Research Article

A Novel Martingale Based Model Using a Smartphone to Detect Gait Bout in Human Activity Recognition

Jonathan Etumusei ⓘ**, Jorge Carracedo Martinez** ⓘ**, and Sally McClean** ⓘ

Ulster University, UK

Correspondence should be addressed to Jonathan Etumusei; etumusei-j@ulster.ac.uk

Gait bout is when an individual performs certain physical activities such as walking or running. In the last few decades, the study of gait bout has led to substantial progress in treating gait impairment (neuropathic, myopathic, and parkinsonian) in a person. Recently, gait bout study has been improved by advancing smartphone technology. To perform gait bout tasks, two different human activity scenarios, such as walking upstairs and standing, are obtained using the axis orientation of a smartphone accelerometer. To capture the pattern of walking upstairs and standing, we utilize a smartphone device attached to the waist of 30 subjects within the age group from 19 to 48 years old. We propose a human activity recognition model known as the multivariate triple exponential weighted moving average of the martingale sequence using particle swarm optimization (MTMS(PSO)) in the experimental setup. MTMS(PSO) utilizes the martingale framework to capture gait bout in human activity recognition data. Firstly, MTMS(PSO) is an unsupervised learning method that uses smoothing techniques such as triple exponential smoothing to remove high-frequency noise from the processed activity times series, making the patterns more visible. Secondly, the activity recognition model involves computing a threshold for identifying gait bout. Thirdly, MTMS(PSO) uses logical precedent and particle swarm optimization to enhance accuracy and precision. As a result, the overall MTMS(PSO) accuracy and G-mean are 95.4% and 96.1%, respectively. In addition, MTMS(PSO) technique independently outperforms other traditional methods such as MRPM(PSO), MGM(PSO), and ELM.

1. Introduction

Ageing is a process that causes physical or physiological decline that could affect people's quality of life, resulting in injuries, decreased mental health, or reduction in physical activity. Human activity recognition (HAR) is an important concept in conventional computing as it applies to actual human condition challenges associated with eldercare and medicare. HAR is a research area that can track gait bout (GB) of an individual through the collection of context information about a user's condition and environment [1]. In this instance, a gait can be the pattern of walking, while GB can be defined as the period of continuous movement (walking or running) [2, 3]. In this case, GB is the time interval between gait initiation and termina-

tion. GB essentially requires repetition of stance-swing cycles [4]. Subsequently, if the foot is on the ground for more than the specified threshold period, it can be considered static, therefore non-GB.

Some research has been successful in discovering GB in HAR time series [5]. However, it is unclear how long an individual can pause movement (walking) within a GB before determining it as a separate bout. A pulse of $1-5$ seconds within GB can potentially alter the amount of GB and its duration included in the analysis [3]. Also, detecting composite gait bout still constitutes some challenges, such as identifying simultaneous movements. For example, an adult with gait impairment (neuropathic, myopathic, and parkinsonian) can be walking toward the kitchen while trying to answer a phone call. These movements are very complex,

making sequential analysis complicated. In addition, gait bout performed by an individual could be interwoven; for instance, a person can be distracted when there is a knock on the door while going to the restroom. After attending to the person at the door, the first person can walk back to the sitting room. Also, there may be ambiguities defining similar activities performed by the elderly. For example, "opening the cupboard" might represent "preparing food" or "tidying."

GB analysis could help understand the impact and challenges of human movement restriction or disability by identifying transitions and estimating the duration of physical activity. This information can be beneficial for investigating and developing new systems that can identify movement limitations. The world population is ageing, giving rise to neurological impairment and inactive conditions (arthritis, Parkinson's disease) that can impede physical activity [6, 7]. The risk associated with these disorders represents a significant challenge to medical practitioners. Therefore, as mentioned earlier, quantitative activity monitoring can help to discover and investigate physical movement [8, 9]. The essential information about the focus and motivation of our work is summarized as follows:

(i) The purpose of studying and monitoring GB in activity recognition is, for example, to detect movement patterns and early signs of physical impairment experienced by an individual. Furthermore, this process could allow clinician(s) to assess movement progress made by people (particularly those with physical impairment) when testing for the solution's efficacy [10, 11].

(ii) The study of GB could help explore and investigate the effect of free-living environments on people with gait disorder [12]. In addition, studying GB could allow medical practitioners to identify this walking disorder from the onset

(iii) Studying GB could enable the tracking of disease progression and test the effectiveness of preventive measures and solutions which can help improve movement irregularities

(iv) The study of GB could also enable clinician(s) to monitor and assist people with gait abnormalities as they may demonstrate unusual movements, which include symmetry of upper and lower limb swinging and other normal joint kinematics [13–15].

(v) By studying and monitoring the gait pattern of people with gait disabilities, proper movement modifications can be advised to enhance their moving style, and long-term well-being [16].

(vi) The study of GB using human activity data obtained from smartphones (see subsequent paragraphs) can offer low-power, low-cost, continuous remote screening tools for pathology identification that would enable adjustment to individual needs for

reducing the burden on clinicians and medical practitioners

The principal objective of this work is to devise ways by which we can detect GB through the use of wireless devices (such as smartphones) to monitor people with movement abnormalities to enable them to live a healthy lifestyle as they perform daily activities. Furthermore, wearable technology such as accelerometers and gyroscope sensors can analyze GB in real-life situations during human movement [12, 17]. Accelerometer sensors estimate the displacement of a mass using a location measuring circuit. That estimation is then transformed into a digital electrical signal for data processing via an analog to digital converter (ADC). On the other hand, a gyroscope sensor measures and maintains orientation and angular velocity. HAR monitoring using these sensors can characterize human movement (e.g., walking) given a set of observations. This process can be achieved by monitoring and analyzing walking information acquired from various sources such as the environment and sensors [18, 19].

Sensors such as accelerometers and gyroscopes are encapsulated in smartphone devices and can be attached to designated segments of the body such as the wrists, waist, chest, and thighs [20, 21]. The study of HAR, through accelerometers and gyroscope sensors, produces crucial information about individual daily movement and lifestyle [22]. Smartphone usage increases tracking GB due to high device processing power with communication handling and sensors such as accelerometers and gyroscope [23]. As previously discussed, accelerometers and gyroscopes in smartphones can reflect the duration of human movement through the measure of velocity and displacement [24]. Smartphones are portable and require no complex architecture to use them. This characteristic enables a smartphone to acquire GB information for HAR [23, 24]. Smartphones (with built-in inertial sensors such as gyroscope and accelerometer) are prevalent nowadays as it makes information and communication services available for individuals as they perform their daily activities. However, sensors embedded in smartphones have drawbacks as a result of random zero bias, and oscillation noise which affect the reading outcome [25].

This paper proposes a heuristic thresholding method called the multivariate triple exponential weighted moving average of the martingale sequence (MTMS) based on previous work [26]. This method can identify GB(s) and estimate their duration using a smartphone attached to the wrist in unsupervised real-world situations. The novelty of this study is as follows:

(i) Our method combines the martingale theory with the triple exponential moving average (TEMA) in a novel way to find a solution to the problem of GB detection for HAR time series

(ii) Unlike many GB detection algorithms that need a sliding window [27, 28], where the window length could be chosen based on the accuracy of the methods, the MTMS algorithm does not require

sliding windows. Therefore, the approach does not need to observe the model accuracy

(iii) Our approach uses optimized parameters to obtain the best performance for detecting GB in HAR sequences

To improve the accuracy and precision of MTMS method, we utilize the concept of optimization. The optimization process of tuning the input parameters is to discover the most efficient parameter value that improves the algorithm's performance [29–33]. An example of such optimization techniques is the particle swarm optimization (PSO). PSO [34] (discussed further in Section 3.3) is applied to optimize the parameters of our MTMS method for improved performance. Optimization methods such as genetic algorithm (GA) [35] can be computationally expensive and demands many iterations. However, in this case, PSO is the best alternative as it can handle complex problems and requires a small number of parameters with a correspondingly lower number of iterations [36, 37]. For our proposed approach, we implement PSO to maximize the G-mean. The reason why we use G-mean instead of F1 as our preferred metric is that it can be used to compute the stability between the classification achievement on the positive and negative classes [38]. G-mean computation can prevent overfitting the negative class and underfitting the positive class [39] when compared to F1. To also handle the challenges of distorted GB points, we implement precedence rules [40]. These precedence rules determine the grouping of bouts in the HAR data set and deal with the challenges of alignment and distortion of bout points. The optimization process and the implementation of the precedence rule are further discussed in Section 3. We benchmark our proposed technique against traditional methods (discussed in Section 3.4) for performance comparison purposes.

The paper structure is as follows: In Section 2, we review the latest work done on GB detection and discuss the stages of the algorithm; in Section 3, we introduce our proposed approach; in Section 4, we show our experimental results and compare them with the existing methods; in Section 5, we link and discuss the result obtained in Section 4 to the healthcare challenges stipulated in Section 1. We summarize and conclude the paper in Section 6, providing insight into the outcomes and the next steps of the research.

2. Related Work

In the last few decades, many gait bout detection techniques have been proposed for human activity recognition; as a result, recent research is now focused on measuring gait speed in clinical areas. For instance, the study of biometric gait pattern classification using an extreme learning machine (ELM) approach can be implemented to detect early gait abnormalities such as brain or neurological disorders in individuals. These abnormalities cannot be discovered using visual observation alone but also by implementing a robust quantitative analysis of the movement of an individual. This procedure can assist in comprehending the neuro-muscular

mechanics related to brain disorders, and this motivated Patil et al. [41] to evaluate the performance of multi-class gait classification using several machine learning methods, namely, KNN, SVM, ELM, and MLP. Experimental results showed that the ELM gave good results (93.54% overall classifications) when used to analyze the neuromuscular mechanics of patients with multiple sclerosis and stroke. However, the model is a classification based supervised learning that can be complex for GB detection.

The acquired push recovery capability for anyone is centred on learning, and the learning mechanism is not defined to us. Different models based on conventional mechanics and controls have been created to explore this mechanism. Nevertheless, these models have limitations. Semwal et al. [42] believe that an efficient computational model centred on learning will be effective in addressing these restrictions and proposed a model that collates humanoid push recovery data by executing the concept of utilizing accelerometer sensors in the smartphone. Experimentation is performed using the proposed HMCD and HLPRDCD solutions to demonstrate knee, hip, and ankle joint angle changes to analyze smartphone data. Results show that smartphone data collection is more accurate than a potentiometer based on HMCD. Also, executing LVQ shows that push recovery capability depends on age, height, weight, sex, race, ambidextrous, etc. The limitation of the model is that it does not implement optimized computational based on hybrid automata that can coordinate biped robot push recovery the same way as in humans.

Gait study is vital to discover a person from afar. However, the major issues associated with human gait-based discovery are high variability, movement obstruction, pose and speed variance, and regular gait cycle detection. These challenges motivate Semwal et al. [43] to develop an algorithm that explores the CASIA A, B, and C data sets to identify view, clothing, and speed invariant human detection. The study comprises a robust approach that utilizes computer vision for human identification. The suggested technique consists of feature extraction procedures. Feature extraction approaches consist of gait energy image (GEI) for cloth invariance, histogram of gradients (HOG) for multiview invariance, and Zernike moment with random transform for cross-view invariance. The following methods, namely, SVM, ANN, and XGBoost-based approaches, are implemented on the CASIA data set. The algorithms achieved 99%, 96%, and 67% detection accuracy independently for the three scenarios of invariance performed such as speed, cloth, and pose. However, the model utilizes some supervised learning methods that can be complex to develop for real-time GB detection and require a lot of computational time.

Most approaches for analyzing HAR involve a robust feature extraction application and time series preprocessing. However, this can entail a lot of human effort that is time-consuming and application-specific. This situation inspired Dua et al. [44] to propose a deep neural network-based model that implemented a convolutional neural network and gated recurrent unit as an end-to-end model. This model executes automatic feature extraction

and categorizing of the HAR scenarios. Experiments were carried out utilizing the wearable sensor's raw data with nominal pre-processing and non-handcrafted feature extraction techniques. The accuracies acquired from the analysis of the UCI-HAR, WISDM, and PAMAP2 data sets were 96.20%, 97.21%, and 95.27% independently. The results show that methods outperform some other related approaches. However, the model's technique implements a supervised learning approach that can be complicated to develop for real-time GB detection and requires a lot of computational time.

Atrsaei et al. [40] suggested an approach based on a single sensor setup. This technique can develop and validate clinical and home areas by attaching a sensor to the lower back. The algorithm uses a method known as naive Bayes classifier (a family of probabilistic classifiers that is centred on applying Bayes' theorem with robust (naive) independence assumptions between features) that can be utilized to discover gait bouts in activities around the home environment. The approach is validated using the data accumulated from 35 patients with movement disabilities such as multiple sclerosis. The method achieved a precision of 0.15m/s for evaluating gait speed with a bias rate of zero. However, the performance is still limited to noise interference from accumulated data obtained via wireless sensors. The technique is capable of assessing unsupervised mobility is reflected in its accuracy of 96.4% and F1 score of 78.6%.

Gait bout study motivated Barrett et al. [45], who suggested a way to improve the evaluation of gait bouts in Fitbit devices. This is achieved by evaluating gait bouts by developing the modelled physical activity level from the Fitbit Flex. The paper's primary purpose is to contrast the "gold standard" ActiGraph to that of the modelled Fitbit Freedson approach and ascertain regular values of expected errors in gait bout identification between two devices and methods. The techniques are proxy methods (the proxy method is a structural design pattern that makes it possible to allow for the replacement of another object) to estimate the actual physical activity levels. The approach uses three techniques to compare bout detection. These techniques are the ActiGraph Freedson method, Fitbit Intensity Score, and the modelled Fitbit Freedson utilizing three varying outcomes. Firstly, the author compares the duration of gait bout achieved from the AntiGraph GT3X technique to that of the baseline of each subject performing physical activity in a day. Secondly, the same procedure determines the intensity score achieved by the Fitbit and the modelled Fitbit Freedson approaches. Lastly, the author contrasted the gait bout discovered from the three methods to a labelled gait bout recorded in a self-report diary for performance evaluation. This process is still restricted by noise, impacting the algorithm's overall performance.

GB detection can also be performed using a wrist-mounted sensor measurement unit. However, the extensive freedom of wrist motion during daily life situations is a severe obstacle to a robust and precise GB analysis. These challenges motivated Soltani et al. [46] to suggest an approach for identifying GB using a wrist-mounted accelerometer. The method uses a Bayes estimator (Bayes estimator is a decision rule that minimizes the posterior expected value of a loss function), least absolute shrinkage, and selection operator (LASSO) (which identifies the optimal possible features to maximize performance on the training data set). The LASSO can select characteristics that include all biomechanical benchmarks (intensity, posture, periodicity, and non-gait dynamicity). The Bayes estimator and LASSO techniques are followed by two physically significant post-categorization steps to handle the problems of movement challenges of the wrist in real-life situations. The proposed approach has been validated against two data sets consisting of healthy young and older people, respectively. The algorithm achieved a satisfactory interquartile range within 80.0% to 98.4% for accuracy, sensitivity, precision, and F1 score in the identification of GB. The algorithm also produces a high correlation of 0.95 between the proposed and reference methods for the total duration of GB discovered. This correlation can be further improved by isolating noise in the HAR sequence.

Ramakrishnan et al. [47] proposed a method, namely, gait asymmetry metric (CGAM), that synthesizes human gait motion. CGAM is weighted by normalizing the data to stabilize each parameter's effect and combines spatial, kinematic, and temporal disproportional parameters. The approach can enhance the performance of gait patterns by assisting mechanized recovery approaches. CGAM also computes the quantifiable thresholds to produce efficient, comprehensive equivalent gait asymmetry. The study for the techniques combines clinical estimation, such as a six-minute walk test (6MWT), timed up and go (TUG), and gait velocity. These combinations are implemented on gait data obtained on individuals with movement disabilities such as stroke before and after recovery. Experimental results show that the CGAM can produce a higher correlation in estimating GB. However, this correlation is still restricted to noise interference in the HAR data set.

The concept of human activity recognition is becoming relevant in the healthcare domain for observing and monitoring disabilities associated with movement. However, some HAR analysis models can identify changes in the HAR data stream but cannot measure the change intensity and duration. This situation motivated Etumusei et al. [48] to suggest an unsupervised learning technique, namely, multivariate exponential weighted moving average of the martingale sequence using genetic algorithm (MEWMS(GA)). The method can discover change duration and intensity by applying the martingale framework in HAR data sets. The model also uses optimization techniques to obtain the optimal parameter value in the weighted average by executing a genetic algorithm. Experimental outcomes show that the proposed approach improves over current martingale techniques. Furthermore, the method does not focus on estimating gait bout(s).

Ho and Wechsler [26] recommended a martingale framework based on the testing for exchangeability property of data values to discover changes in time series. The method comprises a clustering approach and the implementation of a metric called strangeness, which shows how data points differ. Consequently, the strangeness was used to compute

the *P*-value and the randomized power martingale (RPM). This method can detect a transition in a data sequence from the experiment performed. However, there are restrictions to the algorithm's performance. For example, the technique can capture some false positives due to noise. Also, the method can analyze unsupervised multivariate data sequences.

Consequently, in some of our previous work [48–50], we proposed several robust thresholding methods that can analyze time series to detect changes. These methods can be categorized into two stages. Firstly, the methods use smoothing techniques such as a moving median or a Gaussian moving average to isolate and minimize the impact of noise in both univariate and multivariate time series. Secondly, the approaches use optimization methods such as GA and PSO (see Section 1). However, the technique was not designed for GB(s) detection.

2.1. Background of Research. As we can deduce from previous paragraphs, current algorithms can discover GB in HAR data sets acquired from wireless sensor technologies. However, the performance of these approaches is still limited by noise interference which can influence result rendering. Also, some of these approaches are not optimized for maximum performance. Furthermore, some of these methods use a supervised learning approach to detect changes in a HAR data set. However, the supervised learning approach has the following restrictions: it can be complex and expensive to implement, takes a lot of computation time, and cannot be used in real-time. Moreover, it might be challenging to use with dynamic and growing data [51].

To address the limitations stated previously, we present an unsupervised learning model [52] known as a multivariate triple exponential weighted moving average of the martingale sequence (MTMS) method. Further elaboration of the research novelty outlined in Section 1 includes:

(i) The MTMS model is straightforward and not complicated. It does not take much computation time

(ii) Our model can be implemented on dynamic and growing data without restriction

(iii) The MTMS approach can identify GB(s) and estimate their duration by analyzing HAR data obtained using a smartphone attached to the wrist in unsupervised real-world situations

As previously explained in Section 1, the MTMS approach uses the martingale framework and the triple exponential moving average (TEMA) as a smoothing factor [53, 54]. The triple exponential moving average can smooth time series fluctuations, making it easier to discover patterns without the lag related to traditional moving averages (MA) [55, 56]. Informally, it does this by obtaining several exponential moving averages (EMA) of the actual EMA and subtracting some of the lag. TEMA can help to identify trend order and short-term signal alteration. TEMA is more suitable compared to the EMA for this type of analysis because it reacts more adequately to trend variation in time series [54, 57]. In addition, we can apply an optimization algorithm (PSO), which is further explained in Section 3.3. PSO gives the optimal parameter value for enhancing the performance of the proposed algorithm.

2.2. Applicability of the Suggested Approach. The proposed approach can be translated into a real-time healthcare system that can monitor an individual's GB(s) during physical activities. This system can be in the form of a mobile app that can process and analyze accelerometer data set for GB discovery. Our proposed system or application can log a subject's movement throughout the day. This application can also help observe patients with movement disabilities (especially the elderly) or older adults living alone. Furthermore, clinicians can use the information from this health application to provide adaptive healthcare services to users. The relevance of this smartphone system is outlined below:

(i) It is cost-effective as it does not require additional hardware or sensors

(ii) It is efficient in mobility as it can be placed in the pocket

(iii) It is non-obtrusive and does not make a subject feel awkward or uncomfortable

(iv) It does not require supervision as smartphones can be with an individual daily

(v) A smartphone is likely to be with the users during their daily activities. This situation makes our health application effective in monitoring the movement of an individual

Our proposed system involves five distinct stages: the pre-processing of the HAR data set obtained using a smartphone, the analysis of the HAR data sequence using our recommended algorithm, the recognition and the discovery of GB, and finally, the re-initialization of the system.

3. Human Activity Recognition Model

This section explains the proposed approach for analyzing the HAR data set. Our recommended technique aims to identify GB using an accelerometer embedded in a smartphone attached to the wrist in unsupervised real-world situations. In the following paragraphs, we shall discuss the HAR data, pre-processing approach, the martingale concept, threshold computation, logical preferences, particle swarm optimization technique, and applicability of our recommended model.

3.1. HAR Data Set. The HAR data is accessed publicly from the UCI machine learning repository and through the research work of Anguita et al. [1, 58]. The data is then adapted for only the stand to walk of stairs scenarios. The HAR sequence acquired from smartphone devices attached to different participants consisted of 30 volunteers ranging from 19 to 48 years old. Each volunteer performed the experiment protocol twice, and each scenario was at least executed twice on each test to imitate repetition (see

Table 1: Overview of activities in the HAR data set.

Label	Types	Time range
Stand(2)	Non-mobility	29s-70s
Walk upstairs(2)	GB	18s-28 s

Table 1). A timeout of between 29 and 70 seconds in which the subject remains still (stand) was arranged to differentiate between each walk upstairs scenario. This data set is obtained from an accelerometer embedded in a smartphone. The smartphone is mounted in a suitable part of the body (preferably the waist) of the participant to make it possible to monitor or record gait scenarios performed by individuals [1].

A smartphone is recommended for the experiment as there are many limitations to using mobile devices in an unconstrained environment. This limitation includes variation in mobile hardware architecture, low memory capacity, non-robust operating system, embedded sensor quality, and periodicity of acquired data. These constraints make the experiment HAR difficult. However, smartphones can overcome some of these challenges. Moreover, the signals received using on-body sensors such as smartphones are possibly beneficial over signals obtained by video cameras due to the reasons outlined:

(i) Smartphones can mitigate the limitations of environmental constraints and stationary settings that cameras often encounter [59]

(ii) The signal information obtained from on-body sensors such as smartphones are accurate, efficient, and effective [59]

(iii) Smartphones enjoy the advantages of information privacy in contrast to that acquired using video or camera [59]

The Samsung 19100 Gallaxy S II has 32GB of storage (microSDHC), 1GB of RAM, and a dual-core 1.2 Hz cortex - A9 hard drive. These features make it the ideal device for the experiment. The element of these movements comprises triaxial acceleration and action label using video recording. Each participant executes the scenario twice using a smartphone (Samsung 19100 Galaxy S II) attached to the waist (the most prominent smartphone placement for effective accelerometer orientation capture is either the human waist or thigh for the walking upstairs scenario [60]) of the volunteers.

Consequently, smartphones have inserted a triaxial accelerometer that can estimate the subject's acceleration. Figure 1 shows the axis orientation of the inertial sensor of the waist-mounted smartphones executed for the experiments and its casing. Furthermore, Figure 2 shows the orientation of the smartphone attached to the waist. The x-axis estimates the vertical movement, while the y-axis measures the horizontal movement in the lateral direction. Furthermore, the z-axis estimates the action in the posterior-anterior direction. The Acceleration signal was recorded at

a constant rate of 50Hz, which is moderately fast for obtaining human body movement information. [61]. The labelling procedure was achieved manually by choosing the videos recorded during the experiments as ground surveillance and contrasting them with the log files of the inertial signals [1].

Our goal is to identify GB (the continuity of the same movement patterns) in the experiment. As mentioned earlier, GB is the time interval between gait initiation and termination in this case. GB essentially requires repetition of stance-swing cycles [4]. Subsequently, if the foot is on the ground for more than a certain threshold period, it can be considered static, therefore non-GB. The participants perform several actions, such as the stand and walk-up-stairs dynamic scenario. The walking-up-stairs scenario is represented as GB or mobility, while standing is classified as non-mobility or non-GB. .

These actions are implemented several times in the data set [40]. Table 1 represents the scenario discussed.

The encapsulated accelerometer in the smartphone is used to capture triaxial linear acceleration and triaxial angular velocity at a consistent rate of 50Hz which is sufficiently fast for obtaining human body motion information [61]. The multivariate triaxial data accumulated from these activities have been labelled manually utilizing video recording. In addition, the acquired data set has been randomly divided into two sets, where 70% of the participants were chosen for creating the training data and 30% of participants for developing the test data. The randomization of the partition is to make sure that no samples were obtained from the same user in both subsets.

3.2. The Pre-Processing Approach. The sensor signals, acquired through embedded accelerometers in smartphones, are pre-processed by employing various filters [1]. Firstly, the median filter and a low Butterworth filter with a threshold frequency of 20Hz are included to reduce noise in the signal. The Butterworth filter is a signal processing filter devised to procure a response frequency that is as flat as possible in the passband [63]. The threshold frequency was chosen according to work presented in [61] which stipulates that the energy spectrum of the human movement lies within the range of 0Hz and 15Hz. From these procedures, a triaxial total acceleration A was produced. The clean signal is expressed as the total of the two acceleration vectors given as the gravitational component G and the body motion acceleration BA was separated by utilizing another low-pass filter (assuming that the gravitational element only influences the lowest frequencies). The experiment in separating the two signals shows that 0.3Hz was the maximal baseline frequency to attain a constant gravity G. This result was achieved by varying the baseline frequency from 0.1 to 1.0Hz using an increment of 0.025 Hz and measuring the minimum square error of the filtered gravity signal by subtracting the standard gravity constant ($9.81m/s^2$). Consequently, the acceleration time derivative (dA/dt^2), otherwise known as Jerk, was estimated. Consequently, the signal is sampled using a fixed-width sliding window of about 2.56 sec and an overlap of 50% equivalent to 128
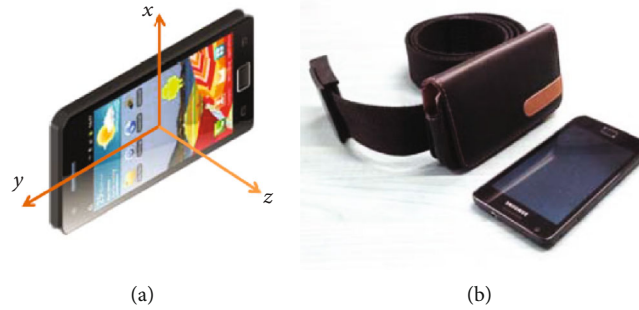
(a) (b)

FIGURE 1: Samsung Galaxy S2 Smartphone. (a) Arrows display axis orientation of the accelerometer. (b) The case and belt of the smartphone utilized for the experiments [1].
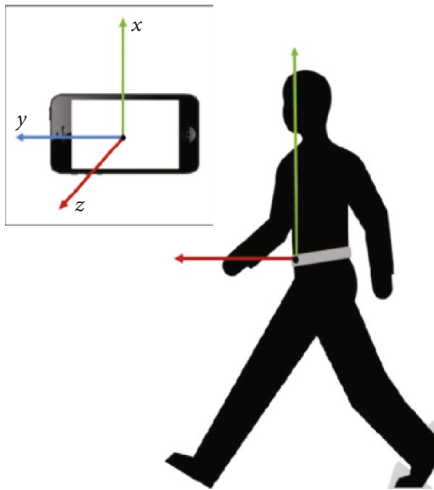

FIGURE 2: Smartphone placed on the waist to capture accelerometer orientation x, y, and z [62].

readings per window. This overlap is effective in executing other HAR methods, such as in work stipulated in [64, 65].

The fixed-width sliding window and overlap were chosen for the following reasons [1]:

   (i) The cadence range of a normal individual walking is between 90 and 130 steps per minute [66] which represents a minimum speed of 1.5 steps per second

   (ii) A minimum walking cycle consists of two practical steps per window sample

   (iii) A minimum speed of 50% is chosen as the average human cadence so that individuals with slower cadence due to disability or age are not excluded from the experiment

   (iv) Frequency domain signals demand the fast Fourier transform (FFT) [67] that maximize the power of two vectors (2.56 sec Ã—50Hz =128cycles)

Furthermore, for each window, a vector of features (accelerometer 3-axial raw signals (tAcc-XYZ) was acquired by computing variables using the time and frequency domain. This feature extraction, namely, signal magnitude area (SMA), mean, standard deviation (STD), entropy,

and signal-pair correlation (Corr), was implemented in previous work [1]. Subsequently, the FFT was utilized to locate the frequency elements for each window. Finally, a normalization technique is used to remove data redundancy [68, 69].

The labelled normalized vector of features standing and walking upstairs scenarios are illustrated in Figure 3 independently. Each plot in Figure 3 represents the accelerometer 3-axial signals X, Y, and Z, separately. The red lines in the plots separate the areas that show the participant walking up the stairs at different time intervals in a dynamic HAR scenario. The rest of the data represents some inactive activity when the participant is only standing. Finally, these patterns are analyzed using an unsupervised learning approach (see Section 3.3) for the discovery of GB. The HAR process is illustrated in Figure 4.

3.3. GB Detection Model. The pre-processed data is analyzed using the martingale framework, which originated from probability theory and was initially observed in gambling. Moreover, the idea has been utilized in diverse sectors such as in finance (for financial asset pricing) [70]. In addition, the martingale theory has been used in domains such as survival analysis, decision-making, and investment optimization [71]. In [49, 50], the research outcomes demonstrate that the martingale idea can be applied to the statistical analysis of time series such as terrestrial data and HAR. These outcomes show the martingale methodology's practicability and efficacy in discovering changes in the data-producing model for time series data streams. Ho and Wechsler [26] suggested a martingale approach, namely, randomized power martingale (RPM). This method can obtain better precision and recall than the traditional sequential probability ratio rate for change identification. We apply the martingale procedure through a heuristic thresholding method to analyze HAR data sets for GB detection. To improve the result's precision and recall, we use several techniques such as TES, logical preferences, and PSO optimization techniques. The advantage of the proposed algorithm is that it can obtain better precision or recall and does not use sliding windows like the RPM method in GB detection.

The profile of the data analysis involves the following steps:

   (i) Compute the martingale point
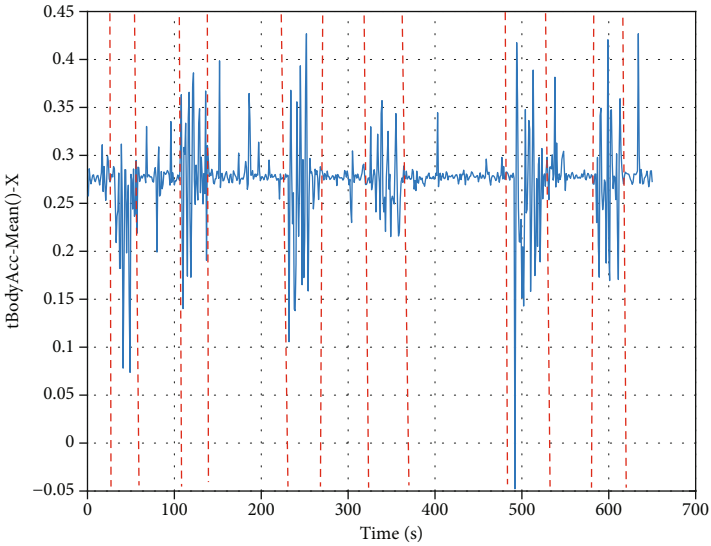
(a) X variable
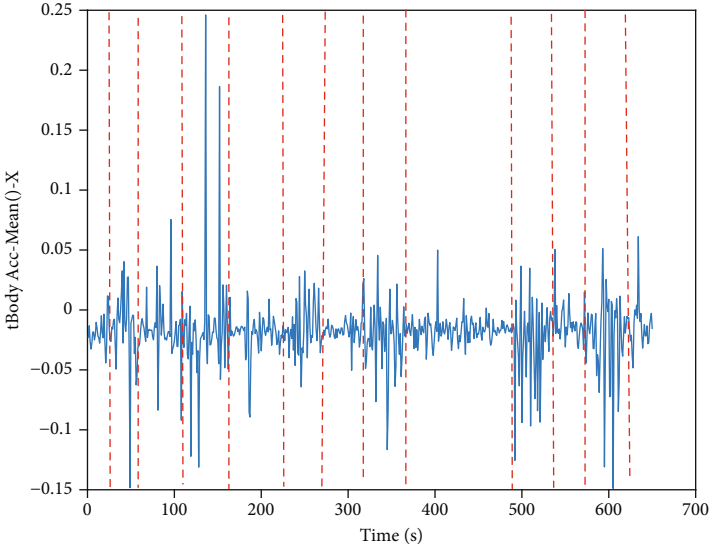
(b) Y variable

Figure 3: Continued.
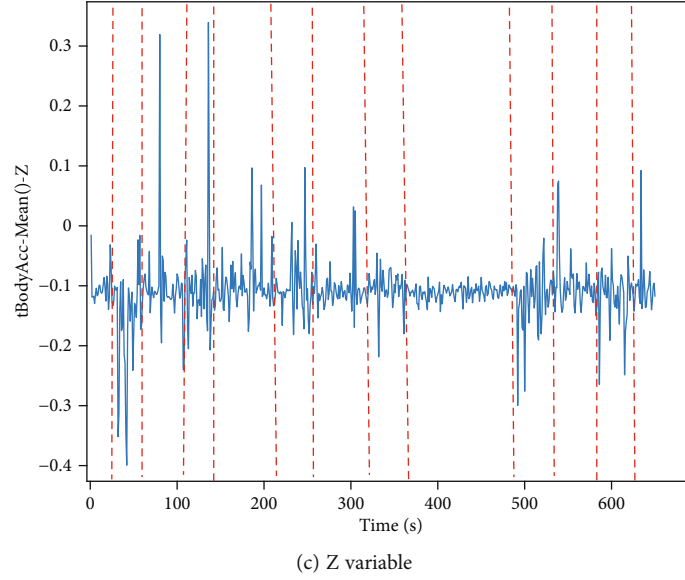
(c) Z variable

FIGURE 3: Accelerometer data.
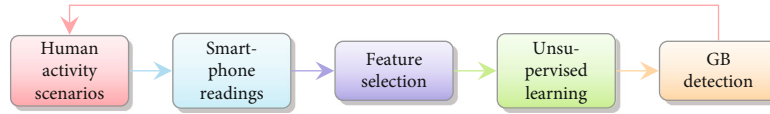

FIGURE 4: Activity recognition flowchart.

(ii) Compute the MTMS point (see Section 3.3)

(iii) Compute the threshold (see Section 3.3)

(iv) Determine the GB(s) that occurs in HAR

(v) Implement logical preferences to improve GB(s) discovery (see Section 3.3)

(vi) Optimize the algorithm's parameters to improve its accuracy and precision further (see Section 3.3). We shall further explain this procedure in the following section

3.3.1. Randomized Power Martingale. Ho and Wechsler [26] suggested an extension to the martingale approach by proposing a metric known as strangeness. Strangeness represents to what extent a new data point differs from the previous one in a time series.

Let us consider a series $Z = \{z_1, \cdots, z_{i-1}\}$, where there is a newly registered point z_i. Let us also examine the situation where the data points in Z have been clustered into k disjoint sets $Y_1, \cdots, Y_k, (k \leq i - 1)$ [72].

Definition 1. The strangeness of z_i is defined as

$$s_i = s(Z, z_i) = \|z_i - C_r\|, \qquad (1)$$

where C_r is the midpoint of the cluster Y_r, for some $r \in \{1, \cdots, k\}$ such that $z_i \in Y_r$. $\|.\|$ represents the selected distance.

The strangeness of z_i is used to compute a "probability" time series where its points are called $\widehat{p}_i$. If for $j = 1, 2, \cdots, i$, s_j is the strangeness of z_j and θ_i is a fixed value in $[0, 1]$ [26, 73], then $\widehat{p}_i$ is computed as follows:

$$\widehat{p}_i(Z \cup z_i, \theta_i) = \frac{\#\{j : s_j > s_i\} + \theta_i \#\{j : s_j = s_i\}}{i}, \qquad (2)$$

where "#" is a function that counts the number of j satisfying the following condition. For instance, $\#(j : s_j > s_i)$ is the number of j satisfying $s_j > s_i$, where s_j is the strangeness estimation stated in equation (1).

Intuitively, $\widehat{p}_i$ measures the probability of being more strange than z_i. Thus, the set of $\widehat{p}_i$ can be utilized to compute a new sequence known as the randomized power martingale.

Definition 2 (see [26]). The randomized power martingale (RPM) is enumerated by $\varepsilon \in [0, 1]$ determined at each time point as

$$M_n^{(\varepsilon)} = \prod_{i=1}^{n} \left(\varepsilon \widehat{p}_i^{\varepsilon-1} \right). \qquad (3)$$

Therefore, for a fixed $\varepsilon \in [0, 1]$, we can compute $M_i^{(\varepsilon)}$. This model will discover GB when

$$M_n^{(\varepsilon)} > t, \qquad (4)$$

where t is the threshold. The threshold will be explained in successive paragraphs.

Let us consider a data sequence $\{X_1, \cdots, X_s\}$, where each point X_i is a j vector, and j is the number of variables of the study. M_n will be computed for each variable at any time point. The next step involves reducing this new multidimensional sequence into a single metric. The M_n is computed for each variable at a given time point. The k-th mean of these values is determined using the equation:

$$N_k = \frac{M_n\left(X_1^{(k)}\right) + M_n\left(X_2^{(k)}\right) + \cdots + M_n\left(X_j^{(k)}\right)}{j}, \quad (5)$$

where $X_i^{(k)}$ is the i-th variable of X_k.

N_k, known as a multivariate randomized power martingale (MRPM), will be our new point for analysis in the multivariate HAR data set. The following section explains the computation of the threshold (t). In the next section, we introduce a method that improves the performance of the previously described martingale approach. The proposed method can estimate GB in HAR.

3.3.2. Multivariate Triple Exponential Weighted Moving Average of the Martingale Sequence (MTMS). In previous work, Alevisakos et al. [74] use time-varying and asymptotic control limits to examine and explore the mathematical properties of the triple exponential weighted moving average (TEWMA) or triple exponential smoothing (TES) chart. The TEWMA chart is more effective when detecting small shifts in the process mean compared to the double exponential weighted moving average (DEWMA) and EWMA charts, respectively. Some instances of TES implementation can be found in the work of Ongiri et al. [75], who use TEM to remove high-frequency noise in hydrological data to make a vital prediction on water demand. Also, Dev et al. [55] suggested a triple exponential smoothing based forecasting methodology that isolates high-frequency noise from time-varying data of estimated solar irradiance to make the time series pattern feasible to make a forecast of up to 20 minutes.

As discussed earlier, this previous work motivates us to propose a thresholding method that implements TES on the martingale framework to discover GB in HAR time series.

Triple exponential smoothing (TES) extends exponential smoothing to support time series patterns that replicate at every k element, where k can be any number other than one [76]. To obtain the new point of the sequence, we use TES to define the exponential weights over time [77]. Finally, TES is applied to signal analysis to remove high frequencies encountered in the signal [77]. Suppose we generate a sequence of observations N_k at a given point with size (l), where $k = 1, 2, \cdots, l$. EWMA is given as:

$$R_k = \alpha(N_k) + (1 - \alpha)R_{k-1}, R_0 = 0, \quad (6)$$

where R_k is the measure of the succeeding value of N_k. α is

TABLE 2: PSO parameter values.

Parameters	Value
InertiaRange	[0.10000, 1.1000]
InitialSwarmSpan	200
MaxIterations	200 ∗ NumberOfVariables
MaxStallIterations	20
MinNeighboursFraction	0.250
SwarmSize:	100
SelfAdjustmentWeight	1.4900
SocialAdjustmentWeights	1.4900

the smoothing factor (SF) in $[0, 1]$. The SF represents the weighting applied to the most recent observation. The higher the SF, the more weight is placed on the current observations, and the less weight is placed on previous observations.

Given a data sequence N_k with replicated pattern cycle variation (L). L is a positive integer that delegates the number of preceding samples. It also refers to the number of data points after which a new season begins. Consequently, TES [77–80] is given as:

$$H_{k+m} = (R_k + mb_k)c_{k-L+1+(m-1)\text{mod}L},$$
$$R_0 = N_0,$$
$$R_k = \alpha \frac{N_k}{c_{k-L}} + (1 - \alpha)(R_{k-1} + b_{k-1}), \quad (7)$$
$$b_k = \beta(R_k - R_{k-1}) + (1 - \beta)_{b_{k-1}},$$
$$c_k = \gamma \frac{N_k}{R_k} + (1 - \gamma)_{c_{k-L}},$$

where H_{k+m} is the measure of N_k at point $k + m(m > 0)$. Note that α is associated with the level of the series, β is associated with the trend, and γ is associated with the repeated pattern factors [77]. For our technique, we express the relation of the smoothing factors as $\beta = \alpha^2$ and $\gamma = \alpha^3$. For a given time, t, modL is the cycle of observation, and c_k represents the expected proportion of the trend that was predicted [55]. To initialize a set of reproduced pattern factors, a minimum of 2L periods is required, which expresses two entire seasons of historical data. $c_{k-L+1+(m-1)\text{mod}L}$ is the counterbalance into the index of the seasonal element against the last set from the observed time series point. Consequently, we apply the TES on the martingale point to obtain the MTMS points. The MTMS point is represented as follows:

$$MTMS\text{point} = \gamma \frac{M_k}{R_k} + (1 - \gamma)_{c_{k-L}}, \quad (8)$$

where M_k is the martingale sequence at a given point with size (l) and $k = 1, 2, \cdots, l$. R_k is the estimator of the succeeding martingale value (M_k). The following section discusses GB detection by presenting a thresholding method using MTMS points.

```
Data: Input (F): HAR multivariate data set
Result: Output: MTMS points
1:   Initialize: M(0) = 1; i = 1;
2:   Set values for cluster group k, ε value, Set the period of historical data (at least 2L periods) for training the TES model;
3:   while do
4:       A new example of z_i is discovered;
5:       if F = { null } then
6:           Set the strangeness of z_i := 0
7:       else
8:           Compute the strangeness of z_i and the data points in F
9:           Compute p̂_i of z_i;
10:          Compute M_i of the vectors using (3);
11:          Translate the dimension into single metrics Compute R_k points;
12:          Compute H_{k+M} points;
13:          Compute MTMS points
14:          Compute threshold t using MTMS points;
15:      end
16:      if MTMS ≥ t then
17:          GB discovered
18:      else
19:          Add z_i into F;
20:      end
21:      if i = i + 1 then;
22:      end
23:  end
```

ALGORITHM 1: The proposed algorithm.

3.3.3. Threshold Computation.

Ley et al. [81] suggested a technique that is defined as $\bar{x} \pm 3 * MAD$, where $\bar{x}$ denotes the mean of the data points and the MAD the median absolute deviation. We compute the threshold using $ME \pm MAD$, where ME denotes the median of our time series window. This threshold will enable the discovery of GB in the HAR data set. The condition for GB detection is

$$MTMS\text{point} > t. \qquad (9)$$

MTMS can determine GB in the data set. If MTMS is more significant than the threshold t, GB is detected, and if MTMS point is lesser than the threshold, GB has not taken place. The interval at which the MTMS point is greater than t is considered mobility (walking upstairs), while the contrary is referred to as non-mobility (standing). In the following section, we shall implement a technique known as logical precedent on the martingale sequence to improve the algorithm effectiveness in bout detection.

3.3.4. Logical Precedent.

To improve the precision of the GB discovery, we apply two basic rules [40] as follows:

(i) The non-mobility intervals that are less than 3 seconds and are between two mobility periods are translated into mobility

(ii) The mobility intervals of less than 3 seconds are translated into non-mobility periods

The neuro-psychological reasons for the two rules are that GB less than 3 seconds are presumed not to be exact GB. In contrast, a non-mobility interval of fewer than 3 seconds connecting two mobility intervals can be acknowledged as a brief, inconsequential recess [40]. Consequently, any gait bout of 3 seconds and above is considered a real walking bout [82]. This understanding will enable us to potentially isolate unwanted GB detection due to noise or unwanted readings captured by sensors in the HAR time series. The following section will explain the optimization technique used to improve the algorithm's parameters for enhanced performance.

3.3.5. Particle Swarm Optimization.

Particle swarm optimization (PSO) is a robust stochastic optimization method that uses fewer parameters to solve complex problems [83, 84]. PSO combines the features of both GA and evolution methods. This combination makes the PSO technique computationally inexpensive in memory consumption and speed. In addition, PSO can manage ongoing optimization problems, where each population is assigned an arbitrary velocity that pushes them through the solution hyperspace. The PSO optimization aim is to locate the optimum function that either maximizes the fitness function or minimizes the loss [36, 85]. PSO involves the following [85, 86]:

(i) For every iteration, measure the fitness value for each particle

(ii) If the fitness value of the particle is less than the Pbest (personal best location discovered by each particle), then the Pbest is updated with the particle's location as it is the current best achievable result at the point in time
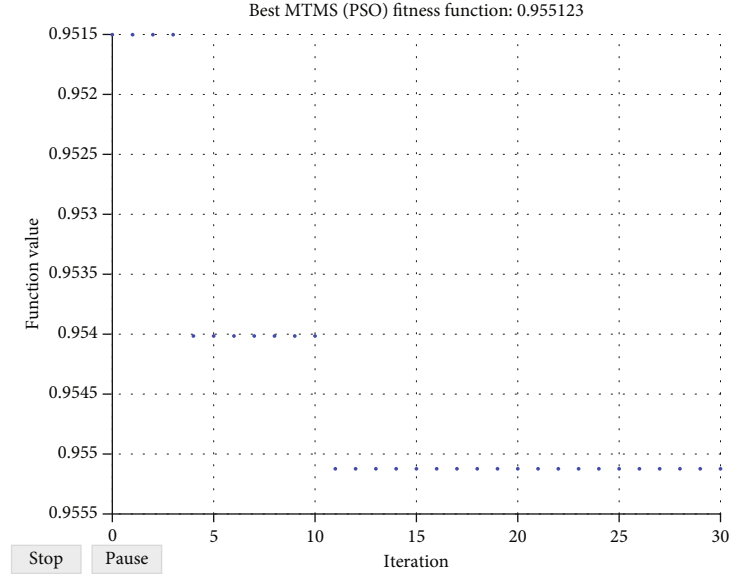
(a) MTMS
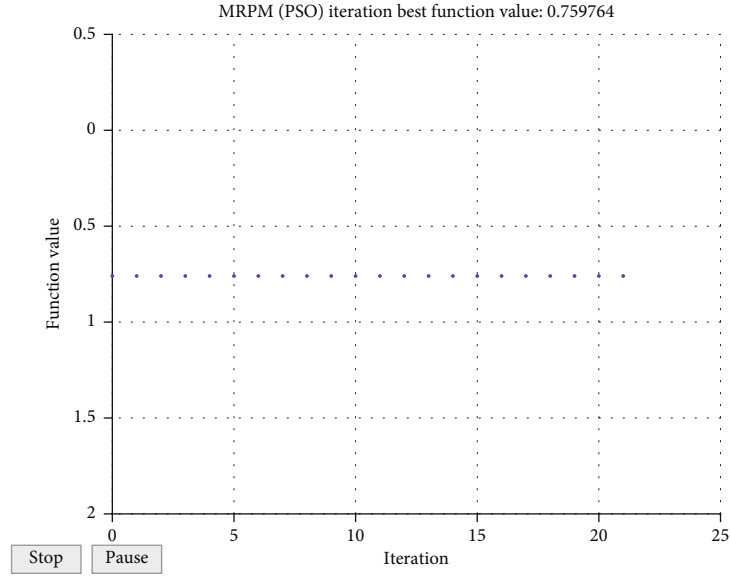
(b) MRPM

FIGURE 5: PSO iteration.

(iii) Identify the particle with the best fitness value from the overall particles as the Gbest (the best global position found by the swarm)

(iv) After discovering the Pbest and Gbest values, the particle's velocity and position are updated. This process is represented below:

$$ve_i = Wv_i + c_i r_i (P_{\text{best}.i} - x_i) + c_2 r_2 (G_{\text{best}} - x_i), \qquad (10)$$

where $x_i = x_i + v_i$, v_i = velocity of the ith particle (in our case, velocity is the mechanism utilized to shift the particle (mar-

tingale points) so as to perform a search for optimal solutions) [87], where $i = 1, 2 \cdots n$, x_i = position of the ith particle, $P_{\text{best}.i}$ = personal best of the ith particle, r_i and r_2 are random numbers, and c_i and c_2 are the acceleration coefficient that determines the amount of the influences on the particles velocity in the path of the global and local optima [88]. W is the inertia weight which ascertain the contribution rate of the particle's preceding velocity to that at the present time point [89]. In the next session, we shall implement the PSO using G-mean metrics.

We use the PSO approach to identify the maximal parameter using G-mean metrics. The PSO utilizes the fitness function (F) to acquire the optimal parameter value of

TABLE 3: Confusion matrix of MTMS (PSO) method on the test set.

	Non-GB	GB	Total (predicted)
Non-GB	428 (99.3%)	22 (10.0%)	450
GB	03 (0.7%)	197 (90.0%)	200
Total (reference)	431(100%)	219(100%)	650

a method. In this case, the fitness function is the maximum G-mean value computed with a specified range of parameters. The fitness function is given as:

It is represented as follows:

$$G-\text{mean}_{\max} = \max_{(\varepsilon,\alpha)}\left(G-\text{mean}_{(MTMS)}\right), \qquad (11)$$

where ε and α range from 0 to 1 for each activity [29]. As we have discussed earlier, PSO uses equation (11), which is the fitness function, to locate the maximum G-mean exploring within the parameter interval values [29]. PSO parameters are chosen to maximize the fitness function and can be seen in Table 2.

The new optimized sequence is known as MTMS(PSO) points. The MTMS method is illustrated in Algorithm 1, which shows the step-by-step implementation of the proposed approach.

3.4. Baseline Methods.
In this section, we will elaborate and discuss the previous heuristic thresholding methods (multivariate randomized power martingale and multivariate geometric moving average martingale) [50, 71]. These methods have been adapted for GB discovery. We selected these baseline methods to implement the martingale framework to detect transitions in the data set. These methods can be compared with our proposed MTMS techniques to evaluate the algorithm's performance within the martingale family.

3.4.1. Multivariate Randomized Power Martingale.
As discussed previously, we adapt the RPM method to discover GB in a multivariate HAR data set. The following step is taken to compute the MRPM(PSO) approach:

(i) Compute the RPM points for the vectors to give a new multivariate sequence

(ii) Translate the new computed multivariate sequence to a single metric

(iii) Compute the threshold of the series

(iv) Optimize the parameter of the algorithm using PSO

3.4.2. Multivariate Geometric Moving Average Martingale (MGM).
Apart from MRPM(PSO), we consider another technique that is based on GB detection called geometric moving average Martingale (GMAM) [71]. The GMAM method makes use of SF described in Section 3.3. Consequently, we adopt this method to analyze multivariate time series, namely, MGM. The MGM algorithm steps are shown below:

(i) Compute the GMAM points of the multivariate HAR data set

(ii) Translate the vectors into single metrics by finding the average of the GMAM vectors

(iii) Compute the threshold of the series

(iv) Optimize the parameters using the PSO technique

3.5. Evaluation Metrics.
This section explains the measure of the performance of our method using popular evaluation metrics [90]. These metrics are evaluated using a confusion matrix (CM) [91]. The CM depicts the predicted class of the activity (GB and non-GB). Consequently, the CM is used to determine the accuracy, precision, recall, harmonic mean ($F1$), and G-mean in terms of GB detection [40]. The accuracy, recall, precision, specificity, sensitivity, and G-mean matric can analyze the optimal selection for MTMS approaches. In this case, the confusion matrix (CM) can evaluate the performance of the algorithm [91].

Accuracy [92] is an intuitive performance metric defined as the ratio of GB(s) correctly detected in HAR to the total observations. The following gives accuracy:

$$\text{Accuracy} = \frac{TP + TN}{TP + FP + FN + TN}. \qquad (12)$$

We define true negatives (TN) as the false GB(s) correctly identified as false. True positives (TP) are the actual GB(s) that are correctly identified, while false positives (FP) are incorrect GB(s) identified as true. False negatives (FN) are the actual GB(s) identified as incorrect. Therefore, precision, recall (also known as sensitivity), F1 score, and specificity are computed as follows:

$$
\begin{aligned}
\text{Precision} &= \frac{TP}{TP + FP}, \\
\text{Recall(sensitivity)} &= \frac{TP}{TP + FN}, \\
\text{F1score} &= \frac{2 * \text{Recall} * \text{Precision}}{\text{Recall} + \text{Precision}}, \\
\text{Specificity} &= \frac{TN}{TN + FP}.
\end{aligned}
\qquad (13)
$$

G-mean [38, 91] is the measure of the correlation and the overall efficiency of the activities. G-mean combines the recall and the specificity. A low G-Mean denotes a poor performance in classifying positive cases irrespective of whether the negative possibilities are precisely classified. The G-mean metric is vital to prevent overfitting the false-negative and underfitting the false-positive class.

G-mean is defined as:

$$G-\text{mean} = \sqrt{\text{Recall} * \text{Specificity}}. \qquad (14)$$

A high G-mean signifies a preferable performance in the positive grading cases, while a low G-mean denote low performance [38]. These performance metrics properly evaluate
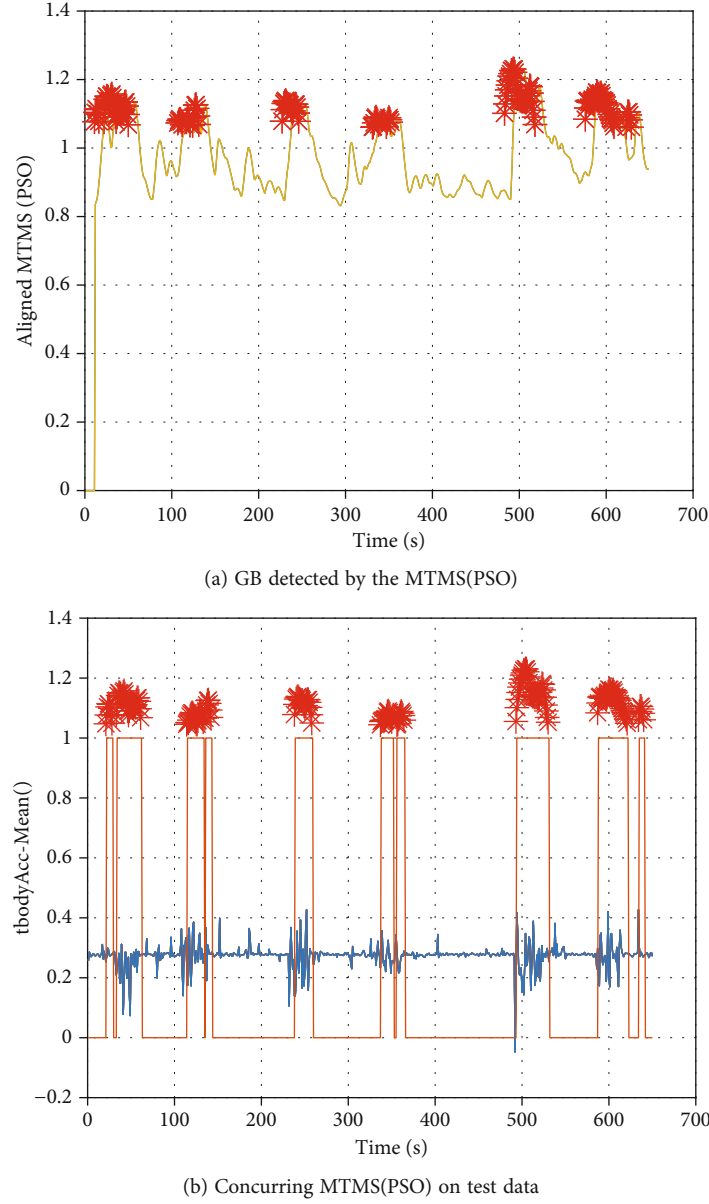
(a) GB detected by the MTMS(PSO)

(b) Concurring MTMS(PSO) on test data

FIGURE 6: MTMS(PSO) implementation.

TABLE 4: Confusion matrix of MGM(PSO) method on the test set.

	Non-GB	GB	Total (predicted)
Non-GB	425 (98.6%)	30 (13.7%)	455
GB	06 (1.4%)	189 (86.3%)	195
Total (reference)	431(100%)	219(100%)	650

the proposed approaches in locating GB(s), mainly on imbalanced HAR data sets. The evaluation performance for our methods is measured using evaluation metrics such as accuracy, precision, recall, harmonic mean (F1), and G-mean [32, 93].

3.6. Statistical Analysis of GB Detection. Our proposed method can estimate the GB(s) in the HAR data set. The

step for identifying the duration of each GB in the sequences are outlined below:

(i) We first detect the number of bouts using our suggested algorithm

(ii) We then estimate GB in HAR data set

The estimation of GB is obtained by computing the time (sec) of its occurrences. As explained earlier, GB is the length of time in seconds the GB takes place. GB is given as:

$$GB = \text{Seconds}(TP). \qquad (15)$$

We also use different statistical techniques (Spearman's rank correlation and mean square error) to estimate the total

GB(s) detected. Spearman's rank-order correlation (SROC) [94, 95] is the non-parametric version of the Pearson product-moment correlation. SROC can measure the capacity and order of relationship between two ranked variables. We can use SROC to estimate the correlation between the actual GB discovered and the predicted GB.

Given two sequences X_i and Y_i, we can compute db_i for every i by firstly converting the two sequences (X_i and Y_I) into rank (rX_i and rY_i). Consequently, we then compute the difference between the two ranks of the sequences. d_i is given as:

$$d_i = r(X_i) - r(Y_i). \tag{16}$$

Subsequently, SROC is given as:

$$\rho = 1 - \frac{6\sum d_i^2}{n(n^2 - 1)}, \tag{17}$$

where n is the number of observations and $i = 1, 2, \cdots, n$. ρ is the Spearman rank correlation coefficient. In our case, the actual GB(s) is represented as X_i while the predicted GB(s) is described as Y_i.

The mean square error (MSE) can estimate the total GB(s) identified in HAR sequence [96]. MSE is the sum of the squared difference between actual and estimated GB(s). MSE is given as:

$$MSE = \frac{\sum_{i=1}^n (X_i - Y_i)^2}{n}. \tag{18}$$

As explained earlier, the SROC value can show how two variables are correlated, while MSE can be referred to as a loss function used to estimate how efficient an algorithm is. The following section discusses the experimental result of our suggested method and previous martingale approaches.

4. Results and Discussion

This section discusses the proposed techniques used to identify GB in HAR. The section also describes the pre-processing methods and results obtained from experimentation using our proposed method and other baseline methods.

4.1. Cross-Validation Procedure. A cross-validation technique is implemented on two different time series called ds1 and ds2. Firstly, we implement our method (MTSM(PSO)) using the training data set (ds1) to obtain the best optimal parameter values that maximize G-mean. These parameter values are subsequently implemented to analyze the test data set (ds2). Both ds1 and ds2 consist of 404 and 650 data points. When we run the algorithms on the training set, our optimal parameter values that maximize G-mean are $\varepsilon = 0.7763$ and $\alpha = 0.9999$. The MTMS(PSO) optimization process for the train set is illustrated in Figure 5(a). We then use these optimal parameter values on a test set (ds2). The confusion matrix (CM) for

Table 5: Confusion matrix of MRPM(PSO) method on test set.

	Non-GB	GB	Total (predicted)
Non-GB	395 (91.6%)	108 (49.3%)	503
GB	36 (8.4%)	111 (50.7%)	147
Total (reference)	431(100%)	219(100%)	650

MTMS(PSO) performance on the test data set is illustrated in Table 3. Consequently, the exact process is repeated for both the MRPM(PSO) (see Figure 5(b)) and MGM(PSO) algorithm, first on the training set and then on the test data. MTMS(PSO) output is illustrated in Figure 6(a) while the corresponding MTMS(PSO) output on the test data set is shown in Figure 6(b). The CM for these approaches is presented in Tables 4 and 5 independently.

4.2. Confusion Matrix Analysis. To evaluate our algorithm performance, it is vital to note that the column of CM (Tables 3–5) shows the reference non-GB and GB, while the row gives the predicted non-GB and GB. Subsequently, we observe from Table 3 that 90% of GB (walk upstairs) are discovered by the MTMS(PSO) approach. Also, the algorithm can identify 99.3% of non-GB (standing). However, the algorithm only identifies 0.7% non-GB as GB and 10% GB as non-GB. Additionally, Table 4 illustrates that the MGM(PSO) technique is able to detect 86.3% of GB and 98.6% of non-locomotion, respectively. Nevertheless, the algorithm can discover 13.7% of non-GB that are GB and 1.4% GB that are non-locomotion independently. Lastly, Table 5 shows that MRPM(PSO) approach can detect 50.7% of GB and 91.6% of non-GB independently. Besides, the algorithm can discover 49.3% of non-GB that are actual movement and 8.4% of GB that are truly non-GB.

To summarize the analysis above, we can conclude that the MTMS(PSO) approach is more efficient as the percentage detection of GB is slightly higher (over 3%) compared to the MGM(PSO) and MRPM(PSO) separately. In addition, the false detection rate of the MTMS(PSO) method is lower than the baseline approaches.

The CM outlined above can be used to compute the evaluation metrics, which is summarized in Table 6 and shows the median and the interquartile range (interquartile range shows the range in values of lower quartile from the value of the upper quartile) of outcomes (accuracy, sensitivity, specificity, precision, F1, and G-mean) for both the training and test set.

4.3. Evaluation Metrics of GB Detection. Table 6 shows the evaluation metrics of our proposed and previous approaches, respectively. The accuracy, sensitivity, specificity, precision, F1 score, and G-mean for both the train and test data set are determined by finding the median of both results and the percentage value. The results in Table 6 are further simplified in Table 7. Table 7 shows the performance improvement (in percentage) of the proposed approach compared to MGM(PSO) and MRPM(PSO) methods,

TABLE 6: Overall performance of the proposed methods under different training and test conditions median (IQR) is reported.

	ds1	ds2	ε	SF	Accuracy (%)	Sensitivity (%)	Specificity (%)	Precision (%)	F1 score (%)	G-mean (%)
MRPM(PSO)	70	84	0.80	—	76.8 [75.6, 77.9]	76.19 [75.5, 76.7]	76.6 [75.3, 78.5]	49.20 [47.3, 50.7]	59.5 [58.5, 60.7]	76.5 [75.9, 77.0]
MGM(PSO)	70	84	0.33	0.13	92.9 [92.0, 94.5]	95.4 [94.7, 96.9]	91.9 [90.7, 93.4]	84.1 [82.7, 86.3]	89.5 [88.3, 91.3]	93.8 [92.7, 95.2]
MTMS(PSO)	70	84	0.31	0.10	95.4 [94.5, 96.2]	98.5 [98.5, 98.5]	93.2 [92.6, 95.1]	87.7 [86.1, 89.9]	92.7 [91.9, 94.0]	96.1 [95.5, 96.8]

TABLE 7: Performance improvement of the proposed MTMS(PSO) approach compared with the baseline methods.

Metrics	MGM(PSO)%	MRPM(PSO)%
Accuracy	2.9	18.6
Sensitivity	3.1	22.3
Specificity	1.3	16.6
Precision	3.6	38.5
F1 score	3.2	33.2
G-mean	2.3	19.6

TABLE 8: Comparison of novel methods.

Approaches	Training accuracy (ds1)	Test accuracy (ds2)	Overall accuracy
ELM	92.0	92.0	92.0

respectively. When we compare our suggested approach to MGM(PSO) method, we can observe that our method gives a better accuracy metric (over 2.0% over) compared to that of the MGM(PSO) approach. Subsequently, our approach offers better sensitivity (over 3.0%) and specificity (over 1.0 %) metrics compared to the MGM(PSO) technique. Also, our proposed approach gives a preferable precision (over 3.0%) compared to the MGM(PSO) method. Consequently, our methods produce an improved F1 (over 3.0%) and G-mean (over 2.0%) measure compared to the MGM(PSO) technique.

Consequently, we compare our suggested approach to the conventional MRPM(PSO) method and observe that our technique gives a better accuracy metric (over 18.0%). In addition, our process produces a preferable sensitivity (over 22.0%) and specificity (over 16.0%) metrics compared to the previous MRPM(PSO) technique. Also, our approach gives a better precision (over 38.0%) compared to that of MRPM(PSO) method. In addition, our methods produce an improved F1 (over 33.%) and G-mean (over 19.0%) measure compared to the traditional martingale technique.

Also, we compare our work to that of Patil et al. [41] who proposed a method (ELM) that evaluates the performance of multi-class gait classification. In Table 8 we present the classification accuracy of classifying GB(s).

Generally, the ELM technique achieved a classification accuracy of 92.0% in detecting the GB(s). However, our method achieved an overall 95.4% accuracy rate in detecting GB in HAR obtained from smartphones. The advantages of our method over the ELM technique are given as follows:

(i) Our model is a practical sequential GB detection approach to capture new HAR data points compared to the ELM approach

(ii) The ELM can only analyze labelled data sets for gait pattern detection. However, our suggested algorithm can analyze unlabelled HAR data sets for GB discovery without supervision

(iii) Our proposed method can adapt to changes in data patterns and

Overall, we can conclude that our method produces a superior outcome compared to the conventional ELM and MGM(PSO), respectively.

4.4. GB Detection Analysis. This section discusses the metric performance results for the evaluation of our proposed algorithm. The two metrics considered are the results of the Spearman test and mean square error (MSE). The Spearman test results between the estimated values of GB detected by our algorithm and that of the reference values were obtained as $R^2 = 0.9141$. Also, the Spearman test for the evaluated value of GB discovered by the MGM(PSO) technique and that of the reference values was acquired as $R^2 = 0.8758$. Consequently, the Spearman test for the GB identified by the MRPM(PSO) method to that of the reference or test value is established as $R^2 = 0.4783$. From the test, we can ascertain that the proposed MTMS(PSO) method gives a higher correlation than the baseline methods. The Spearman test is illustrated in Figures 7(a)–7(c). Furthermore, the MSE obtained between the evaluated values of GB discovered and that of the reference points for MTMS(PSO), MGM(PSO), and MRPM(PSO) is 0.3807, 0.8758, and 0.5072, respectively. The result shows that our proposed MTMS(PSO) is efficient (it produces a low error rate) compared to MGM(PSO) and MRPM(PSO), respectively. The low MSE rate is attributed to the fact that the suggested approach can diminish noise, leading to false positives effectively compared to the baseline methods. These results (SROC and MSE) are outlined in Table 9.

4.5. Computation Time. We also performed several runs (six) for the various algorithms (MRPM(PSO), MGM(PSO), and MTMS(PSO)) and presented the outcome in Table 10. The run iteration time for MRPM(PSO) is computationally efficient compared to MGM(PSO) and MTMS(PSO), respectively, as can be observed in Table 10. This evidence is attributed to the fact that MGM(PSO) and MTMS(PSO) are extensions of MRPM technique that adapt the original martingale method. The results show us that the computation time of executing MTMS(PSO) method is slightly higher (fraction of seconds) compared to the baseline methods (MRPM(PSO) and MGM(PSO)). However, we intend to further validate the proposed algorithm's computational time by experimenting with other HAR data sets.

From experimentation analysis, we can conclude that our proposed MTMS(PSO) method produces better performance in detecting GB compared to MRPM(PSO) and MGM(PSO) approaches. The following section provides further discussion of the results.

5. Discussion

This paper developed an effective algorithm to identify GB(s) and estimate their durations using a smartphone attached to the wrist in unsupervised real-world situations. The goal of this study is to provide a reliable, cost-effective, and efficient detection method that would be able to detect GB(s) in HAR. Two different accelerometer data sets (ds1
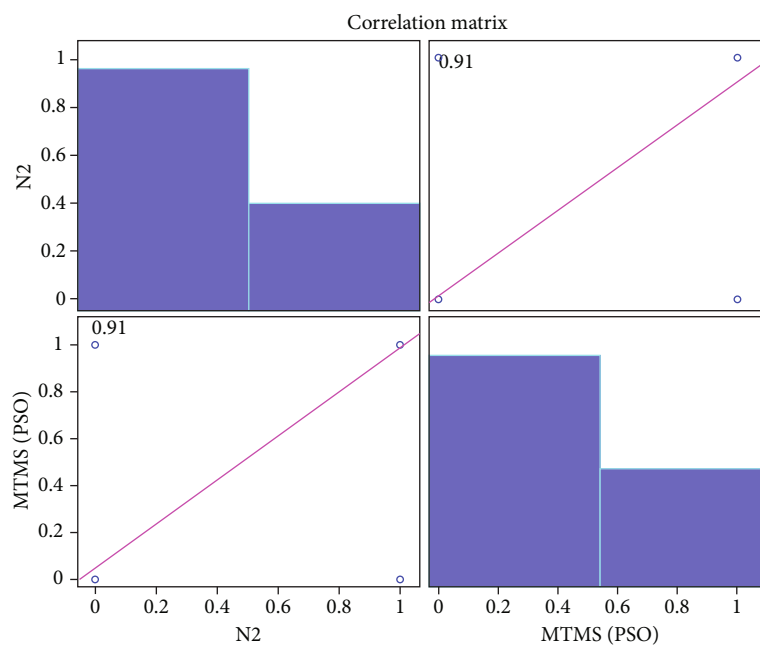
(a) MTMS(PSO)
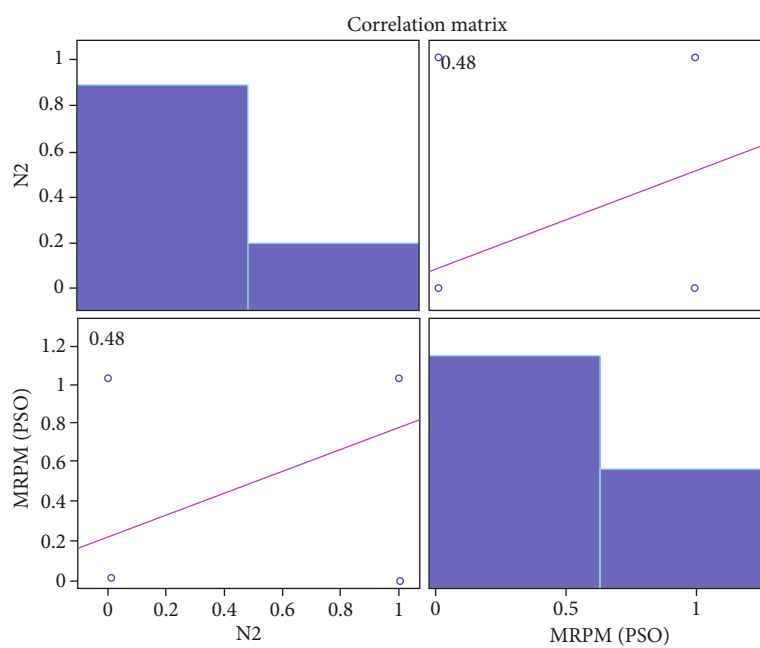
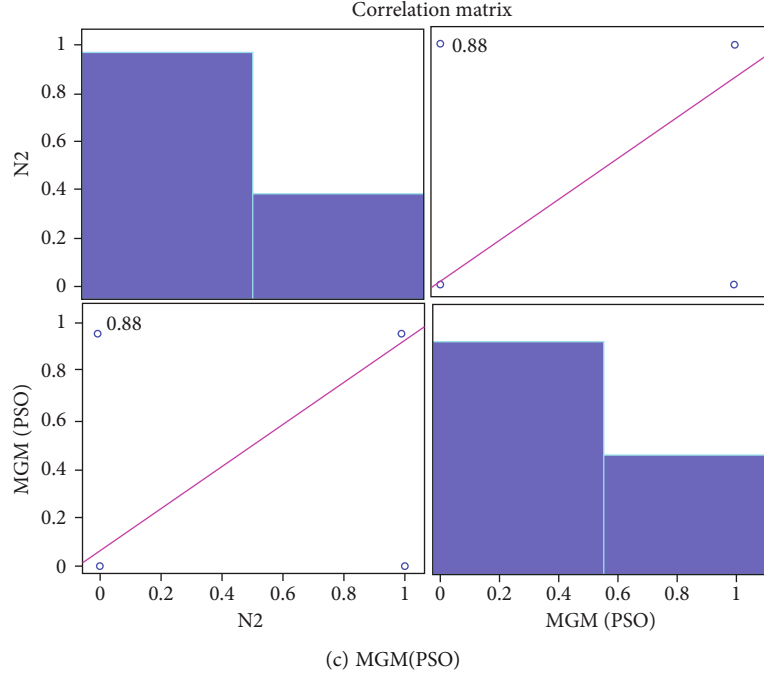(b) MRPM(PSO)

Figure 7: Continued.

Figure 7: Correlation plot.

Table 9: Percentage performance of the proposed MTMS(PSO) compared to MGM(PSO) and MRPM(PSO).

Method	Correlation (R^2)	MSE
MRPM(PSO)	0.4783	0.5072
MGM(PSO)	0.8758	0.3997
MTMS(PSO)	0.9141	0.3807

Table 10: Optimization run time.

Approach	Number of runs	Ave iteration time
MRPM(PSO)	6.0	0.104872 sec
MGM(PSO)	6.0	0.128132 sec
MTMS(PSO)	6.0	0.410360 sec

and ds2) were utilized to validate the effectiveness of our proposed algorithm. Our approach can detect most of the GB(s) in the data set. Our recommended algorithm achieves a high level of accuracy 95% overall and is extremely compelling from a diverse point of view. For instance, GB(s) detection was performed with a significant sensitivity of 98.5%. Our algorithm's discovery of GB(s) detection in HAR proved sufficient for walking categorization. There are very few miscategorizations experienced by our proposed approach between standing and walking upstairs. For instance, the misclassification of standing is 0.7% while the misclassification for walking is 10.0%. The precision, also called the positive predictive value of our proposed approach, is 87.7%, while the F1 and G-mean values for the execution of the proposed method are 92.7% and 96.1% independently (see Table 6). These results confirmed that our suggested approach

is computationally efficient and provides reliable details about GB(s) in HAR, suitable for a cost-effective and real-time computer-aided GB(s) discovery system. This system will enable clinicians to assess the movement of the elderly or persons with movement abnormalities. In Section 1, we presented some such systems and also briefly discussed how GB(s) study can help address these challenges. The following section shall elaborate on how this real-time feedback system can help improve gait impairment. Also, we shall link the results outlined above to these challenges.

5.1. Cost Effectiveness in Terms of Applicability in Discovering GB(s).

Our recommended healthcare systems can incorporate smartphone devices (for accumulating human recognition data), pre-processing procedure (for processing the collected time series), and the proposed algorithm application (for analyzing the time-varying data stream for GB detection). These innovative healthcare monitoring systems are noninvasive, low power, and low cost and will assist clinicians or caretakers in supplying long-term remote monitoring assistance, carrying out an early diagnosis, tracking abnormal movement symptoms over time, and categorizing or predicting pathological states [97].

5.2. Early Diagnosis of Gait Abnormalities.

The possibility of identifying early signs of movement abnormalities, especially for the elderly, could be accomplished using our proposed algorithm. This early diagnosis could be realized by monitoring and studying GB of older individuals using our recommended method. The following outlines could address or manage gait abnormalities in a person (elderly) in a real-time feedback system or application using our suggested algorithm.

(i) Healthcare researchers could track the movement of the elderly in real-time by monitoring and observing the daily activities of the elderly using our proposed health real-time system [2, 98, 99]

(ii) Our suggested system could estimate GB and non-GB activities achieved by an individual [100, 101]

This feedback from our proposed healthcare system could empower medical practitioners to measure the daily gait movement of the elderly [99, 102]. In addition, an unusual observation in the GB study outcome could prompt healthcare professionals to address the root cause of any concerns relating to gait disparity [103, 104]. These concerns might lead to further testing (on muscle strength, muscle tone, and coordination) and the medical solution's implementation to find an immediate solution to the gait issue [105, 106].

5.3. The Effect of a Free-Living Environment on People with Movement Disabilities. Our proposed healthcare solution could enable the practical monitoring of the elderly or persons with gait disorder as a result of neurological or nonneurological causes [104, 107, 108]. These conditions can lead to poor coordination, unsteadiness, and staggering gait for the older generation [109]. Our proposed healthcare system could assess the gait movement of these individuals in their homes and surroundings through sensor devices, as explained earlier. The recommended healthcare application can also identify gait disorders by implementing our proposed healthcare system. Such healthcare systems could discover and analyze the GB of the older person to assess their movement and coordination daily [104]. This solution could enable healthcare workers to adjust their treatment and exercise levels based on these findings.

5.4. Tracking of Parkinson's Disease Progression and Risk Prevention Measure. Gait can be described as a person's pattern of walking. Gait disorder associated with Parkinson's disease (PD) exhibits decrements in acceleration and stride, such as walking with a shortened sequence of two steps and low acceleration. Previously, the unified Parkinson's disease rating scale (UPDRS) motor score was used to study motor symptoms of PD sufferers [110]. However, the score is not related to gait and posture improvement or deterioration. Therefore, there is a need for analysis of gait activity performed by PD patients using our suggested system to assess the degree to which PD affects the movement and posture of the patient. In addition, this gait assessment could assist healthcare professionals to enhance current practices that may aid in symptoms monitoring, rehabilitation processes, therapy strategies, risk assessment, and avoidance. In the following section, we will discuss the conclusion and future work.

6. Conclusion and Future Work

This paper briefly discusses the RPM approach, which can discover abnormalities in time series. We propose a feasible and efficient method in terms of cost and applicability. The approach involves using a smartphone to capture the pattern of human activity scenarios such as walking upstairs and standing. The smartphone is attached to the waist of 30 subjects of the age range from 19 to 48 years old. The accelerometer embedded in this smartphone can pick up movement via its sensitivity to triaxial orientation. Subsequently, we implement an unsupervised learning technique known as the MTMS(PSO) in the experimental setup. The suggested method can reduce frequency noise by adding a smoothing procedure (TEMA) to the processed accelerometer data. The method further used logical precedence and optimization technique to improve the accuracy and precision. The method does not use a window size that can be implemented to measure the accuracy and precision rate of some traditional methods.

The proposed approach can identify GB(s) in HAR data. The proposed MTMS(PSO) gives a slightly higher G-mean value (over 3.0%) compared to the MRPM(PSO). Our suggested algorithm was also compared with the previous MGM(PSO) approach. Experimentation shows that our method produces a better result (over 1.0%) in accuracy, specificity, precision, F1, and G-mean independently. In addition, our method outperforms the ELM approach.

Our recommended methods can study and measure gait irregularities or patterns in a person with movement disabilities. Also, our proposed method can be used to test how efficient deterrent remedies can improve gait disorder, especially in the elderly. Additionally, our system's information on gait patterns analysis could ensure the implementation of effective clinical decisions to assist or monitor such pathological behaviour.

Further study will include the following:

(i) We will also need to validate the approach using a wide range of HAR data streams to characterize GB in the movement of people with disabilities, especially the elderly

(ii) To investigate the nature of GB associated with certain gait disabilities and how they can be distinguished in the HAR data set

(iii) To examine the correlation between the extent of transition and gait abnormalities using MTMS(PSO) value

(iv) To use our proposed healthcare system to assist people with gait abnormalities who demonstrate uncommon movements that involve the exchange of the upper and lower limb swinging in time and other joint dynamics

(v) To implement and validate the algorithm's performance on other time series from a different domain for anomaly or change detection

(vi) To also improve the algorithm, thereby making it computationally efficient

Data Availability

The HAR data is accessed publicly from the UCI machine learning repository through the research work of Anguita

et al. (Davide Anguita, Alessandro Ghio, Luca Oneto, Xavier Parra, and Jorge Luis Reyes-Ortiz. Energy-efficient smartphone-based activity recognition using fixed-point arithmetic. J. UCS,19(9):1295–1314, 2013.)

Disclosure

Conflicts of Interest

The authors declare no conflict of interest.

Acknowledgments

References

[1] D. Anguita, A. Ghio, L. Oneto, X. Parra, and J. L. Reyes-Ortiz, "Energy efficient smartphone-based activity recognition using fixed-point arithmetic," *Journal of Universal Computer Science*, vol. 19, no. 9, pp. 1295–1314, 2013.

[2] V. Vrutangkumar, "Effect of bout length on gait measures in people with and without Parkinson's disease during daily life," *Sensors*, vol. 20, no. 20, p. 5769, 2020.

[3] V. V. Shah, J. McNames, G. Harker et al., "Does gait bout definition influence the ability to discriminate gait quality between people with and without multiple sclerosis during daily life?," *Gait & Posture*, vol. 84, pp. 108–113, 2021.

[4] A. Kharb, V. Saini, Y. K. Jain, and S. Dhiman, "A review of gait cycle and its parameters," *IJCEM International Journal of Computational Engineering & Management*, vol. 13, pp. 78–83, 2011.

[5] L. Ruiz-Ruiz, A. R. Jimenez, G. Garcia-Villamil, and F. Seco, "Detecting fall risk and frailty in elders with inertial motion sensors: a survey of significant gait parameters," *Sensors*, vol. 21, no. 20, 2021.

[6] N. Byl, W. Zhang, S. Coo, and M. Tomizuka, "Clinical impact of gait training enhanced with visual kinematic biofeedback: patients with Parkinson's disease and patients stable post stroke," *Neuropsychologia*, vol. 79, no. Part B, pp. 332–343, 2015.

[7] B. K. Pedersen and B. Saltin, "Exercise as medicine – evidence for prescribing exercise as therapy in 26 different chronic diseases," *Scandinavian Journal of Medicine & Science in Sports*, vol. 25, pp. 1–72, 2015.

[8] M. G. Benedetti, A. Di Gioia, L. Conti et al., "Physical activity monitoring in obese people in the real life environment," *Journal of Neuroengineering and Rehabilitation*, vol. 6, no. 1, pp. 1–9, 2009.

[9] C. M. McDonald, L. M. Widman, D. D. Walsh, S. A. Walsh, and R. T. Abresch, "Use of step activity monitoring for continuous physical activity assessment in boys with duchenne muscular dystrophy," *Archives of Physical Medicine and Rehabilitation*, vol. 86, no. 4, pp. 802–808, 2005.

[10] F. Miller, *Physical Therapy of Cerebral Palsy*, Springer Science & Business Media, 2007.

[11] A. Lobo, *Physical activity and health in the elderly*, Bentham Science Publishers, 2010.

[12] S. Del Din, A. Godfrey, C. Mazza, S. Lord, and L. Rochester, "Free-living monitoring of Parkinson's disease: lessons from the field," *Movement Disorders*, vol. 31, no. 9, pp. 1293–1313, 2016.

[13] J. W. Noble and S. D. Prentice, "Adaptation to unilateral change in lower limb mechanical properties during human walking," *Experimental Brain Research*, vol. 169, no. 4, pp. 482–495, 2006.

[14] B. Roy, I. I. I. Davis, S. Ounpuu, D. Tyburski, and J. R. Gage, "A gait analysis data collection and reduction technique," *Human Movement Science*, vol. 10, no. 5, pp. 575–587, 1991.

[15] R. Baker, "Gait analysis methods in rehabilitation," *Journal of Neuroengineering and Rehabilitation*, vol. 3, no. 1, pp. 1–10, 2006.

[16] M. Chen, B. Huang, and X. Yangsheng, "Intelligent shoes for abnormal gait detection," in *2008 IEEE International Conference on Robotics and Automation*, pp. 2019–2024, Pasadena, CA, 2008.

[17] A. Godfrey, "Wearables for independent living in older adults: gait and falls," *Maturitas*, vol. 100, pp. 16–26, 2017.

[18] I. H. Lopez-Nava and A. Munoz-Melendez, "Human action recognition based on low- and high-level data from wearable inertial sensors," *International Journal of Distributed Sensor Networks*, vol. 15, no. 12, 2019.

[19] A. Farooq and C. S. Won, "A survey of human action recognition approaches that use an rgb-d sensor," *IEIE Transactions on Smart Processing and Computing*, vol. 4, no. 4, pp. 281–290, 2015.

[20] G.-D. Zhou and T.-H. Yi, "Recent developments on wireless sensor networks technology for bridge health monitoring," *Mathematical Problems in Engineering*, vol. 2013, 33 pages, 2013.

[21] Z. Wang, Z. Yang, and T. Dong, "A review of wearable technologies for elderly care that can accurately track indoor position, recognize physical activities and monitor vital signs in real time," *Sensors*, vol. 17, no. 2, p. 341, 2017.

[22] O. D. Lara and M. A. Labrador, "A survey on human activity recognition using wearable sensors," *IEEE Communications Surveys & Tutorials*, vol. 15, no. 3, pp. 1192–1209, 2012.

[23] M. M. Hassan, M. Z. Uddin, A. Mohamed, and A. Almogren, "A robust human activity recognition system using smartphone sensors and deep learning," *Future Generation Computer Systems*, vol. 81, pp. 307–313, 2018.

[24] K. Y. Chen, J. R. David, and R. Bassett, "The technology of accelerometry-based activity monitors: current and future," *Medicine & Science in Sports & Exercise*, vol. 37, no. 11, pp. S490–S500, 2005.

[25] J. Qian, J. Ma, R. Ying, P. Liu, and L. Pei, "An improved indoor localization method using smartphone inertial sensors," in *International Conference on Indoor Positioning and Indoor Navigation*, pp. 1–7, Montbeliard, France, 2013.

[26] S.-S. Ho and H. Wechsler, "A martingale framework for detecting changes in data streams by testing exchangeability," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 32, no. 12, pp. 2113–2127, 2010.

[27] M. H. M. Noor, Z. Salcic, I. Kevin, and K. Wang, "Adaptive sliding window segmentation for physical activity

recognition using a single tri-axial accelerometer," *Pervasive and Mobile Computing*, vol. 38, pp. 41–59, 2017.

[28] N. H. Ghassemi, J. Hannink, C. F. Martindale et al., "Segmentation of gait sequences in sensor-based movement analysis: a comparison of methods in parkinson's disease," *Sensors*, vol. 18, no. 1, 2018.

[29] N. Khan, S. McClean, S. Zhang, and C. Nugent, "Optimal parameter exploration for online change-point detection in activity monitoring using genetic algorithms," *Sensors*, vol. 16, no. 11, 2016.

[30] S. N. Sivanandam and S. N. Deepa, "Genetic algorithms," in *Introduction to Genetic Algorithms*, pp. 15–37, Springer, 2008.

[31] J. H. Holland, "Genetic algorithms," *Scientific American*, vol. 267, no. 1, pp. 66–72, 1992.

[32] M. Romero, J. M. Miranda, H. A. Montes, and J. C. Acosta, "A statistical measuring system for rainbow trout," in *Proceedings of the International Conference on Image Processing, Computer Vision, and Pattern Recognition (IPCV) (p. 1). The Steering Committee of The World Congress in Computer Science, Computer Engineering and Applied Computing (WorldComp)*, Las Vegas, Nevada, USA, 2014.

[33] X. Xuan, D. Lo, X. Xia, and Y. Tian, "Evaluating defect prediction approaches using a massive set of metrics: an empirical study," in *Proceedings of the 30th Annual ACM Symposium on Applied Computing*, pp. 1644–1647, Salamanca Spain, 2015.

[34] X.-S. Yang, *Nature-Inspired Optimization Algorithms*, Academic Press, 2020.

[35] S. Mirjalili and S. Mirjalili, "Genetic algorithm," in *Evolutionary Algorithms and Neural Networks*, pp. 43–55, Springer, 2019.

[36] S. Katiyar, "A comparative study of genetic algorithm and the particle swarm optimization," *International Journal of Technology*, vol. 2, no. 2, pp. 21–24, 2010.

[37] S. Shabir and R. Singla, "A comparative study of genetic algorithm and the particle swarm optimization," *International Journal of Electrical Engineering*, vol. 9, no. 2, pp. 215–223, 2016.

[38] J. Akosa, "Predictive accuracy: a misleading performance measure for highly imbalanced data," *Proceedings of the SAS Global Forum*, vol. 12, 2017.

[39] L. Yang, Y.-m. Cheung, and Y. Y. Tang, "Goboost: G-mean optimized boosting framework for class imbalance learning," in *2016 12th World Congress on Intelligent Control and Automation (WCICA)*, pp. 3149–3154, Guilin, China, 2016.

[40] A. Atrsaei, F. Dadashi, B. Mariani, R. Gonzenbach, and K. Aminian, "Toward a remote assessment of walking bout and speed: application in patients with multiple sclerosis," *IEEE Journal of Biomedical and Health Informatics*, vol. 25, no. 11, pp. 4217–4228, 2021.

[41] K. Prithvi Patil, S. Kumar, N. Gaud, and V. B. Semwal, "Clinical human gait classification: extreme learning machine approach," in *2019 1st international conference on advances in science, engineering and robotics technology (ICASERT)*, pp. 1–6, Dhaka, Bangladesh, 2019.

[42] V. B. Semwal and G. C. Nandi, "Toward developing a computational model for bipedal push recovery–a brief," *IEEE Sensors Journal*, vol. 15, no. 4, pp. 2021-2022, 2015.

[43] V. B. Semwal, A. Mazumdar, A. Jha, N. Gaud, and V. Bijalwan, "Speed, cloth and pose invariant gait recognition-based person

identification," in *Machine Learning: Theoretical Foundations and Practical Applications*, pp. 39–56, Springer, 2021.

[44] N. Dua, S. N. Singh, and V. B. Semwal, "Multi-input cnn-gru based human activity recognition using wearable sensors," *Computing*, vol. 103, no. 7, pp. 1461–1478, 2021.

[45] C. Barrett, G. Dominick, and K. N. Winfree, "Assessing bouts of activity using modeled clinically validated physical activity on commodity hardware," in *2017 IEEE EMBS International Conference on Biomedical & Health Informatics (BHI)*, pp. 269–272, Orlando, FL, USA, 2017.

[46] A. Soltani, A. Paraschiv-Ionescu, H. Dejnabadi, P. Marques-Vidal, and K. Aminian, "Real-world gait bout detection using a wrist sensor: an unsupervised real-life validation," *IEEE Access*, vol. 8, pp. 102883–102896, 2020.

[47] T. Ramakrishnan, S. H. Kim, and K. B. Reed, "Human gait analysis metric for gait retraining," *Applied Bionics and Biomechanics*, vol. 2019, 8 pages, 2019.

[48] J. Etumusei, J. M. Carracedo, and S. I. McClean, "Estimating change intensity and duration in human activity recognition using martingales," in *The 7th International Online & Onsite Conference on Machine Learning, Optimization, and Data Science*, Grasmere, United Kingdom, 2021.

[49] J. Etumusei, J. M. Carracedo, and S. McClean, "Novel martingale approaches for change point detection," in *20th International Conference on Intelligent Systems Design and Applications (ISDA 2020)*, pp. 2113–2127, On world wide web (online), 2020.

[50] J. Etumusei, J. M. Carracedo, and S. I. McClean, "A novel detection technique using martingales for optimal parameters in time series," *Journal of Information Assurance and Security (JIAS)*, vol. 16, no. 2, pp. 59–72, 2021.

[51] D. D. Lewis and J. Catlett, "Heterogeneous uncertainty sampling for supervised learning," in *Machine learning proceedings 1994*, pp. 148–156, Elsevier, 1994.

[52] K. P. Sinaga and M.-S. Yang, "Unsupervised k-means clustering algorithm," *Access*, vol. 8, pp. 80716–80727, 2020.

[53] E. S. Gardner Jr, "Exponential smoothing: the state of the art," *Journal of Forecasting*, vol. 4, no. 1, pp. 1–28, 1985.

[54] P. S. Kalekar, "Time series forecasting using holt-winters exponential smoothing," *Kanwal Rekhi School of Information Technology*, vol. 4329008, no. 13, pp. 1–13, 2004.

[55] S. Dev, T. AlSkaif, M. Hossari, R. Godina, A. Louwen, and W. van Sark, "Solar irradiance forecasting using triple exponential smoothing," in *2018 International Conference on Smart Energy Systems and Technologies (SEST)*, pp. 1–6, Seville, Spain, 2018.

[56] A. Raudys and Ž. Pabarškaitė, "Optimising the smoothness and accuracy of moving average for stock price data," *Technological and Economic Development of Economy*, vol. 24, no. 3, pp. 984–1003, 2018.

[57] P. Goodwin, "The holt-winters approach to exponential smoothing: 50 years old and going strong," *Foresight*, vol. 19, no. 19, pp. 30–33, 2010.

[58] D. Anguita, A. Ghio, L. Oneto, X. Parra, and J. L. Reyes-Ortiz, "A public domain dataset for human activity recognition using smartphones," in *Proceedings of the 21th International European Symposium on Artificial Neural Networks, Computational Intelligence and Machine Learning*, Bruges, 2013.

[59] J. Yang, M. N. Nguyen, P. P. San, X. L. Li, and S. Krishnaswamy, "Deep convolutional neural networks on multichannel time series for human activity recognition," in *Twenty-fourth*

international joint conference on artificial intelligence, Buenos Aires, Argentina, 2015.

[60] A. Prasad, A. K. Tyagi, M. M. Althobaiti, A. Almulihi, R. F. Mansour, and A. M. Mahmoud, "Human activity recognition using cell phone-based accelerometer and convolutional neural network," *Applied Sciences*, vol. 11, no. 24, 2021.

[61] D. M. Karantonis, M. R. Narayanan, M. Mathie, N. H. Lovell, and B. G. Celler, "Implementation of a real-time human movement classifier using a triaxial accelerometer for ambulatory monitoring," *IEEE Transactions on Information Technology in Biomedicine*, vol. 10, no. 1, pp. 156–167, 2006.

[62] M. Jauhiainen, J. Puustinen, S. Mehrang et al., "Identification of motor symptoms related to parkinson disease using motion-tracking sensors at home (KÄVELI): protocol for an observational case-control study," *JMIR Research Protocols*, vol. 8, no. 3, article e12808, 2019.

[63] G. Ellis, *Control System Design Guide: Using your Computer to Understand and Diagnose Feedback Controllers*, Butterworth-Heinemann, 2012.

[64] R. W. DeVaul and S. Dunn, "Real-time motion classification for wearable computing applications," 2001 Project Paper, 2001.

[65] K. Van Laerhoven and O. Cakmakci, "What shall we teach our pants?," in *Digest of Papers. Fourth International Symposium on Wearable Computers*, pp. 77–83, Atlanta, GA, USA, 2000.

[66] C. B. Abdelkader, R. Cutler, and L. Davis, "Stride and cadence as a biometric in automatic person identification and verification," in *Proceedings of Fifth IEEE international conference on automatic face gesture recognition*, pp. 372–377, Washington, DC, USA, 2002.

[67] G. W. Chang, C. I. Chen, Y. J. Liu, and M. C. Wu, "Measuring power system harmonics and interharmonics by an improved fast fourier transform-based algorithm," *IET generation, transmission & distribution*, vol. 2, no. 2, pp. 192–201, 2008.

[68] F. Douglis, P. Kulkarni, J. D. Lavoie, and J. M. Tracey, "Method and apparatus for data redundancy elimination at the block level," US Patent 8, 135, 683, 2012.

[69] S. Xie and Z. Chen, "Anomaly detection and redundancy elimination of big sensor data in internet of things," 2017, https://arxiv.org/abs/1703.03225.

[70] F. Delbaen and W. Schachermayer, "A general version of the fundamental theorem of asset pricing," *Mathematische Annalen*, vol. 300, no. 1, pp. 463–520, 1994.

[71] X. Z. Kong, Y. X. Bi, and D. H. Glass, "A geometric moving average martingale method for detecting changes in data streams," in *International Conference on Innovative Techniques and Applications of Artificial Intelligence*, pp. 79–92, Cambridge, United Kingdom, 2012.

[72] T. Kanungo, D. M. Mount, N. S. Netanyahu, C. D. Piatko, R. Silverman, and A. Y. Wu, "An efficient k-means clustering algorithm: analysis and implementation," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 24, no. 7, pp. 881–892, 2002.

[73] V. Fedorova, A. Gammerman, I. Nouretdinov, and V. Vovk, "Plug-in martingales for testing exchangeability on-line," 2012, https://arxiv.org/abs/1204.3251.

[74] V. Alevizakos, K. Chatterjee, and C. Koukouvinos, "The triple exponentially weighted moving average control chart," *Quality Technology & Quantitative Management*, vol. 18, no. 3, pp. 326–354, 2021.

[75] H. N. Ongiri and P. W. Warutumo, "Use of triple exponential smoothing in the analysis of hydrological data," *International Journal of Statistics and Applied Mathematics*, vol. 5, no. 2, 2020.

[76] E. C. Nwogu, I. S. Iwueze, and V. U. Nlebedim, "Some tests for seasonality in time series data," *Journal of Modern Applied Statistical Methods*, vol. 15, no. 2, pp. 382–399, 2016.

[77] A. Lasek, N. Cercone, and J. Saunders, "Smart restaurants: survey on customer demand and sales forecasting," in *Smart Cities and Homes*, pp. 361–386, 2016.

[78] P. G. Mulloy, "Smoothing data with less lag," *Technical Analysis of Stocks and Commodities*, vol. 12, no. 2, pp. 72–80, 1994.

[79] A. Hull, *How to reduce lag in a moving average*, ActVest Pty Ltd, 2005.

[80] T. Tillson, *Better moving averages*, Digital Filters, Monterrey, 1998.

[81] C. Leys, C. Ley, O. Klein, P. Bernard, and L. Licata, "Detecting outliers: do not use standard deviation around the mean, use absolute deviation around the median," *Journal of Experimental Social Psychology*, vol. 49, no. 4, pp. 764–766, 2013.

[82] C. M. el Achkar, C. Lenoble-Hoskovec, A. Paraschiv-Ionescu, K. Major, C. Büla, and K. Aminian, "Instrumented shoes for activity classification in the elderly," *Gait & Posture*, vol. 44, pp. 12–17, 2016.

[83] F. Cus, U. Zuperl, and V. Gecevska, "High speed end-milling optimisation using particle swarm intelligence," *Journal of Achievements in Materials and Manufacturing Engineering*, vol. 22, no. 2, pp. 75–78, 2007.

[84] D. Wang, D. Tan, and L. Liu, "Particle swarm optimization algorithm: an overview," *Soft Computing*, vol. 22, no. 2, pp. 387–408, 2018.

[85] J. Kennedy and R. Eberhart, "Particle swarm optimization," in *Proceedings of ICNN'95-international conference on neural networks*, Perth, WA, Australia, 1995.

[86] R. Poli, J. Kennedy, and T. Blackwell, "Particle swarm optimization," *Swarm intelligence*, vol. 1, no. 1, pp. 33–57, 2007.

[87] J. Barrera, O. Alvarez-Bajo, J. J. Flores, and C. A. Coello, "Limiting the velocity in the particle swarm optimization algorithm," *Computacion y Sistemas*, vol. 20, no. 4, pp. 635–645, 2016.

[88] H. Ling Chen, B. Yang, S. Jing Wang, G. Wang, H. Zhong Li, and W. Bin Liu, "Towards an optimal support vector machine classifier using a parallel particle swarm optimization strategy," *Applied Mathematics and Computation*, vol. 239, pp. 180–197, 2014.

[89] J. C. Bansal, P. K. Singh, M. Saraswat, A. Verma, S. S. Jadon, and A. Abraham, "Inertia weight strategies in particle swarm optimization," in *2011 Third world congress on nature and biologically inspired computing*, pp. 633–640, Salamanca, Spain, 2011.

[90] C. Goutte and E. Gaussier, "A probabilistic interpretation of precision, recall and f-score, with implication for evaluation," in *European Conference on Information Retrieval*, pp. 345–359, Meylan, France, 2005.

[91] G. H. Nguyen, A. Bouzerdoum, and S. L. Phung, "Learning pattern classification tasks with imbalanced data sets," *Pattern Recognition*, pp. 193–208, 2009.

[92] R. P. Espindola and N. F. F. Ebecken, "On extending f-measure and g-mean metrics to multi-class problems," *WIT Transactions on Information and Communication Technologies*, vol. 35, 2005.

[93] M. Bekkar, H. K. Djemaa, and T. A. Alitouche, "Evaluation measures for models assessment over imbalanced data sets," *Journal of Information Engineering and Applications*, vol. 3, no. 10, 2013.

[94] T. D. Gauthier, "Detecting trends using Spearman's rank correlation coefficient," *Environmental Forensics*, vol. 2, no. 4, pp. 359–362, 2001.

[95] R. A. Armstrong, "Should Pearson's correlation coefficient be avoided?," *Ophthalmic and Physiological Optics*, vol. 39, no. 5, pp. 316–327, 2019.

[96] D. M. Allen, "Mean square error of prediction as a criterion for selecting variables," *Technometrics*, vol. 13, no. 3, pp. 469–475, 1971.

[97] S. M. G. Beyrami and P. Ghaderyan, "A robust, cost-effective and non-invasive computer-aided method for diagnosis three types of neurodegenerative diseases with gait signal analysis," *Measurement*, vol. 156, article 107579, 2020.

[98] M. R. Afzal, M. K. Oh, C. H. Lee, Y. S. Park, and J. Yoon, "A portable gait asymmetry rehabilitation system for individuals with stroke using a vibrotactile feedback," *BioMed Research International*, vol. 2015, 16 pages, 2015.

[99] M. El-Gohary, S. Pearson, J. McNames et al., "Continuous monitoring of turning in patients with movement disability," *Sensors*, vol. 14, no. 1, pp. 356–369, 2014.

[100] I. Hillel, E. Gazit, A. Nieuwboer et al., "Is every-day walking in older adults more analogous to dual-task walking or to usual walking? Elucidating the gaps between gait performance in the lab and during 24/7 monitoring," *European Review of Aging and Physical Activity*, vol. 16, no. 1, pp. 1–12, 2019.

[101] A. Weiss, M. Brozgol, M. Dorfman et al., "Does the evaluation of gait quality during daily life provide insight into fall risk? A novel approach using 3-day accelerometer recordings," *Neurorehabilitation and Neural Repair*, vol. 27, no. 8, pp. 742–752, 2013.

[102] D. Trojaniello, A. Ravaschio, J. M. Hausdorff, and A. Cereatti, "Comparative assessment of different methods for the estimation of gait temporal parameters using a single inertial sensor: application to elderly, post-stroke, Parkinson's disease and Huntington's disease subjects," *Gait & Posture*, vol. 42, no. 3, pp. 310–316, 2015.

[103] J. Jankovic, "Gait disorders," *Neurologic Clinics*, vol. 33, no. 1, pp. 249–268, 2015.

[104] W. Pirker and R. Katzenschlager, "Gait disorders in adults and the elderly," *Wiener Klinische Wochenschrift*, vol. 129, no. 3-4, pp. 81–95, 2017.

[105] A. Erdemir, S. McLean, W. Herzog, and A. J. van den Bogert, "Model-based estimation of muscle forces exerted during movements," *Clinical biomechanics*, vol. 22, no. 2, pp. 131–154, 2007.

[106] C. Kirtley, "Clinical gait analysis: theory and practice," Elsevier Health Sciences, 2006.

[107] P. Mahlknecht, S. Kiechl, B. R. Bloem et al., "Prevalence and burden of gait disorders in elderly men and women aged 60–97 years: a population-based study," *PLoS One*, vol. 8, no. 7, article e69627, 2013.

[108] V. S. Thomas, E. V. Vandenberg, and J. F. Potter, "Non-neurological factors are implicated in impairments in gait and mobility among patients in a clinical dementia referral population," *International Journal of Geriatric Psychiatry*, vol. 17, no. 2, pp. 128–133, 2002.

[109] S. Shroff, C. Kamiya-Matsuoka, and K. Woodman, "An unusual cause of cerebellar ataxia in an immunocompromised elderly patient," *Journal of the Neurological Sciences*, vol. 340, no. 1-2, pp. 218–220, 2014.

[110] M. Iijima, H. Mitoma, S. Uchiyama, and K. Kitagawa, "Long-term monitoring gait analysis using a wearable device in daily lives of patients with parkinson's disease: the efficacy of selegiline hydrochloride for gait disturbance," *Frontiers in Neurology*, vol. 8, p. 542, 2017.