Transfer Learning and Data Fusion Approach to Recognize Activities of Daily Life

Netzahualcoyotl Hernandez
School of Computing
Ulster University
Newtownabbey
Northern Ireland, UK
BT37-0QB
hernandez_cruz-n@ulster.ac.uk

Muhammad Asif Razzaq
Ubiquitous Computing Lab
Kyung Hee University
Global Campus
446-701
South Korea
asif.razzaq@oslab.khu.ac.kr

Chris Nugent
School of Computing
Ulster University
Newtownabbey
Northern Ireland, UK
BT37-0QB
cd.nugent@ulster.ac.uk

Ian McChesney
School of Computing
Ulster University
Newtownabbey
Northern Ireland, UK
BT37-0QB
ir.mcchesney@ulster.ac.uk

Shuai Zhang
School of Computing
Ulster University
Newtownabbey
Northern Ireland, UK
BT37-0QB
s.zhang@ulster.ac.uk

ABSTRACT

Activity recognition is a core domain within intelligent systems that utilizes the sensing devices available in an environment to identify human activity. Conventional solutions rely on machine-learning approaches and the assumption that the target scenario will fit the algorithm training conditions, which raises the cost and effort of labelling data, as daily living environments are dynamic, unpredictable, and exposed to new activities. Hence, we take advantage of the ubiquitous presence of personal gadgets such as smart-watches combined with data fusion approaches to dynamically transfer learned knowledge across devices in a natural environment while performing daily living activities. In this paper, we focus on recognizing walking as an activity, which might enable carers or medical practitioners to monitor the risk of falling or suffering from a chronic disease whose progression is linked to a reduction in movement and mobility. Preliminary results show a 2% increase in activity recognition accuracy on the wearable approach, and a 10% improvement in accuracy when combining features from both wearable and environmental domains.

CCS CONCEPTS

• **Human-centered computing** → *Mobile computing*

KEYWORDS

Transfer learning, activity recognition, data fusion, wearable devices

1 INTRODUCTION

Wearable devices have received much attention in pervasive health due to their ability to enable the observation of relevant signs associated with physical performance, non-invasively and without interrupting the daily living activities of the user [1]. The opportunity to apply this technology in the medical field relies on a variety of applications associated to, for example, chronic diseases like diabetes [2] and obesity [3], that can be managed through monitoring and recognizing physical activities [4], [5]. To further illustrate the applicability of this approach, imagine a scenario in which an 80 years old patient (whose home has been equipped with multiple motion and thermal sensors) visits his medical practitioner every six months to get an assessment of his gait performance. The medical staff

PervasiveHealth '18, May 21–24, 2018, New York, NY, USA

ACM ISBN 978-1-4503-6450-8/18/05 $15.00
https://doi.org/10.1145/3240925.3240949

require the patient to undertake some physical exercises using wearable devices such as fitbit to determine the effectiveness of the prescribed medicine but want to know its effect on daily living activities within the patient's normal environment. In this context, an activity recognition algorithm will need to be able to transfer the patient experience between different physical settings.

Activity recognition is the problem of anticipating human actions given a collection of sensed data. It is a multidisciplinary field drawing upon different areas such as ubiquitous computing, machine learning and data fusion. For example, in ubiquitous computing, activity recognition can take advantage of off-the-shelf technologies available within the user's environment (*e.g.*, smartphones, smart-watches), to enable the collection of data from daily living scenarios. Machine learning, due its capacity to handle differences between sensor readings and features in the domain, can help in the automatic recognition of daily living tasks. Data fusion, which exploits the natural synergy brought about by use of multiple data sources, can help to achieve inferences that could not be feasible from a single sensor [6].

Some approaches to recognize human activities rely on design-time techniques that, after being trained once remain static, and assume that further data is drawn from identical sources and conditions. However, this cannot guarantee accurate prediction within natural environments, since day to day activities are dynamic, exposed to unexpected changes, and independent of the space domain [7].

In this regard, transfer learning is an approach in which a previously learned experience (*i.e.*, knowledge / process) from one domain is used to extract knowledge from another domain [8]. It reuses the existing knowledge and trained model to recognize activities under different conditions, resulting in the leverage of resources such as training and labelling efforts. The benefits of transfer learning positively impact on the time invested learning a new task, the human effort required when training and the diversity of activity to be recognized.

Transferring of recognition capabilities from one sensor to another one can be handled at the classifier level, in which the model is transferred directly, or by mapping from the feature space independent classes to the target domain [9]. In this paper, we propose a transfer learning approach under which an untrained device learns from the experience of an already trained sensor. The knowledge is transferred across devices, relaxing the assumption of the same feature space, and distributed by learning a mapping between different sensor technologies. A data fusion technique is used to explore the benefits of merging device data to recognize activity associated with Activities of Daily Living (ADL), such as walking [10].

In the remainder of the paper we introduce some related work on transfer learning. Then, we explain our approach in detail. Next, we evaluate our methodology by following the teacher-learner approach to better demonstrate the benefit of data fusion. Finally, we conclude by sharing some ideas for directions for future research.

2 RELATED WORK

Transfer learning has been successfully applied in many application areas such as natural language processing [11], sign language recognition [12], adaptive updating of land-cover maps [13], and activity recognition [14]. For example, Diethe *et al.* [8]. examine ADL by using accelerometer sensors to build a model for transfer learning house-to-house. Their results show that by combining active and transfer learning, faster learning with limited labelling resource can be achievable, whilst some limitations focused on space and domain dependency. According to the literature, approaches to transfer learning are categorized based on the means for importing knowledge between the source and target domain, such as instance-based, knowledge-based, and feature-based, as well as dimensions such as cross-participant and cross-device transfer [8].

Researchers have applied transfer learning following various approaches. For example, in TransEMDT they propose a methodology that integrates a decision tree and the k-means clustering technique to adapt a cross-participant activity recognition model [15]. They focus on optimizing the amount of labeled data needed when new participants are presented. TrAdaBoost fits into the category of instance-transfer. It develops a solution based on a Support Vector Machine algorithm in which knowledge from the old domain is taken to solve the task of a new domain [16]. Calatroni *et al.* [5] develop a system capable of transferring knowledge to a new set of wearable sensors based on an already trained model between body motion sensors. They, however, assume that both the source and target domains consisted of similar types of feature, which does not always happen in realistic scenarios since different tasks are handled by the most appropriate and hence often different technologies.

Similarly, Chen *et al.* [6] map the features from two different smart-homes using probabilistic methods to measure the distance between the features. However, the activities concerned are assumed to be detected by a similar feature-space as those presented in the source domain. Hu *et al.* [13] propose a bridge to map sensor devices based on the association of terms, extracted via Web search, that have previously been associated with each sensor device. There are two main assumptions, the activities in the source and target domain are related by the collection of terms, and the underlying feature space is identical between the source and target domain. Zheng *et al.* [29] applied transfer learning when activity recognition labels are different. They also learn the similarity function between the activities in the source and target domain via Web search and specifically transfer the data that have a different label space to the target. Due to the motivation of this paper, we focus on taking advantage of the ubiquitous presence of wearable devices. Considering the resource constraints in this domain, a rapid and power efficient approach is needed.

In this context, data fusion consists on the integration of multi-sensor data that exploits the natural synergy brought by multiple sources to achieve inferences that could not be feasible from a single sensor [6]. There are three categories to consider: the direct fusion of sensor data (*i.e.*, competitive type), representation of sensor data via feature vectors (*i.e.*, complementary type), and processing of each sensor to achieve high-level inference (*i.e.*, cooperative type). For example, in MHTL (Multi Home Transfer Learning) methodology, they follow the cooperative type by combining multiple sources to form different activity recognition models, therefore being able to map target activities based on a more diverse set of source activities [17].

Meaningful scenarios for transfer learning might involve types of device as follows: data from floor sensors in an internal smart floor environment along with pedometer readings when moving outdoors; a smart bed configured with wireless sensor networks and a mobile phone app for sleep monitoring. Regarding the technologies under consideration in our study, video recording provides a fine-grained source of data for activity recognition tasks. One drawback, however, when tracking individuals is the limitation arising from viewing range and fixed positioning [18]. Wearable sensors are another commonly used mechanism to collect activity-related data. These too have challenges such as optimum location and associating their data with meaningful activity recognition outcomes [19]. In this paper, data fusion is used as a strategic way to take advantage of the diversity of data provided by these two-different sets of technology: inertial and thermal vision sensor.

3 METHODOLOGY

Three orthogonally uniaxial accelerometers and Micro Electro Mechanical Systems (MEMS) technology allow sensing components to be miniaturized and have become commonplace in modern technology such as garment devices [20]. In this regard, we focus on detecting walking activity by using accelerometer data retrieved from a Moto 360 2nd generation smartwatch[1]; which for the purposes of this study we will be referring to as the wearable domain. We also use a thermal sensor installed in the Smart Environments Research Group (SERG) laboratory from Ulster University; which for the purposes of this methodology we will be referring to as the environmental domain. The thermal sensor enables real-time processing of frames using an aerial view from the Thermal Vision Sensor (TVS)[2] installed in a 4 x 3.5 meters living space, free of obstacles. Figure 1 shows the view scanned by the TVS detecting a participant in both pre and post-processing frames.

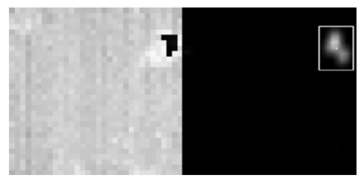

Figure 1: Left: Raw Image Frame from TVS with subject walking. Right: Rectangular Contour encapsulating detected subject.

For the purpose of this study, data is collected from both domains (*i.e.*, wearable and environment) simultaneously, then respective features are extracted in order to build distinct activity recognition models, as illustrated in Figure 2.

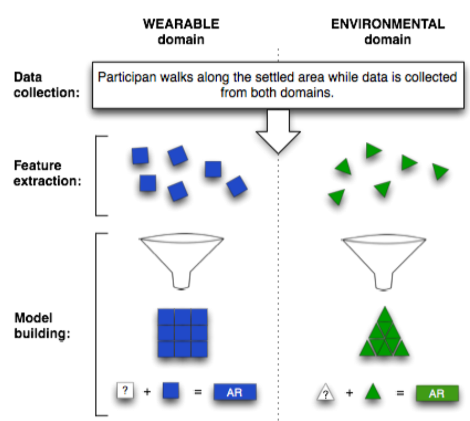

Figure 2: At this stage, features and the activity recognition models themselves are created separately.

[1] https://www.motorola.com.au/products/moto-360

[2] http://www.heimannsensor.com/products_imaging.php

3.1 Wearable domain

The accelerometer sampling rate was set at 50 Hz since we are analyzing short and fine-grained units (*i.e.*, steps). Raw data was pre-processed by removing the DC component and applying a Butterworth filter to clean signal from noise. Due to gravitational values included in the signal, orientation changes of smartwatches can be calculated from the accelerometer readings. Hence, orientation-related features, such as Root Mean Square, are extracted from each accelerometer axis. Non-overlapping fixed length windows of 2 seconds were used.

As a first attempt, we used Weka [21] to explore the modeling of classification to be built. We compared the performance of three activity recognition classifiers: Decision trees (Random Forest) [22], Naive Bayes [23] and Sequential minimal optimization (SMO) [24]. After conducting a single-user training and a cross-validation (10 folds) exercise, it was empirically decided to use the Random Forest classifier, which showed 14% higher performance compared to the other two methods.

3.2 Environmental domain

We have extended a mechanism previously published [26] to calculate the gait velocity while walking along a supervised space observed by the TVS. The vision-based image processing captures every single thermal frame obtained from the TVS and applies binary thresholding which lowers the noise within the threshold. The Gaussian blur technique for smoothing, reducing and filtering out salt-and-pepper noise is applied. Later on, contours are characterized which are necessary to filter out the participant in the thermal frames. A rectangular block is drawn over the contour to keep a track of the moving participant within frames, as shown in Figure 1. The proposed tracking collects real-time inhabitant position using the TVS by monitoring the central location of the inhabitant. These frames are associated with timestamps for real-time tracking. The participant is tracked within the smart room using the TVS for walking and standing activities. All the implementations pertaining to participant monitoring and tracking are developed using the Eclipse IDE and Java-based OpenCV 3.4.1

4 RESULTS AND DISCUSSION

As a proof of concept of the benefit that can be achieved by taking advantage of the synergy created by combining technologies from two different domains (*i.e.*, wearable and environmental) in daily living activities, we have collected data of a participant wearing the smartwatch and walking through the living space with a pre-installed TVS on the ceiling.

Transfer learning has been explored by following the teacher-learner approach [25]. Two initial scenarios, in which each domain plays the role of a teacher have been devised. As illustrated in Figure 3, in the first scenario, we have used the environmental domain's model as a teacher of the wearable domain. Accuracy reaches up to 80% under cross-validation (10 folds). In the second scenario, we use the smart environment activity recognition prediction as a teacher of the wearable model. Accuracy increased 2% compared against the training model itself, when following the similar validation mechanism aforementioned.

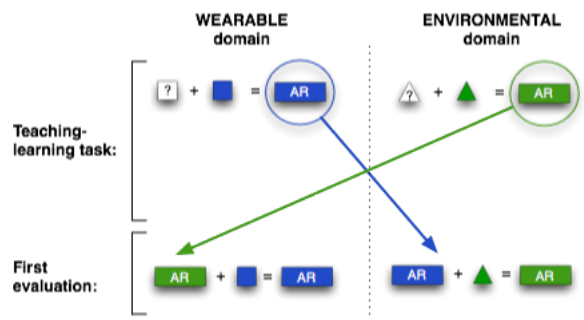

Figure 3: Exploration stage for testing the teacher-learner approach across-technology.

Finally, as per our motivation mentioned in the introductory section of this paper, we have performed a third scenario in which we take advantage of the technology used by fusing data at the feature level. Based on the hypothetical scenario in which a user could have access to both domain's technology, we have built a third model based on the collection of both domain features. This model is then used as teacher to re-train both of the previously created models (*i.e.*, the one consisting of wearable technology and the one consisting of environmental technology). Accuracy of the walking detection improves from 70% up to 80% after the model has been re-trained by combining feature components from the two different technology domains.

5 CONCLUSIONS

In summary, we have presented a proof of concept in which features from heterogeneous technologies collaborate under the teaching-learning approach of transfer learning. Learned experience has been taken from one model to teach the pair, showing an increase of 10% when cross-validating (10 folds) the wearable domain's model. Note that in this paper we have focused on illustrating the benefit of combining different technologies, thus the actual activity recognition model presented in this paper shall be enhanced in further work. In sharing learned knowledge across the two domains, we have faced two main challenges: signal synchronization, due to the fact that for machine learning data is typically organized into discrete values as a basis for feature recognition; and feature selection, due to the inherent dependency on sensor properties and their interpretation in the context of heterogeneous technology.

At this stage of the study, we have identified limited work elsewhere in the literature in which wearable and environmental technology is combined to enhance activity recognition in daily living scenarios. Hence, due the relevance of the topic, as part of our further work we will provide a thorough review of related work in this regard. On the other hand, Also, in addition to providing a more robust activity recognition model, future work will focus on addressing the aforementioned challenges with our main focus on a mechanism to dynamically select those features which will yield higher performance for the models involved.

ACKNOWLEDGMENTS

We thank the reviewers for their feedback and valuable ideas regarding further research.

This project has received partial funding from the European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant agreement No 734355.

REFERENCES

[1] D. M. Karantonis, M. R. Narayanan, M. Mathie, N. H. Lovell, and B. G. Celler, "Implementation of a real-time human movement classifier using a triaxial accelerometer for ambulatory monitoring," *IEEE Trans. Inf. Technol. Biomed.*, vol. 10, no. 1, pp. 156–167, 2006.

[2] G. N. Healy *et al.*, "Objectively measured sedentary time, physical activity, and metabolic risk: the Australian Diabetes, Obesity and Lifestyle Study (AusDiab)," *Diabetes Care*, vol. 31, no. 2, pp. 369–371, 2008.

[3] C. Tudor-Locke, M. M. Brashear, W. D. Johnson, and P. T. Katzmarzyk, "Accelerometer profiles of physical activity and inactivity in normal weight, overweight, and obese U.S. men and women.," *Int. J. Behav. Nutr. Phys. Act.*, vol. 7, no. 1, p. 60, 2010.

[4] A. M. Kriska *et al.*, "Physical activity, obesity, and the incidence of type 2 diabetes in a high-risk population," *Am. J. Epidemiol.*, vol. 158, no. 7, pp. 669–675, 2003.

[5] A. H. Mokdad *et al.*, "Prevalence of Obesity, Diabetes, and Obesity-Related Health Risk Factors, 2001," *JAMA*, vol. 289, no. 1, 2003.

[6] D. L. D. L. Hall, S. Member, and J. Llinas, "An introduction to multisensor data fusion," *Proc. IEEE*, vol. 85, no. 1, pp. 6–23, 1997.

[7] D. Roggen, K. Förster, A. Calatroni, and G. Tröster, "The adARC pattern analysis architecture for adaptive human activity recognition systems," *J. Ambient Intell. Humaniz. Comput.*, vol. 4, no. 2, pp. 169–186, 2013.

[8] S. J. Pan and Q. Yang, "A survey on transfer learning," *IEEE Transactions on Knowledge and Data Engineering*, vol. 22, no. 10. pp. 1345–1359, 2010.

[9] A. Calatroni, D. Roggen, and G. Tröster, "Automatic transfer of activity recognition capabilities between body-worn motion sensors: Training newcomers to recognize locomotion," *Eighth Int. Conf. Networked Sens. Syst.*, 2011.

[10] M. P. Lawton and E. M. Brody, "Assessment of older people: self-maintaining and instrumental activities of daily living," *Gerontologist.*, vol. 9, no. 3, pp. 179–186, 1969.

[11] J. Blitzer, R. McDonald, and F. Pereira, "Domain adaptation with structural correspondence learning," in *Proceedings of the 2006 Conference on Empirical Methods in Natural Language Processing - EMNLP '06*, 2006, p. 120.

[12] A. Farhadi, D. Forsyth, and R. White, "Transfer Learning in Sign language," in *CVPR*, 2007.

[13] L. Bruzzone and M. Marconcini, "Toward the automatic updating of land-cover maps by a domain-adaptation SVM classifier and a circular validation strategy," *IEEE Trans. Geosci. Remote Sens.*, vol. 47, no. 4, pp. 1108–1122, 2009.

[14] A. Bulling, U. Blanke, and B. Schiele, "A tutorial on human activity recognition using body-worn inertial sensors," *ACM Comput. Surv.*, vol. 46, no. 3, pp. 1–33, 2014.

[15] Z. Zhao, Y. Chen, J. Liu, Z. Shen, and M. Liu, "Cross-people mobile-phone based activity recognition," in *IJCAI International Joint Conference on Artificial Intelligence*, 2011, pp. 2545–2550.

[16] W. Dai, Q. Yang, G.-R. Xue, and Y. Yu, "Boosting for transfer learning," in *Proceedings of the 24th international conference on Machine learning - ICML '07*, 2007, pp. 193–200.

[17] D. Rashidi, Parisa and Cook, "Multi Home Transfer Learning for Resident Activity Discovery and Recognition," *Proc. Int. Work. Knowl. Discov. from Sens. Data*, pp. 53–63, 2010.

[18] J. Yang, R. Yan, and A. G. Hauptmann, "Cross-domain video concept detection using adaptive svms," in *Proceedings of the 15th international conference on Multimedia - MULTIMEDIA '07*, 2007, p. 188.

[19] S. J. Pan, I. W. Tsang, J. T. Kwok, and Q. Yang, "Domain adaptation via transfer component analysis," *IEEE Trans. Neural Networks*, vol. 22, no. 2, pp. 199–210, 2011.

[20] C. V Bouten, K. T. Koekkoek, M. Verduin, R. Kodde, and J. D. Janssen, "A triaxial accelerometer and portable data processing unit for the assessment of daily physical activity.," *IEEE Trans. Biomed. Eng.*, vol. 44, no. 3, pp. 136–147, 1997.

[21] M. A. Hall, E. Frank, G. Holmes, B. Pfahringer, P. Reutemann, and I. H. Witten, "The WEKA data mining software: an update," *SIGKDD Explor.*, vol. 11, no. 1, pp. 10–18, 2009.

[22] R. Kohavi, "A Study of Cross-Validation and Bootstrap for Accuracy Estimation and Model Selection," in *Appears in the International Joint Conference on Articial Intelligence (IJCAI)*, 1995, pp. 1–7.

[23] S. Theodoridis, A. Pikrakis, K. Koutroumbas, and D. Cavouras, *Introduction to Pattern Recognition: A Matlab Approach*. 2010.

[24] J. C. Platt, "Sequential Minimal Optimization: A Fast Algorithm for Training Support Vector Machines," 1998.

[25] 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," *Third Int. Conf. Adapt. Self-Adaptive Syst. Appl.*, no. c, pp. 73–78, 2011.