Personalizing Activity Recognition With a Clustering Based Semi-Population Approach

FEDERICO CRUCIANI [1], **CHRIS D. NUGENT** [1], (Member, IEEE), **JAVIER MEDINA QUERO** [2], **IAN CLELAND** [1], **PAUL MCCULLAGH** [1], **KÅRE SYNNES** [3], **AND JOSEF HALLBERG** [3]

[1] School of Computing, Ulster University, Newtownabbey BT37 0QB, U.K.
[2] Department of Informatics, University of Jaen, Campus Las Lagunillas, 23071 Jaen, Spain
[3] Department of Computer Science, Electrical and Space Engineering, Luleå University of Technology, 971 87 Luleå, Sweden

Corresponding author: Federico Cruciani (f.cruciani@ulster.ac.uk)

ABSTRACT Smartphone-based approaches for Human Activity Recognition have become prevalent in recent years. Despite the amount of research undertaken in the field, issues such as cross-subject variability are still posing an obstacle to the deployment of solutions in large scale, free-living settings. Personalized methods (*i.e.* aiming to adapt a generic classifier to a specific target user) attempt to solve this problem. The lack of labeled data for training purposes, however, represents a major barrier. This is especially the case when taking into consideration that personalization generally requires labeled data to be user-specific. This paper presents a novel personalization method combining a semi-population based approach with user adaptation. Personalization is achieved through the following. Firstly, the proposed method identifies a subset of users from the available population as best candidates for initializing the classifier to the target user. Subsequently, a semi-population Neural Network classifier is trained using data from this subset of users. The classifier's network weights are then updated using a small amount of labeled data from the target user subsequently implementing personalization. This approach was validated on a large publicly available dataset collected in a free-living scenario. The personalized approach using the proposed method has shown to improve the overall F-score to 74.4% compared to 70.9% when using a generic non-personalized approach. Results obtained, with statistical significance being confirmed on a set of 57 users, indicate that model initialization using the semi-population approach can reduce the amount of labeled data required for personalization. As such, the proposed method for model initialization could facilitate the real-world deployment of systems implementing personalization by reducing the amount of data needed for personalization.

I. INTRODUCTION

Human Activity Recognition (HAR) finds several potential applications in pervasive systems, varying from Ambient Assisted Living (AAL) to generic home automation scenarios. Recently, smartphone-based solutions for HAR have become prevalent [1]–[5]. Smartphones are perceived as unobtrusive from the user perspective, facilitating activity monitoring in free-living contexts, avoiding the use of more invasive devices [6]. Moreover, they provide a convenient source, both generating and logging data since most users will keep the phone with them almost all day. Despite recent

progress in HAR, there are still important challenges hindering the deployment of solutions in a free-living context [7]–[9]. Among these, inter-person (or cross-subject) variability represents a major hurdle for current solutions [7]. This issue is further compounded by the fact that experiments are primarily conducted using data collected, on a relatively small scale, and in controlled environments [1], [10]. To date, accuracy of smartphone-based HAR solutions, has been reported as ranging between 85% and 95% for detection of simple activities (such as sitting, walking, running and cycling), whereas lower accuracies have been reported for more complex activities (e.g. Activities of Daily Living (ADL)) [2], [10]. Moving solutions to a free-living context has shown to result in accuracy rates decreasing by up to 17% [11].

The associate editor coordinating the review of this manuscript and approving it for publication was Macarena Espinilla [ID].

Consequently, researchers have been highlighting the importance of validating solutions with larger datasets collected in free living scenarios, rather than controlled environments [6], [10]. Furthermore, the conventional approach, has so far been to adopt the 'one-size-fits-all' concept: also known as *population* based approach [12]. A generic classifier is trained offline on the data from the available population, and then used to perform real-time detection on new, unseen subjects. Accuracy for unseen subjects (*i.e.* not used for training), however, has been observed to decrease to between 63% and 86% [10]. Consequently, in recent years, more effort has been targeted towards methods attempting to personalize classifiers (i.e. adapting a generic classifier to a specific user). Personalized solutions (*i.e.* trained on data belonging to the specific user) have shown to improve performance for new subjects [7]. Similarly, semi-population based approaches have been proposed [12], aiming at personalizing models by identifying a subset of users having characteristics that are similar to the target user. Their deployment in a free-living context, however, inherently requires methods to implement online training. Approaching the problem using supervised methods requires availability of a large amount of high-quality labeled data, that must also be user specific. Availability of such datasets therefore represents a significant obstacle. In particular, the annotation phase is a burden that cannot be easily automated. While recent automated Machine Learning (ML) approaches have been proposed (automating part of training process in presence of labeled data) [1], [13], human supervision/intervention is still required to generate reliable labeled datasets [8], [14]–[16]. In the case of smartphone-based solutions, crowd labeling approaches based on experience sampling have been proposed; where the target user directly annotates data points by means of prompts delivered through a mobile app [6], [17], [18]. This approach, however, introduces the problem of reliability of such datasets in terms of label noise [14]. Uncontrolled labeling can introduce label noise (for instance users misinterpreting the meaning of labels or simply as a consequence of distraction) [14], [15]. In this scenario, our approach to personalization is based on a semi-population method identifying similar users in the population dataset. Similarity is measured on a small amount of labeled data from the target user. The same labeled data are then used to complete the model training, performing further adaptation of the semi-population model. This work presents the following contributions: (i) a semi-population based method is proposed to identify a subset of users (from the available training population) as good candidates to initialize the parameters of a personalized model, (ii) an online training mechanism is then proposed to further adapt the model to the target user, and (iii) the results are evaluated using Vaizman's publicly available dataset [6] which provides data collected in free-living and unconstrained conditions using a smartphone.

The remainder of the paper is structured as follows. Section II describes related works and current limitations. Section III describes the proposed approach. Section IV and V describes the experiment and the evaluation methodology. Sections VI and VII report results and discussion, respectively. Finally, Conclusions are drawn in Section VIII.

II. BACKGROUND

Since the early attempts of sensor-based HAR, the most common approach has been the 'one-size-fits-all' or 'population' based approach; i.e., a method aiming at building a generic model trained on the available population data [12]. The need of personalization has long been debated in the research community, with supporters of generic approaches believing that models should be able to generalize, and therefore also to make good predictions on new users [19]. Nonetheless, several studies have highlighted how a population based approach can lower the accuracy rate by a substantial amount, when dealing with unknown users (i.e. not used for training) [9], [12]. Consequently, several studies explored possible solutions to adapt a generic classifier to a new user (a process also referred to as *calibration* [12]). Three main macro categories addressing personalization can be distinguished. The first consists in training a fully personalized model. This method, however, is limited in terms of scalability and therefore is not applicable to a real-world scenario, since it relies on the presence of a substantial amount of labeled data belonging to each target user [12], [16]. An alternative approach is the calibration (or adaptation) of a generic classifier, performed by updating the parameters of a pre-trained model, using a portion of data belonging to the new target user [20], [21]. This approach reduces, at least partially, the amount of labeled data required from the target user for adaptation. Alternatively, *semi-population* approaches attempt to build a model using the available pre-existing population, however, restricting it to users that present similar characteristics to the target users, as presented in [12]. This Section initially provides an overview of relevant work on personalization, which is followed by a discussion on main common limitations of such an approach.

A. OVERVIEW OF RELATED WORK

Among early personalization attempts, in [20] Zhao *et al.* proposed a transfer learning approach to adapt a Decision Tree (DT) classifier by updating its parameters. In their approach, a DT trained on subject A is adapted to subject B using unlabeled data belonging to this new user. The update of parameters is realized through an iterative algorithm using k-means clustering [22]. The original DT of subject A is used to make predictions. K-means clustering is used to identify the centroids for each class on the new data. Finally, selected samples corresponding to high confidence predictions are used to update the parameters of the DT. High confidence samples are selected as the k-nearest samples to the centroid. The process is repeated until convergence is obtained. The approach targeted simple activities ('stationary', 'walking', 'downstairs', 'upstairs' and 'running') using 12 features extracted from accelerometry data. Evaluation was performed on an ad-hoc collected dataset consisting

of 10 users. In [23], classification was performed through a weighted majority voting process, combining predictions produced by a set of expert classifiers. The approach, in this case, requires presence of labeled data belonging to the target user in order to identify the optimal weights that will be used to perform the calibration. On average, results obtained were 89.4% F-score for simple activities and 74.4% for an extended set including a number of complex activities. The approach outperformed the accuracy obtained when using the population based approach. For the worst subject, however, results scored 76.9% and 56.1% F-score, respectively for simple and complex activities sets, indicating that the set of expert classifiers may not be representative for all users.

In [21], a transfer learning algorithm was proposed for adaptation of models based on Reduced Kernel Extreme Learning Machine (RKELM). Similar to [20], the generic model is used to make predictions on new user's unlabeled data. Data points corresponding to predictions with higher confidence are then used to update the model. The approach was evaluated on the UCI-HAR [24], a dataset consisting of 30 users targeting simple activities. In [12], a semi-population approach was proposed. This approach relies on the existence of a set of personalized models, and aims at reusing those models rather than adapting them. Models belonging to users that are similar to the new one, will be re-used to make predictions on the new user. Similarity between users is based on the fitness score of the model itself, measured on accuracy of predictions for new labeled data belonging to the final user. In [7], an online active learning solution based on Random Forests (RF) was proposed for adaptation. The solution allowed to update an RF model based upon availability of new labeled data-points. In this case, when new data points were acquired, the system asks for user-feedback to obtain new labels for the ones classified with lower confidence. User-labeled data points are then used to update the model. The approach has been evaluated over a dataset with 15 participants in realistic contexts, however, controlled conditions (e.g. users were asked to stand still for a few seconds between activities). In [25], k-means algorithm was used to label new data points, then samples with high confidence were used to train a classifier based on Multivariate Gaussian Distribution (MGD). Compared to other studies, the set of targeted activities was rather limited in this case, aiming only at distinguishing between light, moderate and vigorous intensity activity, plus fall detection. Results obtained scored 97.9% F-score for the personalized approach vs 95.4% of the generic approach, although evaluation was performed on a rather limited dataset (10 users) and with the sensor's location constrained to the user's waist. In [26], incremental learning was used to perform adaptation of a generic model when new labeled data became available. User adapted models lead to a 4.6% increase in accuracy compared to the conventional generic approach. Results have been evaluated over a dataset limited to 10 subjects. In [27], similar to [21], the authors used an online learning approach based on RKELM. Adaptation of the models was performed using

TABLE 1. Overview of studies on personalisation: Adaptation method and dataset used for evaluation.

Ref.	Method Approach[+]	Dataset Num. of Subjects	Dataset Data Acquired[*]
Zhao et al. 2011 [20]	A	10	C
Reiss et al. 2013 [23]	SP	9	C
Deng et al. 2014 [21]	A	30	C
Hong et al. 2016 [12]	SP	28	U
Sztyler et al. 2017 [7]	A	15	S
Zhao et al. 2018 [25]	P	10	C
Siirtola et al. 2018 [26]	A	9	C
Hu et al. 2018 [27]	A	8	C
Mannini et al. 2018 [28]	A	33+22	C
Our experiment	SP/A	57	U

[*] Data Acquisition: (C) Controlled, (U) Uncontrolled, (S) Simulated Naturalistic.
[+] Approach: (P) Personalized, (A) Adapted or (SP) Semi-Population.

newly available labeled data belonging to the target user. Data points corresponding to misclassified predictions are used to update the model. Evaluation was performed on a dataset of 8 subjects, including 19 daily and sport activities with 5 minutes of accelerometer and gyroscope data for each activity. In [28], an online personalization strategy was proposed based on a Support Vector Machines (SVM) model. Similar to [7], predictions with high uncertainty are used to solicit users requesting them to provide new labels. The approach was evaluated using two datasets with characteristics: 33 adults in the first, and 22 youths in the second.

Table 1 provides an overview of these studies, and the nature of the datasets used for evaluation. More recently, in [29], the authors tested a personalization approach, comparing its results with a subject independent model. The experiment also analysed results of an adapted version of the model obtained considering physical and sensor similarities with other users. The approach was tested on three datasets, however, mostly on data collected under controlled conditions.

B. COMMON LIMITATIONS
No matter what strategy is used for adaptation, a common method to evaluate performance of personalized approaches is to compare obtained results with the Leave-One-Subject-Out (LOSO) validation [12]. It must be observed, however, that in many cases the sample size of datasets used for evaluation is rather limited (10 subjects or less) [7], [20], [23], [26], and/or data have been collected under controlled conditions [7], [20], [21], [23], [25]–[27].

In some cases, user interaction is not required, since adaptation is performed using samples whose prediction correspond to high confidence [20], [21], [25], this approach however, assumes the initial model is able to produce reasonably good predictions for all target classes. This problem was addressed in [7] by asking user feedback for low confidence samples. On the other hand, this approach assumes that high confidence predictions are mostly correct. In [12], [23], the semi-population approach presents the advantage of not requiring a second (personalized) training phase,
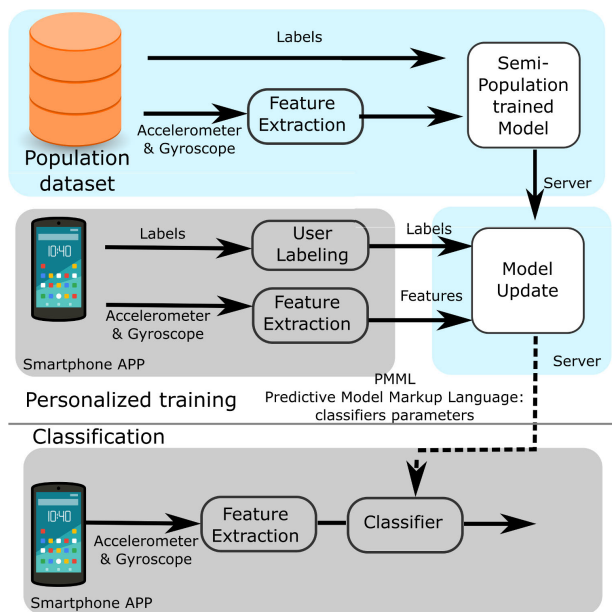
FIGURE 1. Online personalization framework: A subset of the population is identified to generate the training data that are used to initialize the classifier model, the model is then updated using user specific data, finally classification is performed online and locally on the smartphone.

since available models are reused. This approach, however, assumes that available population is universally representative for any user. This limitation can be observed for instance in [23], where for the worst subject, obtained results are significantly lower, indicating that some new users may not be well represented by the available population only. Adaptation using final user labeled data addresses this limitation, introducing, however, at the same time, the problem of reliability of newly obtained labels. Comparison of performances between studies is not an easy task, since dataset nature, sample size and target activities vary significantly. For instance, in many studies the cycling class was not targeted, although it represents a challenging example that has often been reported as conflicting class, e.g. with the walking class [10].

In this work, a novel approach is presented, aiming at combining advantages of the semi-population approach, with the adaptation to the target user. Initialization of the model based on the semi-population approach aims at obtaining a faster convergence on the adaptation phase, thus reducing the amount of labelled data required for adaptation. Results are evaluated over a publicly available dataset of 57 users [6] collected in free-living and uncontrolled conditions.

III. IMPLEMENTATION
Personalization of classifiers requires a framework supporting online training. As in [16], we propose an architecture where the training/adaptation phase is performed remotely on a server. Final classification is performed in real-time and locally on the smartphone, using features extracted from the on-board accelerometer and gyroscope sensors.

Fig. 1 depicts the initial training phase (where a semi-population model is trained using a subset of users from the available population); the personalization phase (in which model parameters are updated using data belonging to the target user); and finally real-time classification of activities on the smartphone using the personalized model. The semi-population model initialization aims at reducing the amount of data required for the latter personalization phase. Personalization aims over time to reduce classification errors due to cross-subject variability, for which the semi-population approach is not able to generalize to properly. The process relies on collecting new labeled data points. Labeling can be performed directly by the target user by means of prompts. Raw-data fragments from the smartphone accelerometer and gyroscope are sent to the server, together with user-provided labels. These new data points belonging to the final user are firstly used to identify similar users from the available population dataset, allowing training the semi-population model. The same data points are then used to update the classifier's parameters (e.g. updating network weights of the Neural Network (NN)) finalizing the model's personalization. A generic model can be used at the outset, when no target user data are available. Parameters of the personalized classifier can then be passed onto a user's mobile app, deploying an adapted and personalized classifier for each user.

Personalization of the model is realized by performing online learning on the server side. Labeled data points (belonging to the specific target user) are then used to update the network's weights. Summarizing, the proposed online architecture is implemented as two steps procedure. First, the semi-population approach is used to initialize the the proposed classifier. The second step consists in further adapting the model, updating its weights using data from the target user. The following Section provides a detailed description of the proposed semi-population based method.

A. SEMI-POPULATION BASED APPROACH
This method aims at identifying a subset of users, out of the available population, exhibiting characteristics similar to the target user. This subset is identified based on similarity between users, which is measured on the feature space, using a clustering procedure.

As presented in Fig. 2, for each user and for each activity, data points labeled with the same class are clustered. Mean shift clustering [30] was used to identify clusters within the set of data points of each activity. Mean shift was chosen considering the algorithm's flexibility, being a non-parametric cluster selecting procedure. It does not require the number of expected clusters as input (compared to other algorithms such as *k*-means), and the required bandwidth parameter can be estimated automatically. Clusters obtained are ordered based on the number of samples they contain. The majority cluster (*i.e.* containing the largest number of points) is used as a reference, and its centroid defines a user's *activity vector*. This is used to measure similarity between users. In this approach, a user is represented by a set of vectors (one for
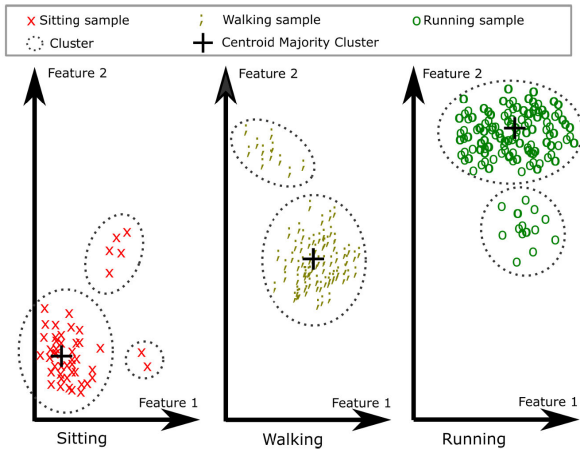
FIGURE 2. Data points are partitioned according to their label. Mean shift is used to cluster data points belonging to the same activity (in the feature space domain). Finally, the centroid of the majority cluster is used as reference to measure similarity between users.

FIGURE 3. Feature's centroids are used to map users into a similarity domain built on the features space. When data from the target user becomes available, the clustering procedure allows the vector locating the user in the feature space to be calculated. Finally, the *n* closest users are used to initialize training weights of the generic model.

each activity) where each vector corresponds to the main cluster centroid. Let $A = [a_1 \cdots a_M]$ be the set of target activities, and N the number of features; all users are represented by M *activity vectors* $uv_{ik} \in \Re^N$, where i identifies i-th user and $k \in [1, 2. \cdots M]$ the target activity class. Best candidates to define the training set are therefore identified as the ones minimizing the Euclidean distance computed on activity vectors $dist(uv_{ik}, uv_{jk})$ i.e. the distance between users i and j, on activity k. The procedure (refer to Algorithm 1) includes the following three steps:

1) For each activity label the set data points with same label are clustered.
2) The centroid of the majority cluster is detected.
3) Finally, the centroids identified are stored as the set of user's activity vectors.

Offline, all activity vectors for the available training population are computed and stored server side.
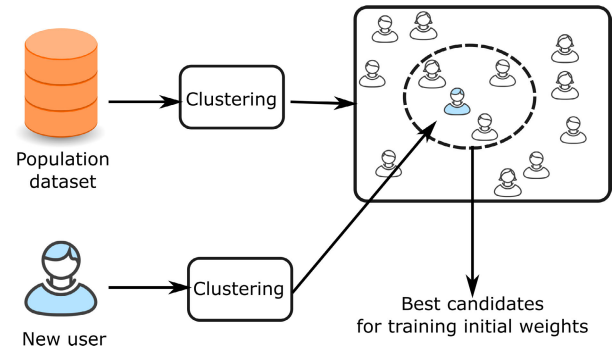

Algorithm 1 Compute User Vector

1: **Input**: target activity labels: $A = [a_1, a_2 \cdots a_M]$
2:         user data: $user\_data_i$
3: **Output**: user vectors: uv_i
4: **procedure** COMPUTEUSERVECTOR(A, USER_DATA)
5:     **for all** $a_k \in A$ **do**
6:        $data = \{\forall p \in user\_data | label(p) = a_k\}$
7:        $clusters = [cl_1 \cdots cl_s] = meanShift(data)$
8:        $max\_cluster = \max_{i=1}^{s}(|cl_i|)$
9:        $uv_{ik} = centroid(max\_cluster)$
10:     **return** $uv_i = [uv_{i1}, uv_{i2}, \cdots uv_{iM}]$

As presented in Fig. 3, when data from a new user have been collected, user provided labels are used to make the partition of data points into the set of target activities. The clustering process is repeated to calculate the activity vectors describing the new target user. Then the n users (for each

activity class) with the closest activity vector are considered as the candidate population to train the initial classifier. Considering that for each activity n users will be identified, in total the training dataset will include data from a number n' of distinct users, with $n \leq n' \leq M \times n$. The rationale of this approach, is that initializing weights on a training dataset (composed of data belonging to users having similar characteristics) can potentially help to converge more rapidly towards an improved personalized solution. The classifier trained on this subset of the population is then adapted for personalization using available data points belonging to the target user.

In principle, the approach can be applied to any supervised learning scenario. Nonetheless, it is important to consider the number of dimensions of the examined feature space. When using Euclidean distance in high dimensional spaces, the relative minimum and maximum distances will tend to converge with the increase of space dimensions [31], [32]. This effect, however, also depends on the nature of the distribution and the number of available samples. Normally distributed datasets for instance are more subject to this phenomenon [32]. On the other hand, the availability of a large number of samples alleviates the effect [32]. Considering our dataset and target activities such as cycling, we want the method to be able to identify macro-cross-subject variabilities, for instance users cycling with their phone in the pockets, as opposed to subjects carrying it on their bag.

Another factor to consider is the number n, defining how many of the nearest subjects will be considered. High values of n in high dimensional spaces will tend to include subjects with much more diverse characteristics. Techniques to reduce the number of dimensions such as Principal Component Analysis (PCA) can also be considered, allowing to compute user vectors in a reduced dimension space.

IV. EXPERIMENT

This Section presents the publicly available dataset used to evaluate the proposed method and the classification approach used in the experiment.

A. EVALUATION DATASET

Evaluation of the proposed approach was conducted using the Extrasensory dataset [6]. The dataset contains data from 60 users (34 females and 26 males) and was collected in free-living using the smartphone, with up to 28 days and an average of 7 days of data collected per user. The dataset includes activity labels provided through user prompts, and includes raw data measurement from the smartphone inertial sensors (accelerometer and gyroscope), watch accelerometer, GPS and microphone. Raw-data from the smartphone inertial sensors were sampled at 40 Hz and consist of a total of 308,306 labeled fragments of recording of 1 minute (with 20 seconds of raw data collected in each fragment) [6] with more than 200,000 labeled fragments for physical activities. Annotations includes in some cases also the phone location (e.g. 'in pocket', 'on table' or 'in bag'). Participants used their own smartphones for data collection, consequently the dataset also covers various smartphone models and platforms with 34 iPhone users and 26 Android users. For the experiment, data from the smartphone accelerometer and gyroscope were used. The set of users was therefore restricted to 57 out of the 60 participants as some of the user's gyroscope data were not available. The dataset presents typical characteristics of free-living scenarios, i.e. smartphone location is not constrained to a specific consistent location, data samples are not collected following a script and therefore are usually highly imbalanced. The majority classes (lying and sitting) provide on average up to 80% of the entire set of data points, followed by the walking class (on average 10-15%), with the cycling or running class normally being the minority class. Class imbalance makes the problem more challenging, an aspect that potentially affects both the training and the evaluation process (implemented strategies to deal with class imbalance will be described in Section V).

As data are collected under uncontrolled labeling, the dataset may potentially include some label noise. Results in [6] show better performance on the running and walking class when using the smartphone and the watch accelerometer (possibly due to situations in which the user labeled a fragment as walking/running activity, however, was not carrying the smartphone). For this reason, from the original datasets some data points were excluded because of the inconsistency between the user selected label and observed signal. In particular, some fragments labeled as running or walking corresponding to a flat signal with no variations measured on the accelerometer signals were excluded, since they potentially correspond to situations where the participants were not carrying the device with them. Analysis of the data also exhibits a significant cross-subject variability. Step cadence for each fragment was calculated using a simple step detector as in [16]. Typical walking pattern can be expected to be in the order of 90-110 steps per minute (spm), whereas values around 160-180 spm are expected for a running pace [16]. Fig. 4 illustrates an example of two users with a statistically significant difference in the step cadence, particularly for the running activity. The figure illustrates a gaussian distribution
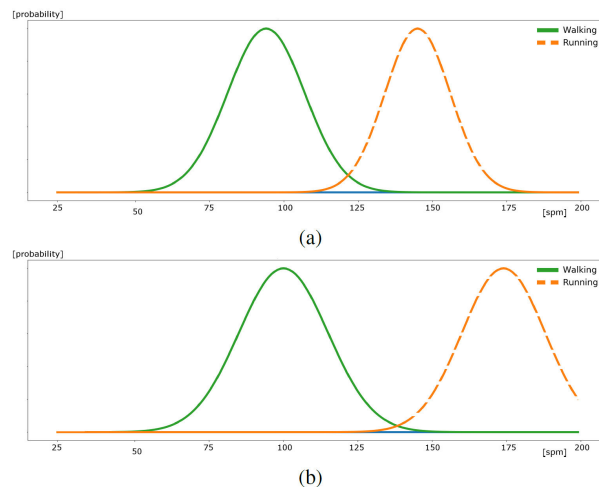
FIGURE 4. An example of cross-subject variability observable on the dataset: a comparison of two users (a) and (b) exhibiting different walking and running cadence.

obtained using the average cadence and standard deviation observed for two different users.

B. CLASSIFICATION APPROACH

Similar to our previous study [16], a NN based on a Multi-Layer Perceptron (MLP) was used, taking as input features data extracted from the accelerometer and gyroscope. NNs offer good support for online training, which is required for personalization purposes. Adaptation of the model can be performed by means of updating network weights on newly available user specific data points. A set of time and frequency domain features (as presented in Table 2) were extracted using a windowing approach with a window size of 3 seconds and a 25% overlap. The set includes features commonly used for this type of classification problem using inertial signals; similarly, a window size between 1-4 seconds is commonly used for segmentation [2], [33]. Time domain features included statistical moments of the magnitude of the acceleration signal (mean, standard deviation, kurtosis, skewness), min and max values, range (max-min) and signal energy. Frequency domain features (also extracted from the 3D magnitude of acceleration) included the number of peaks in the Power Spectral Density (PSD) and location of the highest peak in the 0-20 Hz frequency band. The set of features also included features at single channel level (mean, standard deviation and range of the x, y and z axes). Finally, additional features have been obtained measuring cross-correlation between channels measured using the Pearson coefficient between the pair of axes XY, YZ and XZ. The selected set of features is an extension of the feature set used in our previous experiment [16], with the addition of the cross-correlation features.

The resulting feature set consists of 44 features; 22 for the accelerometer and 22 for the gyroscope signal.

TABLE 2. The set of time and frequency domain features set used in experiment.

Domain	Signal	Table Column Head
Time	3D magnitude	mean, variance, min, max, range, skewness, kurtosis, energy
Time	X,Y,Z axes	mean, variance, range, cross-corr(X,Y) cross-corr(Y,Z),cross-corr(X,Z)[a]
Frequency	3D magnitude	number of peaks in PSD, location of highest peak

[a] cross-correlation is measured using Pearson's coefficient.

An initial partial optimization of hyper-parameters was performed using the 5-fold partition provided with the dataset (48 participants for training, and 12 for validation). The goal was to identify a valid candidate topology and to estimate the average number of epochs before the model's performance starts deteriorating due to overfitting phenomena. The identified NN model consisted of 4 dense hidden layers using Rectified Linear Unit (ReLU) as activation function ($128 \times 256 \times 128 \times 64$), with a dropout layer as the input layer (dropout set to 0.2), using Adam optimizer [34], and softmax as activation for the output layer.

Class imbalance was addressed using Balanced Batch Learning (BBL), ensuring the model is trained on the same number samples for each target activities. At each step, the training samples are extracted taking one random fragment for each target class from the available training data. Training data can be obtained as a random selection of n users from the available population, or using the n closest semi-population approach. The number of steps per epoch was identified based on average data availability (considering the cardinality of minority classes), since days of recording per user vary from 3-4 to 28 days. The approach allows to train using all available training samples from minority classes, preserving availability of larger number samples randomly selected form the majority classes. The model was implemented using Keras [35] with a TensorFlow back-end [36]. For the experiment the default learning rate (0.001) was used.

V. EVALUATION METHODOLOGY
Evaluation of the method was performed using LOSO procedure in two steps. For each user, the first step consisted in training a semi-population model and its comparison with the average of 4 random selections of users. In the second step, some data belonging to the target user are used to update the weights of the classifiers, both for the case of an n-random initialized model, and the proposed n-closest.

Since the dataset is highly imbalanced, common evaluation metrics such as simple accuracy (ratio between correct/wrong predictions) can be misleading. A common metric for imbalanced datasets of this nature is balanced accuracy [6], [37]. Similarly, macro averages (average of precision, recall and F-score for each individual class) can be used to deal with imbalanced datasets, for instance macro-average recall is used as balanced accuracy in [38]. Macro-averages were used to evaluate the approach in the experiment.

The set of target activities was defined as lying, sitting, walking, running and cycling. This set allows to compare

results with a generic approach using the same input sensors as in [6].

A. EVALUATION SEMI-POPULATION APPROACH
The first step aimed at evaluating the semi-population approach described in Section III-A. The validation routine iterates on the set of 57 users comparing the random selections of users (average of 4 random selections of n users out of the remaining 56), with the semi-population based approach (i.e. using data belonging to the n closest users as described in Section III-A) to train a generic and a semi-population model. Regarding the choice of n, it must be considered that low values of n could result in the risk of generating small training datasets. On the other hand, large values of n increase probability of intersection between the two sets (n random and n closest), thereby increasing the risk of making the comparison inconclusive. Consequently, for the experiment the values of $n = 5$ and $n = 10$ have been used.

In the Extrasensory dataset labels are provided per fragment. At verification stage each fragment is segmented, and features are extracted producing multiple samples. The average prediction on the overall fragment is taken as the prediction of the fragment and is compared with the original label.

As aforementioned BBL was used for training. An ad-hoc batch generator was implemented. Batch generators allow to load training data on-the-fly, reducing the amount of memory required during training, and therefore allowing to train multiple models in parallel. Feature extraction during training, was performed starting from a random index (between 0 and 20) instead of a fixed start from position 0 inside the data fragment. The random entry point was introduced since at each epoch, fragments corresponding to minority classes will most likely have already been examined, as opposed to the case of majority classes for which a much larger number of samples is available. The random entry point was introduced as an additional measure to deal with the class imbalance, allowing to reproduce small variations of samples from minority classes.

B. EVALUATION OF ADAPTATION
The second step of evaluation aimed at measuring performance of adapted models, obtained by updating weights of the network using some data of the target user. For the adaptation step, samples of each class have been sorted based on their timestamp. An increasing number of training samples were provided taking the first 10, 20 and 30 fragments of data for each class, and using the remaining fragments as test data. This approach ensures there is no overlap between training and test data during the adaptation, and also allows to simulate a realistic scenario in which the model would evolve in a real-world experiment over time.

VI. RESULTS
The experiment and produced results are divided in two parts. The first part focused on evaluating the proposed semi-population approach. The second focused on evaluating the performance of adaptation: comparing adaptation of

TABLE 3. Precision, Recall and F1 score obtained with a generic model trained on *n* random users (a), on *n* closest users (b), and the personalized version of the classifiers initialized using *n* random users (c), and using *n* closest (d). Values are reported for each target class and as macro-average.

Activity	Precision		Recall		F-Score	
	n=5	n=10	n=5	n=10	n=5	n=10
Lying	0.6185	0.6375	0.5293	0.5749	0.5704	0.6046
Sitting	0.5280	0.5492	0.5819	0.5773	0.5536	0.5629
Walking	0.7651	0.7921	0.7549	0.7872	0.7599	0.7896
Running	0.1435	0.1905	0.6003	0.6511	0.2316	0.2947
Cycling	0.4083	0.4392	0.6176	0.6626	0.5214	0.5283
Average	**0.4927**	**0.5217**	**0.6168**	**0.6506**	**0.5214**	**0.5560**
(a) generic model trained using *n* random users						
Lying	0.6711	0.6802	0.6954	0.6946	0.6830	0.6874
Sitting	0.6298	0.6316	0.5718	0.5962	0.5994	0.6134
Walking	0.8221	0.8320	0.8639	0.8449	0.8425	0.8384
Running	0.2718	0.1993	0.6845	0.6524	0.3891	0.3054
Cycling	0.5164	0.5852	0.6900	0.6905	0.5907	0.6335
Average	**0.5822**	**0.5857**	**0.7011**	**0.6957**	**0.6209**	**0.6156**
(b) generic model trained using *n* closest users						
Lying	0.6215	0.6343	0.6998	0.7332	0.6583	0.6802
Sitting	0.5833	0.6157	0.4927	0.4967	0.5342	0.5499
Walking	0.7452	0.7443	0.8955	0.9008	0.8135	0.8151
Running	0.6840	0.7117	0.8249	0.8516	0.7479	0.7754
Cycling	0.8041	0.8081	0.6369	0.6535	0.7108	0.7226
Average	**0.6876**	**0.7028**	**0.7100**	**0.7272**	**0.6929**	**0.7086**
(c) personalized with initialization on *n* random users						
Lying	0.6368	0.6713	0.7765	0.7686	0.6998	0.7167
Sitting	0.6491	0.6710	0.4693	0.5486	0.5448	0.6037
Walking	0.7561	0.7712	0.8973	0.9000	0.8207	0.8306
Running	0.7294	0.8154	0.8503	0.8503	0.7852	0.8325
Cycling	0.7539	0.7823	0.7042	0.6994	0.7282	0.7386
Average	**0.7051**	**0.7422**	**0.7395**	**0.7534**	**0.7157**	**0.7444**
(d) personalized with initialization on *n* closest users						

FIGURE 5. Macro-average F-score values obtained with an increasing number of data fragments for adaptation.

FIGURE 6. Confusion matrices for the target activities using: (a) 5-random, (b) 5-closest, (c) 5-random adapted (d) 5-closest adapted.
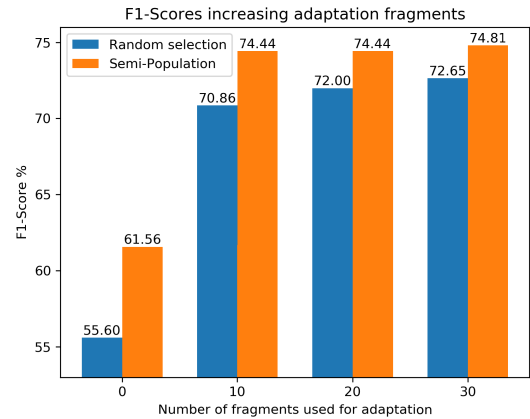
models initialized using the semi-population approach, vs models initialized on a random subset of users.

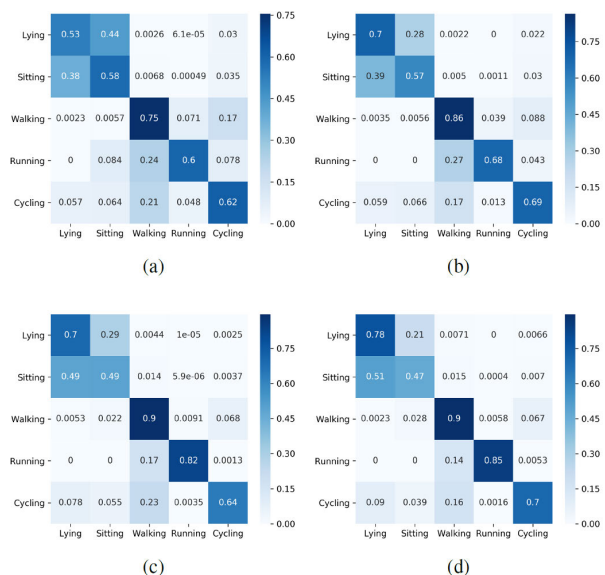## A. EVALUATION OF SEMI-POPULATION APPROACH

This part of the experiment compared different initialization strategies of the classifier, comparing a random selection of users from the population, with the proposed semi-population approach. The random selection of *n* users was repeated 4 times and the average result was compared with the selection of the *n* closest users. Finally, the personalization phase was repeated on the models obtained with these two approaches. Results obtained based on the clustering approach to initialize the weights of a semi-population based classifier are reported in Table 3. Specifically, Table 3-a reports macro-average values for precision, recall and F-score; obtained initializing the weights of the generic classifier using a random selection of users. Table 3-b shows obtained values training the semi-population model obtained using the *n* closest users criterion in the feature space. Finally, Table 3-c and 3-d report metrics obtained performing adaptation starting from random selection of users, and the *n* closest approaches respectively.

B. EVALUATION OF ADAPTATION

Fig. 5 summarizes macro-average F-score values obtained with an increasing number of fragments for adaptation; both for the semi-population and the random selection approach.

The average normalized confusion matrices were calculated on the set of users with the four approaches. Fig. 6 depicts normalized confusion matrices obtained with (a) 5-random users, (b) 5-closest users, (c) 5-random adapted, and (d) 5-closest adapted.

Finally, statistical analysis was performed comparing balanced accuracies obtained with semi-population (*n*-closest) and random selection (*n*-random). Fig. 7 illustrates the boxplot summarizing the analysis and provides a comparison between the proposed semi-population and worst, average and best case for random selection.

The significance of results was also tested by performing a t-test analysis on the 57 subjects. The t-test compared balanced accuracy obtained using the semi-population to the worst, average, and best case of random selection using a threshold of $p = 0.05$ for the null hypothesis. Results are shown in Table 4 confirming that values for the *n*-closest
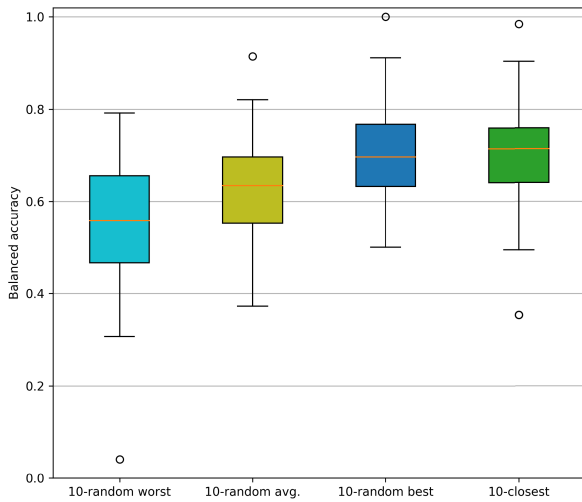
FIGURE 7. Boxplot obtained with balanced accuracy measured on the 57 users comparing the *n*-closest with the worst, average and best case of a random selection.

TABLE 4. Results of the t-test analysis performed using balanced accuracies obtained in the 57 subjects population using the *n*-closest, *n*-random worst, average and best case.

Comparison	p-value	NULL Hypothesis
5-worst vs. 5-closest	1.09731e-11	Rejected
5-average vs. 5-closest	0.000013	Rejected
5-best vs. 5-closest	0.39825	Fail to reject
10-worst vs. 10-closest	1.02969e-9	Rejected
10-average vs. 10-closest	0.000156	Rejected
10-best vs. 10-closest	0.38411	Fail to reject

semi-population method are comparable to the best case of random selection, outperforming the worst and average random choice for both $n = 5$ and $n = 10$.

VII. DISCUSSION

Results in Table 3 show how the semi-population model trained on $n = 5$ and $n = 10$ nearest users scored similar results, both in terms of macro-average precision and recall. After the adaptation step, however, the model initialized on 10 users scored better results, indicating that on average a dataset including more diverse data could be beneficial. For the adapted model obtained from random initialization instead, similar performances were observed for the cases of $n = 5$ and $n = 10$. Adaptation reduces the gap between random selection and proposed semi-population from 7-10% to 2-4% F-score following adaptation. We can expect the two initialization strategies (semi-population and random selection) to converge by increasing the number of new samples provided for adaptation. The semi-population approach, however, exhibits a faster growing learning curve. This gap is particularly relevant, since adaptation requires new user-specific labeled data, and the semi-population approach can reduce the amount of data required for adaptation.

The confusion matrices in Fig. 6 show how error rates are distributed across classes. In particular, walking, running

and cycling classes show a significant gap between the semi-population and the random approach. The gap slightly decreases after adaptation.

We can distinguish two parts of the matrix concerning static classes (sitting, lying) and active classes (walking, running and cycling).

A high conflict was observed between the sitting and lying class. This phenomenon could be related with a phone-off scenario that could happen when leaving the phone on the desk while sitting or lying. This situation is often addressed by considering a combined 'idle' class including lying and sitting, however, in the experiment we kept the classes separated to allow comparison of results in [6]. The cycling class is a problematic class since it can be conflicting with either the sitting or walking class depending on the smartphone location (e.g. in trouser pocket vs. in the bag). The dataset contains only a few labels indicating smartphone location that could help to investigate the confusion, however, the semi-population strategy appears to reduce the conflict both in the adapted and non-adapted case. Semi-population helps to isolate users presumably having a similar consistent behavior in terms of phone location.

Active classes (walking, running, and cycling) appear to benefit more from both semi-population and adaptation, as they could be expected also to be the classes with higher cross-subject variability. For minority classes (running and cycling), classification was more challenging. It must be considered that there were less samples for these activities, and only 25 and 26 users had samples for the running and cycling classes respectively. Therefore these classes suffered also the fact that they were statistically less represented in the population.

Finally, Fig. 7 presents the statistical analysis of balanced accuracy measured in the set of 57 users. The boxplot shows how the semi-population method helps to improve balanced accuracy with respect to the worst and average case of *n*-random selection, confirming that the criterion helps to obtain results in line with the best case.

A. LIMITATIONS

Despite the large size of the dataset there are some limitations to consider. Of the 57 users only a subset had either running or cycling samples. Meaning that cross-subject variability is potentially less represented for these two classes. Nonetheless, this evaluation was performed in a significantly larger dataset, compared to other studies, and in the challenging context of free-living naturalistic conditions. The set of target activities allows to compare results with [6]. On the other hand, a limitation of this target set is that there is no null class considered in the experiment.

Performed hyperparameter optimization was not exhaustive. For instance, learning rate was kept fixed to allow fair comparison between the n-random and n-closest. Nonetheless, different strategies concerning the learning rate could be investigated in the future. Similarly, the experiment evaluated only the Adam optimizer. A comparison with

Stochastic Gradient Descent (SGD) could be also considered in the future, since in some cases SGD has been reported to provide better generalization with respect to Adam [39].

VIII. CONCLUSION

As other studies have highlighted in the past, working with a dataset collected in free-living is a necessary step before deploying solutions in the wild. Working with such datasets, however, introduces the challenge of dealing with the multiple factors contributing to cross-subject variability. To date, ability to deal with this challenge using a generic approach (*i.e.* 'one-size-fits-all' approach) represents a severe obstacle for real-world deployment of models. Personalization of models aims at addressing this challenge adapting classifiers to the target subject. Alas, adaptation usually requires large amounts of user-specific labeled data to solve the problem.

In this work, a novel semi-population based method was proposed. The conducted experiments suggest that a semi-population approach for model initialization can reduce the amount of data required for personalization. The experiment evaluated the proposed method against a conventional approach over a large real-world dataset, simulating the application of an online personalization architecture on a real-world case with 57 subjects. Results obtained confirm that such a method can help reducing the amount data needed for model adaptation, thus paving the way to the deployment of systems implementing incremental personalisation, reducing the time needed for adaptation, while minimizing the required amount of interaction from the end user.

REFERENCES

[1] M. Shoaib, S. Bosch, O. Incel, H. Scholten, and P. Havinga, "A survey of online activity recognition using mobile phones," *Sensors*, vol. 15, no. 1, pp. 2059–2085, Jan. 2015.

[2] J. Morales and D. Akopian, "Physical activity recognition by smartphones—A survey," *Biocybern. Biomed. Eng.*, vol. 37, no. 3, pp. 388–400, 2017.

[3] C. A. Ronao and S.-B. Cho, "Human activity recognition with smartphone sensors using deep learning neural networks," *Expert Syst. Appl.*, vol. 59, pp. 235–244, Oct. 2016.

[4] R. San-Segundo, H. Blunck, J. Moreno-Pimentel, A. Stisen, and M. Gil-Martín, "Robust human activity recognition using smartwatches and smartphones," *Eng. Appl. Artif. Intell.*, vol. 72, pp. 190–202, Jun. 2018.

[5] J.-L. Reyes-Ortiz, L. Oneto, A. Samà, X. Parra, and D. Anguita, "Transition-aware human activity recognition using smartphones," *Neurocomputing*, vol. 171, pp. 754–767, Jan. 2016.

[6] Y. Vaizman, K. Ellis, and G. Lanckriet, "Recognizing detailed human context in the wild from smartphones and smartwatches," *IEEE Pervasive Comput.*, vol. 16, no. 4, pp. 62–74, Oct. 2017.

[7] T. Sztyler and H. Stuckenschmidt, "Online personalization of cross-subjects based activity recognition models on wearable devices," in *Proc. IEEE Int. Conf. Pervasive Comput. Commun. (PerCom)*, Mar. 2017, pp. 180–189.

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

[9] C.-H. Lu, Y.-C. Ho, Y.-H. Chen, and L.-C. Fu, "Hybrid user-assisted incremental model adaptation for activity recognition in a dynamic smart-home environment," *IEEE Trans. Human-Mach. Syst.*, vol. 43, no. 5, pp. 421–436, Sep. 2013.

[10] O. D. Incel, M. Kose, and C. Ersoy, "A review and taxonomy of activity recognition on mobile phones," *BioNanoScience*, vol. 3, no. 2, pp. 145–171, Jun. 2013.

[11] M. Ermes, J. Parkka, J. Mantyjarvi, and I. Korhonen, "Detection of daily activities and sports with wearable sensors in controlled and uncontrolled conditions," *IEEE Trans. Inf. Technol. Biomed.*, vol. 12, no. 1, pp. 20–26, Jan. 2008.

[12] J.-H. Hong, J. Ramos, and A. K. Dey, "Toward personalized activity recognition systems with a semipopulation approach," *IEEE Trans. Human-Mach. Syst.*, vol. 46, no. 1, pp. 101–112, Feb. 2016.

[13] A. Holzinger, "Introduction to machine learning & knowledge extraction (MAKE)," *Mach. Learn. Knowl. Extraction*, vol. 1, no. 1, pp. 1–20, Jul. 2017.

[14] L. Zhao, G. Sukthankar, and R. Sukthankar, "Robust active learning using crowdsourced annotations for activity recognition," in *Proc. Workshops 25th AAAI Conf. Artif. Intell.*, 2011, pp. 74–79.

[15] K. Yordanova, J. Favela, and G. Marcu, "Challenges providing ground truth for pervasive healthcare systems," *IEEE Pervasive Comput.*, vol. 18, no. 2, pp. 100–104, Apr. 2019.

[16] F. Cruciani, I. Cleland, C. Nugent, P. McCullagh, K. Synnes, and J. Hallberg, "Automatic annotation for human activity recognition in free living using a smartphone," *Sensors*, vol. 18, no. 7, p. 2203, Jul. 2018.

[17] I. Cleland, G. Sukthankar, and R. Sukthankar, "Evaluation of prompted annotation of activity data recorded from a smart phone," *Sensors*, vol. 14, no. 9, pp. 15861–15879, Aug. 2014.

[18] M. Stikic, D. Larlus, S. Ebert, and B. Schiele, "Weakly supervised recognition of daily life activities with wearable sensors," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 33, no. 12, pp. 2521–2537, Dec. 2011.

[19] O. D. Lara and M. A. Labrador, "A survey on human activity recognition using wearable sensors," *IEEE Commun. Surveys Tuts.*, vol. 15, no. 3, pp. 1192–1209, 3rd Quart., 2013.

[20] Z. Zhao, Y. Chen, J. Liu, Z. Shen, and M. Liu, "Cross-people mobile-phone based activity recognition," *IJCAI*, vol. 11, no. 3, pp. 2250–2545, 2011.

[21] W.-Y. Deng, Q.-H. Zheng, and Z.-M. Wang, "Cross-person activity recognition using reduced kernel extreme learning machine," *Neural Netw.*, vol. 53, pp. 1–7, May 2014.

[22] G. Hamerly and C. Elkan, "Alternatives to the k-means algorithm that find better clusterings," in *Proc. 11th Int. Conf. Inf. Knowl. Manage. (CIKM)*, 2002, pp. 600–607.

[23] A. Reiss and D. Stricker, "Personalized mobile physical activity recognition," in *Proc. 17th Annu. Int. Symp. Int. Symp. Wearable Comput. (ISWC)*, 2013, p. 25.

[24] D. Anguita, A. Ghio, L. Oneto, X. Parra, and J. L. Reyes-Ortiz, "A public domain dataset for human activity recognition using smartphones," in *Proc. Esann*, vol. 3, 2013, p. 3.

[25] S. Zhao, W. Li, and J. Cao, "A user-adaptive algorithm for activity recognition based on K-means clustering, local outlier factor, and multivariate Gaussian distribution," *Sensors*, vol. 18, no. 6, p. 1850, Jun. 2018.

[26] P. Siirtola, H. Koskimäki, and J. Röning, "Personalizing human activity recognition models using incremental learning," in *Proc. 26th Eur. Symp. Artif. Neural Netw., Comput. Intell. Mach. Learn. (ESANN)*, Apr. 2018, pp. 627–632.

[27] L. Hu, Y. Chen, J. Wang, C. Hu, and X. Jiang, "OKRELM: Online kernelized and regularized extreme learning machine for wearable-based activity recognition," *Int. J. Mach. Learn. Cybern.*, vol. 9, no. 9, pp. 1577–1590, Sep. 2018.

[28] A. Mannini and S. S. Intille, "Classifier personalization for activity recognition using wrist accelerometers," *IEEE J. Biomed. Health Informat.*, vol. 23, no. 4, pp. 1585–1594, Jul. 2019.

[29] A. Ferrari, D. Micucci, M. Mobilio, and P. Napoletano, "On the personalization of classification models for human activity recognition," *IEEE Access*, vol. 8, pp. 32066–32079, 2020.

[30] Y. Cheng, "Mean shift, mode seeking, and clustering," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 17, no. 8, pp. 790–799, Aug. 1995.

[31] C. C. Aggarwal, A. Hinneburg, and D. A. Keim, "On the surprising behavior of distance metrics in high dimensional space," in *Proc. Int. Conf. Database Theory*. Berlin, Germany: Springer, 2001, pp. 420–434.

[32] S. Xia, Z. Xiong, Y. Luo, WeiXu, and G. Zhang, "Effectiveness of the Euclidean distance in high dimensional spaces," *Optik*, vol. 126, no. 24, pp. 5614–5619, Dec. 2015.

[33] M. Janidarmian, A. R. Fekr, K. Radecka, and Z. Zilic, "A comprehensive analysis on wearable acceleration sensors in human activity recognition," *Sensors*, vol. 17, no. 3, p. 529, Mar. 2017.

[34] D. P. Kingma and J. Ba, "Adam: A method for stochastic optimization," 2014, *arXiv:1412.6980*. [Online]. Available: https://arxiv.org/abs/1412.6980

[35] F. Chollet. (2015). *Keras*. [Online]. Available: https://github.com/keras-team/keras

[36] M. Abadi. (2015). *TensorFlow: Large-Scale Machine Learning on Heterogeneous Systems*. [Online]. Available: https://www.tensorflow.org/

[37] H. Carrillo, K. H. Brodersen, and J. A. Castellanos, "Probabilistic performance evaluation for multiclass classification using the posterior balanced accuracy," in *Proc. 1st Iberian Robot. Conf. (ROBOT)*. Cham, Switzerland: Springer, 2014, pp. 347–361.

[38] F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, and J. Vanderplas, "Scikit-learn: Machine learning in Python," *J. Mach. Learn. Res.*, vol. 12, pp. 2825–2830, Nov. 2011.

[39] N. S. Keskar and R. Socher, "Improving generalization performance by switching from Adam to SGD," 2017, *arXiv:1712.07628*. [Online]. Available: http://arxiv.org/abs/1712.07628

IAN CLELAND received the B.Sc. degree in biomedical engineering and the Ph.D. degree from Ulster University, U.K. He is currently a Lecturer of data analytics with the School of Computing, Ulster University. His research interests include the development and evaluation of novel healthcare technologies that incorporate concepts from pervasive computing, biomedical engineering, and behavioral science.

FEDERICO CRUCIANI received the B.Sc. Eng. and M.Sc. Eng. degrees in computer science engineering from the University of Florence, Italy, and the Ph.D. degree from Ulster University, in 2020. From 2006 to 2016, he worked as a Software Project Manager with I+ and Orthokey focusing on Computer Assisted Surgery and eHealth, working on optical and inertial sensors applied to kinematic analysis and rehabilitation. In 2016, he joined Ulster University as a Marie Curie Fellow of computer science. He is currently working as a Research Fellow with Ulster University.

PAUL MCCULLAGH received the B.Sc. and Ph.D. degrees in electrical engineering from the Queen's University of Belfast, in 1979 and 1983, respectively. He is currently a Reader of computing with Ulster University. His research interests include biomedical signal and image processing, data mining, brain–computer interface, and assisted living applications.

CHRIS D. NUGENT (Member, IEEE) received the B.Eng. degree in electronic systems and the D.Phil. degree in biomedical engineering from Ulster University, U.K., in 1995 and 1998, respectively. He is currently a Professor of biomedical engineering with the School of Computing, Ulster University. His research interests include data analytics for smart environments and the design and evaluation of pervasive and mobile solutions within the context of ambient-assisted living.

KÅRE SYNNES is currently a Professor of pervasive and mobile computing with the Luleå University of Technology, Sweden. He is also involved in research within e-health and is developing a new research lab, the activity laboratory, which is a smart home environment for studying assistive technologies in home settings for people with cognitive disabilities. He has a long experience from research projects at a European scale and was a Co-Founder of the startup Marratech, which was procured by Google in 2007.

JAVIER MEDINA QUERO received the M.Sc. and Ph.D. degrees in computer science from the University of Granada, Spain, in 2007 and 2010, respectively. He is currently working as a Postdoctoral Researcher with the University of Jaen. He has published more than 25 articles in impact factor journals and participated in national and international research projects. His research interests include fuzzy logic, e-health, intelligent systems, ubiquitous computing, and ambient intelligence.

JOSEF HALLBERG is currently an Associate Professor of pervasive and mobile computing with the Luleå University of Technology, Sweden. He has over 15 years of experience in e-health related research. His research interests include technology for supporting people in everyday life, including supporting people with chronic disease and preventing disease, tools for increasing and supporting motivation for healthy behavioral change, data analysis, and decision support systems.

• • •