AUTHOR(S): WIJEKOON, A.

TITLE: Reasoning with multi-modal sensor streams for m-health applications.

YEAR: 2018

Publisher citation: WIJEKOON, A. 2018. Reasoning with multi-modal sensor streams for m-health applications. In Minor, M. (ed.) Workshop proceedings of the 26th International conference on case-based reasoning (ICCBR 2018), 9-12 July 2018, Stockholm, Sweden. Stockholm: ICCBR [online], pages 234-238. Available from: http://iccbr18.com/wp-content/uploads/ICCBR-2018-V3.pdf#page=234

OpenAIR citation: WIJEKOON, A. 2018. Reasoning with multi-modal sensor streams for m-health applications. In Minor, M. (ed.) Workshop proceedings of the 26th International conference on case-based reasoning (ICCBR 2018), 9-12 July 2018, Stockholm, Sweden. Stockholm: ICCBR, pages 234-238. Held on OpenAIR [online]. Available from: https://openair.rgu.ac.uk/

Reasoning with Multi-modal Sensor Streams for m-Health Applications

Anjana Wijekoon[1]

School of Computing Science and Digital Media, Robert Gordon University,
Aberdeen AB10 7GJ, Scotland, UK
{a.wijekoon}@rgu.ac.uk

Abstract. Musculoskeletal Disorders have a long term impact on individuals as well as on the community. They require self-management, typically in the form of maintaining an active lifestyle that adheres to prescribed exercises regimes. In the recent past m-health applications gained popularity by gamification of physical activity monitoring and has had a positive impact on general health and well-being. However maintaining a regular exercise routine with correct execution needs more sophistication in human movement recognition compared to monitoring ambulatory activities. In this research we propose a digital intervention which can intercept, recognize and evaluate exercises in real-time with a view to supporting exercise self-management plans. We plan to compile a heterogeneous multi-sensor dataset for exercises, then we will improve upon state of the art machine learning models implement reasoning methods to recognise exercises and evaluate performance quality.

Keywords: Deep Learning · Privileged Learning · Exercise Recognition · Exercise performance Quality

1 Introduction

Maintaining a regular self-managed exercise routine is an essential component when living with Musculoskeletal Disorders(MSDs). Specifically for elderly and people with chronic conditions, it is important to maintain active lifestyle and importantly to adhere to correct execution of exercises. Research on finding technological solutions to support either the prevention or self-management of MSDs has emerged over the last few years. A digital intervention which captures exercises and provides feedback on performance quality at real-time will contribute towards motivating the user to adhere to a regular exercise routine. An effective Digital intervention for self managing MSDs should consist of three main components: intercepting exercises in real-time; recognising exercises; and evaluating performance quality to facilitate personalised feedback generation. In this research we plan to explore each component to design an a optimal digital intervention for self-management of exercises.

Simple sensors on a smart phone are able to identify simple ambulatory activities [11]. Datasets in recent research on HAR [4, 12], Gesture recognition [1], gym exercises [13] and Activities of Daily Living(ADL) [2] use sensors such as inertial sensors, object sensors, pressure sensors and depth sensors. Exercise is a sequence of independent movements of multiple body parts; specifically exercises recommended for low back pain require capturing greater ground surface compared to ADL or ambulatory activities. Hence a smart watch on the wrist is not able to capture an exercise with the level of precision required. Furthermore a wrist sensor is susceptible to noise (due to high freedom of movement) and could temporarily loose data. Aforementioned datasets do not consider these limitations, rendering them inadequate for this research; and it is evident that exercises require capturing different perspectives from multiple sensors. Accordingly we emphasise the need for a data collection in order to identify which sensors can intercept exercises efficiently.

Exercise recognition can be viewed as a special case of Human Activity Recognition (HAR). Research in HAR involves the use of machine learning methods and more recently, deep learning algorithms to reason with sensor data. Many researches shows that deep learning techniques outperform traditional machine learning techniques that use hand-crafted features [10, 11]. Notably most recent research [8, 5, 17] use combinations of deep learning architectures and yield comparatively improved performance. Sensor fusion has been attempted with deep fusion architectures to classify video [6, 9] and reconstruct video and audio [7] by experimenting fusion in different levels of abstraction. Exercise recognition is not widely seen in literature but we find [13] using traditional methods such as Dynamic Time Warping (DTW) and peak counting. Exercise recognition has not been attempted with heterogeneous data streams and with deep learning techniques to the best of our knowledge. Accordingly we first evaluate aforementioned techniques and learn their transferability to the domain of exercises. Next we select advantages techniques from above experiments to build a comprehensive solution for exercise classification with multi-modal data.

Efficient deployment is an important characteristic in any health care digital interventions with direct impact on user acceptance. In this problem domain, the main restriction is the number of sensors. More sensors force the user in to a more restricted setting hence less efficient. We plan to investigate concept Privileged Learning (PL) [14] in order to improve the deployability of reasoning models. PL mimics how humans learn with a teacher; in HAR we interpret PI as deploying a model with less sensors compared to number of sensors available at training. These techniques should contribute towards building robustness in to models to handle missing modalities in real-time which improves usability and efficient deployment.

Performance quality of an exercise can be defined as how much actual execution deviates from correct execution of the exercise performed under the supervision of a physiotherapist. Measuring the deviation is open to interpretation. Recent research in this area show that quality assessment is modelled as a classification task where classes include many wrong executions and a correct execution [3, 15]. This method can be similar to a rule based system; hence unable to locate a problem in real time. We view quality assessment of an exercise execution as a similarity comparison task. We plan to employ similarity based methodologies to compare different representation of exercise executions with correct execution to locate problem areas.

2 Research Aim

The overall goal is to introduce a sustainable digital intervention for self-managing exercise routines. Following review of literature we identity following research questions to achieve aforementioned goal.

1. How to combine multi-modal data streams to improve exercise recognition?
2. How to maintain performance in the presence of noisy and/or missing sensor modalities?
3. How to analyse exercise performance quality by comparing actual and expected multi-modal sensor data?

We outline four objectives in order to answer each research question. First is to compile a multi-modal sensor dataset in the domain of exercises recommended for low back pain. Secondly we will develop a sensor fusion architecture to recognise the most effective sensors and/or features then combine to improve recognition accuracy. Next we implement methodologies to mitigate noise/absence of modalities in deployment. This would enable the network to learn with all modalities but remain robust even with fewer modalities in real time. Finally we plan to introduce a similarity based architecture for comparing sensor data to generate a quality assessment. The resulting solution should localise performance problem to lower level actions of the exercise.

3 Proposed Plan of Research

Comprehensive multi modal sensor data collection for exercises is crucial if we are to address each research question mentioned above. Data collection task will produce a multi-modal sensor dataset for exercise classification and quality assessment. We have selected accelerometers, pressure sensor and a depth sensor to bring together three different modalities of heterogeneous data types; and we have selected seven exercises that are recommended for patients with low back pain.

To achieve Objective 2 first we will investigate how each sensor modality contributes towards accurate classification of exercise movement classes. Next we will explore how a sensor fusion architectures can contribute toward improving previous results. For this we will look at how informed selection of sensors can improve performance. The goal is to create an architecture which will identify the most informative features from different sensors to improve exercise recognition.

Being inspired by the Privileged Learning paradigm [14] we will explore different approaches to address Objective 3. We will investigate how to enforce robustness in sensor fusion model to handle missing modalities; and we will explore generating synthetic data to represent missing modalities at deployment from available sensors.

To address objective 4 we will define a metric to assess exercise performance quality. Specifically the deviation between expected and actual performance will form the basis for this metric. We plan to investigate methods that learn similarities in spatio-temporal data belonging to one classification class to evaluate quality difference. This will call for similarity measures in different abstraction levels of feature embeddings. In order to locate differences in finer detail we will treat exercises as a sequence of primitive actions. Here the idea is to isolate the differences with respect to a primitive action rather than performing a binary evaluation(correct or incorrect).

4 Current Progress

We are at the early stages of data collection task where we compile a multi-modal dataset on exercises for low-back pain. We have identified sensors and exercises we will use is data collection and obtained ethics approval from university ethics committee. We are in the process of recruiting volunteers and collecting data which will continue during the summer of 2018.

A sensor to sensor neural translator for generating missing sensor data was developed and this work is published in ICCBR2018. This work aligns with Objective 3 where we try to minimize number of sensors at deployment for effective deployment. We evaluated this methodology with two datasets (SelfBACK and PAMAP2),both containing ambulatory activities recorded with inertial sensors. Translator method successfully learned dependencies from sensors with different placements and improved k-NN classification accuracy compared to single sensor. These results confirm while we can learn from many sensors, we can re-use these reasoning models in deployment with fewest sensor.

We explored Zero-shot Learning(ZSL) with Matching networks, work presented at The SICSA ReaLX Workshop 2018, where we improved Matching networks[16] to recognise activities never seen during training. We achieve substantially improved performance with modified matching networks compared to original. We will further explore ZSL as it enables a pre-trained network to recognise new activities at deployment. This is desirable when we expand our work from ambulatory activities to exercises where possible number of classes is unmanageably high.

References

1. Bulling, A., Blanke, U., Schiele, B.: A tutorial on human activity recognition using body-worn inertial sensors. ACM Computing Surveys (CSUR) **46**(3), 33 (2014)
2. Chavarriaga, R., Sagha, H., Calatroni, A., Digumarti, S.T., Tröster, G., Millán, J.d.R., Roggen, D.: The opportunity challenge: A benchmark database for on-body sensor-based activity recognition. Pattern Recognition Letters **34**(15), 2033–2042 (2013)

3. Chen, P.C., Huang, C.N., Chen, I.C., Chan, C.T.: A rehabilitation exercise assessment system based on wearable sensors for knee osteoarthritis. In: International Conference on Smart Homes and Health Telematics. pp. 267–272. Springer (2013)
4. Chen, Y., Xue, Y.: A deep learning approach to human activity recognition based on single accelerometer. In: Systems, Man, and Cybernetics (SMC), 2015 IEEE International Conference on. pp. 1488–1492. IEEE (2015)
5. Hammerla, N.Y., Halloran, S., Plötz, T.: Deep, convolutional, and recurrent models for human activity recognition using wearables. In: Proceedings of the Twenty-Fifth International Joint Conference on Artificial Intelligence. pp. 1533–1540. AAAI Press (2016)
6. Karpathy, A., Toderici, G., Shetty, S., Leung, T., Sukthankar, R., Fei-Fei, L.: Large-scale video classification with convolutional neural networks. In: Proceedings of the IEEE conference on Computer Vision and Pattern Recognition. pp. 1725–1732 (2014)
7. Ngiam, J., Khosla, A., Kim, M., Nam, J., Lee, H., Ng, A.Y.: Multimodal deep learning. In: Proceedings of the 28th international conference on machine learning (ICML-11). pp. 689–696 (2011)
8. Ordóñez, F.J., Roggen, D.: Deep convolutional and lstm recurrent neural networks for multimodal wearable activity recognition. Sensors **16**(1), 115 (2016)
9. Pigou, L., Van Den Oord, A., Dieleman, S., Van Herreweghe, M., Dambre, J.: Beyond temporal pooling: Recurrence and temporal convolutions for gesture recognition in video. International Journal of Computer Vision pp. 1–10 (2015)
10. Ronao, C.A., Cho, S.B.: Human activity recognition with smartphone sensors using deep learning neural networks. Expert Systems with Applications **59**, 235–244 (2016)
11. Sani, S., Massie, S., Wiratunga, N., Cooper, K.: Learning deep and shallow features for human activity recognition. In: International Conference on Knowledge Science, Engineering and Management. pp. 469–482. Springer (2017)
12. Stisen, A., Blunck, H., Bhattacharya, S., Prentow, T.S., Kjærgaard, M.B., Dey, A., Sonne, T., Jensen, M.M.: Smart devices are different: Assessing and mitigatingmobile sensing heterogeneities for activity recognition. In: Proceedings of the 13th ACM Conference on Embedded Networked Sensor Systems. pp. 127–140. ACM (2015)
13. Sundholm, M., Cheng, J., Zhou, B., Sethi, A., Lukowicz, P.: Smart-mat: Recognizing and counting gym exercises with low-cost resistive pressure sensing matrix. In: Proceedings of the 2014 ACM international joint conference on pervasive and ubiquitous computing. pp. 373–382. ACM (2014)
14. Vapnik, V., Vashist, A.: A new learning paradigm: Learning using privileged information. Neural networks **22**(5), 544–557 (2009)
15. Velloso, E., Bulling, A., Gellersen, H., Ugulino, W., Fuks, H.: Qualitative activity recognition of weight lifting exercises. In: Proceedings of the 4th Augmented Human International Conference. pp. 116–123. ACM (2013)
16. Vinyals, O., Blundell, C., Lillicrap, T., Wierstra, D., et al.: Matching networks for one shot learning. In: Advances in Neural Information Processing Systems. pp. 3630–3638 (2016)
17. Yao, S., Hu, S., Zhao, Y., Zhang, A., Abdelzaher, T.: Deepsense: A unified deep learning framework for time-series mobile sensing data processing. In: Proceedings of the 26th International Conference on World Wide Web. pp. 351–360. International World Wide Web Conferences Steering Committee (2017)