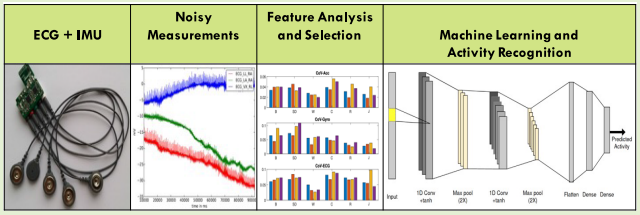
Human Activity Recognition Based on Wireless Electrocardiogram and Inertial Sensors

Sajad Farrokhi, Waltenegus Dargie, *Senior Member, IEEE*, and Christian Poellabauer, *Senior Member, IEEE*

***Abstract*—Wearable devices enable remote, long-term, and unobtrusive monitoring of patients in their everyday living and working environments. Remote health monitoring often involves monitoring physical and cardiac activities (exertions) to establish correlations between the two. With recent advances in sensor technologies and machine learning, the efficiency with which these activities can be recognized has been steadily improving. In this article, we apply convolutional neural networks (CNNs) to measurements taken with wireless electrocardiograms (ECGs) and inertial sensors for human activity recognition (HAR). Experimental results confirm that our approach can recognize a wide range of everyday activities with a high degree of accuracy. Specifically, activities such as jumping, running, and sitting could be recognized with an accuracy exceeding 99%, while activities such as bending over, walking, standing up, and climbing stairs could be recognized with an accuracy exceeding 90%. Overall, the results suggest that the combined use of inertial sensors and ECG leads to better recognition accuracy. Likewise, this article closely examines the contributions of individual sensors and if and to what extent their placement affects recognition accuracy.**



***Index Terms*— Activity recognition, inertial sensors, patient monitoring, wearable computing, wearable sensors, wireless electrocardiogram (ECG).**

I. Introduction

EARLY diagnosis and treatment of diseases are vital in health care. Nevertheless, factors such as work, familial obligations, habitual activities, and financial limitations impede many from seeing doctors regularly [1]. For some patients, getting timely appointments in referral hospitals and advanced clinics is a challenge. Additional factors, such as the COVID-19 pandemic, further complicate matters. The pandemic, on the one hand, overwhelmed hospitals and health personnel, so unexpectedly that many patients were prevented from seeking medical assistance [2], but, on the other hand, caused panic, so that many avoided hospitals and health centers for fear of being infected by the virus [3]. Thus, in 2020, the number of patients visiting cardiac centers for heart-related conditions in the USA dropped by 38% [4].

Wearable computing can alleviate some of these challenges. First, it can enable patients to closely monitor their health at home or in their work environments. As an example, the Wireless Motility Capsule, which has been certified by the U.S. Food and Drug Administration to diagnose gastroparesis [5], enables the simultaneous assessment of regional and whole gut transit [6]. The device concurrently measures intraluminal pH, temperature, and pressure as it traverses the gastrointestinal tract. Besides enabling remote and unobtrusive monitoring, it has the potential to replace painful and expensive clinical diagnostic procedures (such as the use of endoscopy and nuclear medicine) [6]. Second, wearable computing can reduce the cognitive burden of health personnel, so that they can prioritize tasks. Third, it enables long-term diagnosis and monitoring. Besides enabling the early detection of emerging conditions, the latter also enables the collection of reliable statistics, on the basis of which the onset of diseases and their relationship with lifestyle, activity, and habit can be established.

Manuscript received 10 November 2023; revised 19 December 2023; accepted 26 December 2023. Date of publication 3 January 2024; date of current version 29 February 2024. This work was supported in part by the German Research Foundation (DFG) under Project DA 1211/7-1. The associate editor coordinating the review of this article and approving it for publication was Prof. Li-Chia J. Tai. *(Corresponding author: Waltenegus Dargie.)*

This work involved human subjects or animals in its research. Approval of all ethical and experimental procedures and protocols was granted by the Ethics Committee of the TU Dresden (under Application No. EK271072017).

Sajad Farrokhi and Christian Poellabauer are with the Knight Foundation School of Computing and Information Sciences, Florida International University, Miami, FL 33174 USA (e-mail: sfarrokh@fiu.edu; cpoellab@fiu.edu).

Waltenegus Dargie is with the Faculty of Computer Science, Technische Universität Dresden, 01062 Dresden, Germany (e-mail: waltenegus.dargie@tu-dresden.de).

Digital Object Identifier 10.1109/JSEN.2023.3348661

Furthermore, for some health conditions, patients are advised to keep medical journals to establish correlations between symptoms (such as range of motion, episodic events, fatigue, headache, irritability, chest discomfort, breathing difficulty, lightheadedness, dizziness, exhaustion, or anything else the patients may consider relevant) and potential external causes (room temperature, relative humidity, overexertion, etc.). Medical journals, however, are subjective and may be inconsistent and/or incomplete. Wearable sensors can be used to verify and complement journal entries [7].

One of the most important wearable devices is the wireless electrocardiogram (ECG). It is useful for monitoring several cardiovascular conditions (such as atrial fibrillation, atrial tachycardia, and atrial flutter [8]) from remote. Under normal circumstances, there is a correspondence between physical and cardiac activities. During the diagnosis of cardiovascular diseases, patients are often asked to perform certain physical activities (cycle test and cardiopulmonary exercise test) while ECG measurements are taken. The aim is to examine how cardiac responses follow physical exertions. When patients are monitored remotely and cardiologists do not have information about the level of physical exertions, their interpretations of ECG measurements may be inaccurate even though all the vital ECG waves appear to be in their proper places. This article aims to address (quantitatively) the following research concerns.

1) Whether it is possible to estimate physical exertions from ECG measurements.
2) Whether and to what extent ECG measurements complement human activity recognition (HAR) based on inertial measurements.
3) How much the placement of sensors affects the accuracy with which human activity can be recognized.
4) Whether ECG features extracted from arbitrary subjects can be used to develop generalized models.

For this purpose, we: 1) employ a five-lead wireless ECG, a 3-D accelerometer (ACC), and a 3-D gyroscope (Gyro) and 2) identify seven everyday activities that are likely to be performed in home environments—these are bending over (**B**), standing up (**SD**), walking (**W**), climbing (**C**) up or down a staircase, jumping (**J**), sitting (**SI**), and running (**R**).

The remaining part of this article is organized as follows. In Section II, we review related work. In Section III, we discuss the measurement and experimental settings, as well as the preprocessing of the measurement sets. In Section IV, we closely investigate the correlation between the different measurement sets for different activities and discuss activity recognition. In Section V, we describe the convolutional neural network (CNN) we propose for HAR. In Sections VI and VII, we closely examine experimental results and discuss the advantages and limitations of different configurations. Finally, in Section VIII, we give concluding remarks and outline future work.

II. Related Work

HAR is an active research area that promises the support of a diversity of applications. The proposed approaches can be differentiated in terms of the platforms they target; the types of sensors they employ; the sources of raw data; the features they identify; the feature selection process; and the recognition techniques. The primary data sources for HAR are accelerometers and gyroscopes. Similarly, the recognition techniques that are widely employed are support vector machines (SVMs), CNNs, LSTMs, and decision trees (DTs).

Liu et al. [9] employed a wireless ECG and an accelerometer for HAR. Discrimination between the activities is made using a DT. The authors claim to have achieved an overall accuracy of 96.92% based on experiments conducted on 13 volunteers aged between 5 and 68 years. Moreover, the study suggests that the combined use of the accelerometer and ECG improved activity recognition compared to using the ECG alone. Afzali Arani et al. [10] relied on a dataset containing physiological and motion data of 15 subjects [11]. The data were obtained with wrist and chest devices—the accelerometer, ECG, and photoplethysmogram (PPG)—during various activities in real-life conditions. The results show that accelerometer measurements were the most expressive, though ECG measurements improved recognition in some activities, such as walking and ascending/descending stairs. The authors suggest that ECG was better at discriminating between activities that have similar motion patterns but differ in cardiac exertions. Accordingly, combining features from accelerometer and ECG measurements enhanced the classifier's F1-scores by 2.72% and 3.00% for intraperson and interperson classifications, respectively.

Recent advances in rehabilitation robotics have paved the way for personalized gait monitoring. Semwal et al. [12] employed deep learning and individualized gait trajectory graphs to configure rehabilitation systems according to the needs and walking characteristics of individuals with specific walking disabilities. In the study, four models were evaluated using joint angle data of hip, knee, and ankle joints. The models are LSTMs, CNNs, GRU (gated recurrent unit) [13], and a hybrid sequential model combining LSTMs and CNNs. Trained on the dataset of 42 healthy individuals across varied walking speeds, the performance of the LSTM–CNN model stood out. In terms of correlation and R2 Score, this model produced stable gait trajectories within a speed range of 0.49–1.76 m/s and exhibited a significant correlation (0.98) between predicted and actual trajectories.

Jia and Liu [14] presented a technique for identifying human daily activities by fusing an accelerometer and the measurements of a seven-lead ECG. The authors utilize a dataset consisting of four individuals. The proposed approach uses relevance vector machines (RVMs) to classify reduced feature vectors. Experimental results reveal that the fusion of heterogeneous data provided complementary evidence, leading to improved performance (with accuracy as high as 99.57%). Similarly, Mekruksavanich and Jitpattanakul [15] and Celik et al. [16] employed surface electromyography, since physical activities entail strong muscle activities. According to Celik et al. [16], the inclusion of features extracted from the measurements of surface electromyography improved the recognition accuracy by 3.5% when SVM was employed, and by 6.3% when K-nearest neighbors (KNN) was employed. Overall, the authors reported a recognition accuracy exceeding

TABLE I
VARIETY OF NONINVASIVE WEARABLE SENSORS TO MONITOR PHYSICAL ACTIVITIES DURING REHABILITATION

Sensor Type	Purpose	Technology
Accelerometer	Motion tracking + Gait analysis	MEMS
Gyroscope	Posture monitoring	MEMS
EMG	Muscle activity	Surface electrodes
ECG	Cardiac monitoring	Electrode patches
IMU	Spine movement + Joint angle monitoring	MEMS
EEG	Brain activity	Electrode caps
Pressure Sensors	Gait analysis	Force-sensitive resistors

90%. The classifiers discriminated between four everyday activities, namely climbing up a staircase (ascent), climbing down a staircase (descent), walking, and standing.

In summary, the articles we reviewed in this section support our claim that combining features from multiple sensors improves HAR accuracy. Table I summarizes some of the sensors (devices) employed in HAR. This article complements previous approaches in three ways: First, we investigate the contributions of a wireless ECG, a 3-D accelerometer, and a 3-D gyroscope in HAR, both individually and combined. Second, we compare the difference in recognition accuracy of different sensor placements. Our comprehensive investigation provides valuable insights into the optimal selection of sensors and their placements. Third, we examine the correlations between different high-level features to account for the additional gain that can be achieved at higher-level configurations.

III. DATA ACQUISITION AND PREPROCESSING

We employed the Shimmer (version 3) platform[1] to measure the movements and cardiac responses of ten healthy subjects between the ages of 25 and 30. The activities took place in the corridors and one of the staircases of the Faculty of Computer Science at TU Dresden (Germany). Each activity lasted 120 s and the sensors were sampled at 512 samples per second. The sensor platform integrates a five-lead wireless ECG, a 3-D accelerometer, and a 3-D gyroscope. The ECG delivers three channels labeled as left leg–left arm (LL–LA), left leg–right arm (LL–RA), and left arm–right arm (LA-RA). We chose one of the channels (LL–LA) for the task. The six inertial channels measure rectilinear and curvilinear (angular velocity) motions. We considered the total acceleration and the total angular velocity produced by a movement.

A. Preprocessing

The data were subjected to several preprocessing steps to prepare them for further analysis. Initially, interquartile outlier detection was carried out to detect and remove values that lay outside the range specified by the interquartile range (IQR). These values, known as outliers, were identified as those laying beyond the 75th and 25th percentiles of the data

$$\text{IQR} = Q_3 - Q_1 \tag{1}$$

[1]https://shimmersensing.com/product/consensys-ecg-development-kits/ (last visited on December 10, 2023: 11:00 AM CET).

TABLE II
TIME- AND FREQUENCY-DOMAIN FEATURES FOR CLASSIFYING PHYSICAL ACTIVITIES

Domain	Feature
Time	Energy (E) Zero crossing rate (ZCR) Coefficient of variation (CoV) Standard deviation Median Mean
Frequency	Center frequency (FC) Spectral coefficient of variation (SCoV) Spectral energy density (SED) Spectral roll-off (SRO) Spectral centroid (SC)

where IQR is the interquartile range, Q_3 is the 75th percentile, and Q_1 is the 25th percentile. The next step involved reducing noise from the measurements while retaining their essential characteristics, including edges and other important features. This was done through a process known as total-variation denoising [17], which involves minimizing total variation. This is achieved by calculating the absolute differences between adjacent data points and summing them up. The goal of this step was to mitigate the impact of noise on the measurements, preventing at the same time the loss of valuable information [18]

$$\min \left\{ \sum_{i=1}^{n-1} |x_{i+1} - x_i| \right\} \tag{2}$$

where n is the number of data points in the signal. The data was then segmented using a window size of 5120 data points. This corresponds to 10 s of activity (cardiologists often inspect 10-s ECG measurements to establish various cardiac-related conditions, such as sleep apnea and atrial fibrillation [19], [20], [21]; we take this as a reference in our analysis). Furthermore, we differentiated the segments to focus only on the changes in the values introduced by the activities. Following these steps, the data were normalized to eliminate any potential biases that might have been introduced during data collection and preprocessing.

B. Features

We explored several features commonly used in digital signal processing and speech recognition dealing with similar concerns (summarized in Table II). We explored both time- and frequency-domain features. For a detailed description of the features, we refer the reader to [22] and [23].

IV. ANALYSIS

In this section, we present our experimental results feature-by-feature, focusing on the most important features and activities. In the figures that follow, the values on the y-axes need appropriate mapping and scaling to correspond to real physical values, due to the differentiation and normalization steps. They can be regarded as "scores" by which the different activities can be discriminated. In general, a high score does not necessarily correspond to a high physical exertion, since the differentiation step extracts only the difference in

TABLE III
CORRELATION COEFFICIENTS OF THE FEATURES ASSOCIATED WITH THE 3-D ACCELEROMETER, 3-D GYROSCOPE, AND THE ECG

Feature	ρ_{ag}	ρ_{ae}	ρ_{ge}
Energy	-0.7	0.92	-0.57
ZCR	0.88	0.90	0.90
CoV	0.44	0.85	0.24
FC	0.31	0.88	0.42
SED	0.87	0.86	0.74
SCoV	-0.30	0.81	-0.2
SRO	0.99	0.97	0.99
SC	0.99	0.997	0.995

magnitude between adjacent raw samples. If the difference is consistently small, the corresponding score will be small as well, even though the activity involves high physical exertion. Hence, instead of emphasizing the magnitude of a score, it is important to examine the variations in the scores and the correlation coefficients of the scores.

A. Correlation

An essential step toward determining the expressiveness of the features we discussed in Section III-B is to investigate their variance and their correlation with the other features. To visualize these aspects, we plotted the overall average values and the personal average values. The latter refers to the average values of the features based on data collected from a single subject, whereas the former refers to the average values computed based on all the available statistics. The average scores are depicted in blue in Figs. 2–7. To highlight how these scores change from person to person, we plotted the average scores of three of our subjects—in the plots orange, yellow, and purple refer to Subjects 1, 2, and 3, respectively. We sometimes refer to climbing, jumping, and running as robust activities, and to the other activities as slow activities.

In summary, the features associated with the 3-D accelerometer and the ECG exhibit strong correlations, as can be seen from the correlation coefficients in Table III. This suggests that forward acceleration produced by a physical exertion resulted in a corresponding cardiac exertion and both the time- and frequency-domain features captured this aspect. In contrast, the features associated with the gyroscope are moderately correlated with the features associated with the accelerometer and weekly correlated with the features associated with the ECG. The gyroscope consistently ranked bending over and standing high, because these activities involved significant turning even when the body did not exert much. As far as the gyroscope was concerned, the placement of the sensor platform played a crucial role in perceiving body turns. For all the experiments, the sensor platform was placed at the center of the upper torso (Fig. 1). As a result, even though the angular velocities produced by the arms and legs during such activities as jumping and running were high, the gyroscope did not "perceive" them as such due to their placement.

The feature which is the least expressive in terms of the variability of the scores is spectral roll-off (SRO). This feature identifies the frequency up to which a certain portion of the total spectral energy resides. The frequency components of all

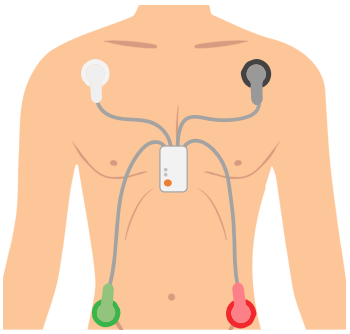


Fig. 1. Placement of the sensor platform and the ECG electrodes.

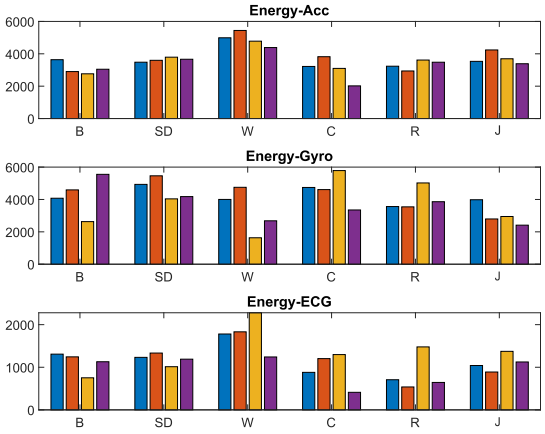


Fig. 2. Variation in the total energy.

the less-robust activities show no significant variations (and consistently scored high), whereas the robust activities exhibit slight variations (and scored less). However, discrimination between the activities based on this feature would not be possible. Likewise, the correlation between the features of the inertial and the ECG measurements is comparatively small (particularly, the correlation between the features of the gyroscope, on the one hand, and the features of the accelerometer and the ECG, on the other), as can be seen in Table III. The second least expressive feature is spectral centroid (SC). Nevertheless, in both cases, all the measurement sets exhibit strong correlations, as can be judged from the correlation coefficients in Table III.

The next least expressive feature is energy (Fig. 2). This is because of the differentiation applied to the raw data; if there is some consistency of physical exertion in an activity, the variation in the energy will be perceived as insignificant. The movement that scores the highest is walking (for the measurements of the 3-D accelerometer and the ECG). Understandably, this activity causes frequent transitions in the body exertion which are captured by the accelerometer and the ECG, as their strong correlation suggests in Table III. As far as the gyroscope is concerned, no conclusion can be made by examining the difference between the scores, both across subjects and across activities.

Zero crossing rate (ZCR) distinguishes bending over and standing unambiguously, based on all the measurement sets (Fig. 3). By contrast, jumping scores the least. ZCR

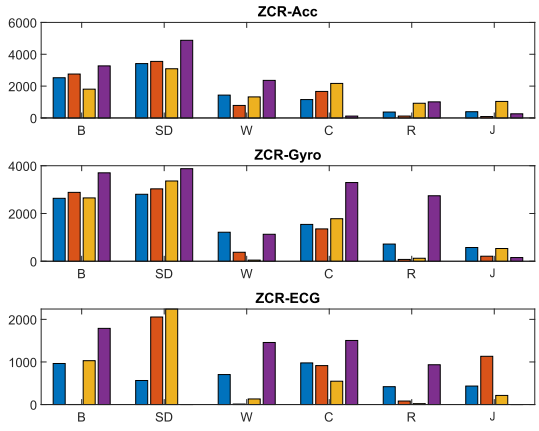


Fig. 3. ZCR of different activities.

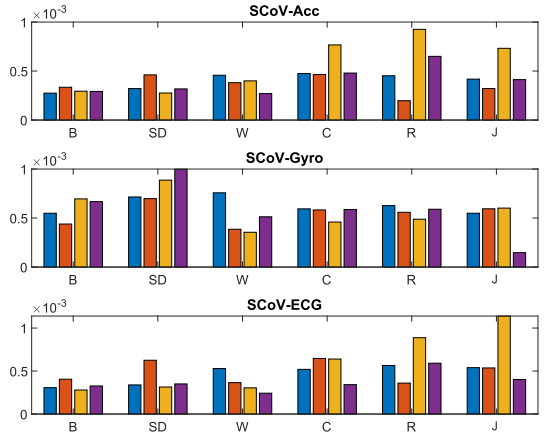


Fig. 5. SCoV.

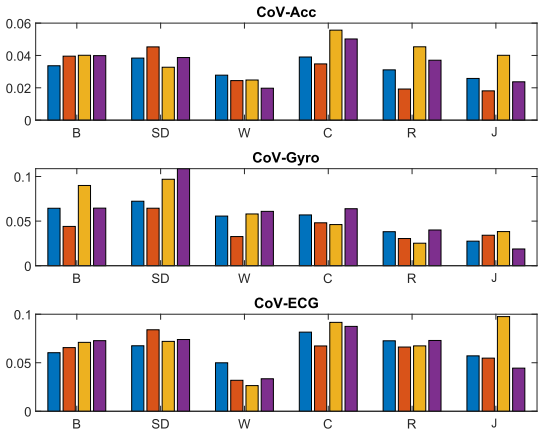


Fig. 4. CoV of different activities.

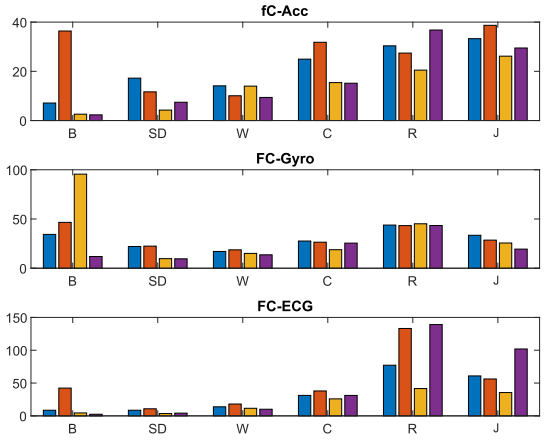


Fig. 6. FC of different activities.

ranks climbing third based on the measurements of the 3-D gyroscope and the ECG, but based on the measurements of the 3-D accelerometer, it ranks walking third. Interestingly, next to SRO and SC, the ZCR of the inertial and ECG measurements exhibit strong correlations as can be judged by the correlation coefficients in Table III. Similarly, coefficient of variation (CoV) consistently ranks bending over, standing, and climbing high based on all the measurement sets. The movement that scores the least is walking. Where the other activities are concerned, the CoV of the accelerometer and the ECG exhibit a strong correlation (corr. coefficient = 0.85), ranking walking the least. However, the CoV associated with the 3-D gyroscope is the most reliable. Accordingly, the more robust a movement is, the smaller is its CoV score. The spectral equivalent of CoV is spectral CoV (SCoV) and expresses the normalized variance between the spectral components. This feature can be taken as the second most reliable feature. Here, however, the more robust an activity is, the higher its value, as can be seen in Fig. 5.

Perhaps, center frequency (FC) is the most reliable feature, based on all the statistics we collected, consistently ranking high the robust activities, as can be seen in Fig. 6. The variance between the different activities for the measurements of the accelerometer and the ECG is appreciably high. If one were merely interested in discriminating between robust and less robust activities, the FC can do the job single-handedly. The correlation coefficient between the measurements of the accelerometer and the ECG for this feature is 0.85.

Finally, spectral energy density (SED) ranks high for those activities that involve sudden changes in body direction. As a result, bending over and standing score comparatively higher than all the other activities based on all the measurement sets. Moreover, the correlation coefficients between the measurement sets are comparatively high. The most reliable SED features in terms of consistency across the different experiments and subjects are those associated with the gyroscope, as can be seen in Fig. 7.

B. Feature Selection

Since all the features are computed based on the same sets of data, it is important to investigate whether they reveal overlapping aspects. Since the features computed for all the activities can be represented by a matrix, dimensionality reduction techniques can be applied to identify the most expressive (independent) features. We chose SVD, as it makes no assumptions regarding the underlying features. The decomposition yielded six singular values, but the first two had values that were far greater than the others suggesting that the matrix

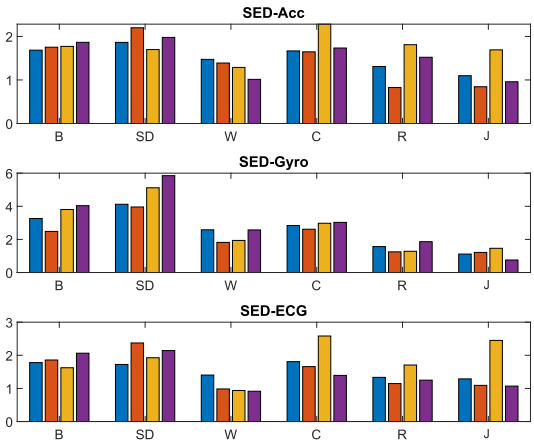


Fig. 7. SED.

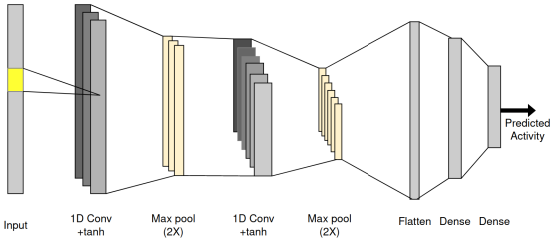


Fig. 8. Proposed CNN model for accurate classification of physical activities based on sensor data.

TABLE IV
NUMBER OF PARAMETERS FOR EACH LAYER OF OUR PROPOSED CNN MODEL

Layer	Conv1	MP1	Conv2	MP2	Flatten	FC1	FC2
Params	160	0	6208	0	0	28736	455

hid two important underlying features. Evaluating the absolute scores (contributions) of the 24 features with respect to the two singular values revealed that some of our features had almost no contributions in describing the singular values, suggesting that they were redundant features. These features were: CoV, SED, and SCoV. As can be seen in Table III, these features were also the ones that yielded weak correlation coefficients. We tested the maximum recognition accuracy that could be achieved with and without these features and the results we obtained suggested that slight improvements could be achieved if they were considered in the subsequent recognition assignments. Hence, despite their poor SVD performance, we decided to retain them.

V. MODEL

In our CNN model, inputs are structured as a 1-D vector, where the length of the vector corresponds to the number of features supplied to the model. The model comprises seven layers, each carefully designed to process input data efficiently and effectively. Specifically, the architecture includes four layers of convolutional and max-pooling (MP) layers that work together to capture and analyze key patterns in the data. The first convolutional layer includes 32 filters with a kernel size of 4, and the second convolutional layer includes 64 filters with a kernel size of 3. These convolutional layers are followed by max-pooling layers with a pool size of 2, respectively, that downsample the output of their corresponding convolutional layers. This architecture captures and extracts relevant patterns and features from the input data while minimizing the number of parameters. Fig. 8 offers a visual representation of our model's architecture.

After passing through the pooling layers, the output is sent to a flattened layer that rearranges the output of the previous layers into a 1-D tensor. This tensor is then supplied to two fully connected layers, where the first layer comprises 64 units with a hyperbolic tangent activation function, and the second layer comprises seven units with a Softmax activation function. By employing the Softmax activation function, the model generates probabilities for each class, ensuring that the sum of probabilities for all classes equals 1. This makes the model fit for multiclass classification.

To evaluate the size of the model, we calculated the number of trainable parameters, which is an indicator of the model's complexity. Table IV shows the number of parameters for each layer of our proposed CNN model.

To enhance the effectiveness of our model, we implement various techniques. Specifically, we utilize the RMSprop optimizer with a learning rate of 0.01 and a decay rate of 0.001 to optimize the model's performance. Additionally, we integrate early stopping to prevent overfitting by discontinuing training if the validation loss did not improve after a certain number of epochs. To evaluate the model's performance, we allocate a 25% holdout set for testing, and we also employ a fivefold cross-validation to ensure the reliability of our results.

VI. RESULTS

We conducted both intraperson and interperson experiments to investigate the impact of individual differences on the accuracy of the model. We used the same dataset for both experiments, but there was a difference in the data usage. For the interperson experiments, data from six participants were used for training, and data from two other participants were used for testing. We performed 27 tests for each sensor combination, considering all possible participant combinations. For the intraperson experiments, data belonging to the same subject were used for both training and testing. The results of all the subjects were then averaged. We split the data into training and test sets in a 75:25 ratio and used fivefold cross-validation to ensure reliable results. We conducted 20 experiments for each sensor placement, considering different sensor combinations to assess their impact on accuracy.

Table V shows the classification accuracy of our model for the intraperson experiments, for different combinations of sensor data. The highest accuracy of 95% and the lowest accuracy of 56.7% were achieved when sensors were placed on the chest. The table also shows the average accuracy and the accuracy range for each sensor placement and combination. We found that the combination of accelerometer (A), gyroscope (G), and ECG (E) sensors placed on the chest yielded

TABLE V
INTRAPERSON CLASSIFICATION ACCURACY OF THE PROPOSED CNN MODEL, FOR VARIOUS SENSOR PLACEMENTS AND SENSOR COMBINATIONS

Sensor Placement	Average Accuracy (%)	Accuracy Range (%)
Chest (A+E+G)	94	91.9 - 95
Back (A+E+G)	93	91.5 - 94.4
Left Arm (A+E+G)	92.2	89.9 - 93.3
Right Arm (A+E+G)	92	90.4 - 92.9
Right Arm (A+G)	92	91.4 - 93.5
Chest (A+G)	92	90.6 - 93.1
Left Arm (A+G)	91	89.5 - 91.9
Left Leg (A+G)	91	88.8 - 92.2
Back (A+G)	91	90 - 92
Left Arm (E+G)	91	88.4 - 92.2
Back (E+G)	91	89.9 - 92.2
Chest (E+G)	90.4	88.2 - 92.3
Right Arm (E+G)	90	89 - 91.9
Left Leg (A+E+G)	89	87.8 - 91.2
Chest (A+E)	89	87.3 - 90.5
Right Arm (G)	89	86.9 - 90.4
Back (A+E)	89	86.7 - 90.5
Chest (G)	88	86.1 - 89.7
Left Arm (A+E)	87	83.9 - 88.8
Left Leg (A+E)	87	85.6 - 88.7
Left Leg (E+G)	86	84 - 86.9
Left Arm (G)	86	84.2 - 87.8
Back (G)	85	82.7 - 86.6
Left Leg (A)	85	83.4 - 86.5
Chest (A)	83	80.7 - 84.2
Back (A)	83	81 - 84.7
Right Arm (A+E)	82	79.7 - 84
Left Arm (A)	82	80.7 - 83.5
Left Leg (G)	82	79 - 84.6
Right Arm (A)	79	77.2 - 80.8
Back (E)	60	58.9 - 62.3
Chest (E)	59	56.7 - 61.4

TABLE VI
INTERPERSON CLASSIFICATION ACCURACY OF THE PROPOSED CNN MODEL FOR VARIOUS SENSOR PLACEMENTS AND SENSOR COMBINATIONS

Sensor Placement	Average Accuracy (%)	Accuracy Range (%)
Right Arm (E+G)	77	71 - 86.9
Left Arm (A+E+G)	75	63.3 - 85.5
Right Arm (A+E+G)	74	54.9 - 83.5
Back (A+E+G)	74	61.6 - 82.1
Right Arm (A+G)	74	60.4 - 84.2
Left Arm (A+G)	73	61.6 - 81.4
Chest (A+G)	73	57.1 - 83.9
Left Arm (E+G)	72	57.2 - 84.2
Back (A+G)	72	54.5 - 81.5
Back (E+G)	72	61.8 - 79.6
Back (A)	72	56.8 - 85.2
Chest (E+G)	72	58.8 - 80.3
Chest (A+E+G)	71	55 - 85
Back (G)	70	57.7 - 79
Left Arm (G)	70	49 - 83
Chest (G)	69	54.2 - 83.2
Right Arm (G)	69	50.9 - 83.1
Left Arm (A+E)	66	52.1 - 76.6
Right Arm (A)	66	53.7 - 78.2
Left Arm (A)	65	47.9 - 77.4
Chest (A+E)	62	52.3 - 62.8
Right Arm (A+E)	62	49.4 - 74.4
Left Leg (A+G)	61	51.4 - 71.3
Left Leg (A+E+G)	60	48.5 - 68.5
Left Leg (A+E)	60	48.7 - 68.5
Left Leg (A)	60	41.4 - 75.7
Left Leg (E+G)	59	49.1 - 65.4
Back (A+E)	59	42.4 - 71.1
Chest (A)	57	43 - 69.9
Left Leg (G)	53	41.2 - 66
Back (E)	41	30.3 - 61.2
Chest (E)	35	23.2 - 44

the highest average accuracy of 94%. Our model consistently delivered accuracy rates ranging between 91.9% and 95% across all 20 experiments, indicating that this combination of sensors is well-suited for predicting the different classes. Conversely, the lowest accuracy was observed when sensors were placed on the chest with only ECG data, with an accuracy of 59%. This indicates that the ECG alone is not sufficient for accurate HAR using CNNs (as discussed below, the problem was not as such with the inadequacy of the ECG in capturing physical exertions but with motion artifacts being included in the ECG measurements when the body was exerting.

Table VI lists the results of interperson experiments. Accordingly, the combination of the ECG and gyroscope placed on the right arm yielded the highest average accuracy of 77% with an accuracy range of 71%–86.9%. On the other hand, the lowest average accuracy of 35% with an accuracy range of 23.2%–44% was observed when the ECG was used together with inertia sensors placed on the chest. These results highlight the importance of sensor placement and the combination of sensor data for accurate HAR across different individuals. Furthermore, the results presented in Tables VII and VIII suggest that the optimal sensor placement depends on the ultimate aim of the model. For instance, the chest was the most suitable place for jumping and bending over in the intraperson experiments, whereas the left arm was optimal for jumping and the right arm, for bending over, in the interperson experiments. This could be because the latter activities resulted in similar arm movement patterns across subjects, enabling features that were amenable to generalization.

TABLE VII
AVERAGE HIGHEST ACCURACY ACHIEVED FOR EACH ACTIVITY DURING THE INTRAPERSON TEST

Activity	Sensor Placement	Accuracy (%)
Jumping	Chest (A+G+E)	99.8
Running	Back (A+G+E)	100
Sitting	Back (G)	99.3
Stairs	Back (A+G)	90
Walking	Right Arm (A+G+E)	91.1
Bending Over	Chest (A+G+E)	91.3
Standing Up	Right Arm (A+G)	94.4

VII. DISCUSSION

A considerable variability of performance can be observed between the intraperson and the interperson experiments. Understandably, this, in part, is due to physical differences. Human movement is not uniform across individuals, and even subtle differences in walking strides and running paces can change measurement statistics appreciably, even though

TABLE VIII
AVERAGE HIGHEST ACCURACY ACHIEVED FOR EACH ACTIVITY DURING THE INTERPERSON TEST

Activity	Sensor Placement	**Accuracy (%)**
Jumping	Left Arm (A+G)	86.8
Running	Left Arm (A+G+E)	98
Sitting	Right Arm (A)	99.3
Stairs	Left Arm (A+G)	79.2
Walking	Left Arm (A)	78.9
Bending Over	Right Arm (G+E)	67.4
Standing Up	Right Arm (A+G)	80.7

the higher-level features were so selected to undermine the effect of such variations. Unlike the interperson experiments, which relied on data coming from different individuals, the intraperson experiments utilized data from the same individual for both training and testing. Consequently, a more consistent performance could be achieved even when a change in the configuration of the sensing system changed. Table V reveals that the ECG excelled in identifying sitting (as also shown in Tables VII and VIII). Indeed, from this observation, we can suspect that it was mainly due to the inclusion of motion artifacts in the ECG measurements that ECG performed poorly. Had the ECG measurements been clean, it could have been possible to map cardiac activities to physical activities. Hence, when the ECG was used together with inertial sensors, the latter implicitly normalized the effects of motion artifacts in addition to providing complementary insights into the underlying physical activities.

As shown in Table V, using an accelerometer alone for chest sensor placement, our model achieved an accuracy of 83% but when used with the ECG and gyroscope, the accuracy increased by 11%, reaching 94%. Overall, the combined use of ECG and IMU features improved accuracy on average by 2%–6%. This indicates that using all three modalities did not always result in a higher accuracy. In the interperson experiments, the combination of the ECG and gyroscope (placement: right arm) yielded the highest accuracy. When, however, all three sensors were used, we observed a 3% decrease in accuracy. This agrees with the study conducted by Afzali Arani et al. [10] which combined features extracted from the measurements of the PPG, ECG, and accelerometer. The authors observed that the combined use of PPG, accelerometer, and ECG features led to a decrease in accuracy.

Our results align well with the results reported in the literature (a summary is given in Table IX). Afzali Arani et al. [10] noted that using the ECG along with a 3-D accelerometer improved the recognition of some activities in ten out of 14 subjects. Mekruksavanich and Jitpattanakul [15] reported a 1% improvement of accuracy when the accelerometer, gyroscope, and electromyography were used in combination. Similarly, Celik et al. [16] reported that the combined use of surface electromyography and inertial sensors resulted in an improvement of accuracy by 3.5%–6.3%. Similar observations were made in [9], [14], and [24].

Our study goes beyond state-of-the-art by exploring the impact of various sensor placements and combinations in greater detail. Table VI presents a summary of prior studies and their corresponding outcomes. To the best of our knowledge, our study is the first to thoroughly investigate the impact of various sensor placements. The confusion matrix for the chest ECG (Table X) shows that the model achieved an accuracy of 96.4% in identifying sitting activity. Similarly, in the confusion matrix for the back ECG (Table XI), the model achieved an even higher accuracy of 98.2% in identifying sitting activity. These results suggest that in the absence of motion artifacts, the ECG effectively mapped cardiac response to physical response.

Similarly, the results in Tables XII and XIII suggest that the chest sensors were more reliable in classifying certain physical activities. The case in point is standing up, bending over, and walking. On the other hand, back placement performed better in identifying climbing stairs and running (achieving

TABLE IX
COMPARISON OF ACTIVITY RECOGNITION STUDIES

Paper	**Sensors**	**Sensor Placement**	**Number of Activities**	**Proposed Algorithm**	**Results**
[9]	ECG ACC	Chest	5	Posture and Activity Classification	Accuracy 96.92%
[10]	ECG ACC PPG	Chest(ECG) + Wrist(ACC + PPG)	8	Random Forest	F1 Score 96.80%
[24]	ECG ACC	Left Thorax	3	Fuzzy Decision Tree	FAR 4.4% FRR 29.5%
[14]	ECG ACC	Waist (ACC) + Chest (ECG)	8	RVM	Accuracy 99.57%
[15]	sEMG ACC Gyro	Forearm	5	DT and MLP	Accuracy 99.97%
[16]	sEMG ACC Gyro	Leg	4	SVM and KNN	Accuracy 99%
This Study	ECG ACC Gyro	Chest + Back + Left Arm + Right Arm + Leg	7	CNN	Accuracy 94%

TABLE X
CONFUSION MATRIX FOR CHEST ECG DATA OVER THE INTRAPERSON TEST

Actual Class	**Predicted Class**						
	J	R	SI	C	W	B	SD
J	67.7	15.0	0.7	3.5	9.3	3.2	0.6
R	10.8	64.0	0	9.1	10.3	3.6	2.2
SI	0.0	0.0	96.4	0.0	0.0	3.6	0.0
C	3.6	2.1	0.0	33.4	32.9	15.9	12.3
W	1.7	3.5	0.0	19.5	59.2	11.4	4.8
B	1.0	0.7	4.7	4.0	11.8	65.1	12.8
SD	0.4	1.0	7.3	19.2	17.8	31.1	23.2

TABLE XI
CONFUSION MATRIX FOR BACK ECG DATA THE INTRAPERSON TEST

Actual Class	Predicted Class						
	J	R	SI	C	W	B	SD
J	60.9	24.9	0.0	5.0	5.9	3.2	0.2
R	14.7	70.5	0.0	9.0	1.9	2.9	1.0
SI	0.0	0.0	98.2	0.0	0.0	1.8	0.0
C	2.5	19.8	0.0	45.3	16.6	3.5	12.2
W	6.1	19.7	0.0	21.5	41.3	5.1	6.2
B	3.7	8.3	3.6	5.7	4.0	67.2	7.5
SD	4.6	7.3	0.0	25.7	4.3	26.1	31.9

TABLE XII
CONFUSION MATRIX FOR CHEST DATA ALL SENSORS THE INTRAPERSON TEST

Actual Class	Predicted Class						
	J	R	SI	C	W	B	SD
J	99.8	0.2	0.0	0.0	0.0	0.0	0.0
R	0.5	99.5	0.0	0.0	0.0	0.0	0.0
SI	0.0	0.0	98.6	0.0	0.0	1.4	0.0
C	0.0	0.0	0.0	86.2	12.0	1.8	0.0
W	0.0	0.0	0.0	9.5	90.4	0.0	0.0
B	0.0	0.0	2.6	0.0	0.2	91.3	6.0
SD	0.0	0.0	0.1	0.0	0.0	9.3	90.5

TABLE XIII
CONFUSION MATRIX FOR BACK DATA ALL SENSORS THE INTRAPERSON TEST

Actual Class	Predicted Class						
	J	R	SI	C	W	B	SD
J	99.6	0.2	0.0	0.1	0.1	0.0	0.0
R	0.0	100.0	0.0	0.0	0.0	0.0	0.0
SI	0.0	0.0	99.3	0.0	0.0	0.7	0.0
C	0.1	0.1	0.0	89.8	9.6	0.2	0.2
W	0.0	0.0	0.0	11.2	88.8	0.0	0.0
B	0.0	0.0	1.9	0.7	0.0	87.9	9.5
SD	0.0	0.0	0.0	0.0	0.0	13.2	86.8

100% accuracy in classifying running). This suggests that back placement is more reliable at detecting and classifying movements that involve the lower body, such as those related to gait and posture.

The results presented in Tables VII and VIII provide further insights into the optimal sensor placements for a reliable HAR. While the chest placement appears to be the best for recognizing jumping and bending over in the intraperson experiments, the left and right arm sensor placements were the best for the interperson experiments. This suggests that the arm placement is better suited for capturing intervariability in activity recognition, while chest and back placements are more reliable for capturing intraperson variability. Our experiments reveal that the left leg placement produced noisy and inaccurate measurements compared to all the other placements, regardless of the participant's age, gender, or physical attributes.

VIII. CONCLUSION

In this article, we employed a CNN for HAR and investigated the contribution of a wireless ECG to recognition accuracy. We considered different configurations to produce the input features, including different sensor placements and sensor combinations. The experimental results indicate that the contribution of the features extracted from the wireless ECG was appreciable; nevertheless, without the inclusion of the features extracted from the inertial sensors, its impact was modest. Additionally, our investigation highlights the significance of sensor placement on recognition accuracy. In general, the features extracted from the sensors placed at the torso were more expressive when the training and test data originated from the same subjects. When the test and training sets originated from different subjects, however, the features extracted from sensors placed at the right and left arms were more expressive.

While our study delivered promising outcomes, it has also left some concerns unaddressed. First, the limited number of subjects involved in the experiments may limit its expressive power. Second, even though our model scored impressive accuracy rates, we did not extensively evaluate its computational complexity and energy cost. Considering the significance of these factors in real-world wearable devices, optimizing the model's efficiency for deployment on resource-constrained devices is an important consideration. Broadening the demographic scope of the model in future studies to ensure the representation of a more diverse set of features and addressing the computational aspects of the model will be the focus of our future research.

REFERENCES

[1] R. Noguchi and J. Shen, "Factors affecting participation in health checkups: Evidence from Japanese survey data," *Health Policy*, vol. 123, no. 4, pp. 360–366, Apr. 2019.

[2] J. M. Templeton, C. Poellabauer, and S. Schneider, "Negative effects of COVID-19 stay-at-home mandates on physical intervention outcomes: A preliminary study," *J. Parkinson's Disease*, vol. 11, no. 3, pp. 1067–1077, Aug. 2021.

[3] M. A. N. Saqib et al., "Effect of COVID-19 lockdown on patients with chronic diseases," *Diabetes Metabolic Syndrome, Clin. Res. Rev.*, vol. 14, no. 6, pp. 1621–1623, Nov. 2020.

[4] S. Garcia et al., "Reduction in ST-segment elevation cardiac catheterization laboratory activations in the United States during COVID-19 pandemic," *J. Amer. College Cardiol.*, vol. 75, no. 22, pp. 2871–2872, Jun. 2020.

[5] M. Camilleri et al., "Gastroparesis," *Nature Rev. Disease Primers*, vol. 4, no. 1, pp. 1–19, 2018.

[6] R. J. Saad and W. L. Hasler, "A technical review and clinical assessment of the wireless motility capsule," *Gastroenterol. Hepatology*, vol. 7, no. 12, p. 795, 2011.

[7] J. M. Templeton, C. Poellabauer, and S. Schneider, "Using wearable devices to mitigate bias in patient reported outcomes for aging populations," in *Proc. 11th EAI Int. Conf. Wireless Mobile Commun. Healthcare (MobiHealth)*, 2022, pp. 362–374.

[8] J. Liu et al., "Early recurrence of atrial tachyarrhythmia during the 90-day blanking period after cryoballoon ablation in patients with atrial fibrillation: The characteristics and predictive value of early recurrence on long-term outcomes," *J. Electrocardiology*, vol. 58, pp. 46–50, Jan. 2020.

[9] J. Liu, J. Chen, H. Jiang, W. Jia, Q. Lin, and Z. Wang, "Activity recognition in wearable ECG monitoring aided by accelerometer data," in *Proc. IEEE Int. Symp. Circuits Syst. (ISCAS)*, May 2018, pp. 1–4.

[10] M. S. Afzali Arani, D. E. Costa, and E. Shihab, "Human activity recognition: A comparative study to assess the contribution level of accelerometer, ECG, and PPG signals," *Sensors*, vol. 21, no. 21, p. 6997, Oct. 2021.

[11] A. Reiss, I. Indlekofer, P. Schmidt, and K. Van Laerhoven, "Deep PPG: Large-scale heart rate estimation with convolutional neural networks," *Sensors*, vol. 19, no. 14, p. 3079, Jul. 2019.

[12] V. B. Semwal, R. Jain, P. Maheshwari, and S. Khatwani, "Gait reference trajectory generation at different walking speeds using LSTM and CNN," *Multimedia Tools Appl.*, vol. 82, no. 21, pp. 33401–33419, Sep. 2023.

[13] R. Dey and F. M. Salem, "Gate-variants of gated recurrent unit (GRU) neural networks," in *Proc. IEEE 60th Int. Midwest Symp. Circuits Syst. (MWSCAS)*, Aug. 2017, pp. 1597–1600.

[14] R. Jia and B. Liu, "Human daily activity recognition by fusing accelerometer and multi-lead ECG data," in *Proc. IEEE Int. Conf. Signal Process., Commun. Comput. (ICSPCC)*, Aug. 2013, pp. 1–4.

[15] S. Mekruksavanich and A. Jitpattanakul, "Exercise activity recognition with surface electromyography sensor using machine learning approach," in *Proc. Joint Int. Conf. Digit. Arts, Media Technol. ECTI Northern Sect. Conf. Electr., Electron., Comput. Telecommun. Eng.*, Mar. 2020, pp. 75–78.

[16] Y. Celik, S. Stuart, W. L. Woo, L. T. Pearson, and A. Godfrey, "Exploring human activity recognition using feature level fusion of inertial and electromyography data," in *Proc. 44th Annu. Int. Conf. IEEE Eng. Med. Biol. Soc. (EMBC)*, Jul. 2022, pp. 1766–1769.

[17] I. Selesnick, "Total variation denoising (an mm algorithm)," NYU Polytech. School Eng. Lect. Notes, Tech. Rep., 2012.

[18] R. M. Al Abdi and M. Jarrah, "Cardiac disease classification using total variation denoising and Morlet continuous wavelet transformation of ECG signals," in *Proc. IEEE 14th Int. Colloq. Signal Process. Appl. (CSPA)*, Mar. 2018, pp. 57–60.

[19] M. E. van den Berg et al., "Normal values of corrected heart-rate variability in 10-second electrocardiograms for all ages," *Frontiers Physiol.*, vol. 9, p. 424, Apr. 2018.

[20] A. H. Khandoker, J. Gubbi, and M. Palaniswami, "Automated scoring of obstructive sleep apnea and hypopnea events using short-term electrocardiogram recordings," *IEEE Trans. Inf. Technol. Biomed.*, vol. 13, no. 6, pp. 1057–1067, Nov. 2009.

[21] W. Shuai et al., "Is 10-second electrocardiogram recording enough for accurately estimating heart rate in atrial fibrillation," *Int. J. Cardiol.*, vol. 215, pp. 175–178, Jul. 2016.

[22] W. Dargie and M. K. Denko, "Analysis of error-agnostic time- and frequency-domain features extracted from measurements of 3-D accelerometer sensors," *IEEE Syst. J.*, vol. 4, no. 1, pp. 26–33, Mar. 2010.

[23] W. Dargie, "Analysis of time and frequency domain features of accelerometer measurements," in *Proc. 18th Int. Conf. Comput. Commun. Netw.*, Aug. 2009, pp. 1–6.

[24] T. Fujimoto et al., "Wearable human activity recognition by electrocardiograph and accelerometer," in *Proc. IEEE 43rd Int. Symp. Multiple-Valued Log.*, May 2013, pp. 12–17.