EMGSense

A Low-Effort Self-Supervised Domain Adaptation Framework for EMG Sensing

Duan, Di ; Yang, Huanqi ; Lan, Guohao; Li, Tianxing ; Jia, Xiaohua; Xu, Weitao

DOI
10.1109/PERCOM56429.2023.10099164

Publication date
2023

Document Version
Final published version

Published in
Proceedings of the 2023 IEEE International Conference on Pervasive Computing and Communications (PerCom)

EMGSense: A Low-Effort Self-Supervised Domain Adaptation Framework for EMG Sensing

Di Duan[1,2], Huanqi Yang[1,2], Guohao Lan[3], Tianxing Li[4], Xiaohua Jia[1,2], Weitao Xu[1,2,*]
[1]City University of Hong Kong Shenzhen Research Institute
[2]Department of Computer Science, City University of Hong Kong
[3]Department of Software Technology, Delft University of Technology
[4]Department of Computer Science and Engineering, Michigan State University

2023 IEEE International Conference on Pervasive Computing and Communications (PerCom) | 978-1-6654-5378-3/23/$31.00 ©2023 IEEE | DOI: 10.1109/PERCOM56429.2023.10099164

***Abstract*—This paper presents *EMGSense*, a low-effort self-supervised domain adaptation framework for sensing applications based on Electromyography (EMG). *EMGSense* addresses one of the fundamental challenges in EMG cross-user sensing—the significant performance degradation caused by time-varying biological heterogeneity—in a low-effort (data-efficient and label-free) manner. To alleviate the burden of data collection and avoid labor-intensive data annotation, we propose two EMG-specific data augmentation methods to simulate the EMG signals generated in various conditions and scope the exploration in label-free scenarios. We model combating biological heterogeneity-caused performance degradation as a multi-source domain adaptation problem that can learn from the diversity among source users to eliminate EMG heterogeneous biological features. To relearn the target-user-specific biological features from the unlabeled data, we integrate advanced self-supervised techniques into a carefully designed deep neural network (DNN) structure. The DNN structure can seamlessly perform two training stages that complement each other to adapt to a new user with satisfactory performance. Comprehensive evaluations on two sizable datasets collected from 13 participants indicate that *EMGSense* achieves an average accuracy of 91.9% and 81.2% in gesture recognition and activity recognition, respectively. *EMGSense* outperforms the state-of-the-art EMG-oriented domain adaptation approaches by 12.5%–17.4% and achieves a comparable performance with the one trained in a supervised learning manner.**

***Index Terms*—EMG sensing, biological heterogeneity, domain adaptation, self-supervised learning.**

I. Introduction

Surface Electromyography (EMG) has been widely used for the measurement of the electrical activity of a muscle through positioned surface electrodes on the skin. Due to its portable and wearable nature, EMG has released a smorgasbord of intelligent applications, such as neurorehabilitation [1]–[3], activity recognition [4], [5], gesture recognition [6]–[8], hand pose reconstruction [9], [10], and virtual reality (VR) [11].

Despite these advances, an overarching challenge in existing EMG systems is to tackle cross-user scenarios. Existing studies have demonstrated that EMG signals can be influenced by various biological factors, such as body fat [12], skin conditions [13] (e.g., moisture [14]), age [15], and fatigue [16]. The heterogeneity of these biological factors across users can significantly degrade the sensing performance when a new user uses the systems. Furthermore, the degradation is exaggerated when the biological factors change over time [13]. For example, recent diet patterns can affect users' body fat percentage. In addition, modeling an exact mathematical model to extract the heterogeneity across users is unfeasible because muscle is a complex fuzzy system [17], and the relationship between muscle contraction and EMG signal is ambiguous.

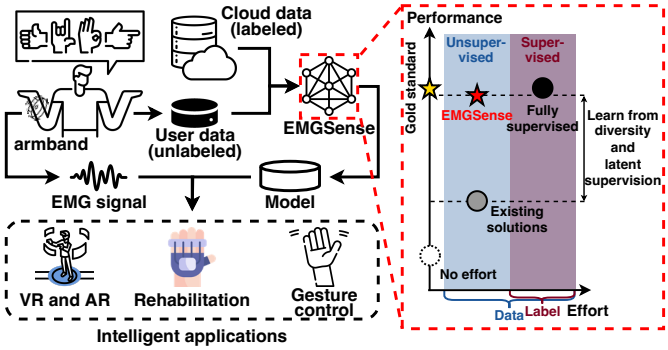


Fig. 1: *EMGSense* enables various applications to adapt to a new user and achieves comparable performance with the solution based on supervised learning in a low-effort manner.

We currently have limited options to deal with such cross-user scenarios. Existing methods either train a model for each user [6], [18], [19] by collecting abundant data and annotating it or leverage transfer learning methods [20]–[23] (e.g., fine-tuning and domain adaptation). However, these methods have several limitations. On the one hand, although the developer can train a model for a new user from scratch or fine-tune the existing model with sizeable labeled data, it introduces high deployment overhead and human efforts. Moreover, training a model in a supervised manner cannot adapt to the time-varying heterogeneity without more labeled data, which is not a silver bullet to combat the EMG heterogeneity challenges. On the other hand, domain adaptation (DA) has emerged as a promising transfer learning technique to solve the performance degradation problem caused by domain shift. It aims to learn a model from source domains (e.g., existing users) that can achieve high performance on a different (but related) target domain (e.g., a new user). Recent works have attempted to combat the performance degradation using conventional DA approaches (e.g., single-source DA) [20]–[22]. However, their methods ignore that the heterogeneity between the source and target domains also exists among different source domains,

* Weitao Xu is the corresponding author.

which limits the conventional DA approaches' performance in EMG sensing. To tackle this issue, Guo et al. [23] attempted to leverage the diversity by modeling it as an unsupervised multi-source domain adaptation problem, but their solution performs inferior in EMG signal modality (about 51% accuracy in gait recognition) due to the lack of supervision. The ensemble of these issues of prior work forms a significant research gap in EMG sensing—*how to train an intelligent model with satisfactory performance in a low-effort manner to solve the biological heterogeneity problem in EMG cross-user scenario*.

To fill the research gap, we propose *EMGSense*—a low-effort self-supervised domain adaptation framework that can achieve low-effort, high-accuracy EMG sensing for new users. As Figure 1 shows, *EMGSense* leverages small-scale unlabeled data from a new user and the pre-collected data from several existing users to train a discriminative model to realize intelligent applications for the new user. Note that the pre-collected data is one-effort labor and can be stored in the cloud to serve all new users. From the principle aspect, *EMGSense* lies at the intersection of domain adaptation and self-supervised learning, which aims at achieving satisfactory performance in a low-effort manner by learning from user diversity (i.e., heterogeneity) and the latent supervision of unlabeled data.

Instantiating our idea in practice requires that we address three main challenges. First, the performance of a well-trained model will degrade as the biological factors change over time. It is difficult to cope with the time-varying heterogeneity in a low-effort manner. Second, extracting sufficient homogeneous biological features to adapt to a new user with acceptable performance is challenging because the biological heterogeneity in every two users is disparate. Third, to the biological signal such as EMG, user-specific biological features are indispensable components to achieving high performance in single-user scenarios. However, the difference in user-specific biological features of different users (i.e., biological heterogeneity) is a lion in the way to cross-user sensing. It is hard to balance the generality (for adapting to a new user) and specificity (for achieving high performance) of the extracted features in cross-user scenarios. Therefore, extracting suitable features in different training stages is a vital task.

EMGSense leverages three core techniques to address these challenges. First, we propose two EMG-specific data augmentation methods to alleviate the burden of data collection. Additionally, we cope with the time-varying heterogeneity using unsupervised learning, which also avoids tedious data annotation. Second, we propose a multi-source domain adaptation approach that can eliminate heterogeneous biological features to solve the EMG domain shift problem by leveraging the diversity among the source domains. Third, we apply self-training techniques to relearn target-user-specific biological features to achieve satisfactory performance on a new user. The contributions of this paper are summarized as follows:

- *EMGSense* is the first low-effort AI-empowered framework that leverages the EMG heterogeneity to cope with the performance degradation caused by inter-user biological heterogeneity in EMG sensing.
- *EMGSense* can adapt to new users in a low-effort manner and resist the performance degradation caused by time-varying biological heterogeneity by combing two EMG-specific data augmentation and label-free scenarios.
- *EMGSense* can seamlessly perform two training stages without modifying any DNN structures. The two training stages can eliminate the heterogeneity features to adapt to a new user and relearn the target-user-specific biological features to achieve a performance that is comparable with the model trained in a supervised learning manner.
- We validate *EMGSense* on two sizable EMG datasets (9,898 gestures and 171 min activities) collected from 13 participants. The result shows that *EMGSense* achieves an average accuracy of 86.6% in new users and outperforms the state-of-the-art approaches by 12.5%–17.4%.

The rest of the paper is organized as follows. We first present an overview of *EMGSense* in Section II. Then, we specify the signal processing and augmentation in Section III, and elaborate the design of *EMGSense* in Section IV. We evaluate the performance of *EMGSense* in Section V. Finally, we discuss related work in Section VI before concluding the paper in Section VII.

II. Framework Overview

This paper considers the scenarios where a model developer has pre-collected a labeled dataset from several existing users (i.e., source domains). The goal of *EMGSense* is to learn a high-performance model for a new user (i.e., target domain) in a low-effort (label-free and data-efficient) manner. Figure 2 illustrates the training process of *EMGSense*, which consists of two stages: a pre-training stage and a self-training stage. The pre-training stage aims to adapt to a new user by eliminating the user-specific biological features, and the self-training stage aims to achieve high performance on the new user by relearning the user-specific biological features.

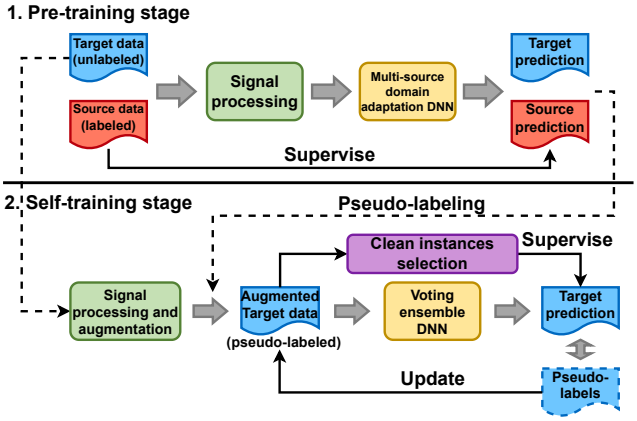


Fig. 2: Framework overview.

Pre-training stage. The pre-training stage aims to initialize a model that can take advantage of the diversity among existing users. To this end, we formulate this stage as a multi-source domain adaptation problem. The pre-training process is shown in the upper figure of Figure 2, during which the labeled

source dataset and the unlabeled target dataset are processed and fed into a multi-source domain adaptation DNN which aims to learn a moderate-performance model for a new user.

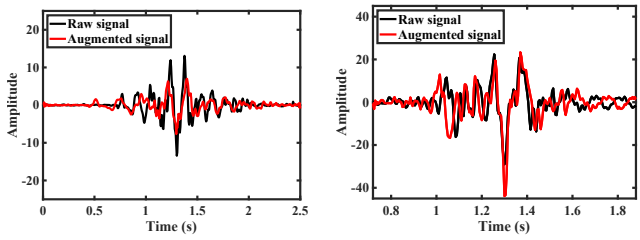
Self-training stage. The self-training stage is proposed to boost the model trained from the pre-training stage to a high-performance model for the new user. The self-training process is shown in the lower figure of Figure 2. We first conduct data augmentation to mitigate the shortage of the target samples and increase the robustness of the trained model. Then, we use the model trained in the previous stage to infer the augmented target samples, and the predictions are used as the initial pseudo-labels. This process is represented as "Pseudo-labeling" in Figure 2. Subsequently, we select the clean instances and corresponding pseudo-labels using a small-loss strategy. The selected instances and pseudo-labels are used to supervise the self-training process, which is performed by the multi-model voting ensemble DNN (the initial parameters are inherited from the pre-training stage, detailed in later sections). Then, we use the improved model to re-infer the augmented target data to reduce the noise (errors) in pseudo-labels. By alternately updating the model parameters and pseudo-labels, the pre-trained model will be boosted to be a high-performance model for the new user.

In a nutshell, *EMGSense* achieves low-effort, high-accuracy EMG sensing in cross-user scenarios. The design details of the framework are elaborated below. Without loss of generality, we use a Myo armband [24] with eight EMG channels as the hardware to illustrate the design details.

III. Signal Processing & Data Augmentation

A. Signal Processing

We first apply a high-pass Butterworth filter with a cut-off frequency 30 Hz on the raw EMG signal to filter out the low-frequency noise. Then, we leverage the threshold-based method [25] to segment the active periods when the user performs gestures/activities. Specifically, we first sum the signals from the eight EMG channels together. Then, we normalize the magnitude of summed EMG signals and apply a 0.1 s sliding window on the summed signals to calculate the standard deviation of each window. After this step, we filter out the segments with a standard deviation lower than a predefined threshold of 0.02. Finally, we can effectively detect and segment the periods when the user performs gestures/activities.



(a) Speed-intensity augmentation (b) Local variation augmentation

Fig. 3: Examples of the proposed augmentation methods: (a) a simulated low-speed signal with lower intensity; (b) a simulated signal's active period with local variations.

B. Data augmentation for EMG

Collecting a large amount of data from different users is labor-intensive. Therefore, we design two data augmentation methods to increase the diversity of training data and improve the model generalization ability and avoid overfitting. We proposed a data augmentation module that mimics the perturbations in EMG signals, which are elaborated below.

Speed-intensity augmentation. As stated in [26], the intensity of the muscle contraction affects the amplitude of the EMG signal, and the speed of the motion affects the length of the EMG signal. The first augmentation method is defined as $Y = \alpha f(X;\beta)$ to simulate EMG signals with different intensities and speeds, where α is a coefficient that changes the amplitude of the raw signal, which simulates the effect of performing the same gesture/activity with different intensities. $f(X;\beta)$ is a function that re-samples the original EMG signal X with a re-sample rate β, which is used to simulate the process that a user may perform the same gesture/activity with varying speeds. Here, α and β are drawn randomly from $[0.5, 1.5]$ and $[0.8, 1.2]$, respectively. Figure 3(a) shows an example of augmented signals. We can see that the proposed speed-intensity augmentation can synthesize signals with different speeds and intensities, which allows *EMGSense* to obtain better generalization capability.

Local variation augmentation. The above method can simulate the effect of different speeds and intensities on the whole EMG signal. However, since a gesture can be decomposed into several sub-gestures [27], it is difficult to ensure the speeds and intensities are always the same for each sub-gesture because of the diversity in user's behaviors. Therefore, we propose another augmentation method that can increase input robustness by simulating randomness (the local changes of speed and intensity) in usage scenarios.

To simulate the local changes of intensity, we first randomly select $\frac{L_X}{3}$ indices from $[1, L_X]$ which are denoted by $I = [\mathrm{i}_1, \mathrm{i}_2, \ldots, \mathrm{i}_{\frac{L_X}{3}}]$. Then, we randomly select $\frac{L_X}{3}$ points from a normal distribution $\mathcal{N}(1, 0.1)$, which are denoted by $S = [s_1, s_2, \ldots, s_{\frac{L_X}{3}}]$. Then, we use the previously generated random points S and indices I to fit a warping curve g, and calculate the corresponding values $G = [g(1), g(2), \ldots, g(L_X)]$. Finally, the EMG signal values are multiplied with the corresponding G values in point-wise manner, namely,

$$X \to X_a = [x_1 \cdot g(1), x_2 \cdot g(2), \ldots, x_{L_X} \cdot g(L_X)]. \quad (1)$$

Since S follows normal distribution, it can ensure that only the amplitude of a small portion of points is changed greatly

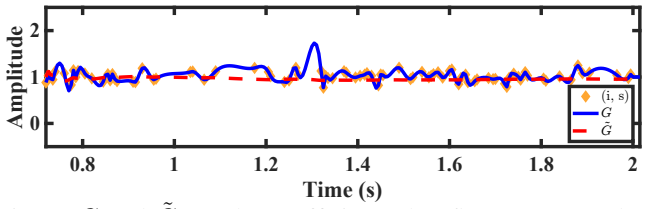


Fig. 4: G and $\tilde{G}$ are the coefficients that fluctuate around 1.

while the amplitude of most points is changed slightly.

To simulate the local changes of speed, we first calculate the cumulative sum $\tilde{G}$ of G as follows:

$$\tilde{G} = \left[\sum_{n=1}^{1} g(n), \sum_{n=1}^{2} g(n), \ldots, \sum_{n=1}^{L_X} g(n) \right]. \quad (2)$$

As Figure 4 shows, both G and $\tilde{G}$ are time series fluctuated around 1. Then, we tune the raw indices with $\tilde{G}$ in an element-wise manner and reorder (represented as Function: $sort$) them to obtain the new indices:

$$I' = sort\left(\left[1 \cdot \sum_{n=1}^{1} g(n), 2 \cdot \sum_{n=1}^{2} g(n), \ldots, L_X \cdot \sum_{n=1}^{L_X} g(n) \right] \right). \quad (3)$$

We use shifted indices I' and magnitude-warped values X_a to fit a new time-warping curve function h. Then, we use the function h to calculate the augmented signal. Figure 3(b) provides an example of the augmented signal. As the evaluation in Section V-D, the proposed two data augmentation methods can significantly improve EMG sensing accuracy.

IV. Framework Design

In this section, we describe the details of the *EMGSense* framework design. The core of the *EMGSense* framework is a well-designed DNN, which consists of a common feature extractor, N domain-specific feature extractors, and N domain-specific classifiers, where N denotes the number of source domains. Each domain-specific feature extractor and the corresponding classifier are combined with the common feature extractor to form a branch.

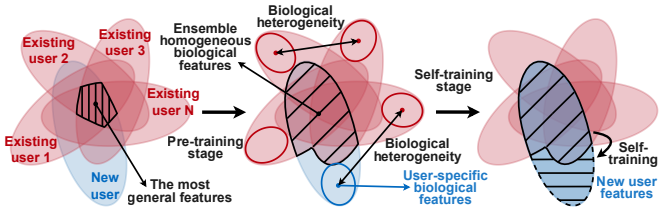


Fig. 5: Venn diagram illustrates the principle of *EMGSense*.

The principle behind this design is the shared common feature extractor aims to ensure the transferability of shallow features. Meanwhile, combinations of domain-specific feature extractors and classifiers are responsible for independently exploring the diversity among the deep features from different source domains. As Figure 5 shows, by eliminating user-specific biological features (e.g., body fat percentage, skin conditions) in extracted deep features, each branch can learn some homogeneous biological features (e.g., myoarchitecture, contracted muscle group) to adapt to the target domain (i.e., a new user) from the corresponding source domain. By ensembling the homogeneous biological features, the model obtains a moderate capability to predict the new user's data, and we can leverage the predictions as pseudo-labels. After purifying the pseudo-labels, the cleaner supervision will force the model to relearn the target-user-specific biological features and improve

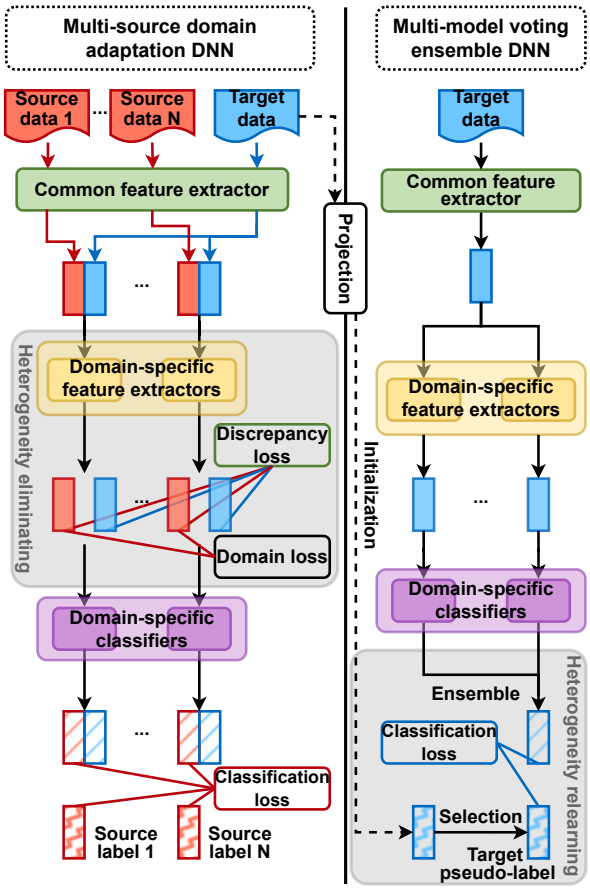


Fig. 6: *EMGSense* training workflow.

the model performance on the new user. Specifically, the *EMGSense* seamlessly conducts two training stages elaborated below—a pre-training stage and a self-training stage, without modifying any DNN structure.

A. Pre-training Stage

1) Design Rationale: Recall in Section II that the goal of the pre-training stage is to train a model that can adapt to a new user with moderate performance as the initialization of the self-training stage. To this end, we treat the samples from N existing users as N independent source domains and combine each existing user and the target user as a subset. As the left part of Figure 6 shows, *EMGSense* minimizes a classification loss and a pre-defined discrepancy loss to learn the homogeneous biological features from each subset. Meanwhile, *EMGSense* maximizes the domain loss calculated by a domain discriminator [28] to confuse the features from different existing users and force the DNN to learn more general features by eliminating heterogeneous biological features.

2) Problem Formulation: The objective function of the pre-training stage consists of three kinds of losses: classification loss $\mathcal{L}_\tau^C$ from the source domains, discrepancy loss $\mathcal{L}_\tau^{dis}$ between the source domains and target domain, domain loss $\mathcal{L}^D$ among source domains. The pre-training stage aims to

solve the following optimization problem:

$$\gamma = \frac{2}{1+\exp(\frac{-100*I}{I_{MAX}})} - 1, \tag{4}$$

$$\min_{\theta_F,\theta_{H_\tau},\theta_{C_\tau}} \max_{\theta_D} \sum_{\tau=1}^{N} \mathcal{L}_\tau^C(\theta_F,\theta_{H_\tau},\theta_{C_\tau}) + \gamma\big(\mathcal{L}_\tau^{dis}(\theta_F,\theta_{H_\tau}) - \mathcal{L}^D(\theta_F,\theta_{H_\tau},\theta_D)\big), \tag{5}$$

where I is the number of iterations, I_{Max} is the max number of iterations. θ_F, θ_{H_τ}, θ_{C_τ}, θ_D denote the parameters of the common feature extractor, domain-specific feature extractors, domain-specific classifiers, and domain discriminator, respectively. $\tau \in [1,\ldots,N]$ is the corner marker that represents the serial number of a source domain, and γ denotes the trade-off parameter for $\mathcal{L}_\tau^{dis}$ and $\mathcal{L}^D$.

3) Design Details: After using the processing method described in Section III, the processed EMG signal X will be used as input directly. First, we employ a common feature extractor (two-layer stacked 1D-CNNs) to extract channel-wise spatial features. Both CNN layers use 1×10 1D filters with a stride length of 1×2. Then, we add a batch normalization layer after each CNN layer to normalize the inputs and accelerate convergence by reducing internal covariate shift [29]. The extracted features are sliced into eight segments and fed into a domain-specific feature extractor (two-layer stacked LSTMs) to extract temporal features. The hidden sizes of the two LSTM layers are 512 and 128, respectively. Overall, the extracted domain-specific features f_τ through the feature extraction function G_{f_τ} can be represented as follows:

$$f_\tau = G_{f_\tau}(X) = \mathrm{LSTM}\big(\mathrm{CNN}(X)\big). \tag{6}$$

The extracted features f_τ are fed into a domain-specific classifier (three fully connected layer stacked) to learn a prediction function $G_{y_\tau}: f_\tau \to [0,1]^K$ that is parameterized by θ_{C_τ}:

$$G_{y_\tau}(f_\tau;\theta_{C_\tau}) = \mathrm{softmax}\left(G_{f_\tau}(X);\theta_{C_\tau}\right), \tag{7}$$

where K is the number of classes. We use ReLU activation function and define the cross-entropy loss function as follows:

$$\mathcal{L}^C = -\sum_{\tau=1}^{N}\sum_{j=1}^{L}\sum_{i=1}^{K} l_{ij\tau}\log\left(p_{ij\tau}\right), \tag{8}$$

where $l_{ij\tau}$ is the truth label, $p_{ij\tau}$ is the softmax probability of the i-th class in j-th sample from the τ-th source user, and L is the batch size. The loss $\mathcal{L}_\tau^C$ will be optimized to achieve high accuracy in single-user (an existing user) scenarios rather than cross-user scenarios. To adapt to the new user, we apply two transfer learning techniques elaborated as follows:

First, we employ Multi-kernel Maximum Mean Discrepancy (MK-MMD) [30] to align the source and target distributions in deep feature space, which aims to learn groups of biological homogeneity features from different subsets. Specifically, we compress the source probability distribution p_τ^S and target probability distribution p_τ^T into a reproducing kernel Hilbert space $\mathcal{H}$ by using Gaussian kernels. Then, we define the discrepancy loss $\mathcal{L}_\tau^{dis}$ as:

$$\mathcal{L}^{dis} = \sum_{\tau=1}^{N} \mathrm{MMD}(p_\tau^S, p_\tau^T, \mathcal{H}) = \sum_{\tau=1}^{N} \|\mu_\tau^S - \mu_\tau^T\|_{\mathcal{H}}, \tag{9}$$

where μ_τ^S and μ_τ^T are the unbiased estimation of expectation $\mathbf{E}_{p_\tau^S}[f_\tau^S]$ and $\mathbf{E}_{p_\tau^T}[f_\tau^T]$ (i.e., mean value). Since each group of homogeneity features enables distinguishing the target data, the ensemble of these groups will contribute to the following training stage by using the voting ensemble structure.

Second, we apply adversarial training to confuse the features from different source domains. Specifically, we add a domain discriminator D with a Gradient Reversal Layer (GRL) [28] before the fully connected layers so that it can reverse the optimization direction. The domain loss $\mathcal{L}^D$ is defined as:

$$\mathcal{L}^D = -\sum_{j=1}^{L}\sum_{\tau=1}^{N} l_{\tau j}^d \log\left(p_{\tau j}^d\right), \tag{10}$$

where $p_{\tau j}^d$ is the domain prediction, $l_{\tau j}^d$ is the domain label of source sample $X_j^{S_\tau}$, L is the batch size. To achieve maximum confusion, the DNN discards the heterogeneous biological features and keeps the homogeneous biological features.

By optimizing the ensemble loss $\mathcal{L} = \mathcal{L}^C + \gamma(\mathcal{L}^{dis} - \mathcal{L}^D)$, the trained model can achieve adaptation performance as envisaged in Section IV-A1.

B. Self-training Stage

1) Design Rationale: The model trained in the pre-training stage has moderate performance in adapting to a new user. Although the model's performance cannot meet the requirements of practical application, we can explore the latent supervision based on this model. The self-training stage aims to boost the model performance trained from the pre-training stage.

To this end, we purify the pseudo-labels obtained from the pre-training stage and use cleaner supervision to relearn the target-user-specific biological features removed during the heterogeneity-eliminating process. Specifically, we use the predictions from the pre-training stage as the initial pseudo-labels. This process is denoted as "Projection" in Figure 6. Since the target-user-specific biological features were eliminated during the heterogeneity-eliminating process and only a few homogeneous biological features are available, the model's performance on the target user is imperfect. Furthermore, noise (errors) in these pseudo-labels will further deteriorate the training process. To tackle this issue, we use the ensemble of homogeneous biological features from different branches to predict the final predictions, which can reduce the noise in the predictions. Moreover, we propose a clean instances selection strategy to further reduce the error rate in pseudo-labels and obtain cleaner supervision. By continuously purifying the pseudo-labels and relearning the target-user-specific biological features, the model will achieve satisfactory performance on the new user.

2) Problem Formulation: We first use the ensemble outputs predicted by the multi-model voting ensemble DNN as

Algorithm 1 Self-training process.

Require: R_{max}: maximum repetition number; I_{max}: maximum iteration number; I_k: iteration constant; θ: DNN parameters; $\mathcal{D}_T$: target dataset.

```
for R = 1, 2, 3, ..., R_max do
    Shuffle D_T;
    for I = 1, 2, 3, ..., I_max do
        Update dynamic ratio R(I) = 1 − min{(I/I_k)λ, λ};
        Fetch mini-batch B_T from D_T;
        Select R(I) ratio clean instances from B_T;
        Train θ with clean instances and pseudo-labels;
    end for
    Update pseudo-labels to reduce noise;
end for
```

pseudo-labels. Then, we sample small-loss instances to train the DNN under the supervision of corresponding pseudo-labels. By updating the pseudo-labels and the parameter of the multi-model voting ensemble DNN alternately, the noise in pseudo-labels will be reduced. To summarise, the problem to solve in the self-training stage can be formulated as:

$$P(y=j \mid x)=\sum_{\tau=1}^{N} \frac{e^{f_\tau^\top w_{\tau j}+b_{\tau j}}}{\sum_{k=1}^{K} e^{f_\tau^\top w_{\tau k}+b_{\tau k}}}/N, \tag{11}$$

$$\min_{\theta_F,\theta_{H_\tau},\theta_{C_\tau}} \sum_{\tau=1}^{N} \mathcal{L}_\tau^C(\theta_F,\theta_{H_\tau},\theta_{C_\tau}), \tag{12}$$

where $P(y = j \mid x)$ denotes the probability of sample x belonging to class $j \in \{1, 2, \ldots, K\}$. Moreover, f_τ denotes the deep features of the sample x extracted by common and τ-th domain-specific feature extractors. w_τ and b_τ denote the τ-th domain-specific classifier's mapping relationships from feature f_τ to prediction.

3) Design Details: There are many techniques to learn from noisy supervision, such as (1) adding a noise adaptation layer [31]; (2) leveraging an additional validation dataset [32]; (3) assigning weights for the training samples [33]; (4) training a label cleaning network with a part of clean instances and use the network to reduce the noise in labels [34]; (5) proposing a joint optimization framework [33]. However, to our knowledge, no attempt has been made to boost the multi-source domain adaptation performance by purifying the noisy pseudo-labels in EMG sensing.

In *EMGSense*, we combine techniques (4) and (5) mentioned above. Specifically, we calculate the target samples' average loss of all branches and select the small-loss instances as the clean instances to train the multi-model voting ensemble DNN. It is because training with small-loss instances instead of all instances improves network performance when learning from noisy supervision [34]. However, there is a foundational problem. Although the small-loss instances have more likelihood of being clean instances, the small-loss instances always belong to the same class or few classes because of the preference of the network. The severe imbalance will greatly corrupt the self-training process. To tackle this problem, we select small-loss instances from each class based on a dynamic ratio $R(I)$. Specifically, the dynamic ratio $R(I)$ is calculated according to the iteration number as follows:

$$R(I)=1-\min\left\{\frac{I}{I_k}\lambda, \lambda\right\}, \tag{13}$$

where I is the number of iterations, I_k denotes a constant to restrain the minimum value of $R(I)$, and $1-\lambda$ denotes the minimum ratio selected from each class. The details of the self-training stage are shown in Algorithm 1.

TABLE I: Gestures and activities we considered in this work.

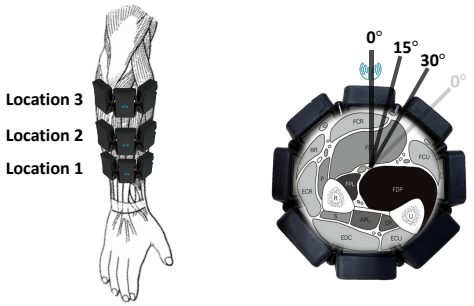
Dataset	Categories	Volume
Gesture	Clap, OK, Snap, Double-tap, Wave-in, and Wave-out [7]	9,898 samples
Activity	Waving, Writing, Typing, Sitting, Running, Applauding, and Boxing	171 min

V. Evaluation

A. Experimental Settings

1) Data Collection: As shown in Figure 7, we leverage the Myo armband which is composed of eight EMG dry sensors for the data collection. The eight sensors measure the action potential of the muscles at a sampling rate of 200 Hz.

To evaluate the performance of *EMGSense* in different EMG cross-user applications, we collected two datasets: a gesture recognition dataset and an activity recognition dataset. Details of the six gestures and seven activities we considered are shown in Table I. A total of 13 healthy volunteers (seven males and six females, aged between 22 and 35, all right-handed) participated in the data collection[1]. As shown in Figure 7, we collected data from three deployment locations at a subject's forearm, with a distance of 5 cm. For each location, we also collected data from three orientations, i.e., 0°, 15°, and 30°. The 0° orientation is defined as the orientation that makes the blue light bar align with the palm of the subject. Since the Myo Armband's eight electrodes are evenly distributed at 360°, the angle between the two adjacent electrodes is 45°. That is, if we rotate the armband by 45°, the electrodes will move to their neighbors' positions, and hence the same EMG signals



(a) Deployment locations (b) Wearing orientations

Fig. 7: Different locations and orientations of MYO Armband.

[1]The study has been approved by our institutional human ethics committee (No. H002554).

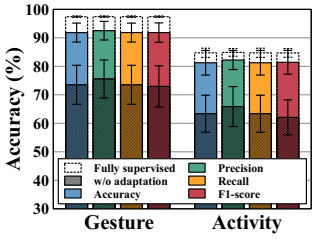


Fig. 8: Overall performance.

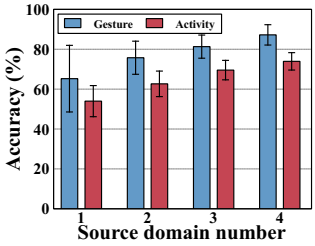


Fig. 10: Impact of N.

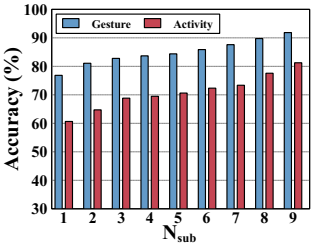


Fig. 11: Impact of N_{sub}.

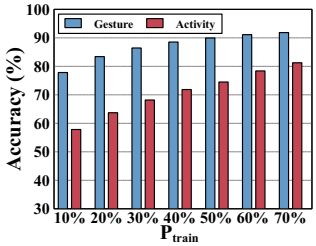


Fig. 12: Impact of P_{train}.

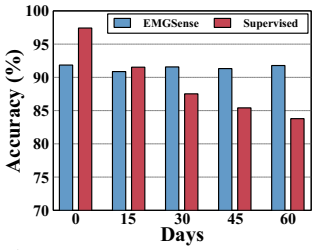


Fig. 13: Long-term accuracy on gesture recognition.

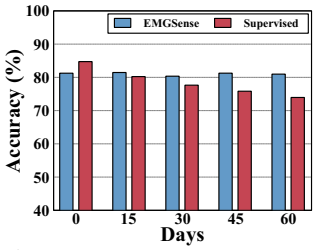


Fig. 14: Long-term accuracy on activity recognition.

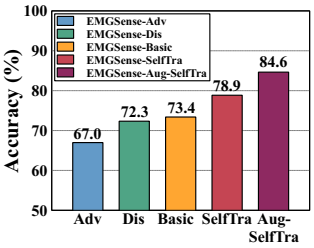


Fig. 15: Ablation study on different versions of *EMGSense*.

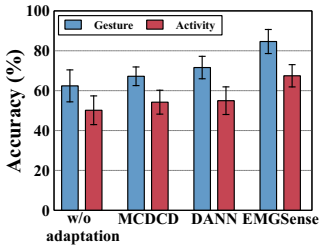


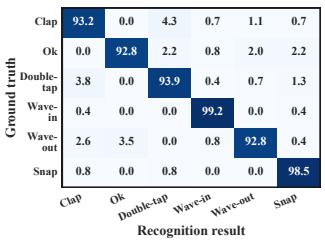
Fig. 16: Comparison on EMG sensing with baselines.

will be recorded. Therefore, we only collected data when the armband is rotated clockwise by 0°, 15° and 30°, respectively.

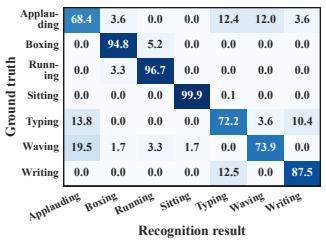
2) Settings: We implement *EMGSense* using the Pytorch framework and train the neural network on a desktop with Nvidia GTX 2080Ti GPU. We adopt the adaptive momentum (Adam) to optimize the model's parameters, and the learning rate of the optimizer is 0.001. The batch sizes are 32 and 256 in the pre-training and self-training stages, respectively. The early stop criterion of the pre-training and self-training stages are 500 and 50 iterations, respectively.

Considering the data collected from different positions, we define the combination of a deployment location and a wearing orientation as a sub-domain. Correspondingly, each user has nine sub-domains (3 locations×3 orientations). For evaluating the performance of *EMGSense*, we define N as the number of source users. Among each user, we define N_{sub} as the number of sub-domains used for training, which indicates how many sub-domains are required to achieve satisfactory performance. Furthermore, we define P_{train} as the ratio between training data and all data in each sub-domain, which indicates how many samples are required to collect for each sub-domain.

In this paper, we conduct the evaluations in a leave-one-out manner by selecting one user as the target domain while the rest users are used as source domains until all the volunteers have been selected as the target domain once. To eliminate random effects, the sub-domains and samples are randomly selected ten times for each setting. The average performance is calculated as the final result. We use Accuracy, Confusion Matrix, Precision, Recall, and F1-Score as the evaluation metrics that comprehensively evaluate the *EMGSense*'s performance.



(a) Gesture recognition



(b) Activity recognition

Fig. 9: Confusion matrices on two EMG sensing tasks.

B. Overall Performance

Since *EMGSense* is a low-effort self-supervised domain adaptation framework working in a data-efficient manner, we evaluate the overall performance of *EMGSense* on whole collected datasets. The results are shown in Figure 8. To intuitively demonstrate the efficiency of *EMGSense*, we also conduct the comparison with the performance achieved by w/o adaptation and fully supervised learning as two different masks in Figure 8. Here, the w/o adaptation mask indicates the inferior overall performance when directly inferring target domain samples using the model trained from source domains; the fully supervised mask indicates the performance upper bound that can be achieved on the collected datasets. The result shows that *EMGSense* model can achieve 91.9% accuracy (92.5% precision) and 81.2% accuracy (82.2% precision) in gesture recognition and activity recognition, respectively. Note that the performance is close to the model trained in a fully supervised manner by 3.5–5.5% below the latter.

Figures 9(a) and 9(b) show confusion matrix examples of *EMGSense* on gesture recognition and activity recognition, respectively. It can be observed that there is no significant difference in the accuracy of different gestures and activities, indicating the robustness of *EMGSense* in different tasks.

C. Micro-benchmark Tests

1) Impact of Source Domain Number: Since the core of *EMGSense* is the multi-source domain adaptation and multi-

model voting ensemble structure, the number of source domains is a vital factor affecting the system performance. Intuitively, the more source domains help the model adapt to a target domain, the better performance the model can achieve. Therefore, we first evaluate the engagement of the number of source domains (users) with the system's accuracy.

Setup: Since the experiment aims to evaluate the influence of source domain number, we fix P_{train} and N_{sub} as 70% (15% for validation, 15% for testing) and 9, respectively. Then, we increase N from 1 to 4 gradually.

Results: As Figure 10 shows, the classification accuracy of *EMGSense* increases gradually with the increment of N. It demonstrates that owing to the multi-source and multi-model voting ensemble structure, *EMGSense* can achieve better performance by increasing the diversity of the source domains.

2) Impact of Source Domain Size: Since *EMGSense* is based on multi-source domain adaptation, the size of source domains plays an important role in the system performance. Intuitively, the more training data are collected from different combinations of locations and orientations, the higher accuracy a system can achieve. However, collecting large amounts of data is labor-intensive and time-consuming. Therefore, we fix the number of source domains and further evaluate the impact of source domain size on the system's accuracy.

Setup: Since the experiment aims to go deep into the domain size, we fix N as the maximum optional value for each dataset. When evaluating N_{sub}, we fix P_{train} as 70% (15% for validation, 15% for testing) and change N_{sub} from 1 to 9 gradually. Similarly, when evaluating P_{train}, we fix N_{sub} as 9 and change P_{train} from 10% to 70% gradually.

Results: Intuitively, as Figures 11 and 12 show, the classification accuracy of *EMGSense* increases gradually with the increment of N_{sub} and P_{train}. Specifically, the accuracy of gesture recognition and activity recognition increases from 76.8% to 91.9% and from 60.7% to 81.2% when N_{sub} increases from 1 to 9. The same pattern can also be observed when P_{train} increases from 10% to 70%.

3) Impact of Time-varying Biological Heterogeneity: Since *EMGSense* is designed for label-free training scenarios, it leaves the potential to optimize the model by leveraging the unlabeled data generated during the using period. Intuitively, as discussed in Section I, the sensing accuracy will degrade as the biological heterogeneity changes over time. Therefore, we evaluate the *EMGSense*'s long-term sensing performance on new users in this part.

Setup: Since the experiment aims to track the models' long-term performance on a new user, we start from the models trained in Section V-B. Then, we test the model accuracy on the new user every fifteen days, lasting for two months. During this period, we let the user use the EMG armband for ten minutes every five days. By leveraging the recorded unlabeled data to optimize the model, we compare the long-term accuracy of a supervised learning model and *EMGSense* model to demonstrate the *EMGSense*'s capability to cope with the time-varying biological heterogeneity.

Results: As Figures 13 and 14 show, the test accuracy of the supervised learning model will decrease with a tendency to fall back to the "w/o adaptation" performance in Figure 8. The reason is that the biological factors of the subject have changed significantly over a long enough period. Therefore, the subject is equal to a new user for the supervised learning model. In contrast, the *EMGSense* model performs robustly to cope with time-varying biological heterogeneity by optimizing the model using unlabeled data generated during use.

D. Ablation Study

Below, we evaluate the efficacy of each system module in *EMGSense*. Specifically, we design five different variants of *EMGSense* by adding different functional modules into the end-to-end system design. Then, we can analyze the impact of different modules in gesture recognition as an example to evaluate the efficacy of each system module in Figure 6. In this experiment, N_{sub} and P_{train} are set to 6 and 30%, respectively. The five variants of *EMGSense* are listed below:

(1) **EMGSense-Adv:** a simplified design of *EMGSense*, which conducts the proposed pre-training stage but only optimizes the classification loss and domain loss.
(2) **EMGSense-Dis:** a simplified design of *EMGSense*, which conducts the proposed pre-training stage but only optimizes the classification loss and discrepancy loss.
(3) **EMGSense-Basic:** the basic design of *EMGSense*, which conducts the proposed pre-training stage and optimizes the classification loss, domain loss, and discrepancy loss.
(4) **EMGSense-SelfTra:** an upgraded version of *EMGSense*-Basic, which conducts the proposed two training stages.
(5) **EMGSense-Aug-SelfTra:** over *EMGSense*-SelfTra, this version further applies data augmentation methods proposed in Section III-B to improve the performance of *EMGSense*-SelfTra. We apply each of the two data augmentation methods four times in this system variant.

The results are shown in Figure 15. Firstly, the comparison among the first two variants (1) and (3) indicates the efficacy of the MK-MMD for aligning the source and target features in deep feature space. Secondly, the accuracy of variants (2) and (3) demonstrate that the domain discriminator can force the DNN to learn more general features and improve the generalization capability. Variant (4) indicates that the proposed self-training stage can cooperate well with the pre-training stage and achieve better performance. Finally, by adding data augmentation to the end-to-end system design, variant (5) can significantly increase the diversity of the training data and can improve the model generalization ability.

E. Comparison with Baselines

1) Baselines: We evaluate *EMGSense* on the collected dataset with three baselines: without adaptation (denoted by w/o adaptation in the rest of the paper), DANN [28] and MCDCD [23]. Here, w/o adaptation means that we use the model trained from source domains to infer the target domain samples directly (recall Section V-B). DANN is a classical domain adaptation approach based on adversarial training.

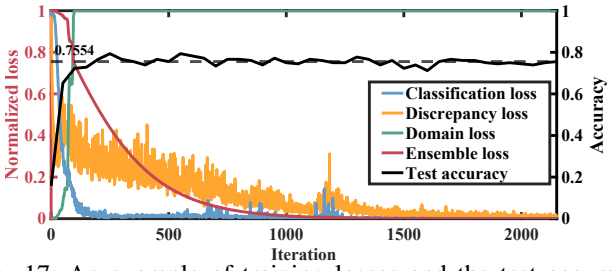


Fig. 17: An example of training losses and the test accuracy of the trained model during the pre-training stage.

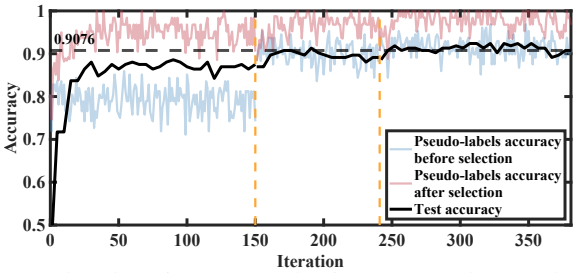


Fig. 18: The clean instances selection strategy boosts the test accuracy during three rounds of the self-training stage.

MCDCD is a multi-source discrepancy-based domain adaptation approach for EMG-based gait detection. For fairness, we employ the same DNN units used in *EMGSense* as the feature extractors and classifiers in baselines. In this experiment, we adopt the same parameter setting as the one in Section V-D.

2) Result: As the results in Figure 16, *EMGSense* achieves an overall accuracy of 84.6% in gesture recognition and 67.5% in activity recognition, respectively. Moreover, *EMGSense* outperforms the existing approaches by 12.5%–17.4%. The result of w/o adaptation is the lowest, demonstrating the significant accuracy degradation caused by user heterogeneity. Additionally, the classical domain adaptation approach (e.g., DANN) and the EMG-oriented solution (e.g., MCDCD) cannot achieve satisfactory results for the following three reasons.

- First, DANN and MCDCD rely on large amounts of data to achieve high performance. However, their performance drops significantly when the number of source domains and the available training data is limited.
- Second, they ignore the heterogeneity among source users and treat all data equally. As a result, some hard-to-transfer samples may not be well treated, leading to inferior performance. However, intrinsic forearm biological factors diversity among different users leads to different domain alignment potentials. Unfortunately, this fact is ignored in existing methods.
- Third, for the same training conditions, the model learned in a supervised manner usually outperforms the one learned in an unsupervised manner. Therefore, how to take full advantage of the unlabeled samples becomes challenging when only unlabeled samples are available. Unfortunately, both DANN and MCDCD fail to explore latent supervision in unlabeled samples.

In contrast to the baselines, we design two EMG-specific data augmentation approaches to increase the variation of EMG signals and reduce the burden of data collection. Furthermore, we formulate the problem as a multi-source domain adaptation problem (the pre-training stage) to leverage the diversity of different subjects. Moreover, we innovatively use self-training techniques (the self-training stage) to explore the latent supervision of unlabeled target samples. The efficacy of the two stages will be evaluated in the next subsection.

F. Pre-training efficacy & Self-training efficacy

Figure 17 shows that the classification loss and the discrepancy loss continuously decline with the increment of iterations, which means the target data is separately aligned to each source data. At the same time, the features of each class are separated from each other in the deep feature space. Contrastively, the domain loss increases rapidly and keeps at the maximum value until the end of the pre-training stage. It indicates that the domain discriminator maximizes the confusion among different source domains, forcing the feature extractors to learn more general features. As a result, the ensemble loss defined in Equation 5 declines smoothly to the minimum value while the test accuracy (i.e., the model performance) rises and trends toward stability. Ultimately, the performance obtained in the pre-training stage will be used as the initialization for the self-training stage.

Figure 18 shows the efficacy of the self-training stage. Specifically, the accuracies of each batch of pseudo-labels before clean instances selection and after selection are shown as the blue line and the red line in Figure 18. It can be observed that the self-training strategy can boost the model to outperform the initial supervision. Three rounds of the self-training stage with an early stop setting are separated by the yellow dashed lines. When each round of the self-training stage stops, the improved model (black line) will be used to initialize the next round's pseudo-labels (blue line) until the improvement is less than a pre-defined threshold. As a result, *EMGSense* achieves a satisfactory performance by combining the proposed augmentation methods, multi-source domain adaptation approach, and self-training strategy.

VI. Related Work

EMG-based sensing. EMG-based sensing has attracted considerable attention in recent years. A variety of novel applications have been proposed, such as authentication [35], activity recognition [4], gesture recognition [6], [7], and hand pose reconstruction [9], [10]. Specifically, Xi et al. [4] proposed a system to identify six activities by wavelet coherence coefficient features and the support vector machine classifier. Zhang et al. [7] designed a gesture recognition system based on an artificial neural network, achieving 98.7% accuracy in classifying five gestures. Additionally, EMG sensors offer rich information for fine-grained hand pose sensing. For example, Liu et al. proposed a system called WR-hand to track 14 skeletal points on the hand [9]. Another group of researchers [10] designed a system named NeuroPose to track finger poses.

Domain adaptation. Domain adaptation is a subcategory of transfer learning whose goal is to train a model on a source

dataset and secure a good accuracy on the target dataset, which is significantly different from the source dataset. It has been widely used to solve the domain shift problem in sensing systems, such as wireless sensing [36]–[38] and wearable sensing [39], [40]. Several attempts have been made to solve the domain shift problem in EMG-based sensing. For example, Côté-Allard et al. [21] proposed a supervised domain adaptation framework for EMG-based hand gesture recognition. The supervised method is further extended to be unsupervised in their recent work [22]. Guo et al. [23] proposed a discrepancy-based approach that leverages the diversity from multiple source domains to adapt to new users.

Different from existing wearable sensor-based solutions, *EMGSense* is a low-effort domain adaptation framework based on EMG-specific data augmentation methods, a multi-source domain adaptation approach, and a self-training strategy. It can leverage the diversity among source domains and the latent supervision from the target domain. Evaluation results show *EMGSense* outperforms existing approaches significantly in two common EMG-based application scenarios.

VII. Conclusions and Future Work

This paper proposes a low-effort self-supervised domain adaptation framework to cope with the heterogeneity challenges in a low-effort (data-efficient and label-free) manner. By integrating the proposed EMG-specific data augmentation methods, the well-designed DNN structure, and the advanced self-supervised techniques, the proposed two training stages can complement each other and train a model which can be adapted to a new user. The model achieves a satisfactory performance close to the model trained in a fully supervised manner and is only 3.5–5.5% below the latter. Comprehensive evaluations indicate that *EMGSense* achieves an average accuracy of 86.6% in two cross-user EMG sensing applications and outperforms the state-of-the-art EMG-oriented approaches by 12.5%–17.4%. *EMGSense* fills the research gap in heterogeneity problems in EMG sensing and will enable a variety of novel EMG-based cross-user applications, such as clinical practice, neurorehabilitation, and human-machine interaction. Future work could research promising topics such as synthesizing EMG signal according to the heterogeneity or generalizing the EMG signal across users, which means the extracted feature can be applied to new users in a zero-effort manner.

Acknowledgment

The work described in this paper was substantially sponsored by the project 62101471 supported by NSFC and was partially supported by the Shenzhen Research Institute, City University of Hong Kong. The work was also partially supported by the Research Grants Council of the Hong Kong Special Administrative Region, China (Project No. CityU 21201420 and CityU 11201422), Shenzhen Science and Technology Funding Fundamental Research Program (Project No. 2021Szvup126), and NSF of Shandong Province (Project No. ZR2021LZH010).

References

[1] J.-Y. Hogrel, "Clinical applications of surface electromyography in neuromuscular disorders," *Neurophysiologie Clinique/Clinical Neurophysiology*, vol. 35, no. 2-3, pp. 59–71, 2005.

[2] H. A. Feldner, D. Howell, V. E. Kelly, S. W. McCoy, and K. M. Steele, ""look, your muscles are firing!": a qualitative study of clinician perspectives on the use of surface electromyography in neurorehabilitation," *Archives of physical medicine and rehabilitation*, vol. 100, no. 4, pp. 663–675, 2019.

[3] Y. Zhao, J. Wang, Y. Zhang, H. Liu, Z. Chen, Y. Lu, Y. Dai, L. Xu, and S. Gao, "Flexible and wearable emg and psd sensors enabled locomotion mode recognition for ioht-based in-home rehabilitation," *IEEE Sensors Journal*, vol. 21, no. 23, pp. 26 311–26 319, 2021.

[4] X. Xi, C. Yang, J. Shi, Z. Luo, and Y.-B. Zhao, "Surface electromyography-based daily activity recognition using wavelet coherence coefficient and support vector machine," *Neural Processing Letters*, vol. 50, no. 3, pp. 2265–2280, 2019.

[5] S. S. Bangaru, C. Wang, S. A. Busam, and F. Aghazadeh, "Ann-based automated scaffold builder activity recognition through wearable emg and imu sensors," *Automation in Construction*, p. 103653, 2021.

[6] J. M. Fajardo, O. Gomez, and F. Prieto, "EMG hand gesture classification using handcrafted and deep features," *Biomedical Signal Processing and Control*, vol. 63, p. 102210, 2021.

[7] Z. Zhang, K. Yang, J. Qian, and L. Zhang, "Real-time surface EMG pattern recognition for hand gestures based on an artificial neural network," *Sensors*, vol. 19, no. 14, p. 3170, 2019.

[8] A. De Silva, M. V. Perera, K. Wickramasinghe, A. M. Naim, T. D. Lalitharatne, and S. L. Kappel, "Real-time hand gesture recognition using temporal muscle activation maps of multi-channel semg signals," in *Proceedings of the International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 2020, pp. 1299–1303.

[9] Y. Liu, C. Lin, and Z. Li, "WR-Hand: Wearable armband can track user's hand," *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies (IMWUT)*, vol. 5, no. 3, pp. 1–27, 2021.

[10] Y. Liu, S. Zhang, and M. Gowda, "NeuroPose: 3D hand pose tracking using EMG wearables," in *Proceedings of the Web Conference (WWW)*, 2021, pp. 1471–1482.

[11] U. Côté-Allard, G. Gagnon-Turcotte, A. Phinyomark, K. Glette, E. Scheme, F. Laviolette, and B. Gosselin, "A transferable adaptive domain adversarial neural network for virtual reality augmented emg-based gesture recognition," *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 29, pp. 546–555, 2021.

[12] P. Bartuzi, T. Tokarski, and D. Roman-Liu, "The effect of the fatty tissue on emg signal in young women," *Acta Bioeng Biomech*, vol. 12, no. 2, pp. 87–92, 2010.

[13] D. G. Park and H. C. Kim, "Muscleman: Wireless input device for a fighting action game based on the emg signal and acceleration of the human forearm," in *Proceedings of the International Symposium on Neural Networks (ISNN)*, 2011.

[14] C. Fleischer, A. Wege, K. Kondak, and G. Hommel, "Application of emg signals for controlling exoskeleton robots," *Biomedical Engineering*, vol. 51, 2006.

[15] R. Merletti, D. Farina, M. Gazzoni, and M. P. Schieroni, "Effect of age on muscle functions investigated with surface electromyography," *Muscle & Nerve: Official Journal of the American Association of Electrodiagnostic Medicine*, vol. 25, no. 1, pp. 65–76, 2002.

[16] J. S. Petrofsky, R. M. Glaser, C. A. Phillips, A. R. Lind, and C. Williams, "Evaluation of the amplitude and frequency components of the surface emg as an index of muscle fatigue," *Ergonomics*, vol. 25, no. 3, pp. 213–223, 1982.

[17] K. Kiguchi, T. Tanaka, and T. Fukuda, "Neuro-fuzzy control of a robotic exoskeleton with EMG signals," *IEEE Transactions on Fuzzy Systems*, vol. 12, no. 4, pp. 481–490, 2004.

[18] M. Wand and T. Schultz, "Towards real-life application of emg-based speech recognition by using unsupervised adaptation," in *Proceedings of the 15th Annual Conference of the International Speech Communication Association (ISCA)*, 2014.

[19] C. Andronache, M. Negru, I. Bădiţoiu, G. Cioroiu, A. Neacsu, and C. Burileanu, "Automatic gesture recognition framework based on forearm emg activity," in *Proceedings of the International Conference on Telecommunications and Signal Processing (TSP)*, 2022, pp. 284–288.

[20] 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," in *Proceedings of the International Conference on Systems, Man, and Cybernetics (SMC)*, 2017, pp. 1663–1668.

[21] U. Côté-Allard, C. L. Fall, A. Drouin, A. Campeau-Lecours, C. Gosselin, K. Glette, F. Laviolette, and B. Gosselin, "Deep learning for electromyographic hand gesture signal classification using transfer learning," *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 27, no. 4, pp. 760–771, 2019.

[22] U. Côté-Allard, G. Gagnon-Turcotte, A. Phinyomark, K. Glette, E. J. Scheme, F. Laviolette, and B. Gosselin, "Unsupervised domain adversarial self-calibration for electromyography-based gesture recognition," *IEEE Access*, vol. 8, pp. 177 941–177 955, 2020.

[23] Y. Guo, X. Gu, and G.-Z. Yang, "MCDCD: Multi-source unsupervised domain adaptation for abnormal human gait detection," *IEEE Journal of Biomedical and Health Informatics*, 2021.

[24] Myo EMG armband. [Online]. Available: https://www.amazon.in/Myo-Gesture-Control-Armband-White/dp/B00O69U344

[25] C.-S. J. Chu, "Time series segmentation: A sliding window approach," *Information Sciences*, vol. 85, no. 1-3, pp. 147–173, 1995.

[26] M. B. I. Reaz, M. S. Hussain, and F. Mohd-Yasin, "Techniques of EMG signal analysis: detection, processing, classification and applications," *Biological Procedures Online*, vol. 8, no. 1, pp. 11–35, 2006.

[27] M. R. Malgireddy, J. J. Corso, S. Setlur, V. Govindaraju, and D. Mandalapu, "A framework for hand gesture recognition and spotting using sub-gesture modeling," in *Proceedings of the 20th International Conference on Pattern Recognition (ICPR)*, 2010, pp. 3780–3783.

[28] Y. Ganin, E. Ustinova, H. Ajakan, P. Germain, H. Larochelle, F. Laviolette, M. Marchand, and V. Lempitsky, "Domain-adversarial training of neural networks," *The Journal of Machine Learning Research*, vol. 17, no. 1, pp. 2096–2030, 2016.

[29] S. Ioffe and C. Szegedy, "Batch normalization: Accelerating deep network training by reducing internal covariate shift," in *Proceedings of International Conference on Machine Learning (ICML)*, 2015, pp. 448–456.

[30] M. Long, Y. Cao, J. Wang, and M. Jordan, "Learning transferable features with deep adaptation networks," in *Proceedings of International Conference on Machine Learning (ICML)*, 2015, pp. 97–105.

[31] S. Sukhbaatar and R. Fergus, "Learning from noisy labels with deep neural networks," *arXiv preprint arXiv:1406.2080*, 2014.

[32] S. Jenni and P. Favaro, "Deep bilevel learning," in *Proceedings of the European Conference on Computer Vision (ECCV)*, 2018, pp. 618–633.

[33] Y. Zhang, Y. Wei, Q. Wu, P. Zhao, S. Niu, J. Huang, and M. Tan, "Collaborative unsupervised domain adaptation for medical image diagnosis," *IEEE Transactions on Image Processing (TIP)*, vol. 29, pp. 7834–7844, 2020.

[34] B. Han, Q. Yao, X. Yu, G. Niu, M. Xu, W. Hu, I. Tsang, and M. Sugiyama, "Co-teaching: Robust training of deep neural networks with extremely noisy labels," in *Proceedings of the Conference on Neural Information Processing Systems (NeurIPS)*, 2018.

[35] B. Fan, X. Liu, X. Su, P. Hui, and J. Niu, "EmgAuth: An emg-based smartphone unlocking system using siamese network," in *Proceedings of the International Conference on Pervasive Computing and Communications (PerCom)*. IEEE, 2020, pp. 1–10.

[36] Y. Zheng, Y. Zhang, K. Qian, G. Zhang, Y. Liu, C. Wu, and Z. Yang, "Zero-effort cross-domain gesture recognition with Wi-Fi," in *Proceedings of the ACM Annual International Conference on Mobile Systems, Applications, and Services (MobiSys)*, 2019, pp. 313–325.

[37] W. Jiang, C. Miao, F. Ma, S. Yao, Y. Wang, Y. Yuan, H. Xue, C. Song, X. Ma, D. Koutsonikolas *et al.*, "Towards environment independent device free human activity recognition," in *Proceedings of the 24th Annual International Conference on Mobile Computing and Networking (MobiCom)*, 2018, pp. 289–304.

[38] T. Gong, Y. Kim, J. Shin, and S.-J. Lee, "MetaSense: Few-shot adaptation to untrained conditions in deep mobile sensing," in *Proceedings of the ACM Conference on Embedded Networked Sensor Systems (SenSys)*, 2019, pp. 110–123.

[39] A. Natarajan, G. Angarita, E. Gaiser, R. Malison, D. Ganesan, and B. M. Marlin, "Domain adaptation methods for improving lab-to-field generalization of cocaine detection using wearable ECG," in *Proceedings of the ACM International Joint Conference on Pervasive and Ubiquitous Computing (UbiComp)*, 2016, pp. 875–885.

[40] A. Akbari and R. Jafari, "Transferring activity recognition models for new wearable sensors with deep generative domain adaptation," in *Proceedings of the ACM International Conference on Information Processing in Sensor Networks (IPSN)*, 2019, pp. 85–96.