2022 18th International Conference on Distributed Computing in Sensor Systems (DCOSS) | 978-1-6654-9512-7/22/$31.00 ©2022 IEEE | DOI: 10.1109/DCOSS54816.2022.00017

Semi-supervised Multi-source Domain Adaptation in Wearable Activity Recognition

Avijoy Chakma*, Abu Zaher Md Faridee*, Raghuveer Rao†, Nirmalya Roy*
*Information Systems, University of Maryland Baltimore County
†Army Research Lab (United States)
*{achakma1, faridee1, nroy}@umbc.edu, †raghuveer.m.rao.civ@army.mil

***Abstract*—The scarcity of labeled data has traditionally been the primary hindrance in building scalable supervised deep learning models that can retain adequate performance in the presence of various heterogeneities in sample distributions. *Domain adaptation* tries to address this issue by adapting features learned from a smaller set of labeled samples to that of the incoming unlabeled samples. The traditional domain adaptation approaches normally consider only a single source of labeled samples, but in real world use cases, labeled samples can originate from multiple-sources – providing motivation for *multi-source domain adaptation (MSDA)*. Several MSDA approaches have been investigated for wearable sensor-based human activity recognition (HAR) in recent times, but their performance improvement compared to single source counterpart remained marginal. To remedy this performance gap that, we explore multiple avenues to align the conditional distributions in addition to the usual alignment of marginal ones. In our investigation, we extend an existing multi-source domain adaptation approach under semi-supervised settings. We assume the availability of partially labeled *target domain* data and further explore the pseudo labeling usage with a goal to achieve a performance similar to the former. In our experiments on three publicly available datasets, we find that a limited labeled *target domain* data and pseudo label data boost the performance over the unsupervised approach by 10-35% and 2-6%, respectively, in various domain adaptation scenarios.**

***Index Terms*—Domain Adaptation, Multi-source, Multi-domain, Activity Recognition, Wearables**

I. Introduction

The advancement of deep learning methodologies has enabled significant progress in various wearable sensor-based activity recognition applications in a number of areas such as health care, assisted daily living, sports analytics, smart buildings, and smart cities. Compared to traditional machine learning models, deep learning approaches are especially good at modeling progressively complex and non-linear relationships from raw unstructured data, providing superior performance. This comes at the prohibitive cost of time-consuming collection, pre-processing, and annotation of a large swathe of labeled data to train the supervised deep learning models. Additionally, variation in on-body placement, user behavior, and overall environmental changes introduce divergence in the raw data distributions among the data sources. The traditional machine learning algorithms assume close similarity between the training and testing data distributions, and these models fail to generalize when faced with such heterogeneities. To handle such cases, specialized transfer learning algorithms (i.e. *domain adaptation*) has been proposed which tries to minimize the divergence between the said two distributions (often termed *source* for the training set and *target* for the testing set) without requiring any extra labeled samples.

Most of these domain adaptation approaches only consider a single source (i.e. single data distribution), but in real life, the samples often originate from multiple sources; hence often after real-world deployment, these single source domain adaptation (SSDA) approaches provide lackluster performance benefit. When multiple source domains are present, state-of-the-art SSDA approaches either choose only a single-source from multiple ones or totally ignore the multi-source information by aggregating samples from all sources prior to the adaptation process. With a more generalized assumption that different subsets of target domain samples are related to each of the source domains in their unique way, a limited number of multi-source domain adaptation (MSDA) approaches have tried to address such perceived limitation of SSDA in wearable sensor-based HAR literature. However, these approaches suffer from a few technical and practical limitations such as (i) only attempting to align the marginal distribution between the source and target domains but completely ignoring the conditional distribution, (ii) limited handling of heterogeneity types (e.g., considering only cross-person heterogeneity while ignoring other heterogeneous scenarios) and (iii) insignificant classification performances.

To overcome the mentioned limitations in MSDA, we investigate and extend an unsupervised MSDA approach under semi-supervised settings with an assumption that labels can be approximated (from derived *pseudo* labels or manual annotations) for a small subset of samples residing in the target domain. Such limited labeled target data would allow aligning both the marginal and conditional data distributions. Apart from the limited target data, we also explore one pseudo label generation mechanism to investigate its impact on MSDA. In our investigation, we aim to discover the following aspects in MSDA:

1) evaluate the performance when both marginal and conditional distribution are aligned in the domain adaptation process
2) measure the amount of labeled target domain data in order to achieve substantial classification performance
3) whether pseudo labeled target data can replace the limited labeled target domain data

Key Contributions: In our study, we make the following contributions -

2325-2944/22/$31.00 ©2022 IEEE
DOI 10.1109/DCOSS54816.2022.00017

- *Semi-supervised multi-source domain adaptation*: We extend an adversarial learning-based MSDA framework for a semi-supervised learning set-up (referred to as SS-MSADA henceforth) in which we assume the availability of a small percentage of labeled target data and exploit it in the network training. In determining the amount of the labeled data, we take motivation from the literature [1] and assume the availability of 10% labeled target domain data where the label space and the feature space are homogeneous across all the domains.
- *Implication of pseudo-label target data in domain adaptation*: We explore one particular approach for pseudo label generation as an additional part of the semi-supervised learning mechanism. In our investigation, we find that generating pseudo labels at the end of the adaptation process increases the performance by 2-3%. In addition, we also discover different implications on pseudo label generating timing and its impact on the network training.
- *Extensive experiments on three publicly available datasets*: We evaluate our proposed framework with three publicly available datasets and report our experimental findings on two domain adaptation settings (a scenario that generates data distribution heterogeneity between two data sources) - cross-person and cross-position heterogeneity.

The paper is organized as follows - we briefly discuss on the MSDA and various semi-supervised mechanisms in wearable activity recognition in section II. We describe our proposed framework in section III. In section IV, we report our experimental findings and analysis. Finally, section V concludes our investigation.

II. Related Works

A. Wearable Multi-source Domain Adaptation

Several approaches have been proposed to tackle data distribution heterogeneity between a single source and target domain in wearable sensor-based activity recognition [1]–[4]. However, limited attempts have been made to incorporate multiple source domains in the adaptation process. [2] considers multiple source domains, and instead of simultaneous processing, explicitly selects the most relevant domain from the multiple source domains based on the cosine similarity with the target domain and uses the selected domain for the domain adaptation process. Whereas *SenseHAR* [5] proposes a data-fusion-based approach to mitigate the heterogeneous data distribution and assign labels to the unlabeled data. Authors of [5] combine multiple sensor data so that each sensor data can complement the other in achieving the intended tasks. [6] proposes an adversarial-based approach to tackle multi-source domain adaptation where multiple source domains are processed concurrently and selects the relevant source domain with the target domain using the noble perplexity scoring mechanism. [7] proposes another adversarial-based approach that employs a domain discriminator to guide the domain-invariant feature learning process, and the proposed approach tackles cross-person heterogeneity in ADL activities whereas [6] considers cross-person and cross-position heterogeneity.

B. Pseudo Labeling in Unsupervised Feature Representation Learning

The label assignment process of the unlabeled data samples is known as the pseudo labeling, and the concept of *pseudo label* is categorized as a semi-supervised learning mechanism [8]. [8] broadly discusses various semi-supervised approaches for attaining the pseudo labels. Nonetheless, pseudo labeling is also exploited in the unsupervised learning approach [9]. In unsupervised learning, target domain data do not have the label information, and therefore, conditional distribution alignment is ignored. Pseudo labeling attempts to fill that gap by assigning label information to the target domain data samples. In wearable activity recognition, there have been a few approaches proposed that leverage pseudo labels during the network training process. [9] applies K-means clustering algorithm over the extracted feature to generate pseudo labels. In this approach, the clustering convergence requires at least 400 epochs. [10] proposes a semi-supervised approach where a small amount of labeled dataset is used to find the label information for the unlabeled dataset. The label information of the unlabeled dataset is predicted based on applying a high confidence sample filtering technique. In literature, there has been very limited exploration made of pseudo labels in domain adaptation problem.

C. Semi-supervised Multi-source Domain Adaptation

[1] provides an analysis of the labeled target domain data usage in the adaptation process. Whereas [11] proposes a non-deep learning-based approach and explores a multi-classifier agreement and threshold-based pseudo label generation mechanism, which helps in conditional distribution alignment. Both these approaches consider a single source and target domain in domain adaptation. [7], authors extended the proposed adversarial approach using weak supervision in the form of the prior class data distribution. This prior class data distribution is exploited in minimizing the KL-Divergence with the network estimated class distribution.

The key difference between our approach to the discussed literature work is that we explore the impact of the small percentage of the labeled target data in multi-source domain adaptation, which has not been studied in the literature work discussed in II-A. Besides, we investigate a pseudo label generation approach and its impacts on the network training. Such investigation allows us to understand whether a similar performance of using partially labeled target data, can be achieved by leveraging pseudo labeled target data and at the same time such investigation also engenders the hidden aspects of pseudo label usage in domain adaptation. In a nutshell, our work studies an unsupervised multi-source domain adaptation under the guidance of two semi-supervised settings - partially labeled and pseudo labeled target data. In the next section, we discuss the explored framework and the training mechanism.

III. Methodology

In this section, we elaborate on the semi-supervised multi-source domain adaptation framework, SS-MSADA, which is an extension of an unsupervised domain adaptation approach, MSADA [6]. The framework is depicted in Figure 1. We describe the problem formulation, architecture details, and the network training mechanism in the following.

A. Problem Formulation

In semi-supervised multi source domain adaptation problem, we assume that there are N source domains with labels and one target domain with limited labeling information. The source domains samples are drawn from N-different data distributions $P_{s_j}(x,y)_{j=1}^{N}$. These data distributions generate the labeled source domain samples, $(X_{s_j}, Y_{s_j})_{j=1}^{N}$ where $X_{s_j} = \{x_{s_{j_i}}\}_{i=1}^{|X_{s_j}|}$ belongs to the source, s_j and $Y_{s_j} = \{y_{s_{j_i}}\}_{i=1}^{|X_{s_j}|}$ is the ground truth for the corresponding samples from the source s_j. We consider $P_t(x,y)$ as the probability distribution that generates the target domain samples, $X_t = \{x_{t_i}\}_{i=1}^{|X_t|}$ with the label space Y_t. We assume X_t is composed of labeled, X_{lt} and unlabeled, X_{ut} target domain samples such that $X_t = X_{lt} + X_{ut}$. We consider that both the source and target domains have homogeneous label spaces and motivated by [1], X_{lt} consists of 10% of the target domain samples which is stratified by the activity labels.

B. Architecture

The adapted MSADA [6] architecture consists of four components - feature extractor, domain discriminator, label classifier and label decider.

Feature Extractor: Feature extractor module consists of three two-dimensional convolution layers, each is followed by a max-pooling layer. The convolutional layers capture the generalizable features, and max-pooling layers help to down-sample the input features into lower dimensions to provide scale-invariant feature representations across multiple datasets. In the proposed framework, the same feature extractor is shared between the *Domain Discriminators* and *Label Classifiers* as shown in Figure 1 with an aim to capture the domain invariant features. Through learning the domain invariant feature across multiple domains, the marginal data distribution is minimized, and leveraging the partially labeled target data, conditional distribution is aligned.

Domain Discriminator: Domain discriminator module consists of the N-source domain-specific discriminators, $\{D_{s_j}\}_{j=1}^{N}$ each tries to predict the origin of the incoming activation from the respective source or target feature distributions. In the framework, each dedicated source-specific domain discriminator D_{s_j} serves as a regular adversarial unit of an adversarial learning mechanism.

In the adversarial learning mechanism, the generator aims to generate domain-invariant features of the source and target domain samples so that the discriminator can not reach its goal of predicting the origin of the incoming features. The overall mechanism resembles a min-max game, and over the process, both generator and discriminator become better at achieving the corresponding goal. In addition to extracting the domain invariant features in the proposed approach, the feature extractor serves the role of a feature generator for the domain discriminators during the adversarial learning phase.

Each of the source-specific discriminators consists of two layers of a fully connected layer with *SELU* in the first layer and a sigmoid activation in the final layer. Each dedicated source-specific domain discriminator process the incoming features from the target domain, $F(x_t)$ and a specific source domain, $F(x_{s_j})$ samples. Besides, providing feedback through an adversarial loss to the feature extractor of its domain-invariant feature generation, each of the domain discriminators serves an additional role of providing a perplexity score for the target domain data to the classifiers. The *perplexity score* acts as a measure of the closeness of the target domain with a specific source domain. Perplexity score serves as a classification *weight* of the corresponding dedicated source domain classifier. The perplexity score is calculated using the following formula:

$$P_{s_j}(x_t; F, D_{s_j}) = -log(1 - D_{s_j}(F(x_t))) \tag{1}$$

Label Classifier: Label classifier consists of N source domain-specific multi-class classifiers $\{C_{s_j}\}_{j=1}^{N}$. Classifier design is similar to the domain discriminator except for the softmax activation in the final layer, which is configured according to the label space of the corresponding source domain j. Each classifier C_{s_j}, along with the shared feature extractor, is pre-trained with labeled data from corresponding j source domains. Further, using the labeled target data, each classifier C_{s_j} is fine-tuned during the adaptation process.

Label Decider: The label decider determines the label for the unlabeled target data. It receives the perplexity scores provided by the source-domain specific discriminators and the classification results, $C_{s_j}(F(x_t))$ provided by the source-domain specific *label classifiers*. Then, the label decider re-weights the classification scores using the perplexity scores and performs a weighted classification to assign a label to the target data.

C. Model Learning

SS-MSADA learns the model leveraging the labeled and unlabeled datasets from both source and target domains, respectively. Feature extractor and the source-domain specific classifiers are pre-trained using the labeled source-domain data. We adapt the pre-trained feature extractor by employing the adversarial objective from the domain discriminators to the feature extractor to align the domain invariant features.

The adversarial objective can be constructed as -

$$\min_{F} \max_{D} V(F, D; \bar{C}) = L_{adv}(F, D) + L_{cls}(F, \bar{C}) \tag{2}$$

$$\begin{aligned} L_{adv}(F, D) = \frac{1}{N} \sum_{j}^{N} & \mathbb{E}_{x \sim X_{s_j}}[\log D_{s_j}(F(x))] \\ & + \mathbb{E}_{x_t \sim X_t}[\log(1 - D_{s_j}(F(x_t)))] \end{aligned} \tag{3}$$

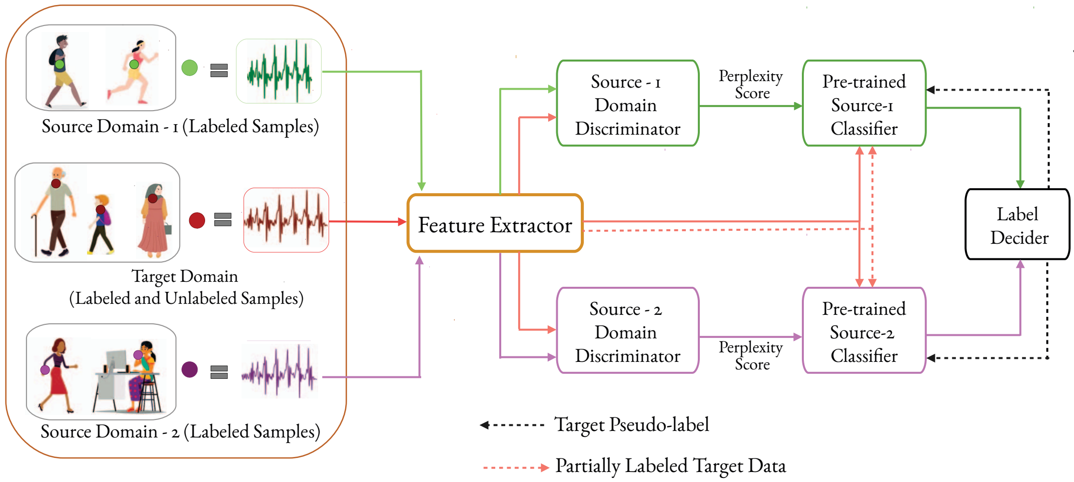

Fig. 1: Overview of our proposed *SS-MSADA* framework that employs pre-trained feature extractor and classifiers. The framework employs a dedicated classifier and domain discriminator for each source domain, where the domain discriminators align the marginal probability distributions using the unlabeled data, and the classifiers leverage the labeled data to help in the conditional distribution alignment. Instead of using labeled target domain data, the *SS-MSADA* can also be extended to generate pseudo labels on the target domain data to help the conditional distribution alignment.

where the first term of the equation (2) defines the adversarial mechanism [12] and the second term denotes the multi-class classification loss. Due to the adverse impact of multiple data distributions on the feature extractor learning process as reported in [12], the optimization works well for D but not F. Therefore, to enable stable learning for the feature extractor, we replace the adversarial loss in equation (2) with the confusion loss as follows -

$$L_{adv}(F,D) = \frac{1}{N}\sum_{j}^{N} \mathbb{E}_{x\sim X_{s_j}} L_{cf}(x;F,D_{s_j}) + \mathbb{E}_{x_t\sim X_t} L_{cf}(x_t;F,D_{s_j}) \tag{4}$$

Here the first and second term determines the domain confusion loss for the source samples and target samples, respectively. Domain confusion loss for both source and target domain samples is calculated using the equation (5).

$$L_{cf}(sample;F,D_{s_j}) = \frac{1}{2}\log D_{s_j}(F(sample)) + \frac{1}{2}\log(1 - Ds_j(F(sample))) \tag{5}$$

In a summary, we train the domain discriminator, D using equation (3) whereas the feature extractor is trained using equation (2) with the adversarial loss is replaced with a confusion loss calculated through equation (4) for a stable learning process. Each classifier is fine-tuned using the partially labeled target data following the adaptation process in each training epochs. Algorithm **??** summarizes the overall *SS-MSADA's* approach. In the following, we discuss our experimental results, and we include the dataset, experimental set-up, implementation details in appendix A, B, C subsections.

Algorithm 1: Learning Algorithm for *SS-MSADA*

Input: N source labeled datasets $\{X_{s_j}, Y_{s_j}\}_{j=1}^{N}$; unlabeled target dataset $X_u t$; partially labeled target dataset $X_l t$; pre-trained feature extractor F and label classifier $\overline{C}$; domain discriminator D; adversarial iteration threshold β

Output: well-trained feature extractor F^*, domain discriminator D^*, label classifier C^*.

```
Pre-train C̄ and F
while not converged do
    for 1:β do
        Sample mini-batch from {X_sj}_{j=1}^N and X_t;
        Update D by Eq. (4);
        Update F by Eq. (2);
        Update C̄ by categorical cross-entropy loss
    using labeled X_lt;
    end for
end while
return F* = F; D* = D; C* = C̄.
```

IV. Results

In this section, we compare the performance of *SS-MSADA* with the baseline approaches on three publicly available datasets under two heterogeneous data distribution settings discussed earlier.

A. Performance Evaluation

Cross-person Heterogeneity: In cross-person domain adaptation on three public datasets, we consider each body position individually within each dataset. Under each position, different users' data operates as the domain, and we perform domain adaptation within different permutations of three persons from all the available persons in a dataset. In the considered persons, two person's body positional data is used as the source domains and the third person's data as the target domain. We ignore such permutations where the two source domains are interchanged in the permutations. For example, consider two user arrangements (a, b, c) and (b, a, c) where User-a and User-b are considered as the source domains and User-c as the target domain. We ignore the settings (b, a, c) in this scenario. We report the average of all the experiments for a particular body position and repeat the procedure for all other positions.

Table I, Table II and Table III reports the cross-person heterogeneity for the OPPORTUNITY, PAMAP2 and DSADS dataset respectively. In the case of the OPPORTUNITY dataset, RevGrad [13] outperforms the *SS-MSADA* and MADA by a large margin, whereas *SS-MSADA* performs superior to the other two in the case of PAMAP2 and DSADS dataset. One key observation is that both RevGrad [13] and MADA [14] utilize a single classifier, whereas *SS-MSADA* uses a dedicated classifier for each source domain which provides a big advantage in terms of achieving high performance, and the classifiers are consist of only two fully connected layers which do not poses a resource burden. For all three datasets, MADA [14] performs lower than the other approaches.

TABLE I: Comparison of Cross Person Domain Adaptation in **OPPORTUNITY** Dataset represented by F1 score

Target Position	RevGrad [13]	MADA [14]	MSADA [6]	SS-MSADA
BACK	**82.07**	75.73	78.01	79.22
RUA	**91.41**	82.71	85.77	85.92
RLA	**88.73**	75.98	78.16	78.47
LUA	**90.81**	81.77	89.16	87.49
LLA	**88.07**	76.6	82.85	77.39
Average	**88.21**	78.55	82.79	81.69

TABLE II: Comparison of Cross Person Domain Adaptation in **PAMAP2** Dataset represented by F1 score

Target Position	RevGrad [13]	MADA [14]	MSADA [6]	SS-MSADA
DA	53.3	49.88	52.75	**77.36**
Torso	62.14	57.01	53.95	**83.53**
DL	56.95	48.49	62.04	**80.16**
Average	57.46	51.79	56.24	**80.35**

TABLE III: Comparison of Cross Person Domain Adaptation in **DSADS** Dataset represented by F1 score

Target Position	RevGrad [13]	MADA [14]	MSADA [6]	SS-MSADA
TORSO	80.46	52.47	59.09	**87.22**
RA	80.33	63.11	65.93	**92.99**
LA	81.45	66.22	71.18	**92.58**
RL	87	58.68	75.33	**91.15**
LL	85.82	53.99	70.6	**91.05**
Average	83.01	58.89	68.43	**91.1**

Cross-position Heterogeneity: Similar to IV-A, we report cross-position heterogeneity for three datasets. Under cross-position domain adaptation, we consider each person individually within each dataset. Different body positional data of each person operates as the domain, and we perform domain adaptation within different permutations of three-body positions from all the available positions in a dataset. Among the considered positions, two body positional data is used as the source domains and the third positional data as the target domain. We ignore such permutations where the two source domains are interchanged in the permutations. For example, consider two different position arrangements (A, B, C) and (B, A, C) where Position-A and Position-B are considered as the source domains and Position-C as the target domain. We ignore the settings (B, A, C) in this scenario. We report the average of all the experiments for each person and repeat the procedure for all other persons.

Table IV, Table V and Table VI presents the cross-position domain adaptation performance for OPPORTUNITY, PAMAP2 and DSADS dataset respectively. Under this setting, *SS-MSADA* performs better than the two baseline approaches. One observation is that similar to cross-person heterogeneity, the performance of *SS-MSADA* on DSADS dataset is close to 90%, but below 75% for both OPPORTUNITY and PAMAP datasets. One significant difference with the DSADS dataset against the other two datasets is that DSADS contains daily activities and sports activities, whereas OPPORTUNITY and PAMAP2 datasets contain only daily livings activities. We assume that the variation in the activities played a significant role in this performance boost.

Insight: Apart from the activity types, we note the scalability aspect of general domain adaptation approaches that deploy a single classifier when the cross-position heterogeneity is considered. When dealing with multiple source domains, either multiple source domain data needs to merge as we did in this paper or select the most relevant one with the target domain before the domain adaptation. In the former case, domain adaptation approaches that deploy a single classifier such as MADA [14], RevGrad [13] performs competitively in cross-person heterogeneity as presented in Table I and Table III. But the performance severely deteriorates when multiple body positional data is combined under the cross-position domain adaptation (refer Table V and Table VI). Such observation highlights the advantage of using dedicated classifiers.

TABLE IV: Comparison of Cross Body Position Domain Adaptation in **OPPORTUNITY** Dataset represented by F1 score

Target User	RevGrad [13]	MADA [14]	MSADA [6]	SS-MSADA
User-1	64.97	54.17	60.76	**77.59**
User-2	**81.39**	67.83	64.79	77.95
User-3	67.33	56.34	60.27	**73.23**
User-4	**74.52**	61.14	63.68	70.25
Average	72.1	59.87	62.37	**74.75**

B. Pseudo Label Investigation

Compared to the [6], a small percentage of labeled target data during the adaptation significantly boosts the perfor-

TABLE V: Comparison of Cross Body Position Domain Adaptation in **PAMAP2** Dataset represented by F1 score

Target User	RevGrad [13]	MADA [14]	MSADA [6]	SS-MSADA
User-1	32.99	21.35	19.8	**74.48**
User-2	38.37	19.88	11.23	**73.7**
User-5	29.89	18.59	14.47	**69.07**
User-8	22.14	19.79	10.94	**73.61**
Average	30.84	19.90	14.14	**72.72**

TABLE VI: Comparison of Cross Body Position Domain Adaptation in **DSADS** Dataset represented by F1 score

Target User	RevGrad [13]	MADA [14]	MSADA [6]	SS-MSADA
User-1	71.67	36.67	49.74	**84.48**
User-2	65.63	34.61	54.23	**86.12**
User-3	59.69	35.5	51.69	**83.15**
User-4	64.09	37.14	49.74	**92.58**
User-5	59.9	39.56	52.53	**96.35**
User-6	65.91	30.26	43.62	**88.98**
User-7	74.4	36.28	52.53	**85.75**
User-8	72.16	42.76	52.11	**91.22**
Average	66.68	36.6	50.77	**88.32**

mance, and to some extent, such performance serves as a baseline. Observing such performance boost due to the small percentage of labeled data, we attempt to generate labels (pseudo label) for the target data using *SS-MSADA* framework and use the pseudo labeled data for the network fine-tuning. We acknowledge that there exist multiple approaches for generating pseudo labels. In this paper, we investigate the network prediction for the labels of the unlabeled target domain as our pseudo label generation process. We discover several aspects of pseudo label generation and its usage in our investigation.

1) *Pseudo label generation during domain adaptation process*

We notice that the general pseudo label generation approaches will provide incorrect pseudo labels because of multiple heterogeneous data distributions. Therefore, one intuitive solution would be to generate at a certain stage of the domain adaptation process. Following our intuition, we attempt to generate the pseudo label at different training epochs and fine-tune the classifiers using the pseudo labeled target data. We consider two different percentages (10% and 100%) of the calculated loss during the backpropagation, but we observe a similar performance in both cases. Figure 2 and Figure 3 presents the performance graph of pseudo label data usage at different training epochs on OPPORTUNITY on cross-person and cross-position heterogeneity respectively. Similar analysis is presented for PAMAP2 dataset in Figure 4 and Figure 5.

2) *Pseudo label generation after the unsupervised domain adaptation*

The stagnant performance in our previous experiment could have several potential reasons. One potential reason could be that the network is not robust even to a small percentage of the incorrect pseudo label data, which hinders the performance

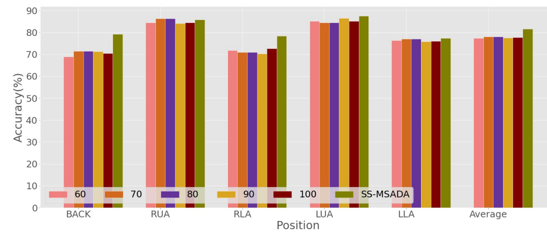

Fig. 2: Performance of cross-person transfer for **OPPORTUNITY** dataset with leveraging pseudo labeled target data at different training epochs and comparing with *SS-MSADA*.

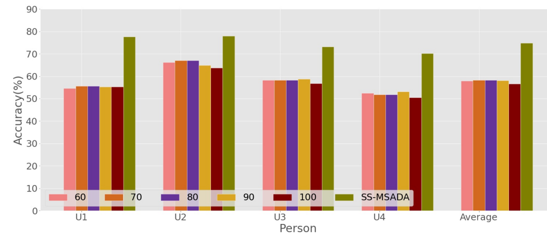

Fig. 3: Performance of cross-position transfer for **OPPORTUNITY** dataset with leveraging pseudo labeled target data at different training epochs and comparing with *SS-MSADA*.

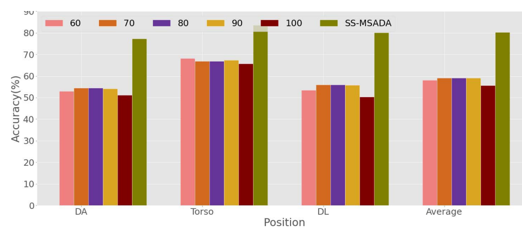

Fig. 4: Performance of cross-person transfer for **PAMAP2** dataset with leveraging pseudo labeled target data at different training epochs and comparing with *SS-MSADA*.

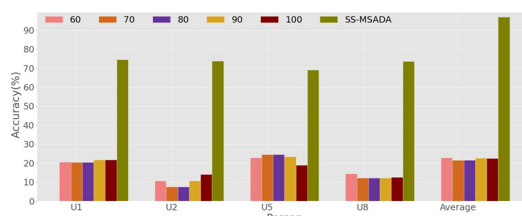

Fig. 5: Performance of cross-position transfer for **PAMAP2** dataset with leveraging pseudo labeled target data at different training epochs and comparing with *SS-MSADA*.

improvement. Following such intuition, we attempt to generate the pseudo label once the unsupervised domain adaptation is completed and further train the network for 30 epochs. We generate pseudo labels for the target domain validation split and train the network using 10% of the calculated loss using the pseudo labeled data. In our experiments on the OPPORTUNITY and DSADS dataset, we observe up to 10% and 7% increase respectively compared to the unsupervised network (MSADA) performance in all the cross-person heterogeneity domain adaptation experiments. Figure 6 and Figure 7 represents the pseudo label generation and its usage under cross-person heterogeneity for OPPORTUNITY and DSADS dataset respectively. The y-axis represents the performance difference between the MSADA and pseudo-label usage. In some cases, the performance does not change and sometimes drops by 4-6%, but overall the performance improves by 2-3%.

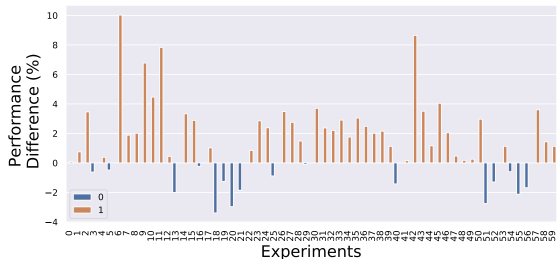

Fig. 6: Performance difference in cross-person transfer in **OPPORTUNITY** dataset between MSADA [6] and after using the MSADA [6]-provided pseudo labels. **1** and **0** in the legend indicate performance improvement and degradation after using pseudo labels respectively.

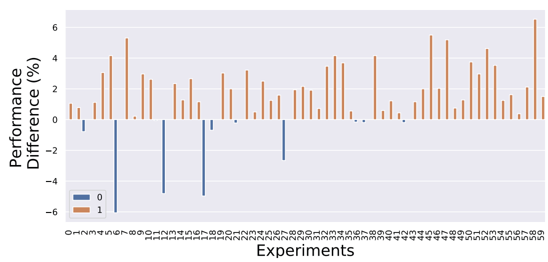

Fig. 7: Performance difference in cross-person transfer in **DSADS** dataset between MSADA [6] and after using the MSADA [6]-provided pseudo labels. **1** and **0** in the legend indicate performance improvement and degradation after using pseudo labels respectively.

C. Combined Impact of Partially Labeled Data and Pseudo Labeling

In our previous investigation, we find that pseudo label data increases the performance over the unsupervised MSADA approach. Here, we investigate the combined impact of using partially labeled target domain data and pseudo label data. Figure 8–13 presents the cross-person and cross-position heterogeneity performance comparison among the unsupervised MSADA, semi-supervised SS-MSADA and semi-supervised SS-MSADA with pseudo label data for OPPORTUNITY, PAMAP2 and DSADS dataset. When the pseudo label data is leveraged along with the partially labeled target data, we observe a 1-4% increase in cross-person heterogeneity, and for cross-position heterogeneity, the performance improves by 4-6%.

We discuss several additional aspects on pseudo label quantity-quantity trade off, our preliminary experience on different attempts to generate pseudo label data and future guideline on pseudo label data in D subsection under appendix.

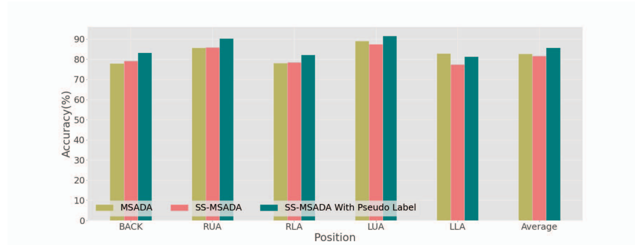

Fig. 8: Performance comparison of cross-person transfer for **OPPORTUNITY** dataset among MSADA [6], *SS-MSADA* and *SS-MSADA* with the pseudo label usage.

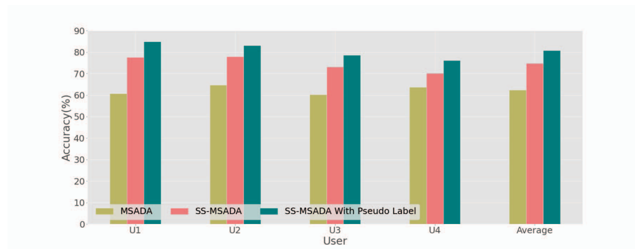

Fig. 9: Performance comparison of cross-position transfer for **OPPORTUNITY** dataset among MSADA [6], *SS-MSADA* and *SS-MSADA* with the pseudo label usage.

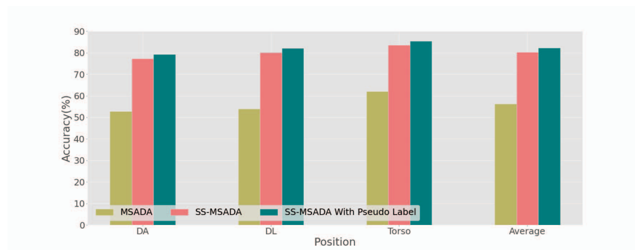

Fig. 10: Performance comparison of cross-person transfer for **PAMAP2** dataset among MSADA [6], *SS-MSADA* and *SS-MSADA* with the pseudo label usage.

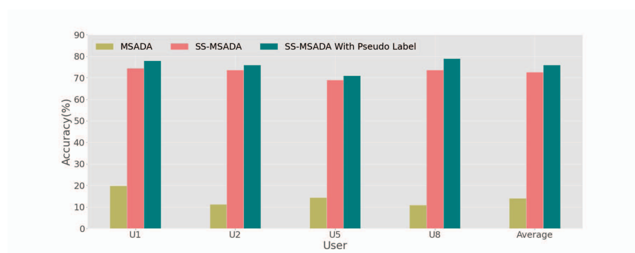

Fig. 11: Performance comparison of cross-position transfer for **PAMAP2** dataset among MSADA [6], *SS-MSADA* and *SS-MSADA* with the pseudo label usage.

V. Conclusion

In this paper, we investigate the multi-source domain adaptation under semi-supervised settings. Our investigation explores two approaches - using 10% labeled target data and pseudo labeled data. We find that using only 10% labeled target domain data provides a substantial performance for cross-person and cross-position heterogeneity. On the other hand, generating and leveraging pseudo labeled data provides

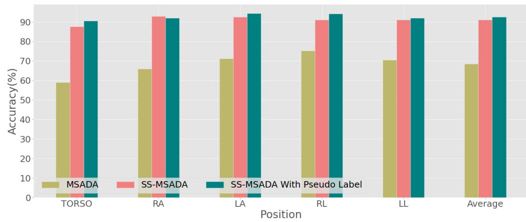

Fig. 12: Performance comparison of cross-person transfer for **DSADS** dataset among MSADA [6], *SS-MSADA* and *SS-MSADA* with the pseudo label usage.

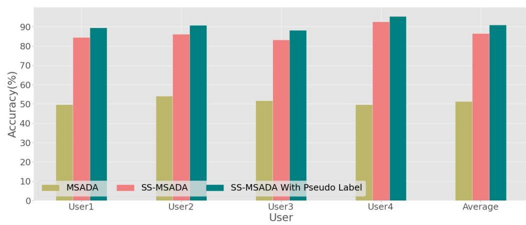

Fig. 13: Performance comparison of cross-position transfer for **DSADS** dataset with among MSADA [6], *SS-MSADA* and *SS-MSADA* with the pseudo label usage.

a marginal performance gain compared to the unsupervised domain adaptation counterpart. We further report some key insight aspects of pseudo labeled data usage in multi-source domain adaptation scenarios, which requires much more investigation in the future.

ACKNOWLEDGMENTS

This work has been partially supported by NSF CAREER Award 1750936 and U.S. Army Grant W911NF2120076.

REFERENCES

[1] M. A. A. H. Khan, N. Roy, and A. Misra, "Scaling human activity recognition via deep learning-based domain adaptation," in *2018 IEEE International Conference on Pervasive Computing and Communications (PerCom)*, pp. 1–9, IEEE, 2018.

[2] J. Wang, V. W. Zheng, Y. Chen, and M. Huang, "Deep transfer learning for cross-domain activity recognition," in *proceedings of the 3rd International Conference on Crowd Science and Engineering*, pp. 1–8, 2018.

[3] A. Z. M. Faridee, M. A. A. H. Khan, N. Pathak, and N. Roy, "Augtoact: scaling complex human activity recognition with few labels," in *Proceedings of the 16th EAI International Conference on Mobile and Ubiquitous Systems: Computing, Networking and Services*, pp. 162–171, 2019.

[4] A. Akbari and R. Jafari, "Transferring activity recognition models for new wearable sensors with deep generative domain adaptation," in *Proceedings of the 18th International Conference on Information Processing in Sensor Networks*, pp. 85–96, ACM, 2019.

[5] J. V. Jeyakumar, L. Lai, N. Suda, and M. Srivastava, "Sensehar: a robust virtual activity sensor for smartphones and wearables," in *Proceedings of the 17th Conference on Embedded Networked Sensor Systems*, pp. 15–28, 2019.

[6] A. Chakma, A. Z. M. Faridee, M. A. A. H. Khan, and N. Roy, "Activity recognition in wearables using adversarial multi-source domain adaptation," *Smart Health*, vol. 19, p. 100174, 2021.

[7] G. Wilson, J. R. Doppa, and D. J. Cook, "Multi-source deep domain adaptation with weak supervision for time-series sensor data," in *Proceedings of the 26th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*, pp. 1768–1778, 2020.

[8] J. E. Van Engelen and H. H. Hoos, "A survey on semi-supervised learning," *Machine Learning*, vol. 109, no. 2, pp. 373–440, 2020.

[9] H. Ma, Z. Zhang, W. Li, and S. Lu, "Unsupervised human activity representation learning with multi-task deep clustering," *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies*, vol. 5, no. 1, pp. 1–25, 2021.

[10] C. I. Tang, I. Perez-Pozuelo, D. Spathis, S. Brage, N. Wareham, and C. Mascolo, "Selfhar: Improving human activity recognition through self-training with unlabeled data," *arXiv preprint arXiv:2102.06073*, 2021.

[11] J. Wang, Y. Chen, L. Hu, X. Peng, and S. Y. Philip, "Stratified transfer learning for cross-domain activity recognition," in *2018 IEEE International Conference on Pervasive Computing and Communications (PerCom)*, pp. 1–10, IEEE, 2018.

[12] E. Tzeng, J. Hoffman, K. Saenko, and T. Darrell, "Adversarial discriminative domain adaptation," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 7167–7176, 2017.

[13] Y. Ganin and V. Lempitsky, "Unsupervised domain adaptation by backpropagation," in *International conference on machine learning*, pp. 1180–1189, PMLR, 2015.

[14] Z. Pei, Z. Cao, M. Long, and J. Wang, "Multi-adversarial domain adaptation," in *Thirty-second AAAI conference on artificial intelligence*, 2018.

[15] D. Roggen, A. Calatroni, M. Rossi, T. Holleczek, K. Förster, G. Tröster, P. Lukowicz, D. Bannach, G. Pirkl, A. Ferscha, *et al.*, "Collecting complex activity datasets in highly rich networked sensor environments," in *2010 Seventh international conference on networked sensing systems (INSS)*, pp. 233–240, IEEE, 2010.

[16] B. Barshan and M. C. Yüksek, "Recognizing daily and sports activities in two open source machine learning environments using body-worn sensor units," *The Computer Journal*, vol. 57, no. 11, pp. 1649–1667, 2014.

[17] A. Reiss and D. Stricker, "Introducing a new benchmarked dataset for activity monitoring," in *2012 16th International Symposium on Wearable Computers*, pp. 108–109, IEEE, 2012.

[18] A. Paszke, S. Gross, S. Chintala, G. Chanan, E. Yang, Z. DeVito, Z. Lin, A. Desmaison, L. Antiga, and A. Lerer, "Automatic differentiation in pytorch," 2017.

[19] F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, J. Vanderplas, A. Passos, D. Cournapeau, M. Brucher, M. Perrot, and E. Duchesnay, "Scikit-learn: Machine learning in Python," *Journal of Machine Learning Research*, vol. 12, pp. 2825–2830, 2011.

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

APPENDIX

We discuss various experimental aspects such as the dataset information, implementation details, and experimental settings in this section. In addition, we also briefly discuss several aspects and guideline for pseudo label usage in domain adaptation process.

A. Datasets

We validate *SS-MSADA* performance with OPPORTUNITY Activity Recognition Dataset [15], Daily and Sports Activities Data Set (DSADS) [16] and PAMAP2 Physical Activity Monitoring Datase [17]. OPPORTUNITY and PAMAP2 datasets contain daily living activities, whereas the DSADS contains a combination of sports activities and daily living activities. DSADS induced a substantial inter-user variation as the participant naturally performed the activities. We consider 4, 10, and 11 activities from the OPPORTUNITY, DSADS, and PAMAP2 datasets, respectively. The detailed dataset description is provided in Table VII and the considered activities are listed in Table VIII.

TABLE VII: Datasets Overview

Factor	Opportunity	PAMAP2	DSADS
Sensors	Acc, Gyr, Mag	Acc, Gyr, Mag	Acc, Gyr, Mag
Positions	BACK, RUA, RLA, LUA, LLA, 2 sensors on shoes	DA, Chest, DL	TORSO, LA, RA, LL, RL
Sampling Frequency (Hz)	32	100	25
Dataset Size	2551	3850505	9120
User	4	9	8

TABLE VIII: Activity List

Dataset	Activities
OPP	Standing, Walking, Sitting, Lying
PAMAP2	Lying, Sitting, Standing, Walking, Running, Cycling, Ascending, Descending, Vacuum, Ironing, Rope jumping
DSADS	Standing, Lying-back, Ascending, Walking-parking-lot, Treadmill-running, Stepper-exercise, Cross-trainer-exercise, Rowing, Jumping, Playing-basketball

B. Implementation Details

a) Runtime Environment

We conduct the experiments on a Linux Server (Ubuntu 18.04) running on Intel Core i7-6850K CPU and 64GB DDR4 RAM, with 4 Nvidia 1080Ti Graphics cards with 44GB VRAM. We implement all the deep learning-related algorithms using an open-source deep-learning library `PyTorch` [18], and for the dataset preprocessing, we use Python and Python-based library `Sci-kit Learn` [19].

b) Dataset Preprocessing

In the initial preprocessing of each dataset, we extract the body position-wise accelerometer data from each participant, followed by standardization. We remove the `NaN` (Not a Number) entries in the extraction process and further split the position-wise extracted dataset into training, validation, and testing set in the ratio of 70-20-10% in such a way that each activity contributes to the mentioned ratio. Further, we consider a window of 128 samples from the 3-axis accelerometer data with a 90% overlap with the consecutive windows for the deep network training.

c) Performance Metrics

In the experiments, we compute the accuracy, micro F1-score, precision and recall using `scikit-learn`'s metrics library [19]. We observe the same results for all these metrics for any particular experiment and, therefore, for the sake of simplicity, report the activity F1 score, which is also widely used in existing domain adaptation literature [1], [3], [11].

d) Baselines

As *SS-MSADA* framework is based on adversarial learning mechanism, we compare the framework with two adversarial approaches - RevGrad [13] and MADA [14]. RevGrad [13] introduces a gradient reversal layer that negates the gradients of the domain discriminator, thus conducting a min-max process against the feature extractor. Whereas MADA [14] extends RevGrad [13] by introducing the same number of domain discriminators equivalent to the number of classes. For the sake of equality, we evaluate the baseline approaches assuming the availability of the 10% labeled target data.

e) Baselines Implementations

As the baseline approaches originally consider a single source domain, we slightly adjust the input data processing. We combine the training splits of the corresponding source domains during the network training and test on the target domain. We also adjust the convolutional layers filter size and input-output channels for the sake of IMU data compatibility. Finally, For the sake of comparison, we train the baselines for similar epochs as we train *SS-MSADA* . We refer to Table IX for detailed hyper-parameters of the framework.

C. Experimental Settings

We consider two scenarios that engender the data distribution heterogeneity - (i) cross-person heterogeneity, (ii) body-position heterogeneity.

a) Cross-person Heterogeneity

Under cross-person heterogeneous settings, participants' data from a particular body position is considered as the `domain`. For example, consider the collected ADL activity data from three participant's similar body position (chest). Here, each participant's data is considered as a domain.

b) Cross-position Heterogeneity

Under cross-position heterogeneous settings, different body-positional data collected from a participant is considered the domain. For example, consider the collected ADL activity data from a participant's multiple body positions such as chest, hand, and ankle. Here, each body positional data is considered as a domain.

TABLE IX: Hyper-parameters of *SS-MSADA* Framework

Hyper-parameters	Values
No. of conv. layers	3
No. of filters in conv. layers	32, 64, 128
Conv. filter dimension	1x9, 1x9, 1x9
No. of fully connected layers in Domain Discriminator	2
No. of units in fully connected layers Domain Discriminator	128, 1
Batch size	32
ADAM [20] optimizer parameter - *Beta1*	0.9
ADAM [20] optimizer parameter - *Beta2*	0.99
Learning rate	0.0001
Pre-training epoch	100
SS-MSADA Training epoch (β)	100

D. Discussion

We explore the multi-source domain adaptation under semi-supervised settings, which allows aligning both marginal distributions as well as conditional distribution. Along with our exploration, we attempt to leverage pseudo-labels to achieve a similar performance using partially labeled target domain data. Here, we discuss a few aspects and insights on pseudo label data.

a) Impact of pseudo-label quantity and quality

We use the network-inferred label as the pseudo label. We design two experiments to see the impacts of pseudo labeling - 1) during the domain adaptation process, 2) after domain adaptation completion. In both use-cases, we use the pseudo labeled data similar way as we use the limited available target data. In both of our test cases, we observe that even though the pseudo-labeling accuracy is high, the network does not provide a high boost in the performance because of the 20-30% incorrect labeling. In our proposed semi-supervised approach, using only 10% labeled target data, the network provides a substantial boost in the performance, but when pseudo labeled data are even 70% correct, the network does not provide a similar performance boost.

b) Clustering-based pseudo label generation

We acknowledge that few proposed approaches deploy a clustering mechanism to generate pseudo label data, but the clustering mechanism has not been investigated for IMU-based ADL domain adaptation. Such clustering mechanisms provide additional advantages in maintaining the intrinsic class semantic separability. In our unreported investigation, we attempt to leverage such a clustering mechanism for pseudo label generation and combine it with the adversarial domain adaptation approach. We have a few observations - 1) clustering objective and adversarial objective act adverse to each other, 2) attaining pseudo labels from clusters requires the clustering training to be saturated, which cause additional network training complexities (when to train the clusters, how much loss to back-propagate, additional training time). In our experiments, we observe at best 30-40% clustering accuracy when combined with adversarial mechanism. Combining the clustering and adversarial mechanism is still challenging.

c) Potential guidelines

In summary, pseudo labeled data may play a trade-off with the partially labeled data depending on some conditions. First, the correctness (quality) of the pseudo labeled data. We observe that even 20% incorrect pseudo label can negate the positive impact of the pseudo labeled data. In case of pseudo-labeled data is used at the end of the adaptation process, then the margin of improvement might not be very significant. Which brings the next condition is the timing of pseudo label generation. Therefore, the pseudo label data can play a vital role if the labels are correctly detected early in the adaptation process. We assume that accurately detecting pseudo labels from very early in the adaptation would be challenging as the different data distributions of different domains become even extreme considering multiple source domains.