ESANN 2024 proceedings, European Symposium on Artificial Neural Networks, Computational Intelligence and Machine Learning. Bruges (Belgium) and online event, 9-11 October 2024, i6doc.com publ., ISBN 978-2-87587-090-2. Available from http://www.i6doc.com/en/.

Human Activity Recognition from Thigh and Wrist Accelerometry

Alejandro Castellanos, Antonio M. López, Diego García, Diego Álvarez, Juan C. Álvarez

Multisensor Systems and Robotics Lab (SiMuR). Electrical Engineering Department, University of Oviedo. C/Pedro Puig Adam, Ed. Torres Quevedo (Departamental Oeste), bloque 2, 2.1.09 33204, Campus de Gijón, Spain.

Abstract. The IMPaCT Cohort (ISCIII, Spain) is expected to collect biomechanical parameters from a wide population (~200,000) over seven consecutive days, using a triaxial accelerometer and a gyroscope positioned on both the wrist and thigh of participants. This will be one of the distinctive features of the Cohort, based on the hypothesis that simultaneous placement of two devices on the wrist and thigh will enable accurate classification of subjects' activity. In this study, we aim to explore this crucial aspect using Deep CNNs and data from publicly available datasets. Our experimental findings demonstrate an 85% accuracy achieved when utilizing data from both the thigh and wrist. The results support the hypothesis that incorporating accelerometry data from both limbs enhances classification, yielding over a 15% increase in accuracy compared to using data from a single limb alone.

Keywords. Convolutional Neural Networks, Human Activity Recognition, Inertial Measurement Unit, Classification.

1 Introduction

In 2020, the Carlos III Health Institute (ISCIII) of Spain established the Infrastructure for Precision Medicine Associated with Science and Technology (IMPaCT), marking a significant milestone with the launch of the Predictive Medicine Program. Inspired by the success of national cohorts established in different countries (e.g. *All of Us* in the USA, *Biobank* in the UK, *CONSTANCES* in France, *LifeGene* in Sweden, *NAKO* in Germany), IMPaCT will address the creation of a population-based cohort comprising approximately 200,000 individuals. Its aim will be to enable the integration of lifestyle data, clinical information, and genetic information for the subsequent generation of predictive models that facilitate the effective implementation of Precision Medicine.

To address population lifestyles, the IMPaCT cohort includes assessments of physical activity, sedentary behavior, and physical fitness of the participating individuals [1]. This will be achieved through a questionnaire assessing lifestyle habits and physical fitness. Additionally, over 7 consecutive days, biomechanical and physiological parameters will be recorded using two wearable monitors integrating a triaxial accelerometer and a gyroscope placed on both the wrist and the thigh of participants. This will be one of the distinctive propositions of the cohort, grounded in the hypothesis that simultaneous placement of two devices on the wrist and thigh enhances the activity classification, thus providing detailed information on specific activities.

ESANN 2024 proceedings, European Symposium on Artificial Neural Networks, Computational Intelligence and Machine Learning. Bruges (Belgium) and online event, 9-11 October 2024, i6doc.com publ., ISBN 978-2-87587-090-2. Available from http://www.i6doc.com/en/.

In this study, our aim is to preliminarily address this crucial aspect using public datasets. To achieve this, we have identified datasets including both thigh and wrist accelerometry information, covering a variety of activities, both daily and sporting. Subsequently, we applied classifiers based on a Deep Convolutional Neural Network (CNN) for classification. The network has been utilized considering accelerations and rotational velocities in the thigh and wrist as inputs. We have also evaluated the performance of networks constructed using only thigh information in one scenario and only wrist information in another.

This study aims to shed light on the feasibility of identifying physical activity from thigh and wrist accelerometry, while also analyzing the effectiveness of simultaneous device placement on the wrist and thigh, providing valuable insights for optimizing the detection of specific physical activities.

2 Methods

2.1 Datasets

In an initial phase, we proceeded to the selection of datasets containing accelerometry information from the thigh and the wrist from subjects performing a broad set of activities, typical of everyday and sporting activity. For this work we considered the datasets analyzed in [2], selecting those including wrist and thigh accelerometer/gyroscope data. The selected datasets were the Realistic Sensor Displacement Dataset [3] (RSD*), and the Daily and Sport Activities Dataset [4] (DSA). The data imported from these datasets comprise accelerations and rotational velocities in three orthogonal spatial axes aligned with the sensor, recorded on both thighs and wrists of the subjects, together with the label of the activity performed at each time instant. Information from 17 and 8 subjects was included in RSD and DSADS respectively. In total, both datasets contain information from up to 50 activities: 'No Activity', 'Walking', 'Walking (treadmill-flat)', 'Walking (treadmill-inclined)', 'Jogging', 'Running', 'Jump up', 'Jump front & back', 'Jump sideways', 'Jump leg/arms open/closed', 'Jump rope', 'Trunk twist (arms outstretched)', 'Trunk twist (elbows bent)', 'Waist bends forward', 'Waist rotation', 'Waist bends (reach foot with opposite hand)', 'Reach heels backwards', 'Lateral bend', 'Lateral bend with arm up', 'Repetitive forward stretching', 'Upper trunk and lower body opposite twist', 'Lateral elevation of arms', 'Frontal elevation of arms', 'Frontal hand claps', 'Frontal crossing of arms', 'Shoulders high-amplitude rotation', 'Shoulders low-amplitude rotation', 'Arms inner rotation', 'Knees (alternating) to the breast', 'Heels (alternating) to the backside', 'Knees bending (crouching)', 'Knees (alternating) bending forward', 'Rotation on the knees', 'Rowing', 'Elliptical bike', 'Cycling', 'Sitting', 'Standing', 'lying on back', 'lying on right side', 'Going up(stairs)', 'Going down(stairs)', 'Standing in an elevator still', 'Moving around in an elevator', 'Exercising on a stepper', 'Exercising on a cross trainer', 'Cycling (horizontal position)', 'Cycling (vertical position)', 'Jumping', 'Basketball'.

---

* RSD incorporates data collected from various sensor configurations. For our experiments, we specifically focused on configurations where a precise sensor placement was implemented, resembling the precise sensor positioning anticipated in the IMPaCT cohort.

ESANN 2024 proceedings, European Symposium on Artificial Neural Networks, Computational Intelligence and Machine Learning. Bruges (Belgium) and online event, 9-11 October 2024, i6doc.com publ., ISBN 978-2-87587-090-2. Available from http://www.i6doc.com/en/.

2.2 Data Preprocessing

The data were aligned setting the corresponding signs for the accelerations with respect to the anatomical as follows (X: anterior-posterior axis, positive forward; Y: medial-lateral axis, positive leftward; Z: vertical axis, positive upward). In the case of the gyroscope signals, the rotation in the clockwise reference axis was considered as positive in all cases. In addition, the data were resampled to 25 Hz.

The segmentation of the experimental data was carried out in windows of two seconds duration with an overlap of 1 second. In total, 51585 and 45443 data windows were obtained from the RSD and DSA datasets respectively. Given that in the time interval covered by a window, by experimental design, the subject may be performing more than one activity, the most frequent value (mode) of the outputs reported as activity in those two seconds in the dataset was assigned as the categorical output of each window of accelerometry signals.

2.3 Convolutional Neural Network

A CNN was designed as an activity classifier. Random Search [5] was used for the optimization of the hyperparameter and the architecture of the CNN. In particular, the search space was defined by: (i) the number of hidden layers of the CNN (2 -7), (ii) the optimizer used during training (ADAM, RMS Prop, Stochastic Gradient Descent), (iii) the activation function of the hidden layers (Relu, tanh, sigmoid), (iv) the size of the normalization mini-batch (10 - 30, with an step of 7), (v) the number of kernels per layer (2, 12, 30), (vi) the size of the convolutional kernels (2, 3, 15) and (vii) the learning rate (from 0.0001 to 0.1, logarithmic sampling). The initial dataset has been divided into a train subset (70 % of the total) and a test subset (30 %). In each iteration of the CNN training for hyperparameter search, a validation process has been carried out with the test data. In order to avoid overtraining of the models, the Early Stopping with patient-5 has been applied.

The Random Search algorithm was run for 150 trials using Keras Tuner implementation. Each trial was repeated 3 times with the selected hyperparameters to reduce the variance of the generated results.

After determining a correct structure for the neural network, we trained three different networks:

- A network (NET_TW) for classification based on one IMU (Inertial Measurement Unit) at the wrist and one at the thigh, processing inputs from 12 features. It was trained, validated and tested with information extracted from all possible pairs of IMUs on thigh and wrist (left thigh and left wrist, left thigh and right wrist, right thigh and left wrist, right thigh and right wrist).
- A network (NET_T) for classification based on only one of the thigh IMUs, processing inputs from 6 features. It was trained, validated and tested with the information extracted from the left and right thigh IMUs.
- A network (NET_W) for classification based exclusively on one of the wrist IMUs, processing 6 feature inputs. It was trained, validated and tested with information extracted from the left and right wrist IMUs.

The training, validation and test datasets were randomly generated, allocating 60%, 25% and 15% of the available windows, respectively.

ESANN 2024 proceedings, European Symposium on Artificial Neural Networks, Computational Intelligence and Machine Learning. Bruges (Belgium) and online event, 9-11 October 2024, i6doc.com publ., ISBN 978-2-87587-090-2. Available from http://www.i6doc.com/en/.

3 Results

After completing the search process[†], the 10 best combinations of hyperparameters were analysed. In all cases there is a similar score ($m \pm s = 0.93250 \pm 0.00066$), so those hyperparameters that simplify the architecture of the model have been selected, aiming to minimize the number of layers of the network, the number of convolutional filters and their size. Final hyperparameters were: 2 layers, RMS Prop optimization, Relu activation functions, 12 filters, size 7, Mini Batch Size of 10 and 0.0004549 for the Learning Rate. This making 7574 trainable parameters in the convolutional neural network. Accuracy values reported for the training, validation and test data are shown in Table 1.

	Train Data	Validation Data	Test Data
NET_TW	0.8734	0.8685	0.8524
NET_T	0.7189	0.7141	0.6868
NET_W	0.7319	0.7316	0.6788

Table 1: Convolutional Neural Networks Accuracy

4 Discussion

Similar levels of activity classification accuracy are observed when utilizing data from just one limb (NET_T, NET_W). This suggests that classification based on data from a single limb (thigh/wrist) yields comparable accuracy (69% and 68% using the thigh and wrist data, respectively). The experimental results support the hypothesis that including accelerometry information from both the thigh and wrist enhances classification compared to using data from a single limb alone. This enhancement is quantified to be over 15% in terms of classification accuracy for the experiments considered, leading to a high classification accuracy, with 85% accuracy achieved when using data from both the thigh and wrist (NET_TW CNN, Table 1).

It is important to note that these performance values represent a lower limit of accuracy due to certain debatable classifications. For instance, Table 2 shows an extract of the classification confusion matrix for the “Walking” target class (NET_TW, test data). As shown, not only instances of data labelled as “Walking” in the dataset are classified into this target class. Other data instances labelled in the dataset as walking flat/inclined, walking on stairs, walking on an elevator, or even basketball, are classified as “Walking” and thus reduce the classification performance in terms of classification accuracy. A similar situation arises for non-atomic activities. Such is the case for example of data instances labelled as basketball in the dataset, that are classified (see Table 3) in varied target classes such as “Walking”, “Running”, “Jumping”, etc. Basketball activity includes episodes of these atomic activities for time periods greater than two seconds, and thus these data segments can be erroneously classified into other target activities different from basketball and may reduce the accuracy of classification. 38% of the instances classified as "No activity" in the test data (NET_TW) are labelled

[†] Computation time: 1 day, 12 hours, 14 minutes and 7 seconds.

ESANN 2024 proceedings, European Symposium on Artificial Neural Networks, Computational Intelligence and Machine Learning. Bruges (Belgium) and online event, 9-11 October 2024, i6doc.com publ., ISBN 978-2-87587-090-2.
Available from http://www.i6doc.com/en/.

differently in the datasets, i.e. they correspond to periods where the user was engaged in another activity. It is unknown if this can be a classification error or perhaps there were commonly no activity episodes during the execution of other activities that could lead to these misclassifications.

	Walking
No Activity	283
Walking	1607
Walking (treadmill-flat)	429
Walking (treadmill-inclined)	49
Knees (alternating) to the breast	10
Going down (stairs)	118
Moving around in an elevator	49
Basketball	19

Table 2: Extract from the confusion matrix (NET_TW, test data) for the target class “Walking” (column). Only classes (rows) for which more than 10 instances from the RSD and DSA datasets have been classified in the target class are included.

	No Activity	Walking	Walking (treadmill-flat)	Walking (treadmill-inclined)	Running	Going up (stairs)	Going down (stairs)	Moving around in an elevator	Exercising on a stepper	Exercising on a cross trainer	Jumping	Basketball
Basketball	189	19	9	11	31	22	23	65	31	30	36	978

Table 3: Extract from the confusion matrix (NET_TW, test data) for the class “Basketball” (row). Only target classes (columns) for which more than ten instances labelled as Basketball in the RSD and DSA datasets were classified in any of them classes are included.

Additionally, it's important to consider that data instances are labeled based on the mode of activity during the two-second sampling period, potentially leading to confusion during transitions between activities and affecting classification performance.
Quantifying these errors is challenging, but it's evident that the reported classification accuracy is a lower bound of the actual classification performance. A more in-depth analysis becomes intricate, specially with 50 different activity classes, perhaps posing

the necessity of establishing a kind of "superclasses" to group equivalent activities for more accurate quantification and eventually classifier refinement. However, to avoid arbitrary definitions of such superclasses, a systematic approach like clustering analysis is needed, identified as future work.
The findings reported in this work shed some light on the expected precision difference of using CNNs to classify human activity from accelerometry from the thigh and wrist compared to only one placed on the thigh or wrist. However, the results should be taken with caution, as the experimental conditions under which the data considered for the study were collected diverge from the real-life conditions in which the subjects monitored by the IMPaCT cohort will operate.

5 Acknowlegments

This study has been funded by Instituto de Salud Carlos III through the project “PMP22/00028” (Next Generation EU funds, Mechanism for Recovery and Resilience).

References

[1] David Martínez Gómez, “Valoraciones de la actividad física y condición física en la cohorte nacional IMPaCT ,” in *VIII Simposio ETERNET: Ejercicio Físico para la salud a lo largo de la vida.*, A. Soriano Maldonado, Rodríguez Pérez M. A., B. Martínez Téllez, and E. García Artero, Eds., Almería: Editorial Universidad de Almería, Oct. 2023.

[2] F. Demrozi, G. Pravadelli, A. Bihorac, and P. Rashidi, “Human Activity Recognition Using Inertial, Physiological and Environmental Sensors: A Comprehensive Survey,” *IEEE Access*, vol. 8. Institute of Electrical and Electronics Engineers Inc., pp. 210816–210836, 2020. doi: 10.1109/ACCESS.2020.3037715.

[3] O. Baños, M. Damas, H. Pomares, I. Rojas, M. A. Tóth, and O. Amft, “A benchmark dataset to evaluate sensor displacement in activity recognition,” in *UbiComp’12 - Proceedings of the 2012 ACM Conference on Ubiquitous Computing*, Association for Computing Machinery, 2012, pp. 1026–1035. doi: 10.1145/2370216.2370437.

[4] K. Altun, B. Barshan, and O. Tunçel, “Comparative study on classifying human activities with miniature inertial and magnetic sensors,” *Pattern Recognit*, vol. 43, no. 10, pp. 3605–3620, Oct. 2010, doi: 10.1016/j.patcog.2010.04.019.

[5] J. Bergstra, J. B. Ca, and Y. B. Ca, “Random Search for Hyper-Parameter Optimization Yoshua Bengio,” 2012. [Online]. Available: http://scikit-learn.sourceforge.net.