sustainability

MDPI

Article

A Robust Hybrid Approach for Real-time Physical Activity Recognition Using Body-worn Inertial Sensors

Ahmet Halıcı[1]

[1] Affiliation 1; ahmethalici2004@gmail.com

Received: date; Accepted: date; Published: date

Abstract:

Real-time physical activity recognition (RPAR) is an important task for researchers pursuing to develop a feasible and sustainable model that precisely recognizes and detects human daily life activities. Additionally, with the rapid development of miniaturized on-body sensor systems, they tend to fuse with inertial sensors to acquire meaningful information. However, in an arbitrary environment, it is challenging to deliver a multi-sensor based robust activity recognition system. This research presents a sustainable data-driven method to recognize physical activities in an indoor-outdoor environment. To improve the recognition of human daily life activities, our proposed model includes pre-processing, hybrid features engineering, GMM based feature encoding, and classification which can recognize and monitor physical activities in real-time. The proposed model is evaluated against three benchmark datasets, namely MotionSense, PAMAP-2, and USC-HAD, respectively. Additionally, the proposed model is also compared with other state-of-the-art methods to evaluate the performance of the RPAR system. The experimental results show that the proposed RPAR system achieves significant results and outperformed current state-of-the-art methods with a recognition rate of 89.93%, 91.16%, and 91.08% over MotionSense, PAMAP-2, and USC-HAD datasets, respectively. The proposed robust system has the capability for exploiting human-machine interface fields such as real-time monitoring, healthcare, surveillance systems, and fitness tracking.

Keywords: Inertial sensors; Micro-electromechanical system (MEMS); Gaussian Mixture Model (GMM); Physical activity recognition; Random Forest (RF)

1. Introduction

In recent years, with the rapid development of sensing technologies, human activity recognition (HAR) has become a hot research topic in ubiquitous computing [1-2]. Human activity recognition via body-worn inertial sensors allows elderly people and their carers to monitor and track physical activities to improve and sustain the quality of the social and personal lives of affected people [3]. With the state-of-the-art enhancements in multi-fused inertial sensors more budget-friendly, features and functionalities are emerging [4]. Additionally, multi-fused body-worn inertial sensors such as accelerometers, gyroscopes, and magnetometers have bagged more research attention and trigger the advancements in robust HAR applications in real-world scenarios [5]. These applications include e-healthcare, internet of things (IoT), smart homes environment, military, security and surveillance, fitness tracking assistance, and biofeedback systems [6-10].

In e-healthcare [11], vital data of patients and elderly people are captured. That particular data is transferred in real-time utilizing multiple IoT-based technologies in order to support online support services. Similarly, smart-based homes can provide real-time consistent physical and

emotional support with cognitive capabilities which help the development of abnormal children and improve their learning strength and brain functionality [12]. The IoT has strong potential in military applications such as can be used in modern warfare. Additionally, it can be deployed on land or even in space to keep monitoring continuously. Similarly, security and surveillance systems can monitor and detect risks and uncertain events in daily surroundings and immediately alert the emergency response team (ERT) [13]. With the assistance of wearable inertial sensors can track and capture a vital part of information about the human body, including dietetic habits and fitness activities which help detect and update the user's wellbeing. Finally, biofeedback is a useful technique via which abnormal people train themselves to attain a set of skills based on the usage of sensors. These feedbacks work with virtual reality in order to quantify body functions such as blood pressure and heart rate. These feedback systems are also useful in stress-relieving techniques [14]. The aim of exploiting body-worn sensors is to deliver real-time continuous monitoring and recognition of motion patterns.

In essence, motion data from multiple dimensions are acquired and examined to learn and recognize different types of postures and movements. These movements can be analyzed by using sophisticated body-worn inertial sensors that can be feasibly attached at different locations of the body in order to acquire vital information about the body for monitoring and detecting the abnormality in human motion patterns in the indoor and outdoor environment [15]. However, body-worn sensors possess many advantages in detecting human physical activities, but due to inconsistent and unreliable contextual information in human motion patterns [16], breaks during data collection, and other constraints like switching activities and resting, etc. In this context, a robust model is needed that can precisely monitor and detect complex human activities.

Micro-electromechanical system (MEMS) based sensors have reshaped the daily lives in healthcare and for daily comfort. These sensors are not facilitating the greater learning of static and dynamic human motion patterns but with the help of multi-fused sensors such as accelerometers, gyroscopes, and magnetometers which combine and extract significant information on the human body in multi-dimensional space [17-19]. Between them, an inertial sensor that comprises triaxial accelerometers, triaxial gyroscopes, and triaxial magnetometers. Like, accelerometers can detect acceleration in 3D space. Gyroscopes can easily measure the rate of variation in the angular motion of a human body. Finally, magnetometers can provide direction information and calculate the variation in magnetic fields at a particular location. However, in order to enhance the performance and functionalities of each sensor, they are integrated into individual units known as inertial measurement units (IMU). The main advantage of using IMU sensors is to provide real-time continuous monitoring of human daily routine activities. The combination of these characteristics makes enhanced classification results which are commonly difficult to achieve with an individual sensor [20].

This paper deals with the real-time physical activity recognition (RPAR) system via multi-fused inertial sensors. To achieve this goal, a robust hybrid approach is proposed to improve the classification process and to achieve a good recognition rate. Initially, multi-fused IMU sensor streaming data is processed with filtering techniques such as a Gaussian filter is utilized to lower the noise ratio from the acquired data. Then different techniques are adopted to extract features from filter data. The features comprise time, frequency, wavelet, and time-frequency domain features respectively. However, due to the high dimensionality of feature vector representation, we optimize the extracted descriptors via GMM based feature encoding method. Finally, a random forest (RF) classifier is utilized to optimize feature vectors to recognize and classify human physical activities with better performance as compared with state-of-the-art methods. Additionally, the proposed system is evaluated against three public benchmark datasets named MOTIONSENSE, Physical Activity Monitoring for Aging People (PAMAP-2), and USC-HAD. The main contributions of our work are as follows:

- Multi-fused feature extraction techniques are utilized from time, frequency, wavelet, and time-frequency domains that enhance the feature optimization process of human daily and physical activities.

- Hybrid features from four main domains are then optimized with a GMM based feature encoding method and random forest classifier that provides more contextual information and classifies activities.
- Detailed comparative evaluation is performed in this research on three public benchmark datasets that provide the significant performance. Moreover, experimental results show a better recognition accuracy for the proposed system when compared with other state-of-the-art techniques.

The remainder of this article is organized as follows. Section 2, briefly presents the literature review of human physical activity recognition. Section 3 addresses the proposed HAR methodology that comprises time, frequency, wavelet, and time-frequency domain features, GMM feature encoding optimizer, and the random forest (RF) classifier. Section 4 analyzes the experimental setting and provides a comparison of the proposed system on three public benchmark datasets. Finally, Section 5 reports the conclusion and future outcomes.

2. Related Work

Body-worn inertial sensor-based HAR has been an emerging research field. Various researchers and analysts have applied varied approaches to detect and recognize physical activity recognition in real-time via different sensors, including body-worn inertial sensors and camera-based sensors. In this section, different aspects have been analyzed in order to improve the performance of HAR-based systems, such as raw data acquisition, feature extraction, feature selection, and classification. Also, different types of sensors have been utilized to enhance the recognition rate and are applicable to various application scenarios. In general, we have grouped different sensing technologies into two categories: vision-based sensors and body-worn inertial-based sensors approach, respectively.

2.1 Analysis based on Vision Sensors

In HAR systems, vision sensing technology mainly involves a camera or video to detect and monitor different human behaviours. Even though vision sensors can provide better recognition and scenes understanding, however, the use of the camera is limited in different indoor and outdoor environments such as illumination variations, occlusions, and even privacy issues. In [21], Sharif et al. presented a framework for HAR-based systems. Their proposed system is distributed into two main steps. Initially, multiple human subjects are detected in the video frame by augmenting EM segmentation and novel uniform. Secondly, multiple descriptors such as texture and shape are extracted from the frames. Then, joint entropy-PCA and a novel Euclidian distance method are applied for optimal feature selection along with the SVM classifier to analyze physical activity recognition in real-time. Additionally, the proposed system is evaluated against multiple datasets. The proposed system shows better results in small datasets. In order to make the proposed system more robust and effective, the database dimension should be increased. In [22], Cui et al. proposed an action recognition-based system. They introduced an end-to-end action analysis system that fused LSTM and CNN in order to deliver two tasks, i.e., human identification and action recognition, which is based on skeleton information. Firstly, geometric descriptors are extracted from local, global, and comprehensive cues-associated data to present action robustly and efficiently. Also, this work is tested against three datasets, and the results have shown better action recognition and human detection.

Wang et al. [23] introduced a probability-based graphical model in order to monitor continuous human motion recognition. They split the system into sub-structured transition and discriminative boundary components. Additionally, they capture the complication of continuous activity segmentation and recognition. However, these systems have been distinguished to work

offline only. In [24], Ji et al. proposed a multi-stage stochastic fusion method that was adopted to deal with the existing issues in human interaction recognition (HIR) systems which comprises indistinct cues representation. In addition, this issue was present only due to the imprecise human body segmentation. However, the proposed work not involve the congenital properties of the interaction. In [25], Wu et al. analyzed a hierarchical scheme to classify multi-view activities in the indoor environment. Using multiple visual descriptors and learning methods, they implemented different fusion techniques including decision methods, spatial and temporal features.

In [26], Liu et al. focused on using RGB-D cameras, author designed human activity monitoring (HAM) in the healthcare environment. They extracted different body movements, static postures, and dynamic actions. The support vector machine (SVM) is used to detect human motion activities. Espinoso et al. in [27], developed a fall detection system with the assistance of multiple vision sensors. They proposed a model with multiple descriptors, including the optical flow method with fixed-time windowing. Finally, a 2D convolution neural network is used to detect abnormal events such as falls. However, the proposed model is evaluated against only one dataset. Anithaa et al. in [28], proposed image processing techniques for human action recognition. Initially, Laplace smoothing transform (LST) is used as a descriptor for HAR. The LST extracted features give the frequency component of the image. These components are later used with the k-nearest neighbour (KNN) classifier to recognize human actions. The limitation of this work is the addition of a KNN classifier which performs poorly with high dimension data.

2.2 Analysis based on Body-worn Inertial Sensors

In [29], Lenardis et al. introduced a wearable inertial sensor-based ADL recognition system to extract multiple features in the wavelet domain. These features include Hilbert-Huang transform (HHT) and empirical mode decomposition (EMD) based approaches that divide inertial sequences into IMF components. These components help to detect human motion patterns. However, due to the misplacement of sensors, missing data is not handled in this paper. Due to the mishandling of streaming data, statistical computations are required to deal with missing data in order to improve recognition accuracy. In [30], Roy et al. proposed a multi-descriptors-based technique to recognize the eight human daily living activities. They acquired data from magnetic IMU sensors which comprise accelerometers, gyroscopes, and magnetometers. In [31], Wan et al. utilized ambient and body-worn sensors, they presented a hybrid approach to detect human activities of daily living in a smart-home environment. The major contribution of this work is to utilize the spatiotemporal constraints in order to improve the recognition accuracy.

In [32], Moriya et al. developed a real-time activity monitoring system with the help of data segmentation strategies. To facilitate dynamic segmentation, sensor and time correlations are introduced. Multiple machine learning approaches have been tested against two datasets which show better recognition rates. The main limitation of this work is that it is restricted to certain activities. Wan et al in [33], developed a real-time activity monitoring system with the help of data segmentation strategies. To facilitate dynamic segmentation, sensor and time correlations are introduced. Multiple machine learning approaches have been tested against two datasets which show better recognition rates. The main limitation of this work is that it is restricted to certain activities. Chen et al. [34] proposed an ensemble of LSTM-RNNs models is fused with scene information to evaluate the accelerometer and gyroscope sensor readings. This method was only tested against one dataset.

In [35], Ceron et al. introduced multiple learning algorithms to recognize the twenty-four ADLs. It involves a diverse set of sensors, including binary sensors, proximity, smart floor sensors, and accelerometer. The highest recognition rate is attained by AdaboostM1 as compared to other state-of-the-art methods. In [36], the author developed a modified weight SVM to detect life-log activities in an indoor environment. Furthermore, multiple problems are handled during the

implementation of algorithms which comprise redundant signal features and misbalance of classes. Their proposed approach is based on the fusion of PCA, weighted SVM, and LDA. To solve the aforementioned issues, weighted SVM is used for every class in the dataset. In [37], Abdine et al. exploited five classifiers including Random Forest, SVM, LMT, Logistic Regression, and Multilayer Perceptron to evaluate the performance. In addition, the fusion of classifiers has shown a good recognition rate.

3. Proposed Methods

The overall architecture of our proposed RPAR system is shown in Figure 1. The proposed RPAR system recognizes 9-Degrees of Freedom (DoF) of human daily activities via three body-worn IMU sensors. These IMU sensors comprise a total of 9-axis varied signal values from accelerometers, gyroscopes, and magnetometers. Initially, inertial signals acquired from IMU sensors are filtered through Gaussian filters in order to eliminate the noise associated with signals. Secondly, the filtered signal is organized into blocks of compatible time duration for feature extraction of signals. Therefore, we proposed a hybrid feature extraction approach that involves descriptors from multiple domains including, time, frequency, wavelet, and time-frequency domains which are quantized using codebook generation for appropriate feature selection. In order to select optimal features from the vector, we utilized GMM feature encoding optimization. Finally, the selected features from GMM are catered to the classifier algorithm to recognize human physical activities in real-time.

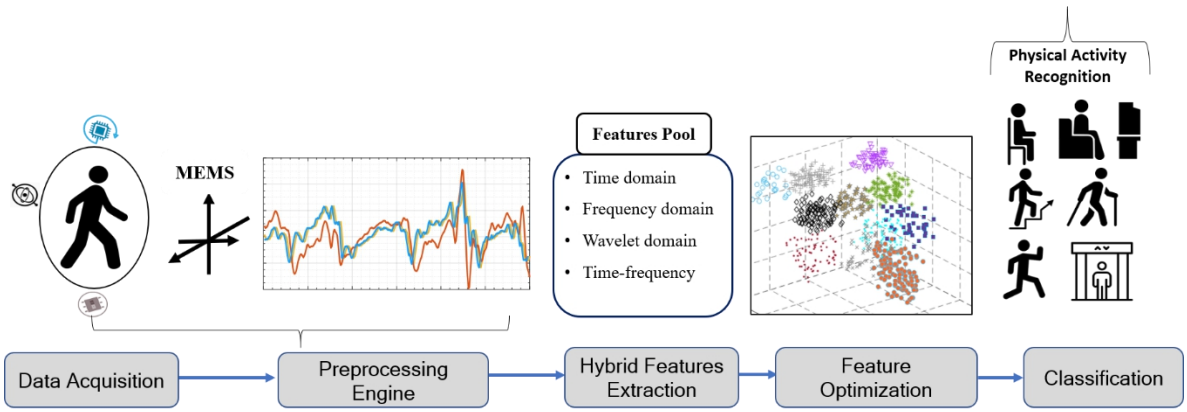


Figure 1. Block diagram of the proposed model

3.1 Preprocessing Engine

Firstly, we acquire inertial signal data from three IMUs for further signal analysis. As inertial sensors are highly prone to even the least amounts of random noise which may affect the signal shape. Therefore, we processed the noisy inertial signal data with a gaussian filter to maintain the shape of the data. The Gaussian filter prevents the outlier and unwanted frequency components without losing the original data. Additionally, the gaussian filter acts as a low pass filter which also smoothens the signal and restores the inertial signals to closely relevant signals. The filtered and noisy inertial components of the Gaussian filter of the inertial signal are shown in Figure 2.

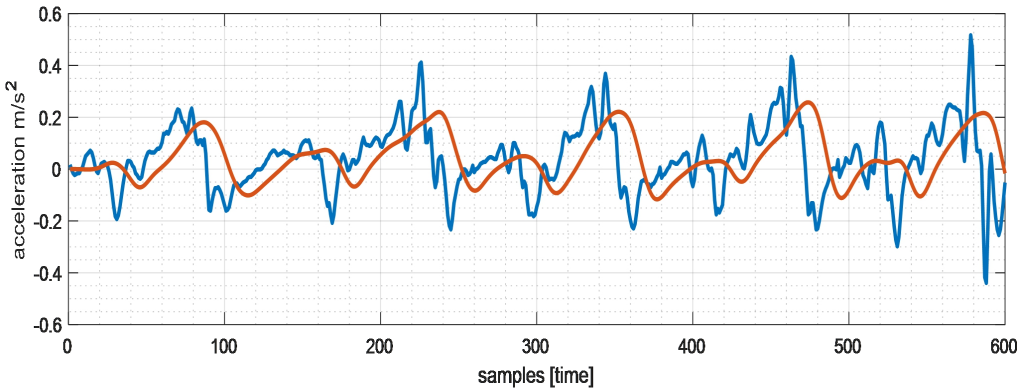


Figure 2. Signal filtering for body-worn inertial signal in the proposed RPAR model

3.2 Hybrid Features Extraction

After the signal acquisition, feature extraction is an important phase in all classification methods, where the significance lies in displaying information with a relevant and significant set of descriptors that balance the whole scenario. In this step, we proposed a hybrid feature model which comprises four major domains such as time, frequency, wavelet, and time-frequency domain descriptors. Algorithm 1 elaborates the hybrid descriptors model.

Algorithm 1 Human Physical Activities Feature Extraction

Input: S_1 **; Output:** F_1
/* $S_{1:}$ IMU data (accelerometer (x,y,z), gyroscope (x,y,z), magnetometer(x,y,z))*/
/* $F_{1:}$ Feature Vector for real-time physical activity recognition (RPAR) system*/
$F_{1 <-}$ []
Method: RPAR(IMU)
    Hybrid Feature Vector <- []
    Denoised_Data <- Gaussain(IMU)
    Frame_Data(Denoised_Data)
While exit condition not satisfied **do**
SC<-Extract_SpectralCentroid(FramedData)
CA<-Extract_CepstralAnalysis(FramedData)
LZC<- Lempel-Ziv Complexity(FramedData)
WT <- Extract_WaveletTree(FramedData)
SG <- Extract_Spectrogram(FramedData)
SFT <- Extract_ShortFourierTransform(FramedData)
return F_1

3.2.1. Spectral Centroid

In the spectral domain, there are different descriptors, including energy which helps in determining the power of the signal at a certain time and also calculates the average frequency, which is weighted by the amplitude of the spectrum. The centroid of a spectral frame is a measure used in signal processing to distinguish the spectrum. It designates where the spectrum center of mass is located [38]. Eq. (1) shows the mathematical model of rotational angular joint features.

$$SC = \frac{\sum_{a=1}^{N} aF[A]}{\sum_{a=1}^{N} F[A]} \quad (1)$$

where $F[A]$ represents the amplitude accompanying to bin a in FFT spectrum. Figure.7 shows the results of spectral centroid over different signal patterns.

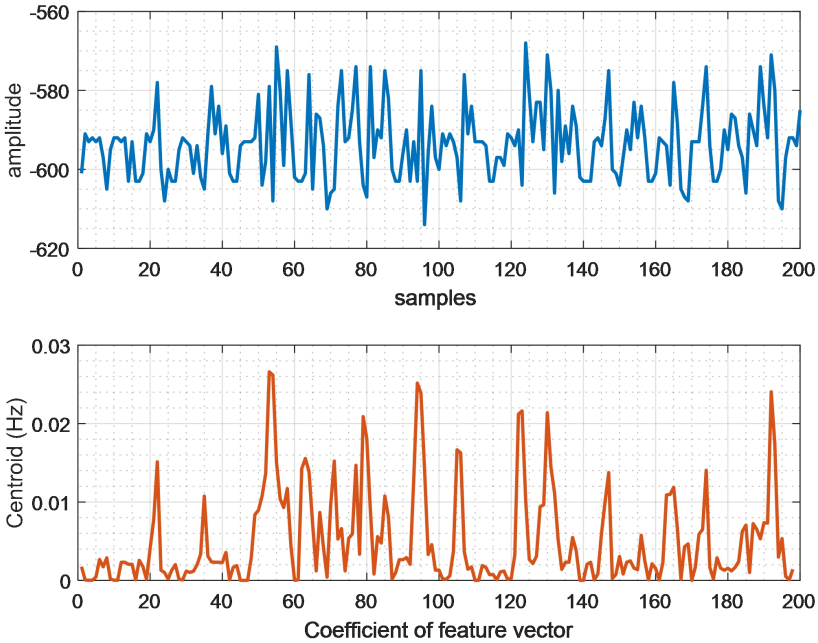


Figure 3. Spectal Centroid descriptor applied over varied signal pattern

3.2.2. Cepstral Analysis

Inertial signal involves motion artifacts during the instant change in the activity. Cepstral analysis [39] is a similar form of signal transformation technique that converts a convolution into an additive correlation which builds it especially conductive for reducing convolutional effects. It has been mostly used in signal processing, such as inertial and speech synthesis. In addition, to mitigate the effects of the multiple paths from the initial signals to the sensors, cepstral analysis helps in modeling the frequency information of the signal which allows us to independently separate the convolution effects. In our model, we have used the cepstral feature which is based on spectral information of the magnitude of the inertial signals. The formation of cepstral analysis using the inertial signal is defined as

$$C[n] = \frac{1}{2\pi}\int_{-\pi}^{\pi} ln\,|Y(e^{ja})|e^{jan}da \quad (2)$$

where $|Y(e^{ja})|$ denotes the spectral magnitude and x[n] is the real cepstral analysis. Figure. 4 shows the cepstral analysis over inertial signal pattern.

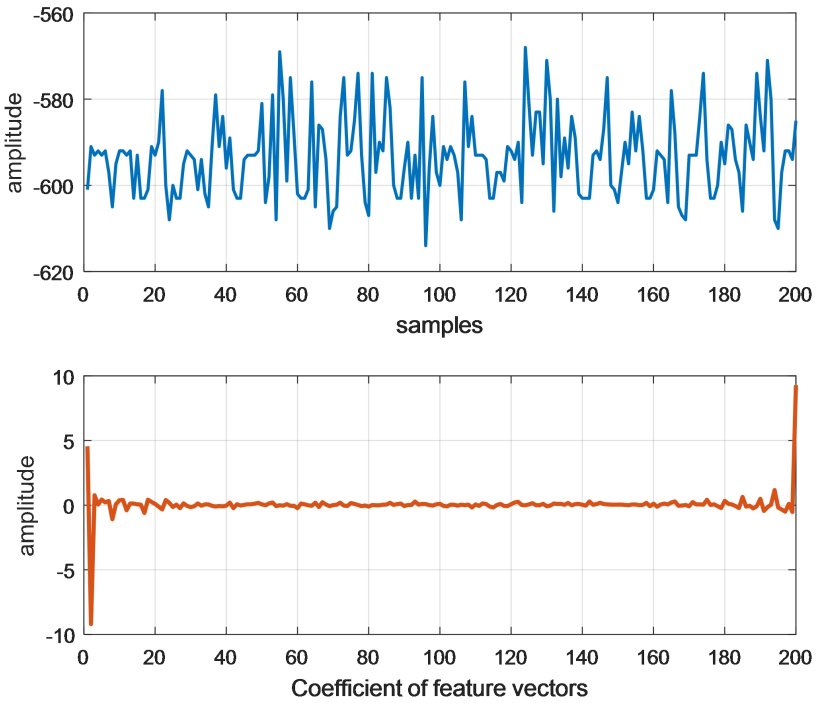


Figure 4. Cepstral Analysis for the inertial signal pattern using the USC-HAD

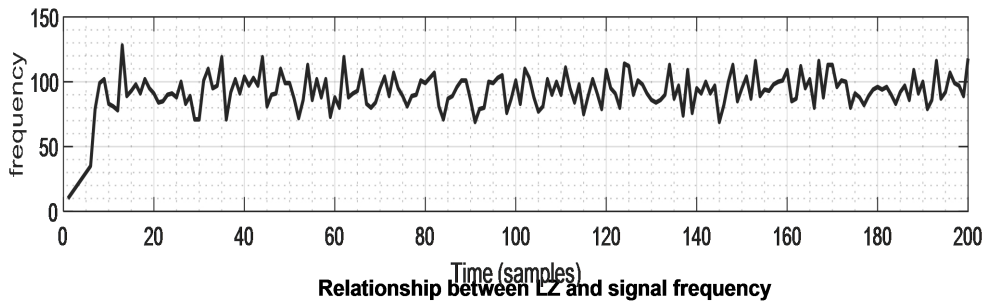
3.2.3. Lempel-Ziv Complexity (LZC)

The Lempel-Ziv Complexity (LZC) method is a symbolic sequence approach based on a measurement of coarse-graining. With LZC, the signal is initially converted into a sequence that comprises a finite number of symbols. In the context of sensor streaming data, the signal is transformed into a binary sequence [40]. In addition, a binary sequence was used with the median threshold Td due to robustness to outliers. The Eq (3) of LZ Complexity with a median threshold can be written as follows:

$$sig(i) = \begin{cases} 0, & if\, s(i) < T_d \\ 1, & if\, s(i) \geq T_d \end{cases}$$

(3)

Figure. 5 shows the relationship between LZ and signal frequency.



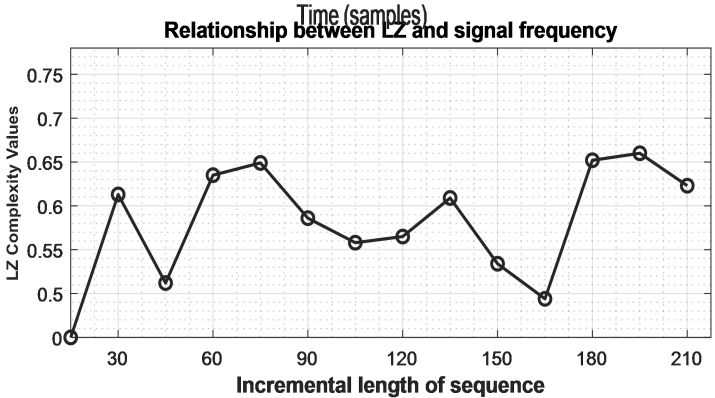


Figure 5. The result of LZ and frequency modulation.

3.2.4. Short-time Fourier Transform (STFT)

The Short-Time Fourier Transform has a significant property that facilitates the analysis of the resultant distribution in both frequency and time domains. Additionally, human activity motion produces time-varying doppler patterns over a limited time period. Due to multiple falls and other activities, the doppler frequencies change rapidly. To deal with this problem, a short-time fourier transform (STFT) is introduced which can be used in different finite time windows [41]. In addition, it schemes signals onto chirp bases. The Eq (4) of STFT can be written as follows:

$$X_a(\omega) = \sum_{n=-\infty}^{\infty} x(n)w(n - aR)e^{-j\omega n} \quad (4)$$

Where x(n) represents the input signal of time n, w(n) denotes the length a window function, and $X_a(\omega)$ is the DTFT of framed data which is centered at aR.

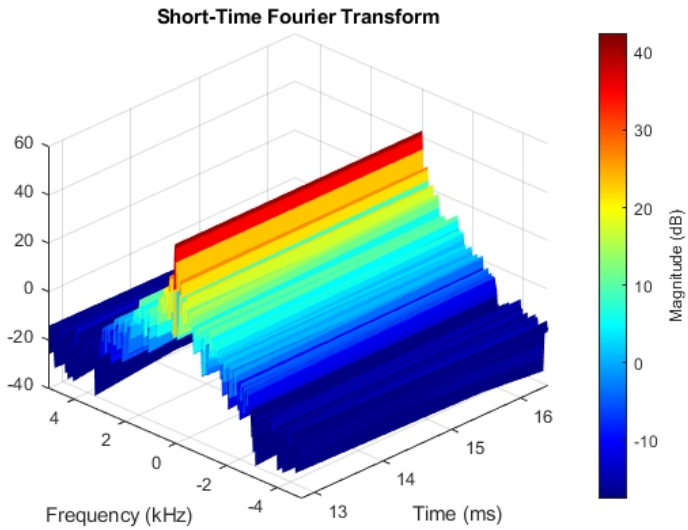


Figure 6. STFT results from the inertial signal over USC-HAD dataset

3.2.5. Wavelet Packet Entropy (WPE)

Wavelet Packet Entropy (WPE) is an effective technique in the time-frequency domain for inertial signals. Initially, WPE split an inertial signal into multiple frequency resolutions, including approximation coefficients and detail coefficients. Additionally, to analyze a particular frequency region, an appropriate tree structure is selected. The mathematical equation (5) of WPE can be defined as:

$$d_a = \begin{cases} d_{0,0}(t) = a(t) \\ d_{i,2j-1}(r) = \sqrt{2}\sum_k h(k) d_{i-1,j}(2r-k) \\ d_{i,2j}(r) = \sqrt{2}\sum_k g(k) d_{i-1,j}(2r-k) \end{cases}$$

(5)

Where h(k) and g(k) denote two filters which are used to extract ACs and DCs, and $d_{i,j}$ represents the reconstruction signals at the ith level and jth node.

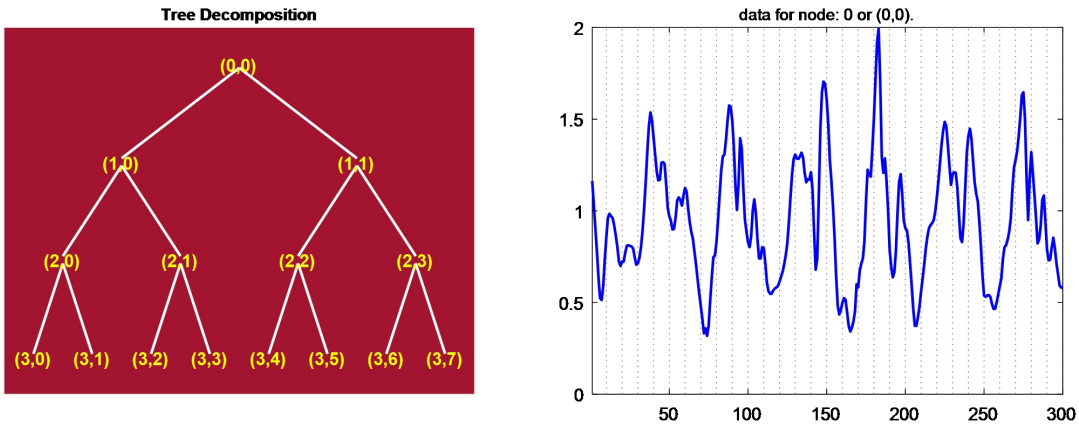


Figure 7. A wavelet decomposition tree for an inertial signal with two level wavlet decomposition

3.3 GMM based Feature Encoding

After combining the hybrid features from all three sensors, the resultant augmented descriptor is further encoded to form a vector that is known as a codebook. The vector representation is unsorted and complexly merged, so there is a need to present it in a simpler and sorted form. Thus, for codebook generation, we applied the fisher encoding method (FVC) which is based on the gaussian mixture model (GMM) [42-43]. Firstly, GMM is applied to evaluate the mean and covariance of each class. Based on these values, each class generates its cluster. Therefore, the probability density function of each class of the cluster represents d dimensionality vector X which is shown in Equation (6).

$$pr(X;\theta) = \sum_{i=1}^{i} w_i N\left(X|\mu_i, \sum_i\right)$$

(6)

where $\theta = \{w_i, \mu_i, \sum_i \mid i = 1,2,3,\ldots,i\}$, w_i is the ith weight of the Gaussian component, i represents the total number of clusters, mean value is shown as μ_i, the $\sum_i$ represents the covariance matrix and d distribution of the dimensional gaussian is defined by N.

In the context of GMM, the next step is fisher vector encoding (FVC) that is performed on fused features. $X = \{x_t, s = 1,2,3\ldots, S\}$ is a input feature set, while the gradient of log probability ∇_θ of X possessing GMM parameter θ is represented as:

$$F_X = \frac{1}{S}\nabla_\theta log\, p(X;\theta)$$

(7)

Finally, all calculated gradient components are fused to form d-dimensional encoded feature vectors. In addition, FVC mitigates the intra-cluster gap and gradually increases the inter-cluster space, that gives a discriminated cluster of each class.

$$\mu_k = \frac{1}{T\sqrt{w_k}} \sum_{t=1}^{T} q_t(k) \frac{x_t - \mu_k}{\sigma_k} \quad (8)$$

Figure (8) shows the formation of cluster by each activity FVC over USC-HAD dataset.

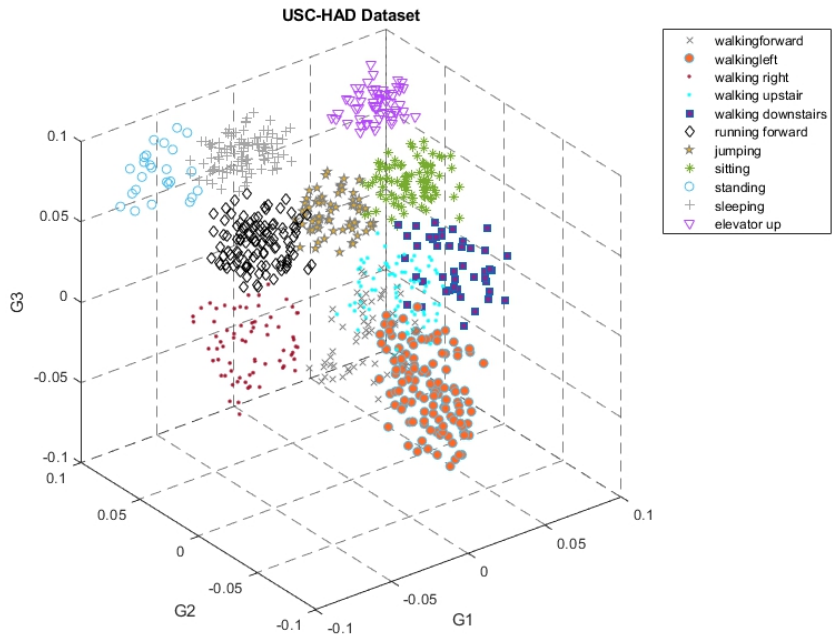


Figure 8. Codebook generation via FVE based GMM model applied to the USC-HAD dataset.

3.4 Random Forest Classification

After the pre-classification step, we evaluated our proposed method against three benchmark datasets that involve human physical activities for diverse and multiple classes. The relevant and optimized descriptors of GMM are classified by a novel classifier Random Forest (RF), which is an ensemble-learning algorithm for regression and data classification [44]. Additionally, the random forest comprises a unique variation of bagged trees, which is a method to generate a training dataset. Bagging selects the examples from training data of all three physical activities datasets including MOTIONSENSE, PAMAP-2, and USC-HAD. A model for all samples is drawn which has been utilized for classification based decisions. Then, all the decisions are combined on the basis of the majority votes to get the final decision. Also, the main advantage of random forest is that it can evaluate the score of each descriptor to supervise the learning impact of each certain feature which is based on its prediction.

$$\hat{f} = \frac{1}{B} \sum_{b=1}^{B} f_b(y') \quad (9)$$

Equation (9), displays the predictions for unseen examples y'. It can be calculted by averaging the predictions from all the the distinct trees on y'. Figure 8 shows the random forest algorithm for human physical activities recognition.

4. Experimental Results and Analysis

All experiments are performed on a system equipped with Intel Core i7-6500U 2.50GHz computation power, 8GB RAM, 2 GB dedicated graphics card Nvidia 940M including x64 bit-based Windows 10 home and exploiting via Google Colab and MATLAB tools. Additionally, a platform has been developed to assess the performance of the proposed RPAR model executed on three benchmark datasets, including, MOTIONSENSE, PAMAP-2, and USC-HAD datasets.

4.1 Datasets Description

First Dataset – The MotionSense dataset [45] is a publicly available open-source dataset that involves 6-Degrees of Freedom (DoF) IMU (accelerometer and gyroscope) sensing data collected from the smartphone iPhone 6s. The smartphone is kept in the subject's front pocket. A diverse group of 24 subjects (14 males and 10 females) performed six static and dynamic activities in different manners, including, sitting, standing, walking, downstairs, jogging, going upstairs.

Second Dataset – The physical activity monitoring for aging people (PAMAP-2) [46] is publicly available open dataset at the University of California, UCI machine learning repository. The PAMAP-2 dataset embodied data from three Colibri wireless IMU sensors integrated with triaxial accelerometers, gyroscopes, and magnetometers positioned on the participant's chest, ankle and wrist regions to perform 18 physical activities. We evaluated twelve activities of this dataset comprising, sitting, standing, walking, Nordic walking, running, cycling, lying down, rope jumping, ironing, ascending stairs, descending stairs and house cleaning. Additionally, this dataset includes repetitive physical activities distinctive for RPAR systems intended to analyze complex motion patterns.

Third Dataset – The University of Southern California Human Activity Dataset (USC-HAD) [47], incorporated with an IMU based MotionNode sensor-fused tri-axial accelerometer and gyroscope sensors sampled at 100 Hz. These sensors are positioned at front right hip. Additionally, the dataset has physical activity data captured with 14 participants performing twelve 12 dynamic and static activities including, sitting, standing, sleeping, walking forward, jumping up, running forward, walking upstairs and downstairs, walking left and right, elevator up and down.

4.2 Experimental Setups and Evaluation

In order to observe the validation performance of the RPAR system, we utilized the leave one-subject-out (LOSO) cross-validation method. In addition, we evaluate the performance of the random forest classifier by providing the optimal and relevant features of time, frequency, time-frequency, and wavelet exploiting the MotionSense, PAMAP-2, and USC-HAD datasets. Figure 9 depicts the confusion matrix over the MotionSense dataset for 6 static and dynamic activities, where a mean accuracy of 89.93 % was attained. Figure 10 illustrates the confusion matrix of the PAMAP-2 dataset for 12 human physical activities, where a recognition rate of 91.16% was achieved. Figure 11 represents a recognition rate of 91.08 % over 12 life-log physical human activities in the USC-HAD dataset.

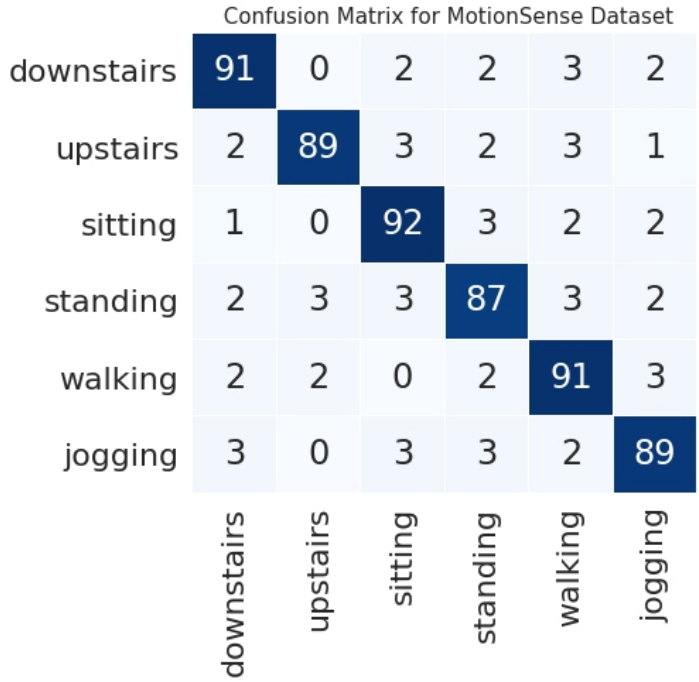


Figure 9. Confusion matrix of 6 diffeent physical activities on the MotionSense dataset using Random Forest Classifier

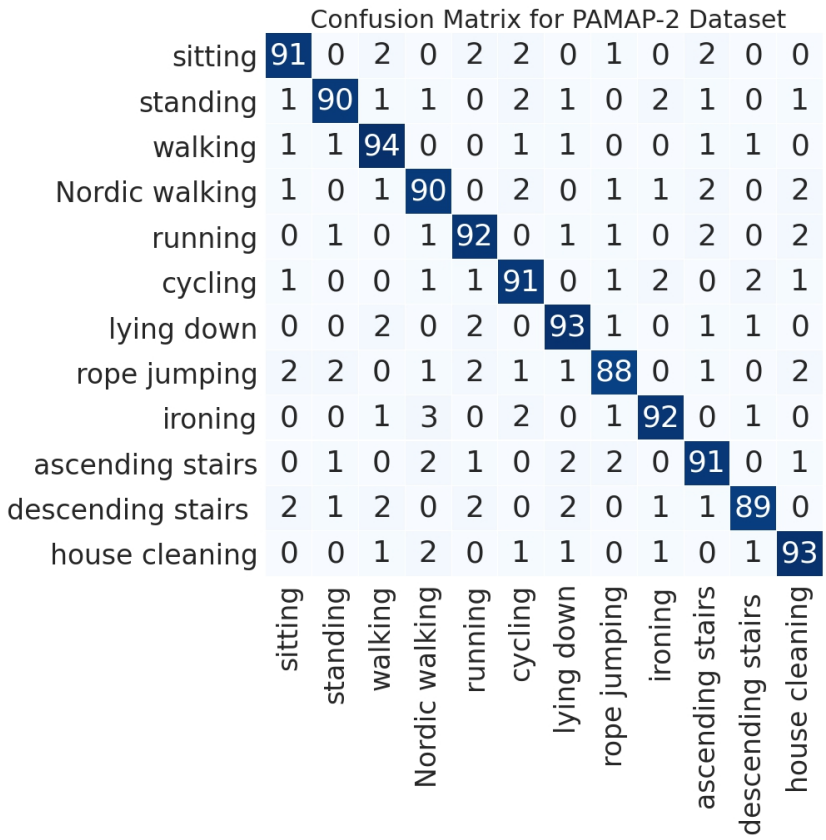


Figure 10. Confusion matrix of 12 different physical and locomotion activities on the PAMAP-2 dataset using Random Forest Classifier

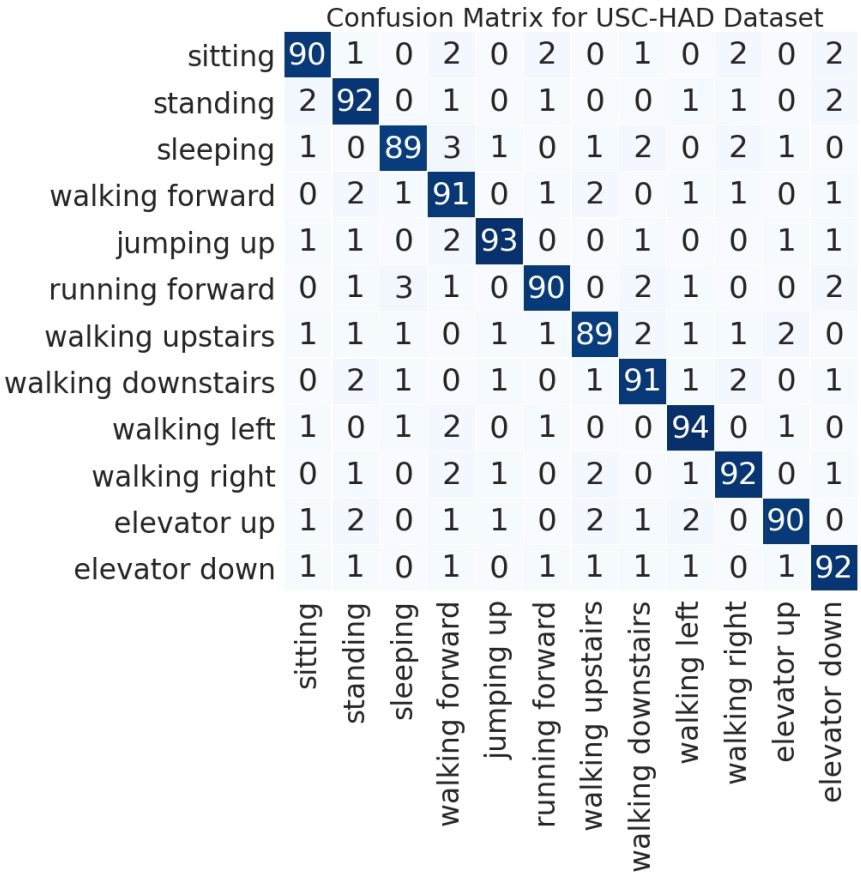


Figure 11. Confusion matrix of 12 life-log physical activities on the USC-HAD dataset using Random Forest Classifier

Table 1 depicts the comparison results between our proposed methodology and different state-of-the-arts methods. In Table 2-4 we validated the performance of proposed system by comparing it with two novel classifiers, including, Artificial Neural Network (ANN), and Support Vector Machine (SVM) via precision using equation (10), recall using equation (11), and F-measure using equation (12) of all classes to detect complex body postures and movements in these datasets: MotionSense, PAMAP-2, USC-HAD.

$$\text{Precision} = \frac{\text{True Positive}}{\text{True Positive+False Positive}} \tag{10}$$

$$\text{Recall} = \frac{\text{True Positive}}{\text{True Positive+False Negative}} \tag{11}$$

$$\text{F1 score} = \frac{2.\,\text{Precision.Recall}}{\text{Precision+Recall}} \tag{12}$$

TABLE I. Comparison of recognition rate (accuracies) of the proposed method with state-of-the-art methods over MotionSense, PAMAP-2 and USC-HAD datasets.

Methods	MotionSense(%)	PAMAP-2 (%)	USC-HAD (%)
Bi-Directional LSTM [48]		64.10%	

Multiclass Adaboost [49]		77.78	
BERT Model [50]	79.86	-	-
Deep Convolutional and LSTM RNN [51]		87.50	-
Bag of features [52]			70
Random Forest [53-54]		90.11	90.70
Multi-fused features [55]	88.25		
Symbolic Approximation [56]			84.3
Proposed Method	**89.93**	**91.16**	**91.08**

Table 2. Comparison of evaluation metrics (precision, recall, and F1 score) of proposed method over MotionSense dataset

Methods	**Random Forest**			**ANN**			**SVM**		
Activities	Precision	Recall	F1 score	Precision	Recall	F1 score	Precision	Recall	F1 score
M1	0.901	0.910	0.905	0.887	0.890	0.888	0.786	0.792	0.788
M2	0.946	0.890	0.917	0.901	0.878	0.889	0.820	0.798	0.808
M3	0.893	0.920	0.906	0.789	0.803	0.795	0.763	0.783	0.772
M4	0.878	0.870	0.874	0.767	0.748	0.757	0.740	0.696	0.717
M5	0.875	0.910	0.892	0.756	0.769	0.762	0.734	0.739	0.736
M6	0.898	0.890	0.894	0.785	0.778	0.781	0.754	0.687	0.718

Table 3. Comparison of evaluation metrics (precision, recall, and F1 score) of proposed method over PAMAP-2 dataset

Methods	**Random Forest**			**ANN**			**SVM**		
Activities	Precision	Recall	F1 score	Precision	Recall	F1 score	Precision	Recall	F1 score
P1	0.919	0.910	0.914	0.870	0.865	0.867	0.847	0.839	0.842
P2	0.937	0.900	0.918	0.898	0.779	0.834	0.856	0.827	0.841
P3	0.903	0.940	0.921	0.858	0.867	0.862	0.839	0.842	0.840
P4	0.891	0.900	0.895	0.874	0.876	0.874	0.858	0.848	0.852
P5	0.901	0.920	0.910	0.859	0.873	0.865	0.821	0.833	0.826
P6	0.892	0.910	0.900	0.897	0.898	0.897	0.861	0.865	0.862
P7	0.911	0.930	0.920	0.865	0.874	0.869	0.843	0.846	0.844
P8	0.916	0.880	0.897	0.871	0.852	0.861	0.849	0.821	0.834
P9	0.929	0.920	0.924	0.891	0.899	0.894	0.851	0.744	0.793
P10	0.892	0.910	0.900	0.876	0.881	0.878	0.848	0.839	0.843
P11	0.936	0.890	0.912	0.895	0.873	0.883	0.852	0.841	0.846
P12	0.911	0.930	0.920	0.866	0.874	0.869	0.826	0.831	0.828

Table 4. Comparison of evaluation metrics (precision, recall, and F1 score) of proposed method over USC-HAD dataset

Methods	**Random Forest**			**ANN**			**SVM**		
Activities	Precision	Recall	F1 score	Precision	Recall	F1 score	Precision	Recall	F1 score
U1	0.918	0.900	0.909	0.889	0.870	0.879	0.835	0.821	0.827
U2	0.884	0.920	0.901	0.865	0.863	0.864	0.823	0.817	0.819
U3	0.927	0.890	0.933	0.891	0.875	0.882	0.839	0.828	0.833

U4	0.858	0.910	0.878	0.846	0.869	0.857	0.811	0.816	0.813
U5	0.948	0.930	0.934	0.910	0.913	0.911	0.845	0.833	0.838
U6	0.927	0.900	0.918	0.891	0.870	0.880	0.841	0.832	0.836
U7	0.908	0.890	0.918	0.874	0.863	0.868	0.826	0.819	0.822
U8	0.900	0.910	0.890	0.865	0.869	0.867	0.820	0.815	0.817
U9	0.912	0.940	0.916	0.879	0.883	0.881	0.829	0.820	0.824
U10	0.910	0.920	0.910	0.876	0.878	0.877	0.827	0.816	0.821
U11	0.937	0.900	0.913	0.895	0.907	0.900	0.841	0.832	0.836
U12	0.901	0.920	0.915	0.872	0.881	0.876	0.823	0.826	0.824

5. Conclusion and Future Works

This paper presents a robust and sustainable framework for real-time physical activity recognition (RPAR) that can precisely recognize the HAR of three benchmark datasets in an indoor-outdoor environment by catering to the inertial signal via hybrid descriptors. These descriptors include spectral centroid, cepstral analysis, Lempel-Ziv Complexity (LZC), Short-time Fourier Transform (STFT), Wavelet Packet Entropy (WPE) to extract relevant features. Such features are optimized by Gaussian mixture models (GMM) and passed through the Random Forest. In addition, this paper also compares the performance of the random forest classifier optimized by GMM with Support Vector Machines (SVM) and Artificial Neural Network (ANN). The proposed model is evaluated on three challenging benchmark datasets that give an accuracy of 89.93%, 91.16%, and 91.08% over MotionSense, PAMAP-2, and USC-HAD datasets, respectively. The experiment results have also revealed that our proposed RPAR model achieved a significant recognition rate over present state-of-the-art methods.

In future work, we will adopt more features from different domains in order to enhance the efficiency of our hybrid features. Additionally, we are planning to create self-annotated datasets for healthcare utilizing via multi-modal sensors.

Funding: This research received no external funding.

Conflicts of Interest: The author declares no conflict of interest.

References

1. Bulling, A.; Blanke, U.; Schiele, B. A Tutorial on Human Activity Recognition Using Body-worn Inertial Sensors. *ACM Comput. Surv.* **2014**, 46, 1–33.
2. Wang, Y.; Cang, C.; Yu, H. A review of sensor selection, sensor devices and sensor deployment for wearable sensor-based human activity recognition systems. In Proceedings of the 10th International Conference on Software, Knowledge, Information Management & Applications, Chengdu, China, 15–17 December 2016; pp. 250–257.
3. Tingting, Y.; Junqian, W.; Lintai, W.; Yong, X. Three-stage network for age estimation. *CAAI Trans. Intell. Technol*. **2019**, 4, 122–126.
4. Alsheikh, M.A.; Selim, A.; Niyato, D.; Doyle, L.; Lin, S.; Tan, H.P. Deep activity recognition models with triaxial accelerometers. In Proceedings of the Workshops at the Thirtieth AAAI Conference on Artificial Intelligence, Phoenix, AZ, USA, 12–13 February 2016.
5. Ahmed, N.; Rafiq, J.I.; Islam, M.R. Enhanced Human Activity Recognition Based on Smartphone Sensor Data Using Hybrid Feature Selection Model. *Sensors* **2020**, 20, 317.
6. Gochoo, M.; Tahir, S.B.; Jalal, A.; Kim, K. Monitoring Real-Time Personal Locomotion Behaviors Over Smart Indoor-Outdoor Environments Via Body-Worn Sensors. *IEEE Access*. **2021**, 9, 70556-70570.
7. Chatterjee, R.; Maitra, T.; Hafizul Islam, S.; Hassan, M.M.; Alamri, A.; Fortino, G. A novel machine learning based feature selection for motor imagery EEG signal classification in Internet of medical things environment. *Future Gener. Comput. Syst*. **2019**, 98, 419–434.

8. Gravina, R.; Alinia, P.; Ghasemzadeh, H.; Fortino, G. Multi-sensor fusion in body sensor networks: State-of-the-art and research challenges. *Inf. Fusion*, **2017**, 35, 68–80.
9. Uddin, M.Z.; Hassan, M.M. Activity Recognition for Cognitive Assistance Using Body Sensors Data and Deep Convolutional Neural Network. *IEEE Sens. J.* **2018**, 19, 8413–8419.
10. Tahir, S.B; Jalal, A; Kim, K. Daily life Log Recognition based on Automatic Features for Health care Physical Exercise via IMU Sensors. In Proceedings of the 18th International Bhurban Conference on Applied Sciences and Technology (IBCAST), Islamabad, Pakistan, 12–16 January 2021.
11. Zhu, C.; Miao, D. Influence of kernel clustering on an RBFN. *CAAI Trans. Intell. Technol*. **2019**, 4, 255–260.
12. Kantoch, E. Human activity recognition for physical rehabilitation using wearable sensors fusion and artificial neural networks. In Proceedings of the 2017 Computing in Cardiology (CinC), Rennes, France, 24–27 September 2017; pp. 1–4.
13. Qin, Z.; Zhang, Y.; Meng, S.; Qin, Z.; Choo, K.-K.R. Imaging and fusing time series for wearable sensor-based human activity recognition. *Inf. Fusion*. **2020**, 53, 80–87.
14. Li, R.; Lai, D.T.H.; Lee, W. A survey on biofeedback and actuation in wireless body area networks (WBANs). *IEEE Rev. Biomed. Eng*. **2017**, 10, 162–173.
15. Yin, J.; Yang, Q.; Pan, J.J. Sensor-based abnormal human-activity detection. *IEEE Trans. Knowl. Data Eng*. **2008**, 20, 1082–1090.
16. Osterland, S.; Weber, J. Analytical analysis of single-stage pressure relief valves. *Int. J. Hydromechatronics* **2019**, 2, 32–53.
17. Casilari, E.; Álvarez-Marco, M.; García-Lagos, F. A Study of the Use of Gyroscope Measurements in Wearable Fall Detection Systems. *Symmetry* **2020**, 12, 649.
18. Bonato, P. Wearable sensors/systems and their impact on biomedical engineering. *Eng. Med. Biol. Mag*. **2003**, 22, 18–20.
19. Rodríguez-Rodríguez, I.; Rodríguez, J.-V.; Elizondo-Moreno, A.; Heras-González, P. An Autonomous Alarm System for Personal Safety Assurance of Intimate Partner Violence Survivors Based on Passive Continuous Monitoring through Biosensors. *Symmetry* **2020**, 12, 460.
20. Zhang, M.; Sawchuk, A.A. A feature selection-based framework for human activity recognition usingwearable multimodal sensors. In Proceedings of the 6th International Conference on Body Area Networks, Beijing, China, 7–8 November 2011; pp. 92–98.
21. Sharif, M.; Khan, M.A.; Akram, T.; Younus, M.J.; Saba, T.; Rehman, A. A framework of human detection and action recognition based on uniform segmentation and combination of Euclidean distance and joint entropy-based features selection. *EURASIP J. Image Video Process*. **2017**, 89.
22. Cui, R.; Hua, G.; Zhu, A.; Wu, J.; Liu, H. Hard sample mining and learning for skeleton-based human action recognition and identification. *IEEE Access* **2018**, 7, 8245–8257.
23. Wang, Z.; Wang, J.; Xiao, J.; Lin, K.H.; Huang, T. Substructure and boundary modeling for continuous action recognition. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Providence, RI, USA, 16–21 June 2012; pp. 1330–1337.
24. Ji, X.; Wang, C.; Ju, Z. A New Framework of Human Interaction Recognition Based on Multiple Stage Probability Fusion. *Appl. Sci*. **2017**, 7, 567.
25. Wu, C.; Khalili, A.H.; Aghajan, H. Multiview activity recognition in smart homes with spatio-temporal features. In Proceedings of the Fourth ACM/IEEE International Conference on Distributed Smart Cameras, Atalanta, GA, USA, 31 August–4 September 2010; pp. 142–149.
26. Liu, X.; Liu, L.; Simske, S.J.; Liu, J. Human Daily Activity Recognition for Healthcare Using Wearable and Visual Sensing Data. In Proceedings of the IEEE International Conference on Healthcare Informatics (ICHI), Chicago, IL, USA, 4–7 October 2016; pp. 24–31.
27. Espinosa, R.; Ponce, H.; Gutiérrez, S.; Martínez-Villaseñor, L.; Brieva, J.; Moya-Albor, E. A vision-based approach for fall detection using multiple cameras and convolutional neural networks: A case study using the UP-Fall detection dataset. *Comput. Biol. Med*. **2019**, 115, 103520.
28. Anithaa, U.; Narmadhab, R.; Sumanthc, R.D.; Kumarc, D.N. Robust Human Action Recognition System via Image Processing. *Procedia Comput. Sci*. **2020**, 167, 870–877.
29. De Leonardis, G.; Rosati, S.; Balestra, G.; Agostini, V.; Panero, E.; Gastaldi, L.; Knaflitz, M. Human Activity Recognition by Wearable Sensors: Comparison of different classifiers for real-time applications. In Proceedings of the 2018 IEEE International Symposium on Medical Measurements and Applications (MeMeA), Rome, Italy, 11–13 June 2018; IEEE: Piscataway, NJ, USA, 2018; pp. 1–6.

30. Roy, N.; Misra, A.; Cook, D. Ambient and smartphone sensor assisted ADL recognition in multi-inhabitant smart environments. *J. Ambient Intell. Hum. Comput.* **2016**, *7*, 1–19.
31. Wan, J.; O'Grady, M.J.; O'Hare, G.M.P. Dynamic sensor event segmentation for real-time activity recognition in a smart home context. *Pers. Ubiquitous Comput.* **2015**, *19*, 287–301.
32. Moriya, K.; Nakagawa, E.; Fujimoto, M.; Suwa, H.; Arakawa, Y.; Kimura, A.; Miki, S.; Yasumoto, K. Daily Living Activity Recognition with Echonet Lite Appliances and Motion Sensors. In Proceedings of the 2017 IEEE International Conference on Pervasive Computing and Communications Workshops (PerCom Workshops), Kona, HI, USA, 13–17 March 2017.
33. Wan, J.; O'Grady, M.J.; O'Hare, G.M.P. Dynamic sensor event segmentation for real-time activity recognition in a smart home context. *Pers. Ubiquitous Comput.* **2015**, *19*, 287–301.
34. Chen, W.-H.; Baca, C.A.; Tou, C.-H. LSTM-RNNs combined with scene information for human activity recognition. In Proceedings of the 2017 IEEE 19th International Conference on e-Health Networking, Applications and Services (Healthcom), Dalian, China, 12–15 October 2017; pp. 1–6.
35. Cerón, J.D.; López, D.M.; Eskofier, B.M. Human Activity Recognition Using Binary Sensors, BLE Beacons, an Intelligent Floor and Acceleration Data: A Machine Learning Approach. Proceedings 2018, 2, 1265.
36. Abidine, B.M.H.; Fergani, L.; Fergani, B.; Oussalah, M. The Joint Use of Sequence Features Combination and Modified Weighted SVM for Improving Daily Activity Recognition. Pattern Anal. Appl. 2018, 21, 119–138.
37. Bayat, A.; Pomplun, M.; Tran, D.A. A Study on Human Activity Recognition Using Accelerometer Data from Smartphones. *Proced. Comput. Sci.* **2014**, *34*, 450–457.
38. Liu, S.; Zeng, Z.; Zhang, Y.D.; Fan, T.; Shan, T.; Tao, R. Automatic human fall detection in fractional Fourier domain for assisted living. In Proceedings of the 41st IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP 2016), Shanghai, China, 20–25 March 2016.
39. Li, M.; Rozgic, V.; Thatte, G.; Lee, S.; Emken, B.; Annavaram, M.; Mitra, U.; Spruijt-Metz, D.; Narayanan, S. Multimodal Physical Activity Recognition by Fusing Temporal and Cepstral Information. *IEEE Trans. Neural Syst. Rehabil. Eng.* **2010**, *18*, 369–380.
40. Yin, J.; Xu, M.; Zheng, H. Fault diagnosis of bearing based on Symbolic Aggregate approXimation and Lempel-Ziv. *Measurement* **2019**, *138*, 206–216.
41. Behnam, H.; Sheikhani, A.; Mohammadi, M.R.; Noroozian, M.; Golabi, P. Analyses of EEG background activity in Autism disorders with fast Fourier transform and short time Fourier measure. In Proceedings of the 2007 International Conference on Intelligent and Advanced Systems, ICIAS 2007, Kuala Lumpur, Malaysia, 25–28 November 2007; pp. 1240–1244.
42. Dutta, A.; Ma, O.; Toledo, M.; Buman, M.P.; Bliss, D.W. Comparing Gaussian Mixture Model and Hidden Markov Model to Classify Unique Physical Activities from Accelerometer Sensor Data. In Proceedings of the IEEE International Conference on Machine Learning and Applications (ICMLA), Anaheim, CA, USA, 18–20 December 2016.
43. Khan, M.H.; Farid, M.S.; Grzegorzek, M.A. Generic codebook based approach for gait recognition. Multimed. Tools Appl. 2019, 78, 35689–35712.
44. Pal, M. Random forest classifier for remote sensing classification. *Int. J. Remote Sens.* **2005**, *26*, 217–222.
45. Malekzadeh, M.; Clegg, R.G.; Cavallaro, A.; Haddadi, H. Mobile Sensor Data Anonymization. In Proceedings of the International Conference on Internet of Things Design and Implementation, Montreal, QC, Canada, 15–18 April 2019
46. Reiss, A.; Stricker, D. Introducing a New Benchmarked Dataset for Activity Monitoring. In Proceedings of the 16th International Symposium on Wearable Computers (ISWC), Newcastle, UK, 18–22 June 2012; pp. 108–109.
47. Zhang, M.; Sawchuk, A.A. USC-HAD: A daily activity dataset for ubiquitous activity recognition using wearable sensors. In Proceedings of the 2012 ACM Conference on Ubiquitous Computing (UbiComp) Workshop on Situation, Activity and Goal Awareness (SAGAware), Pittsburgh, PA, USA, 5–8 September 2012; pp. 1036–1043.
48. Heroy, A.M.; Gill, Z.; Sprague, S. Stationary exercise classification using IMUs and deep learning. *SMU Data Sci. Rev.* **2020**, 3, 1.
49. Reiss, A.; Stricker, D.; Hendeby, G. Confidence-based multiclass AdaBoost for physical activity monitoring. In Proceedings of the 17th Annual International Symposium on International Symposium on Wearable Computers, Zurich, Switzerland, 8–12 September 2013; pp. 13–20.

50. Haresamudram, H.; Beedu, A.; Agrawal, V.; Grady, P.L.; Essa, I. Masked Reconstruction Based Self-Supervision for Human Activity Recognition. In Proceedings of the 24th annual International Symposium on Wearable Computers, Cancun, Mexico, 12–16 September 2020.
51. Murahari, V.S.; Plötz, T. On attention models for human activity recognition. In Proceedings of the 2018 ACM International Symposium on Wearable Computers, Singapore, 8–12 October 2018; pp. 100–103.
52. Jordao, A.; Nazare Jr, A.C.; Sena, J.; Schwartz, W.R. Human activity recognition based on wearable sensor data: A standardization of the state-of-the-art. arXiv preprint 2018, arXiv:1806.05226v3.
53. Priyadharshini, J.M.H.; Kavitha, S.; Bharathi, B. Classification and analysis of human activities. In Proceedings of the International Conference on Communication and Signal Processing (ICCSP), Chennai, India, 6–8 April 2017; pp. 1207–1211.
54. Vaka, P.R. A Pervasive Middleware for Activity Recognition with Smartphones. Master's Thesis, University of Missouri, Columbia, MO, USA, 2015.
55. Jalal, A.; Batool, M.; Kim, K. Stochastic recognition of physical activity and healthcare using tri-axial inertial wearable sensors. Appl. Sci. 2020, 10, 7122.
56. SivaKumar, A. Geometry Aware Compressive Analysis of Human Activities: Application in a Smart Phone Platform. Ph.D. Thesis, Arizona State University, Tempe, AZ, USA, May 2014.