국문 제목

Improving Human Activity Recognition Model with Very Limited Labeled Data using Domain Knowledge Transfer Through Multitask Semi-Supervised Learning

요 약

A key to a well performing human activity recognition (HAR) system through machine learning technique is the availability of substantial amount of labeled data. Collecting sufficient labeled data is an expensive and time-consuming task. To build a HAR system in a new environment (i.e., the *target domain*) with very limited labeled data, it is unfavorable to naively exploit the data or trained classifier model from the existing environment (i.e., the *source domain*) as it is due to the domain difference. While traditional machine learning approaches are unable to address such distribution mismatch, transfer learning approach leverages the utilization of knowledge from existing well-established source domains to help building an accurate classifier in the target domain. In this work we propose a transfer learning approach to create an accurate HAR classifier with very limited data through multitask neural network. The classifier loss function minimization for source and target domain are treated as two different tasks. The knowledge transfer is performed by simultaneously minimizing the loss function of both tasks using a single neural network model. Furthermore, we utilize the unlabeled data in an unsupervised manner to help the model training. Experiment result shows the benefit of knowledge transfer for HAR based on limited data compared to the conventional non-transfer machine learning approach.

1. Introduction

Sensor-based human activity recognition (HAR) has been widely incorporated as the important component in an intelligent environment for various domain, such as healthcare [1], anomaly detection [2], and elder care [3]. To make a good HAR model, sufficient labeled data is needed. However, collecting labeled data is very tedious. Furthermore, it is difficult to merely apply existing data from different domain to make a prediction model for new domain data due to domain difference (e.g., data distribution shift due to difference in sensor placement, users' posture, sampling frequency, etc.).

While traditional machine learning algorithm has not addressed such domain difference, transfer learning comes into rescue to exploit the knowledge of existing information (domain) to help the HAR classifier training. In addition, semi-supervised learning makes use of unlabeled data to improve the training of HAR classifier. Existing work that addressed the utilization of unlabeled sensor data to improve classification with semi-supervised learning approach was designed to only work with a single domain [4]. Therefore, although it can deal with unlabeled data, it still cannot directly take advantage of the availability of labeled data in the existing domains to improve model performance even more.

In this work we propose a transfer learning approach to train a HAR classifier with very limited target domain data through semi-supervised learning multitask neural network (MTLNN). The classifier loss function minimization for source and target domain are treated as two different tasks, and the minimization of loss functions is performed simultaneously using a single network. The unlabeled data are utilized in an unsupervised manner to help the model training process.

2. Problem formulation

Let $\boldsymbol{S} = \{(x_S^1,\ y_S^1), (x_S^2,\ y_S^2), \dots, (x_S^{n_S},\ y_S^{n_S})\}$ be labeled source domain dataset from which we want to transfer knowledge, and n_S is the number of instances of source domain data. The source dataset is assumed to be sufficient enough to train an accurate HAR classifier.

Let $\boldsymbol{T_L} = \left\{(x_T^1, y_T^1), \dots, \left(x_T^{n_{T_l}}, y_T^{n_{T_l}}\right)\right\}$ be labeled target domain dataset and $\boldsymbol{T_U} = \left\{x_T^1, \dots, x_T^{n_{T_u}}\right\}$ be unlabeled target domain dataset. Let n_{T_l} and n_{T_u} denote the number of labeled target domain data and unlabeled target domain data respectively. The number of labeled target domain dataset is assumed to be very small, and too insufficient to train an accurate classifier, and $n_{T_l} \ll n_S$. Typically, $n_{T_l} \ll n_{T_u}$ as well.

Our goal is to exploit the source domain data and unlabeled target domain data in order to make a more accurate classifier for the target domain despite the limited amount of labeled data.

3. Proposed method

This section explains our approach pipeline to leverage knowledge transfer for improving the training of HAR classifier.

3.1. Multitask learning (MTL) for knowledge transfer

MTL can be considered as one of the types of transfer learning. In the MTL setting, a single model is trained to simultaneously solve multiple tasks those are somewhat related to some extent. There are several classification algorithms which can be modified to perform learning algorithm in a multitask fashion, e.g., neural network (through multitask neural network [5]), support vector machine (through multitask-multikernel learning SVM [6]), and logistic regression (through hierarchical Bayesian logistic regression [7]).

In this work we consider multitask neural network (further will be referred as MTLNN) due to its robustness. Furthermore, neural network model family generally produces well-calibrated prediction probability estimate [8] which will be used in the later section for instance selection. Specifically, we treat the learning of source domain objective function as one task, and the learning the target domain objective function as another task.

Figure 1 shows the MTLNN architecture that will be used in this work (denoted as f). Not that in the network, there is a hidden layer where the weights are shared among the tasks. This layer will learn the common characteristics between both tasks. Intuitively, through this shared layer, the learning of one task might also help to boost the learning of another task. Next to the shared layer, there are task-specific layer, which specialize to learn the unique characteristics of each task and give the output based-on for the corresponding task.

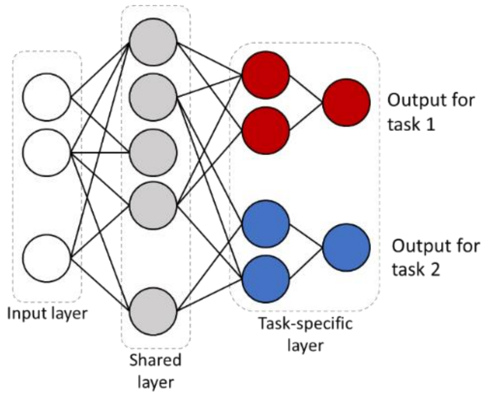

Figure 1 Multitask neural network architecture

The MTLNN will be trained using typical stochastic gradient descent (SGD) with minibatches. However, we need to modify the SGD, i.e., by ensuring that each mini-batch will contain only samples from a single randomly selected task. The purpose is to encourage the network to focus on minimizing the loss function of a single task for every single minibatch. Moreover, by letting minibatch samples to be mixed from source and task domain, the learning process will be biased toward the source domain (i.e., overlearning the source domain data) considering that the amount of source domain data is assumed to be dominating the whole training set. The batch then used to predict label values based on forward propagation through shared weights and subsequently through the appropriate task-specific weights. Cross-entropy loss function (Eq 1) is used to measure the error between ground truth $\boldsymbol{y}$ and the predicted labels $\hat{\boldsymbol{y}}$:

$$l(\boldsymbol{y}, \hat{\boldsymbol{y}}) = -\sum_{i=1}^{m} (y_i \log \hat{y}_i + (1 - y_i) \log(1 - \hat{y}_i)). \qquad (1)$$

3.2. Utilization of unlabeled data

Since both labeled and unlabeled data in target domain very likely come from the same distribution (i.e., due to the same domain), there is a potential that they share common information. Thus, there is also potential of making use of unlabeled data to help the training process. The trained MTLNN can be used to make a prediction on the unlabeled data to give the (pseudo) labels and then include the pseudo-labeled data to update the MTLNN for several next iterations. However, the initial label prediction will tend to be biased since initially the source domain data dominate the model training. Thus, we use the help from the so-called "domain expert classifier" g, $g(x) = \boldsymbol{p}_E$ which is specifically trained only the labeled target domain dataset. It also gives output of class probability estimates, $\boldsymbol{p}_E = \{p_S^1, \dots, p_S^c\}'$. This classifier will help to guide the prediction by providing information purely from target domain. Let $\boldsymbol{p}_S = \{p_S^1, \dots, p_S^c\}'$ and $\boldsymbol{p}_T = \{p_T^1, \dots, p_T^c\}'$ be the probability estimate for activity label as the output of source domain-specific layer and target domain-specific layer, where c is the number of activity classes. The final label prediction can be obtained by Eq 2.

$$\underset{i}{\operatorname{argmax}} \left\{ l_i \mid l_i \in \frac{(\alpha \boldsymbol{p}_S + \beta \boldsymbol{p}_T + \gamma \boldsymbol{p}_E)}{\mathbf{1}'(\alpha \boldsymbol{p}_S + \beta \boldsymbol{p}_T + \gamma \boldsymbol{p}_E)} \right\} \tag{2}$$

The parameter α and β are used to give importance to which domain will contribute the final prediction, $\alpha, \beta \in [0,1]$. In practice we give more weight to β, while still give nonzero value α to get some information from the source domain. Parameter $\gamma \in [0,1]$ is the weight for the domain expert classifier. The γ parameter usually should be less than β since g will tend to overfit the target domain data. Note that only predictions with confidence (indicated with probability estimate) higher than a threshold τ will be selected for the model updating in the next iterations.

4. Experiment setup

This section explains about our experiment using publicly available benchmark dataset for HAR.

4.1. Dataset

To verify the effectiveness of our method, we test it on the standard benchmark dataset for activity recognition: UCI daily and sports dataset (DSADS) and OPPORTUNITY (OPP) dataset. Both DSADS and OPP contains the recording of human activities, such as walking, running, sitting, etc., recorded using wearable sensors (i.e., accelerometer, gyroscope, and magnetometer) attached in users' body parts. Only sensors in users' arms are considered since we want to simulate the real-world situation where the user will only need a smartwatch attached on one of their hand. We also want to simulate the situation where there exists sufficient labeled source domain data. Thus, left-hand data will be used as the source domain, and right-hand data will be used as the target domain. For the target domain, we take 20% of it for testing. Out of 80% remaining, very small amount will be taken as the labeled target domain data. The labels from the rest are excluded to test semi-supervised learning capability of our proposed approach. The feature extraction procedure follows the steps in [9].

4.3. Baseline

Our approach is compared with several standard off-the-shelf supervised classifiers, such as random forest (RF), multilayer perceptron (MLP), and decision tree (DT). Beside the standard supervised classifiers, we also compare our approach with En-Co training (ET) approach proposed in [4] that is able to incorporate unlabeled data for model training.

5. Result and analysis

We prepared two scenarios for experiment. *Scenario 1* is the training of baseline methods only using labeled target data. *Scenario 2* is the training of baseline methods using labeled target data which were combined with labeled source domain data. Table 1 and Table 2 show our experiment result for both scenarios. The number of instance (NI) for labeled target domain data are set to 1, 3, and 5. For scenario 1, our method consistently outperforms the baseline classifiers. ET gets more competitive as NI gets higher. However, most of the time standard RF can still outperform ET.

Table 1. Experiment result with scenario 1

Task	NI	RF	MLP	DT	ET	Proposed
DSADS	1	0.64	0.52	0.25	0.31	**0.82**
	2	0.67	0.64	0.51	0.60	**0.85**
	3	0.78	0.70	0.72	0.76	**0.84**
OPP	1	0.47	0.47	0.34	0.51	**0.63**
	2	0.54	0.63	0.54	0.56	**0.67**
	3	0.67	0.69	0.63	0.61	**0.75**

For scenario 2, the result (as shown in Table 2) suggests that the merging of labeled source and target domain data might help to improve the accuracy to some extent, e.g., when the NI is 1. This is because the source domain data might have almost similar data distribution since they come from the same body part even though from different domain. In this case, our approach still consistently outperforms the base classifiers.

The performance improvement by using our approach indicates that the knowledge transfer is successful. Specifically, the performance benefited from the multitask learning' attention focusing, i.e., it helps the model focus its attention on those features that actually matter as other tasks will provide additional evidence for the relevance or irrelevance of those features. By considering the relevant features and discarding the irrelevant ones, the model training will converge faster and perform better.

Table 2. Experiment result with scenario 2

Task	NI	RF	MLP	DT	EM	Proposed
DSADS	1	0.69	0.71	0.63	0.70	**0.82**
	2	0.64	0.68	0.65	0.76	**0.85**
	3	0.72	0.67	0.69	0.77	**0.84**
OPP	1	0.59	0.61	0.55	0.58	**0.63**
	2	0.61	0.64	0.54	0.56	**0.67**
	3	0.62	0.63	0.58	0.58	**0.75**

We also present the result of instance selection with MTLNN, shown in Figure 2. The number of selected instances from target domain dataset increases over iteration. This is because the network gets more confident in predicting the output because the training loss gets lower. The accuracy also increases because more relevant instances are involved over iterations. This explains that our approach successfully performed the selection of important unlabeled instances for improving training process.

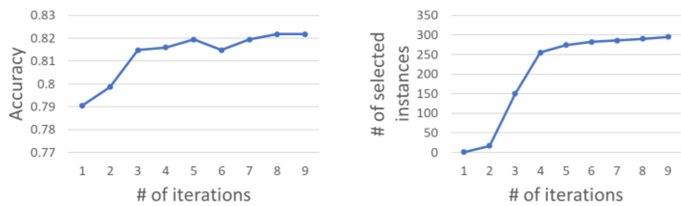

Figure 2. Accuracy improvement and selected instances over iterations (Scenario 2, DSADS, NI=1)

So far, our approach assumes that both domains should be closely related in terms of data distribution. If the distribution is totally different (e.g., transferring from user's leg data to another user's arm data), the knowledge transfer might fail. More investigation is needed to address this issue.

6. Conclusion

We have presented a transfer learning approach to improve the training of HAR classifier with very limited labeled data. The approach takes advantages from existing labeled data from another domain, while simultaneously takes advantages from unlabeled data from the same domain. As a result, our approach only requires very small amount of labeled data to train a HAR classifier. The experiment result also shows that our method yields better performance compared to those of existing approaches. We consider knowledge transfer from different body parts and/or different users as the future work.

References

[1] A. Avci, S. Bosch, M. Marin-Perianu, R. Marin-Perianu, and P. Havinga, "Activity Recognition Using Inertial Sensing for Healthcare, Wellbeing and Sports Applications: A Survey," *Archit. Comput. Syst. (ARCS), 2010 23rd Int. Conf.*, pp. 1–10, 2010.

[2] J. Yin, Q. Yang, J. J. Pan, Jie Yin, Qiang Yang, and J. J. Pan, "Sensor-based abnormal human-activity detection," *IEEE Trans. Knowl. Data Eng.*, vol. 20, no. 8, pp. 1082–1090, 2008.

[3] S. Chernbumroong, S. Cang, and H. Yu, "Genetic algorithm-based classifiers fusion for multisensor activity recognition of elderly people," *IEEE J. Biomed. Heal. Informatics*, vol. 19, no. 1, pp. 282–289, Jan. 2015.

[4] D. Guan, W. Yuan, Y.-K. Lee, A. Gavrilov, and S. Lee, "Activity Recognition Based on Semi-supervised Learning," in *13th IEEE International Conference on Embedded and Real-Time Computing Systems and Applications (RTCSA 2007)*, 2007, pp. 469–475.

[5] R. Caruana, "Multitask Learning," *Mach. Learn.*, vol. 28, no. 1, pp. 41–75, 1997.

[6] M. Kandemir, A. Vetek, M. Gönen, A. Klami, and S. Kaski, "Multi-task and multi-view learning of user state," *Neurocomputing*, vol. 139, pp. 97–106, Sep. 2014.

[7] Y. Xue, X. Liao, L. Carin, and B. Krishnapuram, "Multi-Task Learning for Classification with Dirichlet Process Priors," *J. Mach. Learn. Res.*, vol. 8, no. Jan, pp. 35–63, 2007.

[8] A. Niculescu-Mizil and R. Caruana, "Predicting good probabilities with supervised learning," in *Proceedings of the 22nd international conference on Machine learning – ICML '05*, 2005, pp. 625–632.

[9] L. Hu, Y. Chen, J. Wang, C. Hu, and X. Jiang, "OKRELM: online kernelized and regularized extreme learning machine for wearable-based activity recognition," *Int. J. Mach. Learn. Cybern.*, vol. 9, no. 9, pp. 1577–1590, Sep. 2018.