WIJEKOON, A., WIRATUNGA, N., COOPER, K. and BACH, K. 2020. Learning to recognise exercises for the self-management of low back pain. In Barták, R. and Bell, E. (eds.). Proceedings of the *33rd International Florida Artificial Intelligence Research Society (FLAIRS) 2020 conference (FLAIRS-33)*, 17-20 May 2020, Miami Beach, USA. Palo Alto: AAAI Press [online], pages 347-352. Available from: https://aaai.org/ocs/index.php/FLAIRS/FLAIRS20/paper/view/18460

Learning to recognise exercises for the self-management of low back pain.

WIJEKOON, A., WIRATUNGA, N., COOPER, K. and BACH, K.

2020

The Thirty-Third International
FLAIRS Conference (FLAIRS-33)

Learning to Recognise Exercises for the Self Management of Low Back Pain

Anjana Wijekoon,[1] **Nirmalie Wiratunga,**[1] **Kay Cooper,**[1] **Kerstin Bach**[2]
[1]Robert Gordon University, Aberdeen, UK
[2]NTNU, Trondheim, Norway
{a.wijekoon,n.wiratunga,k.cooper}@rgu.ac.uk, kerstin.bach@ntnu.no

Abstract

Globally, Low back pain (LBP) is one of the top three contributors to years lived with disability. Self-management with an active lifestyle and regular exercises is the cornerstone for preventing and managing LBP. Digital interventions are introduced in the recent past to reinforce self-management where they rely on self-reporting to keep track of the exercises performed. This data directly influence the recommendations made by the digital intervention thus accurate and reliable reporting is fundamental to the success of the intervention. In addition, performing exercises with precision is important where current systems are unable to provide the guidance required. The main challenge to implementing an end-to-end solution is the lack of public sensor-rich datasets to implement Machine Learning algorithms to perform Exercise Recognition (ExR) and qualitative analysis. Accordingly we introduce the ExR benchmark dataset "MEx", which we share publicly to encourage future research. The dataset include 7 exercise classes, recorded with 30 users using 4 sensors. In this paper we benchmark state-of-the-art classification algorithms with deep and shallow architectures on each sensor and achieve performances 90.2%, 63.4%, 87.2% and 74.1% respectively for the pressure mat, the depth camera, the thigh accelerometer and the wrist accelerometer. We recognise the scope of each sensor in capturing exercise movements with confusion matrices and highlight the most suitable sensors for deployment considering performance vs. obtrusiveness.

Introduction

In the Global Burden of Disease Study 2016 (Abajobir and others 2017), low back pain (LBP) was the leading cause of years lived with disability. Clinical guidelines recommend for LBP patients therapeutic exercise, either as a preventive measure or as part of rehabilitation. At first, exercise is supervised by a physiotherapist, and thereafter a self-management programme is created to empower people to manage their own health conditions. Despite programme adherence being important to achieving a positive patient outcome, recent studies show that non-adherence is as high as 70% among self-managing patients (Essery et al. 2017). Here adherence refers to correct performance of exercises as well as following advice on exercise frequency. The innovation opportunity for health technologies is to support people to effectively self-manage through personalised interactive feedback that improves adherence.

Evidence suggests that for a successful large-scale impact, digital health intervention programmes must unobtrusively monitor adherence (Cooper et al. 2017; Bach et al. 2016). Recent advances in pressure sensing mats (Cheng et al. 2016), Depth camera & inertial sensors (Chen et al. 2014) have created new modalities for unobtrusive monitoring, but none have been used to close the loop between sensing, monitoring, and feedback to address the problem of poor adherence to physiotherapy programmes (Franco et al. 2015; Palazzo et al. 2016).

An end-to-end adherence monitoring programme must consider three main components; exercise recognition, performance quality evaluation and feedback generation. In this paper we focus on recognition, and critically review advances made in the related research areas of Human Activity Recognition (HAR) (Yao et al. 2017; Ordóñez and Roggen 2016) and Exercise Recognition (ExR) (Sundholm et al. 2014; Xiao et al. 2018). The absence of transferable research between HAR and ExR is noticeable but this is in part at least attributed to the lack of a publicly available ExR datasets. It is our aim to address these challenges with a view to implementing a comprehensive adherence monitoring programme. Accordingly in this paper we make the following contributions:

- present the multi-modal heterogeneous sensor dataset "MEx" that is publicly available for Exercise Recognition and Human Activity Recognition research;
- provide benchmark measures using three classification algorithms k-NN, SVM and MLP for comparative performance analysis; and
- explore shallow and deep architectures for feature representation learning with the MEx dataset.

Rest of the paper is organised as follows. Related Work section discusses current research in the domain of ExR and next we detail the dataset with pre-processing recommendations in Section "MEx Dataset". In Section "Exercise Recognition with MEx" we present the benchmark performance analysis including evaluation methodology and results and confusion matrices. Finally conclusions with future plans appear in Section "Conclusions".

Related Work

Research in Exercise Recognition (ExR) spans a number of application areas such as callisthenics, weight exercises, yoga and sports. Inertial Measurement Units (IMUs) sensors are widely used in literature (Velloso et al. 2013; Burns et al. 2018; Guo, Wang, and Yang 2018), but some have explored sensors such as Pressure mats (Sundholm et al. 2014; Zhou et al. 2016), Channel State Information (Xiao et al. 2018) and Electrocardiograms (Qi et al. 2018). ExR is often viewed as classification of many discrete labels given a sensor data stream. Often these recognition algorithms use a manual feature extraction pipeline followed by a classification algorithm such as k-NN (Sundholm et al. 2014; Xiao et al. 2018), Random Forest (Velloso et al. 2013), Decision Trees (Zhou et al. 2016) or HMM (Qi et al. 2018).

While Deep Learning methods (CNN and LSTM) are the state-of-the-art in HAR (Yao et al. 2017; Ordóñez and Roggen 2016), literature suggest that these methods are rarely considered with ExR (Burns et al. 2018). For instance, authors of (Burns et al. 2018) use a recurrent architecture to recognise shoulder rehabilitation exercises with wrist worn IMU data streams and achieve 88.9% accuracy; their dataset is not publicly available and their methods cannot be transferred to other exercise domains due to lack of sensors that capture movements from other body parts except the wrist. This proprietary nature of research in ExR domain (algorithms and data) results in non-transferable knowledge and is a barrier to advancing the state-of-the-art.

MEx Dataset

MEx is a sensor-rich dataset collected for 7 exercises with four sensors publicly available at the UCI Machine Learning Repository[1]. In this section we will present details on data collection protocol, sensor specifications, exercises and recommended pre-processing steps.

Data Collection Protocol

The data collection task included 30 participants. 60% of the participants were female and 40% were male. 47% of the group were in the 18-24 age category and the rest were dispersed among the ages from 24 to 54. 8 of the 30 participants had a good understanding of the exercises as they were either physiotherapists or physiotherapy students. A Physical Activity Readiness Questionnaire (PAR-Q) evaluated the physical fitness of each participant prior to data collection. Seven exercises were selected by a physiotherapist for this data collection; 1-Knee Rolling, 2-Bridging, 3-Pelvic Tilt, 4-Bilateral Clam, 5-Repeated Extension in Lying, 6-Prone Punch and 7-Superman. They are frequently used for the prevention or self-management of LBP (Figure1).

At the start of the session, the user was given a sheet with instructions for each exercise. Each exercise is described with a starting position and set of actions. During the session, the researcher demonstrated an exercise to the user and then the user performed the exercise for approximately 60 seconds while being recorded with four sensors. During the recording, the researcher did not provide any advice or counting to enforce rhythm. For exercises where it was suggested to hold a position for 5 or 2 seconds, the user was instructed at the beginning to keep count by themselves to preserve their natural rhythm. Our goal was to capture individual nuances of each user which replicates a scenario where a patient performs these exercises at home without the guidance of the physiotherapist.

Figure 1: Exercises in the MEx dataset

Sensors

We explored the state of the art sensor technologies and were advised by our health partners to select three sensor modalities to capture these movements; Obbrec Astra Depth Camera [2], Sensing Tex Pressure Mat [3] and Axivity AX3 3-Axis Logging Accelerometer [4]. The goal is to explore their capabilities to capture exercises independently as well as an ensemble while considering the obtrusiveness when deploying sensors in the real-world. Accordingly we select the following placements for sensors; two accelerometers on the wrist and the thigh of the user; the pressure mat was used as a exercise mat where the user lays on to perform the exercises; the depth camera was placed above the user facing downwords recording an aerial view. In addition, the top of the depth camera frame was aligned with the top of the pressure mat and the user is asked to position their shoulders such that the face is not recorded in the depth camera or pressure mat data. The tri-axial accelerometers record data at $100Hz$ frequency within the range of $\pm 8g$. The pressure mat and the depth camera record gray scale frames at $15Hz$ frequency and frame sizes are 32×16 and 240×320 respectively. Figure 2 shows a visualisation of each sensor data type. The four sensors, the thigh accelerometer, the wrist accelerometer, the pressure mat and the depth camera will be referred as ACT, ACW, PM and DC in the rest of this paper.

Pre-processing for Supervised Learning

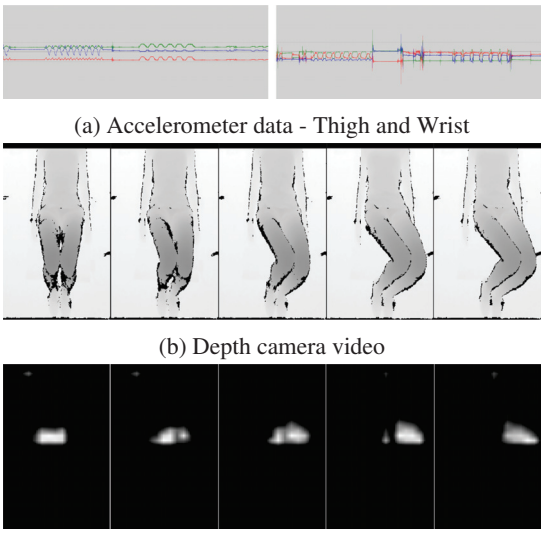
A sliding window method is applied on an individual sensor data stream to create train and test data instances for supervised learning. We use the window size of 5 seconds and an overlap of 3 seconds where each window forms an instance and is labelled with the exercise class. This results in a dataset of 6240 instances (208 instances per user and 30 users) on average per sensor.

[1] https://archive.ics.uci.edu/ml/datasets/MEx

[2] https://orbbec3d.com/product-astra-pro/

[3] http://sensingtex.com/sensing-mats/pressure-mat/

[4] https://axivity.com/product/ax3



(a) Accelerometer data - Thigh and Wrist

(b) Depth camera video

(c) Pressure mat data

Figure 2: Raw data visualisation

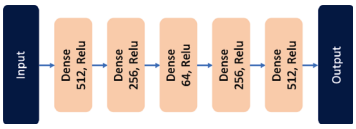
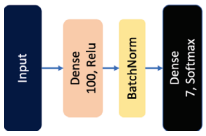
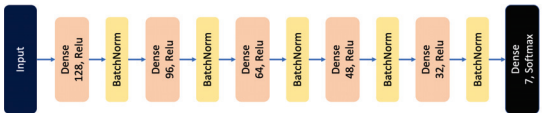


Figure 3: Auto-encoder Architecture



(a) Shallow ANN



(b) Deep ANN

Figure 4: ANN Architectures

Additionally, we recommend a set of pre-processing steps for each sensor modality. A reduced frame rate of 1 frame/second is used with DC and PM data and the DC data frames are compressed from 240×320 to 12×16. The inertial sensor data from ACW and ACT are pre-processed using the Discrete Cosine Transformation (DCT); known to out-perform other feature-transformation methods (Sani et al. 2017). DCT decomposes a signal into constituent cosine waves and returns the ordered sequence of frequency coefficients representing each wave. DCT is applied to each axis of the accelerometer data and the final feature vector of length 180 is formed by appending the 60 most significant coefficients from the 3 axes x, y and z. Above hyper-parameters are recommended following an initial exploratory study [5] while considering performance and computational memory requirements.

Exercise Recognition with MEx

In this section we study the role of both shallow and deep features for classification algorithms for Exercise Recognition (ExR). Accordingly we compare a number of state-of-the-art classification algorithms and feature representation methods for the task of ExR. The three classification algorithms used for this purpose are:

kNN: K-Nearest Neighbours algorithm; we present results with k=1 and k=3;

SVM: Support Vector Machine classifier with a Radial Basis Function kernel; and

MLP: Multi-layer Perceptron as the classifier with a feature representation method followed by a softmax activation layer.

The aim of our evaluation is to find the most optimal feature representation method for each sensor modality. Accordingly we conduct a comparative study with a comprehensive list of shallow and deep feature representation methods that are recognised as the state-of-the-art in recent Deep Learning Literature.

Raw: Raw sensor data (flattened if required)

DCT: For ACT and ACW, the 3 axial inertial data are converted in a DCT feature vector (as described in Section).

AE: For PM and DC visual data, an Auto-encoder (AE) (Figure 3) model learns to reconstruct itself and the mid hidden layer with the lowest dimension is used as the feature representation. Here AE produces an abstract feature vector of size 64.

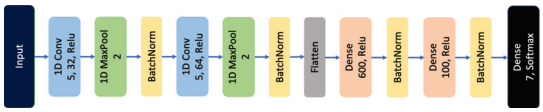
ANN: Artificial Neural Network, comprised of a single or multiple layers of densely connected hidden layers. Each hidden layer is followed by a Batch Normalisation layer to normalise the output and to avoid over-fitting. Both variations below output a feature vector of size 100.

- Shallow-ANN: Consist of one hidden layer densely connected with 100 hidden units and "Relu" activation function (Figure 4a).
- Deep-ANN: Consist of five hidden layers as in Figure 4b.

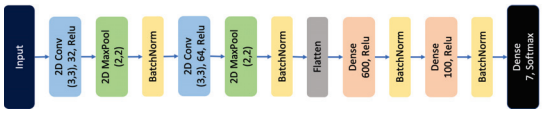
CNN: Convolutional Neural Networks consist of a sequence of blocks where each block is formed by a convolution layer, a max pooling layer and batch normalisation. We explore three variations (that create output vectors of size 100), to suit different sensor modalities as follows:

- DCT-1D: For ACT and ACW; 1-dimensional convolutions (kernel size 5) where the number of channels is 1 and the input is DCT features of length 180. (Figure 5a).
- Raw-1D: Comprised of 1-dimensional convolutions (kernel size 5) as in Figure 5a. For ACT and ACW, the input is a raw data stream of length 500 (5 second window with $100Hz$ frequency) with 3 channels (x, y and

[5] https://arxiv.org/abs/1908.08992



(a) 1D convolution model



(b) 2D convolution model

Figure 5: CNN Architectures

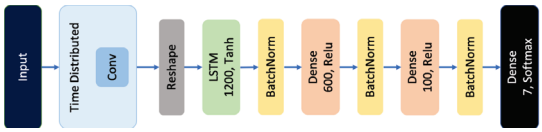


Figure 6: LSTM Architecture

z). For PM and DC data, a frame from each time stamp is flattened and appended together to create the input feature vector with 1 channel.

- 2D: For PM and DC data; 2-dimensional convolutions (kernel size 3×3) (Figure 5b). Frames within a time window are appended to form a 2D vector with 1 channel.

LSTM: The Long-short Term Memory Neural Network; as with the ANN and CNN use Batch Normalisation for regularisation. Additionally a convolutional low-level feature representation is learnt. Specifically we refer to this as a time distributed block where the enclosing "conv" embedding is applied to each frame within the time window simultaneously. Each frame, now represented by these CNN feature representation is the input to the LSTM layer one at a time helping to learn the temporal dependencies within the time window as in Figure 6. Given the differences between our modalities we explore three LSTM variations as follows.

- DCT-1D-CNN: Time distributed 1D Convolution architecture (similar to Figure 5a) for the DCT features of ACT and ACW.
- Raw-1D-CNN: Time distributed 1D Convolution architecture (similar to Figure 5a) suited for all sensor modalities.
- 2D-CNN: Time distributed 2D Convolution architecture (similar to Figure 5a) for PM and DC data.

Evaluation

We adopt the Leave-One-Person-Out (LOPO) evaluation methodology where each fold consists of an individual person's data. Accordingly we train a model with 29 user data and test with 1 user data and repeat for 30 folds. This methodology emulates a real-life deployment setting where end-user data is not available during training. The Mean Macro F1-score (i.e. dataset is class balanced) averaged across all 30 folds is presented as the performance measure. We test for statistical significance at 95% confidence level with Wilcoxon signed-rank test when selecting the best performing architecture.

All MLP and AE models were implemented using Keras and TensorFlow libraries for Python. MLP models are trained end-to-end for 50 epochs, minimising the loss of Categorical Cross-entropy using the AdaDelta optimiser. AE models are trained for 100 epochs, minimising the loss of Mean Squared Error using the AdaDelta optimiser. The kNN, SVM models are implemented using scikit-learn Python libraries.

Classifier	Embedding	ACT	ACW	DC	PM
1-NN	DCT	76.1	45.2	-	-
	Raw	-	-	68.2	56.9
	AE	-	-	72.4	40.9
3-NN	DCT	76.3	46.5	-	-
	Raw	-	-	67.4	56.5
	AE	-	-	73.4	37.1
SVM	DCT	84.7	47.7	-	-
	Raw	-	-	73.3	38.3
	AE	-	-	78.2	70.0
MLP-ANN	ShallowANN	86.7	56.5	61.8	63.9
	DeepANN	84.4	54.1	66.8	67.1
MLP-CNN	DCT-1D	87.9	56.1	-	-
	Raw-1D	72.7	40.1	82.4	70.6
	2D	-	-	**87.2**	69.4
MLP-LSTM	DCT-1D-CNN	**90.2**	**63.4**	-	-
	Raw-1D-CNN	70.4	40.2	83.6	**74.1**
	2D-CNN	-	-	78.2	70.8

Table 1: Results: F1-score(%) for ExR

Results

Table 1 presents the F1-score(%) results for the classifier and feature representation method combinations, obtained by individual sensors. In general MLP classifiers achieved the best performance across different feature representation methods compared to k-NN and SVM. In addition, MLP with deep features performed comparatively better than the shallow features. Overall best performances for ACT, ACW and PM sensors were obtained with the 1D-CNN-LSTM architecture, and for DC sensor with the 2D-CNN architecture (highlighted in bold text). DC which is predominantly a visual sensor as expected benefited from having feature representations that are extracted by the 2D-CNN architecture compared to the 1D-CNN or LSTM architectures. In contrast ACW and ACT sensors with time-series data found learning temporal dependencies with 1D-CNN to be more advantageous. PM data surprisingly had sensor preferences that were closely aligned with ACT and ACW sensors rather than to the DC (despite the similarities with DC data).

Recognition with ACT and ACW data were best with the 1D-CNN-LSTM architecture (with 95% confidence). As expected learning temporal dependencies with LSTM results

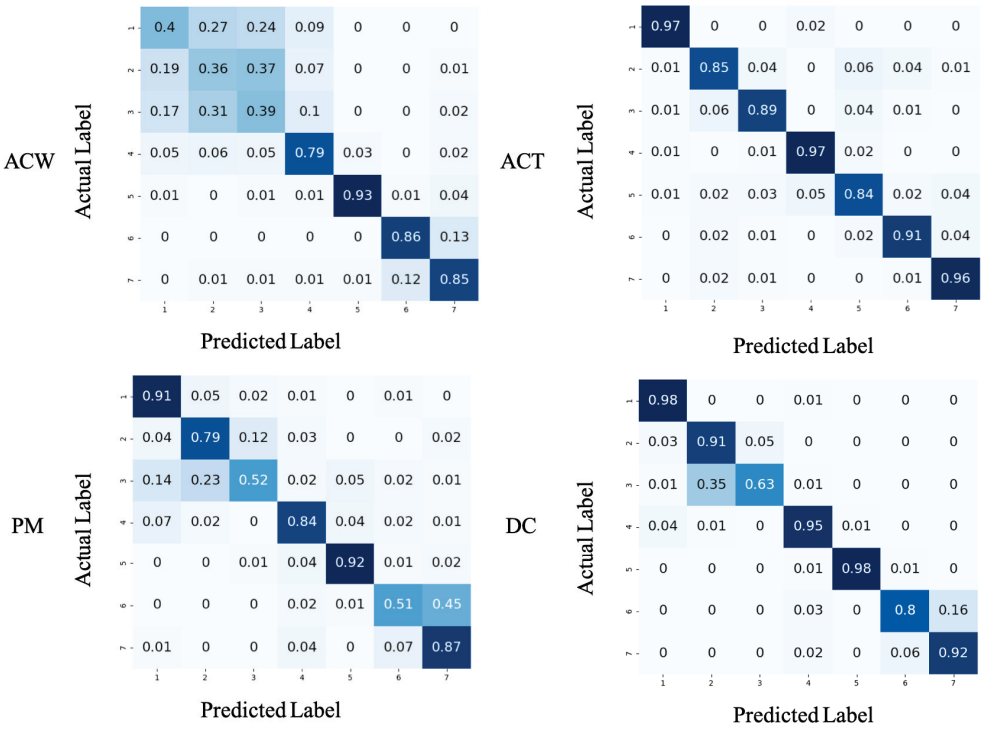


Figure 7: Confusion Matrices

in better feature representation for accelerometer data. Note also with the DCT vs. raw accelerometer data, we found deep architectures using DCT data to significantly outperform those using just the raw data (an increase of 19.71% and 23.19% with LSTM and an increase of 15.20% and 15.92% with CNN respectively for ACT and ACW). These results further confirm the evidence seen in literature comparing raw vs. transformation based feature representation methods (Sani et al. 2017).

Best performances for DC and PM data were achieved with 2D-CNN and 1D-CNN-LSTM architectures respectively where they significantly outperformed (with 95% confidence) all other architectures. k-NN and SVM performs poorly with DC and PM data suggesting the importance of learning feature representations with visual data compared to raw data. DC data achieved a significant performance improvement with the AE reconstruction method over raw data for 1-NN, 3-NN and SVM by 4.20%, 6.02% and 4.95% respectively. For PM data we observe mixed results where a significant performance improvement (31.71%) is seen with SVM but not with kNN. When comparing Deep and Shallow ANN, the results affirm that visual sensor data are best learned with deep architectures.

Confusion Matrices

A confusion matrix visualise the effectiveness of a classification algorithm. We present the confusion matrices for the best performing algorithm for each individual sensor in Figure 7. Each raw represent an exercise class and the correct predictions are on the diagonal. It is not surprising that the ACW sensor is unable to differentiate between exercises 1, 2 and 3 where wrists are kept stationary. Similarly PM and DC sensors find exercise pairs 2, 3 and 6, 7 hard to differentiate because of the similar movements of torso and hands. In contrast, we observe that ACT sensor performs significantly better among all exercise classes. We note with exercise 6, although the thigh remains stationary, it has movements that are unique compared to the other exercise, and so ACT performance is not penalised (compared to ACW sensor).

In summary ExR results emphasised the classification algorithms and feature representation methods that are optimal for each sensor. Importantly we note that sensor modalities with data heterogeneity prefer different feature representation methods to highlight their inherent characteristics that may get overlooked with generic functions. Confusion matrices highlighted the capacity of each sensor to perform ExR as a standalone sensor. ACT sensor successfully identified all seven exercises with high precision but we highlight that the performance may get penalised with a different set of exercises. Considering the obtrusiveness in deployment, DC sensor is the most invasive sensor when installing at a home environment, but wearable sensors and PM sensor which is similar to a yoga mat does not introduce much obstruction. Considering all these factors, we realised the need to reason with multiple sensors simultaneously. Accordingly in future we will look at multi-modal sensor fusion methods to recognise exercises with the MEx dataset.

Conclusions

This paper presents the MEx: Multi-modal Exercises Dataset for Human Activity Recognition and benchmark performance on standard classification algorithms. This dataset was presented with en route to implementing an end-to-end digital intervention for exercise adherence monitoring. The dataset contains 7 exercises recorded with four sensors of heterogeneous data types. Our comparative study suggests that Deep LSTM and CNN models achieve best performances. In addition with confusion matrices we explore the capacity of each sensor to perform independently and highlight the need for sensor fusion methods. Next we plan to explore multi-modal sensor fusion methods with attention mechanisms to improved performance while preserving unobtrusiveness in the sensor setup. This work is contributing towards implementing an exercise recognition algorithm with multiple sensors and further more towards performance quality assessment by comparing exercise performances with recommended guidelines.

Acknowledgments

This work is part funded by SelfBACK. The SelfBACK project is funded by the European Union's H2020 research and innovation programme under grant agreement No. 689043.

References

Abajobir, A. A., et al. 2017. Global, regional, and national incidence, prevalence, and years lived with disability for 328 diseases and injuries for 195 countries, 1990–2016: a systematic analysis for the global burden of disease study 2016. *The Lancet* 390(10100):1211–1259.

Bach, K.; Szczepanski, T.; Aamodt, A.; Gundersen, O. E.; and Mork, P. J. 2016. Case representation and similarity assessment in the self back decision support system. In *Int. Conf. on Case-Based Reasoning*, 32–46. Springer.

Burns, D. M.; Leung, N.; Hardisty, M.; Whyne, C. M.; Henry, P.; and McLachlin, S. 2018. Shoulder physiotherapy exercise recognition: ML the inertial signals from a smartwatch. *Physiological measurement* 39(7):075007.

Chen, C.; Liu, K.; Jafari, R.; and Kehtarnavaz, N. 2014. Home-based senior fitness test measurement system using collaborative inertial and depth sensors. In *2014 36th Annual Int. Conf. of the IEEE Engineering in Medicine and Biology Society*, 4135–4138. IEEE.

Cheng, J.; Sundholm, M.; Zhou, B.; Hirsch, M.; and Lukowicz, P. 2016. Smart-surface: Large scale textile pressure sensors arrays for activity recognition. *Pervasive and Mobile Computing* 30:97–112.

Cooper, K.; Schofield, P.; Klein, S.; Smith, B. H.; and Jehu, L. M. 2017. Exploring peer-mentoring for community dwelling older adults with chronic low back pain: a qualitative study. *Physiotherapy* 103(2):138–145.

Essery, R.; Geraghty, A. W.; Kirby, S.; and Yardley, L. 2017. Predictors of adherence to home-based physical therapies: a systematic review. *Disability and rehabilitation* 39(6):519–534.

Franco, M. R.; Howard, K.; Sherrington, C.; Ferreira, P. H.; Rose, J.; Gomes, J. L.; and Ferreira, M. L. 2015. Eliciting older people's preferences for exercise programs: a best-worst scaling choice experiment. *Journal of physiotherapy* 61(1):34–41.

Guo, M.; Wang, Z.; and Yang, N. 2018. Aerobic exercise recognition through sparse representation over learned dictionary by using wearable inertial sensors. *Journal of Medical and Biological Engineering* 38(4):544–555.

Ordóñez, F., and Roggen, D. 2016. Deep convolutional and lstm recurrent neural networks for multimodal wearable activity recognition. *Sensors* 16(1):115.

Palazzo, C.; Klinger, E.; Dorner, V.; Kadri, A.; Thierry, O.; Boumenir, Y.; Martin, W.; Poiraudeau, S.; and Ville, I. 2016. Barriers to home-based exercise program adherence with chronic low back pain: Patient expectations regarding new technologies. *Annals of physical and rehabilitation medicine* 59(2):107–113.

Qi, J.; Yang, P.; Hanneghan, M.; Waraich, A.; and Tang, S. 2018. A hybrid hierarchical framework for free weight exercise recognition and intensity measurement with accelerometer and ecg data fusion. In *2018 40th Annual Int. Conf. of the IEEE Engineering in Medicine and Biology Society (EMBC)*, 3800–3804. IEEE.

Sani, S.; Massie, S.; Wiratunga, N.; and Cooper, K. 2017. Learning deep and shallow features for human activity recognition. In *Int. Conf. on Knowledge Science, Engineering and Management*, 469–482. Springer.

Sundholm, M.; Cheng, J.; Zhou, B.; Sethi, A.; and Lukowicz, P. 2014. Smart-mat: Recognizing and counting gym exercises with low-cost resistive pressure sensing matrix. In *Proceedings of the 2014 ACM UbiComp.*, 373–382. ACM.

Velloso, E.; Bulling, A.; Gellersen, H.; Ugulino, W.; and Fuks, H. 2013. Qualitative activity recognition of weight lifting exercises. In *Proc. 4th Augmented Human Int. Conf.*, 116–123. ACM.

Xiao, F.; Chen, J.; Xie, X. H.; Gui, L.; Sun, J. L.; and none Ruchuan, W. 2018. Seare: A system for exercise activity recognition and quality evaluation based on green sensing. *IEEE Transactions on Emerging Topics in Computing*.

Yao, S.; Hu, S.; Zhao, Y.; Zhang, A.; and Abdelzaher, T. 2017. Deepsense: A unified deep learning framework for time-series mobile sensing data processing. In *Proceedings of the 26th International Conference on World Wide Web*, 351–360.

Zhou, B.; Sundholm, M.; Cheng, J.; Cruz, H.; and Lukowicz, P. 2016. Never skip leg day: A novel wearable approach to monitoring gym leg exercises. In *IEEE Int. Conf. Pervasive Computing and Communications*, 1–9. IEEE.