

Article

Daily Living Activity Recognition In-The-Wild: Modeling and Inferring Activity-Aware Human Contexts

Muhammad Ehatisham-ul-Haq ^{1,*}, Fiza Murtaza ¹, Muhammad Awais Azam ² and Yasar Amin ³

¹ Department of Creative Technologies, Faculty of Computing and Artificial Intelligence (FCAI), Air University, Islamabad 44000, Pakistan; fizamurtaza@yahoo.com

² Technology and Innovation Research Group, School of Information Technology, Whitecliffe, Wellington 6145, New Zealand; muhammada@whitecliffe.ac.nz

³ Department of Telecommunication Engineering, University of Engineering and Technology (UET), Taxila 47050, Pakistan; yasar.amin@uettaxila.edu.pk

* Correspondence: ehatishamuet@gmail.com

Abstract: Advancement in smart sensing and computing technologies has provided a dynamic opportunity to develop intelligent systems for human activity monitoring and thus assisted living. Consequently, many researchers have put their efforts into implementing sensor-based activity recognition systems. However, recognizing people's natural behavior and physical activities with diverse contexts is still a challenging problem because human physical activities are often distracted by changes in their surroundings/environments. Therefore, in addition to physical activity recognition, it is also vital to model and infer the user's context information to realize human-environment interactions in a better way. Therefore, this research paper proposes a new idea for activity recognition in-the-wild, which entails modeling and identifying detailed human contexts (such as human activities, behavioral environments, and phone states) using portable accelerometer sensors. The proposed scheme offers a detailed/fine-grained representation of natural human activities with contexts, which is crucial for modeling human-environment interactions in context-aware applications/systems effectively. The proposed idea is validated using a series of experiments, and it achieved an average balanced accuracy of 89.43%, which proves its effectiveness.

Keywords: activity recognition; context recognition; human-centric computing; in-the-wild; smart sensor; supervised learning



Citation: Ehatisham-ul-Haq, M.; Murtaza, F.; Azam, M.A.; Amin, Y. Daily Living Activity Recognition In-The-Wild: Modeling and Inferring Activity-Aware Human Contexts. *Electronics* **2022**, *11*, 226. <https://doi.org/10.3390/electronics11020226>

Academic Editor:
Muhammad Muaaz

Received: 26 November 2021
Accepted: 5 January 2022
Published: 12 January 2022

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Human beings are the most integral part of an environment and ecological units that collaborate and make up the urban landscape. Generally, human behavior reflects their surroundings, and varying environments adversely affect human psychology and, thus, their behavior. It is crucial to understand how human beings (as the occupants of an environment) react and acclimate to their surroundings. Naturally, human behavior and activity patterns are chaotic and inconsistent, primarily affected by the variability of environment and contexts. Diverse human contexts may lead a person to behave irrationally, thus giving rise to abnormal user behavior, because of which human physical activity patterns may also get distracted. Therefore, it is crucial to efficiently model and learn human physical activities and interactions in varying contexts for enabling context-dependent systems and applications. Recent advancements in sensing and networking technologies, such as the internet of things (IoT), have provided a ubiquitous platform to develop intelligent systems for context-aware human-centric computing. The accessibility of real-time data through ubiquitous devices (such as smartphones and smartwatches) has resulted in the proliferation of research work in the field of sensor-based activity recognition (AR) [1–3]. The goal of AR is to provide a suitable analysis of human activities from the data acquired from wide-ranging sensors, including video cameras and depth sensors,

infrared (IR) sensors, ambient sensors, and three-dimensional (3D) inertial sensors [4]. In this regard, lots of AR methodologies have been developed and implemented in the past few years using different sensing modalities. Various researchers developed camera-based activity detection/recognition systems for identifying the activities of interest [5–8]. However, camera-based AR approaches are subjected to privacy constraints. They are poorly affected by certain issues, such as camera viewpoint variation and camera motion, light sensitivity, occlusion, and background activities of the people. With the advancement in sensing technologies, the anomalies associated with the camera-based AR methods are addressed using miniaturized sensors that are independent of illumination changes and offer ubiquitous activity monitoring. These sensors are computationally efficient and provide 3D motion representation. Hence, sensor-based AR has now become indispensable for human-centric computing [9,10].

Different sensing modalities have been employed individually as well as in combination for AR tasks. Ambient sensors (including a pressure sensor, infra-red sensor, light sensor, temperature sensor, etc.) are generally adopted for indoor activity monitoring in smart homes and buildings [11–13]. In contrast, wearable and smartphone-embedded sensors enable pervasive and continuous monitoring of human activities in diverse environments [14]. The authors in [15–18] provide a review of the wearable sensing systems and technologies for human activity and health monitoring. Wearable or on-body sensors are mainly advantageous because they can be fitted at multiple body positions to monitor human activities robustly [19–22]. However, these sensors may turn into a cause of disruption for the people during activity execution. As a consequence, the motive of identifying the participants' rational behavior is often disregarded. The authors in [23] proposed "smart garments" for classification of human physical activities as well. The continuous growth in smartphone developing technologies has offered numerous applications for smartphone-embedded sensors in AR studies [24–28]. However, the variations in smartphone position or orientation may poorly affect AR performance. In addition, smartphone-based sensors are not sufficiently adept at tracking the user's activities relating to hand movements or gestures, for example, eating, smoking, drinking, tooth brushing, etc. Henceforth, the researchers emphasized using heterogeneous and multimodal sensors for AR [29–32], which has enhanced AR system performance in most cases. Most of the existing AR methods are generally developed and implemented in restricted environments and settings to learn and identify a particular set of human activities [33]. When collecting data, the subjects are prepared to execute the activities of interest in a predefined manner (following a set of protocols) for training the AR model. As a result, natural user behavior gets disregarded, which poorly affects the performance of such AR models in-the-wild. A few research studies [34–39] have emphasized utilizing either smartphone sensors or heterogeneous sensors for identifying human contexts, e.g., "indoor" vs. "outdoor", "in a car" vs. "in a bus", "sitting" vs. "driving", etc. However, these schemes fail to model human-environment interactions in-the-wild to infer the natural users' context in combination with their primary physical activity. Consequently, the notion of context-aware AR cannot be fulfilled, which is essential to aid context-dependent systems that acclimate to the subject's activities and the associated environments.

This paper presents a novel method for sensor-based activity recognition in-the-wild (ARW) to address the above-discussed challenges and discrepancies. The proposed method incorporates the recognition of human contexts (such as social/behavioral contexts and phone contexts) with the AR task to provide a fine-grained representation of human daily living activities in their natural surroundings. Thus, it enables answering the questions related to human activities and contexts in-the-wild, for example, "what is the person doing?", "where is the person situated/located?", and "why the person is here?". The proposed idea is conceived based on the fact that diverse contexts and surroundings greatly influence human physical activity and behavior patterns. For example, sitting postures in a car and in a meeting are generally quite different. Likewise, walking alone is usually different from walking with a group of friends. Thus, any significant change in the human

behavioral context or environment may lead to ample variations in human physical activity patterns. Likewise, the phone context (i.e., phone position on the human body) during a particular activity execution in-the-wild is also affected as a result of diversity in human behavioral environments. The changes in phone position significantly alter the activity patterns recorded by the device-embedded inertial sensors that are sensitive to phone orientation and placement. These vital differences in the activity patterns (occurring in response to change in human contexts or phone positions) can be effectively modeled based on a supervised machine learning approach to infer the detailed user's contexts associated with different activities. Following this, the proposed ARW model works based on a two-stage supervised classification approach. The first stage identifies the primary physical activities of daily living (PADLs) based on smartphone and smartwatch accelerometers. The second stage entails inferring knowledge about activity-related contexts based on the activity recognized in the first stage, thus providing the notion of activity-aware context recognition. The second stage further consists of two building blocks, i.e., behavioral context recognition (BCR) and phone context recognition (PCR), which independently learn and recognize the specified set of contexts using accelerometer sensors. The outputs from both stages are finally aggregated to form a triplet of information, i.e., {primary physical activity, behavioral context, and phone context}. In this manner, our proposed ARW scheme offers a multi-label and fine-grained/context-aware representation of human daily living activities in-the-wild. The coinciding recognition of the participant's physical activity, behavioral/social context, and phone position is essential for human behavior modeling and cognition in their living environments [40]. Inferring phone positions with activity can be effective for inferring the habitual behavior of a person in different contexts. Thus, the proposed scheme can be extended to detect and recognize normal/abnormal human behavior, which can further be advantageous in predicting/avoiding health-related risks, as discussed in the existing studies [41–43]. In addition, the proposed scheme can serve as a building block for recommender systems and context-aware computing applications.

The experiments for the proposed scheme are conducted using a public domain "ExtraSensory" dataset [38] that involves daily living human activities and the associated contexts in-the-wild. For the AR task in the proposed scheme, six (06) PADLs, including sitting, walking, lying, standing, running, and bicycling, are chosen from the dataset, whereas for context recognition, the fourteen (14) most frequent context labels are selected for identification purpose. These labels provide information regarding the phone positions (such as phone on table or phone in bag/hand/pocket) and the participants' environmental/behavioral aspects (such as participant's location, social context, and secondary activity) during the primary activity execution in-the-wild. Figure 1 shows how different contexts are associated with the selected PADLs for enabling ARW. The relationship between the particular PADLs and the equivalent contexts is established by systematically analyzing the co-occurrences of different activity and context pairs available in the "ExtraSensory" dataset. A boosted decision tree (BDT) is used for evaluating the performance of the proposed ARW model, and the obtained results are compared against those obtained with a neural network (NN) classifier.

This research paper provides the following significant contributions.

- A two-stage model is proposed for ARW, which first identifies the primary physical activity and then uses this label to infer activity-related context information, thus providing a detailed activity representation in-the-wild.
- A methodical approach is conceived and followed to analyze the co-occurrences of different activity-context pairs in the "ExtraSensory" dataset. As a result, a set of ten (10) most frequent human behavioral contexts and four (04) phone contexts/positions are incorporated with six (06) primary PADLs, respectively, for ARW. The approach used to analyze and select activity-context pairs for the proposed ARW scheme is reproducible and can be applied to any multi-label dataset.
- An in-depth exploration of the proposed ARW scheme is conducted for feature selection, model selection (i.e., classifier selection), and classifier hyperparameter optimization.

tion to attain state-of-the-art recognition performance. Finally, based on the best-case experimental observations and parameters, the performance of boosted decision tree and neural network classifiers is further evaluated in detail for the proposed scheme using smartphone and watch accelerometers.

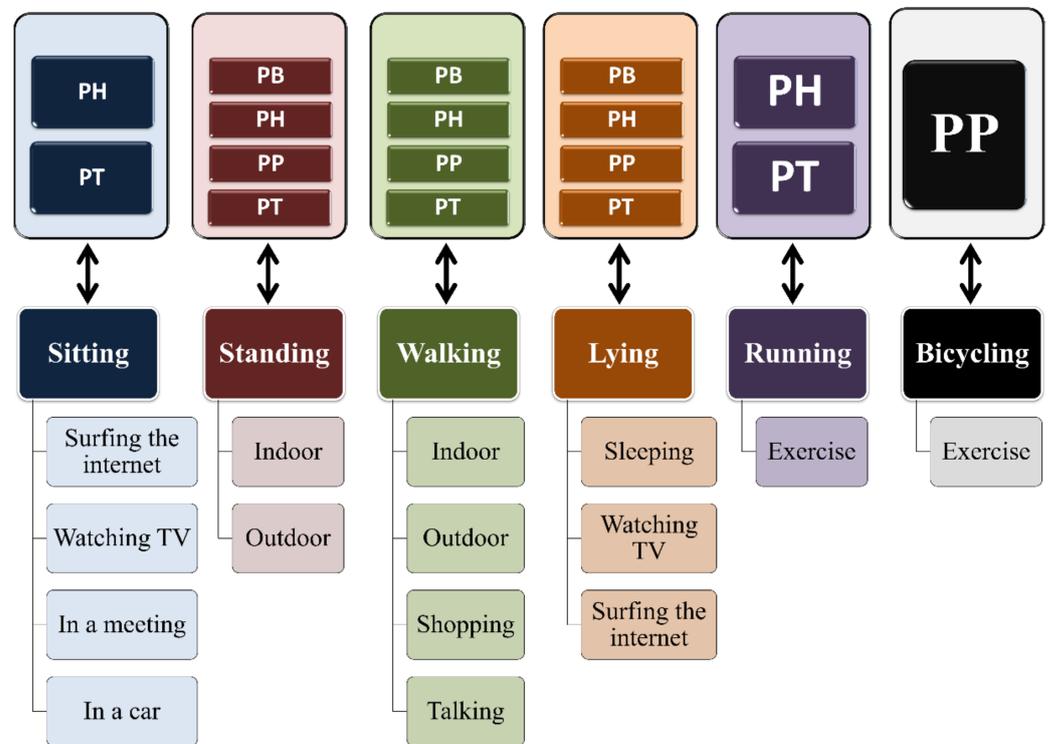


Figure 1. Primary physical activities with corresponding human behavioral contexts and phone positions, which are used for ARW. Here, PB, PH, PP, and PT represent phone in bag, phone in hand, phone in pocket, and phone on table, respectively.

The remaining paper portion is arranged as follows. Section 2 presents the related works for the proposed scheme. Section 3 provides the stepwise explanation of the proposed methodology in detail. Section 4 investigates and discusses the experimental results for our proposed ARW method in detail. Finally, Section 5 summarizes the research outcomes and provides future recommendations for the proposed scheme.

2. Related Works

The upsurge in smart systems with evolving sensing capabilities has made sensor-based AR a significant area of interest for researchers in the field of pervasive computing. Considering this, numerous schemes have been proposed for sensor-based AR, which can be classified as ambient AR, wearable AR, smartphone-based AR, and heterogeneous sensor-based AR approaches. Ambient AR systems aim to collect and process continuous data from various sensors installed in the environments for ambient assisted living (AAL). Numerous research studies have used ambient sensors to recognize PADLs and home tasks [11,12,44–46]. Vanus et al. [13] performed the fusion of gas (i.e., carbon dioxide) and audio sensors with humidity and temperature sensors to detect any person in the smart room. The authors employed a neural network for human detection and achieved accuracy greater than 95%. Ni et al. [47] proposed an ontology-based method for smart home activity monitoring, utilizing a three-layered approach for context-aware activity modeling. Ghayvat et al. [48] proposed an anomaly prediction model for detecting abnormal activity patterns of elderly people in the smart home environment. Likewise, Muheidat et al. [49] proposed a real-time fall detection scheme based on walking activity pattern monitoring using a sensor pad installed under a carpet. The primary advantage of using ambient

AR systems is their high accuracy rate and reliability. However, installing and setting up ambient sensors is a complex and expensive task, and the sensors are restricted to a particular area of monitoring. Hence, it is not possible to monitor natural human activities and behaviors in diverse contexts.

Wearable AR systems entail on-body sensors for recording and monitoring the participant's data. They are advantageous owing to their portability; thus, they can be taken to any place for continuous activity monitoring, including indoor and outdoor environments. The authors in [50–53] utilized the Inertial Measurement Units (IMUs) and other wearable sensors for activity monitoring. In [54], the authors proposed a probabilistic method using Bayesian formulation to recognize transition activities, such as stand-to-sit and sit-to-stand, using wearable sensors. They achieved 100% recognition accuracy for two activities. Mehrang et al. [55] utilized random forests for recognizing a number of daily living activities (including household activities) using wearable sensor data from a wrist-mounted accelerometer. In addition, they also used an optical heart rate sensor for the AR task and achieved an accuracy of $89.6 \pm 3.9\%$. The research work in [56] presented "HuMAN", a wearable AR system for the classification of 21 indoor human activities. In this aspect, the authors recorded data from ten subjects in a home environment using wearable sensors and extracted statistical signal attributes to train their proposed AR system. They used the conditional random field (CRF) classifier for AR task and acquired the best average accuracy around 95%. Anwary et al. [57] utilized wearable sensors (i.e., accelerometer and gyroscope) for monitoring and detecting abnormalities in the gait pattern of the participants. Moreover, wearable sensors have also been utilized for detecting and preventing abrupt human actions, such as falls [20]. The authors in [58] proposed a deep learning model for AR in the mountains using an accelerometer sensor. Nevertheless, wearable sensors often turn out to be a source of disturbance for the subjects in their activity execution, which hinders effective AR performance.

The increasing development in smartphone sensing technologies has offered a ubiquitous platform for sensor-based AR. Consequently, smartphone-based AR systems have been proposed by numerous research studies. The research studies in [59–64] proposed smartphone-based position-dependent and position-independent AR systems, respectively. Moreover, position-aware AR systems [34,65,66] have also been proposed, which employ a two-level or multi-level classification approach to identify a physical activity based on phone position recognition. In [67], Esfahani and Malazi presented "PAMS", a position-aware multi-sensor dataset for an AR task, where they achieved an average precision of approximately 88% in recognizing everyday physical activities in the dataset. Smartphones have also been employed for crowdsourcing and context recognition (such as indoor vs. outdoor, moving vs. stationary, etc.) [35,68–71]. However, smartphone-based AR systems are not sufficiently accomplished to detect or recognize activities involving hand gestures and arm movements. As a result, heterogeneous sensors have been used for AR tasks, which combine multimodal sensors (such as smartphones and wearable sensors) to improve AR performance [29–31,72,73], which is the case for our proposed scheme. With the evolution of deep learning algorithms in recent years, some authors have made use of these algorithms for the automatic extraction of high-level features from the sensor data to achieve promising AR results [74–78]. The survey work in [79,80] investigated the latest trends in sensor-based AR studies based on deep learning models and explained their pros and cons along with the future recommendations/implications. The high computational complexity of deep learning algorithms is a crucial challenge to be addressed in the case of sensor-based AR studies, which makes them ineffective for instant processing on battery-constrained devices, e.g., smartphones and smartwatches. Hence, there is a need for developing such schemes that are computationally efficient and can recognize natural user behavior in varying contexts with high accuracy, which is the main aim of our proposed ARW scheme.

3. Proposed Methodology

Figure 2 provides a block diagram of the proposed methodology for ARW that entails a two-stage classification model, consisting of four crucial steps as follows: (1) data acquisition and preprocessing, (2) feature extraction and feature selection, (3) primary physical activity recognition, (4) activity-aware context recognition. The subsequent sections present the necessary details for each step of the proposed ARW method.

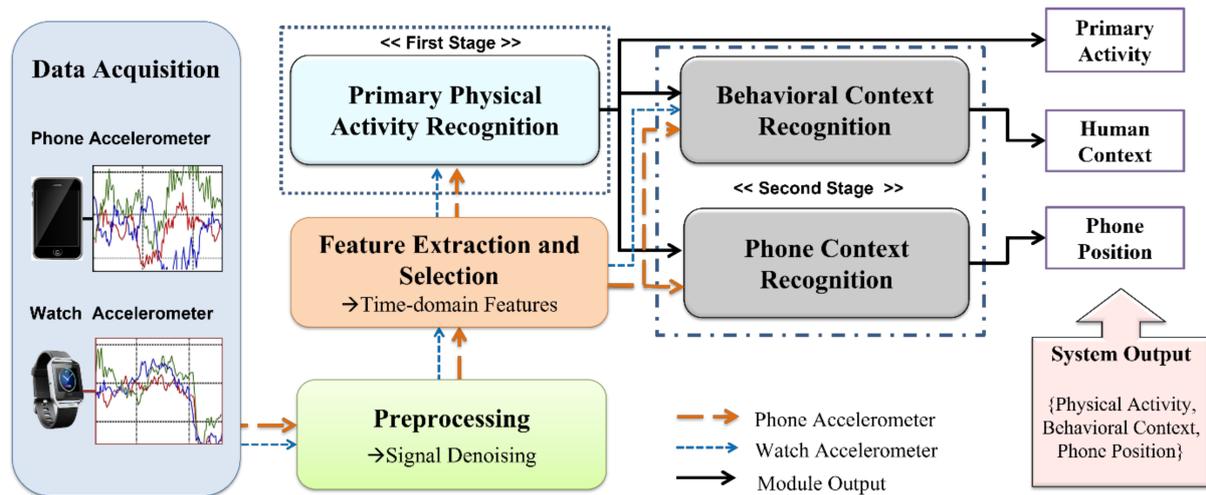


Figure 2. Block diagram of the proposed method for activity recognition in-the-wild (ARW).

3.1. Data Acquisition and Preprocessing

For the implementation and testing of any AR model, the first step is to acquire data concerning the activities of interest. Many researchers have utilized their efforts in collecting sensor-based datasets for AR, which entail data from different sensing modalities, including on-body wearable sensors, smartphone-embedded sensors, and multimodal heterogeneous sensors [81–85]. Generally, these datasets have been recorded in some constrained environments following a specific set of protocols for executing the scripted tasks. Therefore, there is a void of natural user behavior and any information regarding the participant’s context. The “ExtraSensory” dataset [38], presented by Yonatan and Ellis, contains in-the-wild human activity data from 60 subjects. Smartphone and smartwatch-based heterogeneous sensors are used to record natural user behavior regarding six (06) primary PADLs in diverse contexts. As the proposed scheme focuses on recognizing daily living human activities and their context details in-the-wild, the “ExtraSensory” dataset fits well into the proposed pipeline. As a result, we opted to utilize this dataset for the implementation and validation of the proposed ARW model. For the computational efficacy of the proposed method, only smartphone and smartwatch accelerometer data (collected with a sampling rate of 40 Hz and 25 Hz, respectively) are used for ARW. The existing AR studies [61,73] validate the efficient recognition performance of these sensors as compared to other inertial sensors, such as a gyroscope or a magnetometer.

3.1.1. Activity-Context Pairs for ARW: Systematic Analysis and Selection

The “ExtraSensory” dataset contains multiple secondary labels for each activity instance, which demonstrate detail regarding the participant’s context (for example, secondary activity, location, social and/or behavioral context, and phone state/position) during the primary activity execution. However, the context labels for each activity instance are not consistent as the data collection is conducted in-the-wild. To implement ARW, we systematically analyzed the PADLs and corresponding context labels to find out the most frequent activity-context pairs in the “ExtraSensory” dataset. In this regard, for all participants’ data, we counted the frequency of different context labels (including human behavioral contexts and phone positions) that occur in a pair with each of the six

selected daily living activities. In the end, we selected ten (10) and four (04) different human behavioral contexts and phone positions for context recognition, respectively, which had maximum frequencies of co-occurrence with the primary PADLs. Further, we tended to discard the activity instances having secondary labels with very few instances (i.e., less than 100), as they are not sufficient to be trained and tested for context recognition. Neglecting these instances has no adverse effect on the overall system training, due to the remaining instances still being very huge in number, i.e., 51,001.

Algorithm 1 shows the steps followed in extracting the activity-context pairs and their frequencies in the “ExtraSensory” dataset. These steps are reproducible and can be adopted for any multi-label dataset. Table 1 presents the list of primary PADLs and activity-context pairs along with their frequencies, which are finally chosen to validate the proposed ARW method. Two activities, including bicycling and running, are linked to only one behavioral context (i.e., exercise) and phone position (i.e., phone in pocket) as no more context labels exist with these activities in the “ExtraSensory” dataset. Likewise, for lying activity, only two phone positions (i.e., phone in hand and phone on table) are available, which are used in further analysis.

Algorithm 1. Extraction of activity-context pairs and their frequencies per user for ARW

Input: *userID*

Output: *prActCtxLabels* and *freqPrActCtxLabels*

prActCtxLabels and *freqPrActCtxLabels* show the labels and counts for all activity-context pairs per user, respectively.

```

1: Begin % Algorithm starts here
2: userData = readUserData(userID)
3: IN = countInstances(userData)
4: for rowID = 1 : IN % Iterate through all data
   instances per user
5: rowData = readDataChunk(rowID) % Read a data chunk with labels
6: prActLabel(rowID) = extractPrActLabel(rowData) % Extract primary activity
   label
7: ctxLabels(rowID, :) = extractCtxLabels(rowData) % Extract context labels
8: prActCtxLabels = [prActLabel(rowID), ctxLabels(rowID)] % Primary activity and
   context pairs
9: end for
10: prActCount = length(unique(prActLabel)) % Primary activities per user (which
   are generally fixed, i.e., 06)
11: ctxCount = length(unique(ctxLabels)) % Total number of secondary context labels
   for per user
12: freqPrAct = count(unique(prActLabel)) % Frequencies of all primary activities
   per user
13: freqCtxs = count(unique(ctxLabels)) % Frequencies of all secondary context
   labels per user
14: freqPrActCtxLabels = count((prActCtxPairs)) % Frequencies of all activity-context
   pairs per user
15: end

```

3.1.2. Signal De-Noiseing and Segmentation

The raw signals acquired from the accelerometer sensor of the smartphone/smartwatch are exposed to unwanted noise, for example, equipment noise or the noise produced by the subject’s unconscious movements. It is vital to de-noise the acquired signals before any further processing and computation. A lot of signal de-noising techniques have been used in AR literature, including time-domain and frequency-domain filtering methods. In this study, we employed a time-domain averaging filter (with size 1×3) for signal de-noising,

which is computationally cheap and capable of eliminating sudden noise, such as spikes, from the acquired signals.

Table 1. List of primary physical activities and different activity-context pairs utilized for ARW.

Primary Physical Activities			Physical Activities and Behavioral Contexts			Physical Activities and Phone Contexts	
Code	Activity	Count	Code	(Activity, Behavioral Context)	Count	(Activity, Phone Context)	Count
A1	Lying	20,348	A1C1	(Lying, Sleeping)	19,001	(Lying, Phone in Hand)	134
A2	Sitting	15,647	A1C2	(Lying, Surfing the Internet)	1069	(Lying, Phone on Table)	20,214
A3	Standing	7115	A1C3	(Lying, Watching TV)	278	(Sitting, Phone in Bag)	992
A4	Walking	2790	A2C1	(Sitting, Surfing the Internet)	7501	(Sitting, Phone in Hand)	1214
A5	Running	3488	A2C2	(Sitting, In a Car)	1427	(Sitting, Phone in Pocket)	618
A6	Bicycling	713	A2C3	(Sitting, In a Meeting)	1084	(Sitting, Phone on Table)	12,823
-	-	-	A2C4	(Sitting, Watching TV)	5635	(Walking, Phone in Bag)	383
-	-	-	A3C1	(Walking, Indoor)	534	(Walking, Phone in Hand)	768
-	-	-	A3C2	(Walking, Outdoor)	1715	(Walking, Phone in Pocket)	1406
-	-	-	A3C3	(Walking, Shopping)	145	(Walking, Phone on Table)	233
-	-	-	A3C4	(Walking, Talking)	396	(Standing, Phone in Bag)	426
-	-	-	A4C1	(Standing, Indoor)	6477	(Standing, Phone in Hand)	587
-	-	-	A4C2	(Standing, Outdoor)	638	(Standing, Phone in Pocket)	2013
-	-	-	A5C1	(Running, Exercise)	3488	(Standing, Phone on Table)	4089
-	-	-	A6C1	(Bicycling, Exercise)	713	(Running, Phone in Pocket)	3488
-	-	-	-	-	-	(Bicycling, Phone in Pocket)	713

Note: Each count represents a single data instance of 20-s duration in time.

The “ExtraSensory” dataset entails activity instances that are pre-segmented and labeled based on a 20-s time window with mutually exclusive samples. Generally, a fixed-size window of 2 s to 5 s is considered sufficient for simple AR, while complex AR deals with a larger window size having a time duration from 15 s to 30 s or more [28,30,86]. The proposed scheme aims to recognize the natural physical PADLs and in-the-wild activity-aware contexts, thus giving rise to complex AR. Hence, in accordance with the “ExtraSensory” dataset, a segmentation window of 20 s is used for feature extraction and classification in the proposed scheme.

3.2. Feature Extraction

After signal de-noising, features are extracted from the segmented data for further processing. Features are summarized representations of the essential signal attributes, which are fed as input into machine learning algorithms to classify a given chunk of data into one of the selected classes. Based on the existing AR studies [30,87–89], the proposed ARW model involves the extraction of twenty (20) time-domain features corresponding to each segmented data chunk. The extracted features include entropy, maximum signal amplitude, minimum amplitude, mean value, standard deviation of the signal, skewness, kurtosis, peak-to-peak value, peak-to-peak-time, median of the signal, maximum latency, minimum latency, latency-amplitude ratio, energy, signal variance, third moment of the signal, fourth moment of the signal, signal peak-to-peak slope, mean of first difference, and mean of second difference. The features are extracted for 3D data from the phone and watch accelerometer, thus resulting in a feature vector of size 1×60 per sensor.

Feature extraction is followed by feature selection to choose the most discriminating features from the whole set of extracted features. In this regard, we used a filter-based approach for supervised feature selection, which is known as “Correlation-based Feature

Subset Selection" (CfsSubetSel) [90]. This approach assesses the predictive power of each feature individually and finds redundancy between different features to produce the final set of most predictive features. After applying CfsSubetSel, the final subset of obtained features is used for classification in the next stage.

3.3. Primary Physical Activity Recognition

As discussed earlier, the proposed ARW model is based on a two-stage classification approach, where the first stage involves primary physical activity recognition (PPAR), i.e., the classification of six (06) primary PADLs in-the-wild. These activities include lying, sitting, walking, standing, running, and bicycling. Two machine learning classification algorithms, i.e., BDT and NN, are utilized for PPAR in a supervised manner.

A BDT [91] is an ensemble classifier that utilizes a combination of multiple decision trees (instead of using a single decision tree) to boost the output prediction performance. The main objective of the BDT algorithm is to sequentially combine a group of weak learners to create a strong learner. Each subsequent tree performs corrections for the errors in the preceding tree, and the final prediction is made based on the entire set of trees. In general, once aptly configured, a BDT is the easiest method for getting top recognition performances on wide-ranging machine learning tasks.

A NN [92] entails a set of interconnected layers, where the input layer is connected to the output layer using a feed-forward connection based on an acyclic graph consisting of weighted edges and nodes (i.e., neurons). A number of hidden layers can be inserted between the input and output layer; however, usually, one hidden layer is sufficient for most of the predictive tasks. Each node in a layer is connected to all the nodes in the subsequent layer using weighted edges. Each node in the hidden layers participates in generating the output of the network based on a non-linear activation function. This whole process is envisaged as an inspiration from the learning mechanisms of the human brain.

3.4. Activity-Aware Context Recognition

Human physical activity patterns alter with respect to change in their behavioral environments. These variations in the physical activity patterns can be monitored and tracked easily using the 3D accelerometer data from a smartphone/smartwatch to learn and identify the detailed activity contexts, such as human behavioral contexts and phone contexts. As follows, the second stage of the proposed ARW model entails activity-aware context recognition (AACR). The primary objective of AACR is modeling and detecting/recognizing the varying patterns of primary PADLs in diverse contexts to infer details about human behavioral contexts and phone positions in-the-wild. In this manner, the proposed ARW scheme enables activity-related contexts to be inferred based on activity pattern identification. An AACR module comprises two central units, including BCR and PCR. These units individually infer human behavioral context and phone context (i.e., phone position) labels, respectively, based on the activity recognized in the first stage (i.e., PPAR). In this aspect, for each selected primary activity, BDT and NN classifiers are trained to identify the relevant activity contexts (as given in Table 1). These classifiers are fed with the physical activity label (recognized in the first stage) and the final feature vector to train the proposed ARW system context recognition. Both smartphone and smartwatch accelerometers are used for BCR, while in the case of PCR, only a smartphone accelerometer is employed. Overall, for each classifier, four (04) different models are trained for BCR and PCR corresponding to four PADLs (including lying, sitting, walking, and standing). The activities of running and bicycling, which only involve one behavioral context and phone position, are ignored when training the proposed ARW model for AACR.

In the end, the outputs from both BCR and PCR units are aggregated with the output from the first stage (i.e., PPAR) of the model to provide a detailed and in-the-wild representation of daily living human activities. As follows, the proposed ARW scheme is capable of differentiating a large number of context-aware and fine-grained activities produced

as a result of different combinations of primary PADLs, human behavioral contexts, and phone positions.

4. Experimental Results, Performance Analysis, and Discussions

This section discusses the methods of validation and analysis used for assessing the performance of the proposed scheme. In addition, it evaluates and discusses the achieved experimental results in detail, as given in the following sections.

4.1. Method of Validation and Analysis

4.1.1. Model Selection and Hyperparameters Tuning

The proposed ARW scheme is implemented and validated using the Microsoft Azure machine learning tool [93]. The “AutoML” package of Microsoft Azure is used for model selection based on the “ExtraSensory” dataset, where a set of standard machine learning classifiers (including BDT, NN, k-nearest neighbours (K-NN), naïve Bayes (NB), and support vector machine (SVM)) are investigated for the proposed ARW scheme. In this aspect, the finally selected features set (obtained using CfsSubsetSel) is fed as input into machine learning classifiers to assess their performance for PPAR, BCR, and PCR experiments. Table 2 provides the list of finally selected features from each sensor, which are used for experimentation purposes. These features are simply concatenated in the case of sensor fusion.

Table 2. List of finally selected features for different recognition experiments in the proposed scheme.

Experiment Type	Based on (Activity)	Selected Features for Each Sensor Axis	Feature Vector Length per Sensor
PPAR	-	$A_x \rightarrow \{F2, F5, F8, F15, F12, F13, F16\}; A_y \rightarrow \{F2, F5, F8, F10, F13, F14, F15, F16\};$ $A_z \rightarrow \{F3, F4, F5, F8, F9, F10, F12, F13, F14, F15, F16, F17\};$ $W_x \rightarrow \{F1, F4, F5, F10, F14, F15, F16, F17, F18\}; W_y \rightarrow \{F4, F5, F8, F10, F15, F16,$ $F17, F18, F20\}; W_z \rightarrow \{F2, F5, F8, F10, F12, F15, F16, F17, F19, F20\}$	27
			28
BCR	Lying	$A_x \rightarrow \{F1, F2, F4, F6, F8, F10, F11, F20\}; A_y \rightarrow \{F1, F2, F6, F10, F15, F18, F20\};$ $A_z \rightarrow \{F2, F6, F10, F11, F12, F19, F20\};$ $W_x \rightarrow \{F1, F2, F4, F6, F10, F11, F15, F23\}; W_y \rightarrow \{F1, F2, F6, F10, F15, F16, F20\};$ $W_z \rightarrow \{F1, F6, F10, F11, F12, F19, F20\}$	22
			22
			22
	Sitting	$A_x \rightarrow \{F2, F4, F6, F10, F11, F20\}; A_y \rightarrow \{F1, F2, F6, F10, F15, F20\};$ $A_z \rightarrow \{F1, F2, F6, F10, F11, F12, F14, F20\};$ $W_x \rightarrow \{F1, F2, F4, F6, F10, F11, F20\}; W_y \rightarrow \{F1, F2, F4, F10, F15, F16, F20\};$ $W_z \rightarrow \{F2, F6, F10, F11, F12, F14, F20\};$	19
			21
	Walking	$A_x \rightarrow \{F2, F4, F5, F15, F10, F20\}; A_y \rightarrow \{F1, F2, F6, F10, F15, F20\};$ $A_z \rightarrow \{F2, F6, F10, F11, F12, F20\};$ $W_x \rightarrow \{F2, F4, F6, F10, F11, F15, F20\}; W_y \rightarrow \{F1, F2, F6, F10, F15, F20\};$ $W_z \rightarrow \{F1, F2, F6, F10, F11, F12, F14, F20\};$	19
			21
			21
Standing	$A_x \rightarrow \{F2, F4, F6, F8, F10, F11, F14, F20\}; A_y \rightarrow \{F1, F2, F4, F6, F10, F15, F19\};$ $A_z \rightarrow \{F1, F2, F4, F6, F10, F11, F12, F19, F20\};$ $W_x \rightarrow \{F2, F4, F6, F10, F11, F19\}; W_y \rightarrow \{F2, F4, F10, F15, F19\};$ $W_z \rightarrow \{F1, F2, F6, F8, F10, F11, F12, F19, F20\};$	24	
		20	
PCR	Lying	$A_x \rightarrow \{F2, F4, F10, F19\}; A_y \rightarrow \{F2, F6, F10, F19, F20\}; A_z \rightarrow \{F2, F4, F20\}$	12
	Sitting	$A_x \rightarrow \{F2, F4, F10, F11\}; A_y \rightarrow \{F4, F10, F20\}; A_z \rightarrow \{F4, F10, F19\}$	10
	Walking	$A_x \rightarrow \{F2, F4, F10, F19, F20\}; A_y \rightarrow \{F2, F6, F10, F15\}; A_z \rightarrow \{F2, F4, F10, F11\}$	13
	Standing	$A_x \rightarrow \{F2, F10, F20\}; A_y \rightarrow \{F6, F10, F19\}; A_z \rightarrow \{F2, F10, F15\}$	09

Note: **F1:** entropy; **F2:** maximum amplitude; **F3:** minimum amplitude; **F4:** signal mean; **F5:** standard deviation; **F6:** kurtosis; **F7:** skewness; **F8:** peak-to-peak value; **F9:** peak-to-peak-time; **F10:** signal median; **F11:** maximum latency; **F12:** minimum latency; **F13:** latency-amplitude ratio; **F14:** energy; **F15:** signal variance; **F16:** 3rd moment of the signal; **F17:** 4th moment of the signal; **F18:** signal peak-to-peak slope; **F19:** mean of 1st difference of the signal; **F20:** mean of 2nd difference of the signal. ² $A_x, A_y,$ and A_z represent the $x-, y-,$ and z -axis of the smartphone accelerometer, whereas $W_x, W_y,$ and W_z represent $x-, y-,$ and z -axis of the watch accelerometer, respectively. In the case of sensor fusion for PPAR and BCR, the finally selected features from each sensor are combined accordingly.

Following the model selection, BDT is chosen as the first choice for the proposed method implementation. Additionally, the NN classifier is employed to assess its performance for the proposed scheme in comparison to BDT, which has been successfully adopted for numerous sensor-based AR studies [60,87,94]. A one-vs.-all (OVA) classification approach is used for both classifiers, which utilizes an ensemble of C binary classifiers to solve a multiclass problem with C number of classes. The existing research work has demonstrated the effectiveness of using the OVA approach for multiclass classification, provided that the underlying binary classifiers are fine-tuned [95]. Following this, a random parameter sweep is performed on the data using five-fold cross-validation to explicitly learn the optimal hyperparameters of the selected classifiers for different recognition experiments. The maximum number of runs for the parameter sweep is set as 10. Finally, the best-tuned model hyperparameters are chosen for all recognition experiments, providing the best performance for the proposed scheme.

Table 3 presents the optimal hyperparameter values obtained for the selected classifiers regarding PPAR, BCR, and PCR experiments. In the case of BDT, multiple additive regression trees (MART) [91] is used as a decision tree algorithm, whereas gradient descent is used for error estimation. A fully connected (FC) hidden layer is used for the NN classifier with a sigmoid function as the output function. The number of nodes in the hidden layer is set equivalent to the average size of the input and output layer. The size of the input layer for different experiments is equal to the number of input features, which is given in Table 2. The output size represents the number of classes for each NN, which is six (06) and four (04) for PPAR and PCR, respectively. In the case of BCR, the size of the output layer is equal to the number of contexts corresponding to lying, sitting, walking, and standing activities. To evaluate the classification performance, an m -fold cross-validation method (with $m = 5$) is utilized. This validation scheme allows a model to train on multiple splits and uses all the data for training and testing in different iterations, thus ensuring fairness.

Table 3. Classifier hyperparameters for the best-case recognition results using the proposed scheme.

Experiment Type	Based on (Activity)	Sensor(s)	BDT Hyperparameters			NN Hyperparameters		
			No. of Leaves	Minimum Leaf Instances	Learning Rate (α)	No. of Trees	No. of Iterations	Learning Rate (α)
PPAR	-	A	61	17	0.2699	178	114	0.0368
		W	32	06	0.2520	270	88	0.0387
		A + W	26	02	0.2993	358	128	0.0287
BCR	Lying	A	08	50	0.1000	500	82	0.0320
		W	04	06	0.1120	57	40	0.0109
		A + W	86	22	0.0636	87	23	0.0306
	Sitting	A	128	10	0.4000	100	131	0.0135
		W	54	19	0.3364	51	131	0.0135
		A + W	62	02	0.4000	94	51	0.3380
	Walking	A	28	10	0.4000	100	96	0.0396
		W	32	05	0.2520	270	46	0.0144
		A + W	61	17	0.2699	178	97	0.3960
	Standing	A	32	50	0.2000	100	111	0.0158
		W	17	13	0.0629	50	55	0.0315
		A + W	30	16	0.2358	27	135	0.0135
PCR	Lying	A	06	47	0.1400	233	58	0.0101
	Sitting	A	59	27	0.3911	22	109	0.3030
	Standing	A	128	01	0.4000	20	23	0.0309
	Walking	A	36	07	0.3331	182	121	0.0301

Note: Here, A and W denote the smartphone and smartwatch accelerometer, respectively.

4.1.2. Performance Evaluation Metrics for Classification

The performance of BDT and NN classifiers is assessed independently for PPAR, BCR, and PCR experiments, based on accuracy, precision, sensitivity, F1-score, balanced accuracy,

and log loss. The mean value of true positive rate (i.e., sensitivity) and true negative rate (i.e., specificity) is termed as balanced accuracy (BALACC). It is the most crucial measure for assessing the classification performance of a system that entails imbalanced class data [38]. In addition, micro-averaging and macro-averaging metrics (i.e., micro-F1 and macro-F1) are computed for average performance comparison, where micro-precision and micro-sensitivity values are equal to micro-F1 scores. To estimate the classification error, log loss or logarithmic loss is used, which assesses the uncertainty of a model by comparing the output probabilities with ground truths. It expresses the penalty for misclassifications and is measured as a difference of two probability distributions, i.e., the true one and the one enclosed by the proposed model.

4.2. Performance Analysis of Primary Physical Activity Recognition (PPAR)

The first stage of the proposed ARW model incorporates PPAR, where six (06) in-the-wild PADLs are classified based on smartphone and watch accelerometer data, using BDT and NN classifiers. For this purpose, twenty time-domain features are extracted from each sensor channel, which are further subjected to feature selection and reduction using the CfsSubsetSel method. As a result, a set of twenty-nine (29) and twenty-eight (28) features (as shown in Table 2) is obtained related to the phone and watch accelerometer, respectively, which is used for classifier training and testing based on a five-fold cross-validation scheme.

Table 4 provides the average numerical results for PPAR, where the BDT classifier achieves the best results in classifying six selected PADLs. Using the phone and watch accelerometer alone, BDT achieves a macro-F1 score of 82.9% and 75.2%, respectively, for PPAR, which is 20.1% and 19.9% greater than the corresponding scores attained with the NN classifier. Similarly, the values of micro-F1 scores are also improved for the BDT classifier. These results also depict that the phone accelerometer performance is better than the watch accelerometer. With sensor fusion, the macro-F1 score is improved with a 6.4% and 14.1% rate in comparison to that achieved with the individual phone and watch sensor, respectively. Likewise, in the case of the NN classifier, the macro-F1 value is increased to 71.8% with sensor fusion, which is still 17.5% less than the best-case value obtained with BDT. The best error rate (i.e., average log loss of 0.787) for PPAR is also obtained with the BDT classifier using the combination of both sensors. Likewise, the values of other performance measures (i.e., precision and sensitivity) are also better for the BDT as compared to the NN classifier, where the best results are attained with sensor fusion.

Table 4. Average results for primary physical activity recognition (PPAR).

Classifier	Sensor(s)	Accuracy	Precision	Sensitivity	Micro-F1	Macro-F1	Log Loss
BDT	Acc.	0.959	0.856	0.807	0.877	0.829	1.215
	W. Acc.	0.941	0.825	0.708	0.822	0.752	1.414
	Acc. + W. Acc.	0.974	0.907	0.881	0.920	0.893	0.787
NN	Acc.	0.900	0.744	0.579	0.700	0.628	1.284
	W. Acc.	0.861	0.691	0.503	0.583	0.553	1.748
	Acc. + W. Acc.	0.921	0.789	0.677	0.763	0.718	1.018

Note: Acc. and W. Acc. symbolize the smartphone and smartwatch accelerometer, respectively.

Figure 3a compares the BALACC values obtained for PPAR using BDT and NN classifiers. It is evident from the figure that NN underperforms as compared to the BDT classifier in terms of the BALACC value as well. The best BALACC rate of 93.1% is achieved with the BDT classifier using sensor fusion, which is 10.8% more than the best value (BALACC = 82.3%) achieved with the NN. Likewise, using individual sensors for PPAR, the BALACC values achieved with the NN are worse than those obtained with the BDT classifier. These results validate the efficacy of utilizing a BDT classifier to recognize primary PADLs in-the-wild. Furthermore, adding the watch accelerometer with the phone accelerometer tends to achieve the best accuracy rate for PPAR. Generally, natural user behavior involves diverse behavioral contexts and phone positions, which may poorly

affect the activity pattern being recorded by smartphone sensors. For example, in the case of phone on table, it becomes quite impossible to recognize the participant's activity based on a smartphone accelerometer. In such cases, the use of a smartwatch accelerometer may help in learning and identifying the user's activities as the watch is supposed to be worn by the user most of the time.

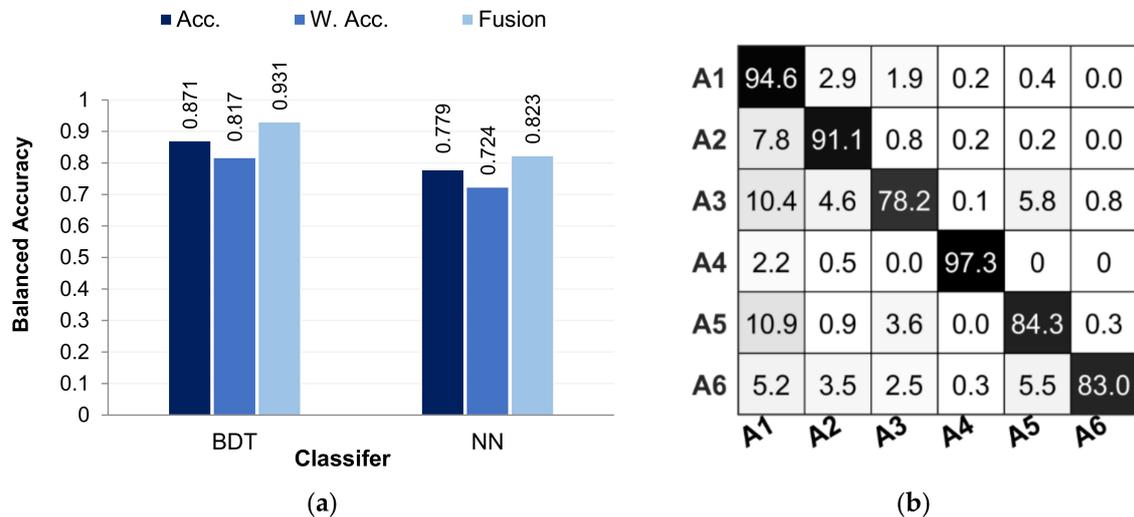


Figure 3. (a) Average results attained for PPAR using BDT and NN classifiers; (b) Confusion matrix (in percentage form) for the best-case PPAR performance accomplished with BDT using the phone and watch accelerometer combination. The labels A1-A6 represent six activities, i.e., lying, sitting, walking, standing, running, and bicycling, respectively.

To demonstrate the per-class recognition performance of the selected PADLs, Figure 3b displays the confusion matrix concerning the best-case PPAR performance (obtained with BDT using sensor fusion). The matrix rows and columns denote the ground truths and predicted outputs, respectively. The labels represent the codes for six (06) primary PADLs as follows: A1: lying, A2: sitting, A3: walking, A4: standing, A5: running, and A6: bicycling. It can be analyzed from the confusion matrix that most of the PADLs are truly classified with a high percentage. Notably, static activities (such as lying/sitting/standing) are truly recognized with a rate of more than 90%. The percentage of truly classified samples for walking, running, and bicycling activities is 78.2, 84.3, and 83, respectively, which shows that identification of the static activities in-the-wild is more comfortable than dynamic activities. This is due to the inconsistency of dynamic activity patterns with respect to diverse behavioral contexts, which gives rise to misclassifications of different dynamic activities in-the-wild. As a result, their recognition accuracies are reduced.

4.3. Performance Evaluation of Activity-Aware Context Recognition (AACR)

The proposed ARW scheme entails the recognition of activity-aware contexts in its second stage, where BDT and NN classifiers are trained explicitly for all PADLs to infer the associated human behavioral contexts and phone contexts (i.e., phone positions). The second stage consists of two parallel units, i.e., BCR and PCR, which take as input features from the sensor(s) data as well as the primary activity label (recognized in the first stage) to identify the corresponding behavioral contexts and phone positions independent of each other. As context labels for the primary activities are not the same, it is vital first to recognize the primary activity and then infer the context information being aware of the primary activity. BCR is performed based on the data from smartphone and smartwatch accelerometers, whereas in the case of PCR, only the smartphone accelerometer sensor is utilized. In this regard, the final subset of features selected for each accelerometer sensor (as shown in Table 2) is used to train and test the chosen classifiers for BCR and PCR using a five-fold cross-validation method. The results are evaluated separately for BCR and PCR

based on four different PADLs, while running and bicycling activities are ignored as they only involve one behavioral context and phone position that requires no classification. The following sections discuss the individual experimental results achieved for BCR and PCR.

4.3.1. Behavioral Context Recognition (BCR) Results and Investigation

The average numerical results obtained for BCR, based on four different PADLs, are presented in Table 5. By investigating these results, it can be stated that the BDT classifier performs significantly better than the NN classifier in recognizing activity-aware behavioral contexts. Furthermore, it can be analyzed that the phone accelerometer better recognizes most of the human behavioral contexts associated with four different physical activities, as compared to the other accelerometer sensor (i.e., watch accelerometer). In the case of BCR based on sitting, walking, and standing activities, the BDT classifier achieved macro-F1 scores of 97.0%, 74.1%, and 98.6%, respectively, using a phone accelerometer. These numerical results are 3.5%, 20.6%, and 1.5% better than the corresponding scores obtained for BCR based on the watch accelerometer, respectively. In contrast, for BCR based on lying activity, the macro-F1 score (i.e., 80.3%) obtained using the watch accelerometer is 4.8% more than the phone accelerometer. The same performance trend is observed in the case of average micro-F1 scores also. However, the fusion of both sensors provides the best-case BCR performance, where the best average macro-F1 scores of 86.8%, 97.8%, 76.4%, and 98.8%, respectively, are achieved for BCR based on lying, sitting, walking, and standing activity, using the BDT classifier. These values are 19.4%, 6.5%, 22.3%, and 2.8% more than the corresponding macro-F1 scores obtained with the NN classifier, respectively, using sensor fusion. The average numerical values of accuracy, precision, sensitivity, and log loss are also better for a BDT as compared to a NN, which proves the efficiency of a BDT classifier over an NN classifier for BCR experiments.

Table 5. Average results for behavioral context recognition (BCR) based on four physical activities.

Activity	Classifier	Sensor(s)	Accuracy	Precision	Sensitivity	Micro-F1	Macro-F1	Log Loss
Lying	BDT	Acc.	0.997	0.794	0.698	0.944	0.755	3.640
		W. Acc.	0.997	0.945	0.719	0.973	0.803	1.502
		Acc. + W. Acc.	0.985	0.940	0.812	0.976	0.868	1.287
	NN	Acc.	0.963	0.716	0.461	0.934	0.514	2.019
		W. Acc.	0.991	0.722	0.514	0.945	0.601	1.650
		Acc. + W. Acc.	0.969	0.813	0.602	0.953	0.674	1.867
Sitting	BDT	Acc.	0.986	0.956	0.960	0.975	0.958	0.178
		W. Acc.	0.970	0.953	0.920	0.939	0.935	0.435
		Acc. + W. Acc.	0.991	0.984	0.971	0.982	0.978	0.147
	NN	Acc.	0.928	0.824	0.869	0.856	0.844	0.465
		W. Acc.	0.911	0.850	0.747	0.821	0.785	0.877
		Acc. + W. Acc.	0.961	0.924	0.902	0.922	0.913	0.321
Walking	BDT	Acc.	0.918	0.854	0.683	0.835	0.741	1.560
		W. Acc.	0.859	0.703	0.485	0.717	0.535	1.980
		Acc. + W. Acc.	0.926	0.847	0.715	0.852	0.764	1.437
	NN	Acc.	0.823	0.441	0.324	0.646	0.318	1.857
		W. Acc.	0.818	0.433	0.316	0.626	0.310	2.296
		Acc. + W. Acc.	0.861	0.688	0.492	0.721	0.536	1.680
Standing	BDT	Acc.	0.996	0.987	0.986	0.996	0.986	0.057
		W. Acc.	0.988	0.973	0.966	0.988	0.970	0.111
		Acc. + W. Acc.	0.996	0.991	0.986	0.996	0.988	0.049
	NN	Acc.	0.983	0.988	0.907	0.983	0.943	0.241
		W. Acc.	0.987	0.880	0.974	0.987	0.920	1.118
		Acc. + W. Acc.	0.971	0.957	0.963	0.971	0.960	0.090

Note: Acc. and W. Acc. denotes the smartphone and smartwatch accelerometer, respectively.

Figure 4 compares the performance of different sensors for activity-aware BCR in terms of BALACC. The average BALACC rate achieved for BCR using the smartphone accelerometer is 71.9%, 92.9%, 78.2%, and 82.3% based on *lying*, *sitting*, *walking*, and *standing* activity, respectively, using the BDT classifier. In the case of BCR based on lying activity, the BALACC value achieved with the watch accelerometer is 7.2% and 6.2% better than that achieved with the phone accelerometer using the BDT and NN classifier, respectively. The combination of phone and watch accelerometers results in an increase in the BALACC values for BCR, particularly for lying and walking activities. The overall average BALACC value achieved for BCR (with sensor fusion) using the BDT classifier is 11% more than the NN classifier. Therefore, based on all these analyses and discussions, it is eminent that the combination of both accelerometers is the best choice for activity-aware BCR using a BDT classifier.

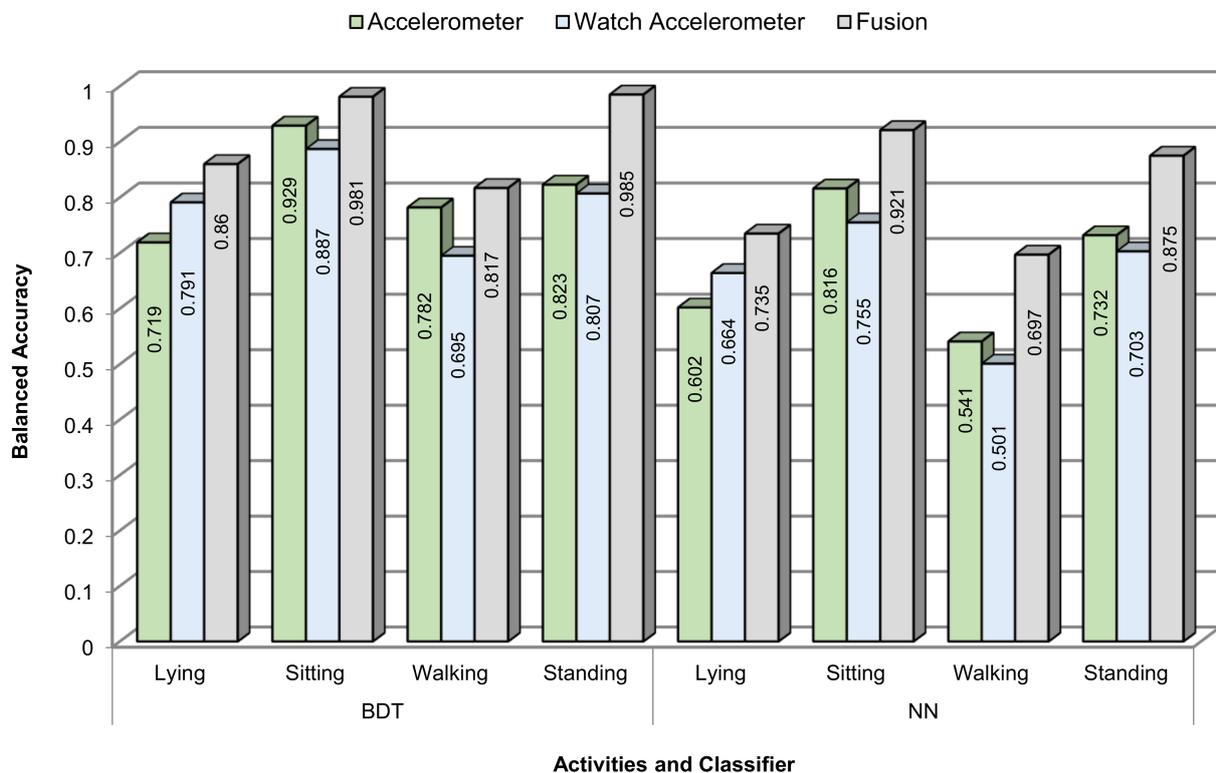


Figure 4. Average balanced accuracies achieved for BCR based on *lying*, *sitting*, *walking*, and *standing* activities using BDT and NN classifiers.

Table 6 provides the confusion matrices for the best-case BCR results obtained with the smartphone and smartwatch combination using the BDT classifier. It can be analyzed from the table that in the case of *sitting* and *standing* activities, the corresponding human behavioral contexts are truly classified with an accuracy of more than 90%. However, the individual recognition rates achieved for behavioral contexts associated with *lying* and *walking* activities are lower. In particular, the percentage of truly classified samples for A1C2 (surfing the internet based on lying), A1C3 (watching TV based on lying), A3C3 (shopping based on walking), and A4C4 (talking based on walking) is 68.8, 75.2, 46.6, and 61.9, respectively. These results depict the difficulty of accurately identifying these human behavioral contexts based on the recognition of associated physical activity patterns.

In general, the *sitting* and *standing* activity patterns of human beings show variations pertaining to different behavioral contexts. For instance, the *sitting* posture for most persons is altered when *working on a personal computer/laptop* or *watching TV*. Likewise, the pattern of *standing indoors* somewhat differs from *standing at some outdoor place*. These differences are realized by the 3D motion sensor (i.e., accelerometer) of smartphone/smartwatch to

efficiently model and recognize different behavioral contexts linked with these activities. As a result, the performance of BCR based on *sitting* and *standing* activity patterns is enhanced. In contrast, the *lying* activity typically attributes to the state of *relaxing*; thus, it does not often entail explicit body movements. Therefore, the recognition of associated human behavioral contexts becomes challenging. In addition, the phone position associated with *lying* is often *on table*, which yields unpredictable and inaccurate results for BCR. Therefore, it is very crucial to use the smartwatch in combination with a smartphone for BCR. In the case of *walking* activity, the recognition of associated human behavioral contexts becomes hard owing to the dynamic motion patterns of an individual in the same or different physical environment. These changes are often triggered as a result of chaotic human behavior and an emotional state that may instinctively alter the gait pattern of a subject. In addition, human behavior varies from one person to another, which makes it impractical to create a general model for BCR based on *walking* activity *in-the-wild*.

Table 6. Confusion matrices obtained for BCR with respect to four physical activities using BDT classifier.

		Predicted Output					Predicted Output			
Lying (A1)		A1C1	A1C2	A1C3	Sitting (A2)		A2C1	A2C2	A2C3	A2C4
Ground Truth	A1C1	99.63%	0.33%	0.04%	Ground Truth	A2C1	99.5%	0.1%	0.0%	0.4%
	A1C2	31.1%	68.8%	0.1%		A2C2	2.6%	96.6%	0.4%	0.4%
	A1C3	10.8%	14.0%	75.2%		A2C3	2.3%	0.3%	94.5%	3.0%
						A2C4	1.9%	0.0%	0.3%	97.8%
		Predicted Output						Predicted Output		
Walking (A3)		A3C1	A3C2	A3C3	A3C4	Standing (A4)		A4C1	A4C2	
Ground Truth	A3C1	82.2%	15.2%	0.4%	2.2%	Ground Truth	A4C1	99.8%	0.2%	
	A3C2	2.4%	94.9%	0.5%	2.3%		A4C2	2.7%	97.3%	
	A3C3	0.0%	47.6%	46.9%	5.5%					
	A3C4	1.5%	35.6%	1.0%	61.9%					

Note: The row and column labels for each confusion matrix represent different behavioral contexts associated with the specific physical activity, which are listed in Table 1.

4.3.2. Phone Context Recognition (PCR) Results and Investigation

Table 7 provides the average numerical results for PCR based on *lying*, *sitting*, *walking*, and *standing* activity patterns. Only the phone accelerometer sensor is used in this regard, which provided the macro-F1 scores of 83.1%, 91.1%, 71.1%, and 97.4% in recognizing different phone positions based on *lying*, *sitting*, *walking*, and *standing* activities, respectively, using the BDT classifier. For the same set of activities, the NN achieved corresponding scores of 49.8%, 34.5%, 31.3%, and 69.8%, respectively, which are quite low as compared to the BDT results. The values of other performance parameters (including *accuracy*, *precision*, *sensitivity*, *micro-F1*, and *log loss*) are also better for the BDT classifier. In addition, Figure 5 compares the PCR results in terms of BALACC, where the best recognition performance is also obtained using the BDT classifier. The overall average BALACC value for PCR based on a BDT is 13.5% more than the NN classifier. Moreover, it can be investigated from the figure that the average performance of PCR based on *sitting* and *standing* activities is better as compared to other activities.

Table 8 provides the confusion matrices for PCR based on four PADLs using the BDT classifier. The individual accuracies of different phone positions based on each physical activity can be computed from these confusion matrices. The row and column labels of the confusion matrices represent different phone positions (i.e., phone in bag (PB), phone in hand (PH), phone in pocket (PP), and phone on table (PT)). There are only two phone positions (i.e., PH and PT) associated with the *lying* activity, which are classified with a

true positive rate of 54.5% and 99.9%, respectively. In the case of sitting and standing activities, PB and PT positions obtained a very high true positive rate of more than 95%, which depicts their easier recognition as compared to other phone positions. Likewise, PP is truly recognized with a more than 95% rate based on standing activity. The recognition of PB and PT positions based on walking activity attained a true positive rate of less than 60%, which shows inferring these phone positions based on in-the-wild gait patterns is very challenging. On the other hand, the recognition of PP based on walking activity is more comfortable, which achieved a true positive rate of 89.4%. In general, the proposed scheme achieves satisfactory performance for activity-aware PCR.

Table 7. Average results for phone context recognition (PCR) based on four physical activities using smartphone accelerometer.

Classifier	Activity	Accuracy	Precision	Sensitivity	Micro-F1	Macro-F1	Log Loss
BDT	Lying	0.996	0.923	0.772	0.996	0.831	1.420
	Sitting	0.982	0.917	0.904	0.964	0.911	0.426
	Walking	0.889	0.753	0.684	0.777	0.711	1.731
	Standing	0.991	0.975	0.974	0.982	0.974	0.076
NN	Lying	0.993	0.497	0.500	0.993	0.498	2.505
	Sitting	0.916	0.473	0.338	0.831	0.345	1.578
	Walking	0.776	0.427	0.336	0.552	0.313	1.356
	Standing	0.861	0.697	0.720	0.721	0.698	1.352

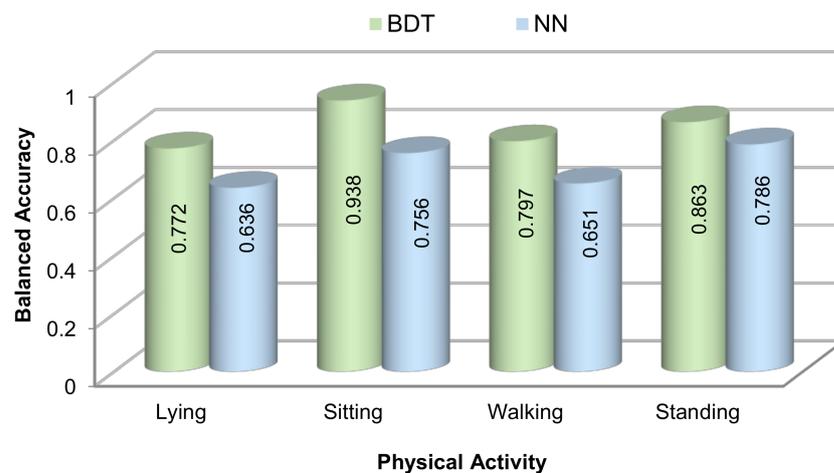


Figure 5. Comparison of balanced accuracy obtained for PCR based on four individual physical activities using BDT and NN classifiers.

4.4. Analysis of BDT vs. NN for Proposed ARW Scheme

As indicated by the results presented and discussed in the previous sections, the performance of the BDT is better for all types of recognition experiments (i.e., PPAR, BCR, and PCR). In contrast, the NN fails to provide satisfactory results for the proposed scheme. The best-case average results (i.e., BALACC values) achieved for PPAR, BCR, and PCR experiments using the BDT classifier are 10.8%, 11%, and 13.5% more than NN results. Generally, the NN classifier fits well for AR tasks, and numerous research studies have successfully utilized different variants of NNs (i.e., deep NN, convolutional NN, and recurrent NN) for AR [96–98]. However, based on the underlying data distribution, there can be certain bottlenecks in achieving effective performance for sensor-based AR-related tasks using a NN. For instance, a massive bulk of data is required for efficient training of NNs to avoid any underfitting/overfitting or regularization issue. When dealing with imbalanced class data (such as the case with our proposed scheme), where the number of samples for certain individual classes is very small, the NN classifier performs below

par due to a lack of training samples. In addition, training or labeling noise, data standardization/normalization, cross-validation strategy, poor hyperparameter optimization strategies, and poor selection of the number of hidden layers and number of nodes in each hidden layer also degrade the performance of the NN. These factors consequently lead to performance degradation of the NN classifier for the proposed scheme. In contrast to the NN, the BDT classifier works well with the smaller datasets by utilizing a combination of multiple decision trees to minimize the prediction error. The trees are connected in a sequential order, where each tree makes up for the prediction error of the preceding trees to boost the overall recognition performance. The final result is based on an ensemble of all decision trees, which may lead to overfitting problems in some cases. However, the final (i.e., best-case) recognition performance of our proposed scheme determines the efficacy of the BDT classifier for ARW, thus making it an optimal choice for such types of experiments. For handling imbalanced class data, resampling techniques, such as the synthetic minority oversampling technique (SMOTE) [99], can be used to achieve good results with classifiers that require a large amount of training data for each class (e.g., NN).

Table 8. Confusion matrices obtained for PCR based on four individual physical activities using BDT classifier.

		Predicted Output				Predicted Output			
Lying (A1)		PH	PT	Sitting (A2)		PB	PH	PP	PT
Ground Truth	PH	54.5%	45.5%	Ground Truth	PB	96.1%	1.5%	0.1%	2.3%
	PT	0.1%	99.9%		PH	1.2%	80.6%	1.6%	16.5%
					PP	1.3%	5.5%	86.6%	6.6%
				PT	0.1%	1.0%	0.4%	98.5%	

		Predicted Output				Predicted Output					
Walking (A3)		PB	PH	PP	PT	Standing (A4)		PB	PH	PP	PT
Ground Truth	PB	59.3%	16.2%	22.7%	1.8%	Ground Truth	PB	99.1%	0.5%	0.0%	0.5%
	PH	7.6%	73.6%	17.7%	1.2%		PH	0.7%	93.2%	1.4%	4.8%
	PP	2.8%	6.5%	89.4%	1.4%		PP	0.0%	0.4%	98.6%	1.0%
	PT	5.6%	14.6%	28.3%	51.5%		PT	0.0%	0.7%	0.6%	98.8%

Note: **PB**: phone in bag; **PH**: phone in hand; **PP**: phone in pocket; **PT**: phone on table.

4.5. Performance Comparison with Existing AR Schemes

Table 9 demonstrates the primary characteristics of some well-known state-of-the-art AR studies and compares them with our proposed ARW method. The comparison is made in terms of activity type, the number of activities recognized, activity occupancy and environment/context, sensing modalities for data acquisition, and machine learning classifiers for AR. It can be investigated from the table that most of the existing AR studies (such as [100–103]) emphasize the recognition of simple (or atomic) daily living activities in certain restricted settings or environments. The occupancy for collecting participants’ data during activity execution generally follows a single location, such as a laboratory or home, where the sensing equipment is installed or carried out to record the participant’s data. The activities to be recognized by the system are performed in a predefined way as scripted tasks. Moreover, there is a lack of diversity in activity-related contexts. As a result, it becomes easier for existing studies to achieve efficient AR performance. However, these schemes fail to adapt to natural user behavior, which is indispensable for real-time applications in diverse environments. Only a few AR schemes (such as [37,104]) worked on learning and identifying natural user activities in indoor and outdoor environments. The authors in [38] recognized diverse single-label human contexts in-the-wild using heterogeneous sensor data from smartphones and smartwatches. However, single-label activity/context information is not adequate for fine-grained AR. Our proposed ARW scheme offers multi-

label activity and context recognition by aggregating outputs from different stages, such as PPAR, BCR, and PCR, and achieves state-of-the-art results in terms of BALACC. In comparison to most of the existing AR studies presented in Table 9, the proposed method demonstrates efficient recognition of six PADLs, ten behavioral contexts, and four phone contexts in-the-wild. Furthermore, the proposed scheme is computationally beneficial and low-cost as it simply depends on the smartphone and smartwatch accelerometer data for recognition. Henceforth, the efficacy of the proposed ARW scheme is justified over state-of-the-art AR schemes.

Table 9. Comparison of the proposed ARW scheme with previous studies.

Study	Activity/ Context Type	No. of Activities/ Contexts	Occupancy/ Environment	Sensing Device/ Sensors	Classifier(s)	Achieved Results
[100]	Daily Living	06	Single/ Controlled Lab	Smartphone (Acc.)	MLP, LR, DT (Decision-level Fusion)	F1-Score = 91.8%
[104]	Daily Living	06	Single/ Controlled Lab	Smartphone (Acc.)	CNN	F1-Score = 97.4%
	Daily Living	07	Multiple/ Indoor and Outdoor	Smartphone (Acc., Gyro.)		F1-Score = 93.1%
[101]	Daily Living	12	Single/-	Smartphone (Acc., Gyro.)	NN, SVM, DBN	Accuracy = 89.61% (DBN)
[102]	Home Task	07	Single/Indoor	Smartphone (Acc., Mic.); Wearable (Acc.)	RF	Accuracy = 94.1%
[37]	Daily Living	09	Multiple/ Indoor and Outdoor	Smartphone (Acc., Gyro., Mag.); Pressure Sensor	DT, NB, SVM, MLP	Accuracy = 92.8% (MLP)
[38]	Human Behavioral Contexts	25	Multiple/In-the-Wild	Smartphone (Acc., Gyro., Mag., GPS); Wearable (Acc.)	LR	BALACC = 80%
[103]	Home Tasks	10	Single/Smart Home	Motion Sensor; Ambient (Temp. Sensor)	NN, HMM, CRF, SVM, CE (using Genetic Algorithm)	F1-Score = 90.1% (CE)
	Home Tasks	11	Single/Smart Home	Motion Sensor; Item; EU; Ambient (Door Sensor, Temp. Sensor, Light Sensor);		F1-Score = 81.9% (CE)
	Home Tasks	15	Single/Smart Home	Motion Sensor; Ambient (Door Sensor, Temp. Sensor);		F1-Score = 85.7% (CE)
[105]	Daily Living and Home Tasks	12	Single/-	Wearable (Acc.)	DT, SVM (Two-level Fusion)	F1-Score = 93.0% (CE)
[106]	Elderly Activities	17	Single/Smart Home	Wearable (Bar., Temp., Acc., Gyro., Mag.); Ambient (PIR)	SVM	Accuracy = 98.32%
Proposed ARW	Daily Living	06	Multiple/In-the-Wild	Smartphone (Acc.); Wearable (Acc.)	BDT, NN	BALACC = 93.1% (BDT)
	Behavioral Contexts	10	Multiple/In-the-Wild	Smartphone (Acc.); Wearable (Acc.)	BDT, NN	BALACC = 91% (BDT)
	Phone Contexts	04	Multiple/In-the-Wild	Smartphone (Acc.);	BDT, NN	BALACC = 84.2% (BDT)

Note: Acc.: accelerometer, BALACC: balanced accuracy, Bar.: barometer, BDT: boosted decision tree, CE: classifier ensemble, CNN: convolutional neural network, CRF: conditional random fields; DBN: deep belief network, DT: decision tree, EU: electricity usage, GPS: global positioning system, Gyro.: gyroscope, HMM: hidden Markov model, LR: logistic regression, Mag.: magnetometer, Mic.: microphone, MLP: multilayer perceptron, NB: naïve Bayes, NN: neural network, PIR: passive infrared sensor, SVM: support vector machine, Temp.: temperature sensor.

5. Conclusions

This research paper demonstrates a novel two-stage model for sensor-based activity recognition in-the-wild. In the first stage, the proposed scheme classifies six (06) primary physical activities, whereas in the second stage, the proposed scheme infers fourteen (14) activity-aware contexts using the “ExtraSensory” dataset. The outputs from both stages are combined for better cognition and understanding of natural human activities in diverse contexts. Three types of experiments are conducted in this paper, including primary

physical activity recognition, behavioral context/environment recognition, and phone context recognition. Smartphone and smartwatch accelerometers are utilized to identify daily living human activities and the associated behavioral contexts. In contrast, phone context recognition only entails a smartphone accelerometer sensor. A boosted decision tree achieves the best experimental results for the proposed scheme. Although the proposed method achieves a reasonable accuracy rate, there are some limitations associated with it. For example, the activities and behavioral contexts considered for experimentation cannot generalize to all use cases in the real world. The proposed scheme thus cannot handle unforeseen activities and contexts. There are some privacy issues with continuous activity/context monitoring of a human being, particularly if an impostor gets access to the device data/output. The continuous monitoring of human beings using smart devices has memory and battery constraints as well.

The limitations of this paper can be improved in future works. In this aspect, our proposed method can be extended to incorporate more sensing modalities for robust detection/recognition of a large number of human activities and contexts, which can be helpful for human-environment interaction modeling. Resampling and data augmentation techniques can be applied to cope with imbalanced class data, particularly for activities/contexts that exist less in-the-wild daily. Likewise, the proposed scheme can be modified to handle unforeseen activities and contexts. The coinciding recognition of a person's physical activity and behavioral/social context can be crucial for human behavior modeling and cognition in their living environments. Thus, the proposed scheme can also be extended to detect/recognize normal and abnormal human behavior for predicting health-related risks. Knowledge-based systems, focusing on human-centered computing, can utilize the proposed method for improved decision-making and recommendations. The correlation between human daily living activities and their social/behavioral contexts can be examined in diverse environments to realize the factors giving rise to abnormal behavior.

Author Contributions: Conceptualization, M.E.-u.-H. and M.A.A.; methodology, M.E.-u.-H. and F.M.; software, M.E.-u.-H. and F.M.; validation, M.A.A. and Y.A.; formal analysis, M.E.-u.-H., F.M., and M.A.A.; investigation, M.E.-u.-H. and F.M.; resources, M.A.A.; data curation, F.M.; writing—original draft preparation, M.E.-u.-H. and F.M.; writing-review and editing, M.A.A. and Y.A.; visualization, Y.A.; supervision, M.A.A. and Y.A.; project administration, M.A.A. and Y.A.; funding acquisition, M.A.A. All authors have read and agreed to the published version of the manuscript.

Funding: This research work is funded by the School of Information Technology, Whitecliffe, Wellington, New Zealand.

Data Availability Statement: The dataset used for validating this research study is publicly available as an "ExtraSensory" dataset that is cited in the paper.

Acknowledgments: This research work is supported by the School of Information Technology, Whitecliffe, Wellington, New Zealand, and Air University, Islamabad, Pakistan.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Liang, Y.; Zhou, X.; Guo, B.; Yu, Z. Activity recognition using ubiquitous sensors: An overview. *Wearable Technol. Concepts Methodol. Tools Appl.* **2018**, *199–230*. [[CrossRef](#)]
2. Roggen, D.; Troster, G.; Lukowicz, P.; Ferscha, A.; Millan, J.D.R.; Chavarriaga, R. Opportunistic human activity and context recognition. *Computer* **2012**, *46*, 36–45. [[CrossRef](#)]
3. Cao, J.; Lin, M.; Wang, H.; Fang, J.; Xu, Y. Towards Activity Recognition through Multidimensional Mobile Data Fusion with a Smartphone and Deep Learning. *Mob. Inf. Syst.* **2021**, *2021*, 1–11. [[CrossRef](#)]
4. Abdallah, Z.; Gaber, M.; Srinivasan, B.; Krishnaswamy, S. Activity Recognition with Evolving Data Streams. *ACM Comput. Surv.* **2018**, *51*, 3158645. [[CrossRef](#)]
5. Murtaza, F.; Yousaf, M.H.; Velastin, S.A.; Qian, Y. Vectors of temporally correlated snippets for temporal action detection. *Comput. Electr. Eng.* **2020**, *85*, 106654. [[CrossRef](#)]
6. Wang, P.; Li, W.; Ogunbona, P.; Wan, J.; Escalera, S. RGB-D-based human motion recognition with deep learning: A survey. *Comput. Vis. Image Underst.* **2018**, *171*, 118–139. [[CrossRef](#)]

7. Zhang, S.; Wei, Z.; Nie, J.; Shuang, W.; Wang, S.; Li, Z. A Review on Human Activity Recognition Using Vision-Based Method. *J. Healthc. Eng.* **2017**, *2017*, 1–31. [\[CrossRef\]](#)
8. Sarabu, A.; Santra, A.K. Human Action Recognition in Videos using Convolution Long Short-Term Memory Network with Spatio-Temporal Networks. *Emerg. Sci. J.* **2021**, *5*, 25–33. [\[CrossRef\]](#)
9. Aguilera, A.A.; Brena, R.F.; Mayora, O.; Molino-Minero-Re, E.; Trejo, L.A. Multi-Sensor Fusion for Activity Recognition—A Survey. *Sensors* **2019**, *19*, 3808. [\[CrossRef\]](#) [\[PubMed\]](#)
10. Xu, Y.; Shen, Z.; Zhang, X.; Gao, Y.; Deng, S.; Wang, Y.; Fan, Y.; Chang, E.I.-C. Learning multi-level features for sensor-based human action recognition. *Pervasive Mob. Comput.* **2017**, *40*, 324–338. [\[CrossRef\]](#)
11. Alshammari, T.; Alshammari, N.; Sedky, M.; Howard, C. SIMADL: Simulated Activities of Daily Living Dataset. *Data* **2018**, *3*, 11. [\[CrossRef\]](#)
12. Alsinglawi, B.; Nguyen, Q.V.; Gunawardana, U.; Maeder, A.; Simoff, S. RFID Systems in Healthcare Settings and Activity of Daily Living in Smart Homes: A Review. *E-Health Telecommun. Syst. Netw.* **2017**, *6*, 1–17. [\[CrossRef\]](#)
13. Vanus, J.; Belesova, J.; Martinek, R.; Nedoma, J.; Fajkus, M.; Bilik, P.; Zidek, J. Monitoring of the daily living activities in smart home care. *Hum. Cent. Comput. Inf. Sci.* **2017**, *7*, 30. [\[CrossRef\]](#)
14. Marques, B.; McIntosh, J.; Valera, A.; Gaddam, A. Innovative and Assistive eHealth Technologies for Smart Therapeutic and Rehabilitation Outdoor Spaces for the Elderly Demographic. *Multimodal Technol. Interact.* **2020**, *4*, 76. [\[CrossRef\]](#)
15. Zhu, Z.; Liu, T.; Li, G.; Li, T.; Inoue, Y. Wearable Sensor Systems for Infants. *Sensors* **2015**, *15*, 3721–3749. [\[CrossRef\]](#)
16. Kristoffersson, A.; Lindén, M. A Systematic Review on the Use of Wearable Body Sensors for Health Monitoring: A Qualitative Synthesis. *Sensors* **2020**, *20*, 1502. [\[CrossRef\]](#)
17. Mukhopadhyay, S.C. Wearable Sensors for Human Activity Monitoring: A Review. *IEEE Sens. J.* **2014**, *15*, 1321–1330. [\[CrossRef\]](#)
18. Schrack, J.A.; Cooper, R.; Koster, A.; Shiroma, E.J.; Murabito, J.M.; Rejeski, W.J.; Ferrucci, L.; Harris, T.B. Assessing Daily Physical Activity in Older Adults: Unraveling the Complexity of Monitors, Measures, and Methods. *J. Gerontol. Ser. A Biomed. Sci. Med. Sci.* **2016**, *71*, 1039–1048. [\[CrossRef\]](#)
19. Ignatov, A. Real-time human activity recognition from accelerometer data using Convolutional Neural Networks. *Appl. Soft Comput.* **2018**, *62*, 915–922. [\[CrossRef\]](#)
20. Hussain, F.; Hussain, F.; Ehatisham-ul-Haq, M.; Azam, M.A. Activity-Aware Fall Detection and Recognition based on Wearable Sensors. *IEEE Sens. J.* **2019**, *19*, 4528–4536. [\[CrossRef\]](#)
21. Xu, L.; Yang, W.; Cao, Y.; Li, Q. Human activity recognition based on random forests. In Proceedings of the 2017 13th International Conference on Natural Computation, Fuzzy Systems and Knowledge Discovery (ICNC-FSKD), Guilin, China, 29–31 July 2017; pp. 548–553. [\[CrossRef\]](#)
22. Fu, Z.; He, X.; Wang, E.; Huo, J.; Huang, J.; Wu, D. Personalized Human Activity Recognition Based on Integrated Wearable Sensor and Transfer Learning. *Sensors* **2021**, *21*, 885. [\[CrossRef\]](#)
23. Esfahani, M.I.M.; Nussbaum, M.A. Classifying Diverse Physical Activities Using “Smart Garments”. *Sensors* **2019**, *19*, 3133. [\[CrossRef\]](#) [\[PubMed\]](#)
24. Gadaleta, M.; Rossi, M. IDNet: Smartphone-based gait recognition with convolutional neural networks. *Pattern Recognit.* **2018**, *74*, 25–37. [\[CrossRef\]](#)
25. Chen, Z.; Zhang, L.; Cao, Z.; Guo, J. Distilling the Knowledge from Handcrafted Features for Human Activity Recognition. *IEEE Trans. Ind. Inform.* **2018**, *14*, 4334–4342. [\[CrossRef\]](#)
26. Lee, K.; Kwan, M.-P. Physical activity classification in free-living conditions using smartphone accelerometer data and exploration of predicted results. *Comput. Environ. Urban Syst.* **2018**, *67*, 124–131. [\[CrossRef\]](#)
27. Incel, O.D.; Kose, M.; Ersoy, C. A Review and Taxonomy of Activity Recognition on Mobile Phones. *BioNanoScience* **2013**, *3*, 145–171. [\[CrossRef\]](#)
28. Shoaib, M.; Bosch, S.; Incel, O.D.; Scholten, J.; Havinga, P.J. A Survey of Online Activity Recognition Using Mobile Phones. *Sensors* **2015**, *15*, 2059–2085. [\[CrossRef\]](#) [\[PubMed\]](#)
29. Dao, M.-S.; Nguyen-Gia, T.-A.; Mai, V.-C. Daily Human Activities Recognition Using Heterogeneous Sensors from Smartphones. *Procedia Comput. Sci.* **2017**, *111*, 323–328. [\[CrossRef\]](#)
30. Shoaib, M.; Bosch, S.; Incel, O.D.; Scholten, J.; Havinga, P.J.M. Complex Human Activity Recognition Using Smartphone and Wrist-Worn Motion Sensors. *Sensors* **2016**, *16*, 426. [\[CrossRef\]](#)
31. Shoaib, M.; Bosch, S.; Scholten, H.; Havinga, P.J.M.; Incel, O.D. Towards detection of bad habits by fusing smartphone and smartwatch sensors. In Proceedings of the 2015 IEEE International Conference on Pervasive Computing and Communication Workshops, PerCom Workshops, St. Louis, MO, USA, 23–27 March 2015; pp. 591–596.
32. Ranieri, C.M.; MacLeod, S.; Dragone, M.; Vargas, P.A.; Romero, R.A.F. Activity Recognition for Ambient Assisted Living with Videos, Inertial Units and Ambient Sensors. *Sensors* **2021**, *21*, 768. [\[CrossRef\]](#)
33. Cao, L.; Wang, Y.; Zhang, B.; Jin, Q.; Vasilakos, A.V. GCHAR: An efficient Group-based Context—aware human activity recognition on smartphone. *J. Parallel Distrib. Comput.* **2018**, *118*, 67–80. [\[CrossRef\]](#)
34. Otebolaku, A.M.; Andrade, M.T. User context recognition using smartphone sensors and classification models. *J. Netw. Comput. Appl.* **2016**, *66*, 33–51. [\[CrossRef\]](#)

35. Fahim, M.; Khattak, A.M.; Baker, T.; Chow, F.; Shah, B. Micro-context recognition of sedentary behaviour using smartphone. In Proceedings of the 2016 6th International Conference on Digital Information and Communication Technology and Its Applications, DICTAP, Konya, Turkey, 21–23 July 2016; pp. 30–34.
36. Ellis, K.; Godbole, S.; Kerr, J.; Lanckriet, G. Multi-sensor physical activity recognition in free-living. In Proceedings of the ACM International Joint Conference on Pervasive and Ubiquitous Computing Adjunct Publication-UbiComp '14 Adjunct, Seattle, WA, USA, 13–17 September 2014; pp. 431–440.
37. Guiry, J.J.; Van De Ven, P.; Nelson, J. Multi-Sensor Fusion for Enhanced Contextual Awareness of Everyday Activities with Ubiquitous Devices. *Sensors* **2014**, *14*, 5687–5701. [[CrossRef](#)]
38. Vaizman, Y.; Ellis, K.; Lanckriet, G.R.G. Recognizing Detailed Human Context in the Wild from Smartphones and Smartwatches. *IEEE Pervasive Comput.* **2017**, *16*, 62–74. [[CrossRef](#)]
39. Safyan, M.; Sarwar, S.; Qayyum, Z.U.; Iqbal, M.; Li, S.; Kashif, M. Machine Learning based Activity learning for Behavioral Contexts in Internet of Things. *Proc. Inst. Syst. Program. RAS* **2021**, *33*, 47–58. [[CrossRef](#)]
40. Marques, G.; Miranda, N.; Bhoi, A.K.; Garcia-Zapirain, B.; Hamrioui, S.; Díez, I.D.L.T. Internet of Things and Enhanced Living Environments: Measuring and Mapping Air Quality Using Cyber-physical Systems and Mobile Computing Technologies. *Sensors* **2020**, *20*, 720. [[CrossRef](#)] [[PubMed](#)]
41. Rehman, A.; Iqbal, M.; Xing, H.; Ahmed, I. COVID-19 Detection Empowered with Machine Learning and Deep Learning Techniques: A Systematic Review. *Appl. Sci.* **2021**, *11*, 3414. [[CrossRef](#)]
42. Klumpp, M.; Hintze, M.; Immonen, M.; Ródenas-Rigla, F.; Pilati, F.; Aparicio-Martínez, F.; Çelebi, D.; Liebig, T.; Jirstrand, M.; Urbann, O.; et al. Artificial Intelligence for Hospital Health Care: Application Cases and Answers to Challenges in European Hospitals. *Healthcare* **2021**, *9*, 961. [[CrossRef](#)]
43. Massaro, A.; Maritati, V.; Savino, N.; Galiano, A. Neural Networks for Automated Smart Health Platforms oriented on Heart Predictive Diagnostic Big Data Systems. In Proceedings of the 2018 AEIT International Annual Conference, Bari, Italy, 3–5 October 2018. [[CrossRef](#)]
44. Fahad, L.G.; Ali, A.; Rajarajan, M. Learning models for activity recognition in smart homes. In *Information Science and Applications*; Springer: Berlin/Heidelberg, Germany, 2015; pp. 819–826. [[CrossRef](#)]
45. Lu, L.; Qing-Ling, C.; Yi-Ju, Z. Activity Recognition in Smart Homes. *Multimed. Tools Appl.* **2016**, *76*, 24203–24220. [[CrossRef](#)]
46. Chahuara, P.; Fleury, A.; Vacher, M.; Chahuara, P.; Fleury, A.; Vacher, M.; Activity, O.H. On-line Human Activity Recognition from Audio and Home Automation Sensors. *J. Ambient. Intell. Smart Environ.* **2016**, *8*, 399–422. [[CrossRef](#)]
47. Ni, Q.; Hernando, A.B.G.; DE LA Cruz, I.P. A Context-Aware System Infrastructure for Monitoring Activities of Daily Living in Smart Home. *J. Sens.* **2016**, *2016*, 9493047. [[CrossRef](#)]
48. Ghayvat, H.; Mukhopadhyay, S.; Shenjie, B.; Chouhan, A.; Chen, W. Smart home based ambient assisted living: Recognition of anomaly in the activity of daily living for an elderly living alone. In Proceedings of the I2MTC 2018-2018 IEEE International Instrumentation and Measurement Technology Conference: Discovering New Horizons in Instrumentation and Measurement, Houston, TX, USA, 14–17 May 2018; pp. 1–5.
49. Muheidat, F.; Tawalbeh, L.; Tyrer, H. Context-Aware, Accurate, and Real Time Fall Detection System for Elderly People. In Proceedings of the 12th IEEE International Conference on Semantic Computing, ICSC, Laguna Hills, CA, USA, 31 January–2 February 2018; pp. 329–333.
50. Esfahani, M.I.M.; Nussbaum, M.A. Preferred Placement and Usability of a Smart Textile System vs. Inertial Measurement Units for Activity Monitoring. *Sensors* **2018**, *18*, 2501. [[CrossRef](#)]
51. Cleland, I.; Kikhia, B.; Nugent, C.; Boytsov, A.; Hallberg, J.; Synnes, K.; McClean, S.; Finlay, D. Optimal Placement of Accelerometers for the Detection of Everyday Activities. *Sensors* **2013**, *13*, 9183–9200. [[CrossRef](#)]
52. Boerema, S.T.; Van Velsen, L.; Schaake, L.; Tönis, T.M.; Hermens, H.J. Optimal Sensor Placement for Measuring Physical Activity with a 3D Accelerometer. *Sensors* **2014**, *14*, 3188–3206. [[CrossRef](#)]
53. Özdemir, A.T. An Analysis on Sensor Locations of the Human Body for Wearable Fall Detection Devices: Principles and Practice. *Sensors* **2016**, *16*, 1161. [[CrossRef](#)]
54. Martinez-Hernandez, U.; Dehghani-Sanij, A.A. Probabilistic identification of sit-to-stand and stand-to-sit with a wearable sensor. *Pattern Recognit. Lett.* **2018**, *118*, 32–41. [[CrossRef](#)]
55. Mehrang, S.; Pietilä, J.; Korhonen, I. An Activity Recognition Framework Deploying the Random Forest Classifier and A Single Optical Heart Rate Monitoring and Triaxial Accelerometer Wrist-Band. *Sensors* **2018**, *18*, 613. [[CrossRef](#)] [[PubMed](#)]
56. Bharti, P.; De, D.; Chellappan, S.; Das, S.K. HuMAN: Complex Activity Recognition with Multi-Modal Multi-Positional Body Sensing. *IEEE Trans. Mob. Comput.* **2018**, *18*, 857–870. [[CrossRef](#)]
57. Anwary, A.R.; Yu, H.; Vassallo, M. Gait Evaluation Using Procrustes and Euclidean Distance Matrix Analysis. *IEEE J. Biomed. Heal. Inform.* **2018**, *23*, 2021–2029. [[CrossRef](#)]
58. Russell, B.; McDaid, A.; Toscano, W.; Hume, P. Moving the Lab into the Mountains: A Pilot Study of Human Activity Recognition in Unstructured Environments. *Sensors* **2021**, *21*, 654. [[CrossRef](#)] [[PubMed](#)]
59. Antos, S.A.; Albert, M.V.; Kording, K. Hand, belt, pocket or bag: Practical activity tracking with mobile phones. *J. Neurosci. Methods* **2013**, *231*, 22–30. [[CrossRef](#)]
60. Khan, A.M.; Siddiqi, M.H.; Lee, S.-W. Exploratory Data Analysis of Acceleration Signals to Select Light-Weight and Accurate Features for Real-Time Activity Recognition on Smartphones. *Sensors* **2013**, *13*, 13099–13122. [[CrossRef](#)] [[PubMed](#)]

61. Shoaib, M.; Bosch, S.; Incel, O.D.; Scholten, J.; Havinga, P.J.M. Fusion of Smartphone Motion Sensors for Physical Activity Recognition. *Sensors* **2014**, *14*, 10146–10176. [[CrossRef](#)] [[PubMed](#)]
62. Shi, D.; Wang, R.; Wu, Y.; Mo, X.; Wei, J. A novel orientation- and location-independent activity recognition method. *Pers. Ubiquitous Comput.* **2017**, *21*, 427–441. [[CrossRef](#)]
63. Martin, H.; Bernardos, A.M.; Iglesias, J.; Casar, J.R. Activity logging using lightweight classification techniques in mobile devices. *Pers. Ubiquitous Comput.* **2012**, *17*, 675–695. [[CrossRef](#)]
64. Coskun, D.; Incel, O.D.; Ozgovde, A. Phone position/placement detection using accelerometer: Impact on activity recognition. In Proceedings of the 2015 IEEE 10th International Conference on Intelligent Sensors, Sensor Networks and Information Processing, Singapore, 7–9 April 2015.
65. Nalepa, G.J.; Kutt, K.; Bobek, S. Mobile platform for affective context-aware systems. *Futur. Gener. Comput. Syst.* **2019**, *92*, 490–503. [[CrossRef](#)]
66. Hoseini-Tabatabaei, S.A.; Gluhak, A.; Tafazolli, R. A survey on smartphone-based systems for opportunistic user context recognition. *ACM Comput. Surv.* **2013**, *45*, 1–51. [[CrossRef](#)]
67. Esfahani, P.; Malazi, H.T. PAMS: A new position-aware multi-sensor dataset for human activity recognition using smartphones. In Proceedings of the 2017 19th International Symposium on Computer Architecture and Digital Systems, CADSD, Kish Island, Iran, 21–22 December 2017; pp. 1–7.
68. Kohavi, R. Scaling Up the Accuracy of Naive-Bayes Classifiers: A Decision-Tree Hybrid. *Proc. Second Int. Conf. Knowl. Discov. Data Min.* **1996**, *7*, 202–207.
69. Lim, H.; An, G.; Cho, Y.; Lee, K.; Suh, B. WhichHand: Automatic Recognition of a Smartphone's Position in the Hand Using a Smartwatch. In Proceedings of the 18th International Conference on Human-Computer Interaction with Mobile Devices and Services Adjunct, Florence, Italy, 6–9 September 2016; pp. 675–681.
70. Yang, Z.; Shanguan, L.; Gu, W.; Zhou, Z.; Wu, C.; Liu, Y. Sherlock: Micro-Environment Sensing for Smartphones. *IEEE Trans. Parallel Distrib. Syst.* **2014**, *25*, 3295–3305. [[CrossRef](#)]
71. Li, X.; Goldberg, D.W. Toward a mobile crowdsensing system for road surface assessment. *Comput. Environ. Urban Syst.* **2018**, *69*, 51–62. [[CrossRef](#)]
72. Wang, A.; Chen, G.; Yang, J.; Zhao, S.; Chang, C.-Y. A Comparative Study on Human Activity Recognition Using Inertial Sensors in a Smartphone. *IEEE Sens. J.* **2016**, *16*, 4566–4578. [[CrossRef](#)]
73. Costa, A.A.M.; Almeida, H.; Lorayne, A.; de Sousa, R.R.; Perkusich, A.; Ramos, F.B.A. Combining Smartphone and Smartwatch Sensor Data in Activity Recognition Approaches: An Experimental Evaluation. In Proceedings of the 28th International Conference on Software Engineering and Knowledge Engineering, Redwood City, CA, USA, 1–3 July 2016; pp. 267–272.
74. Nweke, H.F.; Teh, Y.W.; Al-Garadi, M.A.; Alo, U.R. Deep learning algorithms for human activity recognition using mobile and wearable sensor networks: State of the art and research challenges. *Expert Syst. Appl.* **2018**, *105*, 233–261. [[CrossRef](#)]
75. Monteiro, J.; Granada, R.; Barros, R.C.; Meneguzzi, F. Deep neural networks for kitchen activity recognition. In Proceedings of the International Joint Conference on Neural Networks, Anchorage, AK, USA, 14–19 May 2017; pp. 2048–2055.
76. Mekruksavanich, S.; Jitpattanakul, A. LSTM Networks Using Smartphone Data for Sensor-Based Human Activity Recognition in Smart Homes. *Sensors* **2021**, *21*, 1636. [[CrossRef](#)]
77. Ramos, R.; Domingo, J.; Zalama, E.; Gómez-García-Bermejo, J. Daily Human Activity Recognition Using Non-Intrusive Sensors. *Sensors* **2021**, *21*, 5270. [[CrossRef](#)]
78. Nafea, O.; Abdul, W.; Muhammad, G.; Alsulaiman, M. Sensor-Based Human Activity Recognition with Spatio-Temporal Deep Learning. *Sensors* **2021**, *21*, 2141. [[CrossRef](#)]
79. Wang, J.; Chen, Y.; Hao, S.; Peng, X.; Hu, L. Deep learning for sensor-based activity recognition: A survey. *Pattern Recognit. Lett.* **2019**, *119*, 3–11. [[CrossRef](#)]
80. Hammerla, N.Y.; Halloran, S.; Plötz, T. Deep, convolutional, and recurrent models for human activity recognition using wearables. In Proceedings of the IJCAI International Joint Conference on Artificial Intelligence, New York, NY, USA, 9–15 July 2016; pp. 1533–1540.
81. Anguita, D.; Ghio, A.; Oneto, L.; Parra, X.; Reyes-Ortiz, J.L. A Public Domain Dataset for Human Activity Recognition Using Smartphones. In Proceedings of the 21th European Symposium on Artificial Neural Networks, Computational Intelligence and Machine Learning, ESANN, Bruges, Belgium, 24–26 April 2013.
82. Sucerquia, A.; López, J.D.; Vargas-Bonilla, J.F. SisFall: A Fall and Movement Dataset. *Sensors* **2017**, *17*, 198. [[CrossRef](#)]
83. Vavoulas, G.; Chatzaki, C.; Malliotakis, T.; Padiaditis, M.; Tsiknakis, M. The MobiAct Dataset: Recognition of Activities of Daily Living using Smartphones. In Proceedings of the International Conference on Information and Communication Technologies for Ageing Well and e-Health, Rome, Italy, 21–22 April 2016; pp. 143–151.
84. Chen, C.; Jafari, R.; Kehtarnavaz, N. UTD-MHAD: A multimodal dataset for human action recognition utilizing a depth camera and a wearable inertial sensor. In Proceedings of the 2015 IEEE International Conference on Image Processing (ICIP), Quebec City, QC, Canada, 27–30 September 2015; pp. 168–172. [[CrossRef](#)]
85. Liu, H.; Hartmann, Y.; Schultz, T. CSL-SHARE: A Multimodal Wearable Sensor-Based Human Activity Dataset. *Front. Comput. Sci.* **2021**, *3*, 759136. [[CrossRef](#)]
86. Lima, W.S.; Souto, E.; El-Khatib, K.; Jalali, R.; Gama, J. Human Activity Recognition Using Inertial Sensors in a Smartphone: An Overview. *Sensors* **2019**, *19*, 3213. [[CrossRef](#)] [[PubMed](#)]

87. Vaizman, Y.; Weibel, N. Context Recognition In-the-Wild: Unified Model for Multi-Modal Sensors and Multi-Label Classification. *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.* **2017**, *168*, 22. [[CrossRef](#)]
88. Ehatisham-Ul-Haq, M.; Azam, M.A.; Loo, J.; Shuang, K.; Islam, S.; Naeem, U.; Amin, Y. Authentication of Smartphone Users Based on Activity Recognition and Mobile Sensing. *Sensors* **2017**, *17*, 2043. [[CrossRef](#)]
89. Ehatisham-Ul-Haq, M.; Azam, M.A.; Naeem, U.; Amin, Y.; Loo, J. Continuous authentication of smartphone users based on activity pattern recognition using passive mobile sensing. *J. Netw. Comput. Appl.* **2018**, *109*, 24–35. [[CrossRef](#)]
90. Hall, M.A.; Smith, L.A. Feature subset selection: A correlation based filter approach. In *Progress in Connectionist-Based Information Systems*; Kasabov, N., Kozma, R., Ko, K., O'Shea, R., Coghill, G., Gedeon, T., Eds.; Springer: Berlin/Heidelberg, Germany, 1997; pp. 855–858.
91. Jerome, H. Friedman Greedy Function Approximation: A Gradient Boosting Machine. *Ann. Stat.* **2000**, *29*, 1189–1232.
92. Kothari, S.; Oh, H. Neural Networks for Pattern Recognition. In *Advances in Computers*; Elsevier: Amsterdam, The Netherlands, 1993; pp. 119–166. [[CrossRef](#)]
93. Barga, R.; Fontama, V.; Tok, W.H. *Predictive Analytics with Microsoft Azure Machine Learning*; Springer: Berlin/Heidelberg, Germany, 2015.
94. Dernbach, S.; Das, B.; Krishnan, N.C.; Thomas, B.L.; Cook, D.J. Simple and Complex Activity Recognition through Smart Phones. In Proceedings of the 2012 8th International Conference on Intelligent Environments (IE), Guanajuato, Mexico, 26–29 June 2012; pp. 214–221.
95. Rifkin, R.; Klautau, A. In defense of one-vs-all classification. *J. Mach. Learn. Res.* **2004**, *5*, 101–141.
96. San-Segundo, R.; Blunck, H.; Moreno-Pimentel, J.; Stisen, A.; Gil-Martín, M. Robust Human Activity Recognition using smartwatches and smartphones. *Eng. Appl. Artif. Intell.* **2018**, *72*, 190–202. [[CrossRef](#)]
97. Oniga, S.; Süto, J. Human activity recognition using neural networks. In Proceedings of the 2014 15th International Carpathian Control Conference, ICCCC, Velke Karlovice, Czech Republic, 28–30 May 2014; pp. 403–406.
98. Yao, S.; Hu, S.; Zhao, Y.; Zhang, A.; Abdelzaher, T. DeepSense: A unified deep learning framework for time-series mobile sensing data processing. In Proceedings of the 26th International World Wide Web Conference, WWW, Perth, Australia, 3–7 April 2017; pp. 351–360.
99. Chawla, N.V.; Bowyer, K.W.; Hall, L.O.; Kegelmeyer, W.P. SMOTE: Synthetic minority over-sampling technique. *J. Artif. Intell. Res.* **2002**, *16*, 321–357.
100. Catal, C.; Tufekci, S.; Pirit, E.; Kocabag, G. On the use of ensemble of classifiers for accelerometer-based activity recognition. *Appl. Soft Comput.* **2015**, *37*, 1018–1022. [[CrossRef](#)]
101. Hassan, M.M.; Uddin, Z.; Mohamed, A.; Almogren, A. A robust human activity recognition system using smartphone sensors and deep learning. *Futur. Gener. Comput. Syst.* **2018**, *81*, 307–313. [[CrossRef](#)]
102. Garcia-Ceja, E.; Galván-Tejada, C.E.; Brena, R. Multi-view stacking for activity recognition with sound and accelerometer data. *Inf. Fusion* **2018**, *40*, 45–56. [[CrossRef](#)]
103. Fatima, I.; Fahim, M.; Lee, Y.-K.; Lee, S. A Genetic Algorithm-based Classifier Ensemble Optimization for Activity Recognition in Smart Homes. *KSII Trans. Internet Inf. Syst.* **2013**, *7*, 2853–2873. [[CrossRef](#)]
104. Ravi, D.; Wong, C.; Lo, B.; Yang, G.-Z. A Deep Learning Approach to on-Node Sensor Data Analytics for Mobile or Wearable Devices. *IEEE J. Biomed. Heal. Inform.* **2016**, *21*, 56–64. [[CrossRef](#)] [[PubMed](#)]
105. Aly, H.; Ismail, M.A. UbiMonitor: Intelligent fusion of body-worn sensors for real-time human activity recognition. In Proceedings of the ACM Symposium on Applied Computing, Salamanca, Spain, 13–17 April 2015; pp. 563–568.
106. Wang, Y.; Cang, S.; Yu, H. A Data Fusion-Based Hybrid Sensory System for Older People's Daily Activity and Daily Routine Recognition. *IEEE Sens. J.* **2018**, *18*, 6874–6888. [[CrossRef](#)]