Transfer Learning for Activity Recognition: A Survey

Diane Cook, Kyle D. Feuz, and Narayanan C. Krishnan

Abstract

Many intelligent systems that focus on the needs of a human require information about the activities being performed by the human. At the core of this capability is activity recognition, which is a challenging and well-researched problem. Activity recognition algorithms require substantial amounts of labeled training data yet need to perform well under very diverse circumstances. As a result, researchers have been designing methods to identify and utilize subtle connections between activity recognition datasets, or to perform *transfer-based activity recognition*. In this paper we survey the literature to highlight recent advances in transfer learning for activity recognition. We characterize existing approaches to transfer-based activity recognition by sensor modality, by differences between source and target environments, by data availability, and by type of information that is transferred. Finally, we present some grand challenges for the community to consider as this field is further developed.

1 Introduction

Researchers in the artificial intelligence community have struggled for decades trying to build machines capable of matching or exceeding the mental capabilities of humans. One capability that continues to challenge researchers is designing systems which can leverage experience from previous tasks into improved performance in a new task which has not been encountered before. When the new task is drawn from a different population than the old, this is considered to be *transfer learning*. The benefits of transfer learning are numerous; less time is spent learning new tasks, less information is required of experts (usually human), and more situations can be handled effectively. These potential benefits have lead researchers to apply transfer learning techniques to many domains with varying degrees of success.

One particularly interesting domain for transfer learning is human activity recognition. The goal of human activity recognition is to be able to correctly classify the current activity a human or group of humans is performing given some set of data. Activity recognition is important to a variety of applications including health monitoring, automatic security surveillance, and home automation. As research in this area has progressed, an increasing number of researchers have started looking at ways transfer learning can be applied to reduce the training time and effort required to initialize new activity recognition systems, to make the activity recognition systems more robust and versatile, and to effectively reuse the existing knowledge that has previously been generated.

With the recent explosion in the number of researchers and the amount of research being done on transfer learning, activity recognition, and transfer learning

for activity recognition, it becomes increasingly important to critically analyze this body of work and discover areas which still require further investigation. Although recent progress in transfer learning has been analyzed in [50, 61, 70] and several surveys have been conducted on activity recognition [2, 4, 10, 24] no one has specifically looked into the intersection of these two areas. This survey, therefore, examines the field of transfer-based activity recognition and the unique challenges presented in this domain. For an overview of the survey, see Figure 1 which illustrates the topics covered in this survey and how they relate to each other.

2 Background

Activity recognition aims to identify activities as they occur based on data collected by sensors. There exist a number of approaches to activity recognition [28] that vary depending on the underlying sensor technologies that are used to monitor activities, the alternative machine learning algorithms that are used to model the activities and the realism of the testing environment.

Advances in pervasive computing and sensor networks have resulted in the development of a wide variety of sensor modalities that are useful for gathering information about human activities. Wearable sensors such as accelerometers are commonly used for recognizing ambulatory movements (e.g., walking, running, sitting, climbing, and falling) [31, 40]. More recently, researchers are exploring smart phones equipped with accelerometers and gyroscopes to recognize such movement and gesture patterns [33].

Environment sensors such as infrared motion detectors or magnetic door sensors have been used to gather information about more complex activities such as cooking, sleeping, and eating. These sensors are adept in performing location-based activity recognition in indoor environments [1, 38, 27] just as GPS is used for outdoor environments [36]. Some activities such as washing dishes, taking medicine, and using the phone are characterized by interacting with unique objects. In response, researchers have explored the usage of RFID tags and shimmer sensors for tagging these objects and using the data for activity recognition [45, 52]. Researchers have also used data from video cameras and microphones as well [1].

There have been many varied machine learning models that have been used for activity recognition. These can be broadly categorized into template matching / transductive techniques, generative, and discriminative approaches. Template matching techniques employ a kNN classifier based on Euclidean distance or dynamic time warping. Generative approaches such as naïve Bayes classifiers where activity samples are modeled using Gaussian mixtures have yielded promising results for batch learning. Generative probabilistic graphical models such as hidden Markov models and dynamic Bayesian networks have been used to model activity sequences and to smooth recognition results of an ensemble classifier [35]. Decision trees as well as bagging and boosting methods have been tested [40]. Discriminative approaches, including support vector machines and conditional random fields, have also been effective [11, 27] and unsupervised discovery and recognition methods have also been introduced [58, 22]. The traditional approaches to activity recognition make the strong assumption that the training and test data are drawn from identical distributions. Many real-world applications cannot be represented in this setting and thus the baseline activity recognition approaches have to be modified to

Figure 1: Content map of the transfer learning for activity recognition domain covered in this survey.
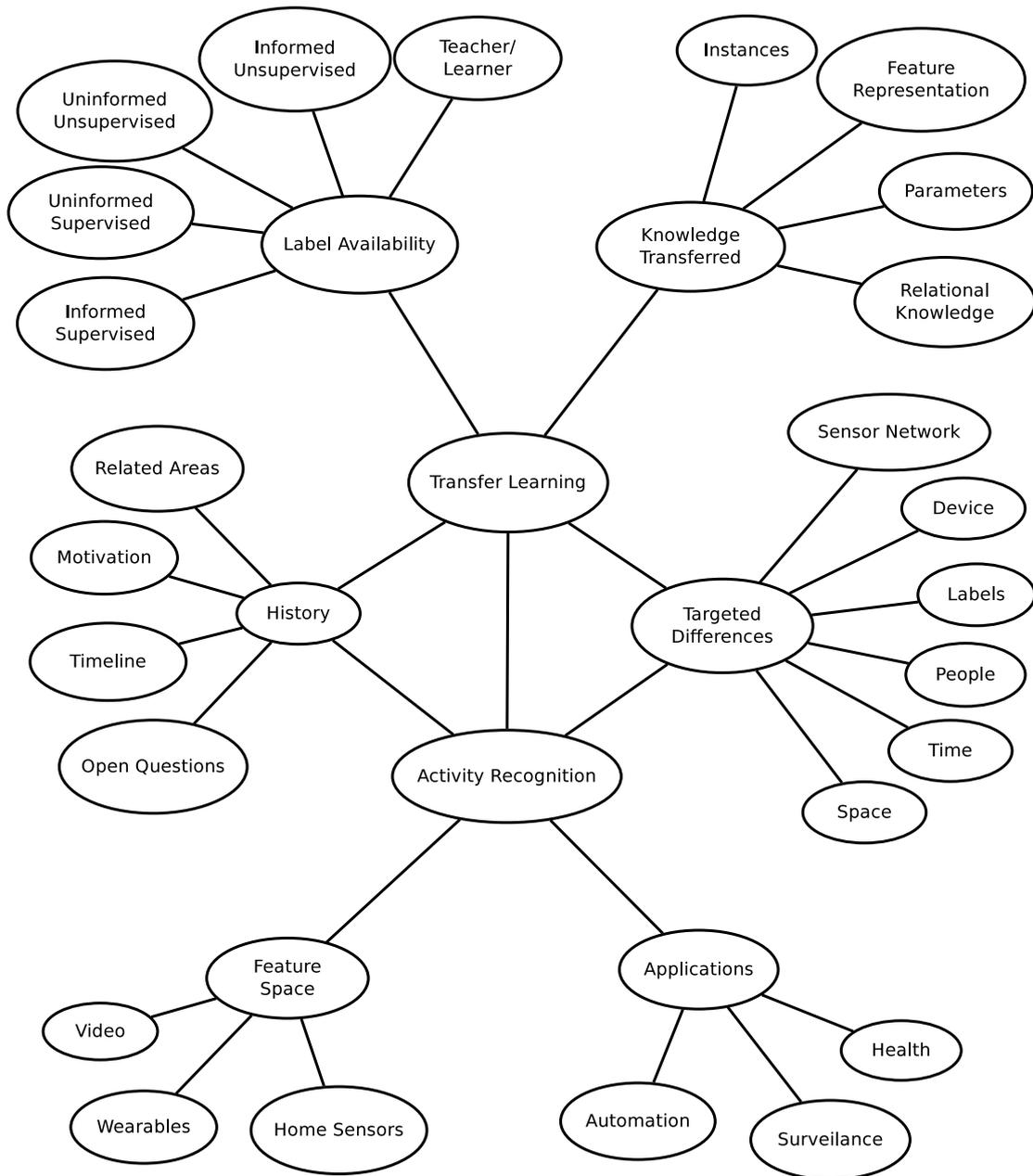
work in these realistic settings. Transfer based activity recognition is one conduit for achieving this.

2.1 Transfer Learning

The ability to identify deep, subtle connections, what we term *transfer learning*, is the hallmark of human intelligence. Byrnes [7] defines transfer learning as the ability to extend what has been learned in one context to new contexts. Thorndike and Woodworth [63] first coined this term as they explored how individuals transfer learned concepts between contexts that share common features. Barnett and Ceci provide a taxonomy of features that influence transfer learning in humans [5].

In the field of machine learning, transfer learning is studied under a variety of different names including learning to learn, life-long learning, knowledge transfer, inductive transfer, context-sensitive learning, and metalearning [3, 19, 64, 65, 70]. It is also closely related to several other areas of machine learning such as self-taught learning, multi-task learning, domain adaptation, and co-variate shift. Because of this broad variance in the terms used to describe transfer learning it is helpful to provide a formal definition of transfer learning terms and of transfer learning itself which will be used throughout the rest of this paper.

2.2 Definitions

This survey starts with a review of basic definitions needed for discussions of transfer learning as it can be applied to activity recognition. Definitions for domain and task have been provided by Pan and Yang [50]:

Definition 1 (Domain). *A domain D is a two-tuple $(\chi, P(X))$. χ is the feature space of D and $P(X)$ is the marginal distribution where $X = \{x_1, ..., x_n\} \in \chi$.*

Definition 2 (Task). *A task T is a two-tuple $(Y, f())$ for some given domain D. Y is the label space of D and $f()$ is an objective predictive function for D. $f()$ is sometimes written as a conditional probability distribution $P(y|x)$. $f()$ is not given, but can be learned from the training data.*

To illustrate these definitions, consider the problem of activity recognition using motion sensors. The domain is defined by a feature space which may represent the n-dimensional space defined by n sensor firing counts within a given time window and a marginal probability distribution over all possible firing counts. The task is composed of a label space y which consists of the set of labels for activities of interest, and a conditional probability distribution consisting of the probability of assigning a label $y_i \in y$ given the observed instance $x \in \chi$.

Using these terms, we can now define transfer learning. In this paper we specify a definition of transfer learning that is similar to that presented by Pan and Yang [50] but we allow for transfer learning which uses multiple source domains.

Definition 3 (Transfer Learning). *Given a set of source domains $DS = D_{s_1}, ..., D_{s_n}$ where $n > 0$, a target domain, D_t, a set of source tasks $TS = T_{s_1}, ...T_{s_n}$ where $T_{s_i} \in TS$ corresponds with $D_{s_i} \in DS$, and a target task T_t which corresponds to D_t, transfer learning helps improve the learning of the target predictive function $f_t()$ in D_t where $D_t \notin DS$ and $T_t \notin TS$.*

This definition of transfer learning is broad and encompasses a large number of different transfer learning scenarios. The source domains can differ from the target domain by having a different feature space, a different distribution of instances in the feature space, or both. The source tasks can differ from the target task by having a different label space, a different predictive function for labels in that label space, or both. The source data can differ from the target data by having a different domain, a different task, or both. However, all transfer learning problems rely on the basic assumption that there exists some relationship between the source and target areas which allows for the successful transfer of knowledge from the source to the target.

2.3   Scenarios

To further illustrate the variety of problems which fall under the scope of transfer-based activity recognition, we provide illustrative scenarios. Not all of these scenarios can be addressed by current transfer learning methods. The first scenario represents a typical transfer learning problem solvable using recently developed techniques. The second scenario represents a more challenging situation that pushes the boundaries of current transfer learning techniques. The third scenario requires a transfer of knowledge across such a large difference between source and target datasets, that current techniques only scratch the surface of what is required to make such a knowledge transfer successful.

2.3.1   Scenario 1

In one home which has been equipped with multiple motion and temperature sensors, an activity recognition algorithm has been trained using months of annotated labels to provide the ground truth for activities which occur in that home. A transfer learning algorithm should be able to reuse the labeled data to perform activity recognition in a new setting. Such transfer will save months of man-hours annotating data for the new home. However, the new home has a different layout as well as a different resident and different sensor locations than the first home.

2.3.2   Scenario 2

An individual with Parkinson's disease visits his neurosurgeon twice a year to get an updated assessment of his gait, tremor, and cognitive health. The medical staff perform some gait measurements and simulated activities in their office space to determine the effectiveness of the prescribed medication, but want to determine if the observed improvement is reflected in the activities the patient performs in his own home. A learning algorithm will need to be able to transfer information between different physical settings, as well as time of day, sensors used, and scope of the activities.

2.3.3   Scenario 3

A researcher is interested in studying the cooking activity patterns of college students living in university dorms in the United States. The research study has to be conducted using the smart phone of the student as the sensing mechanism. The cooking activity of these students typically consists of heating up a frozen snack

from the refrigerator in the microwave oven. In order to build the machine learning models for recognizing these activity patterns, the researcher has access to cooking activities for a group of grandmothers living in India. This dataset was collected using smart home environmental sensors embedded in the kitchen and the cooking activity itself was very elaborate. Thus the learning algorithm is now faced with changes in the data at many layers; namely, differences in the sensing mechanisms, cultural changes, age related differences, different location settings and finally differences in the activity labels. This transfer learning from one setting to another diverse setting is most challenging and requires significant progress in transfer learning domain to even attempt to solve the problem.

These scenarios illustrate different types of transfer that should be possible using machine learning methods for activity recognition. As is described by these situations, transfer may occur across several dimensions. We next take a closer look at these types of transfer and use these descriptors to characterize existing approaches to transfer learning for activity recognition.

2.4 Dimensions of Analysis

Transfer learning can take many forms in the context of activity recognition. In this discussion we consider four dimensions to characterize various approaches to transfer learning for activity recognition. First, we consider different sensor modalities on which transfer learning has been applied. Second, we consider differences between the source and target environments in which data is captured. The third dimension is the amount and type of data labeling that is available in source and target domains. Finally, we examine the representation of the knowledge that is transferred from source to target. The next sections discuss these dimensions in more detail and characterize existing work based on alternative approaches to handling such differences.

3 Modality

One natural method for the classification of transfer learning techniques is the underlying sensing modalities used for activity recognition. Some techniques may be generalizable to different sensor modalities, but most techniques are too specific to be generally applicable to any sensor modality other than that for which they are designed to work with. This is usually because the types of differences that occur between source and target domains are different for each sensor modality. These differences and their effect on the transfer learning technique are discussed in detail in Section 4. In this section we consider only those techniques which have empirically demonstrated their ability to operate on a given sensor modality.

The classification of sensor modalities itself is a difficult problem and indeed creating precise classification topology is outside of the scope of this paper. However, we roughly categorize sensor modalities into the following classifications, video cameras, wearable devices, and ambient sensors. For each sensor modality, we provide a brief description of the types of sensors which are included and a summary of the research works performing transfer learning in that domain. In this section we do not describe the transfer learning algorithms used in the papers as that will be discussed in the other dimensions of analysis.

3.1 Video Sequences

Video cameras are one of the first sensor modalities in which transfer learning has been applied to the area of activity recognition [75]. Video cameras provide a dense feature-space for activity recognition which potentially allows for extremely fine-grained recognition of activities. Spatio-temporal features are extracted from video sequences for characterizing the activities occurring in them. Activity models are then learned using these feature representations.

One drawback of video processing for activity recognition is that the use of video cameras raises more issues associated with user privacy. In addition, cameras need to be well positioned and track individuals in order to capture salient data for processing. Activity recognition via video cameras has received broad attention in transfer learning research [18, 20, 34, 37, 44, 72, 73, 74, 75, 77].

3.2 Wearable sensors

Body Sensor Networks are another commonly used sensing mechanism to capture activity related information from individuals. These sensors are typically worn by the individuals. Strategic placement of the sensors helps in capturing important activity related information such as movements of the upper and lower parts of the body that can then be used to learn activity models. Sensors in this category include, inertial sensors such as accelerometers and gyroscopes, sensors embedded in smart phones, radio frequency identification sensors and tags. Researchers have applied transfer learning techniques to both activity recognition using wearable accelerometers and activity recognition using smartphones but we have not seen any transfer learning approaches applied to activity recognition using RFID tags. This may be due in part to the relatively low use of RFID tags in activity recognition itself.

Within wearable sensors, two types of problems are generally considered. The first is the problem of activity recognition itself [6, 8, 13, 23, 30, 32, 59, 69, 78, 79], and the second is the problem of user localization, which can then be used to increase the accuracy of the activity recognition algorithm [47, 48, 49, 51, 81] . Both problems present interesting challenges for transfer learning.

3.3 Ambient Sensors

Ambient sensors represent the broadest classification of sensor modalities which we define in this paper. We categorize any sensor that is neither wearable nor video camera into ambient sensors. These sensors are typically embedded in an individual's environment. This category includes a wide variety of sensors such as motion detectors, door sensors, object sensors, pressure sensors, and temperature sensors. As the name indicates, these sensors collect a variety of activity related information such as human movements in the environment induced by activities, interactions with objects during the performance of an activity, and changes to illumination, pressure and temperature in the environment due to activities. Researchers have only recently begun to look at transfer learning applications for ambient sensors with the earliest work appearing around 2008 [66]. Since then the field of transfer learning for activity recognition using ambient sensors has progressed rapidly with many different research groups analyzing the problem from several different angles [14, 25, 54, 55, 56, 57, 59, 67, 80].

Table 1: Relationship between formally defined transfer learning differences and the applied meaning for activity recognition.

Formal Definition	Applied Meaning
$\chi_t \neq \chi_{s_i}$ for $0 < i < n$	sensor networks, sensor modality, or physical space
$P(X_t) \neq P(X_{s_i})$ for $0 < i < n$	time, people, devices, or sampling rates
$Y_t \neq Y_{s_i}$ for $0 < i < n$	activities or labels
$f_t(x) \neq f_{s_i}(x)$ for $0 < i < n$	time, people, devices, sampling rates, activities, or labels

3.4 Crossing the sensor boundaries

Clearly, transfer learning within individual sensor modalities is progressing. Researchers are actively developing and applying new techniques to solve a variety of problems within any given sensor modality domain. However, there has been little work done that tries to transfer knowledge between any two or more sensor modalities. Kurz et al. [32] and Roggen et al. [59] address this problem using a teacher/learner model which is discussed further in Section 5. Hu et al. [25] introduce a transfer learning technique for successfully transferring some knowledge across sensor modalities, but greater transfer of knowledge between modalities has yet to be explored.

4 Physical Setting Differences

Another useful categorization of transfer learning techniques is the types of physical differences between a source and target dataset across which the transfer learning techniques can achieve a successful transfer of knowledge. In this section, we describe these differences in a formal setting and provide illustrative examples drawn from activity recognition.

We use the terminology for domain, task and transfer learning defined in Section 2 to describe the differences between source and target datasets. These differences can be in the form of the feature-space representation, the marginal probability distribution of the instances, the label space, and/or the objective predictive function. When describing transfer learning in general, using such broad terms allows one to encompass many different problems. However, when describing transfer learning for a specific application, such as activity recognition, it is convenient to use more application specific terms. For example, differences in the feature-space representation can be thought of in terms of the sensor modalities and sampling rates and differences in the marginal probability distribution can be thought of in terms of different people performing the same activity, or having the activity performed in different physical spaces.

Even when limiting the scope to activity recognition, it is still infeasible to enumerate every possible difference between source and target datasets. In this survey we consider some of the most common or important differences between the source and target datasets including time, people, devices, space, sensor types, and labels. Table 1 summarizes the relationship between each of these applied differences and the formal definitions of transfer learning differences.

Differences across time, people, devices, or sensor sampling rates result in differences in the underlying marginal probability distribution, the objective predictive function, or both. Several papers focus specifically on transferring across time differences [30, 47, 48, 69], differences between people [12, 23, 54, 79], and differences between devices [78, 81].

Differences created when comparing datasets from different spaces or spatial layouts are reflected by differences in the feature-spaces, the marginal probability distributions, the objective predictive functions, or any combination of these. As the number of differences increases, the source and target datasets become less related making transfer learning more difficult. Because of this, current research usually imposes limiting assumptions about what is different between the spaces. Several researchers, for example, assume that some meta-features are added which provide space-independent information [14, 55, 56, 57, 66, 67]. For WiFi localization, Pan et al. [49] assume that the source and target spaces are in the same building. Applying transfer learning to video clips from different spaces usually results in handling issues of background differences [9, 74, 75] and/or issues of camera view angle [37].

Differences in the labels used in the datasets are obviously reflected by differences in the label space and the objective predictive function. Compared to the other differences discussed previously, transferring between differences in the label space has received much less attention in the current literature [25, 34, 72, 77, 80].

One of the largest differences between datasets occurs when the source and target datasets have a different sensor modality. This makes the transfer learning problem much more difficult and relatively little work has been done in this direction. Hu and Yang have started work in this direction in [25]. Additionally, Calatroni et al. [8], Kurz et al. [32] and Roggen et al. [59] take a different approach to transferring across sensor modality by assuming a classifier for the source modalities can act as an expert for training a classifier in the target sensor modality.

5   Data Labeling

In this section we consider the problem of transfer learning from the perspective of the availability of labeled data. Traditional machine learning uses the terms *supervised learning* and *unsupervised learning* to distinguish learning techniques based on the availability and use of labeled data. To distinguish between source and target labeled data availability we introduce two new terms, *informed* and *uninformed*, which we apply to the availability of labeled data in the target area. Thus, informed supervised (IS) transfer learning implies that some labeled data is available in both the target and source domains. Uninformed supervised (US) transfer learning implies that labeled data is available only in the source domain. Informed unsupervised (IU) transfer learning implies that labeled data is only available in the target domain. Finally, uninformed unsupervised (UU) transfer learning implies that no labeled data is available for either the source or target domains. One final case to consider is teacher/learner (TL) transfer learning, where no training data is directly available. Instead a previously-trained classifier (the Teacher) is introduced which operates simultaneously with the new classifier to be trained (the Learner) and provides the labels for observed data instances.

Two other terms that are often used in machine learning literature and may be

Table 2: General relationship between inductive/transductive learning and the availability of labeled data.

Label Availability	Most Common Approach
Informed Supervised	Inductive Learning
Informed Unsupervised	Inductive Learning
Uninformed Supervised	Transductive Learning
Uninformed Unsupervised	Unsupervised Learning

applicable here are *inductive* and *transductive* learning. Inductive learning refers to learning techniques which try to learn the objective predictive function. Transductive learning techniques, on the other hand, try to learn the relationship between instances. Pan and Yang [50] extend the definitions of inductive and transductive learning to transfer learning, but the definitions do not create a complete taxonomy for transfer learning techniques. For this reason, we do not specifically classify recent works as being inductive or transductive in nature, but we note here how the inductive and transductive definitions fit into a classification based upon the availability of labeled data.

Inductive learning requires that labeled data be available in the target domain regardless of its availability in the source domain. Thus, most informed supervised and informed unsupervised transfer learning techniques are also inductive transfer learning techniques. Transductive learning, however, does not require labeled data in the target domain. Therefore, most uninformed supervised techniques are also transductive transfer learning techniques. Table 2 summarizes this general relationship.

Several researchers have developed and applied informed, supervised transfer learning techniques for activity recognition. These techniques have been applied to activity recognition using wearables [6, 29, 46, 48, 69, 81] and to activity recognition using cameras [18, 34, 44, 74, 75, 77].

Research into transfer-based activity recognition using ambient sensors has almost exclusively focused on uninformed supervised transfer learning [14, 25, 26, 54, 66, 67, 69, 80], but a few algorithms are able to take advantage of the labeled target data if it is available [55, 56, 57]. This focus on uninformed supervised transfer learning is most likely due to the allurement of building an activity recognition framework that can be trained offline and later installed into any user's space without requiring additional data labeling effort. Wearables have also been used for uninformed supervised transfer learning research [23, 47, 49, 51, 76, 78, 79] as have cameras[9, 20, 37, 72, 73].

Despite the abundance of research using labeled source data, research into transfer learning techniques for activity recognition in which no source labels are available is extremely sparse. Pan et al. [47] have applied an uninformed unsupervised technique, transfer component analysis (TCA) to reduce the distance between domains by learning some transfer components across domains in a reproducing kernel Hilbert space using maximum mean discrepancy. We are unaware of any other work for uninformed unsupervised transfer-based activity recognition. We are also unaware of any work on informed unsupervised transfer-based activity recognition. The lack of research into informed unsupervised transfer-based activity recognition is not surprising because the idea of having labeled target data available and

10

not having labeled source data is counterintuitive to the general principle of transfer learning. However, informed unsupervised transfer learning may still provide significant benefits to activity recognition.

The teacher/learner model for activity recognition is considerably less studied than the previously discussed techniques. However, we feel that this area has significant promise for improving transfer learning for activity recognition and making activity recognition systems much more robust and versatile. Roggen et al. [59], Kurz et al. [32], and Calatroni et al. [8] apply the teacher/learner model to develop an opportunistic system which is capable of using whatever sensors are currently contained in the environment to perform activity recognition.

In order for the teacher/learner model to be applicable, two requirements must be met. First, an existing classifier (the teacher) must already be trained in the source domain. Second, the teacher must operate simultaneously with a new classifier in the target domain (the learner) to provide the training for the learner. For example, Roggen et al. [59] equip a cabinet of drawers with an accelerometer for each drawer and then a classifier is trained to recognize which drawer of the cabinet is being opened or closed. This classifier becomes the teacher. Then several wearable accelerometers are attached to the person opening and closing the drawers. Now, a new classifier is trained using the wearable accelerometers. This classifier is the learner. When the individual opens or closes a drawer, the teacher labels the activity according to its classification model. This label is given to the learner which can then be used as labeled training data in real-time without the need to supply any manually labeled data.

The teacher/learner model presents a new perspective on transfer learning and introduces additional challenges. One major challenge of the teacher/learner model is that the accuracy of the learner is limited by the accuracy of the teacher. Additionally, the system's only source of a ground truth comes from the teacher and thus the learner is completely reliant upon the teacher. It remains to be explored whether the learner can ever outperform the teacher and if it does so, whether it can convince itself and others of this superior performance. Finally, while the teacher/learner model provides a convenient way to transfer across different domains, an additional transfer mechanism would need to be employed to transfer across different label spaces.

6   Type of Knowledge Transferred

Pan and Yang [50] describe four general classifications for transfer learning in relation to what is transferred, instance transfer, feature-representation transfer, parameter transfer, and relational-knowledge transfer.

6.1   Instance Transfer

Instance transfer reuses the source data to train the target classifier, usually by re-weighting the source instances based upon a given metric. Instance transfer techniques work well when $\chi_s = \chi_t$ i.e., the feature space describing the source and target domains are same. They may also be applied after the feature representation has first been transferred to a common representation between the source and target domains.

Several researchers have applied instance transfer techniques to activity recognition. Hachiya et al. [23] develop an importance weighted least-squares probabilistic classification approach to handle transfer learning when $P(X_s) \neq P(X_t)$ (i.e., the co-variate shift problem) and apply this approach to wearable accelerometers. Venkatesan et al. [29, 68, 69] extend the AdaBoost framework proposed by Freund and Schapire [21] to include cost-sensitive boosting which tries to weight samples from the source domain according to their relevance in the target domain. In their approach, samples from the source domain are first given a relevance cost. As the classifier is trained, those instances from the source domain with a high relevance must also be classified correctly. Xian-ming and Shao-zi apply TrAdaBoost (a different transfer learning extension of AdaBoost) [15] to action recognition in video clips [74] . Lam et al. weight the source and target data differently when training an SVM to recognize target actions from video clips [34]. Training a typical SVM involves solving the following optimization problem:

$$\min_{\vec{w}, \xi} \{ \frac{1}{2} ||\vec{w}||^2 + C \sum_{i=1}^{n} \xi_i \}$$
$$\text{s.t. } y_i(\vec{x_i} \cdot \vec{w} + b) - 1 + \xi_i \geq 0, \xi_i \geq 0$$

(1)

where $\vec{x_i}$ is the ith datapoint and y_i, ξ_i are the label and slack variable associated with $\vec{x_i}$. $\vec{w}$ is the normal to the hyperplane. C is the parameter that trades off between training accuracy and margin size. However, to allow for the different source and target weights, they solve the following optimization:

$$\min_{\vec{w}, \xi} \{ \frac{1}{2} ||\vec{w}||^2 + C_s \sum_{i=1}^{n} \xi_i + C_t \sum_{i=n+1}^{n+m} \xi_i \}$$
$$\text{s.t. } y_i(\vec{x_i} \cdot \vec{w} + b) - 1 + \xi_i \geq 0, \xi_i \geq 0$$

(2)

where the parameters are the same as before except the first n datapoints are from the source data and the last m datapoints are from the target data.

Unlike the previous instance-based approaches which weight the source instances based on similarity of features between the source and target data, Zheng et al. [80] use an instance-based approach to weight source instances based upon the similarity between the label information of the source and target data. This allows them to transfer the labels from instances in the source domain to instances in the target domain using web-knowledge to relate the two domains [25, 26]. Taking a different approach, several researchers [8, 32, 59] use the real-time teacher/learner model discussed in the previous section to transfer the label of the current instance in the source domain to the instance in the target domain.

6.2   Feature-Representation Transfer

Feature-representation transfer reduces the differences between the source and target feature spaces. This can be accomplished by mapping the source feature space to the target feature space such as $f : \chi_s \rightarrow \chi_t$, by mapping the target feature space to the source feature space such as $g : \chi_t \rightarrow \chi_s$, or by mapping both the source and target feature spaces to a common feature space such as $g : \chi_t \rightarrow \chi$ and

$f : \mathcal{X}_s \to \mathcal{X}$. This mapping can be computed manually [66] or learned as part of the transfer learning algorithm[18, 25, 37, 54, 81].

When the mapping is part of the transfer learning algorithm a common approach is to apply a dimensionality reduction technique to map both source and target feature-space into a common latent space[46, 47, 48, 51]. For example, Chattopadhyay et al. [12] use Isomap [62] to map both the source and target data into a common low-dimensional space after which instance-based transfer techniques can be applied.

In some cases, meta-features are first manually introduced into the feature space and then the feature space is automatically mapped from the source domain to the target domain [6, 14, 67]. An example of this is the work of Rashidi and Cook [57]. They first assign a location label to each sensor indicating in which room or functional area the sensor is located. Then activity templates are constructed from the data for both the source and target data, finally a mapping is learned between the source and target datasets based upon the similarity of activities and sensors [55, 56].

6.3   Parameter Transfer

Parameter transfer learns parameters which are shared between the source and target tasks. One common use of parameter transfer is learning a prior distribution shared between the source and target datasets. For example, one technique [9] models the source and target tasks using a Gaussian Mixture Model which share a prior distribution, another algorithm [18] learns a target classifier using a set of pre-trained classifiers as prior for the target classifier, and van Kasteren et al. [66] propose a method to learn the parameters of a Hidden Markov Model using labeled data from the source domain, and unlabeled data from the target domain. Later they extend this work to learn hyperparameter priors for the HMM instead of learning the parameters directly [67].

Another common example of parameter transfer assumes the SVM parameter w can be split into two terms: w_0, which is the same for both the source and target tasks, and v, which is specific to the particular task. Thus $w_s = w_0 + v_s$ and $w_t = w_0 + v_t$. Several works adopt this approach [44, 75].

Using a different approach to parameter transfer, a transfer learning algorithm [49, 51] can extract knowledge from the source domain to impose additional constraints on a quadratically-constrained quadratic program optimization problem for the target domain. Along a similar line of thought, Zhao et al. [78, 79] use information extracted from the source domain to initialize cluster centers for a k-means algorithm in the target domain.

6.4   Relational-Knowledge Transfer

Relational-knowledge transfer applies to problems in which the data is not independent and identically distributed (i.i.d.) as is traditionally assumed but can be represented through multiple relationships [50]. Such problems are usually represented with a network or graph. Relational-knowledge transfer tries to transfer the relationships of in the source domain to the target domain. This type of transfer learning is not heavily explored, and as far as we are able to determine, no research is currently being pursued in transfer learning for activity recognition using

relational-knowledge transfer.

7   Summary

The previous sections analyzed a large body of transfer-based activity recognition research along four different dimensions. Looking at each dimension separately provides an orderly way to analyze so many different papers. However, such separation may also make it difficult to see the bigger picture. Table 3, therefore, summarizes the classification of existing works along these four dimensions.

Table 3: Summarization of existing work based on the four dimension of analysis.

Paper	Sensor Modality	Difference	Labeling	Type of Knowledge Transfer
[6]	wearables	new activities and labels	IS	feature-representation
[8]	wearables	different device, placement	TL	instance-based
[9]	video camera	background, lighting, noise, and people	IS, US	parameter-based
[12]	wearables	people	IS	feature-representation and instance-based
[13]	wearables	people	IS	parameter-based
[14]	ambient sensors	location, layout, people	US	feature-representation
[18]	video camera	web-domain vs consumer domain.	IS	feature-representation and parameter-based
[20]	video camera	view angle	US	feature-space
[23]	wearables	people	US	instance-based
[25]	ambient sensors, wearables	label space, location	US	instance-based and feature-representation
[26]	ambient sensors, wearables	label space	US	instance-based
[29]	wearables	people and setting	IS	instance-based
[32]	wearables	sensors	TL	instance-based
[34]	video camera	labels	IS	instance-based
[37]	video camera	view angle	US	feature-representation
[44]	video camera	activity sets, labels	IS	parameter-based
[46]	wearables	time	IS	feature-representation
[47]	wearables	time	US, UU	feature-representation
[48]	wearables	time	IS	feature-representation
[49]	wearables	space, location	US	parameter-based
[51]	wearables	space, time, device	IS, US	feature-representation and parameter-based
[54]	ambient sensors	people	US	feature-representation
[55]	ambient sensors	layout, sensor network	IS, US	feature-representation
[56]	ambient sensors	layout, sensor network	IS, US	feature-representation
[57]	ambient sensors	layout, sensor network, people	IS, US	feature-representation
[59]	ambient sensors, wearables	devices	TL	instance-based

Table 3: (continued from the previous page.)

Paper	Sensor Modality	Difference	Labeling	Type of Knowledge Transfer
[66]	ambient sensors	location	US	feature-representation and parameter-based
[67]	ambient sensors	location	US	feature-representation and parameter-based
[68]	wearables	people, setting	IS	instance-based
[69]	wearables	people, setting	IS	instance-based
[72]	video camera	labels	US	feature-representation
[73]	video camera	view angle	US	parameter
[74]	video camera	background, people	IS	instance
[75]	video camera	background, video domain	IS	parameter-based
[76]	wearables	space, time, device	IS, US	feature-representation and parameter-based
[77]	video camera	activities performed	IS	distance function
[78]	wearables	mobile device, sampling rate	US	parameter-based
[79]	wearables	people	US	parameter-based
[80]	ambient sensors, wearables	activity labels	US	instance-based
[81]	wearables	devices	IS	feature-representation

8   Grand Challenges

Although transfer-based activity recognition has progressed significantly in the last few years, there are still many open challenges. In this section, we first consider challenges specific to a particular sensor modality and then we look at challenges which are generalizable to all transfer-based activity recognition.

As can be seen in Table 5, performing transfer-based activity recognition when the source data is not labeled has not received much attention in current research. Outside the domain of activity recognition, researchers have leveraged the unlabeled source data to improve transfer in the target domain [16, 53, 71] but such techniques

Table 4: Existing work categorized by sensor modality and the differences between the source and target datasets.

Sensor Modality	$\chi_s \neq \chi_t$	$P(X_s) \neq P(X_t)$	$Y_s \neq Y_t$	$f_s(x) \neq f_t(x)$
Video	[18, 20, 37, 72, 73]	[9, 18, 74, 75, 73]	[34, 44, 77]	[9, 18, 34, 37, 44, 74, 75, 77]
Wearable	[8, 32, 51, 59, 76, 78, 81]	[13, 23, 29, 46, 47, 48, 49, 51, 68, 69, 76]	[6, 25, 26, 80]	[6, 23, 25, 26, 29, 46, 47, 48, 49, 51, 69, 76, 80]
Ambient	[14, 55, 56, 57, 59, 66, 67]	[14, 54, 55, 56, 57, 66, 67]	[25, 26, 80]	[14, 25, 26, 54, 55, 56, 57, 66, 67, 80]

Table 5: Existing work categorized by sensor modality and data labeling.

Sensor Modality	Informed Supervised	Uninformed Supervised	Informed Unsupervised	Uninformed Unsupervised
Video	[9, 18, 34, 44, 74, 75, 77]	[9, 20, 37, 72, 73]	-	-
Wearable	[6, 13, 46, 48, 51, 68, 69, 76, 81]	[23, 25, 26, 29, 47, 49, 51, 76, 78, 79, 80]	-	[47]
Ambient	[55, 56, 57]	[14, 25, 26, 54, 55, 56, 57, 66, 67, 80]	-	-

Table 6: Existing work categorized by sensor modality and the type of knowledge transferred.

Sensor Modality	Instance Based	Feature Representation	Parameter Based	Relational Knowledge
Video	[34, 74]	[18, 20, 37, 72]	[9, 18, 44, 73, 75, 77]	-
Wearable	[8, 23, 25, 26, 29, 32, 59, 68, 69, 80]	[6, 46, 47, 48, 51, 76, 81]	[13, 49, 51, 76, 78, 79]	-
Ambient	[25, 26, 59, 80]	[14, 25, 54, 55, 56, 57, 66, 67]	[66, 67]	-

have yet to be applied to activity recognition.

Another area needing more attention is relational-knowledge transfer for activity recognition as indicated in Table 6. Relational-knowledge transfer requires that there exist certain relationships in the data which can be learned and transfered across populations. Data for activity recognition has the potential to contain such transferable relationships indicating that this may be an important technique to pursue. See [17, 41, 42, 43] for examples of relational-knowledge transfer.

Tables 4-6 also indicate several more niche areas which could be further investigated. For example, in the video camera domain, most of the work has focused on informed supervised parameter-based transfer learning, while the other techniques have not been heavily applied. Similarly, transferring across different labels spaces is a much less studied problem in transfer-based activity recognition. Finally, we note that parameter-based transfer learning is also less studied for the ambient sensor modality.

The current direction of most transfer-based activity recognition is to push the limits on how different the source and target domains and tasks can be. The scenarios discussed in Section 2 illustrate the importance of continuing in this direction. More work is needed to improve transfer across sensor modalities and to transfer knowledge across multiple differences. Instead of transferring learning from one smart home environment to another, can we transfer from a smart home to a smart workplace or smart hospital? We envision one day chaining multiple transfers to-

gether to achieve even greater diversity between the source and target populations.

As researchers continue to expand the applicability of transfer learning, two natural questions arise. First, can we define a generalizable distance metric for determining the difference between the source and target populations? Some domain-specific distances have been used in the past, but it would be useful if we had a domain-independent distance measure. This measure could be used to facilitate comparisons between different transfer learning approaches as well as provide an indication of whether transfer learning should even be applied in a given situations. Such a measure would need to indicate how the source and target data differ (feature-space, marginal probabilities, label-space, and, objective predictive function) as well as quantify the magnitude of the differences. Second, can we detect and prevent the occurrence of negative transfer effects. Negative transfer effects occur when the use of transfer learning actually decreases performance instead of increasing performance. These two questions are actually related, because an accurate distance metric may provide an indication of when negative transfer will occur for a given transfer learning technique. Rosenstein et al. look at the question of when to use transfer learning in [60]. They empirically show that when two tasks are of sufficient dissimilarity, negative transfer occurs. Mahmud and Ray define a distance metric for measuring the similarity of two tasks based on the conditional Kolmogorov complexity between the tasks and prove some theoretical bounds using this distance measure [39].

This survey has reviewed the current literature regarding transfer based activity recognition. We discussed several promising techniques and consider the many open challenges that still need to be addressed in order to progress the field of transfer learning for activity recognition.

References

[1] Rakesh Agrawal and Ramakrishnan Srikant. Mining sequential patterns. In *Proceedings of the International Conference on Data Engineering*, pages 3–14, 1995.

[2] Hande Alemdar and Cem Ersoy. Wireless sensor networks for healthcare: A survey. *Computer Networks*, 54(15):2688 – 2710, 2010.

[3] A. Arnold, R. Nallapati, and W.W. Cohen. A comparative study of methods for transductive transfer learning. In *Data Mining Workshops, 2007. ICDM Workshops 2007. Seventh IEEE International Conference on*, pages 77 –82, oct. 2007.

[4] Akin Avci, Stephan Bosch, Mihai Marin-Perianu, Raluca Marin-Perianu, and Paul Havinga. Activity recognition using inertial sensing for healthcare, well-being and sports applications: A survey. In *Architecture of Computing Systems (ARCS), 2010 23rd International Conference on*, pages 1 –10, Feb. 2010.

[5] S.M. Barnett and S.J. Ceci. When and where do we apply what we learn?: A taxonomy for far transfer. *Psychological bulletin*, 128(4):612–637, 2002.

[6] U. Blanke and B. Schiele. Remember and transfer what you have learned-recognizing composite activities based on activity spotting. In *Wearable Computers (ISWC), 2010 International Symposium on*, pages 1–8. IEEE, 2010.

[7] J.P. Byrnes. *Cognitive development and learning in instructional contexts.* Allyn and Bacon, Boston, 1996.

[8] Alberto Calatroni, Daniel Roggen, and Gerhard Tröster. Automatic transfer of activity recognition capabilities between body-worn motion sensors: Training newcomers to recognize locomotion. In *Eighth International Conference on Networked Sensing Systems (INSS'11)*, Penghu, Taiwan, June 2011.

[9] Liangliang Cao, Zicheng Liu, and T.S. Huang. Cross-dataset action detection. In *Computer Vision and Pattern Recognition (CVPR), 2010 IEEE Conference on*, pages 1998 –2005, June 2010.

[10] Marie Chan, Daniel Estve, Christophe Escriba, and Eric Campo. A review of smart homespresent state and future challenges. *Computer Methods and Programs in Biomedicine*, 91(1):55 – 81, 2008.

[11] Chih-Chung Chang and Chih-Jen Lin. Libsvm: A library for support vector machines. *ACM Transactions on Intelligent Systems and Technology*, 2(3):27:1–27:27, May 2011.

[12] R. Chattopadhyay, N.C. Krishnan, and S. Panchanathan. Topology preserving domain adaptation for addressing subject based variability in semg signal. In *2011 AAAI Spring Symposium Series*, 2011.

[13] H.L. Chieu, W.S. Lee, and L.P. Kaelbling. Activity recognition from physiological data using conditional random fields. Technical report, Singapore-MIT Alliance (SMA), 2006.

[14] D. Cook. Learning setting-generalized activity models for smart spaces. *Intelligent Systems, IEEE*, PP(99):1, 2010.

[15] Wenyuan Dai, Qiang Yang, Gui-Rong Xue, and Yong Yu. Boosting for transfer learning. In *Proceedings of the 24th international conference on Machine learning*, ICML '07, pages 193–200, New York, NY, USA, 2007. ACM.

[16] Wenyuan Dai, Qiang Yang, Gui-Rong Xue, and Yong Yu. Self-taught clustering. In *Proceedings of the 25th international conference on Machine learning*, ICML '08, pages 200–207, New York, NY, USA, 2008. ACM.

[17] Jesse Davis and Pedro Domingos. Deep transfer via second-order markov logic. In *Proceedings of the 26th Annual International Conference on Machine Learning*, ICML '09, pages 217–224, New York, NY, USA, 2009. ACM.

[18] Lixin Duan, Dong Xu, I.W. Tsang, and Jiebo Luo. Visual event recognition in videos by learning from web data. In *Computer Vision and Pattern Recognition (CVPR), 2010 IEEE Conference on*, pages 1959 –1966, June 2010.

[19] Charles Elkan. The foundations of cost-sensitive learning. In *Proceedings of the 17th international joint conference on Artificial intelligence - Volume 2*, IJCAI'01, pages 973–978, San Francisco, CA, USA, 2001. Morgan Kaufmann Publishers Inc.

[20] Ali Farhadi and Mostafa Tabrizi. Learning to recognize activities from the wrong view point. In David Forsyth, Philip Torr, and Andrew Zisserman, editors, *Computer Vision ECCV 2008*, volume 5302 of *Lecture Notes in Computer Science*, pages 154–166. Springer Berlin / Heidelberg, 2008.

[21] Yoav Freund and Robert E Schapire. A decision-theoretic generalization of on-line learning and an application to boosting. *Journal of Computer and System Sciences*, 55(1):119 – 139, 1997.

[22] Tao Gu, Shaxun Chen, Xianping Tao, and Jian Lu. An unsupervised approach to activity recognition and segmentation based on object-use fingerprints. *Data and Knowledge Engineering*, 69(6):533–544, June 2010.

[23] Hirotaka Hachiya, Masashi Sugiyama, and Naonori Ueda. Importance-weighted least-squares probabilistic classifier for covariate shift adaptation with application to human activity recognition. *Neurocomputing*, 80(0):93 – 101, 2012. Special Issue on Machine Learning for Signal Processing 2010.

[24] K.Z. Haigh and H. Yanco. Automation as caregiver: A survey of issues and technologies. In *AAAI-02 Workshop on Automation as Caregiver: The Role of Intelligent Technology in Elder Care*, pages 39–53, 2002.

[25] D.H. Hu and Q. Yang. Transfer learning for activity recognition via sensor mapping. In *Twenty-Second International Joint Conference on Artificial Intelligence*, 2011.

[26] D.H. Hu, V.W. Zheng, and Q. Yang. Cross-domain activity recognition via transfer learning. *Pervasive and Mobile Computing*, 7(3):344 – 358, 2010.

[27] T. L. Kasteren, G. Englebienne, and B. J. Kröse. An activity monitoring system for elderly care using generative and discriminative models. *Personal and Ubiquitous Computing*, 14(6):489–498, Sept. 2010.

[28] Eunju Kim, Sumi Helal, and D. Cook. Human activity recognition and pattern discovery. *Pervasive Computing, IEEE*, 9(1):48–53, 2010.

[29] N.C. Krishnan. *A Computational Framework for Wearable Accelerometer-Based.* PhD thesis, Arizona State University, 2010.

[30] N.C. Krishnan, P. Lade, and S. Panchanathan. Activity gesture spotting using a threshold model based on adaptive boosting. In *Multimedia and Expo (ICME), 2010 IEEE International Conference on*, pages 155 –160, July 2010.

[31] N.C. Krishnan and S. Panchanathan. Analysis of low resolution accelerometer data for continuous human activity recognition. In *Acoustics, Speech and Signal Processing, 2008. ICASSP 2008. IEEE International Conference on*, pages 3337 –3340, April 2008.

[32] M. Kurz, G. Hölzl, A. Ferscha, A. Calatroni, D. Roggen, and G. Tröster. Real-time transfer and evaluation of activity recognition capabilities in an opportunistic system. In *ADAPTIVE 2011, The Third International Conference on Adaptive and Self-Adaptive Systems and Applications*, pages 73–78, 2011.

[33] Jennifer R. Kwapisz, Gary M. Weiss, and Samuel A. Moore. Activity recognition using cell phone accelerometers. In *Proceedings of the Fourth International Workshop on Knowledge Discovery from Sensor Data*, pages 10–18, 2010.

[34] Antony Lam, Amit Roy-Chowdhury, and Christian Shelton. Interactive event search through transfer learning. In Ron Kimmel, Reinhard Klette, and Akihiro Sugimoto, editors, *Computer Vision - ACCV 2010*, volume 6494 of *Lecture Notes in Computer Science*, pages 157–170. Springer Berlin / Heidelberg, 2011.

[35] Jonathan Lester, Tanzeem Choudhury, Nicky Kern, Gaetano Borriello, and Blake Hannaford. A hybrid discriminative/generative approach for modeling human activities. In *Proceedings of the 19th international joint conference on Artificial intelligence*, pages 766–772, San Francisco, CA, USA, 2005. Morgan Kaufmann Publishers Inc.

[36] Lin Liao, Dieter Fox, and Henry Kautz. Location-based activity recognition using relational markov networks. In *Proceedings of the 19th International Joint Conference on Artificial Intelligence*, pages 773–778, San Francisco, CA, USA, 2005. Morgan Kaufmann Publishers Inc.

[37] Jingen Liu, M. Shah, B. Kuipers, and S. Savarese. Cross-view action recognition via view knowledge transfer. In *Computer Vision and Pattern Recognition (CVPR), 2011 IEEE Conference on*, pages 3209 –3216, june 2011.

[38] Beth Logan, Jennifer Healey, Matthai Philipose, Emmanuel Munguia Tapia, and Stephen Intille. A long-term evaluation of sensing modalities for activity recognition. In *Proceedings of the 9th international conference on Ubiquitous computing*, pages 483–500, Berlin, Heidelberg, 2007. Springer-Verlag.

[39] M. M. Mahmud and Sylvian Ray. Transfer learning using kolmogorov complexity: Basic theory and empirical evaluations. In J.C. Platt, D. Koller, Y. Singer, and S. Roweis, editors, *Advances in Neural Information Processing Systems 20*, pages 985–992. MIT Press, Cambridge, MA, 2008.

[40] U. Maurer, A. Smailagic, D.P. Siewiorek, and M. Deisher. Activity recognition and monitoring using multiple sensors on different body positions. In *International Workshop on Wearable and Implantable Body Sensor Networks*, April 2006.

[41] L. Mihalkova, T. Huynh, and R.J. Mooney. Mapping and revising markov logic networks for transfer learning. In *Proceedings of the national conference on artificial intelligence*, volume 22, page 608. Menlo Park, CA; Cambridge, MA; London; AAAI Press; MIT Press; 1999, 2007.

[42] L. Mihalkova and R.J. Mooney. Transfer learning by mapping with minimal target data. In *Proceedings of the AAAI-08 Workshop on Transfer Learning for Complex Tasks*, July 2008.

[43] Lilyana Mihalkova and Raymond J. Mooney. Transfer learning from minimal target data by mapping across relational domains. In *Proceedings of the 21st international jont conference on Artifical intelligence*, IJCAI'09, pages 1163–1168, San Francisco, CA, USA, 2009. Morgan Kaufmann Publishers Inc.

[44] Fabian Nater, Tatiana Tommasi, Helmut Grabner, Luc van Gool, and Barbara Caputo. Transferring activities: Updating human behavior analysis *(both first authors contributed equally)*. In *ICCV WS on Visual Surveillance*, 2011.

[45] Paulito Palmes, Hung Keng Pung, Tao Gu, Wenwei Xue, and Shaxun Chen. Object relevance weight pattern mining for activity recognition and segmentation. *Pervasive and Mobile Computing*, 6(1):43–57, Feb. 2010.

[46] J.J. Pan, Q. Yang, H. Chang, and D.Y. Yeung. A manifold regularization approach to calibration reduction for sensor-network based tracking. In *Proceedings of the National Conference on Artificial Intelligence*, volume 21, page 988, 2006.

[47] S. J. Pan, I. W. Tsang, J. T. Kwok, and Q. Yang. Domain adaptation via transfer component analysis. *IEEE Transactions on Neural Networks*, 22(2):199–210, 2011.

[48] S.J. Pan, J.T. Kwok, Q. Yang, and J.J. Pan. Adaptive localization in a dynamic wifi environment through multi-view learning. In *Proceedings of the National Conference on Artificial Intelligence*, volume 22, page 1108, 2007.

[49] S.J. Pan, D. Shen, Q. Yang, and J.T. Kwok. Transferring localization models across space. In *Proceedings of the 23rd national conference on Artificial intelligence*, volume 3, pages 1383–1388, 2008.

[50] S.J. Pan and Q. Yang. A survey on transfer learning. *Knowledge and Data Engineering, IEEE Transactions on*, 22(10):1345–1359, 2010.

[51] S.J. Pan, V.W. Zheng, Q. Yang, and D.H. Hu. Transfer learning for wifi-based indoor localization. In *Association for the Advancement of Artificial Intelligence (AAAI) Workshop*, page 6, 2008.

[52] Matthai Philipose, Kenneth P, Mike Perkowitz, Donald J. Patterson, Dieter Fox, Henry Kautz, and Dirk Hhnel. Inferring activities from interactions with objects. *IEEE Pervasive Computing*, 3:50–57, 2004.

[53] Rajat Raina, Alexis Battle, Honglak Lee, Benjamin Packer, and Andrew Y. Ng. Self-taught learning: transfer learning from unlabeled data. In *Proceedings of the 24th international conference on Machine learning*, ICML '07, pages 759–766, New York, NY, USA, 2007. ACM.

[54] P. Rashidi and D.J. Cook. Transferring learned activities in smart environments. In *5th International Conference on Intelligent Environments*, volume 2, pages 185–192, 2009.

[55] P. Rashidi and D.J. Cook. Activity recognition based on home to home transfer learning. In *AAAI Workshop on Plan, Activity, and Intent Recognition*, 2010.

[56] P. Rashidi and D.J. Cook. Multi home transfer learning for resident activity discovery and recognition. In *KDD Knowledge Discovery from Sensor Data*, pages 56–63, 2010.

[57] P. Rashidi and D.J. Cook. Activity knowledge transfer in smart environments. *Pervasive and Mobile Computing*, 7(3):331–343, 2011.

[58] P. Rashidi, D.J. Cook, L.B. Holder, and M. Schmitter-Edgecombe. Discovering activities to recognize and track in a smart environment. *IEEE Transactions on Knowledge and Data Engineering*, 23(4):527 –539, April 2011.

[59] Daniel Roggen, Kilian Frster, Alberto Calatroni, and Gerhard Trster. The adarc pattern analysis architecture for adaptive human activity recognition systems. *Journal of Ambient Intelligence and Humanized Computing*, pages 1–18. 10.1007/s12652-011-0064-0.

[60] Michael T. Rosenstein, Zvika Marx, Leslie Pack Kaelbling, and Thomas G. Dietterich. To transfer or not to transfer. In *In NIPS05 Workshop, Inductive Transfer: 10 Years Later*, 2005.

[61] M.E. Taylor and P. Stone. Transfer learning for reinforcement learning domains: A survey. *The Journal of Machine Learning Research*, 10:1633–1685, 2009.

[62] Joshua B. Tenenbaum, Vin de Silva, and John C. Langford. A global geometric framework for nonlinear dimensionality reduction. *Science*, 290(5500):2319–2323, 2000.

[63] E. Thorndike and R.S. Woodworth. The influence of improvement in one mental function upon the efficiency of other functions.(i). *Psychological review*, 8(3):247–261, 1901.

[64] S. Thrun. *Explanation-based neural network learning: A lifelong learning approach.* Kluwer Academic Publishers, 1996.

[65] S. Thrun and L. Pratt. *Learning to learn.* Kluwer Academic Publishers, 1998.

[66] T. van Kasteren, G. Englebienne, and B. Kröse. Recognizing activities in multiple contexts using transfer learning. In *AAAI AI in Eldercare Symposium*, 2008.

[67] T. van Kasteren, G. Englebienne, and B. Krse. Transferring knowledge of activity recognition across sensor networks. In Patrik Floren, Antonio Krger, and Mirjana Spasojevic, editors, *Pervasive Computing*, volume 6030 of *Lecture Notes in Computer Science*, pages 283–300. Springer Berlin / Heidelberg, 2010.

[68] A. Venkatesan. *A Study of Boosting based Transfer Learning for Activity and Gesture Recognition.* PhD thesis, Arizona State University, 2011.

[69] A. Venkatesan, N.C. Krishnan, and S. Panchanathan. Cost-sensitive boosting for concept drift. In *International Workshop on Handling Concept Drift in Adaptive Information Systems 2010*, pages 41–47, 2010.

[70] Ricardo Vilalta and Youssef Drissi. A perspective view and survey of meta-learning. *Artificial Intelligence Review*, 18:77–95, 2002.

[71] Zheng Wang, Yangqiu Song, and Changshui Zhang. Transferred dimensionality reduction. In Walter Daelemans, Bart Goethals, and Katharina Morik, editors, *Machine Learning and Knowledge Discovery in Databases*, volume 5212 of *Lecture Notes in Computer Science*, pages 550–565. Springer Berlin / Heidelberg, 2008.

[72] B. Wei and C. Pal. Heterogeneous transfer learning with rbms. In *Twenty-Fifth AAAI Conference on Artificial Intelligence*, 2011.

[73] Chen Wu, Amir Hossein Khalili, and Hamid Aghajan. Multiview activity recognition in smart homes with spatio-temporal features. In *Proceedings of the Fourth ACM/IEEE International Conference on Distributed Smart Cameras*, ICDSC '10, pages 142–149, New York, NY, USA, 2010. ACM.

[74] Lin Xian-ming and Li Shao-zi. Transfer adaboost learning for action recognition. In *IT in Medicine Education, 2009. ITIME '09. IEEE International Symposium on*, volume 1, pages 659 –664, aug. 2009.

[75] Jun Yang, Rong Yan, and Alexander G. Hauptmann. Cross-domain video concept detection using adaptive svms. In *Proceedings of the 15th international conference on Multimedia*, MULTIMEDIA '07, pages 188–197, New York, NY, USA, 2007. ACM.

[76] Q. Yang. Activity recognition: linking low-level sensors to high-level intelligence. In *Proceedings of the 21st international jont conference on Artifical intelligence*, pages 20–25. Morgan Kaufmann Publishers Inc., 2009.

[77] Weilong Yang, Yang Wang, and Greg Mori. Learning transferable distance functions forhuman action recognition. In Liang Wang, Guoying Zhao, Li Cheng, and Matti Pietikinen, editors, *Machine Learning for Vision-Based Motion Analysis*, Advances in Pattern Recognition, pages 349–370. Springer London, 2011.

[78] Z. Zhao, Y. Chen, J. Liu, and M. Liu. Cross-mobile elm based activity recognition. *International Journal of Engineering and Industries*, 1(1):30–38, 2010.

[79] Z. Zhao, Y. Chen, J. Liu, Z. Shen, and M. Liu. Cross-people mobile-phone based activity recognition. In *Twenty-Second International Joint Conference on Artificial Intelligence*, 2011.

[80] V.W. Zheng, D.H. Hu, and Q. Yang. Cross-domain activity recognition. In *Ubicomp*, volume 9, pages 61–70, 2009.

[81] V.W. Zheng, S.J. Pan, Q. Yang, and J.J. Pan. Transferring multi-device localization models using latent multi-task learning. In *Proceedings of the 23rd national conference on Artificial intelligence*, pages 1427–1432, 2008.