features in the same subspace [26], [27].

B. Data Alignment

Data alignment methods perform alignment between source domain samples and target domain samples in order to mitigate the disparities between the two domains. Recently, the Least Squares Transformation (LST) approach [23] finds a linear transformation relationship among the SSVEP data, effectively reducing the errors between the transformed data from the source SSVEP and the target SSVEP. This method is applicable in multiple cross-domain scenarios, including cross-sessions, cross-subjects, and cross-devices transfer learning for boosting the calibration of training-based SSVEP detection. Furthermore, the property of using time-series output allows data alignment methods to be seamlessly integrated into common SSVEP detection methods such as TRCA [15], and Conv-CA [20]. The utilization of data alignment for SSVEP-based BCI can significantly enhance their feasibility in real-world applications with reduced calibration effort for individual users. Yet, as LST has high flexibility for cross-domain adaptation and desirable simplicity, it can only transform SSVEP signals within the same visual stimulus via linear combination. Further development of data alignment techniques for SSVEP signals is required to tackle the issues of stimulus-independent training and non-linear transformation learning.

III. Methodology

A. SSVEP-DAN

In this study, we assume that there exists a non-linear and channel-wise transformation of SSVEP signals between subjects. We propose a neural network-based transformation to transfer SSVEP signals from an existing subject (source domain) to a set of additional calibration data for a target subject (target domain). The architecture of the proposed neural network is illustrated as in Figure 3. The input to SSVEP-DAN is SSVEP data obtained in the source domain sizes $\mathbb{R}^{N_C \times N_S}$, while the output matches the average of multiple trials corresponding to specific stimuli from the target subject, formatted as $\mathbb{R}^{N'_C \times N_S}$. Here, N_C denotes the number of input channels, $N_{C'}$ is the number of output channels, and N_S represents the number of sampling points. We employ these input and output sets to train our SSVEP-DAN model.

Conventional SSVEP detection methods, such as CCA and TRCA, apply spatial filtering to find a linear combination of channel-wise SSVEP signal [13], [15]. These methods have been demonstrated to improve signal-to-noise ratio (SNR) [36] and enhance SSVEP detection performance. Additionally, SSVEP features time-synchronous signals with prominent oscillatory waveform at specific stimulation frequency and its harmonics, we assume the cross-domain adaptation requires transformation in spatial domain rather than in temporal domain. The signal processing of spatial filtering has been utilized in recently developed neural network-based EEG decoders such as SCCNet [37], where a spatial convolutional layer serves for noise reduction and feature extraction. Therefore, in the first module, we utilize spatial convolution, incorporating N_F spatial filters with a shape of $(N_C, 1)$, to project the original SSVEP data into latent spaces to obtain spatial features, where N_F represents the number of spatial filters. Note that in this study, the number of spatial filters, N_F, is equal to the number of input channels, N_C. After the convolutional layer, the dimensions of the latent features are properly permuted and subjected to batch normalization, where each timestamp and filter channel is normalized independently. Subsequently, channel-wise fully connected layers with dimensions $\mathbb{R}^{N'_C \times N_S}$ are applied, projecting the data from spatial filter component spaces to output channel spaces.

To capture non-linear channel relations between spatial features and target SSVEP templates, we incorporate two channel-wise fully connected layers with a hyperbolic tangent activation function in between. Channel information at each time point is integrated using another channel-wise fully connected layer, projecting the feature into a new latent space. The activation function $tanh$ is then applied to facilitate model fitting. Lastly, another channel-wise fully connected layer is used to project the features into the target domain space, and permutation is applied to obtain transformed SSVEP data in the target domain.

B. Training Strategies

1) Stimulus-Independent Training: Based on the assumption that the transformation of SSVEP data between subjects is irrelevant to the visual stimuli, we propose a stimulus-independent training strategy dedicated for neural network-based SSVEP data alignment to augment the data amount through combining data across different stimuli, aiming for a relatively robust transformation under limited data amount. The stimulus-independent training process involves cross-stimulus training that enables the learning of model parameters across different stimulus frequencies and thereby ensures a relatively reliable learning process for the transformation between subjects, particularly when calibration data is scarce. As data of different stimulus types are merged within individual training batch, the stimulus type between the source/target SSVEP data remains the same.

2) Two-Phase Model Training: Recent studies of deep-learning-based SSVEP detection suggest that fine-tuning a pre-trained model improves the performance as the training data can be fully used during pre-training and the characteristic of individual domain is considered during fine-tuning [37], [38], [39]. We adopt this strategy to handle the training of SSVEP-DAN in the case of multiple source domains with a two-phase procedure consisting of a pre-training phase and a fine-tuning phase, as illustrated in Figure 4. In the pre-training phase, data from all source domains are concatenated and used in the training of a pre-trained model. Next, a fine-tuning phase is performed to fine-tune the pre-trained model separately using the data from individual source domain and acquire a fine-tuned models between each source domain and the target domain. Lastly, all transformed data from the source domains merge into the calibration data for the target subject, as depicted in Figure 5.

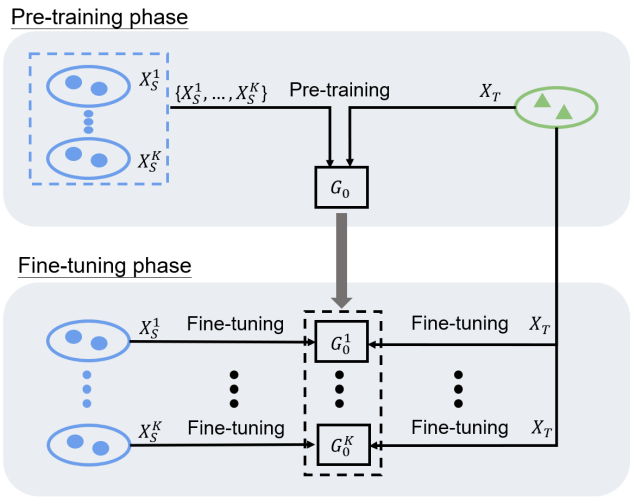


Fig. 4. The two-phase model training proposed in this study includes a pre-training phase and a fine-tuning phase. Initially, a pre-trained model, denoted as G_0, is trained using the target data X_T and the entire source data pool, represented as $X_S^1, X_S^2, \ldots, X_S^K$, where K is the total number of source domains. Subsequently, the pre-trained model G_0 undergoes separate fine-tuning with each individual source dataset, resulting in fine-tuning models $G_0^1, G_0^2, \ldots, G_0^K$.

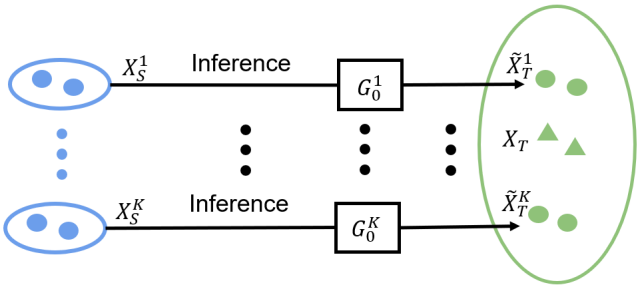


Fig. 5. The inference procedure transforms the data of each source domain into additional data in the target domain. For each source domain D_S^k, $\forall k \in \{1, 2, \ldots, K\}$, the SSVEP-DAN model G_0^k transforms the source data X_S^k into $\tilde{X}_T^k$ for supplement the amount of data in the target domain D_T.

C. Model Fitting

The network is trained using the Adam optimizer [40] with a learning rate set to 5×10^4. The transformation loss L_{trans} is determined by the Mean Square Error (MSE) between the

target SSVEP data X_T and the output SSVEP data $\widetilde{X}_T$:

$$L_{trans}(X_T, X'_T) = \frac{1}{N}\sum_{i=1}^{N} ||X_T - \widetilde{X}_T||_2^2, \quad (1)$$

where N is the batch size. The target SSVEP data X_T is obtained by averaging calibration trials of the same stimulus from the target subject as suggested in [23]. During the pre-training phase, the source subjects are divided into a training set and a validation set using a subject-wise ratio of 8:2. The model is trained for 500 epochs, and the model weights that yield the lowest validation set loss are employed as the fitting result. In the fine-tuning phase, a 150-epoch fitting process is conducted using data from a single source subject, following the same configuration as described above, except for the training/validation data splitting.

D. TRCA-Based Performance Assessment

TRCA is a training-based algorithm that aims to extract task-related components by maximizing the inter-trial coherence of neural activity across multiple trials within each specific task [15]. Furthermore, the combination of TRCA and filter bank analysis facilitates the decomposition of SSVEP signals into multiple sub-band components, thereby effectively extracting independent information embedded within the harmonic components [41]. Finally, an ensemble approach is employed to integrate multiple filters trained using the aforementioned methods.

E. Data

1) Dataset I: SSVEP Benchmark Dataset: [42] used in this study is a publicly available SSVEP dataset prepared by the Tsinghua group. In this dataset, the SSVEP-based BCI experiment involved 35 participants. Each participant participated in 6 blocks of the experiment, each block comprising 40 trials presented in random order. Visual stimuli were presented within a frequency range of 8 to 15.8 Hz with an interval of 0.2 Hz. The phase values of the stimuli ranged from 0, with a phase interval of 0.5π. EEG signals were collected utilizing the Synamps2 EEG system (Neuroscan, Inc.). They were recorded using the extended 10-20 system through 64 channels. We selected EEG data from eight channels (PO3, PO4, PO5, PO6, POz, O1, O2, Oz) in the analysis and performance evaluation. The EEG signals were down-sampled from 1000 to 250 Hz, and a notch filter at 50 Hz was applied to remove the common power-line noise. The data were extracted in [L_1 s, $L_1 + T_{w_1}$ s] of the stimulus onset, where L_1 is the latency delay ($L_1 = 0.14$ s) and T_{w_1} indicates the time-window length ($T_{w_1} = 1.5$ s). In the SSVEP detection based on filter-bank TRCA, we set the number of filter banks as five for this dataset as previously suggested [20].

2) Dataset II: Wearable SSVEP BCI Dataset: The Wearable SSVEP BCI dataset [43] used in this study is another publicly available SSVEP dataset released by the Tsinghua group. In this dataset, 102 healthy subjects participated in the wearable SSVEP-based BCI experiment. The experiment consisted of 10 blocks, each of which contained 12 trials in random order of 12 visual stimuli. Stimulation frequencies ranged from 9.25 to 14.75 Hz with an interval of 0.5 Hz. The phase values of the stimuli started at 0, and the phase difference between two adjacent frequencies was 0.5π. EEG signals were collected utilizing the Neuracle EEG Recorder NeuSen W (Neuracle, Ltd.) system. The 8-channel EEG data was recorded using wet and dry electrodes, and the electrodes were placed according to the international system 10-20. All channels (PO3, PO4, PO5, PO6, POz, O1, O2, Oz) of the EEG signals were used in data analysis and performance evaluation. The EEG signals were resampled at 250 Hz from 1000 Hz. To remove the common power-line noise, a 50 Hz notch filter was applied to the dataset. The data was extracted in [0.5 + L_2 s, 0.5 + L_2 + T_{w_2} s], where 0.5 s denotes stimulus onset, L_2 indicates latency delay ($L_2 = 0.14$ s) and T_{w_2} is the time-window length ($T_{w_2} = 1.5$ s). For this dataset, we set the number of filter banks as three according to [43].

F. Performance Evaluation

The performance evaluation of the SSVEP-DAN was conducted through leave-one-subject-out cross-validation. In this approach, each subject is treated as the target subject, while the remaining subjects serve as source subjects. This setup reflects the real-world scenario of employing SSVEP-based BCI when introducing a new user to the system.

In the leave-one-subject-out cross-validation, we designated a specific set of trials as the test data for each individual target subject. Within each subject's trials from Dataset I, the initial 4 trials were utilized as both the calibration data and the target SSVEP data for the training of data alignment, leaving the remaining 2 trials as the designated test data. For Dataset II, the first 6 trials were allocated for calibration or data alignment training, with the remaining 4 trials serving as the designated test data. We followed the precedent set by a prior study [23] by using a minimum of 2 calibration trials. This minimum is necessitated by the TRCA, which requires at least 2 calibration trials for the acquisition of an average SSVEP template.

To justify the efficacy of this framework, we conducted a performance comparison of SSVEP-DAN against the Baseline (without domain adaptation) and other domain adaptation methods, specifically Concatenation (Concat.) and LST. The performance comparison was based on the evaluation of SSVEP detection performance using filter-bank TRCA with calibration data prepared using these schemes, considering different numbers of calibration trials (Dataset I: 2-4; Dataset II: 2-6) from the target (test) subject. A detailed description of these schemes is provided below.

1) Baseline: This approach utilizes the filter-bank TRCA without any provided source data. The TRCA relies on calibration data collected solely from the target subject.

2) Concat: The Concat. approach employs a simple transfer learning scheme by naively concatenating all SSVEP data from the source domains with the target SSVEP data without any transformation. The concatenated SSVEP data is then used as calibration data for TRCA.

3) LST: The LST approach involves a linear transformation to transfer the source SSVEP data to the target domain based on source subjects and stimuli. The LST-transformed data is

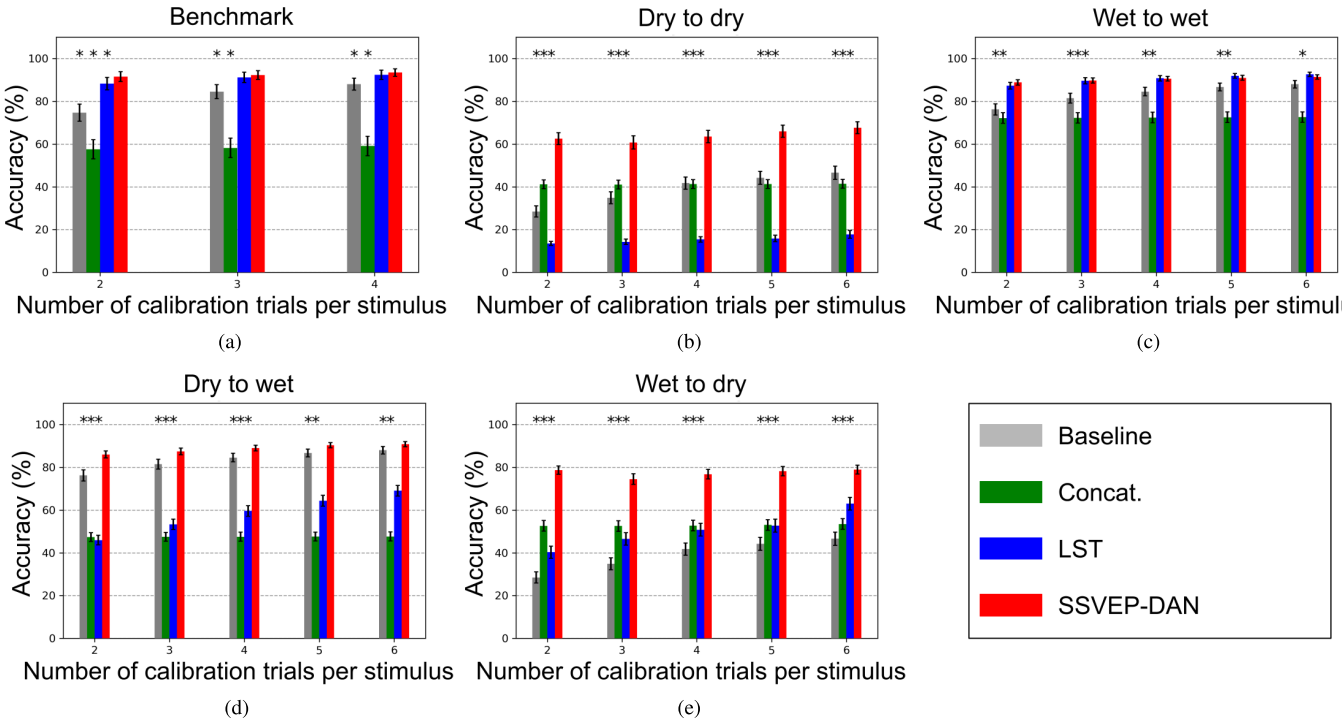


Fig. 6. The performance evaluation (**%**) against number of calibration trials per stimulus in the target domain across the five scenarios. Asterisks indicate significant differences between SSVEP-DAN and other methods. (*p <0.05).

then concatenated with the target SSVEP data to form the calibration data for TRCA.

4) SSVEP-DAN: The SSVEP-DAN approach utilizes a nonlinear transformation to transfer the source SSVEP data to the target domain based on source subjects. The transformed data is then concatenated with the target SSVEP data to construct the calibration data for TRCA.

In addition to the Baseline scheme, for Concat., LST, and SSVEP-DAN schemes, the training and validation sets were randomly partitioned. Furthermore, the parameters of the SSVEP-DAN model were initialized randomly. The decoding performance of each domain adaptation scheme was estimated by averaging ten repeats. We utilized the Wilcoxon signed-rank test to assess the statistical significance of the improvements between the proposed SSVEP-DAN-based method and other domain adaptation schemes.

IV. Experiment Results

This section presents experimental results, evaluating SSVEP-DAN across practical SSVEP-based BCI scenarios. Using two datasets, we conducted five domain adaptation tasks, detailed in Table II. Performance comparisons considered: 1) target subject's calibration trials, 2) source subjects, and 3) SSVEP data time-window length. Additionally, we conducted an ablation study and visualized adaptation results to elucidate SSVEP-DAN's features and training strategy effects.

A. Performance Comparison

Figure 6 illustrates cross-domain performance using different schemes with varying numbers of calibration trials per stimulus. SSVEP-DAN consistently outperforms other schemes across most cases, regardless of the target subject's calibration trials. Interestingly, as the target subject's calibration trials increase, our method consistently enhances the performance. Notably, Concat. scheme sometimes negatively affects TRCA-based methods, but positively impacts them when using dry electrode devices for new subjects. This signal quality discrepancy is prominent between Dataset I and II [21], [44]. Dry-electrode data generally exhibits lower SNR than wet-electrode data due to factors like unstable contact or higher impedance [45], [46]. Within Dataset II, dry-electrode data displays lower signal quality than wet-electrode data [43].

LST-based methods effectively enhance TRCA-based methods in some scenarios but can lead to negative transfer when source data is collected via dry electrode devices [24], [47]. This suggests LST's transformation matrix stability is highly influenced by SSVEP signal quality [23]. While trial averaging enhances SSVEP SNR [15], [48], poor input SSVEP signal quality can yield unreliable transformation matrices, affecting final performance. Our approach consistently outperforms TRCA-based methods across diverse scenarios, particularly

TABLE II
Domain Adaptation Tasks

Task	Source domain	Target domain
Benchmark	Dataset I	Dataset I
Dry to dry	Dataset II - Dry	Dataset II - Dry
Wet to wet	Dataset II - Wet	Dataset II - Wet
Dry to wet	Dataset II - Dry	Dataset II - Wet
Wet to dry	Dataset II - Wet	Dataset II - Dry

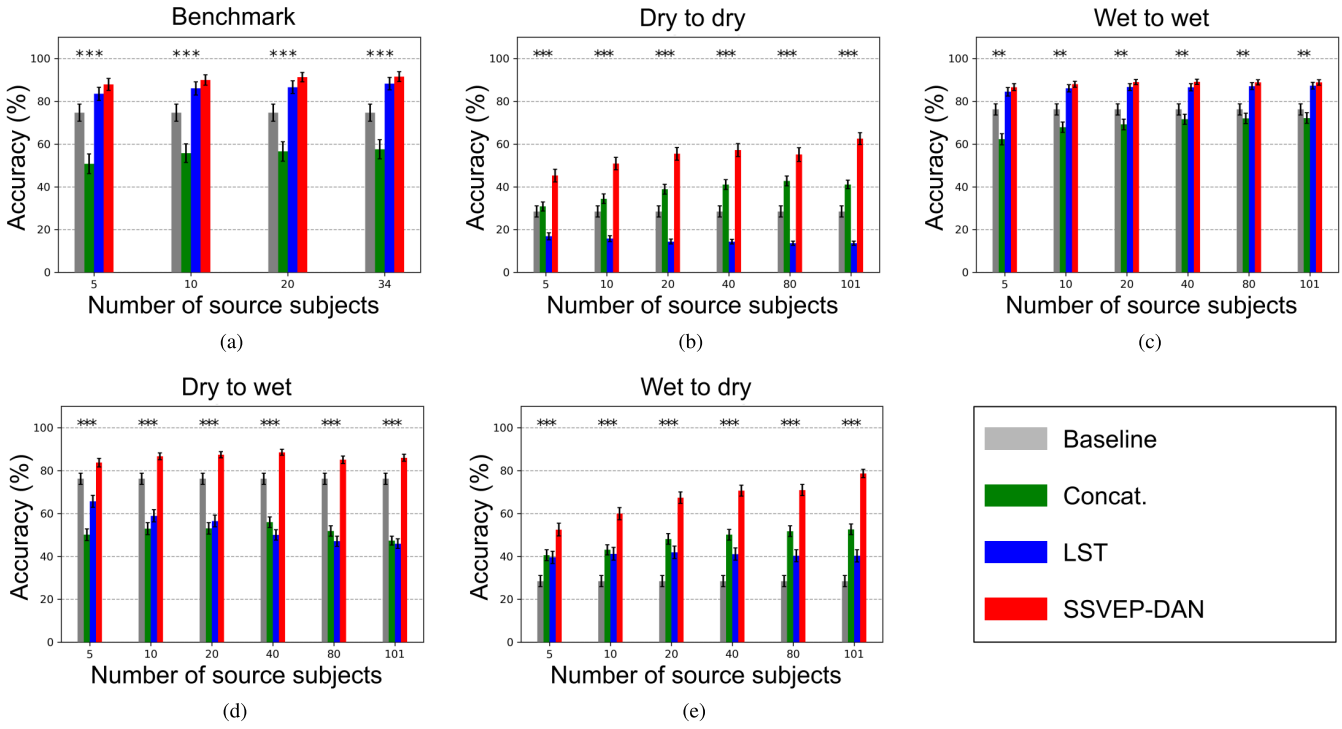


Fig. 7. The performance evaluation (**%**) against number of source subjects across the five scenarios. A fixed amount of target data was applied in the analysis (two trials per stimulus). Asterisks indicate significant differences between SSVEP-DAN and other schemes. (*p <0.05).

with limited calibration data. SSVEP-DAN effectively expands training datasets for decoding algorithms, enhancing performance, notably in real-world scenarios involving dry-electrode SSVEP data [46].

B. Number of Source Subjects

This section investigates the influence of different numbers of source subjects on the performance of SSVEP-DAN. In the Concat., LST, and SSVEP-DAN schemes, we randomly sampled a subset of source subjects from the pool of all available source subjects with ten repetitions. In both Dataset I and Dataset II, we evaluate performance when there is an insufficient number of source subjects, while keeping the amount of target data fixed (two trials). Figure 7 illustrates the cross-domain performance as a function of the number of source subjects. Results consistently demonstrate the superiority of SSVEP-DAN regardless of the number of source subjects in most cases. The influence of source subjects is particularly noticeable in 'dry to dry' and 'wet to dry' tasks, where training-based detection methods heavily rely on supplemented transformed SSVEP data to address the low quality of calibration data. Similar to observations in Fig. 6, Concat. and LST exhibit negative transfer effects. Moreover, we notice that as the quality of source data improves, LST-based methods slightly enhance performance, especially in scenarios like 'benchmark', 'wet to wet', and 'wet to dry'. Conversely, with low-quality source data, the negative impact of LST on TRCA-based methods intensifies, particularly in 'dry to dry' and 'dry to wet' scenarios. In contrast, SSVEP-DAN consistently improves SSVEP detection performance across different numbers of source subjects.

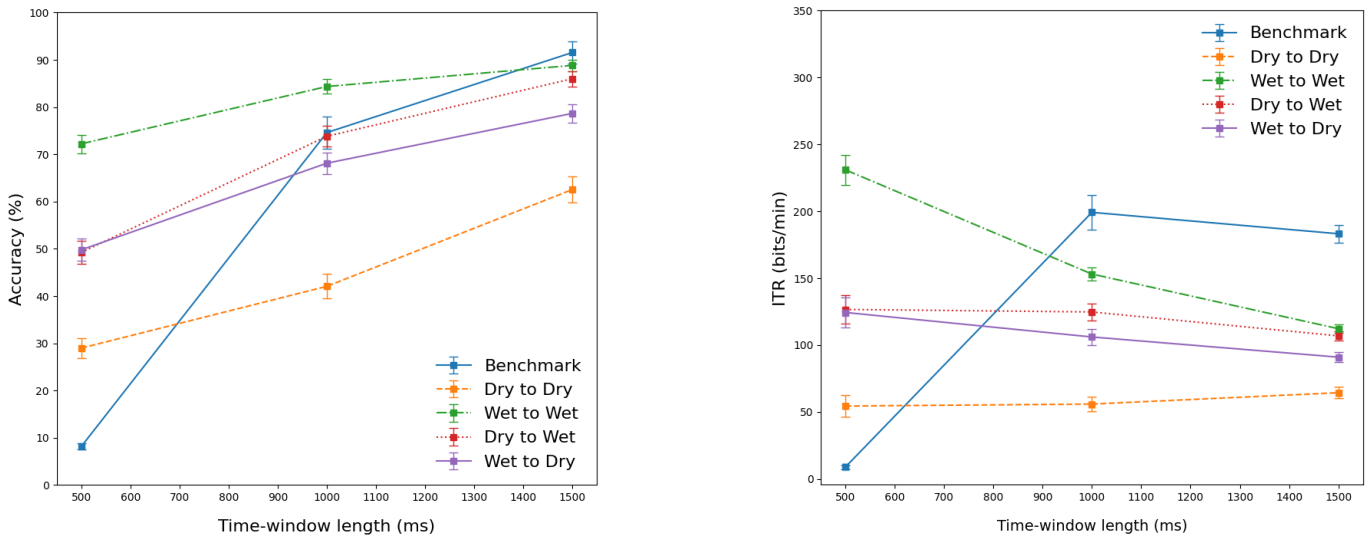
C. Time-Window Length of SSVEP

Figures 8 (a) and (b) depict the accuracy and information transfer rate (ITR) of all subjects across each dataset under varying time-window lengths (T_w) of 0.5, 1.0, and 1.5 seconds. Throughout the experiment, calibration trials were fixed at two. The results across different datasets indicate that TRCA calibrated using SSVEP-DAN transformed data achieves highest accuracy at $T_w = 1.5$s. Among the four sub-datasets under the 'Wearable' category, 'Wet to Wet' consistently demonstrates the highest accuracy, while 'Dry to Dry' exhibits the lowest accuracy due to the relatively better signal quality of wet electrodes compared to dry electrodes. While the accuracy increases with the window length, the overall growth range in terms of ITR is limited. The lower accuracy of SSVEP-DAN with a time-window length of 0.5 seconds can be attributed to several factors, including insufficient features, potential noise, and consequently, lower quality of the aligned source data converted from the source domain.

D. Ablation Study

We performed an ablation study to assess SSVEP-DAN's main training strategies and model components. Table III displays the results, comparing different training methods and model configurations with the proposed method and a baseline (no adaptation).

In the 'w/o stimulus-independent training' approach, SSVEP-DAN was trained to align data of each stimulus specifically. In the 'w/o pre-training phase' approach, we skipped the pre-training phase and trained SSVEP-DAN using source data individually in the fine-tuning phase. In the 'w/o fine-tuning phase' approach, we trained the model using data



(a) (b)

Fig. 8. Performance of the proposed SSVEP-DAN across various time-window lengths (T_w) in (a) accuracy and (b) information transfer rate (ITR).

TABLE III

ABLATION ANALYSIS ON THE EFFICACY OF THE KEY ELEMENTS IN THE PROPOSED SSVEP-DAN FRAMEWORK. THE ASTERISKS INDICATE A SIGNIFICANT DIFFERENCE BETWEEN THE PROPOSED METHOD AND THE COUNTERPART. ($^*p<0.05$, $^{**}p<0.01$, $^{***}p<0.001$)

Methods	Benchmark	Dry to dry	Wet to wet	Dry to wet	Wet to dry
Proposed	91.56	62.55	88.67	86.01	78.67
(w/o) stimulus-independent training	84.79^{***}	39.09^{***}	87.72	81.34^{**}	63.02^{***}
(w/o) pre-training phase	90.60	53.43^{***}	88.28^{***}	82.67^{***}	69.65^{***}
(w/o) fine-tuning phase	89.15^{***}	34.55^{***}	85.24^{***}	75.15^{***}	52.67^{***}
(w/o) non-linear activation function	91.31	44.71^{***}	87.49^{*}	84.75^{*}	69.02^{***}
(w/) temporal convolution	58.03^{***}	28.79^{***}	38.19^{***}	38.35^{**}	25.48^{***}
Baseline (w/o adaptation)	74.68^{***}	28.45^{***}	76.21^{***}	—	—

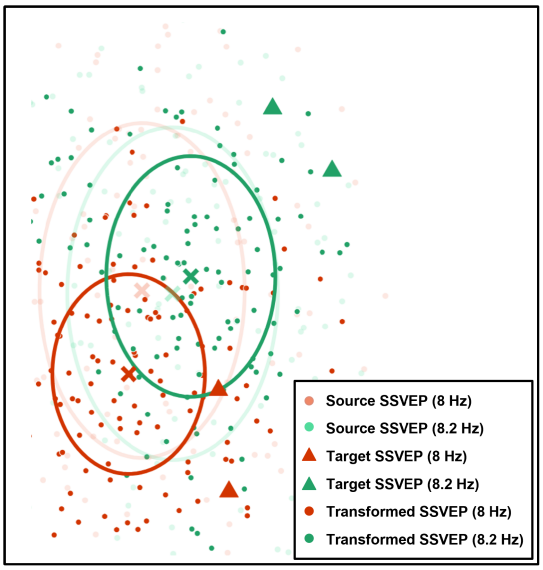
from multiple source domains to obtain the pre-training model without fine-tuning on each source domain. For validating the model architecture, we conducted the 'w/o non-linear activation function' method to reduce non-linearity introduced by the hyperbolic tangent activation layer. Additionally, we included the 'w/ temporal convolution' method, incorporating an additional temporal convolutional layer to increase model complexity.

In the majority of cases for both Dataset I and Dataset II, complete SSVEP-DAN significantly outperforms the training method without stimulus-independent training. This is mainly due to cross-stimulus training, which combines data from different stimuli during training, resulting in a more robust model, particularly with limited calibration data. Furthermore, our proposed training method consistently outperforms each single-stage training method (without pre-training or fine-tuning) in most cases for both datasets. This indicates that pre-training enables SSVEP-DAN to capture common features from SSVEP data and improve generalization, while avoiding overfitting to specific subjects. Moreover, fine-tuning the pre-trained model based on individual subject characteristics allows it to adapt to representations specific to source-target subject pairs. These two primary training methods are indispensable in SSVEP-DAN as they provide data augmentation and robustness, enhancing its performance in aligning SSVEP data across subjects and stimuli.

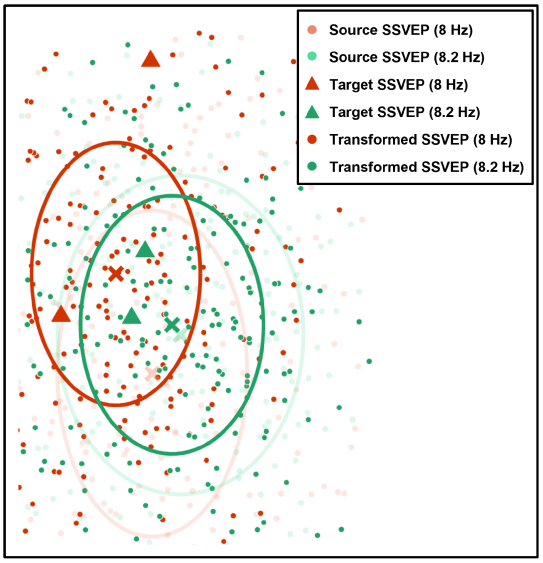
E. Visualization of Data Alignment

We employed t-distributed stochastic neighbor embedding (t-SNE) [49] to visualize SSVEP data before and after alignment by SSVEP-DAN in two-dimensional scatters.

Figure 9 (a) displays the t-SNE visualization of subject S3 as the target subject in Dataset I, with all other subjects as source subjects. Figure 9 (b) illustrates the t-SNE visualization of subject S1 as the target subject wearing a dry electrode device in the 'wet to dry' scenario of Dataset II. These figures reveal that under the same stimulus, the clusters of transformed SSVEP data are smaller and more separated compared to the clusters of source data. This suggests that SSVEP-DAN reduces inter-subject variability and increases inter-stimulus variability, aligning data more effectively.



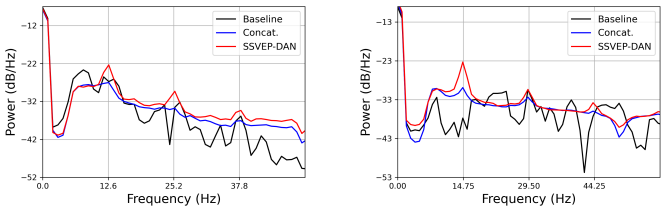
(a) Benchmark



(b) Wet to dry

Fig. 9. Visualization of source, target, and transformed SSVEP data using t-SNE for tasks (a) 'Benchmark' and (b) 'Wet to Dry'. We selected SSVEP data for two stimuli ('benchmark': 8 and 8.2 Hz; 'wet to dry': 9.25 and 9.75 Hz), represented by red and green dots, respectively, across subjects to demonstrate the impact of SSVEP data alignment. Red/green triangles denote target SSVEP data (two trials per stimulus). Light and dark shading differentiate between source and transformed SSVEP data, with class centroids marked by a cross and standard deviations delineated by dotted circles. The shift in SSVEP data clusters illustrates the effect of data alignment by SSVEP-DAN.

Furthermore, power spectrum analysis was conducted to examine the effect of data alignment on power spectral density (PSD) using SSVEP-DAN. Figures 10 (a) depict the PSD of 12.6 Hz SSVEP signals for subject S3 in Dataset I under three schemes (Baseline, Concat., and SSVEP-DAN), with two calibration trials for each stimulus. Similarly, Figures 10 (b) show the PSD of 14.75 Hz SSVEP signals for subject S1 wearing a dry electrode device in the 'wet to dry' scenario of Dataset II, with two calibration trials for each stimulus under the three schemes. In the Concat. scheme, we observe varying outcomes. Figure 10 (a) indicates that on Dataset I,



(a) Benchmark (b) Wet to dry

Fig. 10. Averaged power spectrum density of the SSVEP data obtained under different conditions of domain adaptation in the (a) 'Benchmark' scenario with the 12.6 Hz stimulus and (b) 'wet to dry' scenario with the 14.75 stimulus. 'Baseline': average across two trials of target data only; 'Concat.': average across two trials of target data plus original source data; 'SSVEP-DAN': average across two trials of target data plus transformed data.

the Concat. scheme struggles to produce stable spectra due to high SSVEP trial variability, resulting in less concentrated spectral peaks. Conversely, Figure 10 (b) demonstrates that on Dataset II, where target participants have suboptimal signal-to-noise ratio (SNR), the Concat. scheme can marginally improve SNR by integrating a large amount of higher-quality data, leading to more stable spectra. The SSVEP-DAN scheme consistently shows significant enhancements in peak amplitudes at the target frequency and its harmonics in both figures. These results suggest that SSVEP-DAN effectively reduces inter-trial variability, enabling the utilization of non-target subject trials and increasing SNR. Moreover, these observations align with decoding accuracy trends (Figure 6).

V. Discussion

A. Impact of Data Quality and Variability

In Figure 6, the naive transfer learning method (Concat. scheme) exhibits varying effects on TRCA-based methods in different scenarios, such as 'benchmark', 'dry to wet', or 'wet to wet'. However, it shows a positive impact when dry electrode devices are used for recording new subjects. This observation suggests that both data variability and quality may concurrently influence the performance of training-based detection methods. When new user data quality is high, the impact of data variability becomes more prominent, as high-quality data may be more sensitive to subtle variations. Conversely, when the data quality of new users is relatively low (i.e., dry-electrode EEG data), there is a greater demand for higher-quality data, and simultaneously, the influence of data variability becomes less significant.

Therefore, when training-based detection methods can accurately classify using high-quality data from new users, the negative impact of the Concat. scheme is exacerbated due to the heightened influence of data variability. In contrast, when training-based detection methods struggle to accurately classify using low-quality data from new users, acquiring a larger dataset of higher quality mitigates the negative impact of the Concat. scheme.

B. Impact of Model Architecture

The result of ablation study illustrates that our proposed model design consistently outperforms the design without a non-linear activation function across most scenarios in both

Dataset I and Dataset II. This observed improvement can be attributed to several factors. The utilization of the hyperbolic tangent activation function is beneficial for training performance as it could mitigate signal drift or bias in the data by concentrating data near zero, offering an improvement in managing data variations [50], [51]. Furthermore, it aids in the normalization and constraint of high-amplitude EEG data, thereby enhancing network stability [50]. Regarding the complexity of the model architecture, we observed that incorporating the temporal convolutional layer deteriorates the performance in both Dataset I and Dataset II. This could be relevant to the excessive complexity of the model that leads to overfitting on the training data. This risk is particularly pronounced when the availability of EEG data is limited, making overly complex models susceptible to overfitting [52].

C. Limitation and Future Work

One significant constraint of the proposed SSVEP-DAN stems from the necessity to synchronize stimuli between the source and target domains during SSVEP-DAN training, which hampers its adaptability in scenarios with diverse stimuli. Additionally, the reliance on a small number of calibration trials for new users impedes its calibration-free usability, thereby affecting its user-friendliness.

Regarding future directions, several promising avenues for advancing SSVEP-DAN exist. Firstly, further validation of its effectiveness with state-of-the-art SSVEP classifiers like Compact-CNN [4], Conv-CA [20], DNN [39], and SSVEPNet [53] is essential. Simultaneously, exploring its applicability to other time-locked datasets such as BCI Challenge ERN dataset (BCI-ERN) [54] can enhance its versatility in time-locked signal decoding. Alternatively, SSVEP-DAN can be applied to other scenarios, such as the AR-SSVEP problem described in the paper [55]. The method proposed in this paper can potentially solve this problem. A SSVEP-DAN can be pre-trained to align different brightness levels to the same brightness level in real time. This should improve the performance in this scenario when used with a subsequent SSVEP decoding algorithm. Additionally, researching standardized criteria for source subject selection, considering factors like classification accuracy, SNR, and similarity to target subject, promises to improve SSVEP-DAN's transferability. Lastly, in-depth investigations into the neuroscience underpinnings of model design can provide deeper insights into model interpretation.

VI. Conclusion

In this study, we introduced a novel neural network-based data alignment method, which incorporates innovative training strategies, notably cross-stimulus training and pre-training techniques. Our experimental findings illustrate that the SSVEP-DAN approach enhances subject similarity and boosts SSVEP decoding accuracy by effectively leveraging data from non-target subjects. Moreover, the results of our ablation studies suggest that by employing cross-stimulus training and pre-training techniques, we can further enhance stability in performance improvement. Overall, we anticipate that our proposed SSVEP-DAN will find application in SSVEP-based BCI spellers, where it can improve SSVEP decoding performance by utilizing calibration data from a small number of calibration trials from new subjects and supplementary calibration data from non-target subjects.

References

[1] D. Regan, "Steady-state evoked potentials," *J. Opt. Soc. Am.*, vol. 67, no. 11, pp. 1475–1489, Nov. 1977.

[2] A. M. Norcia, L. G. Appelbaum, J. M. Ales, B. R. Cottereau, and B. Rossion, "The steady-state visual evoked potential in vision research: A review," *J. Vis.*, vol. 15, no. 6, p. 4, May 2015.

[3] N. R. Waytowich and D. J. Krusienski, "Multiclass steady-state visual evoked potential frequency evaluation using chirp-modulated stimuli," *IEEE Trans. Hum.-Mach. Syst.*, vol. 46, no. 4, pp. 593–600, Aug. 2016.

[4] N. Waytowich et al., "Compact convolutional neural networks for classification of asynchronous steady-state visual evoked potentials," *J. Neural Eng.*, vol. 15, no. 6, Dec. 2018, Art. no. 066031.

[5] J. R. Wolpaw, "Brain–computer interfaces (BCIs) for communication and control," in *Proc. 9th Int. ACM SIGACCESS Conf. Comput. Accessibility*, Oct. 2007, pp. 1–2.

[6] N. Birbaumer, "Breaking the silence: Brain–computer interfaces (BCI) for communication and motor control," *Psychophysiology*, vol. 43, no. 6, pp. 517–532, Nov. 2006.

[7] M. Cheng, X. Gao, S. Gao, and D. Xu, "Design and implementation of a brain–computer interface with high transfer rates," *IEEE Trans. Biomed. Eng.*, vol. 49, no. 10, pp. 1181–1186, Oct. 2002.

[8] X. Chen, Y. Wang, M. Nakanishi, X. Gao, T.-P. Jung, and S. Gao, "High-speed spelling with a noninvasive brain–computer interface," *Proc. Nat. Acad. Sci. USA*, vol. 112, no. 44, pp. E6058–E6067, Nov. 2015.

[9] I. Martišius and R. Damaševič ius, "A prototype SSVEP based real time BCI gaming system," *Comput. Intell. Neurosci.*, vol. 2016, pp. 1–15, Jun. 2016.

[10] S.-C. Chen, Y.-J. Chen, I. A. E. Zaeni, and C.-M. Wu, "A single-channel SSVEP-based BCI with a fuzzy feature threshold algorithm in a maze game," *Int. J. Fuzzy Syst.*, vol. 19, no. 2, pp. 553–565, Apr. 2017.

[11] S. M. T. Müller, T. F. Bastos-Filho, and M. Sarcinelli-Filho, "Using a SSVEP-BCI to command a robotic wheelchair," in *Proc. IEEE Int. Symp. Ind. Electron.*, Jun. 2011, pp. 957–962.

[12] A. Güneysu and H. L. Akin, "An SSVEP based BCI to control a humanoid robot by using portable EEG device," in *Proc. 35th Annu. Int. Conf. IEEE Eng. Med. Biol. Soc. (EMBC)*, Jul. 2013, pp. 6905–6908.

[13] Z. Lin, C. Zhang, W. Wu, and X. Gao, "Frequency recognition based on canonical correlation analysis for SSVEP-based BCIs," *IEEE Trans. Biomed. Eng.*, vol. 54, no. 6, pp. 1172–1176, Jun. 2007.

[14] G. Bin, X. Gao, Z. Yan, B. Hong, and S. Gao, "An online multi-channel SSVEP-based brain–computer interface using a canonical correlation analysis method," *J. Neural Eng.*, vol. 6, no. 4, Jun. 2009, Art. no. 046002, doi: 10.1088/1741-2560/6/4/046002.

[15] M. Nakanishi, Y. Wang, X. Chen, Y.-T. Wang, X. Gao, and T.-P. Jung, "Enhancing detection of SSVEPs for a high-speed brain speller using task-related component analysis," *IEEE Trans. Biomed. Eng.*, vol. 65, no. 1, pp. 104–112, Jan. 2018.

[16] K.-J. Chiang, C. M. Wong, F. Wan, T.-P. Jung, and M. Nakanishi, "Reformulating task-related component analysis for reducing its computational complexity," *Biomed. Signal Process. Control*, vol. 86, 2023, Art. no. 105220.

[17] S. Ladouce, L. Darmet, J. J. Torre Tresols, S. Velut, G. Ferraro, and F. Dehais, "Improving user experience of SSVEP BCI through low amplitude depth and high frequency stimuli design," *Sci. Rep.*, vol. 12, no. 1, p. 8865, May 2022.

[18] L. Jiang, W. Pei, and Y. Wang, "A user-friendly SSVEP-based BCI using imperceptible phase-coded flickers at 60Hz," *China Commun.*, vol. 19, no. 2, pp. 1–14, Feb. 2022.

[19] X. Bai, M. Li, S. Qi, A. C. M. Ng, T. Ng, and W. Qian, "A hybrid P300-SSVEP brain–computer interface speller with a frequency enhanced row and column paradigm," *Frontiers Neurosci.*, vol. 17, Mar. 2023, Art. no. 1133933.

[20] Y. Li, J. Xiang, and T. Kesavadas, "Convolutional correlation analysis for enhancing the performance of SSVEP-based brain–computer interface," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 28, no. 12, pp. 2681–2690, Dec. 2020.

[21] X. Zhang, S. Qiu, Y. Zhang, K. Wang, Y. Wang, and H. He, "Bidirectional Siamese correlation analysis method for enhancing the detection of SSVEPs," *J. Neural Eng.*, vol. 19, no. 4, Aug. 2022, Art. no. 046027.

[22] T. Cao, F. Wan, C. M. Wong, J. N. da Cruz, and Y. Hu, "Objective evaluation of fatigue by EEG spectral analysis in steady-state visual evoked potential-based brain–computer interfaces," *Biomed. Eng. OnLine*, vol. 13, no. 1, pp. 1–12, Dec. 2014.

[23] K.-J. Chiang, C.-S. Wei, M. Nakanishi, and T.-P. Jung, "Boosting template-based SSVEP decoding by cross-domain transfer learning," *J. Neural Eng.*, vol. 18, no. 1, Feb. 2021, Art. no. 016002, doi: 10.1088/1741-2552/abcb6e.

[24] S. J. Pan and Q. Yang, "A survey on transfer learning," *IEEE Trans. Knowl. Data Eng.*, vol. 22, no. 10, pp. 1345–1359, Jun. 2010.

[25] P. L. C. Rodrigues, C. Jutten, and M. Congedo, "Riemannian procrustes analysis: Transfer learning for brain–computer interfaces," *IEEE Trans. Biomed. Eng.*, vol. 66, no. 8, pp. 2390–2401, Aug. 2019.

[26] M. Nakanishi, Y.-T. Wang, C.-S. Wei, K.-J. Chiang, and T.-P. Jung, "Facilitating calibration in high-speed BCI spellers via leveraging cross-device shared latent responses," *IEEE Trans. Biomed. Eng.*, vol. 67, no. 4, pp. 1105–1113, Apr. 2020.

[27] B. Liu, X. Chen, X. Li, Y. Wang, X. Gao, and S. Gao, "Align and pool for EEG headset domain adaptation (ALPHA) to facilitate dry electrode based SSVEP-BCI," *IEEE Trans. Biomed. Eng.*, vol. 69, no. 2, pp. 795–806, Feb. 2022.

[28] A. Bleuzé, J. Mattout, and M. Congedo, "Tangent space alignment: Transfer learning for brain-computer interface," *Frontiers Hum. Neurosci.*, vol. 16, pp. 16–1049, Dec. 2022.

[29] N. K. Nik Aznan, A. Atapour-Abarghouei, S. Bonner, J. D. Connolly, N. A. Moubayed, and T. P. Breckon, "Simulating brain signals: Creating synthetic EEG data via neural-based generative models for improved SSVEP classification," in *Proc. Int. Joint Conf. Neural Netw. (IJCNN)*, Budapest, Hungary, 2019, pp. 1–8.

[30] J. Kwon and C.-H. Im, "Novel signal-to-signal translation method based on StarGAN to generate artificial EEG for SSVEP-based brain–computer interfaces," *Expert Syst. Appl.*, vol. 203, Oct. 2022, Art. no. 117574.

[31] Y. Pan, N. Li, Y. Zhang, P. Xu, and D. Yao, "Short-length SSVEP data extension by a novel generative adversarial networks based framework," 2023, *arXiv:2301.05599*.

[32] C. M. Wong, B. Wang, Z. Wang, K. F. Lao, A. Rosa, and F. Wan, "Spatial filtering in SSVEP-based BCIs: Unified framework and new improvements," *IEEE Trans. Biomed. Eng.*, vol. 67, no. 11, pp. 3057–3072, Nov. 2020.

[33] G. Sarafraz, A. Behnamnia, M. Hosseinzadeh, A. Balapour, A. Meghrazi, and H. R. Rabiee, "Domain adaptation and generalization on functional medical images: A systematic survey," 2022, *arXiv:2212.03176*.

[34] Z. Wan, R. Yang, M. Huang, N. Zeng, and X. Liu, "A review on transfer learning in EEG signal analysis," *Neurocomputing*, vol. 421, pp. 1–14, Jan. 2021.

[35] D. Wu, Y. Xu, and B.-L. Lu, "Transfer learning for EEG-based brain–computer interfaces: A review of progress made since 2016," *IEEE Trans. Cognit. Develop. Syst.*, vol. 14, no. 1, pp. 4–19, Mar. 2022.

[36] D. H. Johnson, "Signal-to-noise ratio," *Scholarpedia*, vol. 1, no. 12, p. 2088, 2006.

[37] C.-S. Wei, T. Koike-Akino, and Y. Wang, "Spatial component-wise convolutional network (SCCNet) for motor-imagery EEG classification," in *Proc. 9th Int. IEEE/EMBS Conf. Neural Eng.*, Jun. 2019, pp. 328–331.

[38] A. Ravi, N. H. Beni, J. Manuel, and N. Jiang, "Comparing user-dependent and user-independent training of CNN for SSVEP BCI," *J. Neural Eng.*, vol. 17, no. 2, Apr. 2020, Art. no. 026028.

[39] O. B. Guney, M. Oblokulov, and H. Ozkan, "A deep neural network for SSVEP-based brain–computer interfaces," *IEEE Trans. Biomed. Eng.*, vol. 69, no. 2, pp. 932–944, Feb. 2022.

[40] D. P. Kingma and J. Ba, "Adam: A method for stochastic optimization," 2014, *arXiv:1412.6980*.

[41] X. Chen, Y. Wang, S. Gao, T.-P. Jung, and X. Gao, "Filter bank canonical correlation analysis for implementing a high-speed SSVEP-based brain–computer interface," *J. Neural Eng.*, vol. 12, no. 4, Aug. 2015, Art. no. 046008.

[42] Y. Wang, X. Chen, X. Gao, and S. Gao, "A benchmark dataset for SSVEP-based brain–computer interfaces," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 25, no. 10, pp. 1746–1752, Oct. 2017.

[43] F. Zhu, L. Jiang, G. Dong, X. Gao, and Y. Wang, "An open dataset for wearable SSVEP-based brain–computer interfaces," *Sensors*, vol. 21, no. 4, p. 1256, Feb. 2021.

[44] P. R. A. S. Bassi and R. Attux, "FBDNN: Filter banks and deep neural networks for portable and fast brain–computer interfaces," *Biomed. Phys. Eng. Exp.*, vol. 8, no. 3, May 2022, Art. no. 035018.

[45] V. Mihajlović, G. Garcia-Molina, and J. Peuscher, "Dry and water-based EEG electrodes in SSVEP-based BCI applications," in *Proc. 5th Int. Joint Conf*, 2012, pp. 23–40.

[46] X. Xing et al., "A high-speed SSVEP-based BCI using dry EEG electrodes," *Sci. Rep.*, vol. 8, no. 1, p. 14708, Oct. 2018.

[47] W. Zhang, L. Deng, L. Zhang, and D. Wu, "A survey on negative transfer," *IEEE/CAA J. Autom. Sinica*, vol. 10, no. 2, pp. 305–329, Feb. 2023.

[48] M. Nakanishi, Y. Wang, Y.-T. Wang, Y. Mitsukura, and T.-P. Jung, "A high-speed brain speller using steady-state visual evoked potentials," *Int. J. Neural Syst.*, vol. 24, no. 6, Sep. 2014, Art. no. 1450019.

[49] L. Van der Maaten and G. Hinton, "Visualizing data using t-SNE," *J. Mach. Learn. Res.*, vol. 9, no. 11, pp. 1–18, 2008.

[50] A. Nagarajan, N. Robinson, and C. Guan, "Investigation on robustness of EEG-based brain–computer interfaces," in *Proc. 43rd Annu. Int. Conf. IEEE Eng. Med. Biol. Soc. (EMBC)*, Nov. 2021, pp. 6334–6340.

[51] D. Xu, F. Tang, Y. Li, Q. Zhang, and X. Feng, "An analysis of deep learning models in SSVEP-based BCI: A survey," *Brain Sci.*, vol. 13, no. 3, p. 483, Mar. 2023.

[52] C. He, J. Liu, Y. Zhu, and W. Du, "Data augmentation for deep neural networks model in EEG classification task: A review," *Frontiers Hum. Neurosci.*, vol. 15, Dec. 2021, Art. no. 765525.

[53] Y. Pan, J. Chen, Y. Zhang, and Y. Zhang, "An efficient CNN-LSTM network with spectral normalization and label smoothing technologies for SSVEP frequency recognition," *J. Neural Eng.*, vol. 19, no. 5, Sep. 2022, Art. no. 056014, doi: 10.1088/1741-2552/ac8dc5.

[54] P. Margaux, M. Emmanuel, D. Sébastien, B. Olivier, and M. Jérémie, "Objective and subjective evaluation of online error correction during P300-based spelling," *Adv. Hum.-Comput. Interact.*, vol. 2012, pp. 1–13, Jul. 2012.

[55] R. Zhang et al., "Improving AR-SSVEP recognition accuracy under high ambient brightness through iterative learning," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 31, pp. 1796–1806, 2023.