# Effective Human Activity Recognition through Accelerometer Data

**Vu Thi Thuong**
Faculty of Information Technology and Communication, Phuong Dong University, Hanoi City, Vietnam | Graduate University of Sciences and Technology, Vietnam Academy of Science and Technology, Hanoi City, Vietnam
thuongvt@dhpd.edu.vn

**Duc-Nghia Tran**
Institute of Information Technology, Vietnam Academy of Science and Technology, Hanoi City, Vietnam
nghiatd@ioit.ac.vn

**Duc-Tan Tran**
Faculty of Electrical and Electronic Engineering, Phenikaa University, Hanoi City, Vietnam
tan.tranduc@phenikaa-uni.edu.vn

**Bui Thi Thu**
Institute of Information Technology, Vietnam Academy of Science and Technology, Hanoi City, Vietnam
btthu@ioit.ac.vn

**Vu Duong Tung**
Institute of Information Technology, Vietnam Academy of Science and Technology, Hanoi City, Vietnam
tungvd@ioit.ac.vn

**Nguyen Thi Anh Phuong**
Institute of Information Technology, Vietnam Academy of Science and Technology, Hanoi City, Vietnam
ntaphuong@ioit.ac.vn

**Phung Cong Phi Khanh**
Faculty of Technology Education, Hanoi National University of Education, Hanoi City, Vietnam
phungcongphikhanh@gmail.com

**Pham Khanh Tung**
Faculty of Technology Education, Hanoi National University of Education, Hanoi City, Vietnam
tungpk@hnue.edu.vn

**Manh-Tuyen Vi**
Faculty of Electrical and Electronic Engineering, Phenikaa University, Hanoi City, Vietnam
tuyen.vimanh@phenikaa-uni.edu.vn (corresponding author)

*Received: 24 June 2024 | Revised: 19 July 2024 | Accepted: 23 July 2024*

 *DOI: https://doi.org/10.48084/etasr.8211*

## ABSTRACT

**In recent years, the field of Human Activity Recognition (HAR) has emerged as a prominent area of research. A plethora of methodologies have been documented in the literature, all with the objective of identifying and analyzing human activities. Among these, the use of a body-worn accelerometer to collect motion data and the subsequent application of a supervised machine learning approach represents a highly promising solution, offering numerous benefits. These include affordability, comfort, ease of use, and high accuracy in recognizing activities. However, a significant challenge associated with this approach is the necessity for performing activity recognition directly on a low-cost, low-performance microcontroller. This research presents the development of a real-time human activity recognition system. The system employs optimized time windows for each activity, a comprehensive set of differentiating features, and a straightforward machine learning model. The efficacy of the proposed system was evaluated using both publicly available datasets and data collected in experiments, achieving an exceptional activity recognition rate of over 95.06%. The system is capable of recognizing six fundamental daily human activities: standing, sitting, jogging, walking, going downstairs, and going upstairs.**


***Keywords-accelerometer; classification; wearable computing; activity recognition***

## I. INTRODUCTION

The field of HAR is focused on the automatic detection of everyday activities performed by humans [1]. The method employs time series data recorded by sensors. Over the past decade, there have been notable developments in the field of interconnected sensing technologies, including sensors, the Internet of Things (IoT), cloud computing, and edge computing. Sensors are inexpensive and readily integrated into both portable and non-portable devices [2]. This is the reason why the majority of HAR research has been conducted using sensor technology. The latest research in the field of HAR can be broadly classified into two main categories:

- Ambient sensor-based approaches, which employ fixed environmental sensors, including surveillance cameras, microphones, temperature sensors, and others, to capture contextual cues and recognize daily activities within a specific space. Examples of such applications entail smart homes [3], and rehabilitation centers [4]. Authors in [5], proposed the Residual Deep Convolutional-Gated Recurrent Unit (ResDC-GRU) model, which demonstrates superior performance in learning spatiotemporal features from video data, resulting in more accurate and efficient action recognition. However, the reliance of ambient sensor-based approaches on fixed environments constrains their utility for the analysis of activities beyond the confines of these controlled settings.
- Wwearable sensor-based approaches, which employ the use of wearable devices, such as smartwatches or smartphones, equipped with various sensors, including accelerometers, magnetometers, and gyroscopes, to monitor and collect physiological signals that are generated by the body in a specific location [6]. For example, authors in [7] utilized inexpensive wearable devices equipped with accelerometers and gyroscopes to ascertain the user's location and identify activities, such as sitting, standing, and walking. Similarly, authors in [8] utilized Independent Component Analysis (ICA) and Principal Component Analysis (PCA) to ascertain a person's walking posture based on the accelerometer data.

The machine learning models that can be deployed for the HAR problem include supervised learning, semi-supervised learning, unsupervised learning, and reinforcement learning [9], with the supervised and semi-supervised ones being the most commonly utilized. Semi-supervised learning makes use of the considerable quantity of unlabeled data in conjunction with the limited amount of labeled data, thereby addressing the issue of insufficient activity annotation. To address the challenges associated with web page classification, authors in [10] developed a co-training framework. The framework employs unlabeled data to augment the training set and enhance the efficacy of the recognition algorithm. Authors in [11] proposed a novel approach to semi-supervised learning based on the analysis of discrepancies. This approach entails training multiple classifiers on real-world tasks, with the discrepancies between these classifiers directing the learning process. In a recent study [12], a semi-supervised deep learning approach to HAR utilizing both labeled and unlabeled smartphone sensor data was presented. The method employs a temporal ensemble of deep Long Short-Term Memory (LSTM) networks and combines supervised and unsupervised loss functions to achieve accurate HAR. In conclusion, authors in [13] proposed an annotation strategy that uses sparsely labeled data along with readily available unlabeled data. This strategy allows for unobtrusive and context-aware activity recognition through the use of on-body wearable sensors.

Supervised learning typically attains higher accuracy due to the fact that models are created using fully labeled data, which allows for the establishment of clear relationships between features and labels. Building upon these approaches, Authors in [14] proposed a method that demonstrated superior performance in human detection and activity classification. The method employs a deep Convolutional Neural Network (CNN) to directly extract features and classification boundaries from input videos, thereby achieving superior performance compared to previous techniques. Other authors in [15] addressed the challenges of hidden body parts and self-occlusion by developing a method based on background subtraction and hybrid feature extraction. Meanwhile, the research carried out in [16] addressed issues of light variation, low video quality, diverse expressions, and poses using a hybrid approach. This approach combines cascade head-shoulder features, Haar-like features, and Histogram of Oriented Gradients (HOG) features with an adaptive Gaussian Mixture Model (GMM) for the purpose of human recognition.

Authors in [17] concentrated their efforts on the detection and tracking of multiple humans, achieving an impressive 94.53% accuracy. The method employs rapid template matching and three-dimensional model fitting to enhance performance. Authors in [18] implemented a sparse representation method with a complete scale embedded dictionary, thereby achieving effective detection of humans at various scales, while in [19], authors investigated the potential of sensor fusion by combining RGB and depth images captured by Kinect. This approach resulted in an improvement in detection accuracy, reaching 93.1% at a frame rate of 20 fps. Finally, authors in [20] introduced a graph-based segmentation method using motion vector preprocessing. This method improves classification performance by incorporating Motion Self-Difference (MSD) features. Evaluation and testing of models are easier with supervised learning due to the availability of labeled data for comparison. Supervised learning applications are straightforward and training times are shorter because it does not involve processing unlabeled data. Overall, supervised learning offers higher reliability and precision, which makes it ideal for certain classification and regression tasks.

Some of the above studies used deep learning to solve the HAR problem. The results obtained with deep learning have very high accuracy. However, their inherent computational complexity leads to longer execution times and increased resource consumption [21]. This poses a significant challenge for their reliance on resource-constrained devices such as low-cost, low-power microcontrollers. In such scenarios, machine learning models offer a more practical alternative, enabling real-time activity classification with lower computational requirements and significantly faster execution times [22]. The objective of this study is to exploit the potential of body-worn sensors for HAR. The study uses a supervised machine learning method that combines features suitable for real-time HAR. It aims to develop a low-cost healthcare support device for patients recovering from injuries or for the elderly. Some of the main contributions are:

- The issue of addressing HAR on low-performance devices, with computationally efficient algorithms. The K-Nearest Neighbors (KNN) algorithm is a robust option because it is capable of effectively uncovering patterns directly from sensor data, obviating the necessity for complex feature engineering. In addition, KNN is able to make predictions by utilizing information from the closest data points, and maintains the pace with evolving activity patterns over time. Feature selection plays a significant role in this process, and the proposed features are designed specifically for the system, enabling a real-time operation.
- The objective of a real-time and affordable system for activity identification. The system comprises three principal components: hardware (electronic circuits) for data acquisition, signal processing techniques for feature extraction, and machine learning models for training and activity recognition. The system is configured on low-cost, low-performance microcontrollers. The process comprises a series of stages, such as data collection stage, which involves the gathering of data from the sensors, data preprocessing, which allows data cleaning and preparation for further analysis, feature extraction which identifies the principal characteristics presented in the preprocessed data, which are pertinent to the recognition of activities, classification based on microcontroller, and finally, real time communication, using efficient wireless transmission, which allows the transfer of data to a compact server for subsequent processing or visualization.

The remaining sections of this paper review existing research that is pertinent to the proposed system while describing the system model, including its hardware and software components. The proposed method provides a detailed account of the selected classification method, KNN, and its appropriateness for addressing the activity recognition problem. Furthermore, the paper elucidates the rationale behind the selection of specific features, such as, Variance (Var), Sum of Absolute Values (SAV), Skewness (Skew), and Peak-to-Peak (P2P). Furthermore, a comprehensive evaluation of the system's performance, including a detailed analysis of the simulation results and a rigorous theoretical analysis, is presented. The paper concludes with a summary of the main findings and suggestions for future research directions.

## II. MATERIALS AND METHODS

### A. Activities Recognition Model

In this study, a supervised learning method will be employed with the objective of achieving greater accuracy. The comprehensive system methodology comprises three stages. The three stages of the comprehensive system methodology, as shown in Figure 1, are data gathering, model analysis, and activity identification, which involves both implementation and evaluation. The device under consideration is designed to commence data collection at the outset of the process. The objective is to capture acceleration data along three axes ($x$, $y$, and $z$) for the purpose of classifying activities. Subsequently, the data are subjected to segmentation, whereby they are partitioned into discrete units using a sliding window technique. As the timeline is traversed with precision, a multitude of activity-related data points are recorded within each segment. Subsequently, each vector derived from these segments is subjected to a process of feature extraction, resulting in the combination of numerous features. This initial processing stage prepares the data for the subsequent phase, which involves the meticulous selection of relevant features that will serve as inputs for the classifier. The final stage comprises both the implementation and evaluation of the process. In this phase, the selected features from the preceding step are incorporated into the training process, enabling the development of a robust classification model. The model employs its embedded knowledge to enhance the accuracy of activity recognition.

### B. Data Collection

This study introduces a system comprising a simple accelerometer, a cost-effective Inertial Measurement Unit (IMU), and a wireless connector to enable bidirectional data transmission to a wearable device attached to a beverage can. The user-friendly design permits direct data transfer to a smartphone, specifically an Android smartphone, as the one utilized in this study. A principal objective of the design was

to minimize user inconvenience during the data recording process. The complete system architecture is presented in Figure 2.

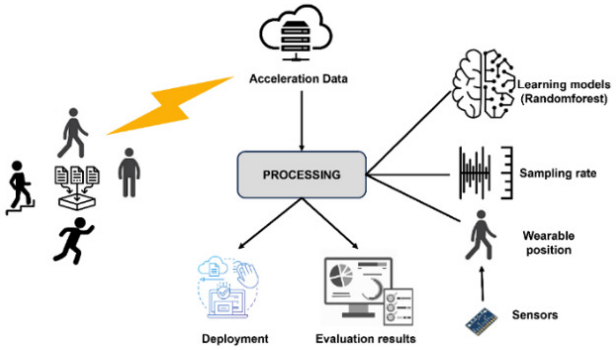
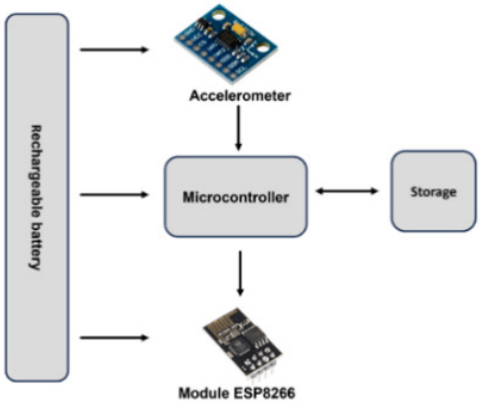



Fig. 1. Steps in activity recognition.

(a)




(b)

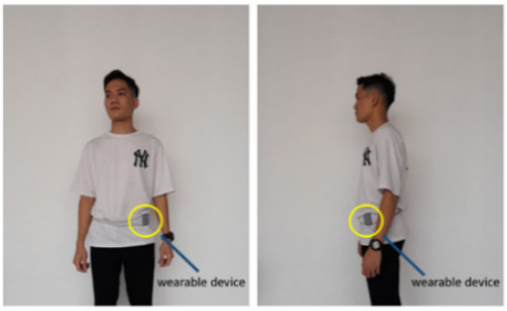



Fig. 2. (a) The proposed system and (b) the developed wearable device's position on volunteer.

The proposed system uses a 3-DOF ADXL345 [23] accelerometer to acquire motion data. To ensure comprehensive activity measurement, the sensor parameters were set to ±4 g full scale, 128 LSB/g sensitivity, and a noise level of 150 µg. This configuration allows for triaxial data acquisition in the *x*, *y*, and *z* axes. The ADXL345 sensor connects to the PIC18F4520 microcontroller via an Inter-Integrated Circuit (I2C) interface. Data transfer is facilitated by the ESP8266, which is connected to the MCU via Universal Asynchronous Receiver Transmitter (UART) communication. The primary function of the ESP8266 is to collect real-time motion data from an accelerometer and transmit them wirelessly to a remote server or connected device for analysis. This enables continuous monitoring and detailed analysis of the user's physical activity patterns. The ESP8266's ability to handle complex calculations and its integrated Wi-Fi module facilitate seamless data transmission, ensuring that the device provides accurate and timely information. The device is powered by a rechargeable 3.7 V, 6,000 mAh battery.

In the proposed device, the accelerometer is used to quantify the movements of volunteers as they perform a series of prescribed actions. In each second, the ADXL345 acquires samples for each axis and transmits them to the PIC18F4520. Consequently, acceleration samples are acquired at one-second intervals. Subsequently, the samples per axis are averaged to produce a single representative sample. Furthermore, a low-pass filter is applied to the data in order to mitigate the effects of high-frequency noise and fluctuations, thereby ensuring a more pristine signal. In the event of the loss of some samples during the collection process, the averaging procedure nevertheless yields a result. This method serves to mitigate the impact of signal loss during the data collection process, therefore enhancing the accuracy and reliability of the captured data. The testing was conducted on a Dell XPS-9310 laptop, which is equipped with a high-performance 4.2 GHz processor and 8 GB of RAM. The data were recorded from a group of 14 students (7 male, 7 female) aged 18-22 years from Phenikaa University in Vietnam. The subjects' heights ranged from 1.7 meters to 1.9 meters, and their weights from 42 to 67 kilograms. To ensure a broader dataset, two additional participants, aged 65 and 72, were included. The subjects' heights were 1.5 meters and 1.6 meters, and their weights were 42 kilograms and 46 kilograms, respectively. The device was positioned at the waist and collected data at a sampling frequency of 50 Hz. Six activities were recorded: standing (40 minutes), sitting (30 minutes), jogging (10 minutes), walking (22 minutes), going downstairs (20 minutes), and going upstairs (20 minutes).

The sampling frequency of 50 Hz is a standard practice in the field of HAR [24, 25]. Furthermore, the selection of 50 Hz is consistent with the Nyquist theorem, which states that the sampling frequency should be at least twice the highest frequency present in the signal to prevent aliasing. Given that the majority of human motion frequencies of interest are well below 25 Hz [26], a 50 Hz sampling rate is deemed appropriate. Prior to data collection, the tri-axial acceleration sensors will undergo calibration and be subjected to a Kalman filter to eliminate noise, superfluous components, and errors. Further details regarding the Kalman filter and calibration can be found in [27]. To evaluate the efficacy of the algorithm, a publicly accessible dataset was employed and a comprehensive account of the activities is presented in Table I.

TABLE I. ACTIVITIES DESCRIPTION

| Activity | Explaination |
|---|---|
| Standing | Standing naturally, stable. |
| Sitting | Sitting on a chair. |
| Jogging | Running with an even pace and slowly. |
| Walking | Normal walking. |
| Downstairs | Walking down a staircase by taking individual steps. |
| Upstairs | Walking up a staircase by taking individual steps. |

In accordance with the model development outlined in Figure 1 and the construction of hardware compliant with the specified requirements, as illustrated in Figure 2, the action classification algorithm is loaded onto the microcontroller within the wearable devices. The PIC18F4520 microcontroller serves as the central processing unit in the proposed device. Once the devices have been installed, they are securely fastened to the volunteers' belts. This research project is focused on the development of optimized algorithms for low-performance, cost-efficient microcontrollers (commonly referred to as "Tiny Machine Learning"). The objective is to create algorithms that are suitable for use in embedded systems and IoT devices. The objective of this study was to embed the algorithm directly into the device in order to enable real-time activity recognition. Consequently, the capacity to program these wearable devices for real-time activity recognition constituted a pivotal element of this research. KNN is capable of adapting effectively to non-linear data and can be readily updated when new data become available. This contributes to an improvement in the accuracy and robustness of the model. Accordingly, the proposed activity classification algorithm for embedding wearable devices is the KNN algorithm. The implementation of the KNN algorithm on devices is accomplished beginning with the determination of the value of parameter *k*, which represents the number of neighbors *n* in the KNN model. The number of neighbors, represented by the parameter *k*, denotes the number of data points deemed most relevant for prediction. A suitable value for *k* is often identified through empirical testing with different parameter values. Furthermore, the weights parameter can be modified to delineate the manner in which the weights of neighboring points impact the prediction. Potential options include "uniform," which is the default setting and assigns equal weight to all points, and "distance," wherein points in closer proximity are given greater weight. In conclusion, the algorithm parameter defines the algorithm utilized to calculate the nearest neighbors. The optimization of these parameters improves the performance and accuracy of the KNN model.

### *C. Feature Selection*

#### *1) Sliding Window*

Subsequently, the data obtained from the accelerometer are divided into smaller segments, each of which encompasses information pertinent to specific activities. The necessity for segmentation arises from the inherent sequential nature of activities, which presents a challenge in defining activity start and end points. This study employs a sliding window technique for segmentation, as it is a simple and effective approach in the context of real-time activity recognition [26]. In this approach, the signal from the 3-DOF ADXL345 accelerometer is divided into fixed-size time windows. The selected window size has a considerable impact on the system's recognition accuracy. To identify the optimal window size, experiments were conducted with windows ranging from 1 second to 20 seconds, with varying overlap percentages between zero and 90%. The preliminary data analysis employed a publicly accessible dataset [28]. The dataset provides movement data for a range of everyday activities, collected from a single device, specifically, a smartphone's accelerometer. The volunteers carried the phone in their pockets while performing the designated actions. The dataset, which was collected from 36 participants, included data on six activities, such as standing, sitting, jogging, walking, going downstairs, and going upstairs. The testing phase involved 16 volunteers wearing the device on their waists, as shown in Figure 2, and performing the activities listed in Table I. Table II presents a comparison of the total number of activity observations obtained from the public data and the collected data. The public data set was extracted using a sliding window size of 15 seconds and a 40% window overlap ratio, resulting in a total of 10,476 observations. With regard to the data collected, the number of observations recorded with a window size of eight seconds and a 40% overlap is 9,745.

TABLE II. ACTIVITY DATA FROM THE PUBLIC DATASET

| Activity pattern | Total public | Total collected |
|---|---|---|
| Standing | 445 | 2,271 |
| Sitting | 548 | 1,769 |
| Jogging | 3,253 | 1,563 |
| Walking | 4,067 | 2,209 |
| Downstairs | 978 | 905 |
| Upstairs | 1,185 | 1,028 |
| Total | 10,476 | 9,745 |

#### *2) Feature Extraction*

In order to achieve optimal classification performance, it is of great importance to select features that contain information regarding their level of activity. Figure 3 displays the data distribution of the initial 400 data samples obtained from the *x*-axis acceleration of each activity. The figure facilitates the identification and selection of suitable features for the purpose of activity classification.

For each volunteer, data are collected separately, while the collection process takes great care to capture each movement in minute detail. Furthermore, participants were selected with great care to ensure a high degree of homogeneity in terms of age, height, and weight. This strategy serves to minimize potential discrepancies between data from male and female groups, as well as between younger and older participants. The values observed in the static states, such as standing and sitting, exhibit a tendency to cluster within a relatively confined range. For example, the data points for the "standing" state cluster around 300, with an approximate value of zero g on the *x*-axis (where 1 g = 9.8 m/s²). Similarly, the data points for the "downstairs" state are concentrated between -0.4 g and 0.3 g on the x-axis. This allows the selection of two features, Var and SAV, to measure data concentration. These features effectively differentiate between static and dynamic states, and they help to classify them further.

In contrast, dynamic states, such as jogging and walking, as well as activities, such as going downstairs and upstairs, exhibit a more extensive range of values in comparison to their static counterparts. In order to capture this difference, the P2P feature was selected, with the objective of determining the maximum and minimum values within the aforementioned dynamic states. Moreover, Figure 3 exhibits the existence of variations within the context of dynamic states. The data set pertaining to

walking exhibits the most significant density of values, with a sparse distribution on both ends of the spectrum. In contrast, the distribution of values for jogging is more uniform than that of walking. The skew was selected as a feature for the analysis of the data distribution relative to the mean, thus enabling the differentiation between these activities. It is noteworthy that the activities of jogging, walking downstairs, and walking upstairs exhibit a higher concentration of values around zero g on the x-axis relative to the activities of sitting and standing. This characteristic enables the Skew feature to be a valuable tool for characterizing these activities. In conclusion, based on the observations in Figure 3, four statistical features, Var, SAV, Skew, and P2P were selected for further analysis. These features are effective in capturing the distinctive characteristics of different activities.

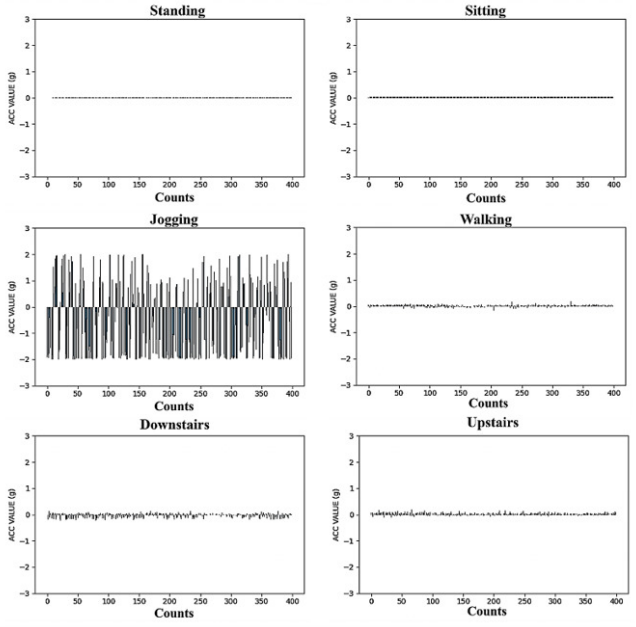



Fig. 3. The histogram depicts a portion of data along the *x*-axis.

Table III presents the formulas for the four features, which were calculated on the *x*-axis, *y*-axis, and *z*-axis using the same formulas. Tables IV, V, and VI present the results of the acceleration data calculated according to the Var, SAV, Skew, and P2P statistical features for operations in all three dimensions (*x*, *y*, and *z*).

TABLE III. FEATURE FORMULAS FOR ACCELERATION DATA IN THE *X*-AXIS

| Feature | Formula | |
|---|---|---|
| Var | $s^2 = \frac{1}{N-1}\sum_{i=1}^{N}(x_i - \overline{x})^2$ | (1) |
| SAV | $S = \sum_{i=1}^{N}\lvert x_i \rvert$ | (2) |
| Skew | $\frac{N}{(N-1)(N-2)}\sum_{i=1}^{N}\left(\frac{x_i - \overline{x}}{s}\right)^3$ | (3) |
| P2P | $xmin_{max}$ | (4) |

where, $x_i$ are data values; $N$ is the number of data values, $\overline{x}$ is the average value of $x_i$, $x_{max}$ is the largest value in the dataset, and $x_{min}$ is the smallest value in the dataset.

Table V reveals notable differences in the activities of standing, sitting, jogging, walking, going downstairs, and going upstairs. The highest Var is observed when the subject is standing (6.4), and the lowest when the subject is walking (0.0016), indicating that the data fluctuate more when the subject is standing in comparison to walking. The highest absolute value is observed in the case of jogging (286.15), while the lowest is seen in the case of standing (0.43). This indicates that fluctuations in jogging activity are larger in comparison to fluctuations in the activity of remaining stationary. The negative Skew observed for the activities of standing, sitting, walking, and going downstairs is indicative of a left-skewed distribution of the data for these activities. In contrast, the data for jogging and going upstairs exhibit a right-skewed distribution. The greatest range in P2P is observed in jogging (3.99), with the lowest occurring in standing (0.02). This indicates that the amplitude of fluctuations in jogging activity is significantly larger than when standing. In general, the data obtained from the standing position exhibit a high variance and low absolute values. The sitting position demonstrates an average variance and a markedly negative skew. The data obtained from jogging display the most pronounced fluctuations, while the walking position exhibits minimal fluctuations and a very low variance. The data obtained from the downstairs and upstairs positions demonstrate moderate and stable fluctuations, respectively. Tables V and VI present the acceleration data in the y-axis and z-axis, respectively.

TABLE IV. ACCELERATION IN THE *X*-AXIS (G)

| Feature | Standing | Sitting | Jogging | Walking | Downstairs | Upstairs |
|---|---|---|---|---|---|---|
| Var | 6.4 | 3 | 2.12 | 0.0016 | 0.007 | 0.025 |
| SAV | 0.43 | 4.38 | 286.15 | 7.41 | 14.3 | 8.7 |
| Skew | -0.58 | -9.67 | 0.45 | -0.54 | -0.46 | 0.19 |
| P2P | 0.02 | 0.027 | 3.99 | 0.19 | 0.37 | 0.3 |

TABLE V. ACCELERATION IN THE *Y*-AXIS (G)

| Feature | Standing | Sitting | Jogging | Walking | Downstairs | Upstairs |
|---|---|---|---|---|---|---|
| Var | 2.29 | 4.22 | 0.003 | 0.0015 | 0.0002 | 0.0026 |
| SAV | 20.4 | 2.43 | 369.16 | 41.2 | 13.21 | 46.4 |
| Skew | -4.75 | -10.1 | -1.68 | 0.17 | -1.24 | -1.25 |
| P2P | 0.01 | 0.074 | 3.07 | 0.17 | 0.08 | 0.232 |

TABLE VI. ACCELERATION IN THE *Z*-AXIS (G)

| Feature | Standing | Sitting | Jogging | Walking | Downstairs | Upstairs |
|---|---|---|---|---|---|---|
| Var | 5.24 | 0.0001 | 0.42 | 0.004 | 0.0016 | 0.006 |
| SAV | 0.93 | 19.8 | 96.67 | 9.4 | 6.17 | 15.86 |
| Skew | -0.15 | 13.75 | -0.51 | 2.05 | 0.003 | 0.12 |
| P2P | 0.02 | 0.2 | 3.07 | 0.42 | 0.32 | 0.41 |

In consideration of the limitations concerning real-time activity classification on low-cost, low-performance microcontrollers, the selected features (Var, SAV, Skew, and P2P) are deemed to be particularly well-suited to the data set under examination. Subsequently, the dataset was divided into two distinct sets: a training set and a testing set. The two sets were created by randomly selecting observations from the

original dataset, maintaining a 60/40 split (60% for training and 40% for testing). In order to gain deeper insights from the training data and prepare them for model building, this study employs the use of t-Distributed Stochastic Neighbor Embedding (t-SNE). This dimensionality reduction technique transforms the high-dimensional data points into a two- or three-dimensional space for visualization purposes. Figure 4(a) illustrates a notable degree of overlap between activities, which suggests a challenging classification task. In contrast, Figure 4(b) shows a more distinct differentiation between activities following the incorporation of all four features (Var, SAV, Skew, and P2P). This underscores the benefit of utilizing multiple features for enhanced classification precision. In general, the use of all four features yielded superior results in activity recognition compared to the use of individual features. It is noteworthy that the classification of activities such as standing (orange area) and sitting (blue area) was relatively straightforward. Furthermore, both jogging (pink area) and walking (green area) demonstrated a certain degree of separability. However, the public dataset demonstrated some degree of overlap between the actions classified as "downstairs" (purple area) and "upstairs" (brown area).

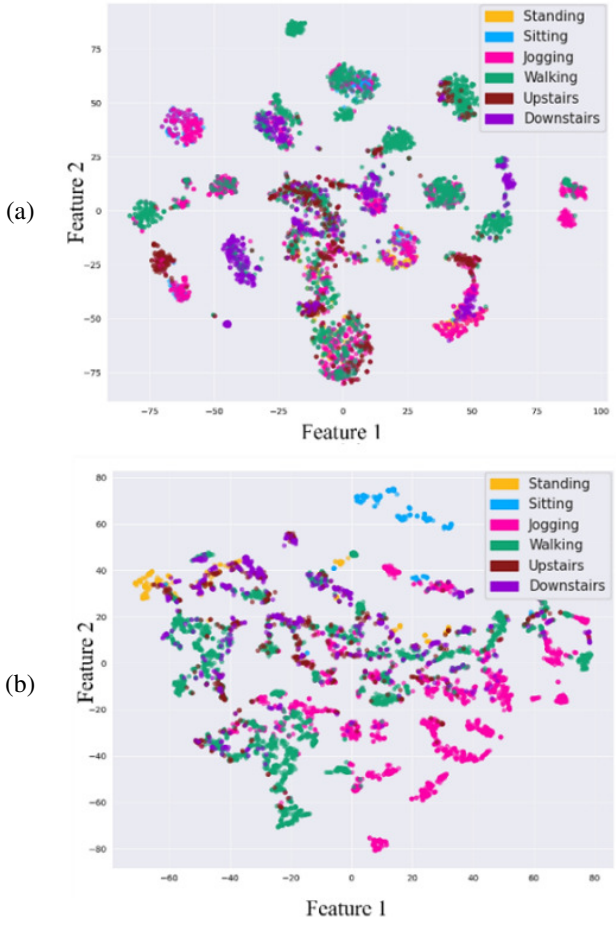



Fig. 4. Public data (a) without features and (b) with features in 2D space.

The same four features (Var, SAV, Skew, and P2P) were applied to the collected dataset in this study, in accordance with the methodology previously outlined. Figure 5 provides a visual representation of the data points' distribution, as generated by the t-SNE algorithm and depicted in a histogram. The figure illustrates a discernible differentiation between the majority of activities, although some degree of overlap persists between jogging (pink area), walking (green area), downstairs (purple area), and upstairs (brown area).

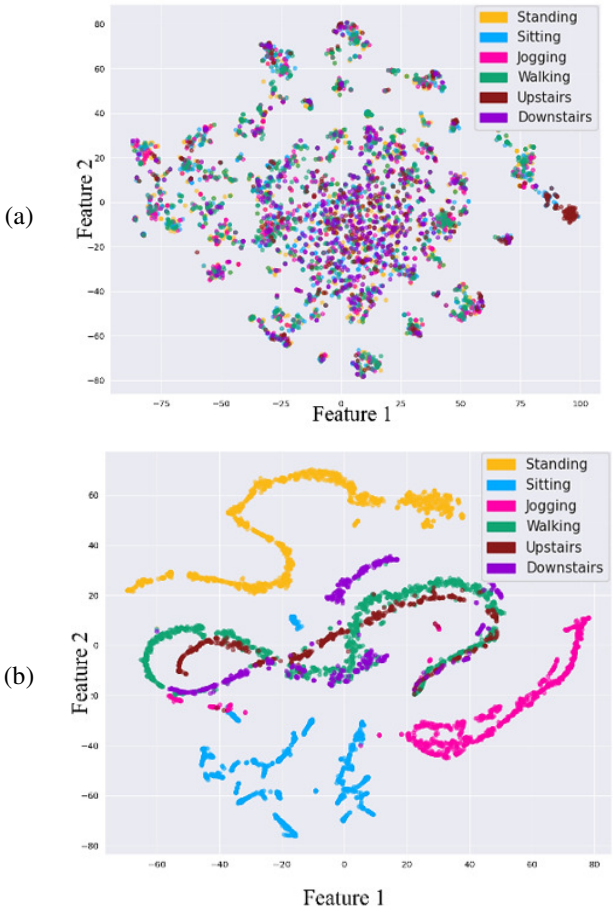



Fig. 5. Collected data without features (a) and with features (b) in 2D space.

### *D. Recognition Activities*

The extracted features (Var, SAV, Skew, and P2P) were used for the training of the machine learning models. In addition to the KNN method, three other common classification algorithms were evaluated: the Decision Tree (DT), SVM, and Logistic Regression (LR). The DT algorithm constructs a model by recursively splitting the data based on feature values, resulting in a tree-like structure of decisions. SVM seeks to identify the optimal hyperplane that maximizes the margin between different classes, employing a range of kernel functions to address non-linear separations. LR is used for binary classification by modeling the probability of a class using a logistic function, resulting in a decision boundary through a linear combination of input features. These

algorithms were implemented using the scikit-learn library (often abbreviated as sklearn) in Python.

In order to evaluate the efficacy of these models, a confusion matrix was employed. The matrix offers a transparent representation of the model's classification efficacy. The following section presents the formulas used for the assessment of the tests:

$$acc_i = \frac{TPs_i + TN_i}{TPs_i + FPs_i + TNs_i + FNs_i} \quad (5)$$

$$sen_i = \frac{TPs_i}{TPs_i + FNs_i} \quad (6)$$

$$PPV_i = \frac{TPs_i}{TPs_i + FPs_i} \quad (7)$$

$$NPV_i = \frac{TNs_i}{TNs_i + FPs_i} \quad (8)$$

where $i$ represents a class (Standing, Sitting, Jogging, Walking, Downstairs or Upstairs), True Positive (*TP*) indicates the number of times the model correctly predicted an activity, False Positive (*FP*) represents the number of times the model incorrectly predicted an activity, False Negative (*FN*) is the number of times the model missed an activity and True Negative (*TN*) is the number of times the model correctly predicted that an activity did not occur.

## III. RESULTS AND DISCUSSION

Figures 6 and 7 show the overall accuracy and sensitivity of each DT, KNN, SVM, and LR, respectively, on the public dataset. The performance of each classifier varies with different window sizes (5 seconds, 10 seconds, 15 seconds, 20 seconds). Among these classifiers, the KNN stands out in terms of classification performance, achieving an accuracy of 96% with a window size of 15 seconds and a sensitivity of 87% with a window size of 5 seconds.

As evidenced in Figure 8 and Figure 9, the collected dataset achieves better accuracy and sensitivity than the publicly available dataset. This improvement can be attributed to several factors, such as:

- Data source: unlike the public dataset, which collected data from phones in pockets, this study's data come from a device worn securely on a belt, minimizing motion artifacts.
- Structured collection: this study's data collection process involves controlled individual actions, resulting in less variation compared to everyday activity variations.
- Homogeneous participants: the proposed system was tested on a group with similar ages and behaviors, further reducing data variability.

Together, these factors contribute to the significantly higher accuracy (100% for the 15-second window KNN classifier) and sensitivity (96% for the 15-second window KNN classifier) observed in the proposed dataset. In simpler terms, the careful control and consistency within this study's data collection process led to these exceptional results.

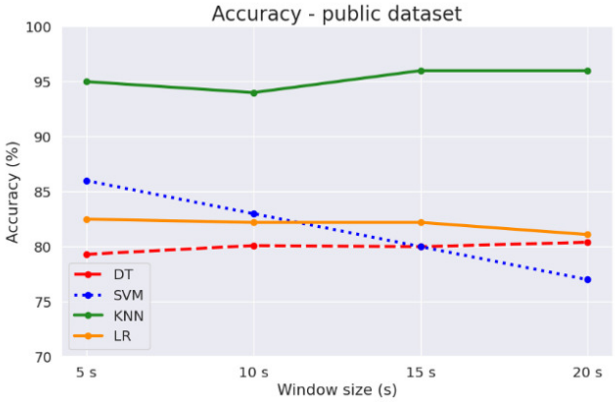



Fig. 6. Accuracy of DT, SVM, KNN, and LR on the public dataset.

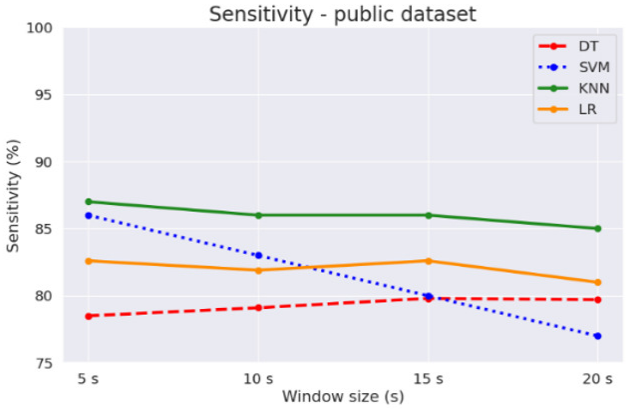



Fig. 7. Sensitivity of DT, SVM, KNN, and LR on the public dataset.

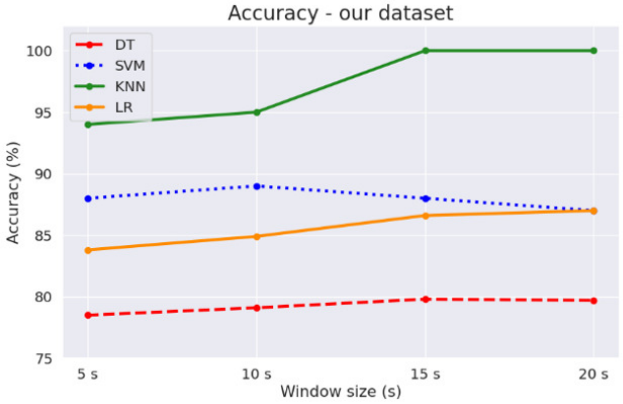



Fig. 8. Accuracy of DT, SVM, KNN, and LR on the collected dataset.

After evaluating the optimal computational objective and its suitability for users who are less prone to abrupt changes in behavior (such as older users or those in recovery), the KNN algorithm emerges as the preferred choice for classifying direct actions on microcontrollers in this study. Tables VII and VIII display detailed results of using the KNN algorithm for the classification of operations on both the public and the collected datasets. As presented in Table VII, the model demonstrates a high degree of accuracy in distinguishing between various

activities, particularly for "standing," "sitting," "jogging," and "walking." Specifically, 141 out of 149 samples for "standing" are correctly classified, 181 out of 184 samples for "sitting," 1,014 out of 1,085 samples for "jogging," and 1,264 out of 1,356 samples for "walking." However, the model demonstrates a greater degree of ambiguity when distinguishing between activities with similar characteristics, particularly between "downstairs" and "upstairs." With regard to the activity designated "downstairs," only 175 out of 327 samples were correctly identified, while for the activity designated "upstairs," 236 out of 396 samples were correctly classified. The primary misclassifications occur between "downstairs," "walking," and "upstairs," indicating a need for improvement in distinguishing between these activities.

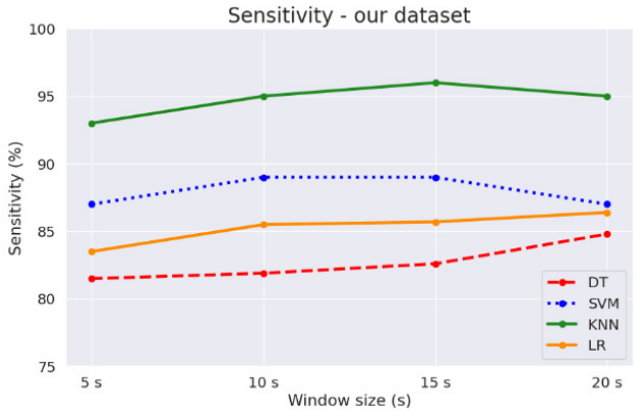



Fig. 9. Sensitivity of DT, SVM, KNN, and LR on the collected dataset.

As with the public dataset, Table VIII demonstrates that the model is highly effective at differentiating between various activities, particularly "standing" and "sitting." The vast majority of samples were correctly classified, with 757 out of 758 instances of "standing" and 589 out of 590 instances of "sitting" having been correctly identified. The model also demonstrates proficiency in differentiating between "jogging" and "walking" activities. It correctly classified 511 out of 522 samples for "jogging" and 686 out of 737 samples for "walking." However, a notable degree of inaccuracy is observed in the differentiation between the terms "downstairs" and "upstairs." Regarding, the "downstairs" category, only 267 out of 302 samples were correctly identified, whereas 314 out of 344 samples were correctly classified in the "upstairs" category. The majority of misclassifications occur between the "downstairs" and "walking" categories, as well as between the "upstairs" and "walking" categories. The accuracy of activity classification is presented in Tables IX and X, which pertain to the public and collected datasets, respectively. It is noteworthy that the KNN classifier demonstrated an accuracy exceeding 90% for all activities in both the public and collected datasets. The results presented here are founded upon the mathematical formulas (5) to (8).

The KNN classifier demonstrated remarkable accuracy and NPV rates exceeding 94% on the public dataset and surpassing 97% on the collected dataset across all operations. The sensitivity and positive predictive value exhibited noteworthy values, exceeding 90% for all classes with the exception of "downstairs" and "upstairs." It is noteworthy that the "standing", "sitting", and "jogging" activities demonstrated near-absolute values on the collected dataset. As can be observed in Tables IX and X, the activities of "downstairs" and "upstairs" displayed markedly inferior classification accuracy in comparison to the remaining activities. This is due to the fact that the acceleration data for descending and ascending stairs are similar. The overall performance of the dataset was evaluated using both the macro-average and micro-average methods, as observed in Table XI.

TABLE VII. CONFUSION MATRIX OF KNN ALGORITHM ON THE PUBLIC DATASET

| **Observed activity** | **Predicted activity** | | | | | | **Total** |
|---|---|---|---|---|---|---|---|
| | Standing | Sitting | Jogging | Walking | Downstairs | Upstairs | |
| Standing | 141 | 1 | 0 | 0 | 5 | 2 | 149 |
| Sitting | 0 | 181 | 1 | 0 | 0 | 2 | 184 |
| Jogging | 0 | 0 | 1,014 | 54 | 13 | 4 | 1,085 |
| Walking | 1 | 0 | 37 | 1,264 | 22 | 32 | 1,356 |
| Down stairs | 5 | 0 | 35 | 71 | 175 | 41 | 327 |
| Upstairs | 3 | 0 | 25 | 89 | 43 | 236 | 396 |
| Total | 150 | 182 | 1112 | 1478 | 258 | 317 | 3,497 |

TABLE VIII. CONFUSION MATRIX OF KNN ALGORITHM ON THE COLLECTED DATASET

| **Observed activity** | **Predicted activity** | | | | | | **Total** |
|---|---|---|---|---|---|---|---|
| | Standing | Sitting | Jogging | Walking | Downstairs | Upstairs | |
| Standing | 757 | 0 | 0 | 0 | 0 | 1 | 758 |
| Sitting | 0 | 589 | 0 | 1 | 0 | 0 | 590 |
| Jogging | 0 | 9 | 511 | 2 | 0 | 0 | 522 |
| Walking | 0 | 0 | 0 | 686 | 19 | 32 | 737 |
| Down stairs | 3 | 0 | 1 | 30 | 267 | 1 | 302 |
| Upstairs | 0 | 0 | 0 | 17 | 13 | 314 | 344 |
| Total | 760 | 598 | 512 | 736 | 299 | 348 | 3,253 |

TABLE IX. MODEL PERFORMANCE ON THE PUBLIC DATASET

| **Activity pattern** | **Algorithm performance** | | | |
|---|---|---|---|---|
| | Accuracy (%) | Sensitivity (%) | PPV (%) | NPV (%) |
| Standing | 99.51 | 94.63 | 94.00 | 99.76 |
| Sitting | 99.89 | 98.37 | 99.45 | 99.91 |
| Jogging | 95.17 | 93.46 | 91.19 | 97.02 |
| Walking | 91.25 | 93.22 | 85.52 | 95.44 |
| Downstairs | 93.28 | 53.52 | 67.83 | 95.31 |
| Upstairs | 93.11 | 59.60 | 74.45 | 94.97 |

TABLE X. MODEL PERFORMANCE ON THE COLLECTED DATASET

| **Activity pattern** | **Algorithm performance** | | | |
|---|---|---|---|---|
| | Accuracy (%) | Sensitivity (%) | PPV (%) | NPV (%) |
| Standing | 99.88 | 99.87 | 99.61 | 99.96 |
| Sitting | 99.69 | 99.83 | 98.49 | 99.96 |
| Jogging | 99.63 | 97.89 | 99.80 | 99.96 |
| Walking | 96.90 | 93.08 | 93.21 | 97.97 |
| Downstairs | 97.94 | 88.41 | 89.30 | 98.82 |
| Upstairs | 98.03 | 91.28 | 90.23 | 98.97 |

TABLE XI. ASSESSMENT OF THE KNN CLASSIFICATION MODEL

| Evaluation Metrics | Micro-average (%) | | Macro-average (%) | |
|---|---|---|---|---|
| | Public dataset | Collected dataset | Public dataset | Collected dataset |
| Accuracy | 86.1 | 96.03 | 85.4 | 95.1 |
| Sensitivity | 82.1 | 95.06 | 82.1 | 95.06 |
| PPV | 85.4 | 95.1 | 85.4 | 95.1 |
| NPV | 85.4 | 95.1 | 82.1 | 95 |

This analysis revealed that the KNN classifier showcased the highest performance. In particular, the micro-average method demonstrates remarkable efficacy, exceeding 85% on the public dataset and surpassing 96% on the collected dataset. A detailed examination of the data reveals that the micro-average accuracy is 86.1% for the public dataset and 96.03% for the collected dataset. Similarly, the sensitivity is 82.1% for the public dataset and 95.06% for the collected dataset. The PPV is 85.4% for the public dataset and 95.1% for the collected dataset. Finally, the NPV is 85.4% for the public dataset and 95.1% for the collected dataset. The macro-average accuracy is 85.4% (public) and 95.1% (collected), while the sensitivity is 82.1% (public) and 95.06% (collected). The PPV is 85.4% (public) and 95.1% (collected), and the NPV is 82.1% (public) and 95% (collected). These findings illustrate the KNN classifier's resilience and efficacy in diverse datasets.

In machine learning, the process of extracting features from raw data represents a pivotal stage in the analytical pipeline. The efficacy of this process is contingent upon the model designer's comprehension of the dataset. While the inclusion of carefully selected features can enhance the accuracy of a model, the incorporation of irrelevant or superfluous features can impose an unnecessary computational burden, potentially impeding performance. This underscores the significance of these fundamental features in attaining optimal performance. Moreover, this study explored the incorporation of supplementary time-domain features into the original feature set. While this resulted in a decline in classification accuracy for the public dataset, the impact on the collected dataset was inconsequential. This indicates that for the particular objective of behavioral classification in the present study, the proposed feature set provides a well-balanced trade-off between classification performance and computational efficiency.

## IV. DISCUSSION

A considerable number of studies have broadened the scope of feature selection in order to more accurately identify human activities. Some studies have combined up to 43 features, as demonstrated by authors in [29], and even 64 features, as examined by authors in [30]. While the aforementioned studies also yielded noteworthy classification outcomes, with all of them exceeding 95%, the current study employed a more parsimonious set of features and yielded results that were commensurate with the data file utilized (e.g., 95.06% on the collected dataset). Moreover, the classification process will be conducted in real time on the microcontroller. The selection of appropriate features for the data is a crucial aspect of the recognition process, as evidenced by previous research [30-33]. Therefore, the preceding four basic features (Var, SAV, Skew, and P2P) are entirely aligned with the objectives of the research. Some researchers employed the use of accelerometer data and deep learning techniques for the purpose of classifying human activities. In [34], authors constructed a Convolutional Neural Network (CNN) model and modified the convolution kernel to adjust the characteristics of the three-axis acceleration signal. The experimental results demonstrated that the CNN is an effective approach, having achieved an average accuracy of 93.8%. A method based on One-Dimensional Convolutional Neural Networks (1D CNNs), which achieved an accuracy of 92.71%, was proposed in [35]. However, this study employed a limited range of human activity data, comprising only walking, running, and stationary poses, collected via smartphone accelerometers. The efficacy of LSTM in classifying six actions: standing, sitting, jogging, walking, going downstairs, and going upstairs, was evaluated. This was conducted on both a public dataset and a dataset collected from the present study's experiments. The test results demonstrate that LSTM achieves an accuracy of greater than 98% and a sensitivity of greater than 96%. These findings substantiate the assertion that LSTM (or deep learning) is a valuable approach for addressing the HAR problem. However, it should be noted that deep learning requires a significant investment of time for both training and testing. Therefore, the deployment of deep learning may not be a cost-effective solution for microcontrollers such as the PIC18F4520.

Some researchers explored the use of built-in accelerometers for activity classification using smartphones and designed a recognition system for activities, such as walking, jogging, jumping, climbing stairs, descending stairs, sitting, standing, and cycling [36]. The system was evaluated using two feature selection methods (OneRAttributeEval and ReliefF AttributeEval) and six classification algorithms (J48, K-Star, Bayes Net, Naïve Bayes, Random Forest, and k-NN). A recognition rate of 94% was achieved by the combination of 15 features and the k-NN algorithm. Authors in [6] deployed the KU-HAR dataset with 18 activities of 90 individuals. They used 66 features and 7 machine learning classifiers: Gradient Boosting (GB), KNN, DT, Random Forest (RF), XGBoost, LightGBM, and Catboost. The study achieved an accuracy and sensitivity of 95.8% with LightGBM. In another study [28], authors collected data from twenty-nine users who carried an Android phone in their pocket while performing six activities, walking, jogging, climbing stairs, descending stairs, sitting, and standing. They evaluated three learning algorithms (logistic regression, J48, and multilayer perceptron), and their results showed an overall accuracy of above 90%, except for the stairs-up vs. stairs-down case, which was more difficult. However, the method using smartphones still suffers from reduced accuracy due to interference from daily activities such as phone calls, web browsing, and texting [23]. With the same operations, the proposed dataset gives better results and fewer features are also utilized. The current research has healthcare applications for the elderly, people recovering from surgery, or people involved in accidents. Therefore, the data are personalized and focused on activity analysis, and the scope is narrowed to these six specific behaviors for greater accuracy. These six behaviors were deemed sufficient for the application while ensuring real-time sorting and cost optimization.

There is a notable discrepancy in the overall accuracy of the public and collected datasets, with the former exhibiting a lower accuracy rate of 82.1% compared to the latter's considerably higher rate of 95.06%. This difference can be attributed to a number of factors. Firstly, the public dataset is derived from a larger pool of data collectors. The public dataset involved a total of 36 participants, whereas this study included only 16. This larger number of participants yields a more diverse data set in the public sample compared to the collected data. Secondly, the data collection process itself differed between the two datasets. In the public dataset, participants were instructed to place their phones in their pockets and rely on the built-in accelerometer. In contrast, the volunteers participating in this study were instructed to securely fasten the device to their belts. The discrepancy in device placement is likely responsible for the enhanced data quality observed in the collected dataset, which in turn contributed to the improved classification performance achieved when utilizing the identical proposed model. It should be noted, however, that the present study is not without limitations. That is, it is limited to a single age group with similar physical characteristics. It is our contention that if the subjects were children or individuals aged 30–50 or older, the activity thresholds would differ. Two activities, "downstairs" and "upstairs," exhibited the lowest accuracy rates in the study's activity classification results, with 88.41% and 91.28%, respectively, using the collected dataset. The specific result is consistent with the findings of previous studies. This can be attributed to the fact that the acceleration data of these two activities are quite similar. In daily activities, ascending and descending stairs typically require an equivalent amount of time. Moreover, the objective of this research is to facilitate cost-effective healthcare for patients recuperating from injuries or for the elderly. It is essential to consider the frequency of activities over an extended period. In the future, there is an intention to combine the two activities of "downstairs" and "upstairs" into a single activity.

## V. CONCLUSIONS

This research has successfully developed a cost-effective, real-time Human Activity Recognition (HAR) system. The incorporation of an effective, low-complexity algorithm into a low-performance microcontroller for the classification of fundamental human activities paves the way for promising new avenues of research in this field. The experimental results demonstrate the efficacy of Decision Tree (DT), K-Nearest Neighbors (KNN), Support Vector Machine (SVM), and Logistic Regression (LR) when utilizing disparate window sizes. The KNN algorithm is particularly well-suited to real-time implementation using digital data, such as those obtained from accelerometers, due to their high computational efficiency. The KNN algorithm exhibits the highest classification performance when the window size is set to 15 seconds. By leveraging four principal features, Variance (Var), Sum of Absolute Values (SAV), Skewness (Skew), and Peak-to-Peak (P2P), the system is capable of accurate classification of fundamental daily activities. The proposed system has the potential to be extended to recognize a wider range of activities, including more complex movements. Furthermore, while this study concentrated on time-domain features derived from the accelerometer data, future iterations could investigate features in the frequency domain. The combination of both time and frequency domain features may result in enhanced performance. The system will be tested in actual healthcare environments to gather practical insights and validate its effectiveness. Collaborations with healthcare providers and institutions could facilitate the integration of this technology into patient monitoring systems, thereby providing valuable support for tailored rehabilitation programs and enhancing the quality of life for the elderly. The objective of this research is to make incremental improvements to the system's capabilities with the aim of contributing to the advancement of activity recognition and its applications in healthcare and other fields.

## ACKNOWLEDGMENT

This work was supported by Hanoi National University of Education under Grant SPHN22-21.

## REFERENCES

[1] R. Singh, A. K. S. Kushwaha, Chandni, and R. Srivastava, "Recent trends in human activity recognition – A comparative study," *Cognitive Systems Research*, vol. 77, pp. 30–44, Jan. 2023, https://doi.org/10.1016/j.cogsys.2022.10.003.

[2] F. Demrozi, G. Pravadelli, A. Bihorac, and P. Rashidi, "Human Activity Recognition Using Inertial, Physiological and Environmental Sensors: A Comprehensive Survey," *IEEE Access*, vol. 8, pp. 210816–210836, 2020, https://doi.org/10.1109/ACCESS.2020.3037715.

[3] N. Gupta and B. B. Agarwal, "Recognition of Suspicious Human Activity in Video Surveillance: A Review," *Engineering, Technology & Applied Science Research*, vol. 13, no. 2, pp. 10529–10534, Apr. 2023, https://doi.org/10.48084/etasr.5739.

[4] P. Gupta and T. Dallas, "Feature Selection and Activity Recognition System Using a Single Triaxial Accelerometer," *IEEE Transactions on Biomedical Engineering*, vol. 61, no. 6, pp. 1780–1786, Jun. 2014, https://doi.org/10.1109/TBME.2014.2307069.

[5] A. Dey, S. Biswas, and D.-N. Le, "Workout Action Recognition in Video Streams Using an Attention Driven Residual DC-GRU Network," *Computers, Materials & Continua*, vol. 79, no. 2, pp. 3067–3087, 2024, https://doi.org/10.32604/cmc.2024.049512.

[6] M. Alanazi, R. S. Aldahr, and M. Ilyas, "Human Activity Recognition through Smartphone Inertial Sensors with ML Approach," *Engineering, Technology & Applied Science Research*, vol. 14, no. 1, pp. 12780–12787, Feb. 2024, https://doi.org/10.48084/etasr.6586.

[7] S.-W. Lee and K. Mase, "Activity and location recognition using wearable sensors," *IEEE Pervasive Computing*, vol. 1, no. 3, pp. 24–32, Jul. 2002, https://doi.org/10.1109/MPRV.2002.1037719.

[8] J. Mantyjarvi, J. Himberg, and T. Seppanen, "Recognizing human motion with multiple acceleration sensors," in *2001 IEEE International Conference on Systems, Man and Cybernetics. e-Systems and e-Man for Cybernetics in Cyberspace*, Oct. 2001, vol. 2, pp. 747–752, https://doi.org/10.1109/ICSMC.2001.973004.

[9] S. Bhattacharjee, S. Kishore, and A. Swetapadma, "A Comparative Study of Supervised Learning Techniques for Human Activity Monitoring Using Smart Sensors," in *2018 Second International Conference on Advances in Electronics, Computers and Communications (ICAECC)*, Feb. 2018, pp. 1–4, https://doi.org/10.1109/ICAECC.2018.8479436.

[10] A. Blum and T. Mitchell, "Combining labeled and unlabeled data with co-training," in *Proceedings of the eleventh annual conference on Computational learning theory*, New York, NY, USA, Jul. 1998, pp. 92–100, https://doi.org/10.1145/279943.279962.

[11] Z.-H. Zhou and M. Li, "Semi-supervised learning by disagreement," *Knowledge and Information Systems*, vol. 24, no. 3, pp. 415–439, Sep. 2010, https://doi.org/10.1007/s10115-009-0209-z.

[12] Q. Zhu, Z. Chen, and Y. C. Soh, "A Novel Semisupervised Deep Learning Method for Human Activity Recognition," *IEEE Transactions*

*on Industrial Informatics*, vol. 15, no. 7, pp. 3821–3830, Jul. 2019, https://doi.org/10.1109/TII.2018.2889315.

[13] M. Stikic, D. Larlus, S. Ebert, and B. Schiele, "Weakly Supervised Recognition of Daily Life Activities with Wearable Sensors," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 33, no. 12, pp. 2521–2537, Dec. 2011, https://doi.org/10.1109/TPAMI.2011.36.

[14] Y. Kim and T. Moon, "Human Detection and Activity Classification Based on Micro-Doppler Signatures Using Deep Convolutional Neural Networks," *IEEE Geoscience and Remote Sensing Letters*, vol. 13, no. 1, pp. 8–12, Jan. 2016, https://doi.org/10.1109/LGRS.2015.2491329.

[15] S. Kamal and A. Jalal, "A Hybrid Feature Extraction Approach for Human Detection, Tracking and Activity Recognition Using Depth Sensors," *Arabian Journal for Science and Engineering*, vol. 41, no. 3, pp. 1043–1051, Mar. 2016, https://doi.org/10.1007/s13369-015-1955-8.

[16] Q. Liu, W. Zhang, H. Li, and K. N. Ngan, "Hybrid human detection and recognition in surveillance," *Neurocomputing*, vol. 194, pp. 10–23, Jun. 2016, https://doi.org/10.1016/j.neucom.2016.02.011.

[17] M. H. Khan, K. Shirahama, M. S. Farid, and M. Grzegorzek, "Multiple human detection in depth images," in *2016 IEEE 18th International Workshop on Multimedia Signal Processing (MMSP)*, Sep. 2016, pp. 1–6, https://doi.org/10.1109/MMSP.2016.7813385.

[18] G. K. Vinay, S. M. Haque, R. V. Babu, and K. R. Ramakrishnan, "Sparse representation-based human detection: a scale-embedded dictionary approach," *Signal, Image and Video Processing*, vol. 10, no. 3, pp. 585–592, Mar. 2016, https://doi.org/10.1007/s11760-015-0781-5.

[19] X. Chen, K. Henrickson, and Y. Wang, "Kinect-Based Pedestrian Detection for Crowded Scenes," *Computer-Aided Civil and Infrastructure Engineering*, vol. 31, no. 3, pp. 229–240, 2016, https://doi.org/10.1111/mice.12163.

[20] S. Zhang, D. A. Klein, C. Bauckhage, and A. B. Cremers, "Fast moving pedestrian detection based on motion segmentation and new motion features," *Multimedia Tools and Applications*, vol. 75, no. 11, pp. 6263–6282, Jun. 2016, https://doi.org/10.1007/s11042-015-2571-z.

[21] N.-T. Nguyen, B.-H. Liu, S.-I. Chu, and H.-Z. Weng, "Challenges, Designs, and Performances of a Distributed Algorithm for Minimum-Latency of Data-Aggregation in Multi-Channel WSNs," *IEEE Transactions on Network and Service Management*, vol. 16, no. 1, pp. 192–205, Mar. 2019, https://doi.org/10.1109/TNSM.2018.2884445.

[22] N.-T. Nguyen, B.-H. Liu, V.-T. Pham, and T.-Y. Liou, "An Efficient Minimum-Latency Collision-Free Scheduling Algorithm for Data Aggregation in Wireless Sensor Networks," *IEEE Systems Journal*, vol. 12, no. 3, pp. 2214–2225, Sep. 2018, https://doi.org/10.1109/JSYST.2017.2751645.

[23] *ADXL345 Data Sheet*. Analog Devices.

[24] S. Ha and S. Choi, "Convolutional neural networks for human activity recognition using multiple accelerometer and gyroscope sensors," in *2016 International Joint Conference on Neural Networks (IJCNN)*, Jul. 2016, pp. 381–388, https://doi.org/10.1109/IJCNN.2016.7727224.

[25] Khimraj, P. K. Shukla, A. Vijayvargiya, and R. Kumar, "Human Activity Recognition using Accelerometer and Gyroscope Data from Smartphones," in *2020 International Conference on Emerging Trends in Communication, Control and Computing (ICONC3)*, Feb. 2020, pp. 1–6, https://doi.org/10.1109/ICONC345789.2020.9117456.

[26] N.-T. Nguyen, B.-H. Liu, and S.-Y. Wang, "Network Under Limited Mobile Sensors: New Techniques for Weighted Target Coverage and Sensor Connectivity," in *2017 IEEE 42nd Conference on Local Computer Networks (LCN)*, Oct. 2017, pp. 471–479, https://doi.org/10.1109/LCN.2017.52.

[27] P. Van Thanh *et al.*, "Development of a Real-Time, Simple and High-Accuracy Fall Detection System for Elderly Using 3-DOF Accelerometers," *Arabian Journal for Science and Engineering*, vol. 44, no. 4, pp. 3329–3342, Apr. 2019, https://doi.org/10.1007/s13369-018-3496-4.

[28] J. R. Kwapisz, G. M. Weiss, and S. A. Moore, "Activity recognition using cell phone accelerometers," *SIGKDD Explor. Newsl.*, vol. 12, no. 2, pp. 74–82, Mar. 2011, https://doi.org/10.1145/1964897.1964918.

[29] C. Catal, S. Tufekci, E. Pirmit, and G. Kocabag, "On the use of ensemble of classifiers for accelerometer-based activity recognition," *Applied Soft Computing*, vol. 37, pp. 1018–1022, Dec. 2015, https://doi.org/10.1016/j.asoc.2015.01.025.

[30] G. Vavoulas, C. Chatzaki, T. Malliotakis, M. Pediaditis, and M. Tsiknakis, "The MobiAct Dataset: Recognition of Activities of Daily Living using Smartphones:," in *Proceedings of the International Conference on Information and Communication Technologies for Ageing Well and e-Health*, Rome, Italy, 2016, pp. 143–151, https://doi.org/10.5220/0005792401430151.

[31] A. Wang, H. Chen, C. Zheng, L. Zhao, J. Liu, and L. Wang, "Evaluation of Random Forest for Complex Human Activity Recognition Using Wearable Sensors," in *2020 International Conference on Networking and Network Applications (NaNA)*, Haikou City, China, Dec. 2020, pp. 310–315, https://doi.org/10.1109/NaNA51271.2020.00060.

[32] Y. Min, Y. Y. Htay, and K. K. Oo, "Comparing the Performance of Machine Learning Algorithms for Human Activities Recognition using WISDM Dataset," *International Journal of Computer (IJC)*, vol. 38, no. 1, pp. 61–72, May 2020.

[33] S. Kiran *et al.*, "Multi-Layered Deep Learning Features Fusion for Human Action Recognition," *Computers, Materials & Continua*, vol. 69, no. 3, pp. 4061–4075, 2021, https://doi.org/10.32604/cmc.2021.017800.

[34] Y. Chen and Y. Xue, "A Deep Learning Approach to Human Activity Recognition Based on Single Accelerometer," in *2015 IEEE International Conference on Systems, Man, and Cybernetics*, Hong Kong, China, Oct. 2015, pp. 1488–1492, https://doi.org/10.1109/SMC.2015.263.

[35] S.-M. Lee, S. M. Yoon, and H. Cho, "Human activity recognition from accelerometer data using Convolutional Neural Network," in *2017 IEEE International Conference on Big Data and Smart Computing (BigComp)*, Jeju, Korea, Feb. 2017, pp. 131–134, https://doi.org/10.1109/BIGCOMP.2017.7881728.

[36] E. Büber and A. M. Guvensan, "Discriminative time-domain features for activity recognition on a mobile phone," in *2014 IEEE Ninth International Conference on Intelligent Sensors, Sensor Networks and Information Processing (ISSNIP)*, Singapore, Apr. 2014, pp. 1–6, https://doi.org/10.1109/ISSNIP.2014.6827651.