Cross-Dataset Activity Recognition via Adaptive Spatial-Temporal Transfer Learning

XIN QIN, Beijing Key Lab. of Mobile Computing and Pervasive Devices, Inst. of Computing Tech., CAS, University of Chinese Academy of Sciences, Beijing, China

YIQIANG CHEN*, Beijing Key Lab. of Mobile Computing and Pervasive Devices, Inst. of Computing Tech., CAS, University of Chinese Academy of Sciences, Pengcheng Laboratory, China

JINDONG WANG, Microsoft Research Asia, China

CHAOHUI YU, Beijing Key Lab. of Mobile Computing and Pervasive Devices, Inst. of Computing Tech., CAS, University of Chinese Academy of Sciences, China

Human activity recognition (HAR) aims at recognizing activities by training models on the large quantity of sensor data. Since it is time-consuming and expensive to acquire abundant labeled data, transfer learning becomes necessary for HAR by transferring knowledge from existing domains. However, there are two challenges existing in cross-dataset activity recognition. The first challenge is source domain selection. Given a target task and several available source domains, it is difficult to determine how to select the most similar source domain to the target domain such that negative transfer can be avoided. The second one is accurately activity transfer. After source domain selection, how to achieve accurate knowledge transfer between the selected source and the target domain remains another challenge. In this paper, we propose an Adaptive Spatial-Temporal Transfer Learning (ASTTL) approach to tackle both of the above two challenges in cross-dataset HAR. ASTTL learns the spatial features in transfer learning by adaptively evaluating the relative importance between the marginal and conditional probability distributions. Besides, it captures the temporal features via incremental manifold learning. Therefore, ASTTL can learn the adaptive spatial-temporal features for cross-dataset HAR and can be used for both source domain selection and accurate activity transfer. We evaluate the performance of ASTTL through extensive experiments on 4 public HAR datasets, which demonstrates its effectiveness. Furthermore, based on ASTTL, we design and implement an adaptive cross-dataset HAR system called Client-Cloud Collaborative Adaptive Activity Recognition System (3C2ARS) to perform HAR in the real environment. By collecting activities in the smartphone and transferring knowledge in the cloud server, ASTTL can significantly improve the performance of source domain selection and accurate activity transfer.

CCS Concepts: • **Human-centered computing → Ubiquitous computing**; • **Computing methodologies → Transfer learning**.

Additional Key Words and Phrases: Human Activity Recognition, Transfer Learning, Domain Adaptation, Cross-Dataset Recognition

---

*Yiqiang Chen is the corresponding author.

Authors' addresses: Xin Qin, Beijing Key Lab. of Mobile Computing and Pervasive Devices, Inst. of Computing Tech., CAS, University of Chinese Academy of Sciences, Beijing, China, qinxin18b@ict.ac.cn; Yiqiang Chen, Beijing Key Lab. of Mobile Computing and Pervasive Devices, Inst. of Computing Tech., CAS, University of Chinese Academy of Sciences, Pengcheng Laboratory, China, yqchen@ict.ac.cn; Jindong Wang, Microsoft Research Asia, Beijing, China, jindong.wang@microsoft.com; Chaohui Yu, Beijing Key Lab. of Mobile Computing and Pervasive Devices, Inst. of Computing Tech., CAS, University of Chinese Academy of Sciences, Beijing, China, yuchaohui17s@ict.ac.cn.

148

Fig. 1. Different sensor reading distributions on different subjects

1 INTRODUCTION

Human activity recognition (HAR) is an important topic in ubiquitous computing. It aims to accurately recognize people's daily activities and behaviors through the collected physiological signals provided by devices. Usually, different human activities tend to have different sensor readings, which can be used to recognize the activities. HAR has been widely applied in various fields to optimize the interaction between human and computer such as gesture recognition [40], fatigue detection [30] and smart home sensing [50].

Accurate HAR relies on the training of machine learning models using massive well-labeled activity data. However, it remains expensive and time-consuming to acquire sufficient labeled data in real applications. For instance, the activities of the elderly are extremely important to help analyze and detect several cognitive diseases such as Parkinson's disease [9, 10]. If we want to recognize the activities for the elder person in Figure 1, it is often difficult to collect sufficient labeled activity data due to health issues and inconvenient situations. Under this circumstance, one may consider reusing the data or models from existing datasets (e.g. the child and the adult in Figure 1) to help learn the model for the elder person. Unfortunately, there is often a large distribution gap in sensor readings from different people as shown in Figure 1, even if they are performing the same activity. Therefore, directly reusing the model trained on one person for a new target will produce unsatisfactory results.

Transfer learning [39, 55, 60] has been a hot research topic in recent years, which has the potential to transfer knowledge from a well-labeled domain (i.e. *source domain*) to an unlabeled domain (i.e. *target domain*). Since the distributions of the two domains are different, the key to successful transfer learning is to find and exploit the similarity between domains. A lot of work [15, 29, 44] on transfer learning based HAR has achieved great success, which can mainly be divided into two categories. The first category is instance weighting [1, 8, 57, 58]. The other category is feature transformation such as Manifold Embedded Distribution Alignment (MEDA) [56], Geodesic Flow Kernel (GFK) [18], and Transfer Component Analysis (TCA) [38]. The essence of these methods is to find a transformation that minimizes the distance between the source domain and the target domain.

By adopting the concepts and methods in transfer learning, we may be able to perform cross-dataset HAR. Unfortunately, most of the above work only focuses on the problem of *how to transfer*, while they fail to consider the problem of *what to transfer*. In our cross-dataset HAR problem, given a target domain and several source domains available, it is important to select the most similar source domain to the target before performing transfer learning. Although several methods [13, 57] tried the source domain selection, they only work in deep neural networks which are computationally constrained in ubiquitous computing. Moreover, as for the problem of how to transfer, existing work fails to consider the temporal relationship between activity features, which is likely to produce negative transfer effect (i.e., the performance of transfer learning is worse than no transfer methods) [39]. It is extremely challenging to design algorithms for both source domain selection and accurate activity transfer in cross-dataset HAR. Firstly, the labeled samples in the target domain are extremely few and a more common situation is that none sample is annotated, making it unfeasible to perform transfer learning. Secondly, it is difficult to know which domain pairs are suitable for transfer, i.e. to perform source domain selection so as to enhance the learning performance on the target domain. Thirdly, after acquiring the most similar source domain, it is still challenging to perform transfer learning between domains.

In this paper, we propose a novel *Adaptive Spatial-Temporal Transfer Learning (ASTTL)* approach to tackle the above challenges. ASTTL is able to conduct both source domain selection and accurate activity transfer for the cross-dataset HAR problem. The key is to learn spatially-adaptive and temporally-adaptive features. Firstly, ASTTL learns temporally-adaptive features in the manifold space [18] by extending the Geodesic Flow Kernel (GFK) with the Markov property [28]. Secondly, ASTTL is capable of learning spatially-adaptive representations by dynamically evaluating the different effects of marginal and conditional probability distributions using the Maximum Mean Discrepancy (MMD) [20, 46]. Based on the spatial-temporal adaptive learning, ASTTL could sequentially perform source domain selection and transfer learning for HAR. (Note that we only consider selecting the top-1 best source domain in this paper for computational efficiency, while there are a lot of research on multiple-source transfer learning). We conduct experiments on 4 public activity recognition datasets (UCI DSADS, UCI-HAR, USC-HAD, and PAMAP2) to evaluate the performance of ASTTL. The results demonstrate that ASTTL is able to achieve accurate source domain selection and activity transfer learning compared to other methods. Furthermore, we design and implement an adaptive cross-dataset HAR system to achieve accurate human activity recognition in real environments.

The main contributions of this paper are summarized as follows:

- We propose ASTTL to perform source domain selection and activity transfer in cross-dataset HAR problems. ASTTL learns both spatial and temporal adaptive features for transfer learning.
- The spatially-adaptive features can be learned by dynamically evaluate the importance of marginal and conditional distributions in transfer learning, while the temporally-adaptive features are learned by adding temporal factors to the geodesic flow kernel. Therefore, ASTTL can avoid negative transfer.
- Extensive experiments on 4 public activity recognition datasets demonstrate the superiority of ASTTL in both source domain selection and activity transfer.
- We design an adaptive cross-dataset HAR system embedding the ASTTL method. The system can achieve accurate human activity recognition in real situations.

The remainder of this paper is organized as follows. Section 2 provides a brief review of the related work. Section 3 introduces the proposed ASTTL and its main process. In Section 4, experimental evaluation and analysis on four public HAR datasets are presented. Section 5 introduces the proposed Client-Cloud Collaborative Adaptive Activity Recognition System (3C2ARS) and provide the evaluation of its performance in real-life human activity recognition. Finally, we present the conclusions and future work in Section 6.

2 RELATED WORK

2.1 Human Activity Recognition

Human activity recognition is a promising research issue since it learns high-level knowledge of human behaviors through the raw sensor data collected in the ubiquitous computing environment, which can promote the development of medical, smart home and safe driving fields, etc.

Generally, in existing methods, the procedure of human activity recognition can be regarded as a standard time series classification problem by utilizing machine learning methods [9]. This procedure subsequently solves data collection and preprocessing, feature extraction, machine learning model construction, and classification in the end. Some existing methods achieve HAR by exploiting the deep understanding of signal data [61], or focusing on machine learning methods such as traditional methods of active learning [25], similarity graph [43] and deep learning methods [14, 22, 24, 31, 53]. However, most of these methods assume the training and testing data are in the same distribution, which results in relatively weak generalization when faced with novel data.

2.2 Transfer Learning

Transfer learning has been successfully applied in various fields, such as visual image classification [13], natural language processing [17], sentiment classification [45], intelligent planning [65], etc. According to the whole transfer process, and the given situation, transfer learning can be divided into two categories: 1) From the perspective of source, determine what component to transfer, i.e. source selection; 2) From the perspective of transfer methods to achieve knowledge transfer, determine how to transfer, i.e. transfer method.

2.2.1 Source Domain Selection. As for what to transfer, problem situation set as that given a target task and several acquired source knowledge, select which domain or component as the source to transfer can achieve the best performance. Methods mainly can be divided into two categories, domain-based and instance-based method. Domain-based methods aim to select the most appropriate domain from the high level of the domain. Much research has been done with this kind of methods. Schultz L R *et al.* proposed a linear combination of distance metrics to select the most similar source domain [45]. Afridi M J *et al.* extended to deep learning field to select a pre-trained Convolutional Neural Network (CNN) by automatically ranking source CNNs before utilizing them [1]. Instance-based methods weight the source instances to determine the probability of each instance to transfer. Based on this kind of methods, extension method like double-selection process [32] was proposed to deal with a wider range of transfer learning scenarios.

Noticing that distance metrics are usually used as the estimation for similarity. Distance metrics including Kullback-Leibler divergence (KL) [64], Jensen-Shannon divergence (JS) [33, 42], $\mathcal{A}$-distance [6] and CORrelation Alignment loss (CORAL-loss) [48] etc. are widely used techniques. One of the most promising distance metrics in transfer learning is Maximum Mean Discrepancy (MMD). Gretton *et al.* [21] developed the spirit of MMD to measure possibility distributions divergence between the mean embeddings of two distributions based on the corresponding RKHS distance. The MMD is a theoretically useful technique that can be formally shown to always detect a difference if one occurs [7], thus it is taken as the basic distance metric in our proposed global-local distance metric. Although there are several distance metrics, few of them comprehensive consider the domain divergence and class divergence, which leads to the neglect of global or local properties. Besides, most of these methods only work in deep neural networks. Our ASTTL is a traditional transfer learning method, which considers both the global and local properties while performing transfer learning.

2.2.2 Transfer Learning Method. How to transfer is the most important topic of transfer learning according to literature research. Much work has been done to explore the best transfer method according to the task. The goal of these methods is to minimize the distance between the source and target domains by mapping features into the same or closer subspace, or exploiting the correlations between features [6], or building the middle domain to

narrow the source and target gap [49]. The most commonly used method is the feature-based method. Maximum mean discrepancy embedding (MMDE) [37] learns features in the Reproducing Kernel Hilbert Space (RKHS), while it is computationally prohibitive for the requirement of solving a semidefinite programming (SDP). Joint distribution adaptation (JDA) [35] proposed by Long *et al.* is based on minimizing joint distribution between domains, however, it limits by focusing on marginal distribution. Geodesic flow kernel (GFK) [18] exploits the low-dimensional structure to integrate the domains. However, GFK only considers the spatial feature space while neglects the temporal property. Our method is based on GFK, making a modification by combining the spatial and temporal property.

2.3 Human Activity Recognition with Transfer Learning

There is much prior work focusing on HAR with transfer learning. Rey V F *et al.* [43] discussed the case that the new domain just happened to contain the old one. Hu D H *et al.* [26] developed a bridge between the activities in two domains by learning a similarity function via Web search for HAR, while it only considered the distance between classes and ignored the global distance. Huynh T *et al.* [27] focused on unsupervised zero-shot learning, although it didn't require labels, the output was a set of an unnamed cluster which cannot be used for activity recognition. Liang *et al.* [31] achieved domain adaptation between online videos and real-world collections with a deep framework, while this was relatively computing consuming with a deep network. Chen *et al.* [11] trained a classifier on the source and classified the target to make a soft tag, then iterative trained model with high confident samples being added into the training set. However, this required a large computational cost. While our work can reduce the computing consumption by a preprocess of selecting the closest source and can guarantee the classification accuracy by spatial-temporal feature learning.

3 ADAPTIVE SPATIAL-TEMPORAL TRANSFER LEARNING

In this section, we present the proposed Adaptive Spatial-Temporal Transfer Learning (ASTTL) approach in detail. We will start with the problem definition and notations. Then, we introduce the main idea of our method. Subsequently, we show how to learn spatially-adaptive features and the learning process of temporally-adaptive features. Finally, we describe how to perform source domain selection and activity transfer using the proposed ASTTL approach.

3.1 Problem Definition and Notations

In the cross-dataset HAR problem, we are given an unlabeled dataset as the target domain $\mathcal{D}_t = \left\{\mathbf{x}_{t_j}\right\}_{j=1}^{m}$. In order to accurately recognize the activities in $\mathcal{D}_t$, we usually have K fully-labeled datasets as the source domain: the k-th source domain can be denoted as $\mathcal{D}_s^k = \left\{\mathbf{x}_{si}^k, y_{si}^k\right\}_{i=1}^{n_k}$. We often assume that both the source and the target domains involve the same kinds of sensors and activities, i.e. the feature space and the label space are the same: $\mathcal{X}_s = \mathcal{X}_t$ and $\mathcal{Y}_s = \mathcal{Y}_t$. In real applications, different datasets tend to have *different* probability distributions, i.e. $P(\mathbf{x}_s) \neq P(\mathbf{x}_t)$ [56, 57]. In order to accurately recognize the activities in $\mathcal{D}_t$, we need to firstly select the most similar source domain $\mathcal{D}_s^{\hat{k}}$ from the available source domains which has the minimum distribution distance to the target domain. Then, we design novel algorithms to perform knowledge transfer such that the unlabeled samples in $\mathcal{D}_t$ can be annotated with good accuracy. Note that in our work, we only select the $top - 1$ source domain for computational efficiency, while it is easier to select the $top - k$ source domains and apply multiple-source transfer learning according to existing work, which is not our focus.

3.2 Main Idea

The problem of cross-dataset HAR is a specific transfer learning related application. According to existing transfer learning works [18, 38, 39, 57], the key to successfully transfer is to learn good and transferable feature
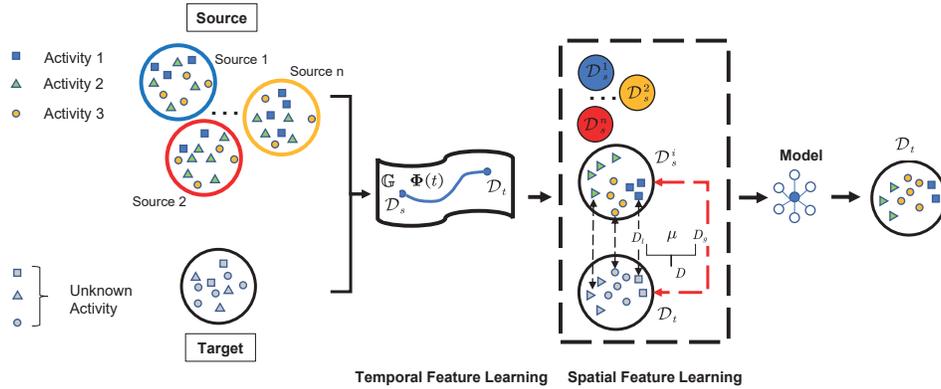
Fig. 2. The learning process of ASTTL for cross-dataset HAR

representations such that the distribution discrepancy between domains can be reduced. To this end, existing work either proposes to perform distribution adaptation, which can explicitly reduce the distribution divergence [34, 51, 52, 61], or proposes to learn transferable features in the manifold space in which manifold distance can be adopted to minimize the distribution distance [18, 19, 48]. Recent work focuses on deep transfer learning [57] which is beyond the scope of this research.

Despite the great success achieved by existing work, there are still two critical challenges regarding the cross-dataset HAR problem. Firstly, few of them pay attention to source domain selection, which is extremely important to the knowledge transfer. If we perform transfer learning between two highly-dissimilar domains, there will be negative transfer [39]. Secondly, existing work fails to consider the temporal relationship during the feature learning process. Currently, there is no existing work that can handle the two challenges together.

In this work, we propose Adaptive Spatial-Temporal Transfer Learning (ASTTL) to resolve the above two challenges. ASTTL learns both spatial and temporal adaptive features for transfer learning problems. Specifically, the spatial features are determined by the marginal probability distributions and conditional probability distributions between domains. Thus, we adaptively and quantitatively evaluate their relative importance by extending existing work [56] with computational efficiency in Reproducing Kernel Hilbert Space (RKHS). As for the temporal feature learning, we take inspirations from Markov property to temporally learn manifold features in the Grassmann manifold [18] through extending geodesic flow kernel. Finally, by integrating the spatial and temporal adaptive feature learning, ASTTL can perform source domain selection and activity transfer for the cross-dataset HAR problem.

Technically, the learning objective of ASTTL (i.e. the prediction function f) can be formalized as:

$$f = \arg\min \sum_{i=1}^{m} \ell(g(\mathbf{x}_i), y_i) + \lambda Spatial(\mathcal{D}_s, \mathcal{D}_t) + \beta ||f||_{\mathcal{H}}^2, \tag{1}$$

where $\ell(\cdot, \cdot)$ denotes the loss function on the source domain, function $g(\cdot)$ denotes the temporal feature learning, and $Spatial(\cdot, \cdot)$ denotes the spatial feature learning which explicitly reduces the distribution divergence. λ, β are the trade-off parameters and $|| \cdot ||_{\mathcal{H}}^2$ denotes the norm in Hilbert space $\mathcal{H}$.

The learning process of ASTTL is illustrated in Figure 2. In next sections, we will introduce ASTTL in detail.

3.3 Spatially-Adaptive Feature Learning

The Spatially-adaptive transfer learning aims at capturing the spatial relationship between the source and the target domains (i.e. the $Spatial(\cdot, \cdot)$ function in Equation (1)). Inspired by the recent work [56], the distribution

divergence in transfer learning is caused by marginal ($P(\mathbf{x})$) and conditional ($Q(y|\mathbf{x})$) distributions. Therefore, it is important to consider the relative importance of these two distributions. In this paper, we treat the divergence of marginal distributions as the *global* distance between domains, while the divergence of conditional distributions as the *local* distance.

According to [56], the global and local distances can be dynamically weighted as:

$$Spatial(\mathcal{D}_s, \mathcal{D}_t) = D(\mathcal{D}_s, \mathcal{D}_t) = (1 - \mu) D_g(\mathcal{D}_s, \mathcal{D}_t) + \mu \sum_{c=1}^{C} D_l\left(\mathcal{D}_s^c, \mathcal{D}_t^c\right), \tag{2}$$

with

$$\mu = 1 - \frac{D_g}{D_g + \sum_{c=1}^{C} D_l}, \tag{3}$$

where μ is the adaptive factor which ranges from 0 to 1, and $c \in \{1, \cdots, C\}$ is the class indicator. $D_g(\mathcal{D}_s, \mathcal{D}_t)$ is the global distance between the source and the target domains, and $D_l\left(\mathcal{D}_s^c, \mathcal{D}_t^c\right)$ denotes the local distance between classes in the source and target domains, correspondingly.

In previous work [56], the adaptive factor μ is calculate by the proxy $\mathcal{A}$-distance [5]. The proxy $\mathcal{A}$-distance is computed by building a binary classifier and calculating the error of classifying whether the samples are from the source or the target domain. Therefore, we need to build one binary classifier to learn the global distance and C classifiers to learn the local distance. Moreover, it requires several rounds of iterations to learn the value of the adaptive factor. This process is computationally intensive and it is time-consuming to fine-tune parameters of these classifiers in real applications.

In this paper, we propose ASTTL to learn the spatially-adaptive features by using the distribution divergence measurement to calculate the adaptive factor. Compared to existing work [56] which is overparameterized since it needs to learn multiple parameters through training classifier iteratively, our spatially-adaptive feature learning does not need to build classifiers. Instead, we use the distribution information to help learn the relationships between the marginal and conditional distributions. Since the target data is unlabeled, we use the soft labels for the target by following existing work [34, 56]: train a base classifier on the source and make a prediction on the target in the initialization stage.

Specially, we use Maximum Mean Discrepancy (MMD) [20] as the basic distance metric. MMD is a distance metric which embeds distributions in Reproducing Kernel Hilbert Space (RKHS), using the difference between these embeddings as the test statistic. The global-local distance between two domains $\mathcal{D}_s$ and $\mathcal{D}_t$ can be respectively formulated as

$$D_g(\mathcal{D}_s, \mathcal{D}_t) = \left\| \frac{1}{n_s} \sum_{\mathbf{x}_{s_i} \in \mathcal{D}_s} \phi(\mathbf{x}_{s_i}) - \frac{1}{n_t} \sum_{\mathbf{x}_{t_j} \in \mathcal{D}_t} \phi(\mathbf{x}_{t_j}) \right\|_{\mathcal{H}}^2, \tag{4}$$

$$D_l(\mathcal{D}_s, \mathcal{D}_t) = \left\| \frac{1}{n_s^{(c)}} \sum_{\mathbf{x}_{s_i} \in \mathcal{D}_s^{(c)}} \phi(\mathbf{x}_{s_i}) - \frac{1}{n_t^{(c)}} \sum_{\mathbf{x}_{t_j} \in \mathcal{D}_t^{(c)}} \phi(\mathbf{x}_{t_j}) \right\|_{\mathcal{H}}^2. \tag{5}$$

Thus, the distribution distance calculated by Equation (2) can be denoted as:

$$D(\mathcal{D}_s, \mathcal{D}_t) =$$

$$(1 - \mu) \left\| \frac{1}{n_s} \sum_{\mathbf{x}_{s_i} \in \mathcal{D}_s} \phi(\mathbf{x}_{s_i}) - \frac{1}{n_t} \sum_{\mathbf{x}_{t_j} \in \mathcal{D}_t} \phi(\mathbf{x}_{t_j}) \right\|_{\mathcal{H}}^2 + \mu \sum_{c=1}^{C} \left\| \frac{1}{n_s^{(c)}} \sum_{\mathbf{x}_{s_i} \in \mathcal{D}_s^{(c)}} \phi(\mathbf{x}_{s_i}) - \frac{1}{n_t^{(c)}} \sum_{\mathbf{x}_{t_j} \in \mathcal{D}_t^{(c)}} \phi(\mathbf{x}_{t_j}) \right\|_{\mathcal{H}}^2, \tag{6}$$

where $\phi(\cdot)$ is the feature map which maps the original instances into the RKHS $\mathcal{H}$. $n_s = |\mathcal{D}_s|$, $n_t = |\mathcal{D}_t|$ are number of instances in the source and the target domains, respectively. Similarly, $n_s^{(c)} = \left| \mathcal{D}_s^{(c)} \right|$ and $n_t^{(c)} = \left| \mathcal{D}_t^{(c)} \right|$ are numbers of instances in source class and target class.

The relative importance of these two distance is weighted by μ. When $\mu \to 0$, the distribution between the source and target is relatively large, the global distance is dominated. When $\mu \to 0.5$, it declares that both global and local distances are treated equally. When the value of μ further increases close to 1, the distribution between domains is relatively small, thus the local distance plays a more important role in the global-local distance. In real applications, we can find a kernel $\mathcal{K}(\cdot, \cdot)$ to map the inputs to the RKHS and then learn the spatially-adaptive features.

3.4 Temporally-Adaptive Feature Learning

In this part, we articulate the learning process of the temporally-adaptive transfer feature learning. The goal here is to learn the feature transformation function $g(\cdot)$ in Equation (1). Inspired by the work of Geodesic Flow Kernel (GFK) [18], ASTTL focuses on the temporal effect during feature learning by extending GFK.

In general, GFK learns the transferable features gradually in the Grassmann manifold $\mathbb{G}$ [23] to preserve the geometrical structures of the domain. The Grassmann manifold $\mathbb{G}$ can facilitate domain adaptation [23] and classifier learning [3]. Geodesic Flow Kernel (GFK) is one of the most efficient and widely used methods among the various approaches [3, 19] to learn $g(\cdot)$, thus transform the raw data into $\mathbb{G}$. The Grassmann manifold $\mathbb{G}$ is a collection of all d-dimensional subspaces, among which we denote the source and target subspace after performing PCA are $\mathbf{P}_s$ and $\mathbf{P}_t$, respectively. Find a geodesic flow $\Phi(t)$ to continuous transform $\mathbf{P}_s = \Phi(0)$ to $\mathbf{P}_t = \Phi(1)$, the transformed feature can be represented as $\mathbf{z} = g(\mathbf{x}) = \Phi(t)^T \mathbf{x}$. According to [18], using kernel trick, the inner product of two transformed vector $\mathbf{z}_i$ and $\mathbf{z}_j$ is:

$$\langle \mathbf{z}_i, \mathbf{z}_j \rangle = \int_0^1 \left(\Phi(t)^T \mathbf{x}_i \right)^T \left(\Phi(t)^T \mathbf{x}_j \right) dt = \mathbf{x}_i^T \mathbf{G} \mathbf{x}_j. \tag{7}$$

Note that the original GFK only considers to gradually transfer the subspace of the source domain to the subspace of the target domain. In the beginning, the features are in the subspace of the source domain. Along with this GFK process, the features are more likely to be in the target subspace. Therefore, the temporal factor is extremely important during this process. However, GFK ignores this fact by treating all temporal factors as equally important ($t = 1$ through the integral). This may result in less successful transfer feature learning.

In this work, we extend the idea of GFK by adding *temporal* factors to the transfer feature learning. Different from the original GFK, we believe that time should be an important factor in feature transformation. When GFK gradually transforms the source domain into the subspace of the target domain, the latter transformation has better influence. For time $t_1 < t_2$, the transformation at t_2 should also have larger influence to the final result than t_1. We call this *temporally-adaptive* feature learning for our problem. This can also be explained from the perspective of the non-reversal time property in Markov models: in time series, the closer to the target, the more important the feature is.

The geodesic flow can be parameterized as

$$\Phi(t) = \mathbf{P}_s \mathbf{U}_1 \Gamma(t) - \mathbf{R}_s \mathbf{U}_2 \Sigma(t) = \left[\begin{array}{cc} \mathbf{P}_s & \mathbf{R}_s \end{array} \right] \left[\begin{array}{cc} \mathbf{U}_1 & 0 \\ 0 & \mathbf{U}_2 \end{array} \right] \left[\begin{array}{c} \Gamma(t) \\ -\Sigma(t) \end{array} \right], \tag{8}$$

where $\mathbf{R}_s \in \mathbb{R}^{D \times (D-d)}$ presents the orthogonal complements to $\mathbf{P}_s$. $\mathbf{U}_1 \in \mathbb{R}^{d \times d}$ and $\mathbf{U}_2 \in \mathbb{R}^{(D-d) \times d}$ are two orthonormal matrices that can be computed by

$$\mathbf{P}_s^{\mathrm{T}} \mathbf{P}_t = \mathbf{U}_1 \Gamma \mathbf{V}^{\mathrm{T}}, \mathbf{R}_s^{\mathrm{T}} \mathbf{P}_t = -\mathbf{U}_2 \Sigma \mathbf{V}^{\mathrm{T}}. \tag{9}$$

According to GFK [18], the geodesic flow kernel $\mathbf{G}$ can be calculated by

$$\mathbf{G} = \begin{bmatrix} \mathbf{P}_s\mathbf{U}_1 & \mathbf{R}_s\mathbf{U}_2 \end{bmatrix} \begin{bmatrix} \mathbf{\Lambda}_1 & \mathbf{\Lambda}_2 \\ \mathbf{\Lambda}_2 & \mathbf{\Lambda}_3 \end{bmatrix} \begin{bmatrix} \mathbf{U}_1^\mathrm{T}\mathbf{P}_s^\mathrm{T} \\ \mathbf{U}_2^\mathrm{T}\mathbf{R}_s^\mathrm{T} \end{bmatrix}, \tag{10}$$

where $\mathbf{\Lambda}_1, \mathbf{\Lambda}_2, \mathbf{\Lambda}_3$ are three diagonal matrices with elements

$$\lambda_{1i} = \int_0^1 \cos^2(t\theta_i)dt = 1 + \frac{\sin(2\theta_i)}{2\theta_i},$$
$$\lambda_{2i} = -\int_0^1 \cos(t\theta_i)\sin(t\theta_i)dt = \frac{\cos(2\theta_i) - 1}{2\theta_i}, \tag{11}$$
$$\lambda_{3i} = \int_0^1 \sin^2(t\theta_i)dt = 1 - \frac{\sin(2\theta_i)}{2\theta_i}.$$

In our temporally-adaptive feature learning, Equation (7) should be extended with the temporal factor. Therefore, the elements can be calculated by

$$\lambda_{1i} = \int_0^1 t\cos^2(t\theta_i)dt = \frac{1}{4} - \frac{1}{4\theta_i^2}\sin^2\theta_i + \frac{1}{4\theta_i}\sin 2\theta_i, \tag{12}$$

$$\lambda_{2i} = -\int_0^1 t\cos(t\theta_i)\sin(t\theta_i)dt = \frac{\cos(2\theta_i)}{4\theta_i} - \frac{\sin(2\theta_i)}{8\theta_i^2}, \tag{13}$$

$$\lambda_{3i} = \int_0^1 t\sin^2(t\theta_i)dt = \frac{1}{4} + \frac{1}{4\theta_i^2}\sin^2\theta_i - \frac{1}{4\theta_i}\sin 2\theta_i, \tag{14}$$

where Equation (12) to (14) can learn more target-related features than the original GFK. After calculating the geodesic flow kernel $\mathbf{G}$, we can get the transformed features by $\mathbf{z} = g(\mathbf{x}) = \sqrt{\mathbf{G}}\mathbf{x}$.

Finally, we can summarize all equations into Equation (1) to reformulate and compute the classifier f.

3.5 ASTTL for Cross-dataset HAR

After learning the spatial and temporal features, ASTTL can be used for transfer learning problems by substituting $Spatial(\cdot, \cdot)$ and $g(\cdot)$ in Equation (1). In this section, we show how to perform source domain selection and activity transfer for cross-dataset HAR problem using the proposed ASTTL approach. Given a labeled activity dataset as the source domain, in order to maximize its utility to enhance the learning performance on the target dataset, the first task is to determine the most similar (with the closest distribution distance) source domain. There are various categories of distance metrics. However, many of them are parametric or require an intermediate density estimate [38]. Another existing problem is that most of the prior work ignores the local property between classes in source and target domain [9].

In this paper, we adopt the proposed ASTTL to learn the distance between domains, which is capable of leveraging the importance of global domain distribution and local class distribution to calculate the domain distance. On the other hand, ASTTL can also facilitate knowledge by combining its spatial and temporal feature learning. We adopt a *greedy* method to select the source domain with the closest distance to the target domain. We calculate the global and local distances respectively according to Equation (4) and (5), and leveraging the importance of these two distances by μ in Equation (3), finally get global-local distance according to Equation (6). Noticing that it is difficult to solve Equation (6) as the existing of determined feature mapping $\phi(\cdot)$, thus we adopt kernel method Radical Basis Function (RBF) to solve this problem.

After the spatially-adaptive and temporally-adaptive feature learning and calculation of the distribution distance, we get the most similar source domain to the target domain and the mapped source and target features, thus we can build a transfer model on them. The overall cross-dataset HAR with ASTTL is described in Algorithm 1.

Algorithm 1 ASTTL for Cross-dataset HAR

Input: K Source datasets $\mathcal{D}_s^1, \cdots, \mathcal{D}_s^K$, and target domain $\mathcal{D}_t$, λ, β
Output: Classifier f.
 1: Initialize a source domain set $S = \{\}$;
 2: **for** $i = 1$ to K **do**
 3:     Learn temporal feature transformation, and get manifold feature $\mathbf{z} = \sqrt{\mathbf{G}}\mathbf{x}_i$;
 4:     Calculate the global distance in Equation (4) and local distance in Equation (5), respectively;
 5:     Calculate the spatially-adaptive factor μ according to Equation (3), and get the global-local distance between each source candidate and the target;
 6:     Add $D_i\left(\mathcal{D}_s^i, \mathcal{D}_t\right)$ into S;
 7: **end for**
 8: Select the source domain index j which has the smallest value in S: $D_j = \min\{S\}$;
 9: Train classifier f with $\mathcal{D}_s^j$ and $\mathcal{D}_t$ and apply prediction on $\mathcal{D}_t$ using Equation (1).
10: **return** Classifier f.

---

4 EXPERIMENTAL EVALUATION

In this section, we evaluate the performance of the proposed ASTTL approach via extensive experiments on public activity recognition datasets.

4.1 Datasets and Preprocessing

We adopt four large public activity datasets. Table 1 summarizes the statistics of the four datasets, while Figure 3 illustrates the sensor locations on the subject in each dataset. In the following, we briefly introduce the basic information of each dataset, and detailed descriptions can be found in their original papers.

UCI daily and sports dataset (DSADS) [4] collects 19 activities through 8 subjects (four males and four females, age from 25 to 28) wearing body-worn sensor units including triaxial accelerometer, triaxial gyroscope, and triaxial magnetometer on 5 body parts. Subjects were asked to perform the activities in their style and were not restricted to how the activities should be performed. In this way, the dataset can reflect natural human activities. UCI Smartphone (HCI-HAR) [2] contains 6 activities collected from 30 subjects and each of the subject wears a smartphone on the waist. 3-axial linear acceleration and 3-axial angular velocity at a constant rate of 50Hz are captured by the smartphone embedded accelerometer and gyroscope. USC Human Activity Dataset (USC-HAD) [62] consists of 12 activities collected from 14 subjects (7 males, 7 females), each wears a motion node at front right hip. This dataset also recorded four external factors of gender, age, height and weight of subjects

Table 1. Statistical information of four public activity recognition datasets

Dataset	Subject	Activity	Sample	Position
DSADS	8	19	1.14M	Tarso, Right Arm, Left Arm, Right Leg, Left Leg
UCI-HAR	30	6	1.31M	Waist
USC-HAD	14	12	2.81M	Front Right Hip
PAMAP	9	18	2.84M	Wrist, Chest, Ankle

as a reference. PAMAP2 dataset (PAMAP) [41] is made up of 9 subjects performing 18 activities with 3 inertial measurement sensor units on 3 body parts, and a heart rate monitor, collecting data in continuous time.

Note that the dataset collection protocols of each dataset are somewhat different, such as activities and sensor location among four datasets. Thus, we need to unify our experimental setup for all datasets. To be more specific, we adopt the same sensor devices and activities categories. And we utilize the common body positions from these datasets (i.e. the dominant hand in Figure 3). We choose data from Right Arm, Waist, Front Right Hip and Right Wrist form these four datasets, respectively. Four kinds of activities in the four datasets, including LYING (L), WALKING (W), WALKING-UPSTAIRS (U) and WALKING-DOWNSTAIRS (D) are to be recognized. These four activities are also the common activities of daily living. As for sensor devices, we utilize accelerometer and gyroscope since they are both used in all datasets. Each sensor provides 3-axial data (x-, y-, and z-axis), thus we combine them by $\alpha = \sqrt{x^2 + y^2 + z^2}$. Then, we extract features according to the [54] and obtain 27 features in both time and frequency domain. Thus, the original data from accelerometer and gyroscope totally have 54 features. In experiments, we perform unsupervised transfer learning on the target domain, the true labels for the target domains are only used for evaluation.

4.2 Evaluation of Source Domain Selection

Given a target dataset, the first and most vital task is to select the most similar source domain. In this section, we evaluate the performance of ASTTL on source domain selection by calculating the domain distances between the target and each source domain. This distance will be utilized as the similarity metric between the two domains. However, it is almost impossible to acquire the ground-truth for the distance since we can never explicitly determine the most similar source domain to the target. Therefore, we use the classification accuracy on the target domain after transfer learning as the ground-truth for domain similarity, i.e., for the same target domain C, if the transfer model built from source domain A acquires better accuracy than source domain B, then source domain A is more similar to the target domain C. The classification accuracy of all tasks after transferring with ASTTL are shown in Figure 4 (a), the horizontal axis represent that four domains are chosen in turn as the target domain. When the source dataset is the target itself, we set the classification accuracy to 100%, which is only the ideal state and has no meaning in real practice. According to this assumption, the domain which should be the most similar source to the target is noted in red arrow. From Figure 4 (a), we can extract the ground-truth for source domain selection by selecting the transfer task that achieves the best accuracy. It should be noted that we train classifiers here for each transfer task just to evaluate the efficient of source selection by using ASTTL distance metric. In the real implementation of ASTTL, it only need to build a classifier for the selected transfer
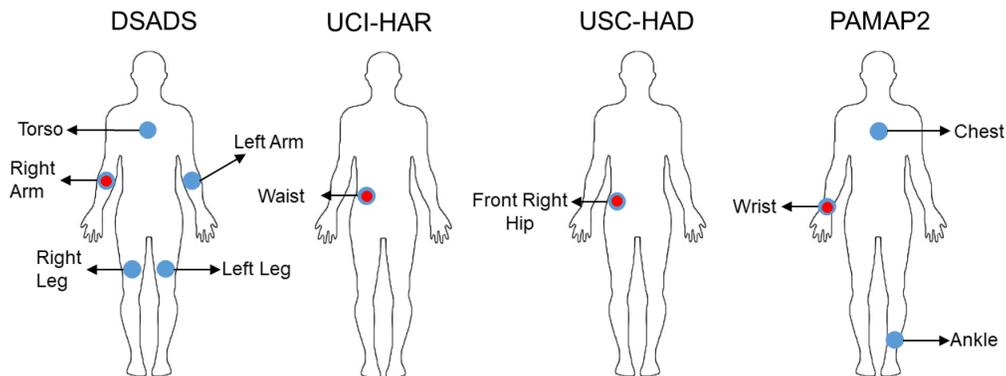
Fig. 3. Location of sensor devices in four datasets

(a) Transfer with ASTTL (ground-truth)

(b) No transfer with RF

(c) CORAL

(d) MEDA

(e) ASTTL
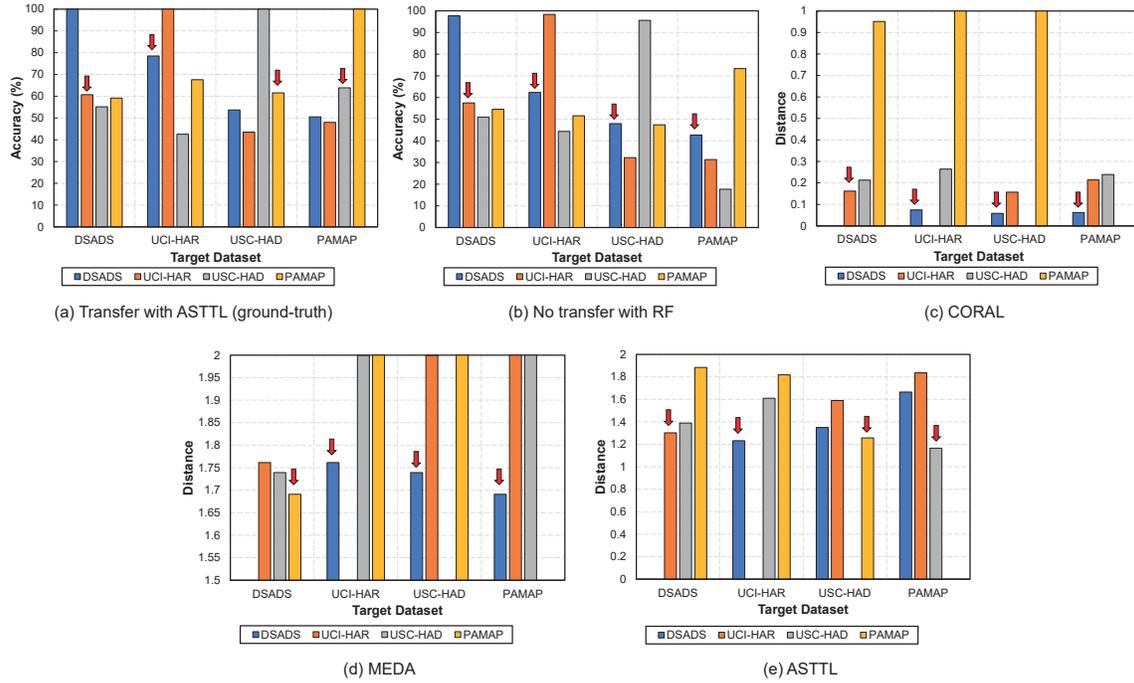
Fig. 4. Results of source domain selection

task. Apparently, computation of the domain distance is much more efficient than building classifiers. Therefore, the distance can serve as the measure for source domain selection.

To verify the effectiveness of ASTTL on source domain selection, we also provide the classification accuracy using a base classifer i.e. Random Forest trained on the original source domain and tested on the target domain as a comparison, it is another common used metric for similarity [9, 54]. The results are shown in Figure 4 (b), the highest accuracy is acquired by the model trained on the same dataset as the target, which is just for comparison. In addition, we compare the performance of our method with two state-of-the-art methods CORAL (CORrelation alignment) [47] and MEDA [56]. We treat each dataset as the target domain and perform source domain selection using different methods. The results of distances calculated by these methods are shown in Figure 4 (c), Figure 4 (d) and Figure 4 (e), where the most similar source domain to the target domain (which has the smallest distance) is marked with red arrows. Due to the differences in calculation methods, the results obtained by different methods may vary greatly, so we normalized the results of CORAL to 0-1. The distance 0 denotes the target domain itself is acting as the source domain, which is only the ideal state. Besides, the source domain selection results are listed in Table 2, the wrong selections i.e. the selections that are not the same as the ground-truth are marked with (×), and the last column summarizes the total number of correct selections by each method.

Combine all figures in Figure 4 ignoring the ideal states and Table 2, we can find : 1) ASTTL is superior to CORAL and MEDA, as ASTTL can accurately select the closest source domain which is consistent with the ground-truth, so that the transfer task can achieve the highest classification accuracy in transfer learning, while CORAL and MEDA makes several bad selections (the cross mark in Table 2). Therefore, it indicates the superiority of ASTTL in selecting the most appropriate source domain to achieve the most successfully transfer learning.

Table 2. Source selection results of comparison methods

Target	DSADS	UCI-HAR	USC-HAD	PAMAP	#Correct
Ground-truth	UCI-HAR	DSADS	PAMAP	USC-HAD	4
RF	UCI-HAR	DSADS	DSADS (×)	DSADS (×)	2
CORAL	UCI-HAR	DSADS	DSADS (×)	DSADS (×)	2
MEDA	PAMAP (×)	DSADS	DSADS (×)	DSADS (×)	1
ASTTL	UCI-HAR	DSADS	PAMAP	USC-HAD	**4**

2) Training base classifiers under no transfer condition like Figure 4 (b) shows, two of the task with highest accuracy (i.e. *DSADS → USC-HAD, DSADS → PAMAP*) can't perform the best transfer. Thus it is not accurate to use classification accuracy before transferring as the similarity metric between domains.

4.3 Evaluation of Activity Transfer

4.3.1 Classification Performance. Our goal is to learn a model from the given closest source domain $\mathcal{D}_s$ to gain an accurate classification of the novel target domain $\mathcal{D}_t$. In this section, we make a further evaluation of the Spatial-Temporal transfer method by comparative experimenting with several classic and state-of-the-art methods.

These methods are listed as follows:

- KNN, which is a popular traditional classification algorithm and only serves as the baseline method.
- Transfer Component Analysis (TCA) [38], which focuses on domain alignment with the assumption that the marginal distribution is different.
- Geodesic Flow Kernel (GFK) [18], which performs domain shift by integrating an infinite number of subspace in manifold.
- CORrelation Alignment (CORAL) [47], which achieves unsupervised domain adaptation by aligning the second-order statistics.
- Balanced Distribution Adaptation (BDA) [52], which tackles the issue of class imbalance by adaptive weighting the marginal and conditional distribution discrepancies.
- Manifold Embedded Distribution Alignment (MEDA) [56], performs dynamic marginal and conditional distributions alignment for manifold domain adaptation.

Experimental settings are as follows. All of the methods conduct the same transfer tasks in which each dataset is regarded as the target domain, and perform transfer learning to the most similar source domain. By following the results of source domain selection in Section 4.2, we report the performance of transfer learning between the most similar source-target pairs: UCI-HAR → DSADS, DSADS → UCI-HAR, PAMAP → USC-HAD, and USC-HAD → PAMAP. ASTTL only uses 1NN (K=1 in KNN) to get soft labels for the target domain in the first iteration, and iteratively refines the prediction according to Equation 1. We use RBF kernel for ASTTL and set $\lambda = 0.1, \beta = 1$. The parameters are tuned according to transfer cross-validation [63]. The other parameters of each method are also set to achieve the best performance according to the corresponding literature for evaluation settings. For a fair study, each experiment is repeated 10 times and then report the average performance. In order to extensively evaluate the method, we use both classification accuracy and F1 score as the evaluation metrics.

The results of classification accuracy and mean-F1 score of all methods are shown in Table 3. From these results, we have the following observations: 1) The proposed ASTTL achieves the best classification accuracy and mean-F1 score on all tasks. To be concrete, the average accuracy on 4 tasks is **66.3%**, compared to the second best method MEDA, it significantly improves the average accuracy by **8.7%**. ASTTL also has the best F1 score compared to other comparison methods. This indicates the effectiveness of ASTTL. 2) Other transfer learning methods such as BDA and MEDA can achieve good classification results on some tasks while performing worse
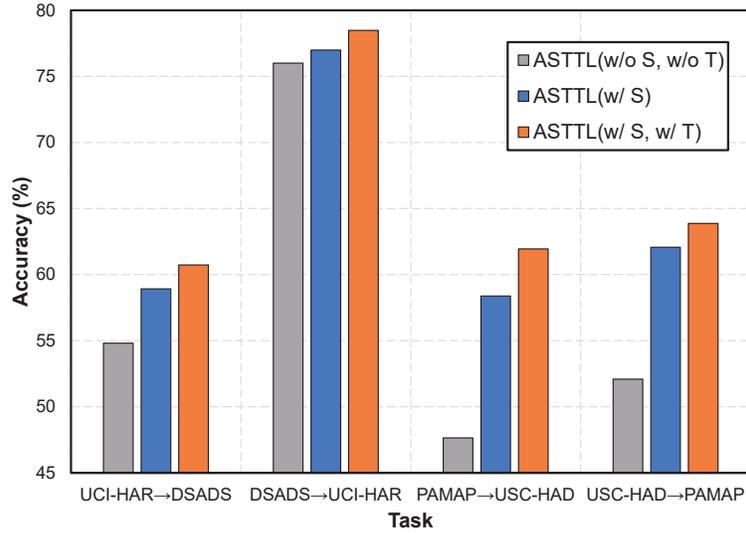
Fig. 5. Ablation study of ASTTL ('w/'=with, 'w/o'=without)

on specific tasks. This may be because that they only consider domain adaptation from the spatial aspect. Other methods have different limitations, the domain shift is still too large to achieve successful transfer after domain adaptation. 3) From all the results, we can see that the performances can drop in transfer learning scenario when the source and target domains have different distributions. This indicates the necessity of transfer learning in these situations. Among those transfer learning methods, our proposed ASTTL achieves the best performance.

4.3.2 Ablation Study. ASTTL consists of two important components, namely, spatial and temporal feature learning. In this section, we evaluate the importance of each component by ablation experiments. Specifically, we use *ASTTL(w/o S, w/o T)* to denote the ASTTL algorithm with no spatial and temporal feature learning procedures, and use *ASTTL(w/ S)* to denote the situation with only spatial feature learning. *ASTTL(w/ S, w/ T)* denotes the full ASTTL method. The ablation study is shown in Figure 5. From this figure, we can draw some conclusions: 1) Each feature learning component makes important contribution to the performance of ASTTL. 2) The proposed ASTTL(w/ S) shows better performance than ASTTL(w/o S, w/o T), this is because the spatial distribution alignment narrows the distance between the source and target feature space. 3) The proposed ASTTL(w/ S, w/ T) outperform the ASTTL(w/ S), indicating the temporal property can help improve the classification accuracy. By combining the two feature learning processes, ASTTL can achieve the best performance.

Table 3. Accuracy (%) and mean-F1 (shown in the bracket) of ASTTL and other comparison methods

Task	KNN	TCA	GFK	CORAL	BDA	MEDA	ASTTL
UCI-HAR → DSADS	54.5 (0.55)	54.7 (0.52)	54.4 (0.52)	36.9 (0.44)	60.6 (0.56)	54.8 (0.52)	**60.7 (0.56)**
DSADS → UCI-HAR	62.7 (0.61)	62.2 (0.61)	63.0 (0.61)	32.1 (0.12)	65.0 (0.62)	76.0 (0.71)	**78.5 (0.73)**
PAMAP → USC-HAD	45.6 (0.43)	47.2 (0.46)	45.4 (0.43)	47.7 (0.47)	53.3 (0.51)	47.6 (0.48)	**61.9 (0.56)**
USC-HAD → PAMAP	46.0 (0.49)	43.1 (0.49)	47.3 (0.50)	44.9 (0.48)	42.0 (0.48)	52.1 (0.51)	**63.9 (0.61)**
Average	52.2 (0.52)	51.8 (0.52)	52.5 (0.52)	40.4 (0.38)	55.2 (0.54)	57.6 (0.55)	**66.3 (0.61)**

Table 4. Precision, recall, and F1 score of ASTTL and other comparison methods

	GFK			BDA			MEDA			ASTTL		
Activity	P	R	F1	P	R	F1	P	R	F1	P	R	F1
L	1.00	0.99	1.00	0.98	0.98	0.98	1.00	1.00	1.00	**1.00**	**1.00**	**1.00**
W	0.53	**0.68**	0.59	0.52	0.43	0.47	0.72	0.60	0.66	**0.87**	0.63	**0.73**
U	0.43	**0.69**	0.53	0.50	0.48	0.49	0.67	0.55	0.60	**0.74**	0.65	**0.69**
D	0.00	0.00	0.00	0.51	0.65	0.57	**0.61**	**0.86**	**0.71**	0.56	0.83	0.67
Average	0.49	0.59	0.53	0.63	0.64	0.63	0.75	0.75	0.74	**0.79**	**0.78**	**0.77**
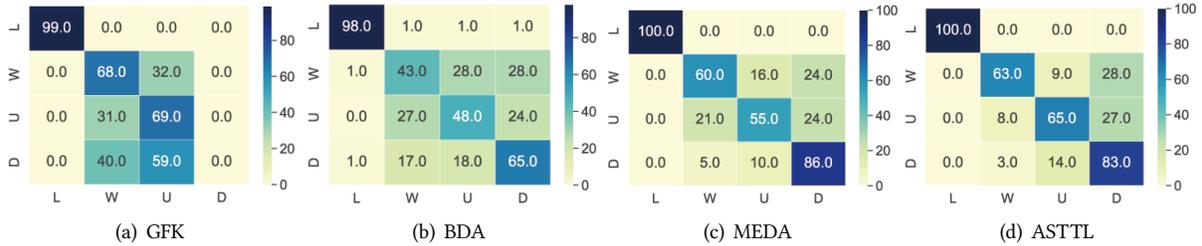
Fig. 6. Confusion matrices of ASTTL and other comparison methods

4.3.3 Detailed Analysis. In order to thoroughly evaluate the superiority of ASTTL, we provide fine-grained ROC curves [16] and compute the AUC on each categories of each transfer task in Figure 7, where class 0 to 3 denote the four kinds of activities. Macro and Micro average are also provided for a thorough view of each task. From the ROC curves, we can see that ASTTL can make smoothly classification on all activities of each transfer task, and the macro and micro AUC are all higher than **0.79**, which indicates ASTTL can perform effective classification. From the fine-grained results of each category, ASTTL can achieve an ideal classification accuracy on class 0 (i.e. LYING) with an ROC area > **0.99** and good classification on the other three dynamic activities (WALKING, WALKING UPSTAIRS, WALKING DOWNSTAIRS). Classification is more difficult on these dynamic activities due to the high similarity of the motions patterns of these activities.

Furthermore, noticing that the classes and samples are imbalanced in our tasks which is also common in real applications, we extensively provide horizontal comparison with other methods from fine-grained level by confusion matrix with the precision (P), recall (R), and macro F1-scores (F1) of each activity. Due to space limitation, we only report the results on task DSADS → UCI-HAR, and other tasks follow the same tendency. Figure 6 shows the confusion matrices of ASTTL and other three state-of-the-art methods GFK, BDA and MEDA. The value of the i_{th} row and j_{th} column represents the percentage of the samples with true label i while being classified into label j. Combining the results of Table 4 and Figure 6, we can observe that: GFK only achieves good classification on the first category while fail in the last one. BDA also recognizes one category well and performs worse in the other three categories. Compared with MEDA, ASTTL achieves better classification with the accuracy of each category is above **63.0%**. Among all comparison methods, ASTTL achieves the highest or the second highest value in all category, and has the highest average value in precision, recall and F1-score. To sum up, ASTTL achieves the best results in all categories, indicating its significant superiority in cross-dataset HAR tasks especially in imbalanced situations which is common in real applications.

4.3.4 Parameter Sensitivity. In this section, we evaluate the parameter sensitivity of ASTTL. It is worth noting that ASTTL achieves the best performance, and it contains fewer parameters than the recent state-of-the-art method MEDA [56]. We empirically evaluate the sensitivity of the subspace dimension d of manifold learning and parameter λ, β with a wide range on all transfer tasks. The results are shown in Figure 8 (a)~ (c). We observe

(a) UCI-HAR→ DSADS

(b) DSADS→ UCI-HAR

(c) PAMAP→ USC-HAD

(d) USC-HAD→ PAMAP
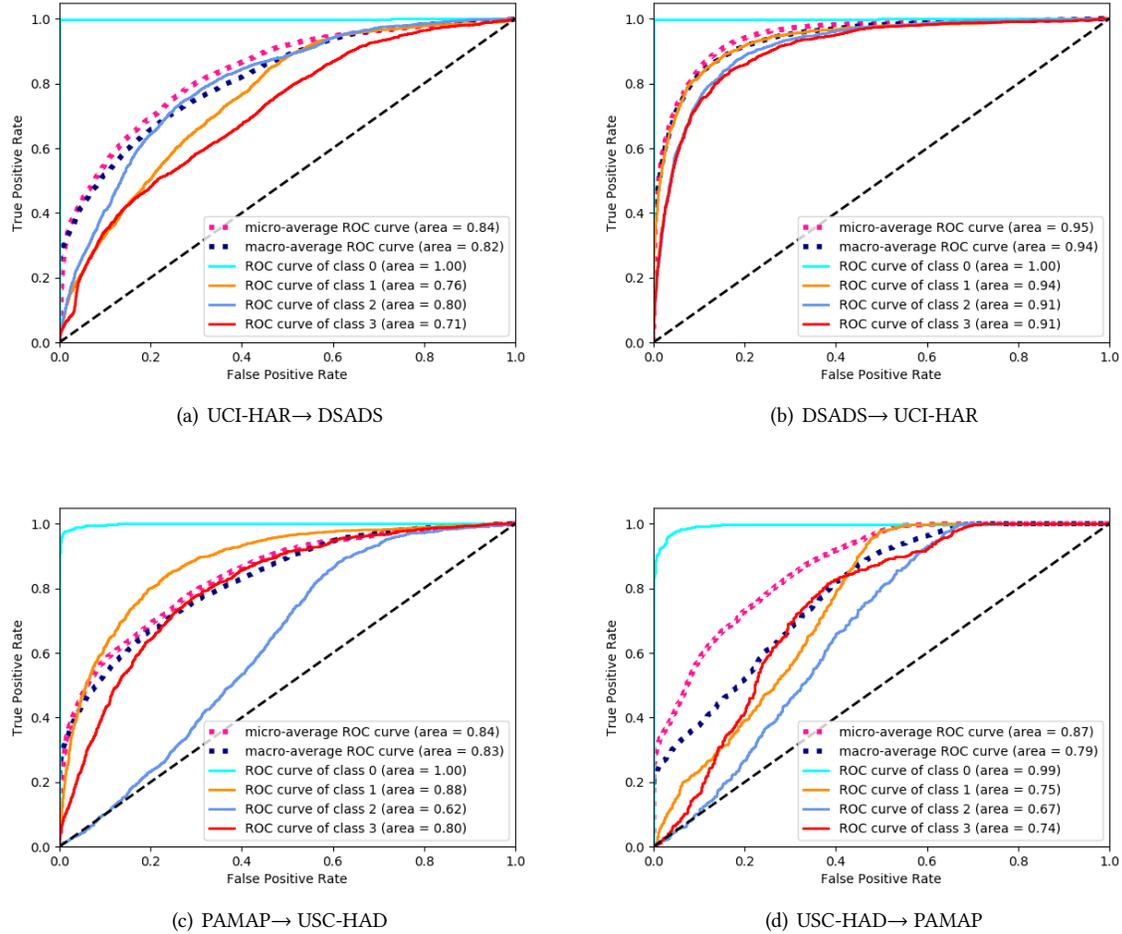
Fig. 7. ROC curves for each task of ASTTL

that the classification accuracy varies smoothly, which means that ASTTL is robust in human activity recognition with respect to a large range of parameter choices.

4.3.5 Convergence and Time Complexity. ASTTL is able to iteratively learn the final classifier. In this section, we investigate its convergence and time complexity. The convergence results in Figure 8 (d) indicate that ASTTL reaches a steady performance within fewer than 20 iterations. This shows the efficiency of training ASTTL in real applications. We analyze the time complexity of ASTTL using the big-O notation. Denote d as the number of non-zero features for the domains, m and n as the number of samples in source and target domains. The complexity of ASTTL mainly comes from two sources: 1) Temporal feature learning takes the complexity of $O\left(d(n+m)^2 + d^3\right)$, which mainly comes from the learning of the geodesic kernel $\mathbf{G}$. The complexity of spatial feature learning is $O(n^2)$, which is omitted compared to the temporal complexity. 2) The solving process of f takes the complexity of $O(n+m)^3$. Constructing the MMD matrix will take $O\left(C(n+m)^2\right)$ time.

To sum up, the time complexity of ASTTL is $O\left((n+m)^3 + (d+C)(n+m)^2 + d^3\right)$. Note that after manifold feature learning, d is often a small value. On the other hand, it is also possible to accelerate ASTTL using
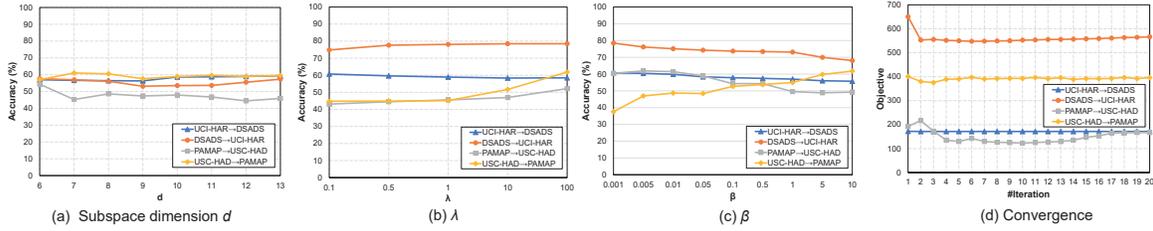
Fig. 8. (a)∼(c): Classification accuracy w.r.t d, λ and β, respectively. (d) Convergence analysis

conjugate gradient methods [56]. In addition, matrix multiplication is the very common operation in most of transfer methods whose time complexity is $O(n^{2.3757})$ by the Coppersmith-Winograd method [12].

4.4 Discussion

Compared to the recent popular deep transfer learning methods, ASTTL is a traditional transfer learning approach, it can improve the transfer learning performance on HAR comparing to the state-of-the-art traditional methods. It can be easily applied on the small resource constraint devices such as smart phones and wearables, while deep learning methods often require large computing resources and have to be compressed manually for small devices. One possible limitation of ASTTL maybe that it relies on certain feature extraction methods. As for human activity dataset, we probably need to extract time and frequency domain features. ASTTL is a generic transfer learning method which allows it to be easily extended to deep model by using the deep features extracted by deep neural networks such as CNNs [59] and LSTM [36]. We will leave this part for future research.

5 AN ADAPTIVE CROSS-DATASET ACTIVITY RECOGNITION SYSTEM

In this section, we design an adaptive cross-dataset activity recognition application in the real environment based on the proposed ASTTL method. We call the system Client-Cloud Collaborative Adaptive Activity Recognition System (3C2ARS). It is composed of two components: 1) the software on a smartphone for activity data collection, and 2) the software on a server for cross-dataset transfer learning. In the device side, using the smartphone can achieve efficient data collection. In the server side, based on the existing labeled data, the server can provide sufficient computation for continuous training of ASTTL and feedback to the client, thus to achieve real-time human activity recognition. The proposed application achieves satisfying performance in real activity recognition scenarios.
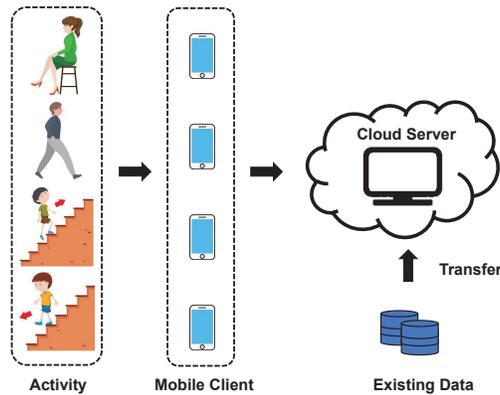
Fig. 9. Overview of the framework
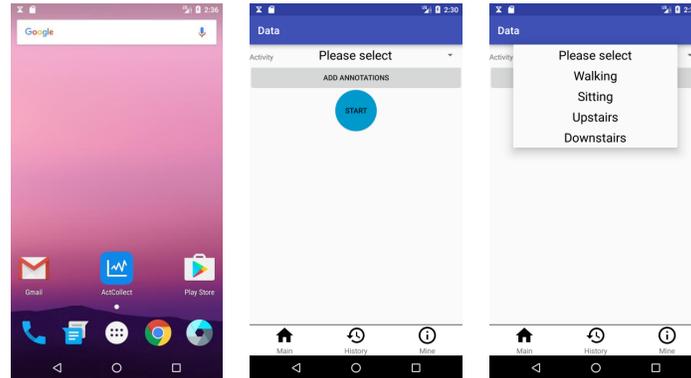
Fig. 10. Startup and data collection interfaces of the client application
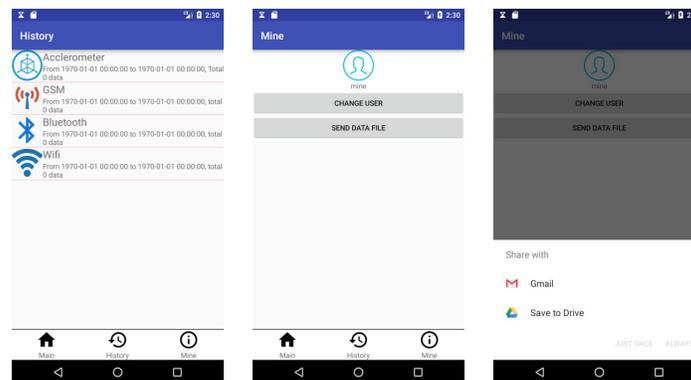
Fig. 11. History and transport interfaces of the client application

5.1 Framework

The framework of the system is shown in Figure 9. We develop an Android mobile application to collect the inertia information of people's daily activities in a natural and undisturbed environment. After that, the data are transferred to the cloud server, which can continuous perform data analysis and model building with high performance. Since the collected data is usually inadequate in quantity, the cloud server trains activity recognition model by transferring knowledge from the existing labeled datasets to the newly collected data with our proposed method ASTTL.

5.2 Mobile Client

As the most widely used wearable device in people's daily life, the smartphone makes it easy to access the activity signals when people are performing activities by utilizing the corresponding hardware facility of the smartphone. We develop an application called *ActCollect* to collect people's motion information including acceleration and gyroscope. In the meantime, WIFI, Bluetooth and GSM data can also be recorded in the motion environment. In the process of data collection, user can select the category of activity to assign a label, then the labeled data is saved into a different table. The collected data can also be easily exported and sent by the inner-built function of e-mail. The main interfaces of ActCollect is shown in Figure 10.

5.2.1 Activity Data Collecting Module. Generally speaking, most of the motion information of human can be accurately described by the inertia devices e.g. acceleration and gyroscope. Meanwhile, the information of the ubiquitous environment constructed by WIFI, Bluetooth and GPS can also reflect the motion state of human from the interaction between human and environment. In this module, people's motion data (i.e. acceleration, GSM, WIFI, and Bluetooth) is collected and saved into a lightweight database in the smartphone for further analysis. The interface of activity data collecting module is shown in Figure 10. In the *home page* of Figure 10, users can click the button of *Please select* to select the category of activity, and then users can click *Start* to start the activity data collection process. The collected data is visible in this interface in real-time, and the collection process can be stopped or paused by clicking this button again.

5.2.2 History Browsing Module for Activity Data . In order to check the accuracy of data collection in real-time, the history browsing module can provide a simple review of the collected data. As shown in Figure 11, by clicking the tab of *History*, the basic information of the activities including the category of data, time of the beginning and ending, and the number of samples can be observed. In this module, users can view all the behavior data that are already collected.

5.2.3 Activity Data Transport Module. For further data analysis, transport module provides the function of wireless data transport. By clicking the tab of *Mine* as shown in Figure 11, users can transport the activity data through the built-in e-mail application to the cloud. Thus the cloud server can further train a transfer learning model to recognize these activities.

5.3 Cloud Sever

Given a collected activity dataset, we use a high-performance computer as the remote server. Users can upload real-time activity data to the cloud by sending e-mails as aforementioned, or they can set a time schedule such as automatically uploading new data every one hour. The cloud sever trains transfer learning models continuously and update the model every 12 hours. The main interface of the *Cross-dataset Activity Recognition* cloud server is shown in Figure 12. In this paper, the whole transfer learning process is based on the proposed method ASTTL. It is noteworthy that this system can be extended by integrating other transfer learning approaches. This software is composed of four modules:

- `Source Domain Selection`: The software contains several public activity datasets to serve as the source domains for a new target task. Specifically, it contains four public activity recognition datasets as mentioned in Section 4. When a new target dataset comes, users can choose any of them as the source domain. Furthermore, in order to avoid negative transfer caused by selecting a wrong dataset, our software provides an *Auto-Selection* function. This function can automatically calculate the distance between the target domain and each source domains to select a most similar source domain.
- `Data visualization`: In order to visually display the different sensor readings before transfer learning, this function can give a quick preview of the sensor readings of certain activities.
- `Transfer Learning`: This is the core function of the software. Our system is based on the proposed ASTTL algorithm. After source domain selection, users can click the *Begin Transfer* button. Then, the system starts to train an ASTTL model for cross-dataset HAR.
- `Online Service API`: Our system provides an online service API for cross-data HAR using transfer learning algorithms to enhance the learning performance of activity recognition. Users can either upload their data to our server through http request for automatic labeling, or they can download our pretrained models to their local devices for transfer learning. We believe this will make our system more ubiquitous.
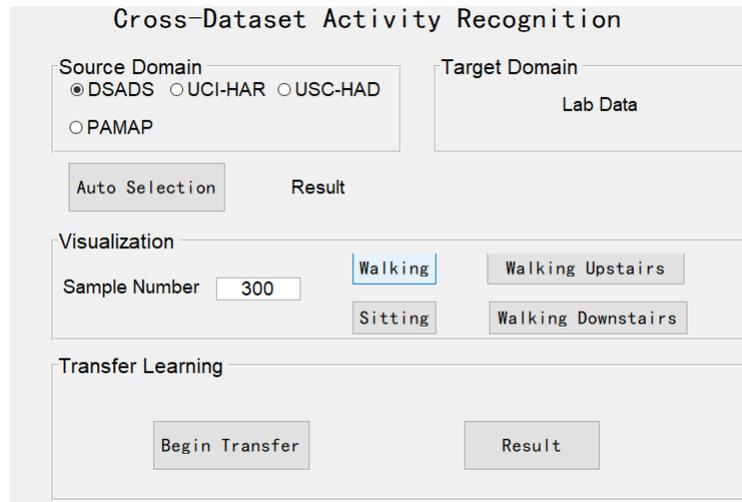
Fig. 12. Interface of the software in the cloud server

5.4 Effectiveness Evaluation of the System

In this section, we evaluate the effectiveness of our adaptive cross-dataset HAR system in real applications. Firstly, we introduce the process of data collection and brief information about the collected data. Secondly, we conduct experiments on cross-dataset HAR transfer learning from the existing datasets to the newly collected activity data and analyze the experimental results.

5.4.1 Data Collection. Using the smartphone software, we collect a dataset comprised by four activity categories including: *Walking*, *Sitting*, *Upstairs* and *Downstairs*. Eleven subjects participated in the data collection process who were asked to behave each activity for 5 times in a state of natural relaxation. Considering that most people prefer putting their phone around the front hip, the participants were asked to put their phones in their trouser picket in our experiments. We successfully collected over 20 M of sensor readings which is larger than most of the existing public activity recognition datasets. The statistics of the subjects contain age, height and weight factors. Subjects age from 20 to 30 years old and average at 26, their height is in a range of 155 to 185 centimeters, with an average of 163 centimeters and their weight is between 45 and 80 kilograms, with an average of 61.5 kilograms. We make effort to cover a wider range of population according to these three factors.

5.4.2 Accuracy and Analysis. In this section, we evaluate the accuracy of the system w.r.t. classification accuracy based on transfer learning. In order to verify the superiority of the proposed system, we compare our method with other four classic or state-of-the-art methods. For the collected target data, we firstly conduct data preprocessing and feature extraction using the methods mentioned in Section 4. Then, we calculate and get the most similar source domain as a comparison case to other transfer tasks. In this experiment, the most similar source domain to the collected lab data is USC-HAD, which is in consistency with the ground truth. Other settings are consistent with those in Section 4.3. We compare the classification accuracy of activity recognition using each method and present the results in Table 5.

From Table 5, we can draw the following conclusions:

- The classification performance of transfer learning is better than no transfer method (KNN and RF), this is due to the large distribution discrepancy between two domains which can produce worse results if reusing the models directly.

Table 5. Accuracy (%) of ASTTL and other comparison methods in real applications

Source Domain	KNN	RF	CORAL	BDA	ASTTL
DSADS	61.3	43.0	60.8	57.4	**61.6**
UCI-HAR	44.2	38.1	36.7	57.4	**57.8**
USC-HAD	59.9	59.8	59.3	56.3	**63.0**
PAMAP	44.7	48.8	44.5	51.8	**54.7**
Average	52.5	47.4	50.3	55.7	**59.3**

- Given the same collected data, the situation that USC-HAD acts as the source domain achieves the best performance. This can in turn verify that USC-HAD is the most similar source domain to the target (as introduced above).
- Our system achieved the highest accuracy in the activity recognition tasks when given the newly collected data. To be more specific, our proposed ASTTL significantly outperforms the best comparison method BDA by **3.6%** in average accuracy. This indicates that ASTTL is effective for cross-dataset HAR and can be applied in real-life activity recognition.

5.5 User Satisfaction of the System

This section provides an overall evaluation of the system after making a survey on 24 users who have used this system. We make this survey online, where the full score of the evaluation is 5 points, indicating that the user is very satisfied, and the lowest score is 1 point, indicating not satisfied. The last question is an open question, requiring the user to give feedback on other opinions and suggestions of the system.

Firstly, from the perspective of overall satisfaction, we receive an average of 4.71 points, in which 70.83% users give a full score, and 29.17% give a high score of 4 points. This illustrates that our system can solve the pivotal issues of activity recognition with adaptive transfer learning, thus the users give positive reflection. User satisfaction w.r.t. the overall system, mobile client, and the cloud server is shown in Figure 13. Additionally, the usability of applications on the smartphone and server is evaluated by receiving a relatively high score of 4.71 and 4.63, which reflects user acceptance and approval.

In addition, we collect the users' satisfaction to usability, interface design, and operating conditions. These are important measurements for an application which can reflect the performance of our system in detail.

- `Usability`: Usability is a key issue in application design. The operating difficulty for new users, the ability of interface description, the presentation of system language are all considerations for usability.
- `Interface Design`: The unity of the system interface, button and panel description need to be considered in terms of interface design.
- `Operating Condition`: Problems such as long delay, breakdown, etc. can influence the usability of the system.

Table 6 shows the statistic results. The numbers in this table indicate the number of users that give specific scores. From Table 6, it can be concluded that users are well satisfied with both the mobile client and cloud server applications. We also collect the comment of these three evaluation aspects for client and server form users, results are shown in Figure 13. We can see that the vast majority of users are satisfied with the proposed system.

6 CONCLUSIONS AND FUTURE WORK

Label scarcity is a common problem in Human Activity Recognition. Given a new unlabeled HAR dataset, it is extremely difficult to train machine learning models for activity recognition. To this end, we propose an Adaptive Spatial-Temporal Transfer Learning (ASTTL) method. ASTTL learns the spatially and temporally
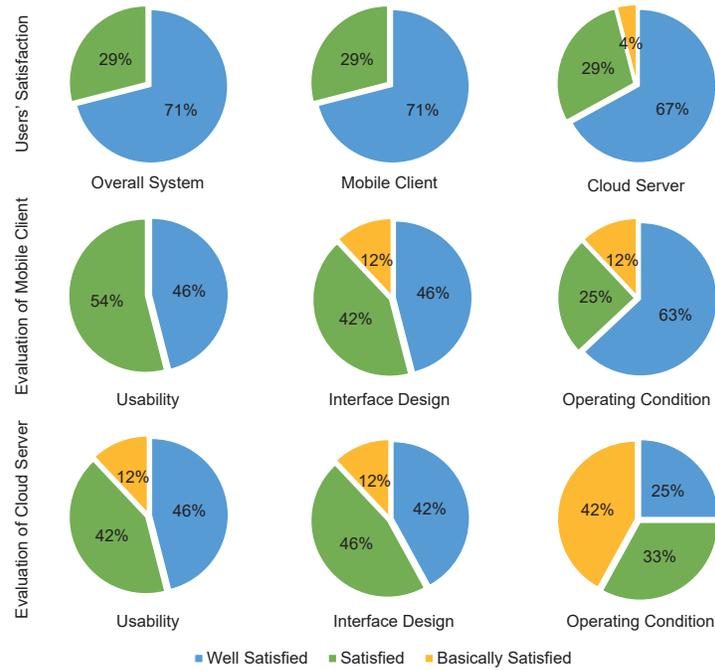
Fig. 13. User satisfaction survey to the system

Table 6. Satisfaction survey of the mobile client and cloud server

	Mobile Client			Cloud Server		
Score	3	4	5	3	4	5
Usability	0	13	11	3	10	11
Interface Design	3	10	11	3	11	10
Operating Condition	3	6	15	10	8	6

adaptive transferable features in a unified framework. Then, it can be applied to source domain selection and activity transfer in cross-dataset HAR problems. Comprehensive experiments demonstrate that ASTTL has superior accuracy in source domain selection and knowledge transfer, which greatly enhances the ability of activity recognition in the target task. In addition, based on ASTTL, we design a Client-Cloud Collaborative Adaptive Activity Recognition System (3C2ARS) to achieve the real-life application of activity recognition with satisfactory performance.

In the future, we expect to increase the performance of HAR by extending ASTTL using deep neural networks. We also intend to apply ASTTL to other fine-grained activity recognition problems. We believe that our work can be widely applied to various fields of ubiquitous computing, such as fatigue detection, medical diagnosis, and smart home sensing.

ACKNOWLEDGMENTS

REFERENCES

[1] Muhammad Jamal Afridi, Arun Ross, and Erik M Shapiro. 2018. On automated source selection for transfer learning in convolutional neural networks. *Pattern Recognition* 73 (2018), 65–75.

[2] Davide Anguita, Alessandro Ghio, Luca Oneto, Xavier Parra, and Jorge L Reyes-Ortiz. 2012. Human activity recognition on smartphones using a multiclass hardware-friendly support vector machine. In *International workshop on ambient assisted living.* Springer, 216–223.

[3] Mahsa Baktashmotlagh, Mehrtash T Harandi, Brian C Lovell, and Mathieu Salzmann. 2014. Domain adaptation on the statistical manifold. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition.* 2481–2488.

[4] Billur Barshan and Murat Cihan Yüksek. 2014. Recognizing daily and sports activities in two open source machine learning environments using body-worn sensor units. *Comput. J.* 57, 11 (2014), 1649–1667.

[5] Shai Ben-David, John Blitzer, Koby Crammer, and Fernando Pereira. 2007. Analysis of representations for domain adaptation. In *Advances in neural information processing systems.* 137–144.

[6] John Blitzer, Ryan McDonald, and Fernando Pereira. 2006. Domain adaptation with structural correspondence learning. In *EMNLP.* Association for Computational Linguistics, 120–128.

[7] Nicholas Carlini and David Wagner. 2017. Adversarial examples are not easily detected: Bypassing ten detection methods. In *Proceedings of the 10th ACM Workshop on Artificial Intelligence and Security.* ACM, 3–14.

[8] Qingchao Chen, Yang Liu, et al. 2018. Re-weighted adversarial adaptation network for unsupervised domain adaptation. In *CVPR.* 7976–7985.

[9] Yiqiang Chen, Jindong Wang, Meiyu Huang, and Han Yu. 2019. Cross-position activity recognition with stratified transfer learning. *Pervasive and Mobile Computing* (2019).

[10] Yiqiang Chen, Jindong Wang, Chaohui Yu, Wen Gao, and Xin Qin. 2019. FedHealth: A Federated Transfer Learning Framework for Wearable Healthcare. *arXiv preprint arXiv:1907.09173* (2019).

[11] Yiqiang Chen, Zhongtang Zhao, Shuangquan Wang, and Zhenyu Chen. 2012. Extreme learning machine-based device displacement free activity recognition model. *Soft Computing* 16, 9 (2012), 1617–1625.

[12] Don Coppersmith and Shmuel Winograd. 1990. Matrix multiplication via arithmetic progressions. *Journal of symbolic computation* 9, 3 (1990), 251–280.

[13] Yin Cui, Yang Song, Chen Sun, Andrew Howard, and Serge Belongie. 2018. Large scale fine-grained categorization and domain-specific transfer learning. In *Proceedings of the IEEE conference on computer vision and pattern recognition.* 4109–4118.

[14] Mariella Dimiccoli, Juan Marín, and Edison Thomaz. 2018. Mitigating Bystander Privacy Concerns in Egocentric Activity Recognition with Deep Learning and Intentional Image Degradation. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* 1, 4 (2018), 132.

[15] Renjie Ding, Xue Li, et al. 2019. Empirical Study and Improvement on Deep Transfer Learning for Human Activity Recognition. *Sensors* 19, 1 (2019), 57.

[16] Tom Fawcett. 2006. An introduction to ROC analysis. *Pattern recognition letters* 27, 8 (2006), 861–874.

[17] George Foster, Cyril Goutte, and Roland Kuhn. 2010. Discriminative instance weighting for domain adaptation in statistical machine translation. In *EMNLP.* Association for Computational Linguistics, 451–459.

[18] Boqing Gong, Yuan Shi, Fei Sha, and Kristen Grauman. 2012. Geodesic flow kernel for unsupervised domain adaptation. In *2012 IEEE Conference on Computer Vision and Pattern Recognition.* IEEE, 2066–2073.

[19] Raghuraman Gopalan, Ruonan Li, and Rama Chellappa. 2011. Domain adaptation for object recognition: An unsupervised approach. In *2011 international conference on computer vision.* IEEE, 999–1006.

[20] Arthur Gretton, Karsten M Borgwardt, Malte J Rasch, Bernhard Schölkopf, and Alexander Smola. 2012. A kernel two-sample test. *Journal of Machine Learning Research* 13, Mar (2012), 723–773.

[21] Arthur Gretton, Karsten M. Borgwardt, Malte J. Rasch, Bernhard Schölkopf, and Alexander J. Smola. 2007. A Kernel Method for the Two-Sample-Problem. In *Conference on Advances in Neural Information Processing Systems.*

[22] Yu Guan and Thomas Plötz. 2017. Ensembles of deep lstm learners for activity recognition using wearables. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* 1, 2 (2017), 11.

[23] Jihun Hamm and Daniel D Lee. 2008. Grassmann discriminant analysis: a unifying view on subspace-based learning. In *Proceedings of the 25th international conference on Machine learning.* ACM, 376–383.

[24] HM Hossain, MD Al Haiz Khan, and Nirmalya Roy. 2018. DeActive: scaling activity recognition with active deep learning. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* 2, 2 (2018), 66.

[25] HM Sajjad Hossain, Md Abdullah Al Hafiz Khan, and Nirmalya Roy. 2017. Active learning enabled activity recognition. *Pervasive and Mobile Computing* 38 (2017), 312–330.

[26] Derek Hao Hu, Vincent Wenchen Zheng, and Qiang Yang. 2011. Cross-domain activity recognition via transfer learning. *Pervasive and Mobile Computing* 7, 3 (2011), 344–358.

[27] Tâm Huynh, Mario Fritz, and Bernt Schiele. 2008. Discovery of activity patterns using topic models.. In *UbiComp*, Vol. 8. 10–19.

[28] John G Kemeny and J Laurie Snell. 1976. *Markov Chains*. Springer-Verlag, New York.

[29] Md Abdullah Al Hafiz Khan, Nirmalya Roy, and Archan Misra. 2018. Scaling human activity recognition via deep learning-based domain adaptation. In *PerCom*. IEEE, 1–9.

[30] Jianmei Lei, Fangli Liu, Qingwen Han, Yunyang Tang, Lingqiu Zeng, Min Chen, Lei Ye, and Li Jin. 2018. Study on Driving Fatigue Evaluation System Based on Short Time Period ECG Signal. In *2018 21st International Conference on Intelligent Transportation Systems (ITSC)*. IEEE, 2466–2470.

[31] Dawei Liang and Edison Thomaz. 2019. Audio-Based Activities of Daily Living (ADL) Recognition with Large-Scale Acoustic Embeddings from Online Videos. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* 3, 1 (2019), 17.

[32] Di Lin, Xing An, and Jian Zhang. 2013. Double-bootstrapping source data selection for instance-based transfer learning. *Pattern Recognition Letters* 34, 11 (2013), 1279–1285.

[33] Jianhua Lin. 1991. Divergence measures based on the Shannon entropy. *IEEE Transactions on Information theory* 37, 1 (1991), 145–151.

[34] Mingsheng Long, Jianmin Wang, Guiguang Ding, Jiaguang Sun, and Philip S Yu. 2013. Transfer feature learning with joint distribution adaptation. In *Proceedings of the IEEE international conference on computer vision*. 2200–2207.

[35] Mingsheng Long, Jianmin Wang, Jiaguang Sun, and S Yu Philip. 2015. Domain invariant transfer kernel learning. *IEEE Transactions on Knowledge and Data Engineering* 27, 6 (2015), 1519–1532.

[36] Francisco Ordóñez and Daniel Roggen. 2016. Deep convolutional and lstm recurrent neural networks for multimodal wearable activity recognition. *Sensors* 16, 1 (2016), 115.

[37] Sinno Jialin Pan, James T Kwok, and Qiang Yang. 2008. Transfer Learning via Dimensionality Reduction.. In *AAAI*, Vol. 8. 677–682.

[38] Sinno Jialin Pan, Ivor W Tsang, James T Kwok, and Qiang Yang. 2011. Domain adaptation via transfer component analysis. *IEEE Transactions on Neural Networks* 22, 2 (2011), 199–210.

[39] Sinno Jialin Pan and Qiang Yang. 2010. A survey on transfer learning. *IEEE Transactions on knowledge and data engineering* 22, 10 (2010), 1345–1359.

[40] Salyer B Reed, Tyson RC Reed, and Sergiu M Dascalu. 2019. Spatiotemporal recursive hyperspheric classification with an application to dynamic gesture recognition. *Artificial Intelligence* 270 (2019), 41–66.

[41] Attila Reiss and Didier Stricker. 2012. Introducing a new benchmarked dataset for activity monitoring. In *2012 16th International Symposium on Wearable Computers*. IEEE, 108–109.

[42] Robert Remus. 2012. Domain adaptation using domain similarity-and domain complexity-based instance selection for cross-domain sentiment analysis. In *2012 IEEE 12th international conference on data mining workshops*. IEEE, 717–723.

[43] Vitor F Rey and Paul Lukowicz. 2017. Label Propagation: An Unsupervised Similarity Based Method for Integrating New Sensors in Activity Recognition Systems. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* 1, 3 (2017), 94.

[44] Seyed Ali Rokni and Hassan Ghasemzadeh. 2018. Autonomous training of activity recognition algorithms in mobile sensors: a transfer learning approach in context-invariant views. *IEEE Transactions on Mobile Computing* 17, 8 (2018), 1764–1777.

[45] Lex Razoux Schultz, Marco Loog, and Peyman Mohajerin Esfahani. 2018. Distance Based Source Domain Selection for Sentiment Classification. *arXiv preprint arXiv:1808.09271* (2018).

[46] Alex Smola, Arthur Gretton, Le Song, and Bernhard Schölkopf. 2007. A Hilbert space embedding for distributions. In *International Conference on Algorithmic Learning Theory*. Springer, 13–31.

[47] Baochen Sun, Jiashi Feng, and Kate Saenko. 2016. Return of frustratingly easy domain adaptation. In *Thirtieth AAAI Conference on Artificial Intelligence*.

[48] Baochen Sun and Kate Saenko. 2016. Deep coral: Correlation alignment for deep domain adaptation. In *European Conference on Computer Vision*. Springer, 443–450.

[49] Ben Tan, Yu Zhang, Sinno Jialin Pan, and Qiang Yang. 2017. Distant domain transfer learning. In *AAAI*.

[50] Praneeth Vepakomma, Debraj De, Sajal K Das, and Shekhar Bhansali. 2015. A-Wristocracy: Deep learning on wrist-worn sensing for recognition of user complex activities. In *2015 IEEE 12th International Conference on Wearable and Implantable Body Sensor Networks (BSN)*. IEEE, 1–6.

[51] Jindong Wang, Yiqiang Chen, Wenjie Feng, Han Yu, Meiyu Huang, and Qiang Yang. 2019. Transfer Learning with Dynamic Distribution Adaptation. *arXiv preprint arXiv:1909.08531* (2019).

[52] Jindong Wang, Yiqiang Chen, Shuji Hao, Wenjie Feng, and Zhiqi Shen. 2017. Balanced distribution adaptation for transfer learning. In *2017 IEEE International Conference on Data Mining (ICDM)*. IEEE, 1129–1134.

[53] Jindong Wang, Yiqiang Chen, Shuji Hao, Xiaohui Peng, and Lisha Hu. 2019. Deep learning for sensor-based activity recognition: A survey. *Pattern Recognition Letters* 119 (2019), 3–11.

[54] Jindong Wang, Yiqiang Chen, Lisha Hu, Xiaohui Peng, and S Yu Philip. 2018. Stratified transfer learning for cross-domain activity recognition. In *2018 IEEE International Conference on Pervasive Computing and Communications (PerCom)*. IEEE, 1–10.

[55] Jindong Wang, Yiqiang Chen, Han Yu, Meiyu Huang, and Qiang Yang. 2019. Easy Transfer Learning By Exploiting Intra-domain Structures. *arXiv preprint arXiv:1904.01376* (2019).

[56] Jindong Wang, Wenjie Feng, Yiqiang Chen, Han Yu, Meiyu Huang, and Philip S Yu. 2018. Visual Domain Adaptation with Manifold Embedded Distribution Alignment. In *ACM Multimedia Conference*. ACM, 402–410.

[57] Jindong Wang, Vincent W Zheng, Yiqiang Chen, and Meiyu Huang. 2018. Deep transfer learning for cross-domain activity recognition. In *ICCSE*. ACM, 16.

[58] Feng Xu, Jianfei Yu, and Rui Xia. 2018. Instance-based Domain Adaptation via Multiclustering Logistic Approximation. *IEEE Intelligent Systems* 33, 1 (2018), 78–88.

[59] Jianbo Yang, Minh Nhut Nguyen, Phyo Phyo San, Xiao Li Li, and Shonali Krishnaswamy. 2015. Deep convolutional neural networks on multichannel time series for human activity recognition. In *Twenty-Fourth International Joint Conference on Artificial Intelligence*.

[60] Chaohui Yu, Jindong Wang, Yiqiang Chen, and Meiyu Huang. 2019. Transfer Learning with Dynamic Adversarial Adaptation Network. *arXiv preprint arXiv:1909.08184* (2019).

[61] Fusang Zhang, Kai Niu, Jie Xiong, Beihong Jin, Tao Gu, Yuhang Jiang, and Daqing Zhang. 2019. Towards a Diffraction-based Sensing Approach on Human Activity Recognition. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* 3, 1 (2019), 33.

[62] Mi Zhang and Alexander A Sawchuk. 2012. USC-HAD: a daily activity dataset for ubiquitous activity recognition using wearable sensors. In *ACM UbiComp*. ACM, 1036–1043.

[63] Erheng Zhong, Wei Fan, Qiang Yang, Olivier Verscheure, and Jiangtao Ren. 2010. Cross validation framework to choose amongst models and datasets for transfer learning. In *Joint European Conference on Machine Learning and Knowledge Discovery in Databases*. Springer, 547–562.

[64] Fuzhen Zhuang, Xiaohu Cheng, Ping Luo, Sinno Jialin Pan, and Qing He. 2015. Supervised representation learning: Transfer learning with deep autoencoders. In *IJCAI*.

[65] Hankz Hankui Zhuo and Qiang Yang. 2014. Action-model acquisition for planning via transfer learning. *Artificial intelligence* 212 (2014), 80–103.