The Empirical Study On Deep Convolutional Network Transferring Among Users Within Activity Recognition[†]

Jiazhen Li, Renjie Ding[‡], Lanshun Nie[§], Xue Li, Xiandong Si, Lulu Wang, Dianhui Chu[**], Dechen Zhan

{lijiazhen_smile@163.com, renjie_ding_hitwh@163.com, nls@hit.edu.cn }

Harbin Institute of Technology, Dalian University

Abstract. Human activity recognition based on sensor data is one of the significant problems in pervasive computing. In recent years, deep learning has become the main method in this field due to its high accuracy. However, it's difficult to recognize activities of user B with the model trained for user A. The effect of transferring the model (among different users) is the key that restricts activity recognition in practice. At present, there is still little research on the transferring of deep learning model in this field. Its effect, principle and influential factors remain to be studied. Therefore, we carried out the empirical study on the transferring of deep learning model among users. We visualized the features extracted from CNN and studied its distribution. We compared the feasibility, strengths and weaknesses of typical unsupervised and semi-supervised transferring methods. The observations and insights in this study have deepened the understanding of transferring in activity recognition field and provide guidance for further research.

Keywords: Sensor-based data; Human activity recognition; CNN; Transfer learning

1 Introduction

Human activity recognition is rapidly emerging in the field of pervasive computing. The main data sources are divided into two categories: sensor data and audio image data. Audio image data is easy to recognize, but it is restricted to scenes, the layout of device, and personal privacy issues. With the rapid development of sensors, the smaller, more precise and accurate sensors are incorporated into people's lives. Sensors become a mainstream tool in the field of activity recognition because they are easy to install, easy to collect and easy to integrate quickly. The research study analyzes the activity recognition based on sensor data. In the area of activity recognition, we study the transferring of deep learning model among different users.

When we used sensor data based on different people for training model, we found that if training set and test set belonged to the same set of people, we could get a better accuracy. If they belonged to different sets, the accuracy would drop dramatically. In [1], L1O method was used. (L1O method refers to the test set which is one person's activity data, and this person's data do not participate in the training of the model.) When L1O method is used to evaluate the model, the accuracy of classification will fall about 5% to 20%. Therefore, we hope that the model trained by the data of training set can well identify the activities of people who do not participate in training. At the same time, we hope that through the transferring method, we can improve the accuracy of recognizing different people in the test set.

The application of transferring learning in activity recognition includes the transfer in different data sets and the transfer among different people in activity recognition. [2][3] are studying of a transfer in different data sets of activity recognition, [4] implements the transfer scenes based on different settings in Wi-Fi location, [5][6][7] implement the transfer among users. All of these papers used traditional machine learning with transferring methods.

However, the traditional machine learning methods of activity recognition have the following disadvantages. Most machine learning methods rely on manual feature extraction heavily, which is often limited by professional knowledge. Moreover, people generally only know some shallow features of

---

[‡] The first two authors contributed equally to this work.

[§] Corresponding author

[**] Corresponding author

feature extraction. Currently, in the field of deep learning model, convolutional neural networks mainly rely on three important properties, including local connection, weight sharing and translation invariant[8][9] to extract complex features, so that convolutional neural networks have achieved good research results in the field of activity recognition.

Therefore, this paper is focused on transfer among different people with the convolutional neural network based on deep learning model. To the best of our knowledge, there are very few studies on transfer among different people based on deep learning. [10] explored the ability of deep convolutional features to transfer among Domain, Modalities and Locations.

Using the deep transfer methods, we think it can be divided into unsupervised and semi-supervised transfer according to usage of labeled data. From two aspects of unsupervised and semi-supervised, we apply transfer methods from other fields such as computer vision, to activity recognition among different people and study their effects and characteristics. The main contributions of this paper are as follows:

1. This paper explored the distribution characteristics of CNN for the extracted features from sensor data, and analyzed the change of the CNN model before and after the transfer among different people and its impact.

2. This paper explored the effects, limitations and influencing factors of typical unsupervised transfer methods and semi-supervised transfer methods on activity recognition, and compared unsupervised and semi-supervised transfer methods.

3. In the field of activity recognition, a systematic and empirical study of the problem of deep convolutional network transfer was carried out. The obtained insight deepens the understanding of the problem and provides guidelines for future model and method design.

2   Related Work

2.1   Activity Recognition

With their wide spread as well as the increasing number of sensors in them and the constantly improving functions, smartphones have greatly assisted in the recognition of human activities in the field of life, medical field and sports field. The body's gait monitoring was performed in [11] to detect the degree of alcohol consumption. In [12], environmental sensors and object sensors, such as RFID tags, were combined with commonly used items in daily life to study the daily activity recognition in smart home. [13] used the RFID tag technology to observe the patient's physical reaction or the symptom of disease, and to assist the emergency personnel to carry on the corresponding treatment measure. [14] mainly evaluated the performance of a smartphone-based activity recognition classifier for monitoring healthy and stroke patients. [15] surveyed basketball activities in detail. [16] mainly studied the location of action and its recognition, and selected the video of the real sports competition as the data set.

2.2   Transfer Learning

In the real world, it is difficult for people to get lots of labeled data. Therefore, the researchers hope that the model trained by the source domain (a large amount of labeled data) can transfer to the target domain (unlabeled data or a small amount of labeled data), and the model can well classify in the target domain. In recent years, with the continuous development of machine learning and deep learning, transfer learning has made great achievements in many fields of application. For example, [17] studied the adaptive problem of structured language in the field of natural language processing, and the adaptive problem in image classification in computer vision [18].

Transfer learning methods are mainly divided into instance-based transfer learning, feature-based transfer learning, and relation-based transfer learning. On the model side, the method of transfer learning is model-based transfer learning. [19], [20] and [21] realized the instance-based transfer learning method, adjusted the weights of instances based on the source domain, screened the data with high similarity with the target domain, and then conducted training and learning. [22] and [23] proposed that feature-based transferring learning method is to transform the data of two domains into the same feature space through feature transformation, and then carry on the traditional deep learning or machine learning. [24] proposed the maximum mean difference (MMDE) to learn potential features in RKHS. [22] extended MMDE through Transmission Component Analysis (TCA) while learning a core in RKHS. In [25] and [26], a relation-based transfer learning method was implemented by using the source domain to learn the logical relation grid and then applying it to the target domain. In [27] and [28], a model-based transfer method was implemented. The model learned from the source domain was applied to the target domain and a

new model was learned from the target domain.

2.3 Transfer Learning based on Activity Recognition

In [29], a combination of naive Bayesian and Support Vector Machine (SVM) was proposed, which had better prediction accuracy for new users in activity recognition test set. [30] proposed a combination of EM algorithm and CRF, using physiological data to assist in the transfer in different data sets. In [3], GMM was used to model the data and then GMM parameters were transferred. A decision tree was proposed in [5] to train a model for the first person, then identified the second person as a new user, and clustered to provide a pseudo-label for the second person through the clustering result to improve the accuracy of identifying new users constantly and iteratively. [10] proposed to apply deep learning and transfer learning in the field of activity recognition. In this paper, a 3-layer convolution neural network combined with a 1-layer lstm structure was proposed, through which fixed-layer fine-tuning method was applied to explore feature transfer ability.

3 Questions

This paper is mainly about the question of the deep learning model transferring among different people in the field of activity recognition. We found and raised some questions, and we wanted to answer these questions experimentally.

1. The research in this paper is based on the sensor data. Sensor data is not as intuitive as the image or audio data, people cannot effectively distinguish them. So what are the characteristics of sensor data's feature distribution extracted by CNN?

2. Why is the recognition of the user B ineffective with using the user A activity recognition model?

3. What are the advantages and limitations of unsupervised transfer in activity recognition?

4. Are the models trained with sensor data, like those in the field of computer vision, also having some layers with general capabilities? Will these layers provide good help when the target domain has only a very small amount of labeled data?

5. What are the advantages and limitations with semi-supervised transfer in activity recognition?

6. What is the impact of increasing the amount of data on the source domain for unsupervised methods? How does the amount of target domain's labeled data affect semi-supervised transfer?

4 Methods

In order to delve deeper into the transfer among different users and answer the questions above. We designed the following combination of methods to experiment. The research steps in this paper are shown in Fig. 1, and according to the three steps, we answered six questions in Section 3.
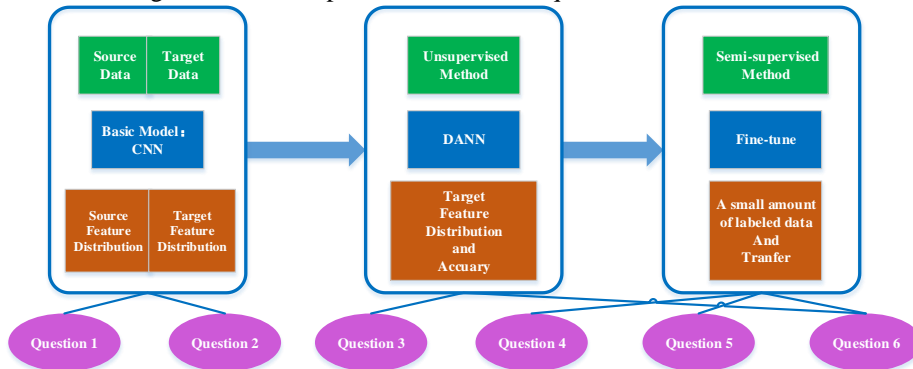

Fig. 1. The overall flow chart of the method

1. Firstly, we use the convolutional neural network [8][9] which is commonly used in activity recognition as basic model, we want to study the distribution of features extracted by this basic model.

2. Secondly, DANN[31], which does not rely on target's labeled data, is used as an unsupervised transfer method. We choose this method because source and target in this question share the same classes. We want to research the change of accuracy and feature distribution before and after transferring. The

limits, weaknesses and effect factors of this method are also very important.

3.   Finally, we use the classical method called fine-tune with freezing layers of basic model. We want to study the transfer of CNN models with very little labeled target data and its effect and deficiencies.

4.1   Basic Deep Learning Network

This paper used convolutional neural network as the basic model. Its structure is shown in Fig. 2, which contains convolution layer, pooling layer, fully connected layer, and Softmax classifier. We regarded a variety of sensor data as two-dimensional single-channel image data, that is, the input data is as follows, the data of different sensors are spliced into a matrix and taken as the input of the convolutional network, and then features are extracted through convolutional layers and pooling layers. Finally, Softmax classifier classifies human activities.
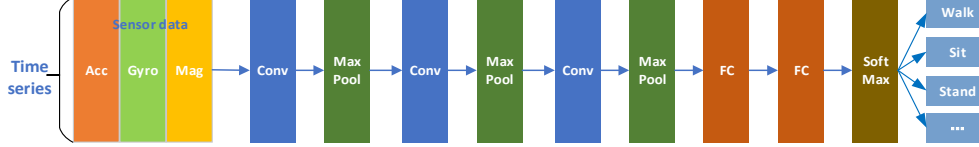

Fig. 2. Structure of basic model. Conv: convolutional layer. Max Pool: pooling layer. FC full connected layer.

4.2   Unsupervised Transfer Learning

In unsupervised transfer learning, we used the proposed DANN [31] model. It used the confusion of Source and Target fields to extract domain invariant features and classify them by using domain invariant features. Its structure is shown below.
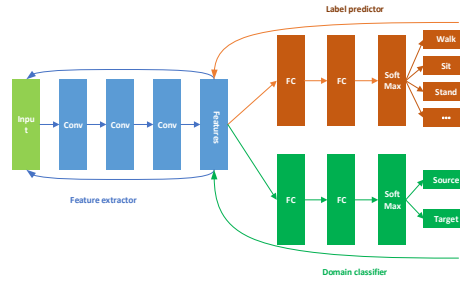

Fig. 3. DANN model structure

In Fig 3, Features is the feature extractor f=Gf(x,θf), the classifier Gy(.,θy), and the domain classifier Gd(.,θd). The training process of DANN network is firstly training the parameter θf and θd so that it can maximize the confusion of domain classifier loss in the field, so as to ensure that the extracted features have the domain invariance, and then the training label classifier minimizes the label classifier loss. So we get the loss function:

$$E(\theta f, \theta y, \theta d) = \sum_{\substack{i=1..N \\ di=0}} Ly(Gy(Gf(\mathrm{x}i, \theta f)\theta y), yi) - \lambda \sum_{i=1..N} Ld(Gd(Gf(\mathrm{x}i, \theta f)\theta d), yi). \tag{1}$$

$$= \sum_{\substack{i=1..N \\ di=0}} Lyi(\theta f, \theta y) - \lambda \sum_{i=1..N} Ldi(\theta f, \theta d) \quad . \tag{2}$$

In the above formula, *Ly* is the loss function of the classifier and *Ld* is the loss function of the domain classifier. In the forward propagation, the gradient reverse layer is an identity transformation. In the process of back propagation, the parameter is multiplied by -λ and then transmitted to the previous layer.

$$\theta f = \theta f - \mu\left(\frac{\partial Lyi}{\partial \theta f} - \lambda \frac{\partial Ldi}{\partial \theta f}\right)、\quad \theta y = \theta y - \mu \frac{\partial Lyi}{\partial \theta y},\quad \theta d = \theta d - \mu \frac{\partial Ldi}{\partial \theta d}. \tag{3}$$

In the above formula   μ   is the learning rate, and λ is a hyper-parameter to control the classifier and domain classifier. Simply, DANN uses a lot of labeled data from Source to train tag classifiers. Put a Source tag (such as 0) for the Source data and a Target tag (1) for Target's unlabeled data. Use two-part data to train domain classifiers. We want the label classifier to have as small loss function value as possible to achieve the correct classification. We hope domain classifier loss function values as much as possible, in order to achieve confusion in the field. Until the model cannot distinguish whether the data belongs to Source or Target, we can extract the common features of the two domains. The two losses

come adversarial. The basic CNN model we use is the one presented in Section 4.1.

4.3 Semi-supervised Transfer Learning

In the semi-supervised learning method, we used the most basic Fine-tune. [10] used a similar method to study the ability of CNN + LSTM to transfer among different people and used more labeled data.

We fixed the first layer of CNN, the second layer, and all three layers respectively, and non-fixed completely. The fixed layer is not involved in learning. In other similar experiments, unfixed layers tend to use random initialization, and we were not optimistic about the ability to train entirely new models with very little data. Therefore, we still used the parameters of Source's basic model, and these unfixed layers are involved in Fine-tune training.
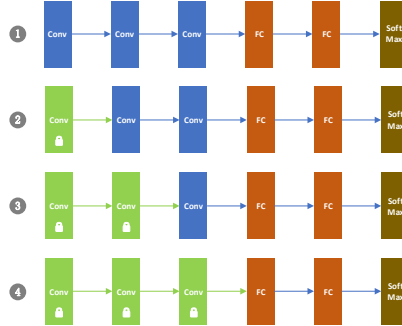

Fig. 4 Fixed convolution diagram. Green indicates that the layer parameters are fixed and do not participate in learning.

5 Distribution Of Features Extracted By CNN

With the visualization approach, we study the distribution of the sensor-based data extracted by CNN. We used UCI daily and sports dataset [32] and USC-HAD [33] dataset. UCI has a set of 8 users with 19 kinds of daily activities and sports. USC-HAD combines the data of 12 kinds of activities which are collected by acceleration and gyro from 14 users. Referring to [34], we combine the elevator up and down activities into one activity. Therefore, we get 11 kinds.

We picked out two users from UCI and USC-HAD as U1(UCI) and H1(USC). With their respective 80% data, we trained two basic models separately. We changed the number of nodes in the last fully connected layer into 2. Therefore, the two-dimensional features of each activity were extracted through this layer, and their distribution can be observed in the two-dimensional coordinate axes. Admittedly, there is a little drop in accuracy. Sometimes only one user's data is very easy to get overfitted model.
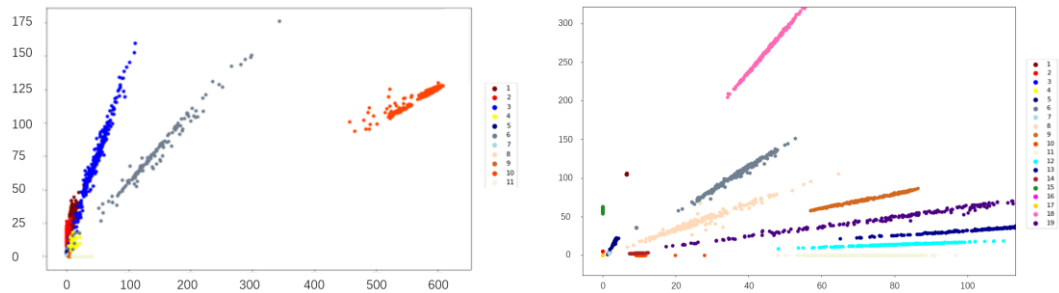
Fig. 5. Distribution of sensor-based data. H1 on the left and U1 on the right. Points are colored by the real labels.

From **Fig. 5.**, we can see: Several kinds of activity features are distributed very close and some have intersections. This can be considered that the distance between classes is too small. Here we give some explanations: We do not think this distance can simply be defined as the distance between feature centers. We see that the features are strip-shaped. We consider it as the distance between the center of close edges of the two types of features. Because here is a two-dimensional feature, Euclidean metric can be used to measure. The features extracted by CNN are often strip-shaped and the distribution of the same kind of activity tends to pull very long, which is the distance within classes(center-loss[35]).

The following conclusions are drawn:

Observation 1: The features of activity recognition data show that the distance between classes is small,

and the distance within classes is large.

Observation 2: Too small distance between classes may lead to the mixing of features and wrong recognition. Too large distance within the same class will make features on the edge more similar to other classes. These two aspects increase the difficulty of recognition.

Observation 3: We cannot judge the similarity of activities by subjective feelings of people.

We also randomly selected U2 from UCI. U1 as Source, U2 as Target.
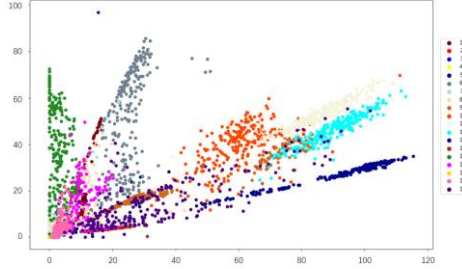
Fig. 6. Features of U2 extracted by U1 model.

From **Fig. 6.**, we can see that the features extracted from U2 data by model trained with U1 are quite mixed and interlaced, which cannot be accurately classified. The lower accuracy is an inevitable result. Therefore, it is necessary to carry out the transfer learning among different users.

Observation 4: Because of the differences between data from different users, the models trained on A usually extract mixed features for most B and are difficult to recognize.

6 Analysis on Results of DANN

We picked out five people from each of the two data sets. Five of UCI are P1, P2, P4, P7, P8 and five of USC are P5, P6, P7, P8, P9.

6.1 One-To-One Transfer

We experimented with the individuals to perform one-to-one transfer within the same dataset. Everyone acted as source and target domain. We used 80% of data from source to train the label classifier and 80% of data from target as unlabeled data to make domain fully confused. Every one-to-one experiment was repeated five times. We also averaged the results which shared the same target users.
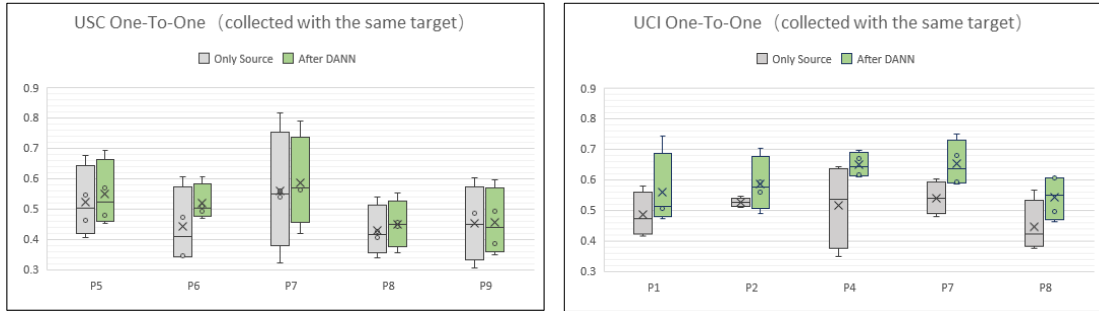
Fig. 7. Mean accuracy of one-to-one experiments, collected with the same target.

In terms of the average and median of different source with the same target, there is an improvement overall, but the accuracy of DANN with different people is quite different, sometimes negative transfer occurs. And even using the same users, the results are not stable.

Observation 5: DANN can improve the accuracy of target domain in one-to-one experiments, but selecting different people has a great impact on the result. It is a big limitation for practical use

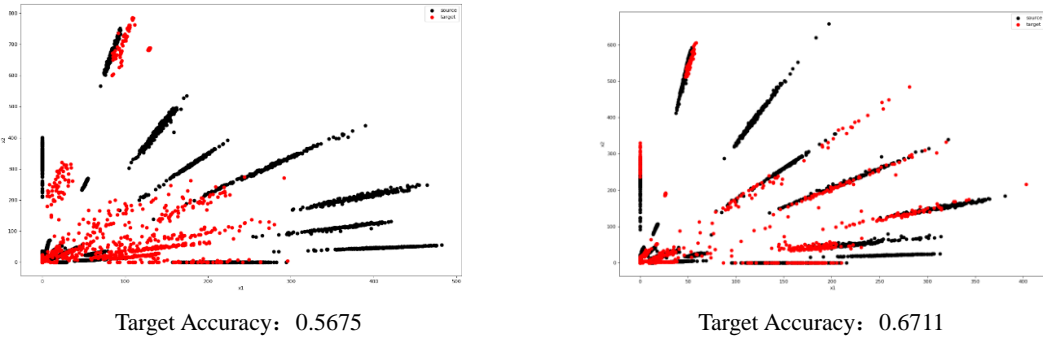We chose two users from UCI to take a deep look on the features. We call them Source and Target.

Target Accuracy：0.5675   Target Accuracy：0.6711

Fig. 8. Distribution of features. Left (only-source), right (DANN). Black (source), red (target).
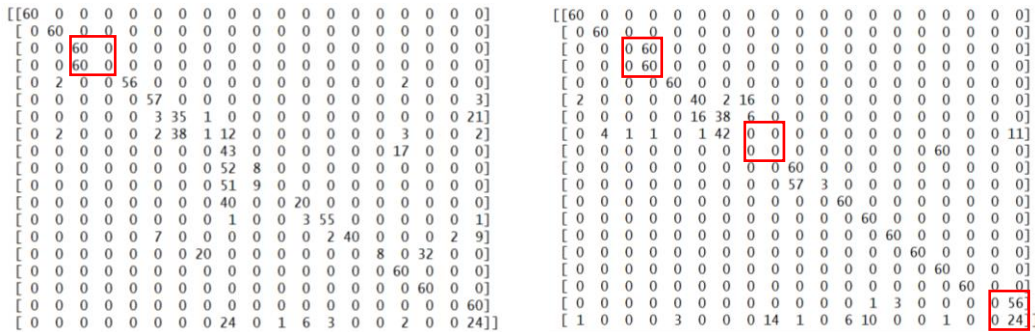

Fig. 9. Confusion matrix of target test data. Left is test by only source model. Right is test after DANN.

After adaptation, the model learnt the general features across source and target. It will classify target to the source classes where they have the most similar or closest features. However, the model cannot always make the right decisions because of small distances between classes (see the lower left corner in **Fig. 8**) and difference between users' actions. From **Fig. 8**, it's clear that after DANN there are some classes of dark spots on which there are no red spots attached, while some others have a large number of red spots. From the confusion matrix(**Fig. 9**), there are still several entire classes of activities which are totally wrong recognized after the adaptation. Negative migration may occur when it is difficult to recognize multiple types of activities.

Observation 6: We think it is quite difficult to correct the confusion of the entire class. We suspect it is the common phenomenon in unsupervised transfer learning and it's one of the weaknesses.

6.2 Several-To-One Transfer

We hope to explore the impact of the number of source users on transfer. Therefore, we increased the amount of source users. Firstly, in each dataset, we chose two users as source from the five individuals listed above. We simply doubled the data of the target to make the domains fully confused.
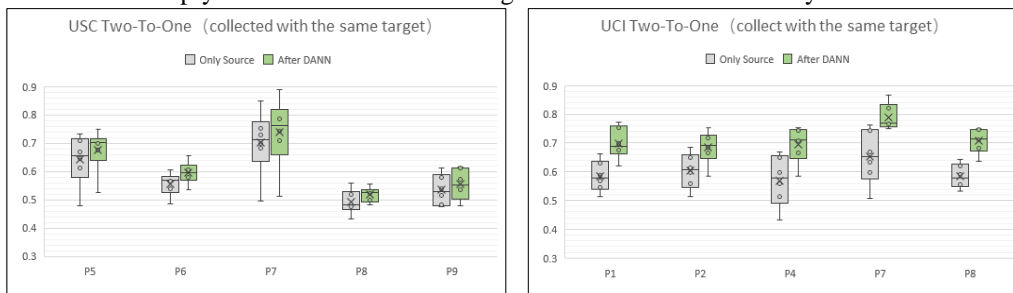

Fig. 10. Accuracy of two-to-one experiments, collected with the same target.

We can see that by increasing the number of users in source, the mean and median of accuracy from the two datasets before and after DANN improved. But The result really depends on the users you selected. We continue to increase the number of users in source to 4.

Since there are only one combination of four-to-one for every target in the five users, we ploted the histogram. We repeat twice for each combination and calculated the mean accuracy.
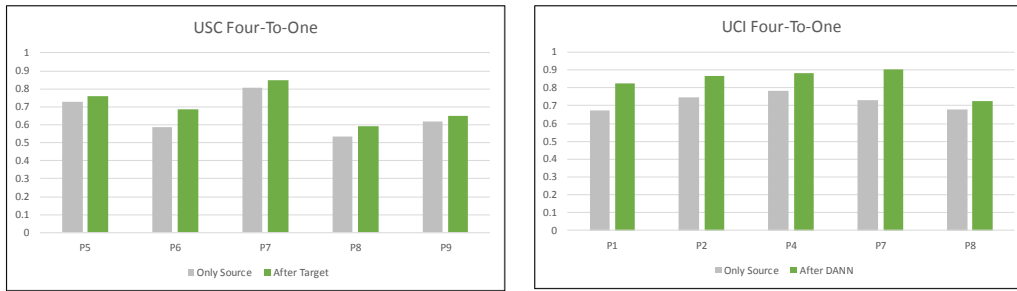
Fig. 11. Accuracy of two-to-one experiments.

After using four users in source, the accuracy are further improved before and after transferring. ***Observation 7:*** Increasing the number of users in source can help the adaptation of DANN. However, the selection of users in source can have a huge impact on results. We think this is a big disadvantage.

7 Analysis on Results of Fine-tune

7.1 Study Under Very Few Data From Target

We still make use of UCI daliy and sports and USC-HAD. First of all, three users are selected from UCI (P1\P4\P7), Everyone acts as source and target. We use 80% data of source to train the base model. Use five target data of every class (19 classes have 95 data, occupies 1/60 of Target) as labeled data to fine-tune. The rest are used for test. Transferring among three people within the same dataset has six combinations. We calculated the mean accuracy of the six combinations for each frozen way.
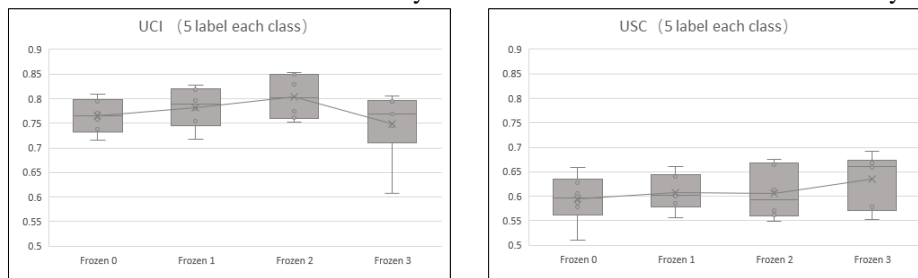

Fig. 12. The accuracy of different number of frozen convolutional layers. The polyline is connected with mean.

We repeated the experiment with USC-HAD dataset. Unlike UCI, it seems to perform best when the three layers are all frozen. Through the two results, we found that with the increase of the number of frozen convolutional layers, in fact, the change of accuracy is not large. We also find that the accuracy is not very stable, and there is a fluctuation from time to time.

Observation 8: All the first three convolutional layers in the model can be general.

Observation 9: Fine-tuning with few labeled data cannot be compared with DANN simply on accuracy. But when using a small number of labeled data, the transfer is less affected by the selection of users, which is better than unsupervised transfer.

### 7.2 Data Size and Change of Features Distribution

We set target to have 5,10,15 tags per class. The accuracy is shown below (**Fig .12**). The general accuracy is rising.
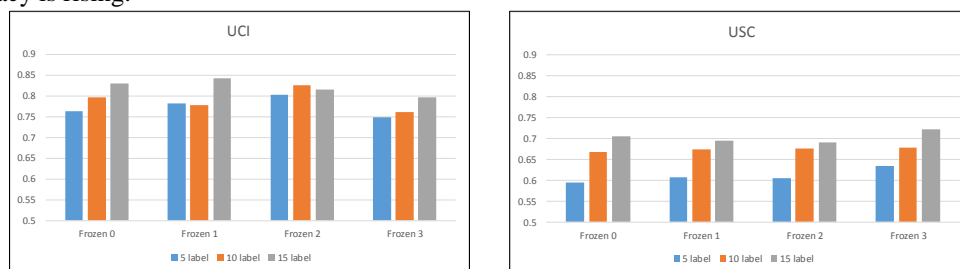
Fig. 13. Add the size of labeled data.

We changed the number of nodes in the former FC layer into 2 to observe the distribution features. We used 1/10 and 1/5 of data from target as the labeled data. We still used the source and target (U1\U2) in section 5 and froze the first two convolutional layers for experiment.

1/10 （acc 0.721）　　　　　　　　1/5 （acc 0.813）
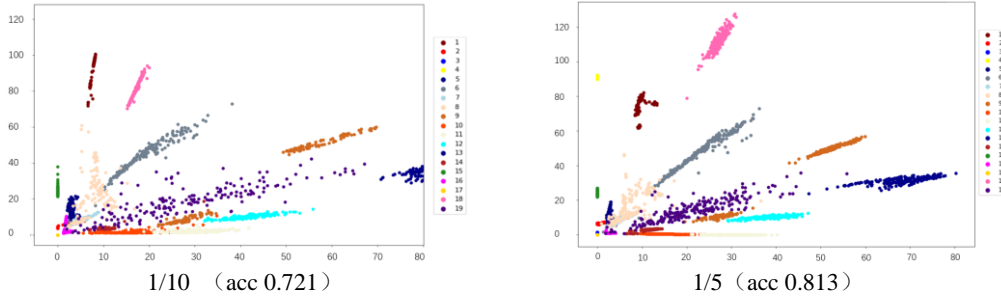
Fig. 14. Distribution of target domain data feature before and after Fine-tuning.

After adding 1/10 tagged data to training, the situation of mixed feature has been improved, but its distribution is still partly mixed, which restricts the recognition accuracy rate.

We used 1/5 of data. Due to the help of two frozen layers from source model, the feature distribution of target is somewhat similar to source (**Fig. 5**). By Comparing **Fig. 6** with **Fig. 14**, we found that the distance between classes was increased, the distribution of features within the cluster became aggregated, and the degree of mixing lowered. Therefore, the difficulty of identification has been reduced.

Observation 10: Increasing the size of labeled data will improve accuracy. Fine-tuning helps the model get the features of target data, changes it into a cluster within the class, separating from other classes.

If you are interested in our experiment, you can contact us by email, we will open our experimental results and code.

8　Conclusions and Outlook

In this paper, we focused on transferring across different users within HAR, experimented with unsupervised and semi-supervised methods, and studied the change of distribution of features before and after the transfer. We came to the following insights to the questions in Section 4.

Insight 1: We find that the sensor-based data extracted by CNN has small distance between classes and large discrepancy within the same class, which makes it difficult to recognize. (*Observation 1 2 3*)

Insight 2: Due to the difference of actions among different people, the feature of user B which is extracted by model trained for user A is quite mixed and overlapped, so it is difficult to classify accurately. Transfer of the model is necessary. (*Observation 4*)

Insight 3: The unsupervised approach does indeed have the effect on the transfer among different people. However, we also found that this type of method often leads to the wrong classification of the whole class. We think that this kind of error cannot be corrected only by narrowing the gap between source and target. It needs tagged data. And as we know, the unsupervised approach is not always effective, so the selection of user in source also plays an important role. Moreover, in the field of activity recognition, it is not easy to get a large amount of untagged data, which also requires human acquisition and noise processing. (*Observation 5 6 7*).

Insight 4: With the increase of fixed number of layers in fine-tune, the change of accuracy is not large and the accuracy is often unstable. We think that the first three convolutional layers have some general abilities of extracting features within HAR. (*Observation 8*)

Insight 5: Freezing general layers and fine-tuning can be the foundation of semi-supervised methods; in the use of a small amount of tagged data in transferring, it can achieve the accuracy which is not inferior to that in the unsupervised methods. Moreover, its training speed and efficiency is much higher. It's far less affected by the selection of users. In HAR, we consider the cost of acquiring small amounts of labeled data far less than large amount of unlabeled data. We think that it is more appropriate to transfer with small amount of tagged data. (*Observation 9*)

Insight 6: In the unsupervised DANN approach: We find that increasing the amount of data in the source domain has a significant contribution to the transfer. In fine-tuning, increasing the amount of tagged data will undoubtedly improve the transfer. (*Observation 7 10*)

We find that, it will sometimes be unstable both in unsupervised and semi-supervised methods. We think this is related to the hyperparameters and the amount of data, so it still remains to be studied. Within the human activity recognition, we still insist that transferring across users is one of the most valuable approaches. From our conclusions, we argue that semi-supervised methods in transfer learning

is more appropriate and efficient than unsupervised methods. This paper mainly focuses on experimental conclusions. In the future, we will study the theoretical basis about transferring among users.

References

[1] Altun Kerem, Barshan Billur, Tuncel Orkun.: Comparative study on classifying human activities with miniature inertial and magnetic sensors.Pattern Recognition, 43(10). pp. 3605-3620 (2010)

[2] Wang Chenglong, Wu Yue, Liu Zongtian.: Hierarchical boosting for transfer learning with multi-source. International Conference on Artificial Intelligence and Robotics. pp. 723-731 (2016)

[3] Cao Liangliang, Liu Zicheng, Huang Thomas S.: Cross-Dataset Action Detection. IEEE Computer Society Conference on Computer Vision and Pattern Recognition. pp. 1998-2005 (2010)

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

[5] Zhao Zhongtang, Chen Yiqiang, Liu, Junfa, Shen Zhiqi.: Cross-People Mobile-Phone Based Activity Recognition. IJCAI International Joint Conference on Artificial Intelligence. pp. 2545-2550 (2011)

[6] Wang, Zheng, Song, Yangqiu, Zhang, Changshui.: Transferred Dimensionality Reduction. Machine Learning and Knowledge Discovery in Databases - European Conference. pp. 550-565 (2008)

[7]Hong J H, Ramos J, Dey A K.: Toward Personalized Activity Recognition Systems With a Semipopulation Approach. IEEE Transactions on Human-Machine Systems.46(1). pp. 101-112 (2016)

[8] Jiang, W., Yin, Z.: Human activity recognition using wearable sensors by deep convolutional neural networks. Proceedings of the 23rd ACM international conference on Multimedia. pp. 1307–1310 (2015)

[9] Ravi, D.,Wong, C., Lo, B., Yang, G.Z. ,et al.: Deep learning for human activity recognition: A resource e_cient implementation on low-power devices. Wearable and Implantable Body Sensor Networks (BSN), IEEE 13th International Conference IEEE. pp. 71–76 (2016)

[10] Morales F J O, Roggen D.: Deep convolutional feature transfer across mobile activity recognition domains. sensor modalities and locations Proceedings of the 2016 ACM International Symposium on Wearable Computers. pp. 92-99 (2016)

[11] Kao, Hsin-Liu Cindy.: Phone-based Gait Analysis to Detect Alcohol Usage. UbiComp2012 - Proceedings of the 2012 ACM Conference on Ubiquitous Computing. pp. 661-662 (2012)

[12] Yao, L., Sheng, Q.Z. Compressive representation for device-free activity recognition with passive RFID signal strength. IEEE Transactions on Mobile Computing, 17(2), pp.293-306 (2018)

[13] Parlak Siddika ,Marsic Ivan, Burd, Randall S.: Activity Recognition for Emergency Care using RFID. BODYNETS 2011 - 6th International ICST Conference on Body Area Networks.pp. 40-46 (2012)

[14] Y Kwon, K Kang, C Bae.: Evaluation of a smartphone human activity recognition application with able-bodied and stroke participants. Journal of Neuro Engineering and Rehabilitation (2016)

[15] Nguyen, Le Nguyen Ngu, Rodríguez-Martín, Daniel, Catal, Andreu.: Basketball activity recognition using wearable inertial measurement units. ACM International Conference Proceeding Series. pp. 59-69 (2015)

[16] K. Soomro and A.R. Zamir.: Action Recognition in Realistic Sports Videos. Springer International Publishing Switzerland. (2014)

[17] J. Blitzer, R. McDonald, and F. Pereira.: Domain adaptation with structural correspondence learning. Proceedings of the 2006 conference on empirical methods in natural language processing. pp. 120–128 (2006)

[18] Patel, Vishal M , Gopalan, Raghuraman , Li, Ruonan , Chellappa, Rama.: Visual Domain Adaptation. A Survey of Recent Advances. pp 53-69 (2015)

[19] Dai W, Yang Q, Xue G R, et al.: Boosting for transfer learning. Proceedings of the 24th international conference on Machine learning. pp.193-200 (2007)

[20] Gretton A, Smola A, Huang J, et al.: Covariate shift by kernel mean matching. Dataset shift in machine learning. 3(4): 5. (2009)

[21] Sugiyama M, Suzuki T, Kanamori T.: Density ratio estimation in machine learning. Cambridge University Press. (2012)

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

[23] Long M, Wang J, Sun J, et al.: Domain invariant transfer kernel learning. IEEE Transactions on Knowledge and Data Engineering, 27(6), pp.1519-1532 (2015)

[24] S. J. Pan, J. T. Kwok, and Q. Yang. Transfer learning via dimensionality reduction .AAAI, vol. 8. pp. 677–682. (2008)

[25] Mihalkova L, Huynh T, Mooney R J.: Mapping and revising Markov logic networks for transfer learning . AAAI . pp. 608-614 (2007)

[26] Davis J, Domingos P.: Deep transfer via second-order markov logic. Proceedings of the 26th annual international conference on machine learning. pp. 217-224 (2009)

[27] Oquab, Bottou, Laptev, Sivic.: Learning and Transferring Mid-Level Image Representations using Convolutional Neural Networks. CVPR 2014. pp. 1717-1724 (2014)

[28] Yao, L., Nie, F., Sheng, Q.Z., Gu, T. Learning from less for better: semi-supervised activity recognition via shared structure discovery.pp. 13-24 (2016)

[29]Hong J H, Ramos J, Dey A K.: Toward Personalized Activity Recognition Systems With a Semipopulation Approach. IEEE Transactions on Human-Machine Systems, 46(1). pp. 101-112 (2016)

[30] HaiLeongChieu1, WeeSunLee2, Leslie PackKaelbling.: Activity Recognition from Physiological Data using Conditional Random Fields.(2006)

[31]Yaroslav Ganin,Evgeniya Ustinova, Hana Ajakan,etc.: Domain-Adversarial Training of Neural Networks.Journal of Machine Learning Research 17 pp. 1-35 (2016)

[32] K. Altun:Human activity recognition using inertial magnetic sensor units. Proceedings First International Workshop on Human Behavior Understanding (in conjunction with the 20th Int. pp.38-51 (2010)

[33] Zhang Mi, Sawchuk, Alexander A.: USC-HAD: A Daily Activity Dataset for Ubiquitous Activity Recognition Using Wearable Sensors. UbiComp 12. pp. 1036-1043 (2012)

[34] Jian Cui, Bin Xu.:Cost-Effective Activity Recognition on Mobile Devices. BODYNETS 2013 - 8th International Conference on Body Area Networks.pp. 90-96 (2013)

[35] Yandong Wen, Kaipeng Zhang, Zhifeng Li.: A Discriminative Feature Learning Approach for Deep Face Recognition. European Conference on Computer Vision(ECCV).pp. 499-515 (2016)