Abstract

Purpose: Physical activity classifiers are typically trained on data obtained from sensors at a set orientation. Changes in this orientation (such as being on a different wrist) result in performance degradation. This work investigates a method to obtain sensor location and orientation invariance for classification of wrist mounted accelerometry via a technique known as domain adaption.

Methods: Data was gathered from 16 participants who wore accelerometers on both wrists. Physical activity classification models were created using data from each wrist and used to predict activities when using data from the opposing wrist. Using subspace alignment domain adaption, this procedure was then repeated to align the training and testing data before the classification stage.

Results: Prediction of activity when using data where the wearer's wrist was incorrectly specified resulted in a significant (p=0.01) decrease in performance of 12%. When using domain adaption this drop in performance became negligible, mean difference <1%, p = 0.73.

Conclusion: Domain adaption is a valuable method for achieving accurate physical activity classification independent of sensor orientation in wrist worn accelerometry.

Introduction

Accelerometers are now being utilized in large scale observational studies as well as controlled trials of physical activity (PA) interventions to more precisely estimate levels of physical activity compared to self-report [1,2]. Such studies use body mounted accelerometers to gather data about the accelerations that the sensor undergoes; which is used as a surrogate measure for the participants acceleration. Converting the raw acceleration data into an output that has behavioural meaning permits inferences to be made about the relationship between PA behaviours and health. The two common forms of accelerometer output are: `counts`, a form of filtered aggregated data [3] and the raw acceleration itself [4]. It is inevitable that aggregation causes information loss; consequently methods focusing on the raw data are becoming more commonplace. As acceleration is only a surrogate of the true PA performed, methods for converting the data into behavioural metrics are required. A technique of growing popularity is activity classification via machine learning, which creates a model, based on some training data, that can predict the activity type performed based on the acceleration data [3,4].

The earliest examples of using machine learning to classify activity via acceleration used networks of accelerometers placed around the body (wrists, ankles, chest and waist). While these were able to achieve a high performance, such networks were unfeasible to wear in in everyday scenarios [4]. Additionally the cost was considerably higher than using a single accelerometer. With the increasing classification performance obtainable using a single accelerometer [5], it is now more common to just use one, with the device being worn either on the wrist or hip [6]. While the hip appears to allow for a higher performance in some areas, using the wrist is associated with higher compliance [7] and is becoming the location of choice in many large studies such as the UK Biobank study [1].

While usage of the wrist has become more common, there are no clear guidelines on which wrist should be used (and at which orientation the device should be), an issue further exacerbated by poor participant adherence to device wear and orientation guidelines. Several studies exist making use of dominant/non-dominant wrists [5] as well as others that use left/right wrists [8], severely limiting the interoperability of these systems. Different wrists obtain different acceleration values for classifying the intensity of the same activity on the same participant [9,10]. Furthermore, using activity classifiers to classify activity using data where the wearer's wrist is incorrectly specified has been shown to reduce performance by up to 12% [8], compared to using the specified wrist. A related issue occurs when the device is placed upside down, which reverses all acceleration values gathered; needless to say this also decreases the performance. Therefore it is clear that a method that can either identify which wrist the data was gathered on or allows for location/orientation invariance is required.

Typically, reductions in performance when applying a classification model to different wrists is not remarked upon, and it is assumed that all participants will wear the device on the same wrist or location as advised [1]. This assumption is not guaranteed, especially when the participant places the sensor themselves, and the violation of this assumption may partly explain the poor performance of activity classifiers outside of their training data [11]. A possible solution is to make use of features derived from the raw accelerations that are orientation invariant, meaning that the features will be identical regardless of the orientation of the sensor [12]. This prevents the performance reduction, but limits the available features that can be used in the classification, which in turn may compromise performance. Another more effective approach by Gjoreski et al [8] trained the classifier with data from both wrists, resulting in a higher performance than using a single wrist. However, data from the opposing wrist may confuse classification, reducing the performance. An additional limitation of this method, is that data from both wrists must be collected in the training stage, increasing both the cost of the data gathering process and the burden to the participant.

Standard machine learning aims to create a classifier based on labelled data from a training dataset that can correctly classify unlabelled data collected in the same way. Here the training data is said to come from the *source domain*, while the unlabelled data for classification is said to come from the *target domain*; in standard machine learning the source and target domains are identical. When the target data comes from a different domain, the classification performance drops due to the classifier only being suited to the source domain. Domain adaption (DA) methods seek to adapt data from the target domain to the source domain so that good performance is achieved [13]. Clearly, the source and target domains must be related for DA to be successful. In this paper we investigate whether DA can be used to adapt data measured from the “wrong” wrist (the target domain) to the training data in order to achieve good classification performance regardless of the wrist on which accelerometer is worn. DA approaches only require acceleration data from one wrist, as well as having no limitations of the features used, which is a substantial advantage over other approaches.

As mentioned above, DA involves allowing a classification model created on a source domain to be applied to a different (but related) target domain without suffering performance reduction. DA specifically deals with techniques for moderating the performance reduction when classifying over different distributions, making it well suited to attenuating the performance drop from using differing wrists. DA has seen some use in activity classification and accelerometry studies [14], however no work has been found that allowed for location/orientation invariance using DA, hence this work is novel. The similarity between visual data and time series data (their local correlations and innate structure) allows for DA algorithms designed for visual applications to function well on time series data. Although DA is strongly linked with another field known as Transfer Learning, and most literature uses these terms interchangeably, DA will be used in this work [13].

Therefore this paper will examine DA as a possible solution to achieving location/orientation invariance in activity classification via accelerometry. If successful, this will allow for all data to be used for training classifiers instead of only the data that is obtained from the same wrist or location.

Method

All code is available in the GitHub repository [https://github.com/PARIS-Uni/Domain-Adaption].

Procedures

The data for this study was collected from 16 people (mean age 25 ± 4.7, mean BMI 23 ± 2.6) wearing a wrist-worn tri-axial GENEActiv [10] accelerometer on each wrist, set at a sampling rate of 100 Hz.

Each participant followed a given exercise protocol while supervised consisting of a variety of activities of daily living (see supplement 1 for full details). Activities that are highly similar tend to be confused for one another by many activity classification methods [3-6], e.g., watching television and working at a desk. Activities were grouped according to posture and how the activity may be performed:

- Lying: Lying
- Desk work: Watching TV, working at a desk
- Walking: Slow/moderate/brisk and stair climbing
- Household tasks: Vacuuming/washing dishes/folding laundry
- Standing: Standing still

The University of [redacted] ethics committee approved the data collection protocol (20/4/2017) and informed consent was obtained before participation.

Comparison with non-DA approaches:

In order to evaluate the effect of DA on the performance reduction observed when applying activity classification models to the 'wrong wrist' a series of comparisons were made between different DA and non-DA approaches. Each of the approaches is described below and summarised in Table 1.

Five DA and non-DA approaches were evaluated:

- The Criterion approach: This refers to creating and testing the classifiers on the same wrist. This approach serves as the gold standard for performance.
- The DA approach: In the approach a classifier was trained on data of one wrist. DA was used to adapt the target data collected from the opposite wrist data to the source domain so that the trained classifier could be applied to the

adapted target data. This served as the main focus of this study to identify if DA reduced the performance drop.

- The Not-Domain Adaption (NDA) approach: Here a classifier was trained from the data of one wrist. The resultant classifier was then used to classify the data from the opposite wrist with no modification. This method served as the control.
- The Not Applicable (NA) approach: Here a classifier was trained from the data of one wrist. DA was then used with the same wrist data serving as the target domain. The resultant domain adapted classifier was then used to classify the data from the same wrist. This method served to investigate the effect of using DA when it is not required, in circumstances where the wrist placement of the accelerometer is unknown.
- The Amalgam approach (AMAL): Here a classifier was trained on data from both wrists, similarly to [8]. The resultant classifier was then used to classify the data from just one wrist with no modification to examine if using data from both wrists allowed for wrist invariance without the need for DA.

TABLE 1

Classification Procedure

There were six steps to the classification procedure: these were data pre-processing, feature extraction, normalization, feature reduction, Domain Adaption and classification.

1. Pre-processing: In this step the training acceleration data at each time point was annotated with the corresponding activity label. The data was divided into fixed non-overlapping windows 1.6 seconds in length, based on common window sizes identified by Banos [15].

2. Feature extraction: Features were then extracted from each window. The features extracted from the windows were those identified by Trost et al [5], and represent some of the highest performing features as well as having been validated on both lab-based and free-living data [5,11]. There were 39 features in total, representing various aggregate statistical and frequency measures about the acceleration (see supplement 2 for full details).

3. Normalization: Normalization is a procedure used to ensure that all features have a similar variance, meaning that all features have an equal weighting in the data set. This procedure is not required for one of the classification methods used in this work (Random Forest), however it is required for the DA step. As such should be performed regardless of classification method utilised.

4. Feature Reduction (FR): This step is concerned with reducing the number of features used in the classification model. Often a larger number of features will lead to a decreased ability of the model to generalise to unseen data, resulting in a lower performance. FR techniques are commonly used in activity classification work [16]. The DA method used in this work makes use of a FR stage and an adaptive stage. In order to ensure comparability between all methods the same FR stage was performed regardless of whether DA was utilised. This ensures that the effect of the adaptation stage is not masked by FR. The specific form of FR used in this work is Principal Component Analysis (PCA). PCA works by projecting high dimensional data (in this case 39) into a smaller number of dimensions---a subspace---while preserving as much variance as possible. The resulting low-dimensional features are linear combinations of the original, high-dimensional features. This technique has commonly been used in activity classification and a more detailed explanation can be found in the work of Lever et al [17]. PCA requires a parameter (*k*) to be chosen, which is the dimension of the lower dimensional subspace. In this work, *k* is chosen to be 12 (as determined with cross validation) unless stated otherwise, although the performances for all values of *k* were evaluated.

5. The pre-processing, feature extraction, normalisation and feature reduction steps are applied to both the source and target data. As described below, in the DA step, the features from the target data were aligned with the source features before prediction of the labels (physical activity class) of the aligned target features. Typically the target data will represent the data from a single participant which is then aligned to the source data which is comprised of multiple participants. However it is worth noting that aligning the source data to the target data on a participant by participant basis is a viable option, although no difference in performance was identified.

6. Classification: Two classification models were used: a Random Forest [18] and a support Vector Machine (SVM). These are machine learning techniques which given data and labels construct a predictive model to assign labels to previously unseen data, in this case assigning an activity label to acceleration data. RF's have been validated on lab and free-living based data [5,11], and SVM's see common usage in activity classification [16]. The Random Forest comprised 250 decision trees, a commonly used value in this field. The SVM used a radial basis function kernel, with kernel parameters chosen by LOSOCV, as described below. All implementations were done using the Python SKlearn package [19].

Classifier performance evaluation

To evaluate the performance of the activity classifier the Leave One Subject Out Cross Validation (LOSOCV) technique was used [5,11]. It works by training the classifier on all but one participant, and then evaluating the performance on the left out participant. This procedure is repeated for all participants and then the evaluation metric averaged over all left-out participants to give a representation of how well the classification will perform on unseen data.

The evaluation metric used was accuracy of the classification.

Domain Adaption

The domain adaption method used in this paper is a straightforward modification of the Subspace Alignment (SA) algorithm [20]. SA was selected because it does not require target labels, does not drastically increase computational load and, unlike other DA methods, it is insensitive to the precise value of the single parameter that must be chosen. The underlying idea of the SA algorithm introduced in [20] is to rotate the source data so that it best aligns with the target data; a classifier is then trained on the aligned source data in order to be able make accurate predictions on the target/test data. The dimension of the features in both source and target sets is reduced in a feature reduction (FR) step prior to the alignment.

In this form SA requires the source data to be rotated to align with each new set of target data, which may represent a substantial computational burden because the classifier must be retrained for each newly-aligned set of training data. Since training the classifier is computationally expensive compared with the cost of classifying new examples with the trained classifier, in this paper the *target* data is rotated to align with the training/source data, as illustrated in Figure 1. This means that the expensive training of the classifier is done once (using unrotated source data), after which each new target set for classification is aligned with the source data (a computationally cheap step) allowing it to be classified.

Algorithm 1 summarises the main steps in classification using subspace alignment. Prior to alignment, the source and target data are each projected into a subspace defined by a smaller dimension subspace defined by their principal components. This often has the beneficial effect of discarding noise, improving classification performance [16], and identifies the principal directions in the data ($\boldsymbol{P}_s$ and $\boldsymbol{P}_t$ in Algorithm 1) that should be aligned by rotation. Alignment of the dimension reduced target features is then accomplished by multiplication by the (orthogonal) matrix $\boldsymbol{P}_t^T\boldsymbol{P}_s$ in step 6, after which the trained classifier may be used to predicted labels for the target features that have been rotated into alignment.

$\boldsymbol{X}_s$ and $\boldsymbol{X}_t$ denote the feature matrices of the source and target data respectively; each row represents an observation and each column one of the M = 39 features. $\boldsymbol{P}_s$ and $\boldsymbol{P}_t$ respectively denote the M by k (orthonormal) matrices of principal components of the source and target feature matrices, and $\boldsymbol{P}^T$ denotes the transpose of matrix $\boldsymbol{P}$.

Algorithm 1: Subspace Alignment (SA)

Input: Source features $\mathbf{X}_s$, target features $\mathbf{X}_t$, subspace dimension k

1: $\boldsymbol{P}_s \leftarrow \mathrm{PCA}(\boldsymbol{X}_s, k)$ // Generate k principal components of $\boldsymbol{X}_s$

2: $\boldsymbol{X}_s^a = \boldsymbol{X}_s\boldsymbol{P}_s$ // Reduce dimension of $\boldsymbol{X}_s$

3: Train classifier using $\boldsymbol{X}_s^a$ and corresponding labels

4: Collect target features $\boldsymbol{X}_t$

5: $\boldsymbol{P}_t \leftarrow$ PCA($\boldsymbol{X}_t$, k) // Generate k principal components of $\boldsymbol{X}_t$

6: $\boldsymbol{X}_t^a = \boldsymbol{X}_t\boldsymbol{P}_t\boldsymbol{P}_t^T\boldsymbol{P}_s$ // Reduce dimension of $\boldsymbol{X}_t$ and align with source

7: Use classifier to predict labels for aligned features $\boldsymbol{X}_t^a$

FIGURE 1

Analysis:

The comparative performance of the different approaches across different folds/subjects were tested for statistical significance using the Wilcoxon Signed Rank test, which tests the null hypothesis that two related paired samples come from the same distribution. A low p value ($p<0.05$) indicates that the results are statistically significantly different from one another with high confidence.

Results

TABLE 2:

Table 2 shows the Accuracy for all methods over both wrists. R, L indicates an approach that was trained on the right, left wrist data respectively. The Criterion approaches achieved the highest performance, as was expected because this refers creating and testing the classifiers using data collected from the same wrist and is the gold standard for classification. The NDA approaches experienced an average performance reduction of 12% compared to the Criterion approaches, whereas the DA approaches experienced an average reduction of 1%. As such it is clear that using DA allowed PA classification without significant reduction in performance regardless of which wrist the accelerometer was worn on. The AMAL approaches, which use orientation invariant features, similarly did not have a significant performance reduction, but had a slightly worse performance than the DA method.

Figure 2 shows the performance of the RDA approach (RF) for varying dimensions of the dimensional subspace (represented by k). As can be seen the performance begins low, at 0.44 for one dimension and then increases with the introduction of more dimensions until it reaches, 0.83, at 7 dimensions. The accuracy remains relatively stable for all values of k, until 34 dimensions, where the performance starts dropping. This shows that the choice of the k parameter does not have a great effect on the performance of the approaches as long as it is in the range 7-34. Hence the number of subspace dimensions was chosen to be 12.

FIGURE 2

Due to the similarity of the SVM and RF model performance, only the RF were analysed. When using DA approaches there were not significant differences between the scores compared to the Criterion approaches (F = 0.11, 0.07, p = 0.73,0.80). When comparing Criterion approaches to non-DA approaches (RNDA & LNDA in Table 2), significantly different scores were observed, highlighting that the DA is responsible for the stopping the performance reduction.

TABLE 3

Discussion:

This study set out to evaluate five approaches to achieving wrist/orientation invariance in activity classification via accelerometry and specifically examined the efficacy of domain adaptation.

The results showed that without DA, classifying data with a classifier trained on the opposing wrist leads to an average performance drop of 12% compared to using a classifier trained and evaluated on the same wrist. However, when using DA the performance drop was reduced to <1%. The performance of DA approaches was not different to Criterion approaches, whereas there was a difference in performance between Criterion and approaches not using domain adaption (NDA). Additionally, it was the case that the DA approaches outperformed the NDA approaches.

Furthermore, the DA approaches did not cause a reduction in the performance of the classifier and similarly the two amalgamated approaches (AMAL), which require accelerometers to be worn on both wrists, did not lead to performance reduction. Gjoreski et al [8] also found that the AMAL method outperformed even the Criterion approaches. The DA and AMAL approaches were equally effective at attenuating the performance reduction associated with NDA approaches, but because DA only required acceleration data from one wrist, the cost and participant burden of gathering the data, as well as the computational cost of creating the classification model, makes them preferable.

The results of this work are consistent with the work of Montoye et al. [12], who found that by making use of features that were invariant to orientation it was possible to reduce the performance drop to a negligible amount (<2%). However, their method limits the features available to the classifier. Specifically their work only made use of the ENMO feature, which is an aggregate measure of the three axis of accelerations. As mentioned in the introduction, this aggregation leads to a loss in information, which explains the low performance of that approach compared to others [5]. Additionally, making use of features that are invariant to orientation makes the implicit assumption that orientation itself is not a useful feature. Since orientation is used as a feature in some classification models [21], this assumption may be invalid.

Strengths and limitations

PCA was utilised both as a feature reduction (FR) method and as part of the subspace alignment (SA) algorithm. Use of PCA even when not using SA ensured that any decrease in the performance drop could be attributed to the DA, and not simply to the reduction in dimensions, a strength of this study. The number of dimensions in the projected subspace (k) can be an important parameter. In this study when k was low or high there were clear effects on the performance however the effect of k over a wide range between these limits was negligible. Although there are methods for automatically identifying an optimal k value (e.g., [20]), it was not necessary as the aim of this work is to evaluate the effectiveness of DA not PCA and performance was unaffected across a wide range of values.

A major strength of using SA is its simplicity. The DA parts of the classification algorithm amount to only 4 lines, it does not depend upon a particular classification model and existing classification schemes are easily augmented with it. This combined with the fact that it does not decrease the performance if alignment is not required means that there is no reason not to use this technique when there is any uncertainty about the location and orientation of an accelerometer [20]. Moreover, unlike other techniques, SA allows for all data to be used in the classification as opposed to just data from one wrist.

Furthermore, unlike other approaches, there are no restrictions on the potential features that can be used, linking well with methods of automatic feature extraction

[22] where it may be impossible to ensure that the extracted features are rotation invariant.

Some potential weaknesses of this study are as follows: Both classification models utilised in this work are ‘black-box’ approaches. Thus, it is very difficult to assess if the approach is generalizable to other populations and especially to the data collected in a free-living environment. However this weakness only applies to the classification methods utilised, not the DA method used. SA can be interpreted as a simple rotation in the feature domain, a far more interpretable method than other DA techniques. As the focus on this work was the DA not the classification, it is not felt this is a significant issue.

Only a single data set was used in the classification. It would have been preferable to test the use of DA with multiple data sets to ensure that the success is not simply due to the particular activities being symmetric about the body (and therefore having similar acceleration on either hands). This is unlikely, however, as classification performance does drop when domain adaption is not used. Nevertheless, running similar experiments on more data sets to ensure results can be replicated would be valuable. More data sets including free-living data would verify that the method is robust across a broader range of activities including more wrist dominant activities.

The modification of the SA procedure to align the target data with source data obviates the expensive retraining of the classifier for each new set of target data. This greatly decreases the computational and data storage burden compared with the original SA algorithm and enables rapid classification of new data. However, sufficient target data must be collected to characterise the principal directions before the rotation that best aligns it with the source data can be identified. This means that the algorithm cannot in its present form be used for online PA classification. Nonetheless, it is envisaged that online classification could be achieved after a delay in which data to characterise the rotation is collected; a further enhancement would be to track and update the necessary rotation through a non-stationary version of PCA.

A limitation of this study was the fact that all of the data used was gathered in a laboratory setting. It has been shown that data gathered in such a setting does not generalise well to more naturalistic free living data [11]. Data obtained in free-living conditions tends to be more variable than that gathered in a laboratory setting. For example a person may walk holding a bag or a coffee. In such cases both activities would be labelled the same but would have considerably different acceleration profiles. In a laboratory setting participants are encouraged to perform the activity in identical ways with no confounding factors, a circumstance that tends to artificially increase the classification performance.

Random forests provide the ranking of feature importance, which can be beneficial for both feature reduction and investigation into the importance of individual features. A limitation of using PCA as part of SA is that the modification to the data renders these importances more difficult to interpret. It is possible however to gain some notion of importance. Each Principal Component (PC) gives the ‘contribution’ of each original feature. By normalising this and then combining it with the feature importance found with the RF (of the PC’s) this should give some idea of the original feature importance. However this step would have to be done before the alignment, and would therefore show nothing about the DA. As the focus of this study is on the efficacy of the DA methodology any inquiry into feature importance is beyond the scope of this paper.

Some potential areas of study would be to investigate the effect of using DA on the raw acceleration data, before feature extraction. This may allow for methods of automatic feature extraction such as Convolutional Neural Networks, which have shown success in this field [22], to be used with DA.

Conclusion

Most physical activity classification models utilising wrist worn accelerometers experience performance degradation when they are applied to data extracted from accelerometers located on a different wrist or differently oriented from which the model was trained. Domain adaption, specifically Subspace Alignment, overcomes this problem as it allows for wrist/orientation invariance. The method is simple and can easily be incorporated in existing classification schemas with no loss in accuracy even if DA is not required. Further work is required to fully validate the techniques in free-living data where activities that are wrist dominant (e.g. Tennis) are more likely to be encountered.

References:

[1] Allen, N., Sudlow, C., Downey, P., Peakman, T., Danesh, J., Elliott, P., ... & Sprosen, T. (2012). UK Biobank: Current status and what it means for epidemiology. Health Policy and Technology, 1(3), 123-126.

[2] Pavey, T. G., Gomersall, S. R., Clark, B. K., & Brown, W. J. (2016). The validity of the GENEActiv wrist-worn accelerometer for measuring adult sedentary time in free living. Journal of science and medicine in sport, 19(5), 395-399.

[3] Staudenmayer, J., Pober, D., Crouter, S., Bassett, D., & Freedson, P. (2009). An artificial neural network to estimate physical activity energy expenditure and identify physical activity type from an accelerometer. Journal of Applied Physiology, 107(4), 1300-1307.

[4] Bao, L., & Intille, S. S. (2004, April). Activity recognition from user-annotated acceleration data. In International Conference on Pervasive Computing (pp. 1-17). Springer, Berlin, Heidelberg.

[5] Chowdhury, A. K., Tjondronegoro, D., Chandran, V., & Trost, S. G. (2017). Ensemble Methods for Classification of Physical Activities from Wrist Accelerometry. Medicine and science in sports and exercise, 49(9), 1965-1973.

[6] Trost, S. G., Zheng, Y., & Wong, W. K. (2014). Machine learning for activity recognition: hip versus wrist data. Physiological measurement, 35(11), 2183.

[7] Fairclough, S., Noonan, R., Rowlands, A., Van Hees, V., Knowles, Z., & Boddy, L. (2016). Wear compliance and activity in children wearing wrist and hip mounted accelerometers. Medicine & Science in Sports & Exercise, 48(2), 245-253.

[8] Gjoreski, M., Gjoreski, H., Luštrek, M., & Gams, M. (2016). How accurately can your wrist device recognize daily activities and detect falls?. Sensors, 16(6), 800.

[9] Prince, S. A., Reed, J. L., Mark, A. E., Blanchard, C. M., Grace, S. L., & Reid, R. D. (2015). A comparison of accelerometer cut-points among individuals with coronary artery disease. PloS one, 10(9), e0137759.

[10] Esliger, D. W., Rowlands, A. V., Hurst, T. L., Catt, M., Murray, P., & Eston, R. G. (2011). Validation of the GENEA Accelerometer. Medicine and science in sports and exercise, 43(6), 1085.

[11] Pavey, T. G., Gilson, N. D., Gomersall, S. R., Clark, B., & Trost, S. G. (2017). Field evaluation of a random forest activity classifier for wrist-worn accelerometer data. Journal of science and medicine in sport, 20(1), 75-80..

[12] Montoye, A. H., Pivarnik, J. M., Mudd, L. M., Biswas, S., & Pfeiffer, K. A. (2016). Wrist-independent energy expenditure prediction models from raw accelerometer data. Physiological measurement, 37(10), 1770.

[13] Csurka, G. (Ed.). (2017). Domain adaptation in computer vision applications (pp. 1-35). Springer.

[14] Cook, D., Feuz, K. D., & Krishnan, N. C. (2013). Transfer learning for activity recognition: A survey. Knowledge and information systems, 36(3), 537-556.

[15] Banos, O., Galvez, J. M., Damas, M., Pomares, H., & Rojas, I. (2014). Window size impact in human activity recognition. Sensors, 14(4), 6474-6499.

[16] Zhang, S., Rowlands, A. V., Murray, P., & Hurst, T. L. (2012). Physical Activity Classification Using the Genea Wrist-worn Accelerometer. Medicine & Science in Sports & Exercise, 44(4), 742-748.

[17] Lever, J., Krzywinski, M., & Altman, N. (2017). Points of Significance: Principal component analysis. Nature Methods, 14(7), 641-642.

[18] Breiman, L. (2001). Random forests. Machine learning, 45(1), 5-32.

[19] Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., ... & Vanderplas, J. (2011). Scikit-learn: Machine learning in Python. Journal of machine learning research, 12(Oct), 2825-2830.

[20] Fernando, B., Habrard, A., Sebban, M., & Tuytelaars, T. (2013). Unsupervised visual domain adaptation using subspace alignment. In Proceedings of the IEEE international conference on computer vision (pp. 2960-2967).

[21] Incel, O. D. (2015). Analysis of movement, orientation and rotation-based sensing for phone placement recognition. Sensors, 15(10), 25474-25506.

[22] Yang, J., Nguyen, M. N., San, P. P., Li, X., & Krishnaswamy, S. (2015, July). Deep Convolutional Neural Networks on Multichannel Time Series for Human Activity Recognition. In Ijcai (Vol. 15, pp. 3995-4001).

TABLE 1:

Approach	**Description**	**Training wrist data**	**Testing wrist data**	**Uses DA**
Criterion R	Single, same wrist non-DA	Right	Right	No
Criterion L	Single, same wrist non-DA	Left	Left	No
RDA	Single wrist DA	Left	Right	Yes
LDA	Single wrist DA	Right	Left	Yes
RNDA	Single wrist non-DA	Left	Right	No
LNDA	Single wrist non-DA	Right	Left	No
RNA	Single, same wrist DA	Right	Right	Yes
LNA	Single, same wrist DA	Left	Left	Yes
RAMAL	Both wrists non-DA	Left+Right	Right	No
LAMAL	Both wrists non-DA	Left+Right	Left	No

Table 1. Summary of classification methods using DA and alternatives.

TABLE 2:

Approach	**Description**	**Average (RF)**	**Average (SVM)**
Criterion R	Single, same wrist non-DA	0.83 (0.17)	0.80 (0.19)
Criterion L	Single, same wrist non-DA	0.82 (0.15)	0.81 (0.17)
RDA	Single wrist DA	0.81 (0.1)	0.80 (0.15)
LDA	Single wrist DA	0.84 (0.13)	0.79 (0.16)
RNDA	Single wrist non-DA	0.72 (0.11)	0.69 (0.14)
LNDA	Single wrist non-DA	0.68 (0.14)	0.69 (0.14)
NAR	Single, same wrist DA	0.81 (0.15)	0.80 (0.12)
NAL	Single, same wrist DA	0.81 (0.15)	0.80 (0.21)
LAMAL	Both wrists non-DA	0.80 (0.14)	0.79 (0.11)
RAMAL	Both wrists non-DA	0.81 (0.14)	0.78 (0.1)

Table 2: Performance results (Accuracy) of classification approaches using DA and alternatives for each participant

TABLE 3:

Comparison	*F*	*p*
ANOVA(Criterion L, LDA)	0.11	0.73
ANOVA(Criterion R to RDA)	0.07	0.80
ANOVA(Criterion L to LNDA)	7.3	0.01
ANOVA(Criterion R to RNDA)	5.0	0.03
ANOVA(Criterion LDA to LNDA)	9.5	0.04
ANOVA(Criterion RDA to RNDA)	6.2	0.02
ANOVA(Criterion L, LNA)	0.15	0.69
ANOVA(Criterion R to RNA)	0.14	0.85
ANOVA(Criterion L to LAMAL)	0.03	0.86
ANOVA(Criterion R to RAMAL)	0.12	0.72

Table 3: Results of a Wilcoxon Signed Rank test, comparing the performances if five different methods for statistical significance.

FIGURE 1

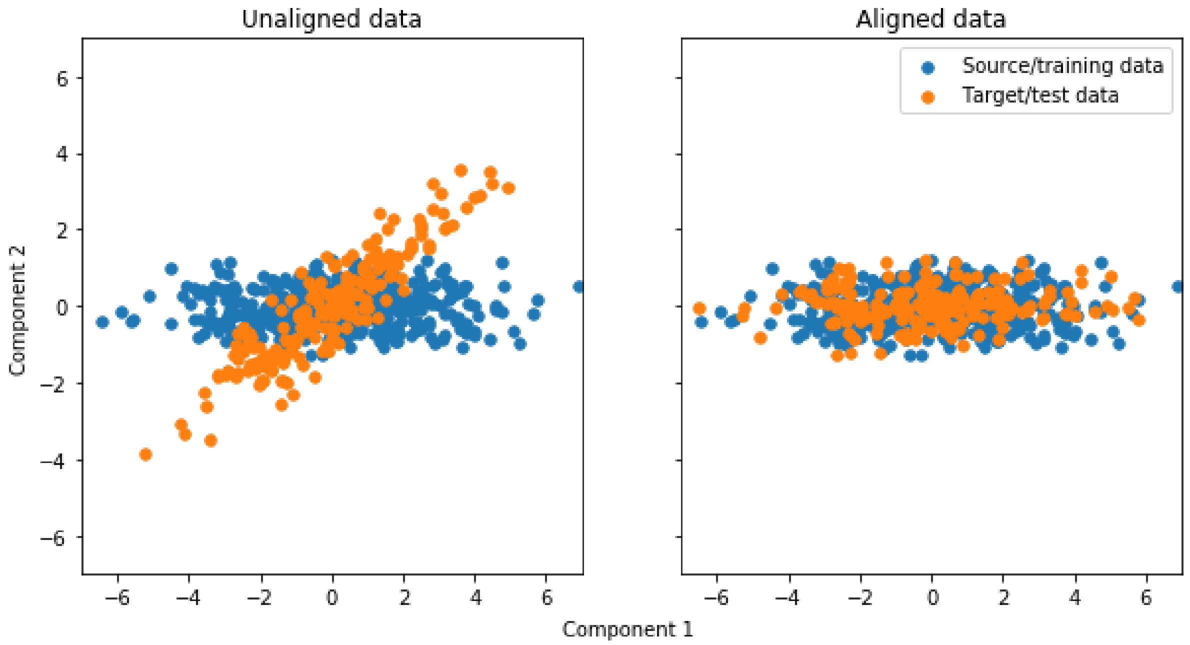

Figure 1: Example data set before and after SA. First the data is reduced to a two dimensional subspace (k = 2), in which the principal directions of the source data are aligned with the coordinate axes (left panel), then the data sets are aligned by rotating the target data (right panel).

FIGURE 2:

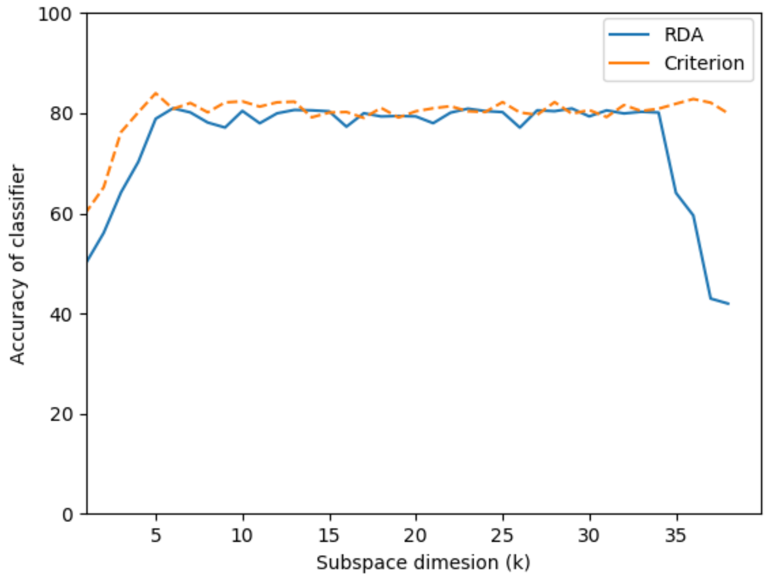

Figure 2: Performance (accuracy) of the RDA approach versus subspace dimension, *k*.