





Article

Human Activity Recognition Based on Embedded Sensor Data Fusion for the Internet of Healthcare Things

Mohamed E. Issa¹, Ahmed M. Helmi^{1,2} , Mohammed A. A. Al-Qaness^{3,4,*} , Abdelghani Dahou⁵,
Mohamed Abd Elaziz^{6,7,8}  and Robertas Damaševičius^{9,*} 

- ¹ Computer and Systems Engineering Department, Faculty of Engineering, Zagazig University, Zagazig 44519, Egypt; Mamohamedali@eng.zu.edu.eg (M.E.I.); amhm162@gmail.com (A.M.H.)
- ² College of Engineering and Information Technology, Buraydah Private Colleges, Buraydah 51418, Saudi Arabia
- ³ State Key Laboratory for Information Engineering in Surveying, Mapping and Remote Sensing, Wuhan University, Wuhan 430079, China
- ⁴ Faculty of Engineering, Sana'a University, Sana'a 12544, Yemen
- ⁵ LDDI Laboratory, Faculty of Science and Technology, University of Ahmed DRAIA, Adrar 01000, Algeria; dahou.abdghani@univ-adrar.edu.dz
- ⁶ Faculty of Computer Science and Engineering, Galala University, Suez 435611, Egypt; abd_el_aziz_m@yahoo.com
- ⁷ Artificial Intelligence Research Center (AIRC), College of Engineering and Information Technology, Ajman University, Ajman 346, United Arab Emirates
- ⁸ Department of Mathematics, Faculty of Science, Zagazig University, Zagazig 44519, Egypt
- ⁹ Department of Applied Informatics, Vytautas Magnus University, 44404 Kaunas, Lithuania
- * Correspondence: alqaness@whu.edu.cn (M.A.A.A.-Q.); robertas.damasevicius@vdu.lt (R.D.)



Citation: Issa, M.E.; Helmi, A.M.; Al-Qaness, M.A.A.; Dahou, A.; Abd Elaziz, M.; Damaševičius, R. Human Activity Recognition Based on Embedded Sensor Data Fusion for the Internet of Healthcare Things. *Healthcare* **2022**, *10*, 1084. <https://doi.org/10.3390/healthcare10061084>

Academic Editors: Giner Alor-Hernández, Jezreel Mejía-Miranda, José Luis Sánchez-Cervantes and Alejandro Rodríguez-González

Received: 24 May 2022

Accepted: 9 June 2022

Published: 10 June 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/>).

Abstract: Nowadays, the emerging information technologies in smart handheld devices are motivating the research community to make use of embedded sensors in such devices for healthcare purposes. In particular, inertial measurement sensors such as accelerometers and gyroscopes embedded in smartphones and smartwatches can provide sensory data fusion for human activities and gestures. Thus, the concepts of the Internet of Healthcare Things (IoHT) paradigm can be applied to handle such sensory data and maximize the benefits of collecting and analyzing them. The application areas contain but are not restricted to the rehabilitation of elderly people, fall detection, smoking control, sportive exercises, and monitoring of daily life activities. In this work, a public dataset collected using two smartphones (in pocket and wrist positions) is considered for IoHT applications. Three-dimensional inertia signals of thirteen timestamped human activities such as Walking, Walking Upstairs, Walking Downstairs, Writing, Smoking, and others are registered. Here, an efficient human activity recognition (HAR) model is presented based on efficient handcrafted features and Random Forest as a classifier. Simulation results ensure the superiority of the applied model over others introduced in the literature for the same dataset. Moreover, different approaches to evaluating such models are considered, as well as implementation issues. The accuracy of the current model reaches 98.7% on average. The current model performance is also verified using the WISDM v1 dataset.

Keywords: Internet of Healthcare Things; human activity recognition; smart technologies for healthcare; m-Health; mobile devices; digital healthcare

1. Introduction

1.1. Motivation

Smart solutions for Internet of Healthcare Things (IoHT) [1], also known as Healthcare Internet of Things [2], Internet of Medical Things [3], or Medical Internet of Things [4], systems have extensively emerged since the Industry 4.0 revolution [5], making use of digital devices, in particular wearable sensors and smart handheld devices. In the new phase of the industrial revolution, termed Industry 5.0, collaborative interaction between machines and people is coming back to the forefront [6]. Unlike aiming to find the best

ways to connect devices together—in the first place—which was a goal of Industry 4.0, there is great interest in moving toward personalization in Industry 5.0. This means that creative thinking and smart usage of the entities of smart systems are expected to increase the productivity and benefits of emerging IoT-based solutions [7]. The guidelines of Industry 5.0—under the umbrella of IoT—open up a new window to the development and enhancement of existing smart IoHT systems, in particular, during present-day circumstances, such as the spread of COVID-19, and ehealth and telehealth services can be provided without in-person visits [8,9], while decision support provided by artificial intelligence methods can facilitate doctors' decisions [10]. Numerous applications are categorized under IoHT applications. For example, indoor localization and IoT applications inside smart buildings such as keeping social distances have been used since the COVID-19 pandemic began [11]. In addition, such applications are used for traditional tasks such as the monitoring of daily life activity [5,12,13], fall detection [14] and assisted living [15–18], bad habits (such as smoking) detection and control [19], monitoring of industrial workers' activity [20], monitoring the heart rate of vehicle drivers [21], using wearable sensors to monitor heart activity [22,23], mHealth Apps for Self-Management [24], gait detection for people with Parkinson's disease [25,26], and many others.

The implementation of IoHT systems starts with data acquisition, followed by a preprocessing and feature-extraction phase, and finally arrives at the decision-making stage. Most known approaches in the literature can be categorized as video-based, WiFi-based, and sensory-based. Video-based human activity monitoring approaches may provide rich information via videos and images for indoor activities when there are no ad hoc cameras in outdoor environments such as walking tracks, parks, traditional malls, and swimming pools. Conversely, both wearable sensors and smart handheld devices are very suitable for the environment-invariant Human Activity Recognition (HAR) models. Another concern is that maintaining the privacy of individuals is questionable in vision-based approaches [27], while dealing with data fusion from sensors presents no such compromise. However, WiFi-based recognition of activities of daily life [28–30] has the advantage of using the fixed WiFi devices, but such approach has no applicability in outdoor environments.

A great interest is devoted to employing wearable sensors (e.g., accelerometer units), embedded sensors in smart devices (e.g., accelerometer, gyroscope, and magnetometer), and Kinect sensors [31] to develop HAR models [12,32]. Currently, smart devices such as smartphones and smartwatches are receiving much attention in such IoHT applications for obvious reasons [5]. On the other hand, a special-purpose real-time health monitoring device may have concerns regarding the efficient implementation in terms of power consumption [33]. When data acquisition is performed through many sensors and/or devices, there is a need for a suitable IoT framework to be able to move to the preprocessing stage. In preprocessing stage, the tri-axial activity signals registered by the sensors usually first need noise filtration, then segmentation in window length that ranges from <1 to 30 s [5,34] with more focus on reasonable small lengths (e.g., 2–10 s) in order to simulate real-time situations. Furthermore, feature extraction can follow the traditional approach of handcrafting a set of fine features selected in the time domain (mean, standard deviation, min, max, Pearson coefficients, etc.) and the frequency domain (energy, entropy, FFT coefficients, etc.), or they may follow the modern trend of deep learning networks [16,34]. In the latter approach, features are implicitly extracted as the encodings of hidden layers of the network, while outer layers such as fully connected layers together with softmax layer are responsible for the decision-making (i.e., classification and recognition). Following the feature engineering approach, the Random Forest (RF) algorithm [35], Multilayer Perceptron (MLP) [36] (one variant of artificial neural networks), Support Vector Machines (SVMs) [37], and Naive Bayes (NB) [38] are among the well-known shallow classifiers.

However, deep learning models perform well for many available human activity datasets in the literature [34], but the RF algorithm, for example, performs better than a single LSTM classifier for a specific dataset addressed in [16]. In addition, the recent studies in [17,39–41] in IoT applications depend on shallow classifiers. Recently, hybrid ensemble

approaches that make use of shallow classifiers in addition to deep convolutional layers are significantly bullish [28].

The limitation of existing approaches concerning a dataset collected by two smartphone units (in pocket and wrist positions) of human activities and gestures introduced by Shoaib et al. [42] motivates improving the state-of-the-art results. In this paper, an interesting and challenging dataset of thirteen activities is addressed. Activities are divided into two groups: the first group consists of hand gestures such as eating, smoking, drinking coffee, typing, and writing, and the other group consists of biking, jogging, standing, sitting, walking, walking upstairs, and walking downstairs. As a classification problem, the whole dataset is handled at a time in the training and testing processes. Using a feature set that is adequate to sensors' positions on the human body, an impartial comparison between the aforementioned shallow classifiers is conducted. The RF algorithm shows outstanding performance compared to previous models in the literature according to both subject-dependent and stratified k-fold cross-validation evaluation metrics. Furthermore, for testing the model generalization, another dataset, namely WISDM v1 [43], is used to examine the applied model performance.

1.2. Related Work

In the literature, numerous human activity datasets were collected from smartphones and/or smartwatches, e.g., WISDM v1 and v2, UCI-HAR, and UniMiB SHAR; see the survey by Demrozi et al. [44] for complete details. Shoaib et al. published a public dataset in [42] using two smartphone units. Below, we shed light on some closely related studies that addressed this dataset. In [42], a simple feature set of mean, standard deviation, median, min, max, semi-quartile, and the sum of the first ten FFT coefficients were extracted from each sensor stream, and the magnitude of its 3-dimensional signal was applied to the NB classifier. Since the readings of the accelerometer, linear accelerometer, gyroscope, and magnetometer sensors in both smartphones were registered, the focus in [42] was to evaluate the combination of sensors and device positions on the body, besides determining the effect of the window length from 2 to 30 s. The accelerometer and the gyroscope from both devices' positions gave the best performance. Baldominos et al. [45] performed a comparative study between different machine learning techniques (deep and shallow). Readings of the four sensors mentioned above were used. For shallow techniques, handcrafted features such as the mean and the standard deviation of raw signals and skewness, kurtosis, and the lower and upper quartiles of real coefficients of FFT of each dimension were obtained. The ensemble of randomized decision trees (ET) outperformed both shallow classifiers such as RF, MLP, NB, and K-nearest neighbors and convolutional neural networks (CNN). Alo et al. [46] examined two deep learning models, namely deep-stacked autoencoders (DSAE) and deep belief neural networks (DBNN). Only signals of the accelerometer are considered in both devices. Besides raw signals, the magnitude vector and the vectors of pitch and roll angles are used for training the models. The DSAE showed notable performance over both DBNN and the shallow classifiers (with the time-domain features in [42]) such as SVM, NB, and linear discriminant analysis. There are also deep learning models proposed for HAR using wearable sensors. For example, in [47], a combination of long short-term memory (LSTM) and a conventional neural network (CNN) was proposed to solve the HAR problem. In [48], a new HAR model was developed based on convolutional and LSTM recurrent units. In [49], a new model called iSPLInception was developed based on the Inception-ResNet framework from Google. It showed acceptable performance using different HAR datasets. In [50], the authors studied the applications of several deep learning methods. They found that the hybrid CNN-BiGRU showed the best results. Among the aforementioned studies, stratified k-fold evaluation criteria were applied by Shoaib et al. [42], while dataset samples were divided into train/test sets with a subject-dependent measure in [45,46]. Moreover, there is a variance between the different studies about the most suitable sensors for this task. Finally, there is some confusion about the superiority of conventional machine learning approaches versus deep learning models for this specific dataset.

To solve such conflicts, this paper proposes an individual model that proves superior according to both evaluation criteria. In addition, an impartial comparison between previous approaches and the current one has been performed.

1.3. Contribution of Current Work

- Presenting a light human-activity-recognition system using wearable sensors.
- Implementing a robust real-time model based on the Random Forest algorithm that outperforms other known classifiers and deep learning models.
- Handling a complex dataset of thirteen different human activities and gestures and improving the state-of-the-art results according to both subject-dependent and stratified k-fold cross-validation measures and using a different dataset, namely WISDM v1, for verifying model performance.
- Conducting sensitivity analysis for the applied model parameters (Random Forest size and depth).

1.4. Paper Organization

This document is organized as follows: Section 2 introduces the applied IoHT system framework. Section 3 presents the experimental results within the discussion. Section 4 handles the effect of important parameters on model performance. Section 5 provides a comparison with previous related studies. A different dataset is used to verify model performance in Section 6. The discussion of obtained results is given in Section 7. Section 8 includes conclusions, limitations, and future extensions of this work.

2. The Applied Approach

2.1. Dataset

Table 1 presents the generic information of dataset addressed here. Activity signals were recorded at a frequency of 50 Hz from the accelerometer, linear accelerometer, gyroscope, and magnetometer sensors of two Samsung Galaxy S2 smartphones. One device was put in the right pocket, and the other was placed on the right wrist. Ten subjects were asked to perform thirteen activities following a protocol; see Table 2 for the duration of each activity performed for each subject. This data set comprises six activities involving hand gestures, namely eating, smoking, drinking coffee, typing, and writing, and seven activities involving full-body motions, namely biking, jogging, standing, sitting, walking, walking upstairs, and walking downstairs. The total number of observations was 1,170,000. Activity signals were successfully registered, and there were no missing values. More details about the settings of collecting activities can be reviewed in [42].

Table 1. Dataset collection configuration.

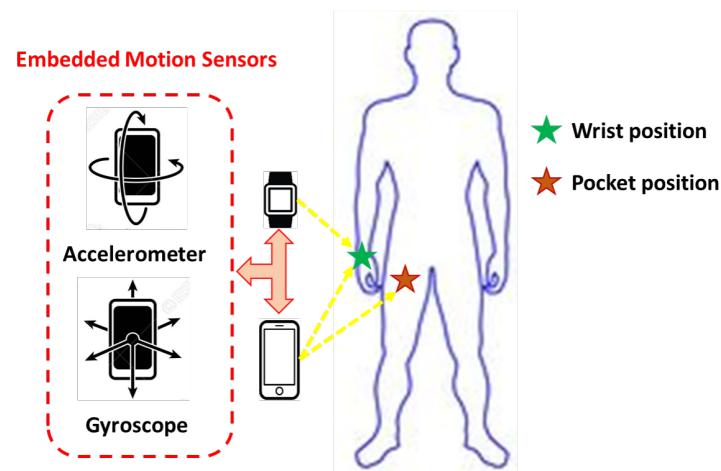
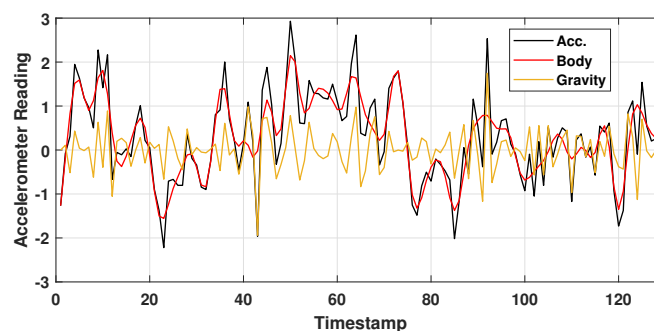
Parameter	Information
# Subjects	10
# Activities	13
Total # Observations	1,170,000
Missing values	NO
Device	Two Samsung Galaxy S2 smartphones
Position on Body	Right pocket and right wrist
Sensors	Accelerometer, Linear Accelerometer, Gyroscope and Magnetometer
Frequency	50 Hz

Table 2. Dataset activities.

Activity	Abbreviation	Duration (min)
Biking	BK	3
Having Coffee	CO	5
Walking Downstairs	DS	3
Eating	ET	5
Jogging	JO	3
Sitting	ST	3
Smoking	SM	5
Standing	SN	3
Giving a Talk	TK	5
Typing	TP	5
Walking Upstairs	UP	3
Walking	WK	3
Writing	WR	5

2.2. Sensory Data Processing

The applied model makes use of the readings of accelerometer and gyroscope sensors, where the acceleration and angular velocity of body limbs are sufficient for characterizing the activities performed. This point of view coincides with the well-known study of Anguita et al. [51]. Figure 1 clarifies the sensors' positions on the human body in order to acquire activity signals. Figure 2 shows the signal separation into body and gravity components using the Butterworth filter. Figure 3 presents the IoHT framework applied here. When applying the model, it is suggested to connect devices through Bluetooth technology. Then, the processing takes place at one central point (i.e., smartphone) as shown in Figure 3.

**Figure 1.** Activity signal acquisition from handheld smart devices.**Figure 2.** Accelerometer signal separation into body and gravity components using the Butterworth filter with a corner frequency of 20 Hz.

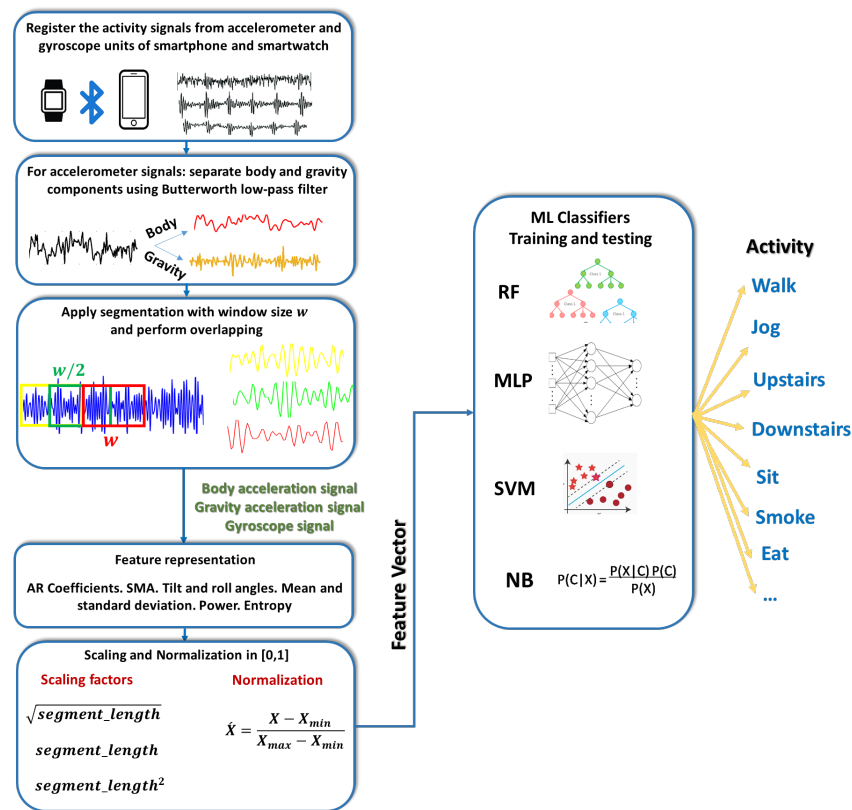


Figure 3. The composition of the applied IoHT system.

Activity Signal Preprocessing. According to previous studies, e.g., [51–53], it is preferred separate body and gravity components of accelerometer signals using, for example, a fourth-order Butterworth low-pass filter with a corner frequency of 20 Hz to filter out the body-acceleration component, since signals were collected at 50 Hz. For real-time considerations, signals were segmented using a window length of 2.56 s (i.e., 128 data points) with an overlap of 50% [51]. Figure 2 presents an illustrating example of acceleration signal separation for the walking activity in a time interval of 2.56 s. Thus, there is a fusion of six time-series signals: body acceleration, gravity acceleration, and gyroscope readings of both devices.

Feature Representation. The features for smartphone-based activity signals (with the numerical participation in the feature vector in parenthesis) are listed as follows:

- (F1-12) Mean and standard deviation (STD) of each of the acceleration signal (AS) and its jerk signal (JS)
- (F13-24) Autoregressive (AR) model coefficients for AS
- (F25) Signal magnitude area (SMA)
- (F26) Tilt angle (TA)
- (F27-30) Roll angle (RA) Equation (1): mean, STD, entropy of JS, and power
- (F31) Angle of x-component of AS Equation (2)
- (F32-34) Entropy of JS
- (F35-37) Power of AS

$$\text{Roll angle} = \arctan(-BA_z, -BA_y) \tag{1}$$

where BA_y and BA_z are body acceleration in y and z dimensions, respectively.

$$\text{Angle of x-component of AS} = \text{real} \left(\arccos \left(\max \left(\min \left(\frac{B_x \cdot G_m}{\|B_x\| * \|G_m\|}, 1 \right), -1 \right) \right) \right) \tag{2}$$

where the only real part of the resulting quantity is used; B_x and G_m are body acceleration in the x -axis and the mean of gravity component in 3D, respectively; and the denominator represents the multiplication of the 2-norm of each vector. For the rest of the features, the

readers can review [51]. Such a feature set is sensitive to body kinematics (e.g., wrist and leg motion in action). Thus, the 3D signals of each of the four operating sensors are represented by 37 features. Furthermore, combining the extracted features results in a 222-dimensional feature vector where the separation of body and gravity components of the accelerometer is performed.

Scaling and Normalization. The numerical values of the feature vector have a great variance in magnitude; e.g., SMA can reach a value that is a few hundred times that of the power of AS and the STD of acceleration JS. In order to eliminate the negative effect on the classification task, scaling is performed in terms of the segment length (*slen*). The coefficients of the AR model, TA, mean, and STD of AS, mean of JS, mean of RA, and power of RA are scaled by \sqrt{slen} , while the angle of the x-component of AS is scaled by *slen*, and finally the scaling factor $slen^2$ is applied for SMA. The rest of the features are used without scaling. This treatment is heuristically examined. After that, the whole feature vector is normalized in [0, 1] as illustrated in Figure 3.

Classification Layer. Commonly applied classification algorithms in human-activity-recognition tasks are referred to here as RF, MLP, SVM, and NB. RF [35] is a voting-based classifier where a decision tree is created for each sample inside a random subset of features. Then, the decision is taken for the sake of the class that is the most voted for. Thus, the most important parameters of the RF classifier are the number of decision trees and the maximum depth of the tree. MLP [36] contains interconnected processing units called neurons in one or more layers. Each neuron is characterized by its activation function, that is, a function of the weights of the preceding layer. The training algorithm, which is responsible for finding the best weights, plays a vital role in the network performance. In addition, the number of layers, number of neurons, and type of activation function are the most important parameters for the MLP. SVM [37] depends on finding the best hyperplanes that achieve the maximal margin between the nearest examples in high-dimensional spaces of two different classes. For a multiclass problem, $n * (n - 1) / 2$ binary SVM models are generated to distinguish *n* classes. NB [38] is a simple classifier that makes use of Bayes' rule to determine the class with the highest posterior probability.

3. Experimental Results and Analysis

3.1. Setup

Well-known machine learning (ML) classifiers in the IoT area, namely RF, MLP, SVM, and NB, are examined in an impartial comparison in order to clarify the most suitable one for this specific application. Since subject-dependent evaluation is usually easier than k-fold cross-validation in human-activity-recognition applications [54], the outstanding classifier according to the first mentioned criteria is examined in the later one. ML algorithms are referred to under the Scikit-learn framework in Python. Table 3 illustrates the parameters of each classifier during the experiments conducted here.

Table 3. Classifiers settings and parameter values.

Classifier	Function Call	Settings and Parameters
RF	<i>RandomForestClassifier()</i>	# estimators = 200, max. depth = 25, min. samples split = 2
MLP	<i>MLPClassifier()</i>	solver: quasi-Newton method, # hidden neurons = 75, activation function: tanh, max. # iterations = 1000, momentum = 0.9, initial learning rate is 0.01, validation ratio = 15%
SVM	<i>svm.SVC()</i>	kernel: radial basis function, polynomial degree is 3
NB	<i>GaussianNB()</i>	μ and σ parameters of Gaussian distribution are estimated using maximum likelihood

Performance of the examined ML algorithms is evaluated according to four metrics, namely the classification accuracy (Equation (3)); the F1-measure, which is the average of

precision and recall of classification; (Equations (4) and (5)); execution time; and size on the disk.

$$\text{Accuracy} = \frac{TP}{TP + TN + FP + FN} \quad (3)$$

$$\text{Precision} = \frac{TP}{TP + FP} \quad (4)$$

$$\text{Recall} = \frac{TP}{TP + FN} \quad (5)$$

where TP represents the true-positive, TN is the true-negative, FP is the false-positive and FN is the false-negative classification rate. The best settings for each classifier are used in experiments after examining various training options. Experiments run on a computer machine with 10 GB RAM and 2.60 GHz i5 CPU.

3.2. Subject-Dependent Evaluation

The samples of each class are randomly separated, with 70% in the training and validation set and 30% in the testing set. The test samples are never introduced training any of examined classifiers, but samples of the same subject may appear in both training and testing sets. For impartial comparison, the simulation procedure was repeated by 10 independent runs, where each time, the same training/testing data are provided to each classifier. The average classification rates for activity recognition are presented in Figure 4.

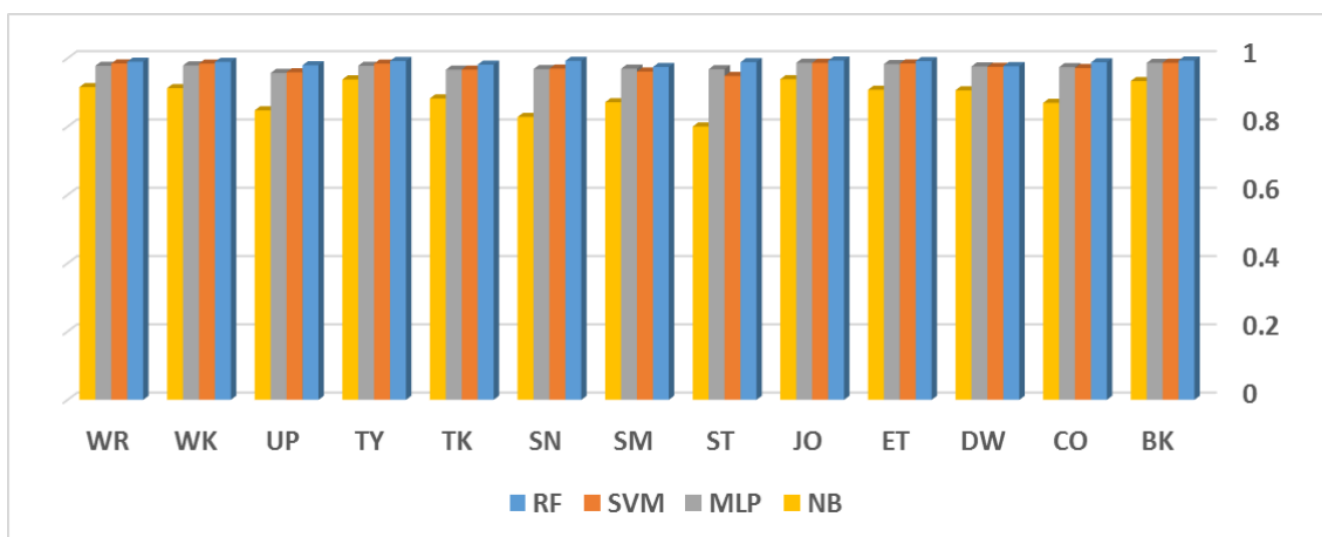


Figure 4. Average F-measure scores of the activities: Bike (BK), Coffee (CO), Downstairs (DW), Eat (ET), Jog (JO), Sit (ST), Smoke (SM), Stand (SN), Talk (TK), Type (TY), Upstairs (UP), Walk (WK) and Write (WR). Compared classifiers are evaluated under subject-dependent criteria.

Figure 4 shows the average classification rates for different activities per classifier. RF has the highest rate for each activity. Biking, eating, jogging, sitting, typing, and writing activities are successfully recognized with a rate $> 99\%$. The activities walking downstairs, walking upstairs, and smoking are the least recognized by the RF classifier with a rate slightly less than 98% . Such behavior can be justified by reading the confusion matrix shown in Figure 5. On average, eight examples of walking downstairs were misclassified as walking upstairs, and vice versa for 11 examples of walking upstairs. Another notable conflict occurred for nine examples between smoking and giving a talk. It was noticed that conflicts occurred between very close activities, which is likely expected in such applications. However, the overall performance of the current model (employed sensors + preprocessing + features + classifier) is accepted, and it can be further improved by providing more training examples.

True Classes	BK	418	0	0	0	0	0	0	0	0	0	0	0	
	CO	0	413	0	0	0	4	0	0	0	0	0	1	
	DW	0	0	408	0	0	0	0	0	0	8	2	0	
	ET	0	0	0	417	0	0	0	0	1	0	0	0	
	JO	0	0	1	0	417	0	0	0	0	0	0	0	
	ST	0	1	0	1	0	415	0	0	0	0	0	1	
	SM	0	0	0	0	0	0	408	1	9	0	0	0	
	SN	0	0	0	0	0	0	7	410	1	0	0	0	
	TK	0	0	0	0	0	0	9	0	409	0	0	0	
	TY	0	0	0	0	0	0	0	0	0	417	0	1	
	UP	0	0	11	0	0	0	0	0	0	0	406	1	
	WK	0	0	2	0	0	0	0	0	0	3	413	0	
	WR	0	3	0	0	0	1	0	0	0	0	0	414	
			BK	CO	DW	ET	JO	ST	SM	SN	TK	TY	UP	WK
		Predicted Classes												

Figure 5. Confusion matrix for the RF classifier under subject-dependent evaluation criteria.

Table 4 provides a summary of comparing different ML algorithms, as well as important implementation issues. On average, the accuracy (and F-measure) of RF reaches 98.72%, which exceeds the accuracy of each of SVM, MLP, and NB by 1.3%, 1.27%, and 11.1%, respectively. MLP takes a notably long training time of 90.41 s, while NB training occurred quickly at less than one second, and RF needed about 29.3 s to announce its decisions. RF occupies about 22.68 MB of the disk, which is the largest size, while NB needs only 0.046 MB space. To improve the readability of comparative results of all classifiers, Figure 6 presents an illustrative radar plot.

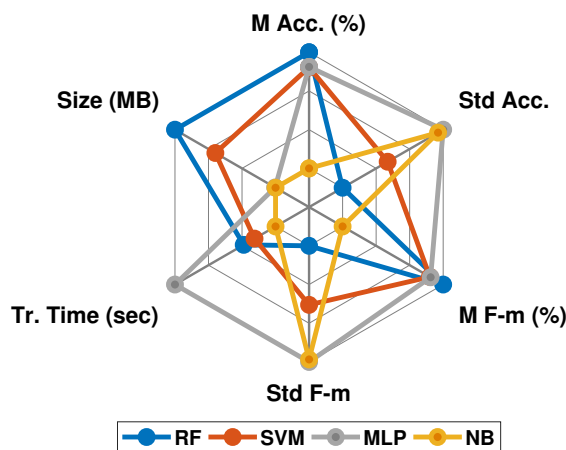


Figure 6. Radar plot for compared classifiers according to mean and standard deviation of accuracy (M Acc (%) and Std Acc), respectively; mean and standard deviation of F-measure (M F-m (%) and Std F-m), respectively; raining time in sec. (Tr. Time (sec)); and Size on disk in MB (Size (MB)).

3.3. Stratified k-Fold Cross-Validation

In the experimental settings of collecting this dataset, a controlled protocol was performed by each of the 10 participants. Each participating subject performed the same set of activities within the same permitted time duration. Thus, by chance, for this particular dataset, 10-fold cross-validation implicitly involved the stratified 10-fold validation followed in Shoaib et al. [42]. Moreover, the common evaluation criterion for human activity recognition models, i.e., leave-one-subject-out, can also be applied via the 10-fold cross-validation for this particular dataset. The latter measure criteria are of interest where the dataset provides subject-independent evaluation, and hence it examines the model’s

of generalization ability for newly introduced data. The average accuracy of the applied RF-based model here is equal to 92.54%.

Table 4. Performance of compared classifiers for subject-dependent evaluation.

	Accuracy		F-Measure		Training Time (sec)	Size on Disk (MB)
	Mean	Std	Mean	Std		
RF	Mean	98.72	Mean	98.72	29.3	22.683
	Std	0.1015	Std	0.1015		
SVM	Mean	97.43	Mean	97.42	19.69	13.593
	Std	0.2279	Std	0.2398		
MLP	Mean	97.47	Mean	97.49	90.41	0.143
	Std	0.3837	Std	0.3736		
NB	Mean	88.82	Mean	88.87	1	0.046
	Std	0.3693	Std	0.3677		

4. Sensitivity Analysis for Model Parameters

The performance of the RF algorithm is tremendously sensitive to both the number of decision trees (known as RF size) and the longest path from a tree head to the leaves (known as RF depth). For RF depth ≥ 15 , with a suitable RF size ≥ 50 , the applied RF-based model can provide notable recognition performance under subject-dependent evaluation measure; see Figure 7a. Moreover, increasing the RF size up to 400 trees has a slight improvement in the model accuracy. Conversely, under 10-fold cross-validation evaluation, the model accuracy grows by 1% when increasing both RF size and RF depth from (50, 10) to (15, 200); see Figure 7b. Moreover, increasing the RF size to 400, for example, will not enhance the model accuracy as much as the notable increment in processing time in this case. From Figure 7, we can conclude that with an RF depth between 15 and 25 and an RF size equal to 200, an efficient recognition model can be implemented for these kinds of IoHT systems that make use of sensory data from smartphones.

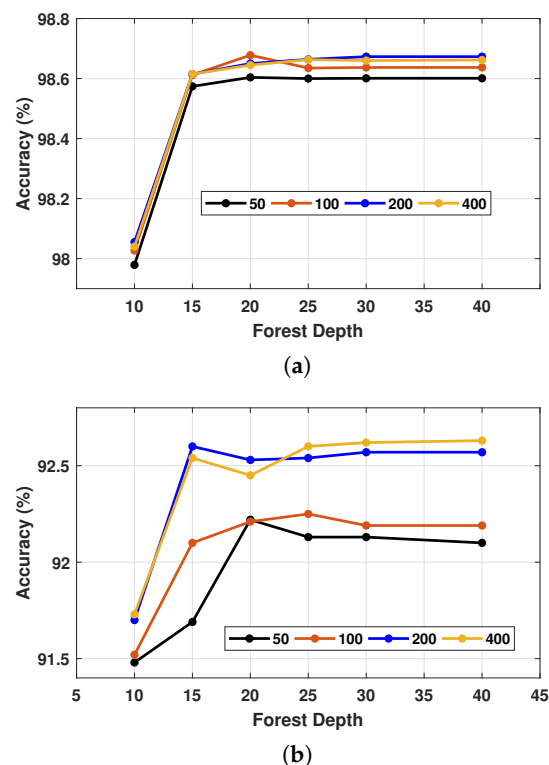


Figure 7. Model performance for different values of both forest size (50, 100, 200, 400) and forest depth (15, 20, 25, 30, 40) for (a) subject-dependent and (b) 10-fold cross validation criteria.

5. Comparison with Previous Studies

Different studies in the literature have addressed this dataset according to different evaluation measures. Table 5 provides the previous best recognition rates according to subject-dependent evaluation. Baldominos et al. [45] have tested shallow techniques against the deep CNN model. Only raw signals are used in 60 s segments. The ensemble of randomized decision trees (ET), with a set of handcrafted features, provides an average overall accuracy of 95.3%, while the accuracy of the CNN-based approach decreases to 85%. Stacked autoencoders provided better results than deep belief networks, where the accuracy reached 97.13% according to Alo et al. [46]. In a later study, besides raw activity signals, the magnitude vector and the vectors of pitch and roll angles were provided to deep networks in segments with a length of 2 s.

Table 5. Recognition rates of each activity for different models under subject-dependent validation criteria. ET: ensemble of randomized trees, FC: fully connected layer, AE: autoencoders, and DBN: deep belief networks.

Reference	Input Signals	Segment Length (s)	Feature Extraction	Classifier	Accuracy (%)
Baldominos et al. [45]	Raw signals	60	Handcrafted	ET	95.3
Baldominos et al. [45]	Raw signals	60	CNN hidden layers	FC layer	85
Alo et al. [46]	Raw signals, magnitude vector, pitch and roll vectors	2	Sparse AE layers	FC layer	97.13
Alo et al. [46]	Raw signals, magnitude vector, pitch and roll vectors	2	DBN hidden layers	DBN output layer	91.57
Current model	Raw signals	2.56	Handcrafted	RF	98.7

The proposed DL model was able to outperform the conventional classifiers such as support vector machines (SVM), Naive Bayes (NB), and linear discriminant analysis (LDA); however, the RF classifier was not included in this comparison. The current RF-based model presents the best recognition results among related studies. However, samples of the same person may appear in both the training and testing sets, but the experimental findings are still useful for seeking good models since registered data points occurred at different timestamps.

Moreover, the current model improves the recognition rates obtained by Shoaib et al. [42]. Table 6 shows the rates of each activity when stratified 10-fold cross-validation criteria are applied. Numerical values of Shoaib et al.'s model were computed from the confusion matrix in Figure 2c in [42]. The applied classifier was NB, but features were extracted from segments with a length of 5 s, and only accelerometer and gyroscope signals were used. Because of the suitable feature set used within the current model, the activities that directly depend on hand movement are well-recognized. The improvements in the rates of activities are as follows: having coffee (0.83 to 0.92), eating (0.89 to 0.99), smoking (0.82 to 0.95), giving a talk (0.86 to 0.97), typing (0.95 to 0.98), writing (0.89 to 0.97). For the other activities, the current model performs worse than or equal to Shoaib et al.'s model. In conclusion, the average overall accuracy is improved by 1.4%.

Table 6. Recognition rates of each activity for different models under 10-fold cross validation criteria.

Ref.	BK	CO	DW	ET	JO	ST	SM	SN	TK	TY	UP	WK	WR	Accuracy (%)
Shoaib et al. [42]	0.99	0.83	0.98	0.89	1	0.90	0.82	0.92	0.86	0.95	0.96	0.85	0.89	91.2
Current model	0.99	0.92	0.91	0.99	0.99	0.76	0.95	0.94	0.97	0.98	0.83	0.80	0.97	92.54

6. Applied Model Performance for WISDM Dataset

In this section, the validation of the applied framework is extended to the WISDM dataset [43]. It is one of the most addressed datasets in the HAR literature. WISDM v1 contains a total of 1,098,207 examples of activities that have been collected by 29 subjects. Six activities, namely walking (37.2%), jogging (29.2%), upstairs (12.0%), downstairs (10.2%), sitting (6.4%), and standing (5%), were registered via a smartphone in the front pants pocket (see Figure 1) of each subject. Walking and jogging activities were the most represented in this dataset. Activity signals were registered using the embedded accelerometer of the smartphone at a 20 Hz sampling rate. In the experimental settings, a window size of 10 s (according to the original study [43]) with 50% overlapping was applied to raw signals. The proposed feature set was generated for each activity segment, where the feature vector was 74 dimensions; since only the accelerometer signals are available, the RF classifier is called. Using the best settings, e.g., RF size and depth (200, 25), gave acceptable classification rates for this dataset. For 10-fold cross-validation criteria, the applied model gave an average accuracy of 94%, while for the subject-dependent evaluation (i.e., 70% training and 30% testing), the average accuracy reached 98.56%. This model performance regarding this dataset is comparable to many recent related studies in the literature, as summarized in Table 7.

Among the compared studies that appear in Table 7, using a window of 5 s for segments in [55] is more challenging than using longer segments, but a deep learning model was able to achieve 94.2% accuracy under 10-fold cross-validation. Moreover, an accuracy value of 98.85% was obtained in [56], but applying 95% overlapping when doing segmentation, and this is questionable in such a HAR study (i.e., overlapping usually ranges from 0 to 50%). In addition, for a 70%/30% split, using a more efficient RF such as (50, 20) gives an average accuracy of 98.34%, which is still close to the best performance obtained. However, under 10-fold cross-validation, using an RF with (50, 20) does not degrade the accuracy by more than 0.02%.

Summing up, the applied framework shows good performance for the WISDM v1 dataset under different evaluation criteria, while usually, only one of them is used in previous related studies. This model behavior reflects the robustness and suitability of both the feature set and the classifier algorithm for real-time HAR applications.

Table 7. Applied model results for WISDM dataset. MLP: multi-layer perceptron. LR: logistic regression. Stat. Feat.: statistical features. Att. M.: attention mechanism. R. B.: residual block. LSTM: Long short-term memory.

Evaluation	Reference	Segment Length (s)	Feature Extraction	Classifier	Accuracy (%)
10-fold cross validation	Kwapisz et al. [43]	10	Handcrafted	MLP	91.7
	Garcia-Ceja et al. [55]	5	CNN	FC layer	94.2
	Catal et al. [57]	10	Handcrafted	Ensemble of (LR, MLP, j48)	91.62
	Ignatov [58]	10	CNN + Stat. Feat.	FC layer	93.32
	Current model	10	Handcrafted	RF	94
70%/30% split	Gao et al. [56]	10	CNN + Att. M.	FC layer	98.85
	Suwannarat et al. [59]	8	CNN	FC layer	95
	Abdel-Basset et al. [60]	10	CNN + R. B. + LSTM + Att. M.	MLP	98.90
	Zhang et al. [61]	11.2	CNN	FC layers	96.4
	Zhang et al. [62]	10	CNN + Att.	FC layer	96.4
	Current model	10	Handcrafted	RF	98.56

7. Discussion

The applied framework introduces one example of an IoHT system that is examined using two datasets with different settings. Shoaib's dataset contains thirteen activities gathered by 10 subjects at a sampling rate of 50 Hz, while WISDM v1 has six activities collected by 29 subjects at a sampling rate of 20 Hz. Such a variety of activity signal resources constitutes a strong test for any proposed HAR model. Applying the different common evaluation criteria of HAR models in the same study is highly recommended to ensure its superiority. Later observation is missing in most studies in the literature.

More evidence is needed for the use of the dense production of deep learning models in the HAR field. Such models have thousands of parameters learned during training (tremendous computational load). However, they should at least outperform the conventional shallow approaches. Classical handcrafted features are meaningful and interoperable to a great extent, while the interpretation of most deep models, in particular in the HAR field, is still in its infancy.

In [46], the applied DL model required the help of extra inputs such as magnitude and pitch and roll signals, together with the raw 3D acceleration signals, in order to improve the performance. On the other hand, features extracted implicitly from DL models may need refinement via feature selection approaches in order to eliminate illusive features of classifiers. Recent studies such as [63] and others have emphasized the role of applying feature selection with DL models. On the other hand, the RF algorithm performs feature selection as one of the steps performed to achieve its classification result. One important observation is the degradation of accuracy when moving from the subject-dependent to 10-fold cross-validation criteria. For the WISDM v1 dataset, the misclassification is relatively high between upstairs and downstairs in comparison to other activities, in addition to the difficulty when applying 10-fold cross-validation (i.e., different subjects are used for training and testing). The later result has also been reported by different previous models such as [43,55,60], which can probably be attributed to the sensor position on subjects' bodies. A similar notation also holds for Shoaib's dataset, where in Figure 5, the confusion matrix shows that the majority of false predictions take place between the activities of walking upstairs and walking downstairs.

8. Conclusions and Future Trends

In this work, an efficient model for an IoHT system is introduced through a set of carefully handcrafted features and a shallow classifier such as Random Forest for the dataset of Shoaib et al. [42]. Participants used to collect this dataset followed a specific protocol, which may be called a controlled environment. Similarly to related studies, using accelerometers and gyroscope sensors in smartphones is convenient for such applications. Moreover, inducing features (e.g., statistics of the roll angle vector and the angle of the x-component of body acceleration with a gravity vector) that depend on body kinematics (e.g., wrist and leg motion) improve the model performance. The presented model provides state-of-the-art results under both subject-dependent and 10-fold cross-validation criteria. Moreover, the current model performance was verified by another dataset, namely WISDM v1 [43] under both aforementioned evaluation criteria.

Author Contributions: All authors contributed equally to this paper. All authors have read and agreed to the published version of the manuscript.

Funding: This work was supported by National Natural Science Foundation of China (Grant No. 62150410434).

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The data are publicly available as described in the main text.

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

References

1. Baker, S.B.; Xiang, W.; Atkinson, I. Internet of Things for Smart Healthcare: Technologies, Challenges, and Opportunities. *IEEE Access* **2017**, *5*, 26521–26544. [\[CrossRef\]](#)
2. Qadri, Y.A.; Nauman, A.; Zikria, Y.B.; Vasilakos, A.V.; Kim, S.W. The Future of Healthcare Internet of Things: A Survey of Emerging Technologies. *IEEE Commun. Surv. Tutor.* **2020**, *22*, 1121–1167. [\[CrossRef\]](#)
3. Kamruzzaman, M.M.; Alrashdi, I.; Alqazzaz, A. New Opportunities, Challenges, and Applications of Edge-AI for Connected Healthcare in Internet of Medical Things for Smart Cities. *J. Healthc. Eng.* **2022**, *2022*, 2950699. [\[CrossRef\]](#) [\[PubMed\]](#)
4. Dimitrov, D.V. Medical internet of things and big data in healthcare. *Healthc. Inform. Res.* **2016**, *22*, 156–163. [\[CrossRef\]](#) [\[PubMed\]](#)
5. Wang, Y.; Cang, S.; Yu, H. A survey on wearable sensor modality centred human activity recognition in health care. *Expert Syst. Appl.* **2019**, *137*, 167–190. [\[CrossRef\]](#)
6. Pillai, S.G.; Haldorai, K.; Seo, W.S.; Kim, W.G. COVID-19 and hospitality 5.0: Redefining hospitality operations. *Int. J. Hosp. Manag.* **2021**, *94*, 102869. [\[CrossRef\]](#) [\[PubMed\]](#)
7. Demir, K.A.; Döven, G.; Sezen, B. Industry 5.0 and human–robot co-working. *Procedia Comput. Sci.* **2019**, *158*, 688–695. [\[CrossRef\]](#)
8. Vanagas, G.; Engelbrecht, R.; Damaševičius, R.; Suomi, R.; Solanas, A. eHealth Solutions for the Integrated Healthcare. *J. Healthc. Eng.* **2018**, *2018*, 3846892. [\[CrossRef\]](#) [\[PubMed\]](#)
9. Hernández-Chan, G.S.; Ceh-Varela, E.E.; Sanchez-Cervantes, J.L.; Villanueva-Escalante, M.; Rodríguez-González, A.; Pérez-Gallardo, Y. Collective intelligence in medical diagnosis systems: A case study. *Comput. Biol. Med.* **2016**, *74*, 45–53. [\[CrossRef\]](#)
10. Rodríguez-González, A.; Torres-Niño, J.; Mayer, M.A.; Alor-Hernandez, G.; Wilkinson, M.D. Analysis of a multilevel diagnosis decision support system and its implications: A case study. *Comput. Math. Methods Med.* **2012**, *2012*, e0148991. [\[CrossRef\]](#)
11. Barsocchi, P.; Calabrò, A.; Crivello, A.; Daoudagh, S.; Furfari, F.; Girolami, M.; Marchetti, E. COVID-19 & privacy: Enhancing of indoor localization architectures towards effective social distancing. *Array* **2021**, *9*, 100051.
12. Kiran, S.; Khan, M.A.; Javed, M.Y.; Alhaisoni, M.; Tariq, U.; Nam, Y.; Damaševičius, R.; Sharif, M. Multi-Layered Deep Learning Features Fusion for Human Action Recognition. *Comput. Mater. Contin.* **2021**, *69*, 4061–4075. [\[CrossRef\]](#)
13. Şengül, G.; Ozelcik, E.; Misra, S.; Damaševičius, R.; Maskeliūnas, R. Fusion of smartphone sensor data for classification of daily user activities. *Multimed. Tools Appl.* **2021**, *80*, 33527–33546. [\[CrossRef\]](#)
14. Şengül, G.; Karakaya, M.; Misra, S.; Abayomi-Alli, O.O.; Damaševičius, R. Deep learning based fall detection using smartwatches for healthcare applications. *Biomed. Signal Process. Control.* **2022**, *71*, 103242. [\[CrossRef\]](#)
15. Khan, M.F.; Ghazal, T.M.; Said, R.A.; Fatima, A.; Abbas, S.; Khan, M.; Issa, G.F.; Ahmad, M.; Khan, M.A. An IoMT-Enabled Smart Healthcare Model to Monitor Elderly People Using Machine Learning Technique. *Comput. Intell. Neurosci.* **2021**, *2021*, 2487759. [\[CrossRef\]](#)
16. Farsi, M. Application of ensemble RNN deep neural network to the fall detection through IoT environment. *Alex. Eng. J.* **2021**, *60*, 199–211. [\[CrossRef\]](#)
17. Moualla, S.; Khorzom, K.; Jafar, A. Improving the performance of machine learning-based network intrusion detection systems on the UNSW-NB15 dataset. *Comput. Intell. Neurosci.* **2021**, *2021*, 5557577. [\[CrossRef\]](#)
18. Maskeliūnas, R.; Damaševičius, R.; Segal, S. A review of internet of things technologies for ambient assisted living environments. *Future Internet* **2019**, *11*, 259. [\[CrossRef\]](#)
19. Agac, S.; Shoaib, M.; Incel, O.D. Context-aware and dynamically adaptable activity recognition with smart watches: A case study on smoking. *Comput. Electr. Eng.* **2021**, *90*, 106949. [\[CrossRef\]](#)
20. Patalas-maliszewska, J.; Halikowski, D.; Damaševičius, R. An automated recognition of work activity in industrial manufacturing using convolutional neural networks. *Electronics* **2021**, *10*, 2946. [\[CrossRef\]](#)
21. Bharti, R.; Khamparia, A.; Shabaz, M.; Dhiman, G.; Pande, S.; Singh, P. Prediction of heart disease using a combination of machine learning and deep learning. *Comput. Intell. Neurosci.* **2021**, *2021*, 8387680. [\[CrossRef\]](#)
22. Girčys, R.; Kazanavičius, E.; Maskeliūnas, R.; Damaševičius, R.; Woźniak, M. Wearable system for real-time monitoring of hemodynamic parameters: Implementation and evaluation. *Biomed. Signal Process. Control.* **2020**, *59*, 101873. [\[CrossRef\]](#)
23. Olmedo-Aguirre, J.O.; Reyes-Campos, J.; Alor-Hernández, G.; Machorro-Cano, I.; Rodríguez-Mazahua, L.; Sánchez-Cervantes, J.L. Remote Healthcare for Elderly People Using Wearables: A Review. *Biosensors* **2022**, *12*, 73. [\[CrossRef\]](#)
24. Cruz-ramos, N.A.; Alor-hernández, G.; Colombo-mendoza, L.O.; Sánchez-cervantes, J.L.; Rodríguez-mazahua, L.; Guarneros-nolasco, L.R. mHealth Apps for Self-Management of Cardiovascular Diseases: A Scoping Review. *Healthcare* **2022**, *10*, 322. [\[CrossRef\]](#)
25. Yang, Z. An Efficient Automatic Gait Anomaly Detection Method Based on Semisupervised Clustering. *Comput. Intell. Neurosci.* **2021**, *2021*, 8840156. [\[CrossRef\]](#)
26. Priya, S.J.; Rani, A.J.; Subathra, M.S.P.; Mohammed, M.A.; Damaševičius, R.; Ubendran, N. Local pattern transformation based feature extraction for recognition of parkinson’s disease based on gait signals. *Diagnostics* **2021**, *11*, 1395. [\[CrossRef\]](#)
27. Bokhari, S.M.; Sohaib, S.; Khan, A.R.; Shafi, M. DGRU based human activity recognition using channel state information. *Measurement* **2021**, *167*, 108245. [\[CrossRef\]](#)
28. Cui, W.; Li, B.; Zhang, L.; Chen, Z. Device-free single-user activity recognition using diversified deep ensemble learning. *Appl. Soft Comput.* **2021**, *102*, 107066. [\[CrossRef\]](#)
29. Al-qaness, M.A. Device-free human micro-activity recognition method using WiFi signals. *Geo-Spat. Inf. Sci.* **2019**, *22*, 128–137. [\[CrossRef\]](#)

30. Al-Qaness, M.A.; Abd Elaziz, M.; Kim, S.; Ewees, A.A.; Abbasi, A.A.; Alhaj, Y.A.; Hawbani, A. Channel state information from pure communication to sense and track human motion: A survey. *Sensors* **2019**, *19*, 3329. [[CrossRef](#)]
31. Ryselis, K.; Petkus, T.; Blažauskas, T.; Maskeliūnas, R.; Damaševičius, R. Multiple Kinect based system to monitor and analyze key performance indicators of physical training. *Hum.-Centric Comput. Inf. Sci.* **2020**, *10*, 51. [[CrossRef](#)]
32. Al-Qaness, M.A.; Dahou, A.; Abd Elaziz, M.; Helmi, A. Multi-ResAtt: Multilevel Residual Network with Attention for Human Activity Recognition Using Wearable Sensors. *IEEE Trans. Ind. Inform.* **2022**. [[CrossRef](#)]
33. Siam, A.I.; Almaiah, M.A.; Al-Zahrani, A.; Elazm, A.A.; El Banby, G.M.; El-Shafai, W.; El-Samie, F.E.A.; El-Bahnasawy, N.A. Secure Health Monitoring Communication Systems Based on IoT and Cloud Computing for Medical Emergency Applications. *Comput. Intell. Neurosci.* **2021**, *2021*, 8016525. [[CrossRef](#)] [[PubMed](#)]
34. Sousa Lima, W.; 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](#)]
35. Svetnik, V.; Liaw, A.; Tong, C.; Culberson, J.C.; Sheridan, R.P.; Feuston, B.P. Random forest: A classification and regression tool for compound classification and QSAR modeling. *J. Chem. Inf. Comput. Sci.* **2003**, *43*, 1947–1958. [[CrossRef](#)]
36. Priddy, K.L.; Keller, P.E. *Artificial Neural Networks: An Introduction*; SPIE Press: Bellingham, WA, USA, 2005; Volume 68.
37. Wang, L. *Support Vector Machines: Theory and Applications*; Springer Science & Business Media: Berlin/Heidelberg, Germany, 2005; Volume 177.
38. Murphy, K.P. Naive bayes classifiers. *Univ. Br. Columbia* **2006**, *18*, 1–8.
39. Khan, R.; Kumar, S.; Srivastava, A.K.; Dhingra, N.; Gupta, M.; Bhati, N.; Kumari, P. Machine Learning and IoT-Based Waste Management Model. *Comput. Intell. Neurosci.* **2021**, *2021*, 5942574. [[CrossRef](#)]
40. Alsamhi, S.H.; Almalki, F.A.; Al-Dois, H.; Ben Othman, S.; Hassan, J.; Hawbani, A.; Sahal, R.; Lee, B.; Saleh, H. Machine learning for smart environments in B5G networks: Connectivity and QoS. *Comput. Intell. Neurosci.* **2021**, *2021*, 6805151. [[CrossRef](#)]
41. Atalaa, B.A.; Ziedan, I.; Alenany, A.; Helmi, A. Feature Engineering for Human Activity Recognition. *Int. J. Adv. Comput. Sci. Appl.* **2021**, *12*, 160–167 [[CrossRef](#)]
42. Shoaib, M.; Bosch, S.; Incel, O.D.; Scholten, H.; Havinga, P.J. Complex human activity recognition using smartphone and wrist-worn motion sensors. *Sensors* **2016**, *16*, 426. [[CrossRef](#)]
43. Kwapisz, J.R.; Weiss, G.M.; Moore, S.A. Activity recognition using cell phone accelerometers. *ACM SigKDD Explor. Newsl.* **2011**, *12*, 74–82. [[CrossRef](#)]
44. Demrozi, F.; Pravadelli, G.; Bihorac, A.; Rashidi, P. Human activity recognition using inertial, physiological and environmental sensors: A comprehensive survey. *IEEE Access* **2020**, *8*, 210816–210836. [[CrossRef](#)]
45. Baldominos, A.; Cervantes, A.; Saez, Y.; Isasi, P. A comparison of machine learning and deep learning techniques for activity recognition using mobile devices. *Sensors* **2019**, *19*, 521. [[CrossRef](#)]
46. Alo, U.R.; Nweke, H.F.; Teh, Y.W.; Murtaza, G. Smartphone Motion Sensor-Based Complex Human Activity Identification Using Deep Stacked Autoencoder Algorithm for Enhanced Smart Healthcare System. *Sensors* **2020**, *20*, 6300. [[CrossRef](#)]
47. Xia, K.; Huang, J.; Wang, H. LSTM-CNN architecture for human activity recognition. *IEEE Access* **2020**, *8*, 56855–56866. [[CrossRef](#)]
48. Ordóñez, F.J.; Roggen, D. Deep convolutional and lstm recurrent neural networks for multimodal wearable activity recognition. *Sensors* **2016**, *16*, 115. [[CrossRef](#)]
49. Ronald, M.; Poulouse, A.; Han, D.S. iSPLInception: An inception-ResNet deep learning architecture for human activity recognition. *IEEE Access* **2021**, *9*, 68985–69001. [[CrossRef](#)]
50. Mekruksavanich, S.; Jitpattanakul, A. Deep convolutional neural network with rnns for complex activity recognition using wrist-worn wearable sensor data. *Electronics* **2021**, *10*, 1685. [[CrossRef](#)]
51. Anguita, D.; Ghio, A.; Oneto, L.; Parra, X.; Reyes-Ortiz, J.L. A public domain dataset for human activity recognition using smartphones. *Esann* **2013**, *3*, 437–442.
52. Hamäläinen, W.; Järvinen, M.; Martiskainen, P.; Mononen, J. Jerk-based feature extraction for robust activity recognition from acceleration data. In Proceedings of the 2011 11th International Conference on Intelligent Systems Design and Applications, Córdoba, Spain, 22–24 November 2011; pp. 831–836.
53. Quiroz, J.C.; Banerjee, A.; Dascalu, S.M.; Lau, S.L. Feature selection for activity recognition from smartphone accelerometer data. *Intell. Autom. Soft Comput.* **2017**, 1–9. [[CrossRef](#)]
54. Dehghani, A.; Sarbishei, O.; Glatard, T.; Shihab, E. A quantitative comparison of overlapping and non-overlapping sliding windows for human activity recognition using inertial sensors. *Sensors* **2019**, *19*, 5026. [[CrossRef](#)] [[PubMed](#)]
55. Garcia-Ceja, E.; Uddin, M.Z.; Torresen, J. Classification of recurrence plots' distance matrices with a convolutional neural network for activity recognition. *Procedia Comput. Sci.* **2018**, *130*, 157–163. [[CrossRef](#)]
56. Gao, W.; Zhang, L.; Teng, Q.; He, J.; Wu, H. DanHAR: Dual attention network for multimodal human activity recognition using wearable sensors. *Appl. Soft Comput.* **2021**, *111*, 107728. [[CrossRef](#)]
57. Catal, C.; Tufekci, S.; Pirmitt, E.; Kocabag, G. On the use of ensemble of classifiers for accelerometer-based activity recognition. *Appl. Soft Comput.* **2015**, *37*, 1018–1022. [[CrossRef](#)]
58. Ignatov, A. Real-time human activity recognition from accelerometer data using Convolutional Neural Networks. *Appl. Soft Comput.* **2018**, *62*, 915–922. [[CrossRef](#)]
59. Suwannarat, K.; Kurdthongmee, W. Optimization of Deep Neural Network-based Human Activity Recognition for a Wearable Device. *Heliyon* **2021**, *7*, e07797. [[CrossRef](#)]

60. Abdel-Basset, M.; Hawash, H.; Chakraborty, R.K.; Ryan, M.; Elhoseny, M.; Song, H. ST-DeepHAR: Deep learning model for human activity recognition in IoHT applications. *IEEE Internet Things J.* **2020**, *8*, 4969–4979. [[CrossRef](#)]
61. Zhang, Y.; Zhang, Z.; Zhang, Y.; Bao, J.; Zhang, Y.; Deng, H. Human activity recognition based on motion sensor using u-net. *IEEE Access* **2019**, *7*, 75213–75226. [[CrossRef](#)]
62. Zhang, H.; Xiao, Z.; Wang, J.; Li, F.; Szczerbicki, E. A novel IoT-perceptive human activity recognition (HAR) approach using multihead convolutional attention. *IEEE Internet Things J.* **2019**, *7*, 1072–1080. [[CrossRef](#)]
63. Fan, C.; Gao, F. Enhanced human activity recognition using wearable sensors via a hybrid feature selection method. *Sensors* **2021**, *21*, 6434. [[CrossRef](#)]