Across Sessions and Subjects Domain Adaptation for Building Robust Myoelectric Interface

Wei Li, Xinran Zhang, Ping Shi, Sujiao Li, Ping Li, Hongliu Yu*

***Abstract*— Gesture interaction via surface electromyography (sEMG) signal is a promising approach for advanced human-computer interaction systems. However, improving the performance of the myoelectric interface is challenging due to the domain shift caused by the signal's inherent variability. To enhance the interface's robustness, we propose a novel adaptive information fusion neural network (AIFNN) framework, which could effectively reduce the effects of multiple scenarios. Specifically, domain adversarial training is established to inhibit the shared network's weights from exploiting domain-specific representation, thus allowing for the extraction of domain-invariant features. Effectively, classification loss, domain diversence loss and domain discrimination loss are employed, which improve classification performance while reduce distribution mismatches between the two domains. To simulate the application of myoelectric interface, experiments were carried out involving three scenarios (intra-session, inter-session and inter-subject scenarios). Ten able-bodied subjects were recruited to perform sixteen gestures for ten consecutive days. The experimental results indicated that the performance of AIFNN was better than two other state-of-the-art transfer learning approaches, namely fine-tuning (FT) and domain adversarial network (DANN). This study demonstrates the capability of AIFNN to maintain robustness over time and generalize across users in practical myoelectric interface implementations. These findings could serve as a foundation for future deployments.**

***Index Terms*— Myoelectric interface, surface electromyography, domain adaptation, robustness, domain adversarial training.**

I. Introduction

WITH the increasing prominence of the Internet-of-Things systems in the lives of human beings, human-computer interaction is evolving in order to achieve better synergy between human intention and machine behavior. Hands, as the most dexterous limbs in the human locomotor system, can provide a natural and intuitive way for human-computer interaction through various hand gestures. While computer vision, WiFi, and data gloves have been acknowledged as methods, it is recommended to depend on biosignals recorded from the body for continuous use [1], [2]. Surface electromyography (sEMG) is an improved option since it has several characteristics, including high signal-to-noise ratio, human-machine friendliness, and low interference noise. sEMG is very suitable for hands-free interaction in virtual reality and augmented reality, control of robotic arms and prostheses, and has broad applications [3]. Prospects for achieving a bridge between signal and robot input guidance rely on artificial intelligence, specifically deep learning (DL) [4]. Studies have shown that DL has been successfully applied in many fields, including face recognition, speech recognition, human pose estimation, and indoor localization. However, sEMG signals pose additional challenges for DL.

The application of myoelectric interface requires a long learning curve or a high concentration of subjects due to the inherent instability caused by physiological factors and external environment, such as muscle fatigue, cross-talk, sweating and electrode shift, leading to insufficient robustness in application. In real life, the cumbersome classifier training and time-consuming repeated recalibration procedures seriously hinder the deployment of myoelectric interface [5]. In recent years, researchers have focused on resolving the challenge, trying to make the myoelectric interface know when and how to adapt properly and shorten recalibration time. Notably, deep learning techniques have demonstrated significant potential in the myoelectric interface, with the most successful application is convolutional neural network (CNN) [6] [7] [8] [9] [10]. From the perspective of machine learning, the most direct approach to improve the robustness and generalization is to incorporate sufficient data to train a classifier that encompasses all variations, a process that is both time-consuming and labor-intensive, and may ultimately prove unfeasible [11].

Recently, transfer learning (TL) has been utilized to prevent degradation arising from potential changes in sEMG signals [12]. There are many approaches for transfer learning, such as ensemble learning, common spatial-spectral analysis, covariate shift adaptation technique and deep domain adaptation [13] [14] [15]. Compared with deep domain adaptation, other approaches heavily rely on feature engineering, which leads to their limited capacity for multi-domain data learning and uncertain efficacy for high-performance inter-subject learning model [44]. In domain adaptation, the change of session or subject is referred to as domain shift, while different sessions or subjects are considered different domains [5]. Deep domain adaptation is not training a new model, but using insufficient recalibrated data to learn the consistent knowledge between the original (source domain $\mathcal{D}_S$) and new data space (target

W. Li, X. R. Zhang, P. Shi, S.J. Li, P. Li and H. L. Yu* are with Institute of Rehabilitation Engineering and Technology, University of Shanghai for Science and Technology, Shanghai, 200093, P.R. China (*corresponding author: Hongliu Yu, e-mail: yhl98@hotmail.com).

This work was supported by the National Key R&D Program of China (2020YFC2007902).

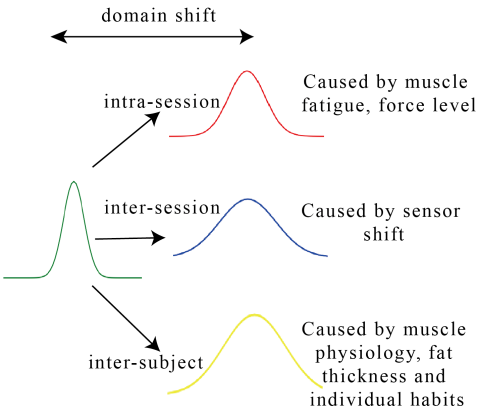


Fig. 1: Domain shift in case of different scenarios

domain $\mathcal{D}_T$), so as to quickly enhance the original model and adapt to the new data distribution. Domain adaptation is a prevalent deep transfer learning scheme, which aims at learning a discriminative predictor by reducing the discrepancy between different distributions of $\mathcal{D}_S$ and $\mathcal{D}_T$ and aligning feature distributions of $\mathcal{D}_S$ and $\mathcal{D}_T$. For the prosperous deployment and commercial application of myoelectric interface, robustness across time and generalizability across users are both critical factors. As opposed to previous studies that only focused on a single scenario, a comprehensive and systematic analysis of performance across multiple scenarios is necessary to provides guidance for future studies.

The main contribution of this research is introducing a novel domain adaptation scheme called the Adaptive Information Fusion Neural Network (AIFNN), which is based on CNN and Siamese architecture. This approachenables learning a distinct representation across multiple domains simultaneously without relying on the target domain label, while explicitly penalizing any domain-variant representation, maximizing the discovery and exploitation of common features of sEMG signals across multiple sessions or multiple subjects and diminishing any domain transfer effects. To tune CNN weights effectively, three types of loss functions are employed, including the classification loss for supervised learning in $\mathcal{D}_S$ that has a clear label while $\mathcal{D}_t$ does not, a domain divergence loss to reduce marginal distribution mismatches between two domains in the latent space, and a domain discrimination loss for $\mathcal{D}_S$ and $\mathcal{D}_T$ to reduce conditional distribution mismatches between two domains.

Another significant contribution of this research is the release of a new long-term multi-subject sEMG-based gesture classification dataset(https://github.com/tinker1017/long-term-multi-subject-sEMG-dataset), which contains 16 gestures performed by 10 able-bodied participants. Using the dataset, this research investigates three scenarios: (1) intra-session myoelectric interface, instability mainly stemming from muscle fatigue or force changes [17], (2) inter-session myoelectric interface, instability mainly due to electrode shift [18], [19] and (3) inter-subject myoelectric interface, instability related to muscle physiology, fat thickness and individual habits [20], [21], as indicated in Fig. 1.

The rest of this paper is structured as follows: Section II presents relevant research relevant for this work. Section III demonstrates the design and conduct of long-term sEMG signal experiments, Section IV describes the design of AIFNN, including the associated signal preprocessing, training methods and comparison methods, Section V presents and analyzes the related experimental results, and in Section VI, the related discussion of the experimental results is presented.

II. Related Work And Mathematical Notations

This work is mainly related to sEMG-based gesture recognition using transfer learning. Therefore, we sort out the currently available related works and summarize the main contributions and limitations.

For intra-session variations, the application of myoelectric interface is based on the fundamental assumption that the sEMG mode is repeatable for the same movement and separable for different movements. Previous work has assumed that over 90% of the overall accuracy is well feasible [4]. However, the clinical applications and user acceptance of this scenario are still limited. Several studies have attempted to use an optimized CNN to enhance the generalizability of the system, but these efforts either have limited improvements or use small datasets that are insufficient to demonstrate their effectiveness [45] [23].

To address session-to-session nonstationarity, Zhai *et al.* [24] proposed an automatically updating self-recalibration strategy that could significantly improve intra-session accuracy for both amputees and able-bodied subjects [25], yet such strategy would increase latency. Du *et al.* [26] handled the intra-session and inter-session problem using a multi-stream extension of AdaBN, which is an unsupervised domain adaptation learning applying to both sparse multi-channel and high-density sEMG [27]. Côté-Allard *et al.* [28] proposed the self-calibrating asynchronous domain adversarial neural network (SCADANN) to improve the performance of multi-session system, but pseudo-labeling heuristic are easy to accumulate error effects and need to be improved in long-term experiments.

For high variability between different subjects, in [9], a Continuous Wavelet Transform-based progressive neural network (PNN) architecture was applied on a complex gesture dataset and trained to solve inter-subject problems. Côté-Allard *et al.* [29] proposed the adaptive domain adversarial neural network (ADANN) to improve the inter-subject performance, and compared the topological structures of deep learning-based features with hand-crafted features. Campbell *et al.* [30] proved that ADANN could as an novel multi-subject classification model to realize advanced performance with very little retraining data, which would outperform the state-of-the-art model. Bao *et al.* [31] proposed a regression supervised domain adaptation (SDA) approach for wrist kinematics estimation using sEMG signal, which effectively reduced the burden of recalibration.

For simultaneous recognition of multi-session and multi-subject, Sosin *et al.* [32] employed RNN and adversarial domain adaptation (ADA) to estimate continuous gestures, which improved the accuracy between subjects but reduced the

TABLE I: List of Mathematical Notations.

Notation	Definition
$\mathcal{D}_S$	Source domain dataset
$\mathcal{D}_T$	Target domain dataset
n_s,n_t	Length of the source and target domain dataset
$\boldsymbol{X}$	sEMG dataset
y	Gesture labels
M, N	Data length of the source domain, target domain, $N \ll M$
W	Length of sliding window
C	Number of electrodes
λ	Penalty for gradient reversal layer(GRL)
$R_\lambda(r)$	The forward output of GRL
L,$\mathcal{L}_c$, $\mathcal{L}_d$, $\mathcal{L}_l$	The total loss, classification loss, domain divergence loss, and domain discrimination loss
λ_u, λ_l	Penalty for domain divergence loss, domain discrimination loss
$\boldsymbol{\theta}$	CNN weights
$\boldsymbol{\Theta}_c^S$	Weights used to optimise classification loss
$\boldsymbol{\Theta}_d^S$, $\boldsymbol{\Theta}_d^T$	Weights used to optimise domain divergence loss between two domains
$\boldsymbol{\Theta}_l^S$, $\boldsymbol{\Theta}_l^T$	Weights used to optimise domain discrimination loss in two domains

accuracy between sessions. Côté-Allard *et al.* [33] proposed PNN with a multi-stream AdaBatch scheme to extract stable and general features across subjects as much as possible. To achieve multi-day and multi-subject zero training, Sheng *et al.* [34] proposed common spatial-spectral analysis for signal processing. However, the multi-day verification and multi-subject verification in the paper are carried out on two different datasets.

Besides, we introduce and summarize the mathematical notations that we used in this paper with Table I.

III. Experiment Methods

A. Experiment Setup

Ten able-bodied subjects (mean age in years: 22(± 2.8), 5 males and 5 females) were recruited. Exclusion criteria were any neurological pathology or musculoskeletal complaints interfering with study outcomes. Each subject received written informed consent before the experiment. All participants conducted the experiment with their right hand. Participants had no relevant professional knowledge of sEMG and had not received corresponding training before the experiment.

B. Acquisition Protocol

sEMG signals were recorded using the commercial wearable gForcePro Armband, developed by OYMotion Technologies Co., Ltd (Shanghai, China, https://www.oymotion.com/), which are consists of eight dry electrodes with a sampling frequency of 1000 Hz and a nine-axis inertial measurement unit (IMU). In this work, only sEMG was utilized to study. The gForcePro armband is worn on the forearm, about 3cm from the elbow. Before wearing the armband, the skin of the forearm needs to be cleaned with alcohol. The subjects sat comfortably in the office chair with arms used in the experiment on the table and were expected to complete each trial

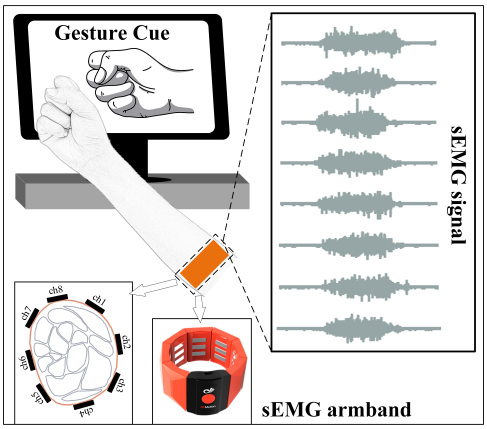


Fig. 2: Schematic of experimental environment

following the imitation stimulus displayed on the computer monitor, as shown in Fig. 2. To provide a more security and reliable system, all movements were recorded under the most appropriate contractile force of subjects.

The experiment involved the signal acquisition of 16 movements. Each trial is collected with 5 seconds of motion execution and 3 seconds of rest as well as 10 repetitions in a single session. All experimental sessions was completed over 10 consecutive days, and two adjacent sessions was conducted with a 24-hour interval. Markers were set to ensure that the position of the armband is similar among different sessions. The hand movements included Thumb up (TU), Extension of index and middle (EIM), Flexion of ring and littlt finger (FRL), Thumb opposing base of little fingers (TOL), Flexion of all fingers (FAF), Fingers flexed together in fist (FFTF), Pointing index (PI), Adduction of extended fingers (ADEF), Abduction of extended fingers (ABEF), tripod pinch (TP), all fingers pinch(AFP), Extension of thumb and index (ETI), Contact with thumb and index finger (CTI), Contact with thumb and middle finger (CTM), Contact with thumb and ring finger (CTR) and Contact with thumb and little finger (CTL), as illustrated in Fig. 3.

IV. Data Analysis

In this section, we would introduce how to preprocess the signal, the design of AIFNN, and how to evaluate the performance of AIFNN.

A. Signal preprocessing

The purpose of the experiment is to give a pre-study of the real-time sEMG control system, so the input delay is the key factor to be considered. Related studies have shown that the optimal guidance delay is between 150 and 250 ms, and the partially overlapping sliding window method can be used as a data enhancement technique to segment sEMG data, therefore, the sEMG data were segmented into 200 ms [35] with an increment of 50 ms [36] (75% overlap). Besides, sEMG is a noisy signal. To keep the useful motion information unchanged, we used a fourth-order Butterworth band-pass filters (20Hz-495Hz) to remove motion artifacts and high-frequency noise. Because the handcraft feature extraction of

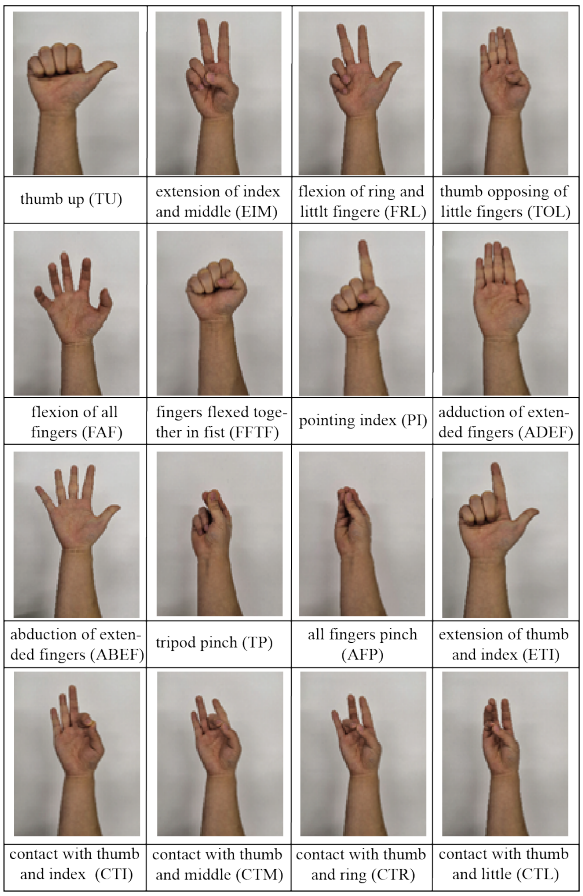


Fig. 3: The gestures used in experiment.

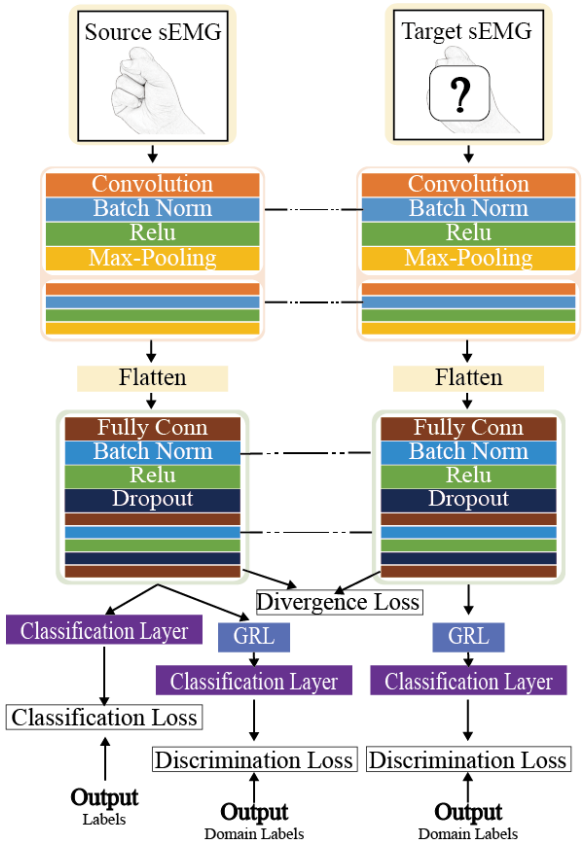


Fig. 4: Framework of AIFNN

sEMG is time-consuming, and CNN can automatically extract high-level and internal features from the raw data, this paper takes the raw sEMG as the input of the algorithm. Deep learning has been shown to yield promising results [9] [37] [38] [39] in learning features from sEMG data to perform motor intent recognition without feature engineering. A review can be found in [40], [41].

B. CNN

The AIFNN proposed in this paper consists of almost the same two streams, both based on the same CNN architecture, and the differences would be unfolded later in the domain adaptation framework. A sketch of the structure of AIFNN is shown in Fig.4, and Algorithm 1 is its details. The CNN consists of two main parts: feature learning block and classification block. The feature learning block takes two convolution layers as the core to learn the motion features of sEMG. There are 16 kernels in 1^{st} convolutional layer and 32 kernels in 2^{nd}, whilst two convolution layers uses a kernel size of $(3, 1)$. A batch normalization layer, a ReLU layer, a max-pooling layer and a dropout layer are added subsequently to each convolution layer. The classification layer divides the previously learned features into different motions. In order to realize this function, the classification layer is composed of two blocks and a full connection layer, and each block has a fully connection layer, a batch normalization layer, a ReLU layer and a dropout layer. ADAM is employed to optimize the CNN with an initial learning rate of 0.0023. Due to the inherent limitations of sEMG-based gesture recognition, CNN algorithms are susceptible to overfitting. To address the issue, Batch Normalization, Dropout and early stopping are employed.

C. Design of AIFNN

Multiple scenarios motion recognition system contains three kinds of data: $\mathcal{D}_S = \left\{\left(\boldsymbol{X}_m^S, y_m^S\right)\right\}_{m=1}^{M}$, $\mathcal{D}_T = \left\{\left(\boldsymbol{X}_n^T\right)\right\}_{n=1}^{N}$, $\mathcal{D}_S$ comes from the source domain, while $\mathcal{D}_T$ from the target domain. They have the same dimensionality of features and label spaces, i.e. $\boldsymbol{X}_m^S, \boldsymbol{X}_n^T \in \mathbb{R}^d$, d is the dimensionality, and $y \in \mathcal{Y}_m = \mathcal{Y}_n$ is label space. M and N represents the length of the data, where $N \ll M$. In the gesture recognition system, the data of $\mathcal{D}_S$ is obtained through the intentional training session. Correspondingly, the data of $\mathcal{D}_T$ represents the data obtained in the actual normal usage. As depicted in Fig. 4, a pairwise sample $\left\{\left(\boldsymbol{X}_m^S, y_m^S\right), \left(\boldsymbol{X}_n^T\right)\right\}$ is imported into domain adaptation, in which the first stream operates x_m^S and the second operates x_n^T separately. The construction of pairwise samples allows each target sample $\left(\boldsymbol{X}_n^T\right)$ to be paired with all source samples $\left(\boldsymbol{X}_m^S, y_m^S\right)$, which is able to effectively align the entire source data with the few target data. This process can also be regarded as the Cartesian product of two datasets.

Algorithm 1 Adaptive Information Fusion Neural Network

Input: Source domains with label $\mathcal{D}_S$ and target domains without label $\mathcal{D}_T$, numbers of epochs E

Output: Parameters of the depth domain adaptation network P and predicted labels for the target domain samples $\tilde{y}^T$

begin:
  Define Classifier $f_c\left(\cdot \mid \mathbf{\Theta}_c^S\right)$, Domain divergence head $f_d\left(\cdot \mid \mathbf{\Theta}_d^S, \mathbf{\Theta}_d^T\right)$, domain discrimination head $f_l\left(\cdot \mid \mathbf{\Theta}_l^S, \mathbf{\Theta}_l^T\right)$
  Initialize parameters $\boldsymbol{\theta}$, $\mathbf{\Theta}_c^S$, $\mathbf{\Theta}_f^S$, $\mathbf{\Theta}_f^T$, $\mathbf{\Theta}_d^S$, $\mathbf{\Theta}_d^T$
  Set $i = 0$
repeat:
  $i = i + 1$
  while *True* **do**
    Sample B^S from $\mathcal{D}_S$, B^T from $\mathcal{D}_T$
    $z_c = f_c\left(B^S \mid \mathbf{\Theta}_c^S\right)$
    $z_d = f_d\left(B^S, B^T \mid \mathbf{\Theta}_d^S, \mathbf{\Theta}_d^T\right)$
    $z_d = f_L\left(B^S, B^T \mid \mathbf{\Theta}_L^S, \mathbf{\Theta}_L^T\right)$
  end while
  Compute loss $(\mathcal{L}_c)$, $(\mathcal{L}_d)$, and $(\mathcal{L}_l)$:
  Update $\mathbf{\Theta}_c^S$ to minimize $(\mathcal{L}_c)$ by Eq.4
  Update $\mathbf{\Theta}_d^S$, $\mathbf{\Theta}_d^T$ to minimize $(\mathcal{L}_d)$ by Eq.5
  Update $\mathbf{\Theta}_l^S$, $\mathbf{\Theta}_l^T$ to maximize $(\mathcal{L}_l)$ by Eq.8
until L without reduction over 50 epochs **or** $i > E$
return P, $y^T = \tilde{y}^T$

The main idea of reducing the domain shift is by aligning the feature distribution of $\mathcal{D}_S$ and $\mathcal{D}_T$. Specifically, we build a Siamese architecture that can share the weights of the two-stream CNN to maximize the extraction of domain-invariant features in the source and target data. Outputs of the two-stream CNN will be utilized as deep features for three uses: the first is to calculate gesture classification results with the classification layer, the second is to minimize the distribution mismatch of two different domains, and the third is to calculate domain differences with the adaptive domain adversarial training strategy. Adaptive domain adversarial training strategy involves learning a discriminative predictor from $\mathcal{D}_S$ and $\mathcal{D}_T$, which makes it hard to distinguish between different distributions. Its most prominent feature is the use of the gradient reversal layer (GRL), which reverses the loss of all layers automatically by multiplying a negative constant $(-\lambda)$ and maintains identity transformation in the forward pass. Such operation allows for simultaneous extraction of domain-invariant features and discriminative features. The relevant mathematical expression is defined as the following equation:

$$R_\lambda(r) = r \tag{1}$$

$$\frac{dR_\lambda}{dr_i} = -\lambda I \tag{2}$$

where $R_\lambda(r)$ is the forward output of GRL, r is the input, λ is penalty and I is the identity matrix. After several verification, when λ is set to 1, the appropriate penalty could be observed. Utilizing these operation, it can effectively improve the robustness in different domains while maintaining the accuracy of the model.

Apart from the model structure and domain adaptation techniques, the loss function is also vital important to the performance. There are three loss functions added to tune CNN weights $\boldsymbol{\theta}$: classification loss $(\mathcal{L}_c)$, domain divergence loss $(\mathcal{L}_d)$, and domain discrimination loss $(\mathcal{L}_l)$, where $(\mathcal{L}_c)$ is used for supervised learning, $(\mathcal{L}_d)$ ia used to align the feature distribution of source and target streams, and $(\mathcal{L}_l)$ is to describe the difference in edge distribution between two domains. The optimal weight $\boldsymbol{\theta}^*$ could be obtained by optimizing the total loss, which could be calculated as follow:

$$L\left(\mathbf{\Theta} \mid \boldsymbol{X}^S, y^S, \boldsymbol{X}^T, d\right) = \mathcal{L}_c + \lambda_d \mathcal{L}_d + \lambda_l \mathcal{L}_l \tag{3}$$

where $\boldsymbol{\theta}$ is the CNN total weight, including five parts: $\mathbf{\Theta}_c^S$, $\mathbf{\Theta}_f^S$, $\mathbf{\Theta}_f^T$, $\mathbf{\Theta}_d^S$ and $\mathbf{\Theta}_d^T$, of which weight $\mathbf{\Theta}_c^S$ is obtained by optimizing the weight of classification loss in $\mathcal{D}_S$, $\mathbf{\Theta}_f^S$, $\mathbf{\Theta}_f^T$ are obtained by optimizing the weight of domain divergence loss between two domains in the latent space, $\mathbf{\Theta}_d^S$ and $\mathbf{\Theta}_d^T$ are obtained by optimizing the weight of domain discrimination loss in $\mathcal{D}_S$ and $\mathcal{D}_T$ respectively.

$\mathcal{L}_c$ is obtained by optimizing the most commonly used cross entropy cost function. It is the distance between the real gesture label and the prediction:

$$\mathcal{L}_c = \frac{1}{n_s} \sum_{i=1}^{n_s} c\left(\mathbf{\Theta}_c^S \mid \boldsymbol{X}_i^S, y_i^S\right) \tag{4}$$

Since only the data of $\mathcal{D}_S$ has labels, only $\mathcal{D}_S$ is used here.

$\mathcal{L}_d$ draw support from maximum mean discrepancy (MMD), which is used to measure the average distance between two domains of samples in a Reproducing Kernel Hilbert Space (RKHS).

$$\mathcal{L}_d = r_u\left(\mathbf{\Theta}_f^S, \mathbf{\Theta}_f^T \mid \boldsymbol{X}^S, \boldsymbol{X}^T\right) \tag{5}$$

Among them, the empirical estimation of square MMD is calculated by the following formula:

$$MMD^2\left(\boldsymbol{X}^S, \boldsymbol{X}^T\right) = \left\| \sum_{i=1}^{n_s} \frac{\varphi\left(\boldsymbol{x}_i^S\right)}{n_s} - \sum_{j=1}^{n_t} \frac{\varphi\left(\boldsymbol{x}_j^T\right)}{n_t} \right\|_{\mathcal{H}}^2 \tag{6}$$

Where $\boldsymbol{x}_i^S$ and $\boldsymbol{x}_j^T$ represent the output of the subsequent FC block in the source stream and target stream, respectively,and can be further expressed using kernel tricks:

$$\begin{aligned} r_u\left(\mathbf{\Theta}_f^S, \mathbf{\Theta}_f^T \mid \boldsymbol{X}^S, \boldsymbol{X}^T\right) = & \frac{\sum_{i,i*}^{n_s} k\left(\boldsymbol{x}_i^S, \boldsymbol{x}_{i^*}^S\right)}{\left(n_s\right)^2} \\ & - \frac{\sum_{i,j}^{n_s,n_t} k\left(\boldsymbol{x}_i^S, \boldsymbol{x}_j^T\right)}{n_s \times n_t} \\ & + \frac{\sum_{j,j*}^{n_t} k\left(\boldsymbol{x}_j^T, \boldsymbol{x}_{j^*}^T\right)}{\left(n_t\right)^2} \end{aligned} \tag{7}$$

For $\mathcal{L}_l$, CNNs can only achieve accurate classification of source domain data, and to achieve the classification task in the target domain, it is necessary to confuse the source domain dataset and the target domain dataset to maximize the domain classification error. The domain discriminant loss is to achieve this purpose, which is defined as:

$$\mathcal{L}_l = \left[\frac{1}{n_s} \sum_{i=1}^{n_s} c\left(\mathbf{\Theta}_d^S \mid \boldsymbol{X}_i^S, d_i\right) + \frac{1}{n_t} \sum_{j=1}^{n_t} c\left(\mathbf{\Theta}_d^T \mid \boldsymbol{X}_j^S, d_j\right) \right] \tag{8}$$

This function is used in conjunction with GRL, and since AIFNN is a two-stream design, this loss function also includes two components. The use of $\mathcal{L}_l$, as opposed to using a deterministic neuron for each domain in the training set, allows for a higher degree of certainty in the differentiation of domains, thus producing a more appropriate penalty term that guarantees approximately equal representation during testing.

D. Comparison Methods

To demonstrate the effectiveness of AIFNN, we further compare it with several baseline approaches. The descriptions of these approaches are as follows.

1) Raw CNN: For Raw CNN, we employ the same structure as described in Sec. IV-B, and only labeled data in $\mathcal{D}_S$ is are utilized to train for supervised learning. For multi-scenario learning, the data of $\mathcal{D}_T$ are utilized for testing, so the model is unaffected by $(\mathcal{L}_d)$ and $(\mathcal{L}_l)$. In this work, Raw CNN is regarded as the lower-bound recognition performances to compare with other algorithms.

2) Fine-Tuning (FT): As aforementioned, FT is the most basic but most commonly-used TL approach. Following previous research [12] and according to the structure of the CNN in Sec. IV-B, the feature learning block is transferred directly from the CNN that was pre-trained using the data in $\mathcal{D}_S$, as the initial value for training the new model in $\mathcal{D}_T$. And only the weights of the classification block are fine-tuned using the data in $\mathcal{D}_T$. Theoretically, this would give rise to errors in $\mathcal{D}_T$, causing catastrophic forgetting [10].

3) DANN: DANN is considered as a possible solution for multi-subject/cross-subject pattern recognition [29]. DANN learns the invariant features across participants by simultaneously training the network across multiple domains. However, the performance of DANN in multi-session experiment needs to be further tested. In this paper, we mainly use single-stream CNNs with two output neurons, one is the motion classification output neuron, and the other is the domain classification output neuron. The motion classification output neuron is to learn the discriminative ability to identify category of gesture, which is similar to the traditional deep learning. The domain classification output neuron is to allow extracting domain-invariant with GRL, as described in Sec. IV-C.

E. Performance Metrics

1) Classification Accuracy: Accuracy was defined as the ratio of correct predictions over total predictions:

$$Acc = \frac{correctpredictions}{totalpredictions} \times 100\% \tag{9}$$

Error represents the pooled standard variation in accuracy for different gestures. The accuracy and error information on the figure axis are expressed as a percentage.

2) Statistical Analysis: To test the effectiveness of the proposed approach, we conducted statistical analysis via the t-test and the Wilcoxon rank sum test. In each figure, No sign means no significant difference, $*$ means $p-value < 0.05$, $**$ means $p - value < 0.01$ and $***$ means $p - value < 0.001$.

3) Experimental Settings: All experiments are conducted on a microprocessor (Intel i7-11700K CPU) and a GPU (NVIDIA GeForce RTX 3060) environment with Python 3.8.1.

4) Performance Evaluation: K-fold Cross validation(CV) is a common technique used by researchers to evaluate offline performance. However, a study by Sugiyama *et. al.* showed that CV was biased in non-stationary environments (e.g., brain-computer interface) [42]. Relatively speaking, time serial related validation (TSV) can transferable offline performance to online performance [43]. The training set of TSV is performed in time sequence without iterations, meaning that trials 1 to (K-1) are the training set and the last trial is the test set, which has the potential to evaluate the effectiveness of user learning. Inspired by this approach, we adopt an incremental training protocol to analyze the performance drop caused by algorithms differences to never-seen scenario experiments, where 1 to half of the sessions/subjects are divided into $\mathcal{D}_S$, and the rest as $\mathcal{D}_T$.

For intra-session, to validate the effect of muscle fatigue on the system, the first 3 repetitions were used as the source domain, 4-6 repetitions as the target domain, and the remaining repetitions as the validation set in our incremental training protocol, as opposed to using a comparison approach. For inter-session, to verify the effect of electrode shift and the number of sessions trained on the system, $\mathcal{D}_S$ was sequentially increased from one day to five days, and after determining the size of the training set, each of the remaining days would be used sequentially as a separate $\mathcal{D}_T$. For inter-subject, to verify the effect of individual factors and the number of trained subjects on the system, $\mathcal{D}_S$ was sequentially increased from one to five individuals, and each of the remaining individuals was used as a separate $\mathcal{D}_T$.

V. Experimental Results and Performance Analysis

A. Scenario one: intra-session classification performance

Previous work has assumed that over 90% of the overall accuracy indicates good feasibility of the algorithm [4]. Fig. 5 depicts the comparison of the specific average classification performance over ten sessions between different transfer learning approaches. For the three algorithms, we have achieved high classification accuracy $(> 92.10\%)$ when testing on the run not used for training in the same session. As we can see, the differences in accuracy and error indicate that there may be significant differences of sEMG in different sessions.

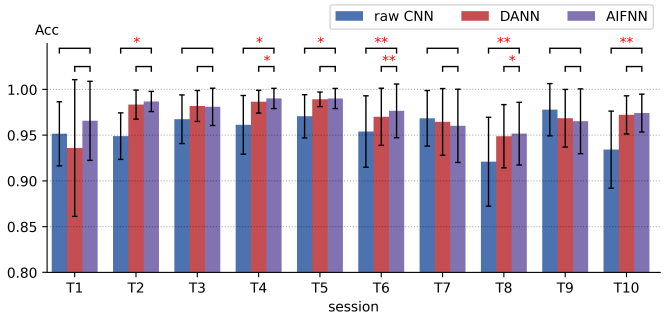
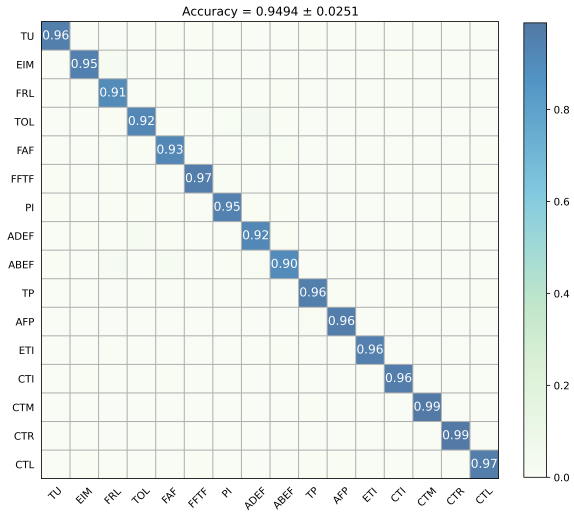


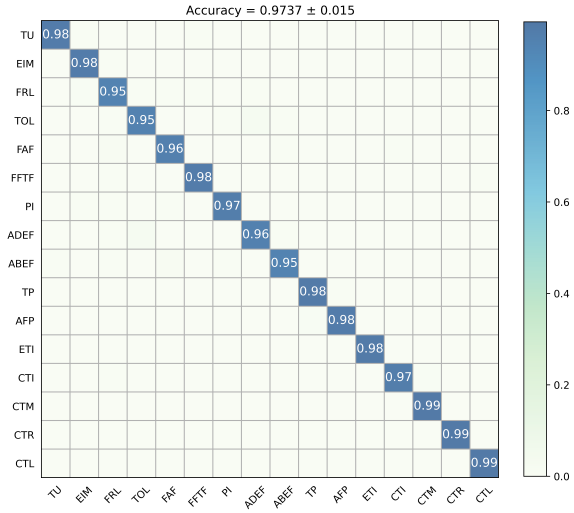
Fig. 5: Comparison of the performance among raw CNN, DANN and AIFNN for each session. T1-T10 indicate the corresponding experimental dates.

Comparing the performance of the three algorithms, it can be found that AIFNN is the best in multiple sessions, and there are significant differences between AIFNN and the other two algorithms in statistical analysis. For session T3, T7 and T9, AIFNN is not optimal, but there is no significant difference in statistical analysis.

In addition, we also analyze the confusion matrix of all sessions. Fig. 6 illustrates the average value of the confusion matrix related to sixteen gestures over 10 individual sessions under raw CNN and AIFNN schemes, as evaluated by the average accuracy and error. By analyzing the prediction performance of all single motions in the confusion matrix, the effectiveness of these algorithms in multi-gesture classification approaches is verified. From the analysis of the average confusion matrix of raw CNN in Fig. 6a, it can be found that the recognition accuracy of FRL, TOL, BG, ADEF and ABEF was less than 95% $(91.00\%, 92.1\%, 92.59\%, 92.16\%, 90.20\%)$, which is significantly worse than AIFNN. However, for the recognition performance of CTM and CTR, the two confusion matrices are similar(raw CNN: $98.57\%, 98.84\%$, AIFNN: $99.07\%, 99.39\%$), which shows that the raw CNN scheme performs poorly on some gestures, while the AIFNN scheme performs well on all gestures. From the analysis aforementioned, we can learn that an excellent adaptive scheme is helpful to build a high-precision gesture recognition system.



(a)



(b)

Fig. 6: Average confusion matrix of 10 individual sessions. (a) Confusion matrix of raw CNN scheme. (b) Confusion matrix of AIFNN scheme.

B. Scenario two: inter-session classification performance

The performance of different transfer learning approachs in the inter-session experiment is depicted in Fig. 7, as evaluated by the overall accuracy and standard error. Regardless of the approach used, the classification performance is on a downward trajectory as the number of never-seen sessions increased. Specifically, when the transfer learning approach was not used, the accuracy of the algorithm deteriorated rapidly, with accuracies dropping by 82.15%, 77.85%, 77.13%, 77.93%, 72.56%, which could not be used in practical application. The performance of raw CNN technique confirmed, compared to the intra-session test, that the variations of sEMG signals during different sessions would seriously degrade the classification performance of sEMG signals completed within ten separate days. FT have improved the classification accuracy (with accuracies drop of 49.70%, 48.74%, 47.45%, 53.09%, 48.40%), but the performance decline is also very obvious. For DANN and AIFNN, the performance degradation is reduced less than 15%, but AIFNN performs better (accuracies drop of DANN by 11.08%, 8.75%, 7.37%, 6.77%, 6.22%, accuracies drop of AIFNN by 5.34%, 7.34%, 6.42%, 6.45%, 7.69%), which is maintained at more than 90%, except for the eighth session. It is an unusual phenomenon that the performance decreased significantly during the test in the eighth session. This may be related to the environment on the day of the experiment, which needs further discussion in the future.

Furthermore, for the pre-training performance analysis from one session to five sessions, it is noteworthy that the performance of AIFNN increases by 2.1%, while the performance of raw CNN, TL and DANN show a fluctuating state. For the performance analysis of the inter-session test from one session to five sessions, another discrepancy between different transfer

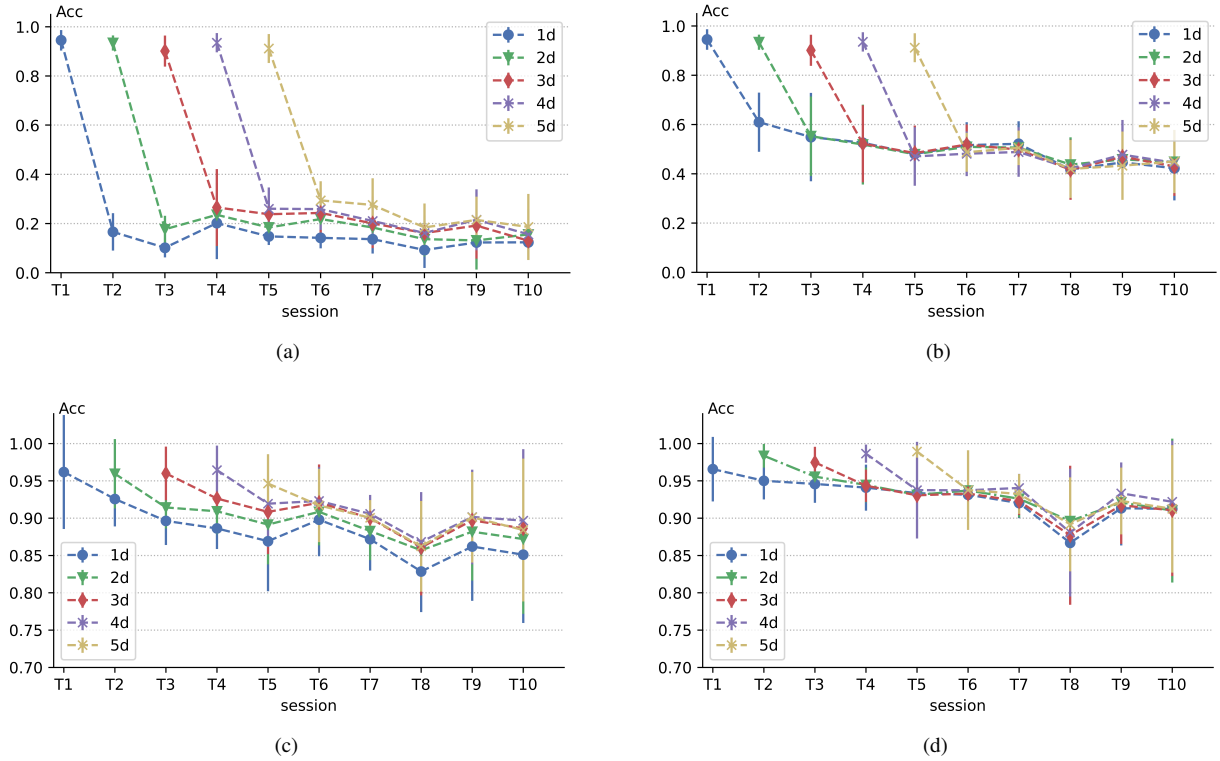


Fig. 7: The classification accuracy of different transfer learning approachs on the different sessions. The 5 multi-session training procedures gradually improve the generalization ability to never-seen sessions. 1d stands for using the data obtained in the first day as the pre-training set, 2d stands for using the data obtained in the previous two days as the pre-training set, and so on. T1-T10 indicate the corresponding experimental dates. (a) Classification accuracy curve of raw CNN. (b) Classification accuracy curve of FT. (c) Classification accuracy curve of DANN. (d) Classification accuracy curve of AIFNN.

learning techniques can be obtained that the performance of raw CNN and DANN is affected by the amount of data during pre-training, and the performance is better on the never-seen sessions with the increase of sessions used in pre training, while the performance of TL and AIFNN is hardly affected. This means that AIFNN can achieve more robust representation in heterogeneous samples, which is extremely beneficial for the application of myoelectric interface.

The statistical analysis for comparing the classification accuracy among the four approaches is shown in TableII. Note that, the performance of AIFNN is significantly different from that of raw CNN and FT. For DANN, when the data of the first three days or the first four days are pre-training data respectively, there is no significant difference between the test results on the eighth day with AIFNN, which seems to be related to the abnormality of the test results of AIFNN model on the eighth day. When the data of the first three days or the first four days are pre-training data respectively, there is no significant difference between the test results on the eighth day with AIFNN, which seems to be related to the abnormality of the test results of AIFNN model on the eighth day. Using the pre-training data of the first five days and the data of the sixth day as the test data, there is no significant difference between DANN and AIFNN, which shows that the increaseing of the amount of heterogeneous data can improve the robustness of

TABLE II: Statistical analysis for comparison of classification accuracy between four approaches. T2-T10 indicate the corresponding experimental dates.

session		T2	T3	T4	T5	T6	T7	T8	T9	T10
CNN vs. AIFNN	1d	***	***	***	***	***	***	***	***	***
	2d		***	***	***	***	***	***	***	***
	3d			***	***	***	***	***	***	***
	4d				***	***	***	***	***	***
	5d					***	***	***	***	***
FT vs. AIFNN	1d	***	***	***	***	***	***	***	***	***
	2d		***	***	***	***	***	***	***	***
	3d			***	***	***	***	***	***	***
	4d				***	***	***	***	***	***
	5d					***	***	***	***	***
DANN vs. AIFNN	1d	**	**	**	**	**	*	*	**	**
	2d		***	**	***	***	***	***	***	***
	3d			*	*	*	*	–	*	*
	4d				*	*	*	–	*	*
	5d					–	*	*	*	*

$*$ means $p-value < 0.05$, $**$ means $p-value < 0.01$ and $***$ means $p-value < 0.001$.

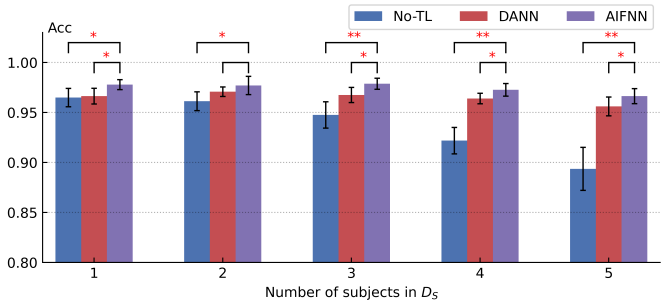


Fig. 8: Statistical analysis of raw CNN, DANN and AIFNN for each source subject

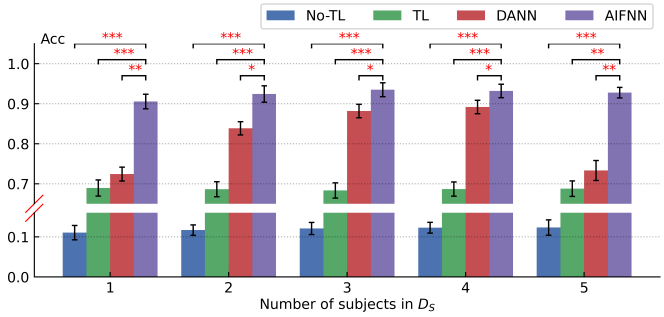


Fig. 9: Statistical analysis of raw CNN, FT, DANN and AIFNN for each target subject

DANN, but the significant difference in the performance of the next few days shows that AIFNN is still better than DANN.

C. Scenario three: inter-subject classification performance

The third scenario is more complex than the first two because it adds differences between subjects based time variability. First, we use three different pre-training approaches to train on different numbers of subjects, and the results are shown in Fig. 8. The average accuracy of AIFNN and DANN is not only high but also stable ($> 96.64\%$, $> 95.61\%$), which shows that the two approaches can classify hand movements more accurately and effectively. The accuracy of the raw CNN decreases with the increase of the number of subjects, indicating that the raw CNN cannot maximize the extraction of similar features in different subjects, and the gradual increase of error seems to explain this problem. Fig. 9 preliminarily shows the test results with different number of subjects in the source domain, and the results have some different with pre-training. Not surprisingly, the raw CNN is difficult to learn the feature of new subjects. Even FT is used, the accuracy is not very high ($< 68.97\%$). DANN changes greatly ($72.43\% - 89.18\%$), and it seems that there would be over fitting when the number of subjects is small or large. Although the accuracy of AIFNN has decreased, it is basically maintained at more than 90.57%, which shows the effectiveness of our proposed approach. This signify that the proposed AIFNN shows a less loss of performance than other approaches.

Furthermore, Fig. 10 illustrates in detail the performance differences among different transfer learning techniques in the transfer process $\mathcal{D}_S \longrightarrow \mathcal{D}_T$. Through analysis, it can be found that the proposed AIFNN is not always effective. The worst performance in the test trial was observed in S4, with the best accuracy of 92.44% and the worst of 78.98%. Meanwhile, the worst performance of S1, S2 and S7 was also less than 90% (87.56%, 85.92%, 87.55%). Based on these results, we can assume that the domain adaptation framework proposed in this study cannot improve the accuracy of subjects in all cases. At present, further experiments are needed to solve this problem.

TABLE III: Average Computational Cost (Millisecond) of Each Time Window.

Stage	CNN	FT	DANN	AIFNN
Training	138.9		206.4	257.9
Retraining		50.3	155.3	176.6
Prediction	129.1	133.4	137.5	139.2

D. Computational Cost

Traditionally, the sEMG-based gesture recognition literature has primarily focused on increasing signal information density with feature engineering for improving recognition accuracy. This approach is computationally costly for the algorithm [44], [45]. This paper employs a deep learning approach, shifting the paradigm from feature engineering to feature learning, which learn an embeddings from inputs to facilitate classification recognition. Correspondingly, deep networks tend to be computationally expensive. Therefore, an important issue in algorithm processing is time consumption. And we counted the average computational cost consumed by the algorithm at each stages and presented in Table III, including training, retraining, and prediction in each time window. Since for each stage the preprocessing is the same, we cover this process in the statistics for each stage. From the table we can see that the computational cost of training and retraining changes dramatically as the complexity of the algorithm increases. However, the computational cost of prediction does not change much, which is the most important for us, as well as meeting the latency requirements of the system.

VI. Discussion

Domain shift issues are very prevalent and difficult to solve in multi-scenario sEMG-based motion estimation, which results in a great challenge for the integration of the myoelectric interface into the advanced system [46], [47]. To be effective in commercial applications, adaptive approaches need to be easy to train and can achieve better robustness with the change of time or subjects. However, to the best of our knowledge, preliminary research on myoelectric interface has rarely focused on adaptive schemes. In this paper, we propose a novel AIFNN to improve robustness and reduce the damage of domain shift on CNN performance, while most studies focus on enhancing performance in relatively ideal environments.

In the result analysis, although the performance of DANN is very good, there are two dominant disadvantages: greatly

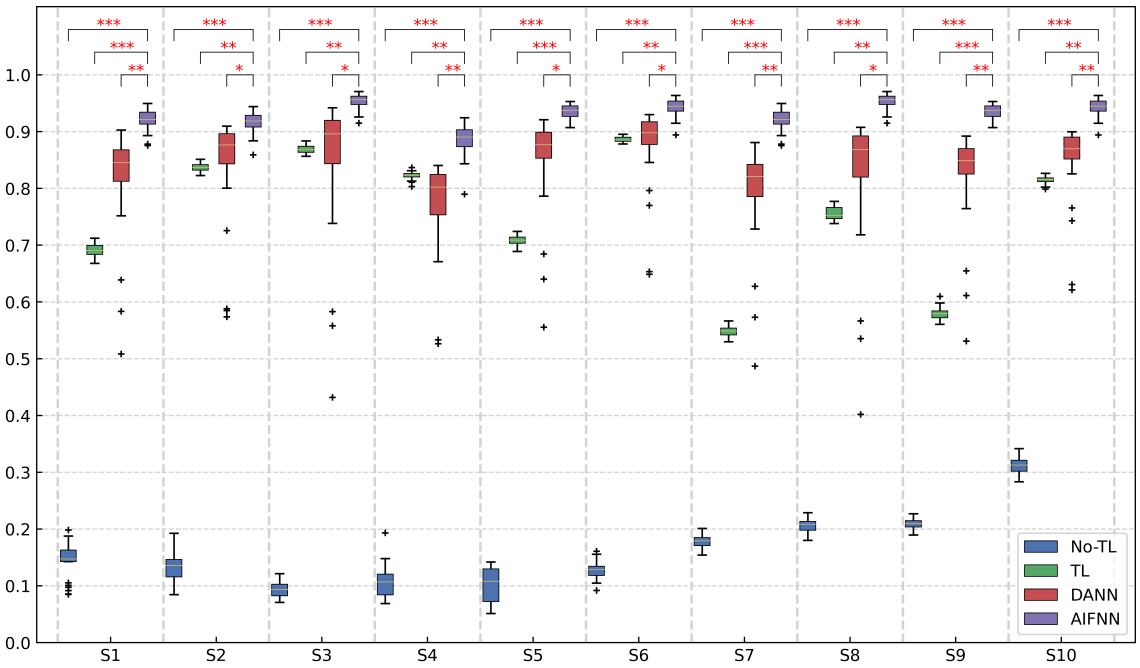


Fig. 10: Boxplots of the overall recognition caccuracies of $\mathcal{D}_S \longrightarrow \mathcal{D}_T$. S1-S10 indicate different subjects.

affected by the amount of data, and the accuracy of different gestures is different. The proposed AIFNN approaches could realize effective robust myoelectric interface without or little retraining the classifier, promoting the practical application in a real-life situation. This is mainly due to several designs of AIFNN: 1) different from raw CNN, which intrinsically learns the best features, and FT, which is vulnerable to catastrophic forgetting, AIFNN could achieve good robustness by exploiting information of both $\mathcal{D}_S$ and $\mathcal{D}_T$; 2) employing the multi-domain adversarial training could inhibit the shared network's weights from learning domain-specific representation, improving robustness of the model to different domains [48]; 3) using the domain discrepancy losses, the network can distinguish different domains with higher certainty, so as to produce a more appropriate penalty [49], [50]; 4) using the label correlation loss, the marginal distribution of $\mathcal{D}_S$ and $\mathcal{D}_T$ could be align, which ensure that the parameters of the $\mathcal{D}_T$ without any label are still updated to output non-zero probability [51]. This design enhances the relationship between gesture categories, which can not only align the features between the two domains, but also adjust $\mathcal{D}_T$ with unlabeled data according to the conditional distribution of $\mathcal{D}_S$ [52], [53], enables the network to be quickly achieved stabilised under conditions where each gesture needs to be repeated only once in a short period of time. These features make the network suitable for the target session/subject while minimising the possibility of over- or under-fitting under the limited data conditions. As the network weights are modified during the adaptation process but the validation set is not available, compared to AIFNN, the overfitting/underfitting of CNN and DANN networks cannot be mitigated. It can be seen from the experimental results that different algorithms have differences in different influencing factors. However, in multi-subject long-term experiment, the performance of AIFNN outperformed other algorithms. The study has shown that the proposed AIFNN could help to learn better representation between the two domains and adapt model to the new environment with the unlabeled data of sEMG signal. In typical domain adaptation approaches, AIFNN can greatly reduce the burden of manual labeling, which is very meaningful in many related works such as intelligent prostheses, wearable exoskeleton robots and other research fields [12], [54].

The experimental results of this paper only provide an experimental example of the advantages of AIFNN proposed in this study. Fig. 7 and Fig. 9 illustrate that the increase of heterogeneous data seems to have little impact on the performance of AIFNN, which may be due to the fact that there are only 10 subjects, so do with the abnormality in the experimental results [55]. The common features obtained from these subjects may not be enough to simulate the multi-session/multi-subject interface, which requires further research these hypotheses. At present, the computational cost during recalibration and real-time application are still a limitation of the proposed AIFNN [56]. In addition, there are still some deficiencies in this work; we addressed the small displacements that occur when subjects put on and take off the electrodes, but not the larger shift, such as 1 cm. Studies have shown that this could lead to more severe performance degradation. Another major issue is that although the existing work proposes well performing algorithms, the training process is not sufficiently uniform, a more complete pre-trained model should be built at a later stage using the existing data for adaptive real time testing. Considering the confounding factors, we will collect sEMG data containing external conditions (variation in muscle contraction force, limb position and subject mobility) [57]–[59], and reduce the training burden by using class-incremental

techniques to add more gestures without having to retrain the classifier from the scratch.

VII. Conclusion

In this study, we propose a novel AIFNN to improve robustness and reduce the damage of domain shift on motion recognition performance in the cross-scenario circumstance. The two-stream structure and adaptive domain adversarial training can effectively reduce the training burden while maintaining high accuracy. By adding the domain discrepancy loss and domain discrimination loss in model training, model can maximally extract the common features of multiple domains, and enhance the robustness and generalization of the model while maintaining the good performance in the original domain. This work can realize feasible rapid training and adaptive myoelectric interface, which have great significance to promote the practical deployment.

References

[1] J. Wu and R. Jafari, "Orientation independent activity/gesture recognition using wearable motion sensors," *IEEE Internet Things J.*, vol. 6, no. 2, pp. 1427–1437, Apr. 2019.

[2] Y. Zhang, Y. Chen, H. Yu, X. Yang and W. Lu, "Learning effective spatial–temporal features for semg armband based gesture recognition," *IEEE Internet Things J.*, vol. 7, no. 8, pp. 6979-6992, Aug. 2020.

[3] L. Qian, J. Y. Wu, S. DiMaio, N. Navab and P. Kazanzides, "A review of augmented reality in robotic-assisted surgery," *IEEE Trans. Med. Robot. Bionics*, vol. 2, no. 1, pp. 1-16, Feb. 2020.

[4] E.J. Scheme and K. Englehart, "Electromyogram pattern recognition for control of powered upper-limb prostheses: State of the art and challenges for clinical use," *J. Rehabil. Res. Develop.*, vol. 48, no. 6, pp. 643-659, 2011.

[5] D. Farina et al., "The extraction of neural information from the surface EMG for the control of upper-limb prostheses: Emerging avenues and challenges," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 22, no. 4, pp. 797-809, Feb. 2014.

[6] M. Atzori, M. Cognolato and H. Müller, "Deep learning with convolutional neural networks applied to electromyography data: A resource for the classification of movements for prosthetic hands," *Frontiers Neurorobot.*, vol. 10, pp. 1-10, Sep. 2016.

[7] W. Wei, Q. Dai, Y. Wong, Y. Hu, M. Kankanhalli and W. Geng, "Surface-electromyography-based gesture recognition by multi-view deep learning," *IEEE Trans. Biomed. Eng.*, vol. 66, no. 10, pp. 2964-2973, Oct. 2019.

[8] W. Geng, Y. Du, W. Jin, W. Wei, Y. Hu, and J. Li, "Gesture recognition by instantaneous surface emg images," *Sci. Rep.*, vol. 6, no. 1, pp. 1–8, 2016.

[9] U. Côté-Allard, C. L. Fall, A. Drouin, A. Campeau-Lecours, C. Gosselin, K. Glette, et al., "Deep learning for electromyographic hand gesture signal classification using transfer learning," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 27, no. 4, pp. 760-771, Apr. 2019.

[10] K. -T. Kim, C. Guan and S. -W. Lee, "A Subject-Transfer Framework Based on Single-Trial EMG Analysis Using Convolutional Neural Networks," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 28, no. 1, pp. 94-103, Jan. 2020.

[11] L. Hargrove, K. Englehart and B. Hudgins, "A training strategy to reduce classification degradation due to electrode displacements in pattern recognition based myoelectric control," *Biomed. Signal Process. Control*, vol. 3, no. 2, pp. 175-180, 2008.

[12] A. Ameri, M. A. Akhaee, E. Scheme and K. Englehart, "A deep transfer learning approach to reducing the effect of electrode shift in EMG pattern recognition-based control," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 28, no. 2, pp. 370-379, Feb. 2020.

[13] Q. Li, A. Zhang, Z. Li and Y. Wu, "Improvement of EMG Pattern Recognition Model Performance in Repeated Uses by Combining Feature Selection and Incremental Transfer Learning." *Front. Neurorobot.*, vol. 15, Jun. 2021.

[14] X. Sheng, B. Lv, W. Guo and X. Zhu, "Common spatial-spectral analysis of EMG signals for multiday and multiuser myoelectric interface," *Biomed. Signal Process. Control*, vol. 53, pp. 101572, 2019.

[15] M. M.-C. Vidovic, H.-J. Hwang, S. Amsüss, J. M. Hahne, D. Farina and K.-R. Müller, "Improving the robustness of myoelectric pattern recognition for upper limb prostheses by covariate shift adaptation," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 24, no. 9, pp. 961-970, Sep. 2016.

[16] T. Bao, S. Q. Xie, P. Yang, P. Zhou and Z.-Q. Zhang, "Toward robust adaptive and reliable upper-limb motion estimation using machine learning and deep learning–a survey in myoelectric control", IEEE J. Biomed. Health Inform., vol. 26, no. 8, pp. 3822-3835, Aug. 2022.

[17] F. Demir, V. Bajaj, M. C. Ince, S. Taran, and A. Şengür, "Surface EMG signals and deep transfer learning-based physical action classification," *Neural Comput. Appl.*, vol. 31, no. 12, pp. 8455–8462, Dec. 2019.

[18] A. J. Young, L. J. Hargrove, and T. A. Kuiken, "Improving myoelectric pattern recognition robustness to electrode shift by changing interelectrode distance and electrode configuration," *IEEE Trans. Biomed. Eng.*, vol. 59, no. 3, pp. 645–652, Mar. 2012.

[19] S. Kanoga, A. Kanemura and H. Asoh, "Are armband sEMG devices dense enough for long-term use?—Sensor placement shifts cause significant reduction in recognition accuracy," *Biomed. Signal Process. Control*, vol. 60, Jul. 2020.

[20] L. Guidetti, G. Rivellini, and F. Figura, "EMG patterns during running: Intra-and inter-individual variability," *J. Electromyography Kinesiol.*, vol. 6, no. 1, pp. 37–48, 1996.

[21] E. A. Biddiss and T. T. Chau, "Upper limb prosthesis use and abandonment: A survey of the last 25 years," *Prosthetics Orthotics Int.*, vol. 31, no. 3, pp. 236–257, 2007.

[22] X. Jiang _et al._, "Optimizing the Cross-Day Performance of Electromyogram Biometric Decoder", IEEE Internet Things J., vol. 10, no. 5, pp. 4388–4402, Mar. 2023.

[23] Y. Wen, S. J. Kim, S. Avrillon, J. T. Levine, F. Hug, and J. L. Pons, 'Toward a generalizable deep CNN for neural drive estimation across muscles and participants', J. Neural Eng., vol. 20, no. 1, p. 016006, Feb. 2023.

[24] X. Zhai, B. Jelfs, R. H. M. Chan and C. Tin, "Self-recalibrating surface EMG pattern recognition for neuroprosthesis control based on convolutional neural network," *Frontiers Neurosci.*, vol. 11, pp. 379, Jul. 2017.

[25] M. Atzori, A. Gijsberts, C. Castellini, B. Caputo, A.-G.-M. Hager, S. Elsig, et al., "Electromyography data for non-invasive naturally-controlled robotic hand prostheses," *Sci. Data*, vol. 1, no. 1, pp. 140053, Dec. 2014.

[26] Y. Du, W. G. Jin, W. T. Wei, Y. Hu and W. D. Geng, "Surface EMG-based inter-session gesture recognition enhanced by deep domain adaptation," *Sensors*, vol. 17, no. 3, pp. 458, Mar. 2017.

[27] Y. H. Li, N. Y. Wang, J. P. Shi, X. D. Hou and J. Y. Liu, "Adaptive batch normalization for practical domain adaptation," *Pattern Recogn.*, vol. 80, pp. 109-117, Aug. 2008.

[28] U. Côté-Allard, G. Gagnon-Turcotte, A. Phinyomark, K. Glette, E. J. Scheme, F. Laviolette, et al., "Unsupervised domain adversarial self-calibration for electromyography-based gesture recognition," *IEEE Acc.*, vol. 8, pp. 177941-177955, Sep. 2020.

[29] U. Côté-Allard, E. Campbell, A. Phinyomark, F. Laviolette, B. Gosselin and E. Scheme, "Interpreting deep learning features for myoelectric control: A comparison with handcrafted features," *Front. Bioeng. Biotechnol.*, vol. 8, pp. 158, Mar. 2020.

[30] E. Campbell, A. Phinyomark and E. Scheme, "Deep Cross-User Models Reduce the Training Burden in Myoelectric Control." *Front. Neurosci.*, vol. 15, pp.657958, May. 2021.

[31] T. Bao, S. A. R. Zaidi, S. Xie, P. Yang and Z. -Q. Zhang, "Inter-Subject Domain Adaptation for CNN-Based Wrist Kinematics Estimation Using sEMG," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 29, pp. 1068-1078, 2021.

[32] I. Sosin, D. Kudenko and A. Shpilman, "Continuous gesture recognition from sEMG sensor data with recurrent neural networks and adversarial domain adaptation," *Proc. 15th Int. Conf. Control Autom. Robotics and Vision*, pp. 1436-1441, 2018.

[33] U. Côté-Allard, C. L. Fall, A. Campeau-Lecours, C. Gosselin, F. Laviolette and B. Gosselin, "Transfer learning for sEMG hand gestures recognition using convolutional neural networks," *Proc. IEEE Int. Conf. Systems Man and Cybernetics*, pp. 1663-1668, 2017.

[34] X. Sheng, B. Lv, W. Guo and X. Zhu, "Common spatial–spectral analysis of EMG signals for multiday and multiuser myoelectric interface," *Biomed. Signal Process. Control*, vol. 53, pp. 101572, Aug. 2019.

[35] L. H. Smith, L. J. Hargrove, B. A. Lock and T. A. Kuiken, "Determining the optimal window length for pattern recognition-based myoelectric control: Balancing the competing effects of classification error and

controller delay," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 19, no. 2, pp. 186-192, Apr. 2011.

[36] A. M. Simon, L. J. Hargrove, B. A. Lock and T. A. Kuiken, "Target achievement control test: Evaluating real-time myoelectric pattern-recognition control of multifunctional upper-limb prostheses," *J. Rehabil. Res. Develop.*, vol. 48, no. 6, pp. 619, 2011.

[37] M. Atzori, M. Cognolato and H. Müller, "Deep learning with convolutional neural networks applied to electromyography data: A resource for the classification of movements for prosthetic hands," *Front. Neurorobot.*, vol. 10, pp. 9, Sep. 2016.

[38] A. Ameri, M. A. Akhaee, E. Scheme and K. Englehart, "Real-time simultaneous myoelectric control using a convolutional neural network," *PLoS ONE*, vol. 13, Sep. 2018.

[39] A. Ameri, M. A. Akhaee, E. Scheme and K. Englehart, "Regression convolutional neural network for improved simultaneous EMG control," *J. Neural Eng.*, vol. 16, no. 3, 2019.

[40] A. Phinyomark and E. Scheme, "EMG pattern recognition in the era of big data and deep learning," *Big Data Cognit. Comput.*, vol. 2, pp. 21, Sep. 2018.

[41] W. Li, P. Shi, and H. Yu, "Gesture recognition using surface electromyography and deep learning for prostheses hand: State-of-the-art, challenges, and future," *Front. Neurosci*, vol. 15, 2021.

[42] M. Sugiyama, B. Blankertz, M. Krauledat, G. Dornhege and K. R. Müller, "Importance-weighted cross-validation for covariate shift," *Proc. Joint Pattern Recognit. Symp.*, pp. 354-363, Sep. 2006.

[43] B. Lv, G. Chai, X. Sheng, H. Ding and X. Zhu, "Evaluating User and Machine Learning in Short- and Long-Term Pattern Recognition-Based Myoelectric Control," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 29, pp. 777-785, 2021.

[44] R. N. Khushaba, A. H. Al-Timemy, A. Al-Ani, and A. Al-Jumaily, "A Framework of Temporal-Spatial Descriptors-Based Feature Extraction for Improved Myoelectric Pattern Recognition," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 25, no. 10, pp. 1821–1831, Oct. 2017, doi: 10.1109/TNSRE.2017.2687520.

[45] D. Xiong, D. Zhang, X. Zhao, Y. Chu, and Y. Zhao, "Learning Non-Euclidean Representations With SPD Manifold for Myoelectric Pattern Recognition," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 30, pp. 1514–1524, 2022, doi: 10.1109/TNSRE.2022.3178384.

[46] O. W. Samuel et al., "Intelligent EMG Pattern Recognition Control Method for Upper-Limb Multifunctional Prostheses: Advances, Current Challenges, and Future Prospects," *IEEE Access*, vol. 7, pp. 10150-10165, 2019.

[47] H. Xu and A. Xiong, "Advances and disturbances in sEMG-based intentions and movements recognition: A review," *IEEE Sensors J.*, vol. 21, no. 12, pp. 13019-13028, Jun. 2021.

[48] Y. Ganin et al., "Domain-adversarial training of neural networks," *J. Mach. Learn. Res.*, vol. 17, no. 1, pp. 2030–2096, May 2015.

[49] M. Long, H. Zhu, J. Wang, and M. I. Jordan, "Unsupervised domain adaptation with residual transfer networks," in *Adv. Neural Inf. Process. Syst.*, pp. 136–144, 2016.

[50] M. Long, H. Zhu, J. Wang, and M. Jordan, "Deep transfer learning with joint adaptation networks," in *International conference on machine learning*. PMLR, pp. 2208–2217, 2017.

[51] E. Tzeng, J. Hoffman, T. Darrell and K. Saenko, "Simultaneous deep transfer across domains and tasks," *Proc. IEEE Int. Conf. Comput. Vis.*, pp. 4068-4076, Dec. 2015.

[52] R. Shu, H. H. Bui, H. Narui, and S. Ermon, "A DIRT-T approach to unsupervised domain adaptation," 2018, *arXiv:1802.08735*. [Online]. Available: http://arxiv.org/abs/1802.08735

[53] Rafael Müller, Simon Kornblith and Geoffrey E Hinton, "When does label smoothing help?," *Advances in Neural Information Processing Systems*, vol. 32, 2019.

[54] B. Xue, L. Wu, K. Wang, X. Zhang, J. Cheng, X. Chen, and X. Chen, "Multiuser gesture recognition using semg signals via canonical correlation analysis and optimal transport," *Comput. Biol. Med.*, vol. 130, p. 104188, 2021.

[55] J. Liu, "Adaptive myoelectric pattern recognition toward improved multifunctional prosthesis control," *Med. Eng. Phys.*, vol. 37, no. 4, pp. 424-430, Apr. 2015.

[56] M. Scherer, M. Magno, J. Erb, P. Mayer, M. Eggimann and L. Benini, "TinyRadarNN: Combining spatial and temporal convolutional neural networks for embedded gesture recognition with short range radars", *IEEE Internet Things J.*, vol. 8, no. 13, pp. 10336-10346, Jul. 2021.

[57] M. Coscia et al., "The effect of arm weight support on upper limb muscle synergies during reaching movements," *J. NeuroEng. Rehabil.*, vol. 11, no. 1, pp. 22, 2014.

[58] D. Yang, Y. Gu, L. Jiang, L. Osborn and H. Liu, "Dynamic training protocol improves the robustness of PR-based myoelectric control," *Biomed. Signal Process. Control*, vol. 31, pp. 249-256, Jan. 2017.

[59] D. Farina et al., "The extraction of neural information from the surface EMG for the control of upper-limb prostheses: Emerging avenues and challenges," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 22, no. 4, pp. 797-809, Jul. 2014.