



Measurement method of determining natural and unnatural gaits using autocorrelation coefficients

Sangjin Park¹ · Sangil Choi²

Received: 11 December 2020 / Revised: 5 March 2021 / Accepted: 5 May 2021 /

Published online: 13 May 2021

© The Author(s) 2021

Abstract

Walking is the most common physical activity in humans, and gait can be used as a measure of human health. If the gait is unnatural or uncomfortable, it indicates a problem inside or outside the person's body. In particular, for the elderly, walking is used as an important indicator of their health status. In this study, we developed an algorithm that can determine whether human walking is natural or unnatural, by comparing the autocorrelation coefficients of the left and right foot. We used F1-scores to measure the accuracy of the gait result determined by the algorithm. Natural walking was accurately distinguished with 80% accuracy, and unnatural with 60% accuracy. Owing to the splint attached to one foot to express unnatural walking, both feet affected gait, resulting in slightly lower accuracy than natural walking. As a future study, it is possible to devise a method to improve accuracy by extracting various gait features that can be obtained through gait and using artificial intelligence algorithms such as machine learning or deep learning.

Keywords Human gait · Gait analysis · Autocorrelation coefficient · Gait cycle

1 Introduction

Walking is the most common human physical activity; especially for children or young people, it is a natural daily activity. However, for various reasons, elderly people may have difficulty

✉ Sangil Choi
schoi@gwnu.ac.kr

Sangjin Park
idsjsj2@sehan.ac.kr

¹ Department of Police Administration, Sehan University, Dangjin-si, Chungnam, South Korea

² Department of Computer Science & Engineering, Gangneung-Wonju National University, Wonju-si, South Korea

walking. The gait of the elderly with gait problems is not only clearly different from that of healthy people but also easily observed with the naked eye. However, it may be difficult to identify gait abnormalities because there is no clear medical standard or definition of normal gait. Gait abnormalities in the elderly are accompanied by a risk of fracture or brain damage due to a fall, limiting their independent activities. Furthermore, they may have to rely on their family members or other caregivers for assistance.

One of the methods traditionally used to assess gait abnormalities is a direct examination of a person's gait. The most representative way to evaluate gait abnormalities in medical fields is to evaluate gait with a method called "Timed up and go test." In this method, a subject stands up from a chair, walks 3 m, turns around, and returns. An examiner records the time it took to sit back down on the chair and observes the gait. There are a few problems with this test. First, there is a high probability that the subjective opinion of the examiner will be involved. Second, there is no objective data for determining gait abnormality, and finally, there is no clear standard for concluding that there is a problem with the gait.

Researchers have been considering a variety of gait analysis methods that would solve the problem of subjective human observation in determining gait abnormalities. In a study of gait analysis, normal and abnormal gait were differentiated using an optical motion capture system in a controlled experimental environment. After attaching a number of markers to specific parts of the subject's body, the subjects walked a certain distance. After capturing the subject's gait using several infrared cameras installed in the laboratory, the collected gait data were analyzed to determine gait abnormality. The advantage of this method is that it shows high accuracy in determining gait abnormalities. However, considerable effort is required to collect and analyze human gait data this way. In addition, a disadvantage of laboratory environment is that it is not realistic because it is far removed from the daily environment in which humans walk. Moreover, markers are attached to test subjects for experimentation, which may cause inconvenience in walking, hindering the collection of natural gait data.

According to a recent report on healthcare wearable devices, their market share in the healthcare sector is expected to increase rapidly in the future. Based on the recent statistics on senior citizens released by the National Statistical Office, there are 70,700,000 people aged 65 or older, accounting for 13.8% of the total population. It is estimated that by 2060, the proportion will reach approximately 40% of the total population, and the growing proportion of older adults is a global trend. Interest in the use of smart health devices is increasing across all age groups. Furthermore, owing to the recent interest in the Internet of Things (IoT) and the rapid development of wearable sensor devices, studies conducting gait analysis using inertial sensors, such as three-axial accelerometers, gyroscopes, or insole pressure sensors, are actively underway. The use of sensors is widely seen as an alternative to address the shortcomings of human observation and the optical motion capture system mentioned above. Wearable sensors have low power consumption, are easy to carry, and do not require a controlled experimental environment. The exclusion of the subjective opinion of the experimenter may also be an advantage of using wearable sensors in gait analysis study.

The purpose of this study is to develop an algorithm that can measure unnatural human gait using wearable devices. To achieve this goal, the research is divided into three parts: investigation and analysis of the latest gait inspection methods, gait data collection and feature extraction using IoT wearable devices, and unnatural gait measurement algorithm development using an autocorrelation coefficient.

This paper is organized as follows. First, studies that measured gait abnormalities are summarized in Section 2, and specific methods for distinguishing natural and unnatural gait

using an autocorrelation coefficient are described in Section 3. Next, Section 4 describes in detail the test subjects and experimental methods for the gait analysis and the experimental results of the gait analysis algorithm using the autocorrelation coefficient devised in this study. Finally, the research results are summarized and future research directions are presented in Section 5.

2 Related works

Unnatural gait is very common in the elderly: 63% of the elderly in nursing homes have a gait disorder, and 8–19% of those who are not hospitalized in a medical institution have a similar problem [22]. Figure 1 shows mediolateral sway (internal and external axis) and stride velocity (stride width) using wearable sensors (body-attached sensors) targeting healthy community-based groups of all ages. In Fig. 1, it can be seen that as the person ages, the regression of the gait function becomes clear [11]. Gait abnormalities in the elderly are related to diseases and act as a limiting factor in their quality of life. Therefore, the field of geriatric medicine should measure gait abnormalities in the elderly, identify the causes of the abnormal gait, and provide pharmaceutical or interventional treatment accordingly.

For example, a study of dementia among the elderly indicated that one of the common antecedents of dementia were gait abnormalities, such as a change in the shape of the gait or slowing down, and the incidence of dementia actually increases [23]. Therefore, gait analysis can help predict future dementia. The graphs in Fig. 2 show the gaits of the healthy elderly and dementia patients, and it is clear that the two graphs are different [30].

Recent gait analysis studies have assumed that human gait is typically a series of repetitive patterns. Therefore, after determining the patterns within the gait cycle, efforts have been made to observe and analyze temporal characteristics according to the patterns [5, 12, 26, 28]. The human gait cycle begins by lifting one foot away from the ground and moving forward; this is called toe off (TO). The next step occurs when the foot touches the ground; this is called heel strike (HS). In the case of TO, force to move forward is applied, so the speed increases. In contrast, in the case of HS, the speed decreases as the foot touches the ground. As soon as one foot touches the ground, the other foot performs TO. These two states occur in succession. One

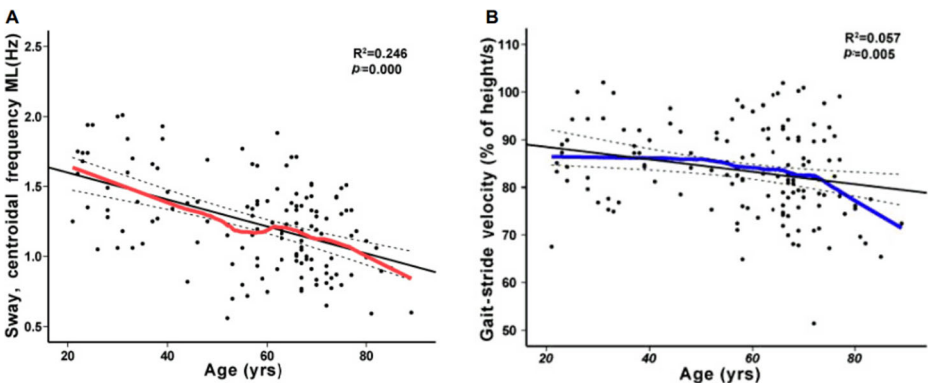


Fig. 1 (A) mediolateral sway (ML) and (B) stride velocity measured using wearable sensors. The black dotted lines are linear regression, and the dark solid lines represent the data values determined by a non-parametric regression analysis [11]

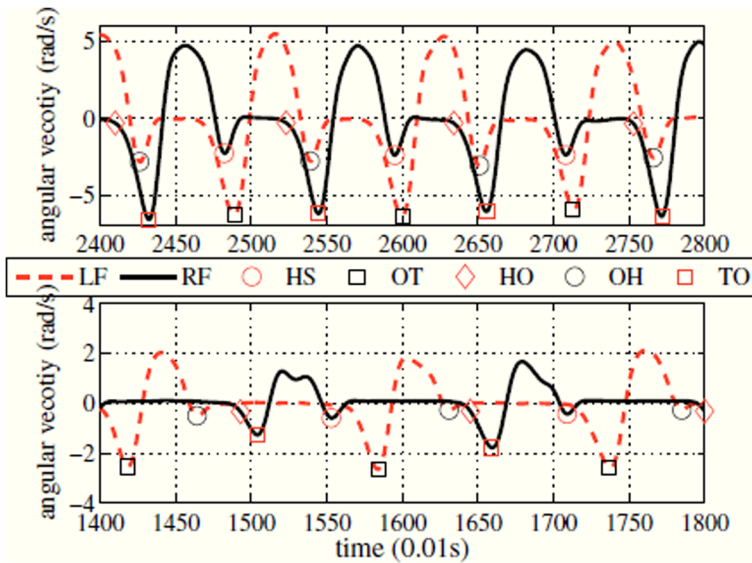


Fig. 2 Gait sequences of healthy person (upper graph) and dementia patients (lower graph) [30]

gait cycle starts with the TO state of one foot and ends with the next TO state of the same foot. Acceleration sensors are accepted as suitable for representing the change in speed in TO and HS; thus, many researchers are actively using accelerometers for gait analysis.

A peak detection algorithm was applied to detect specific events occurring within the gait cycle [20, 25]. In another study, hidden Markov models were adapted [14]. In these studies, the researchers calculated gait features related to time, such as swing time, double stance time, and stance phase time, and used them for gait analysis as gait features or characteristics. Swing time is defined as the time from TO to HS of one foot, double stance time refers to the time when both feet are on the ground, and the stance phase time is the time from HS to TO of one foot. According to gait analysis studies, obtaining the abovementioned gait characteristics using inertial sensors produces relatively accurate results.

Walking speed has also been studied as an important gait feature of humans. For people over 65 years of age, walking speed can be seen as a distinct physical change due to decreased physical capacity [21, 24, 27, 33]. Therefore, calculating walking speed accurately is a goal of many researchers in the gait analysis field. Studies investigating walking speed applied methods of human gait model [13], integration calculation [3], or machine learning approach [7, 8, 10, 31].

Another approach to extracting gait characteristics is to use the ground reaction forces as gait features measured by insole sensors [1, 6, 17, 18]. Before insole sensors were used, the ground reaction force generated each time the subject walked on a force plate was calculated. However, this method is expensive, and it is difficult to analyze gait if an experimental environment is not provided. The main disadvantage is that it is difficult to obtain information on natural gait in daily life. On the other hand, if the insole sensor is used, walking data can be obtained from the sensor by attaching it to the test subject's shoe.

In gait analysis studies, gyro sensors tend to be used in addition to the acceleration and insole sensors. The inertial sensors are attached to the subject's body to acquire gait data, and their positions vary from study to study. The most commonly used positions are waist, thighs,

shanks, and feet. The number of sensors used in the experiments varies from one to six. The primarily measured gait features are temporal characteristics, such as stride length, stride time, swing time, stance time, cadence, step time, and step length. In addition, the pressure values from the insole sensor and the angle values from the gyro sensor were also extracted and used.

In most gait studies, data related to specific events within the gait cycle are used. A study was conducted with a pressure sensor to analyze the gait motion that occurs between toe off and heel strike, an acceleration sensor to measure the acceleration component, and a gyro sensor to determine the rotational speed [9]. In another study, four inertial measurement units were attached to the thigh and shank area of each leg. The joint angular trajectory and gait cycle during walking were obtained to analyze the gait [32]. As shown in these studies, to accurately obtain the characteristics of walking, at least three to four sensors were used to record various data. However, this is inconvenient, especially for the elderly or patients with gait abnormalities, and must be resolved for commercialization. To address this problem, we conducted research on gait abnormality measurement with one acceleration sensor. An acceleration sensor was chosen because it is often used in combination with user information for various behavior recognition tasks, such as calculating the amount of activity or finding out a life pattern.

The most recent research trend on gait pattern analysis is mainly applying machine learning or deep learning techniques [2, 4, 15, 16, 19, 29]. Beyond the conventional method that directly designed the gait analysis algorithm to obtain the desired result by detecting the gait cycle from the raw data collected from the sensor and analyzing the characteristics of the gait cycle for each individual or group, efforts to obtain meaningful results by using a large amount of walking data to learn through various machine learning and deep learning algorithms are in full swing. Accelerometers, gyro sensors, and insole sensors used in existing gait analysis studies are still used for gait data collection and analysis in machine learning oriented gait research. The main gait features obtained in these studies using machine learning techniques show different patterns depending on the final results to be obtained in the study. In addition, it can be seen that there are various types of machine learning and deep learning algorithms used in each study. Gait analysis using machine learning is expected to receive more and more researchers' interest in the near future.

3 Proposed method

3.1 Definition of natural and unnatural walking

Walking refers to the process of moving the body forward while maintaining the center of the body. The start of the gait cycle is defined as the point at which the heel strikes the ground (initial contact in Fig. 3). During the first part (approximately 0% to 60%) of the gait cycle, one foot is touching the ground (the right foot in Fig. 3). Then, one foot is kept in contact with the ground while the other foot moves to move the upper body forward. In this process, the other foot is in the air. Finally, the foot touches the ground, and the foot that was previously supporting the ground is no longer in contact with the ground (initial swing in Fig. 3). When one foot first touches and then falls off the ground, it is called the stance phase. According to Fig. 3, the right foot first touched the ground. Therefore, the right foot is initially in the stance phase. The next step is called the swing phase because the foot that has fallen off the ground remains in the air. At this stage, the body moves forward, and acceleration occurs. Finally, the

foot that was in the air lands on the ground, causing deceleration. In this way, both feet go through the stance phase and swing phase once periodically and continuously, and this process is called the gait cycle.

From the definition of the gait cycle, we can infer that people who walk naturally will have a repetitive and constant gait pattern created by the left and right feet in the stance phase and swing phase. Furthermore, it is conceivable that a certain pattern of acceleration and deceleration, which occurs as the body moves, especially in the swing phase, will be created. Based on this, natural walking refers to a gait in which the medial side of the leg is in a straight line while maintaining the proper stride length in an upright position. On the other hand, unnatural walking is out of the range of natural walking and occurs in various combinations, such as decrease in walking speed, decrease in stride length, or foot drag. For example, gait abnormalities caused by arthritis or muscle disease break the center of gait by focusing only in one direction depending on the location of the damaged area. It also acts as a cause of disrupting the gait cycle defined earlier. Figures 4 and 5 show typical raw 3-axis acceleration values of two different human subjects. In these pictures, *Accx* stands for the acceleration of a sideways direction (left and right), *Accy* represents the one of an up-down motion, and *Accz* indicates that of forward and backward movement.

In this paper, instead of using the terms normal and abnormal gait, which has been commonly used in gait analysis studies, the terms corresponding to our study were newly defined. Walking without any problems was defined as *natural walking* (NW). The gait contrary to natural walking was defined as *unnatural walking* (UW). The reason for defining a new term for this study is that in the case of NW, it is not significantly different from the normal walking we talked about earlier, but in the case of UW, it is a walking method that we artificially devised for this study, thus avoiding confusion for researchers. UW will be described in detail in the next section.

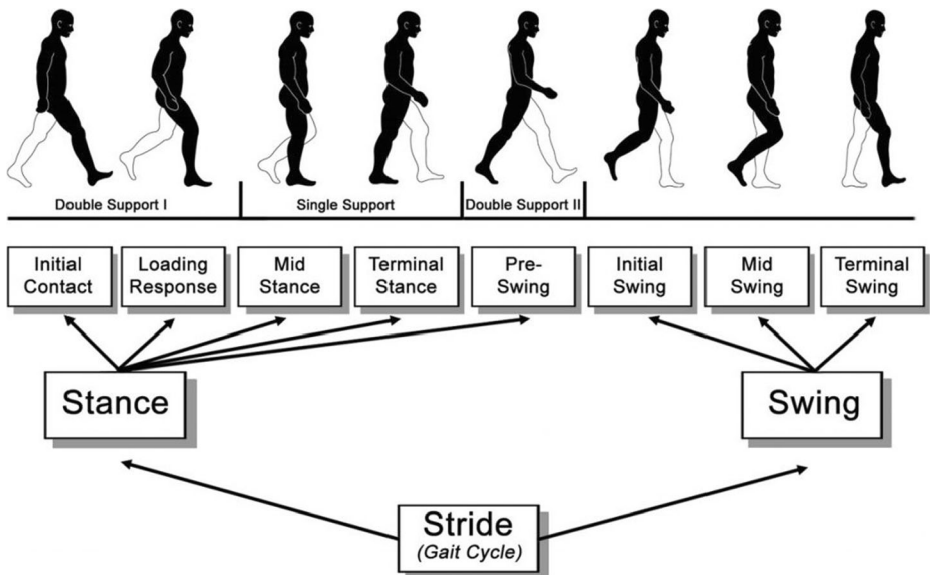


Fig. 3 Human gait cycle (Figure from: Stöckel, Tino & Jacksteit, Robert & Behrens, Martin & Skripitz, Ralf & Bader, Rainer & Mau-Moeller, Anett. (2015). The mental representation of the human gait in young and older adults. *Frontiers in Psychology*. 6. 943. <https://doi.org/10.3389/fpsyg.2015.00943>)

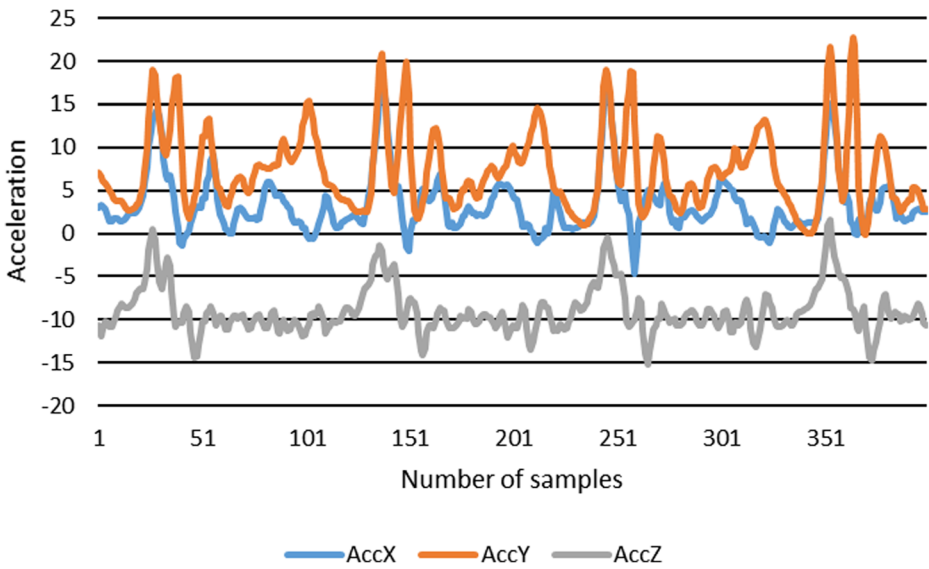


Fig. 4 Raw acceleration data of subject I

3.2 Reasons for using autocorrelation coefficients

The correlation coefficient is defined as a numerical value for the correlation between two variables x and y . Since the correlation coefficient was proposed by *Karl Pearson*, it is also referred to as *Pearson’s correlation coefficient*. In our study, walking data of participants were collected while they walked a given distance for a certain amount of time. The walking data were collected using an acceleration sensor according to a predetermined sampling rate. Therefore, the walking data are defined as the value of the 3-axis accelerometer data collected

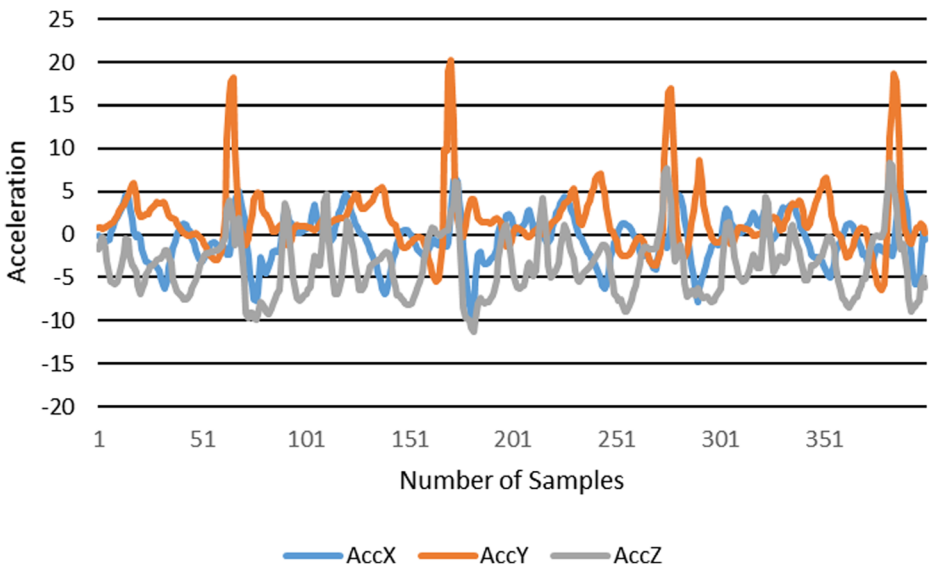


Fig. 5 Raw acceleration data of subject II

according to the change in time. The formula for the correlation coefficient according to the sample of walking data collected using the acceleration sensor is as follows:

$$r_{xy} = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^n (x_i - \bar{x})^2} \sqrt{\sum_{i=1}^n (y_i - \bar{y})^2}}, \quad (1)$$

where:

n is a total sample size in a given experimental time
 x_i, y_i are the individual acceleration samples of gait at time t

$$\bar{x} = \frac{\sum_{i=1}^n x_i}{n}$$

$$\bar{y} = \frac{\sum_{i=1}^n y_i}{n}.$$

According to Eq. (1), two variables x and y are required. In other words, the above formula is necessary when investigating the correlation between two distinct variables x and y , but the desired correlation in this study should be based on the correlation between the same acceleration values over time, rather than between two different variables. Therefore, it is difficult to apply the general correlation formula presented above to our study. Unlike general correlation coefficients, autocorrelation coefficients are suitable for elucidating correlations between the same variables over time. In this study, the correlation between the acceleration values from the left and right feet is investigated to determine whether walking is natural or artificial. Thus, it was judged that it is more appropriate to use an autocorrelation coefficient, which can obtain a correlation between the same data (acceleration values in our research) rather than a general correlation coefficient that identifies the correlation between different variables x and y .

The correlation between the value x_{t_1} at time t_1 and the value x_{t_2} at time t_2 is called autocorrelation with x_t as an arbitrary process. The autocorrelation function defines this relationship. The value obtained through it is the autocorrelation coefficient, which indicates the degree of autocorrelation between different data, and is expressed as a value between -1 and 1 . If this value is close to 1 , the two data points are very close with the same characteristics. Otherwise, there is no meaningful relationship between two data points.

As a result, in this study, a gait analysis was conducted using the autocorrelation coefficient defined in Eq. (2) without applying the general correlation coefficient presented in (1). The autocorrelation coefficient r_h applied in this study is defined as follows:

$$r_h = \frac{\sum_{t=h+1}^n (x_t - \bar{x})(x_{t-h} - \bar{x})}{\sum_{t=1}^n (x_t - \bar{x})^2}, \quad 0 \leq h \leq n-1, \quad (2)$$

where:

n is a total sample size in a given experimental time.
 x_t is the individual acceleration sample of gait at time t .
 \bar{x} the average of x_1, x_2, \dots, x_n .

The autocorrelation function is effective in calculating whether data values at a specific time are related to each other or not when compared to previous data values. In the formula given in (2), the autocorrelation coefficient r_h is calculated by the autocorrelation function over time. This function returns a value between -1 and 1 for a given input sample. Therefore, the autocorrelation coefficient for a specific sample h is calculated by formula (2), and we use this autocorrelation coefficient value to distinguish between NW and UW.

3.3 Algorithm for determining natural and unnatural gaits using autocorrelation coefficients

The proposed algorithm consists of five steps. In the first step, the values from the acceleration sensor built into the smartphone are collected using a mobile application created by us for this study. People participating in the experiment first run the mobile application to check whether the acceleration sensor value is displayed properly. Then they put the smartphone in the right pocket of their pants and start the experiment according to the instructions of the experimenter. Details related to gait data collection are described in Section 4.

The second step is to refine the gait data collected in the first step. In most cases, the initial acceleration value contains noise due to various environmental factors. Therefore, it is necessary to remove this noise. In this study, it was removed using the low pass filter (LPF) algorithm. LPF is applied as shown in Fig. 6 to remove abnormally high or low values among the 3-axis acceleration sensor data. In particular, Fig. 6 only shows the acceleration value of a forward and backward direction (Z-axis). As shown in the graph, we move on to the next step after removing the noise. Figure 6 shows the noise removal results for the z-axis, not all acceleration sensor data on the x, y, and z axes. The reason for this is that in Section 3.1, the act of walking is explained as an activity that makes the whole body move forward with the force that occurs when two legs alternately touch the ground. The x, y, and z axes of the 3-axis acceleration sensor are each determined in a direction for generating acceleration. In particular, the z-axis represents acceleration produced through movement in the forward and backward directions. Therefore, the acceleration created in the z-axis can best represent the walking behavior of a person. For this reason, in this study, the gait analysis was only conducted using the acceleration value from the z-axis.

After the filtering process, the gait data of all test subjects were normalized to a value between 0 and 1. The reason for this is that the ranges of acceleration values of different subjects are diverse, as shown in Figs. 10 and 11. By performing normalization, it is possible to analyze the gait data within a common range. In addition, to compare the data, they were normalized by substituting the data distribution with a value between 0 and 1.

After normalization is performed, the next step applies the autocorrelation coefficient to the normalized gait data to differentiate NW and UW. As a criterion for determining natural and unnatural gait, the autocorrelation coefficients of the left and right footsteps were calculated, and if they were above a certain threshold, it was determined as NW. If they were lower than the threshold value, the walking was determined to be UW. To this end, data were distributed for each step of the left and right feet, and after removing the data from the front and back, which measured unnecessary actions, 800 data samples of 8 s were used. To determine the optimal amount of data per one step, the accuracy was obtained by extracting the results for 1 to 200 data samples for one step.

For instance, if the sampling rate of an acceleration sensor is 50 ms, it can be possible to get 20 data samples per second. We identified the most appropriate number of data samples to

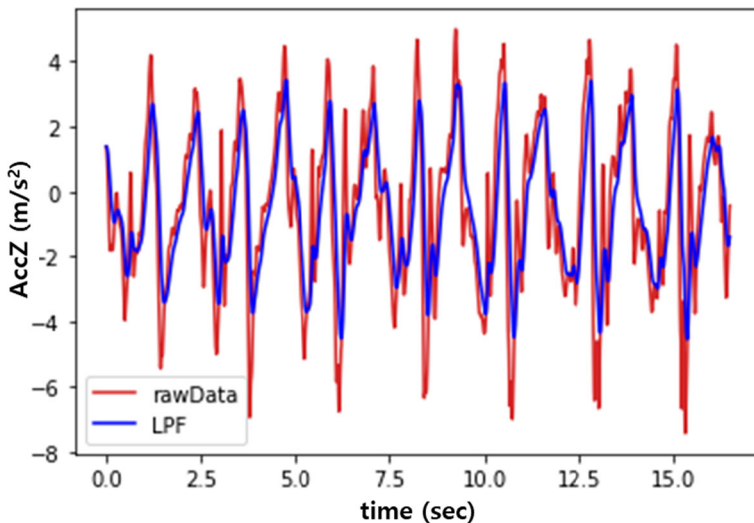


Fig. 6 Low pass filter (LPF) applied to acceleration data

represent a single step for all of the participants in the experiment. Figure 7 shows that, as a result of analyzing the gait of all participants in 200 data samples, when the experiment was conducted by defining 67 data samples as one step, the accuracy of distinguishing between NW and UW was the highest. Based on this result, the original data collected from the experimenters' walking was analyzed by grouping 67 data samples into one step.

The final step was to determine the accuracy of the proposed algorithm. All gait analysis processes were terminated with measuring the accuracy using the F1-score method. Figure 8 depicts the process of gait analysis by processing gait data in a CSV format, and Fig. 9. shows the pseudocode of our proposed algorithm. We discuss the estimation of the accuracy using the F1-score method in Section 4.

4 Experiment

4.1 Experimental settings

Gait data were collected from 25 healthy men and women in their twenties without trauma or disease. They each performed NW and UW five times. We conducted experiments with the approval of the Institutional Review Board (IRB, GWNUIRB-2020-33) for human experiments, and explained all possible risks to the participants. Each participant signed an informed consent form approved by the IRB. The experiment was conducted by installing a splint approximately 5 cm high on the sole of the participant's foot for UW. By doing this, the participants in the experiment had no choice but to walk unnaturally and differently from their NW. When the subjects walk in this manner, the balance of their body will be unnatural, and accordingly, it is expected that they will be slower than their usual pace. As a result, it was predicted that the values recorded by the acceleration sensor would be smaller than those recorded during NW. To collect NW data, the participants were instructed to walk comfortably at their usual walking speed. Data were collected while the participants walked on a flat

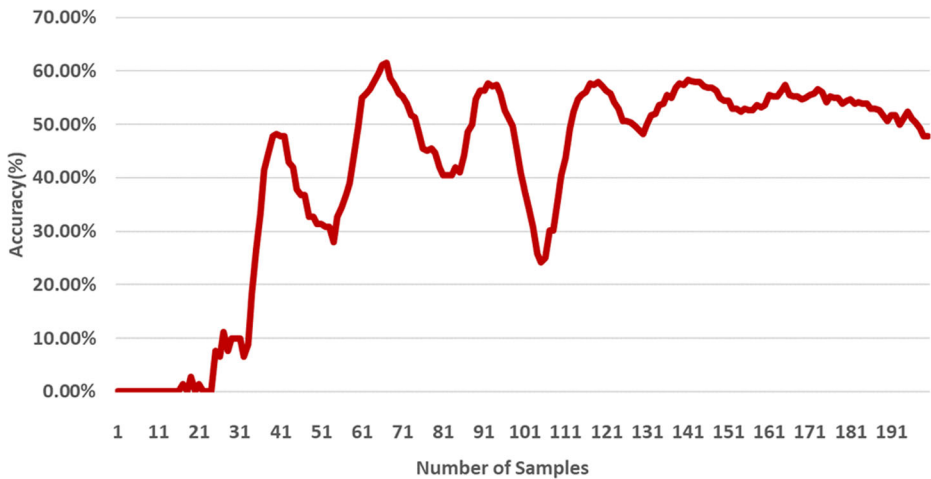


Fig. 7 F1-score based on the number of data samples

surface at their usual walking pace. As can be seen in Figs. 10 and 11, the acceleration value measured by the acceleration sensor in the case of NW was greater than that in the case of UW. Figures 10 and 11 show the NW and UW of two different participants. During NW, there are no restrictions, so the acceleration value reflects the speed of one's walking. On the other hand, during UW, with a splint on the sole of the foot, normal acceleration is not produced because the walking is unnatural. Therefore, as shown in the figure, the acceleration value is lower than during NW. In the initial setting, all the participants wore tight jeans, sneakers, and the smartphone was placed in the right jean pocket. The experiment was conducted in an indoor corridor, and the walking distance was 10 m. The smartphone application collected acceleration data at 100 Hz. To measure the natural gait, the participant walked the 10 m distance five times with the smartphone application running. To measure unnatural gait, approximately 5-cm splint was attached to the sole of the shoe by wrapping a compression bandage. Then, the

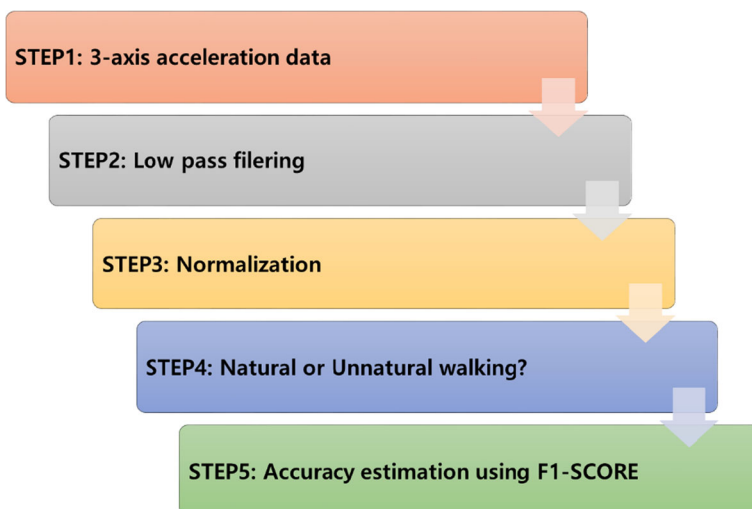


Fig. 8 Gait analysis procedure

Algorithm: determining natural and unnatural gaits

```

1: procedure Differentiating NW and UW
2:   DATA  $\leftarrow$  Raw acceleration data
3:   LPF  $\leftarrow$  a low pass filtering algorithm
4:   MIN  $\leftarrow$  the minimum function
5:   MAX  $\leftarrow$  the maximum function
6:   /* Apply LPF to DATA for noise cancellation */
7:   LPF_DATA  $\leftarrow$  LPF(DATA)
8:   /* Normalization of LPF_DATA */
9:   NORM_DATA  $\leftarrow$   $\frac{LPF\_DATA_i - MIN(LP\_DATA)}{MAX(LP\_DATA) - MIN(LP\_DATA)}$  for all LPF_DATA
10:  CNT  $\leftarrow$  the number of data samples per one step
11:  while all  $\varphi \in$  NORM_DATA
12:    do
13:      STEP_COUNT[ ] =  $\frac{\varphi}{CNT}$ 
14:    end
15:  SCOUNT  $\leftarrow$  size of STEP_COUNT[ ]
16:  /* Calculate autocorrelation coefficient ACF */
17:  CORR_COEF  $\leftarrow$  correlation function
18:  while  $i <$  SCOUNT
19:    do
20:      ACF = CORR_COEF(STEP_COUNT[ $i-1$ ], STEP_COUNT[ $i$ ])
21:       $i \leftarrow i + 1$ 
22:    end
23:    if ACF > threshold then Natural Walking
24:    else Unnatural Walking
25:  end

```

Fig. 9 Proposed gait detection algorithm

smartphone application was running while the participants walked the 10 m a total of five times.

4.2 Gait data collection

4.2.1 Mobile applications

Walking time and values in the x (left-right), y (up-down), and z (forward-backward) directions of the 3-axis acceleration sensor built into the smartphone were saved together. The smartphone used in this study was a Samsung Galaxy Note 9 (SM-N960F), and the walking data were collected using the acceleration sensor built into this smartphone. For data collection, we made our own mobile application using Android Studio. Figure 12 shows the running screen of our application designed to collect the participants' walking data in real time. The collected data were saved using Google's Firebase real-time database so that they were saved whenever the sensor value changed. In addition, Firebase authentication was used to authenticate data identity. Therefore, the UID of the participants was obtained according to the authentication request, and their data values corresponding to the child node are stored. Here, the UID is a value that is used as the user's unique identifier every time the application is installed, so that the subject corresponding to the data can be identified. However, this method required the participants to install or subscribe to the application. Therefore, as shown in Fig. 12, the experiment step was simplified by changing the method to receive an ID from each participant before collecting their data.

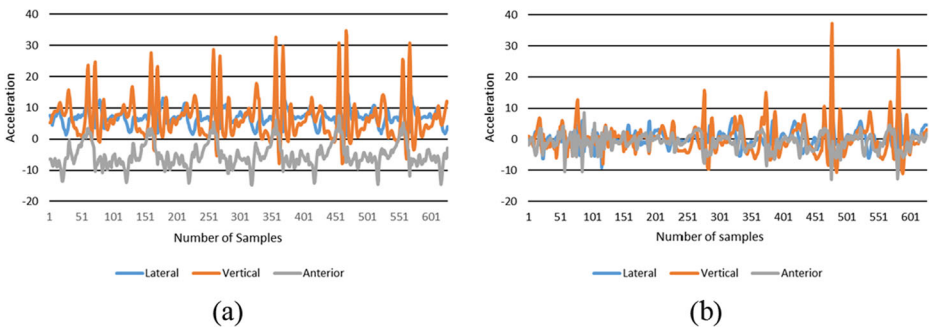


Fig. 10 Natural (a) and unnatural(b) raw acceleration data of subject A

4.2.2 Data collection and analysis

The process for collecting and analyzing gait data consisted of four steps. First, by using the self-produced mobile application mentioned in Section 4.2.1, the walking data of the participants were collected in real time. Second, the collected walking data were automatically stored in the Firebase database in Jason format. Third, we extracted the data stored in the Jason format in Firebase for gait data analysis. Finally, by using the API provided by the Pandas library, the gait data format in Jason format is changed to the CSV format. The gait analysis algorithm is applied to the gait data in CSV format to obtain the autocorrelation coefficient mentioned in Section 3 and the data are classified into normal gait and abnormal gait. The gait analysis is completed by calculating the accuracy of the classified result using F1-score. Figure 13 shows the configuration of the system and the procedure for data collection.

4.3 Experimental results

4.3.1 F1-score

F1-score was used to measure the accuracy of the algorithm for distinguishing between normal and abnormal walking. For example, suppose there are a total of 100 gait samples. Of these, 50 samples are of normal walking and 50 for abnormal walking. If we provide a total of 100 samples as input values to our proposed algorithm, and if the algorithm classifies all 50 normal walking samples as normal walking and all 50 abnormal walking samples as abnormal

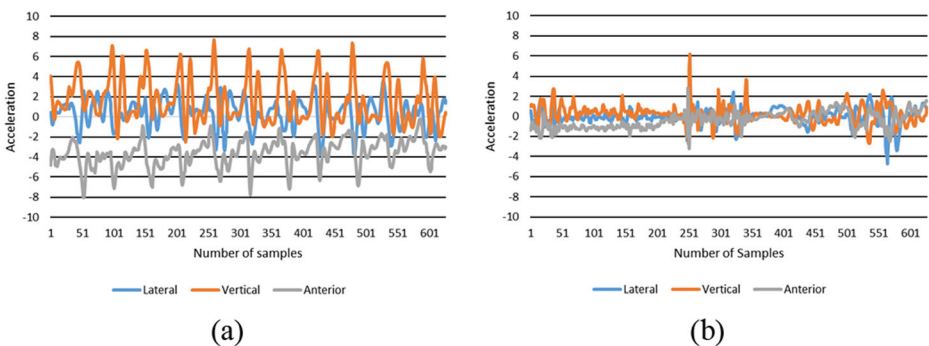


Fig. 11 Natural (a) and unnatural (b) raw acceleration data of subject B

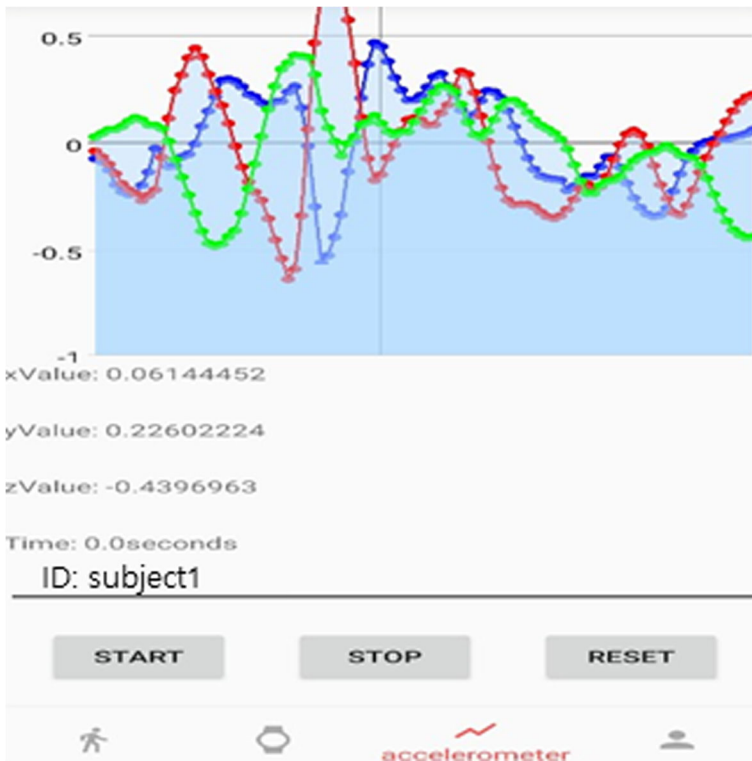


Fig. 12 Values of 3-axis acceleration from the acceleration sensor in the smartphone as shown in the application

walking, the accuracy of the algorithm is 100%. F1-score is calculated using the formula in (3). In Eq. (3), *precision* is the ratio of what is actually classified as *true* in the model, and *recall* is the ratio of what the model predicts as *true* among what is actually *true*. *Precision* and *recall* are shown in Eqs. (4) and (5). In the formulas of *precision* and *recall*, *TP* stands for “true positive,” *FP* for “false positive,” and *FN* for “false negative.”¹

$$F1\text{-SCORE} = \frac{2}{\text{recall}^{-1} + \text{precision}^{-1}} \quad (3)$$

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

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

¹ <https://en.wikipedia.org/wiki/F-score>

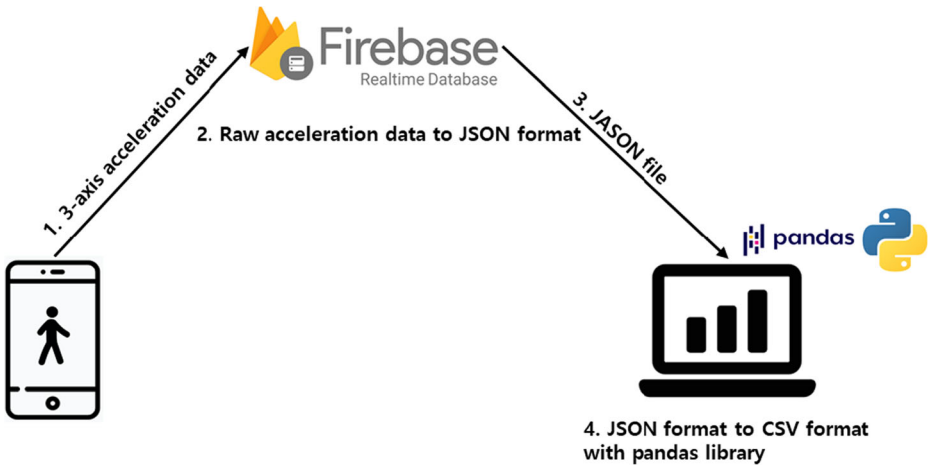


Fig. 13 System organization for data collection and analysis

4.3.2 Accuracy calculation

To calculate the accuracy of the algorithm proposed in this study, the accuracy was computed by dividing one step to up to 200 points of data. Using 67 points showed the highest accuracy at 61.5% (Fig. 7). Therefore, based on a total of 800 data samples, 67 were grouped into left and right foot data, and the autocorrelation coefficient between the data divided into left and right feet was calculated and used as input of the algorithm. The same process was performed for normal and abnormal walking data. Through this process, the classification of walking determined by the proposed algorithm was calculated. The next step was to use the F1-score method to calculate the accuracy between the actual value and the results of the algorithm. Table 1 shows the results of this calculation as a confusion matrix, and Fig. 14 shows the pie chart results for accuracy.

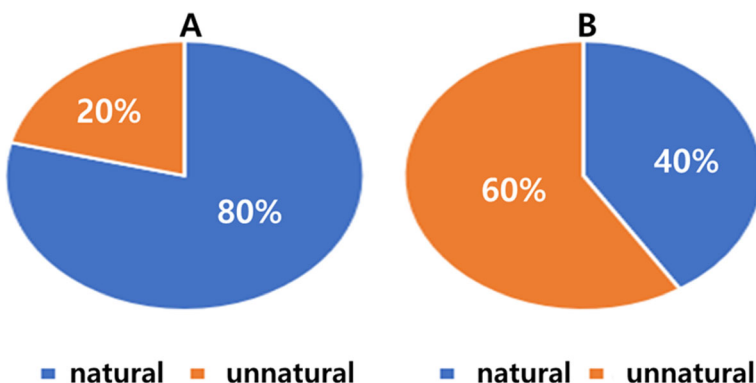
When normal walking data are provided as input to the algorithm, the probability of recognizing normal walking is 80% and the probability of determining abnormal walking is 20% (Fig. 14A). When abnormal data are entered into the algorithm, the probability of determining abnormal walking is 60%, and the probability of determining normal walking is 40% (Fig. 14B). The probability that the proposed algorithm judges normal walking as normal is 80%, whereas the probability of determining abnormal walking as abnormal is 60%, which is relatively low accuracy. The reason for the low accuracy in determining abnormal walking is that there may be various environmental and other factors that we did not predict in advance. However, we learned through experiments that both normal and abnormal walking show a certain pattern. In addition, especially in the case of abnormal walking, due to the splint attached to only one foot, there may be an element that prevents the other foot from being able to walk normally. If so, it is possible to speculate that the accuracy in determining the abnormal walking was relatively low because a periodic walking pattern was preserved. This could cause the autocorrelation coefficient to be above the threshold, even if the pattern was different than the pattern in normal walking.

Table 1 Confusion matrix

		Original label	
		Natural	Unnatural
Label produced by proposed algorithm	Natural	80%	40%
	Unnatural	20%	60%

5 Conclusion

Walking is the most common and repetitive physical activity that a person performs in daily life, and human gait contains much information. Various attempts have been made to develop technologies that can be used in the healthcare field by using ultra-small sensor devices equipped with advanced information technologies, such as IoT, big data, cloud computing, and artificial intelligence. The era has arrived in which information on human gait can be obtained without using an experimental environment equipped with expensive imaging equipment and special facilities. If sensor devices are used effectively, it is expected that in the future, instead of administering treatments after the disease occurs, the disease would be predicted in advance and the disease would be prevented before it becomes serious. Three conditions need to be satisfied to achieve this goal. First, micro-sensor devices usually operate at low power and have limited computing capacity. Therefore, gait analysis should be possible while consuming minimal power with a small amount of calculation. Second, it is necessary to collect walking data using a minimum number of sensors so that there is no inconvenience in daily life. Lastly, the results of gait analysis should be accurate and reflect the current state of health. In this study, we created an algorithm to discriminate between normal and abnormal walking using the autocorrelation coefficient. Smartphones used by ordinary people every day were used, and walking data were collected using the acceleration sensor built into the smartphone. The gait analysis consisted of three steps (filtering, normalization, and calculation of autocorrelation coefficient), and the proposed algorithm is designed to obtain accurate results with the least possible computation. F1-score was used to calculate the accuracy of the algorithm. Normal walking data were identified with 80% accuracy and abnormal walking

**Fig. 14** Accuracy of the proposed algorithm

data with 60% accuracy. As mentioned in the experimental results section, in the case of abnormal walking, the foot with a splint prevented normal walking, affecting the walking of the other foot, resulting in relatively low accuracy. Compared with other gait analysis studies, our study tried to distinguish between normal and abnormal walking, the most basic task for gait analysis, with a relatively simple methodology. However, a clear definition of normal and abnormal gait must be established as an important task to be preceded first in the gait research field.

As a future research task in gait analysis, it is meaningful to conduct a study to discriminate between normal and abnormal gait using machine learning or deep learning techniques, which are currently receiving much attention. Similar to other algorithms or programs, artificial intelligence techniques require clean input data. In other words, the extent to which the input data reflect the phenomenon is the key to success. In addition, research to extract features that well represent the features of human walking should also be conducted.

Acknowledgements This study was supported by the National Research Foundation of Korea (NRF) grant funded by the Korean government (MSIT) (No. 2020R1G1A1013937).

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

1. Adam MH, Toshiki K, Heather AHKBF, Stacy JMB (2013) Kinetic gait analysis using a low-cost insole. *IEEE Transaction Biomedical Engineering* 60(12):3284–3290
2. Adrian R-S, George P, Carlos B-P, Uriel M-H (2019) A multilayer interval type-2 fuzzy extreme learning machine for the recognition of walking activities and gait events using wearable sensors. *Neurocomputing* 389:42–55
3. Anastasopoulou P, Shammam L, Hey S (2012) Assessment of human gait speed and energy expenditure using a single triaxial accelerometer. In: *Proceedings of International Conference of Wearable Implantable Body Sensor Network*, pp. 184–188.
4. Clara PB, Lea G, Miguel AGB, Jerome N, Fabian S, Martin S, Tim A, Simone T (2020) In-ear accelerometer-based sensor for gait classification. *IEEE Sensors J* 20(21):12895–12902
5. Del Din S, Elsbehabi M, Galna B, Hobert MA, Warmerdam E, Suenkel U, Brockmann K, Metzger F, Hansen C, Berg D, Rochester L, Maetzler W (2019) Gait analysis with wearables predicts conversion to Parkinson disease. *Ann Neurol* 86(3):357–367
6. Drew BF, Pierre A, Noam YH, Irina A-B, Stephen DK, William AB, Ann MS (2013) Vertical ground reaction force-based analysis of powered exoskeleton-assisted walking in persons with motor-complete paraplegia. *J Spinal Cord Med* 36(4):313–321
7. Eladio M (2011) Novel method for stride length estimation with body area network accelerometers. In: *proceedings of IEEE Topical Conference on Biomedical Wireless Technologies, Networks, and Sensing Systems (BioWireless)*, pp. 79–82.
8. Eladio M, Victor S, Ruzena B (2011) Determination of a patient's speed and stride length minimizing hardware requirements. In: *Proceedings of International Conference on Body Sensor Network*, pp. 144–149.

9. Geun Sub H, Seung Han Y, Sang Ryong L, Jong Gyu L, Choon-Young L (2012) A study on particular abnormal gait using accelerometer and gyro sensor. *Journal of the Korean Society for Precision Engineering* 29(11):1199–1206
10. Harshvardhan V, Adar E, Donna S-Me, Gaurav SS (2010) Toward free-living walking speed estimation using Gaussian process based regression with on-body accelerometers and gyroscopes. In: *Proceedings of 4th International Conference on Pervasive Computer Technologies for Healthcare*, pp. 1–8.
11. Jeong-Ho P (2017) Characteristics of gait in the elderly: Normal vs. Abnormal. *J Korean Neurol Assoc* 35(4):1–4
12. Jinxiao L, Zhelong W, Xin S, Sen Q, Hongyu Z, Ming G (2018) Quantitative analysis of abnormal and Normal gait based on inertial sensors. In: *proceedings of the 2018 IEEE 22nd International Conference on Computer Supported Cooperative Work in Design*, pp 365–370.
13. Kaushik N, Nicholas RG, Roozbeh J (2011) Modeling human gait using a Kalman Filter to measure walking distance. In: *Proceedings of 2nd Conference of Wireless Health*, pp. 34–1–34-2.
14. Mannini A, Sabatini AM (2011) A hidden Markov model-based technique for gait segmentation using a foot-mounted gyroscope. In: *Proceedings of Annual International Conference of IEEE Eng Med Biol Soc*, pp. 4369–4373.
15. Mc Ardle R, Silvia DD, Brook G, Alan T, Lynn R (2019) Differentiating dementia disease subtypes with gait analysis; feasibility of wearable sensors. *Gait Posture* 76:372–376
16. Mikihisa N, Sung-Gwi C, Ming D, Gustavo AGR, Jun T, Tsukasa O (2019) Prediction of plantar forces during gait using wearable sensors and deep neural networks. In: *Proceedings of 41st Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, pp. 3629–3632.
17. Muniz AMS, Liu H, Lyons KE, Pahwa R, Liu W, Nobre FF, Nadal J (2010) Comparison among probabilistic neural network, support vector machine and logistic regression for evaluating the effect of subthalamic stimulation in Parkinson disease on ground reaction force during gait. *J Biomech* 43(4):720–726
18. Murad A, Thompson S-G, Huiying Y, Olac F, Richard B, Amr A (2011) Automatic classification of pathological gait patterns using ground reaction forces and machine learning algorithms. In: *Proceedings of Annual International Conference Engineering in Medicine & Biology Society*, pp. 453–457.
19. Robert L, Timothy M (2020) Network centric therapy for machine learning classification of hemiplegic gait through conformal wearable and wireless inertial sensors. In: *Proceedings of the 8th IEEE International Conference on E-Health and Bioengineering (EHB)*, pp. 1–4.
20. Salarian A, Russmann H, Vingerhoets FJ, Dehollain C, Blanc Y, Burkhard PR, Aminian K (2004) Gait assessment in Parkinson's disease: toward an ambulatory system for long-term monitoring. *IEEE Trans Biomed Eng* 51(8):1434–1443
21. Salarian A, Burkhard PR, Vingerhoets FJ, Jolles BM, Aminian K (2013) A novel approach to reducing number of sensing units for wearable gait analysis systems. *IEEE Trans Biomed Eng* 60(1):72–77
22. Sang-Hyun L (2007) Evaluation of elderly patients with gait and balance disorders. *Proceedings of the 8th Conference on the Korean Academy of Clinical Geriatrics* 8(1):346–351
23. Seung-Wook C, Byul-Nim L, Jae-Moon L (2009) A study on gait ability and balance ability of dementia-high-risk senior citizens. *The Korean Society of Sports Science* 18(2):1117–1124
24. Shanshan C, John L (2013) Nonlinear feature for gait speed estimation using inertial sensors. *Proceedings of International Conference of Body Area Network*, In, pp 185–188
25. Shanshan C, Christopher LC, John L, Bradford CB (2011) Extracting spatio-temporal information from inertial body sensor networks for gait speed estimation. In: *Proceedings of International Conference on Body Sensor Networks*, pp. 71–76.
26. Sithi F, Wahida B (2016) Abnormal walk identification for systems using gait patterns. *Biomed Res*, pp 112–117.
27. Studenski S, Perera S, Patel K, Rosano C, Faulkner K, Inzitari M, Brach J, Chandler J, Cawthon P, Connor EB, Nevitt M, Visser M, Kritchevsky S, Badinelli S, Harris T, Newman AB, Cauley J, Ferrucci L, Guralnik J (2011) Gait speed and survival in older adults. *JAMA* 305(1):50–58
28. Tanigawa A, Morino S, Aoyama T, Takahashi M (2018) Gait analysis of pregnant patients with lumbopelvic pain using inertial sensor. *Gait Posture* 65:176–181
29. Tobias S, Sandro W, Ingrid B, Garlos MT (2020) Analyzing gait symmetry with automatically synchronized wearable sensors in daily life. *Microprocess Microsyst Vol* 77.
30. Xiaoli M, Haoyong Y, Ming Po Th (2013) Gait phase detection in able-bodied subjects and dementia patients. In: *proceedings of the 35th annual international conference on the IEEE Engineering in Medicine and Biology Society (EMBC)*, pp 4907–4910.
31. Xiaoyu X, Maxim AB, William JK, Bruce D (2011) Robust hierarchical system for classification of complex human mobility characteristics in the presence of neurological disorders. In: *Proceedings of International Conference on Body Sensor Network*, pp. 65–70.

32. Xingchen W, Danijela R-D, Matthias S, Axel G (2017) Gait assessment system based on novel gait variability measures. In: Proceedings of International Conference on Rehabilitation Robotics (ICORR), pp 467-472.
33. Zhiqiang Z, Charence W, Benny L, Guang-Zhong Y (2015) A simple and robust stride length estimation method using foot-mounted micro gyroscopes. In: Proceedings of International Conference of Body Sensor Network.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.