An On-Node Processing Approach for Anomaly Detection in Gait

Guglielmo Cola, Marco Avvenuti, Alessio Vecchio, Guang-Zhong Yang, and Benny Lo

***Abstract*—A novel method is proposed for capturing deviation in gait using a wearable accelerometer. Previous research has outlined the importance of gait analysis to assess frailty and fall risk in elderly patients. Several solutions, based on wearable sensors, have been proposed to assist geriatricians in mobility assessment tests, such as the Timed Up-and-Go test. However, these methods can be applied only to supervised scenarios and do not allow continuous and unobtrusive monitoring of gait. The method we propose is designed to achieve continuous monitoring of gait in a completely unsupervised fashion, requiring the use of a single waist-mounted accelerometer. The user's gait patterns are automatically learned using specific acceleration-based features, while anomaly detection is used to capture subtle changes in the way the user walks. All the required processing can be executed in real-time on the wearable device. The method was evaluated with 30 volunteers, who simulated a knee flexion impairment. On average, our method obtained $\sim 84\%$ accuracy in the recognition of abnormal gait segments lasting ~ 5 s. Prompt detection of gait anomalies could enable early intervention and prevent falls.**

***Index Terms*—Activity Monitoring, Anomaly Detection, Fall-risk assessment, Gait Analysis, Wearable sensors.**

I. Introduction

Gait changes, such as reduced stability or speed variations, are often used as early indicators of cognitive impairment [1]–[3]. In addition, reduced gait ability plays an important role in fall risk assessment [4]–[6]. These findings have brought increasing interest into the design of systems and algorithms for gait analysis. Several methods have been proposed to estimate the most relevant gait parameters and to assist geriatricians in the assessment of gait ability [7]–[10]. However, most of these solutions can only be applied to short walks or to specific tests such as the Timed Up-and-Go (TUG) test [8]. Though these systems offer significant help in making the assessments less subjective, they do not allow continuous monitoring of gait during daily activities.

One common approach to continuously monitor gait consists in instrumenting the environment with vision-based systems and/or passive infrared motion sensors [11]–[13]. These systems achieve high precision in capturing important gait parameters, such as stride duration and walking speed. On the other hand, the major drawbacks are the prohibitively expensive set-up costs, some privacy concerns, and the limitation of laboratory based assessment. These issues have hindered the widespread use of these approaches.

Manuscript received XXXXX XX, 2015; revised XXXXXX XX, XXXX.

G. Cola, M. Avvenuti and A. Vecchio are with the Dip. di Ingegneria dell'Informazione, University of Pisa, Pisa, Italy (e-mail: {g.cola, m.avvenuti, a.vecchio}@iet.unipi.it).

G.Z. Yang and B. Lo are with the Hamlyn Centre, Imperial College London, London, UK (e-mail: {g.z.yang, benny.lo}@imperial.ac.uk).

Another approach is the one based on wearable devices equipped with inertial sensors like accelerometers and gyroscopes [14]–[18]. Wearable devices enable continuous monitoring of users regardless of the environment, and offer fast system set up at relatively low costs [19], [20]. Nevertheless, wearable systems are often affected by usability issues, which make their adoption impractical for continuous monitoring. A relevant example is the number of sensors that the user is required to wear at different body positions: wearable systems that aim to precisely estimate gait parameters usually require the use of two or more sensors [21]. Another issue is caused by the limited computational resources of miniaturized devices. Most of the algorithms used for analyzing raw sensor signals cannot be executed in real-time on the wearable device. Hence, the user must always be in the transmission range of some external device or server, to which the collected samples are streamed via radio.

This paper presents a method to continuously monitor and capture changes in the user's gait patterns. Tri-axial acceleration samples are collected and analyzed using a single wearable device placed at the waist. Instead of directly estimating all the relevant gait parameters, which may be impractical using one accelerometer, our method aims to detect deviation in the gait patterns. To this purpose, eleven acceleration-based features are extracted and provided as inputs to an anomaly detection algorithm. The anomaly detection algorithm can be trained without any supervision, and is designed to automatically learn the user's gait patterns during the first few days of use. After the gait patterns are learned, the algorithm is able to detect if previously unseen patterns (anomalies) occur. The on-node processing algorithms that compose our method were designed with low complexity and computational requirements for real-time analysis. Hence, these algorithms can be executed on the wearable device without the need or support of another device or a backend server. Early detection of subtle gait changes could prompt geriatricians to introduce preemptive measures to reduce the risk of falling and prevent future falls. In addition, gait changes could be used to monitor degenerative diseases.

The rest of the paper is organized as follows. Section II summarizes the state of art related to the use of wearable inertial sensors for fall risk assessment and abnormal gait detection. The major contributions of this work are also highlighted. The proposed method is described in detail in Section III, while the experimental setting and the evaluation procedure are described in Section IV. Results are then presented and discussed in Section V. The energy consumption on the wearable device is evaluated in Section VI, using

different implementation strategies. Finally, conclusions and future work are presented in Section VII.

II. Related work and contributions

The use of inertial sensors to help clinicians to assess the risk of fall has been widely studied [9], [22]–[24]. These systems generally derive a model to classify users as fallers or non-fallers, on the basis of some relevant gait parameters. Gait parameters are estimated from the data collected while users are walking or performing some specific and predefined activity, such as, for example, the Timed Up-and-Go (TUG) test [8].

An extensive review related to this area of research is found in [25], where 40 systems based on wearable accelerometers and/or gyroscopes have been compared. According to this study, the use of accelerometers is more common compared to gyroscopes: accelerometers have been the only inertial sensor in 70% of the reviewed articles. The most common body position for attaching the sensor is at the subject's lower back, which is an approximation of the center of mass and is considered an acceptable location for long-term use [26]. Typically, the evaluation of fall risk predictions has been performed using clinical assessment tools or past fall history. Both of these approaches have relevant flaws according to [25]. The clinical assessment, which is observational-based, may include incorrect evaluations, thus leading to an incorrect validation of the fall risk classifier. On the other hand, an evaluation based on patients' fall history may be affected by inaccurate reports or by gait changes occurred in the patient because of the fall itself. The most robust evaluation method relies only on the falls occurring after fall-risk assessment. This approach, however, has been used only in 15% of the reviewed studies.

A framework to detect abnormal gait is presented in [27]. This system is characterized by the use of an ear worn sensor and does not rely on predefined gait parameters. Instead, multi-resolution wavelet analysis with margin feature selection has been proposed to distinguish normal walking from impaired walking. Different experiments have been performed using a treadmill: mild impairment has been simulated with a lower limb involvement (tubigrip), while moderate impairment has been simulated using a knee brace or abdominal brace system. The k-Nearest Neighbors algorithm has been used for classification, using both normal and impaired gait segments for training the classifier. Results from the ten subjects study have shown that normal walking and mild impairment classes tend to merge, while a clear separation is achieved between the moderate trunk and lower limb impairment classes.

The idea of continuous monitoring of gait for detecting abnormal changes has been suggested in [28]. Dynamic Time Warping has been used to obtain two indexes related to stability and symmetry of gait. The proposed system requires to stream the samples to a base station for off-line analysis. Moreover, no validation of the method on abnormal gait detection has been supplied.

The main characteristics of our method and the major contributions with respect to the the state-of-the-art techniques mentioned above can be summarized as follows:

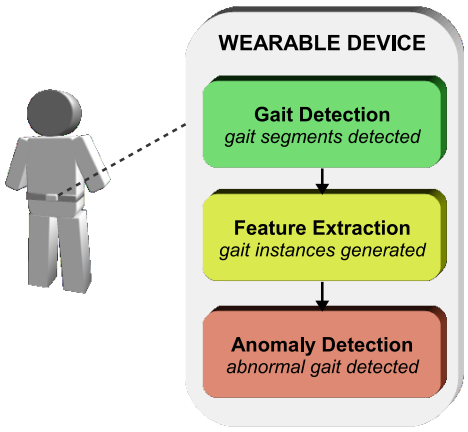


Fig. 1. Schematic view of the proposed method. The device (Shimmer 2r) is worn at the user's lower back. All the processing is executed on-node. Acceleration segments containing gait patterns are automatically detected by the gait detection algorithm. Feature extraction is applied to each gait segment, in order to obtain a vector with acceleration-based features (gait instance). Finally, the anomaly detection algorithm classifies each gait instance as normal or abnormal.

- A method is proposed to continuously monitor deviation in gait using a single wearable accelerometer. Instead of predicting fall risk, the proposed approach detects subtle changes in the way the user walks. Early detection of gait changes may be used to introduce preemptive interventions to prevent falls, and/or to monitor degenerative diseases.
- The method consists of three algorithms: gait detection, feature extraction, and anomaly detection. These algorithms were designed for miniaturized sensor devices with low computational resources. Hence, they can be executed directly on the wearable device, eliminating the need of any gateway or server for post processing or detection.
- The anomaly detection algorithm learns the user's gait patterns and detects if an anomaly is emerging. The algorithm is able to learn gait patterns in an *unsupervised* fashion and no prior training is required with abnormal gait examples. The classification accuracy was tested by off-line analysis to ensure repeatable evaluation. As a corollary contribution, the efficacy of the proposed feature set was evaluated also with supervised classification, where both normal and abnormal gait patterns were used to train the classifier.
- The method was implemented entirely on the wearable device to prove that real-time detection of anomalies can be performed on-node. The energy consumption was then estimated: a comparison was made between performing all the processing on-node and transmitting the raw acceleration samples to an external device. Results demonstrated that on-node processing is key to preserve battery life.

III. Proposed Method

A schematic description of the proposed method is given in Figure 1. For capturing acceleration reliably and continuously during daily activities, the wearable device is equipped with a

TABLE I
SHIMMER 2R: DEVICE CHARACTERISTICS

Subsystem	Model	Characteristics	Current consumption
Microcontroller	TI MSP430F1611	• 16 bit, up to 8 MHz • 10 Kbyte RAM, 48 Kbyte Flash • 2 DAC outputs, 8 12 bit A/D inputs	• High load: 5 mA • Idle: 0.02 mA
Communication	CC2420	• IEEE 802.15.4 compliant • Max data rate: 250 kbit/s	• Active (tx/rx): $\sim$ 20 mA • Idle: 426 µA
Sensing	Freescale MMA7361 Accelerometer	• Three-axial • Range: $\pm$ 1.5 / 6 g	• Active mode: 400 µA • Sleep Mode: 3 µA
Storage	Micro SD card	• Up to 2 GB	• Peak: 47 mA • Typical: 3 mA
Battery	Li-ion	• Capacity: 450 mA h • Voltage: 3.7 V • Regulator output: 3.0 V	

tri-axial accelerometer and it is fixed with a belt at the subject's lower back (i.e. near the person's center of mass). Gait patterns are detected in real-time by a gait detection algorithm. Whenever a new segment of acceleration data containing gait patterns (*gait segment*) is detected, a feature extraction algorithm is executed directly on the wearable device. The result of feature extraction forms a *gait instance*, which is a vector with eleven features. The gait instance is then classified as normal or abnormal, using an anomaly detection algorithm based on k-Nearest Neighbors analysis. Anomaly detection is designed as an unsupervised method, where the user's typical gait patterns (normal class) are learned automatically from the gait instances collected during the first few days of use. The following subsections describe the wearable device and the algorithms that compose our method.

A. Wearable device

A Shimmer 2r device was used for the experiments presented in this paper. The Shimmer is equipped with a TI MSP430 microcontroller, a tri-axial accelerometer, and an IEEE 802.15.4 radio. The device also includes a Micro SD slot for up to 2 GB of data storage. The hardware characteristics of the device are summarized in Table I. The size of the device is 53x32x15 mm, weight is 25 g. Additional details about the Shimmer 2r architecture can be found in [29].

B. Gait detection

To the purpose of real-time gait detection, a modified version of the walking recognition algorithm proposed in [30] was used. This algorithm was specifically designed for miniaturized wearable devices with limited resources. Tri-axial acceleration is sampled continuously at 51.2 Hz. Gait is detected by analyzing the peaks in the acceleration magnitude, which are generated by the ground reaction force when the foot hits the ground. A new gait segment is detected when 8 consecutive steps are found. The algorithm ensures that only homogeneous gait segments are detected: a filter based on standard deviation is applied using the duration of even and odd steps as inputs. In addition, the first two and the last two steps in a sequence of consecutive steps are discarded, as they tend to be highly irregular.

TABLE II
SELECTED FEATURES

AC_C1	AC_DP2	AAV vertical	AAV horizontal
Duration	Mean	Median	P2P
RMS	Standard Deviation		ZCR

C. Preprocessing

To reduce noise, the acceleration signal of each gait segment is low-pass filtered at 20 Hz using a second-order Butterworth filter. After, three values are found for each tri-axial acceleration sample: (i) the Euclidean norm of the acceleration vector (*acceleration magnitude*), (ii) the acceleration along the direction of gravity (*vertical acceleration*), (iii) the Euclidean norm of the acceleration vector on the horizontal plane (*horizontal acceleration magnitude*). Vertical acceleration and horizontal acceleration magnitude are calculated as indicated in [31].

D. Feature extraction

The features used are listed in Table II. Feature selection was performed by means of a greedy heuristic approach, starting from a set of 43 features. The optimized metric was the average classification accuracy obtained by the anomaly detection algorithm.

Mean, *median*, *Peak-to-Peak amplitude (P2P)*, *RMS*, *standard deviation*, and *Zero Crossing Rate (ZCR)* are statistical measures which have been widely used for activity recognition purposes [32]. *Duration* is the duration of the gait segment.

The *Average absolute Acceleration Variation (AAV)* has been previously proposed to improve the accuracy of fall detection systems [33], [34]. It is found according to the following equation:

$$AAV = \sum_{i=1}^{N-1} \frac{|x_{i+1} - x_i|}{N},$$

where N is the number of samples in the segment; x_i is the i-th sample in the segment.

AC_C1 is the autocorrelation coefficient at the first dominant period, while *AC_DP2* is the second dominant period of autocorrelation. Unbiased autocorrelation coefficients are

calculated as follows:

$$AC_k = \frac{1}{N-k} \sum_{i=1}^{N-k} x_i * x_{i+k},$$

where AC_k is the k-th unbiased autocorrelation coefficient; N is the number of acceleration samples in the gait segment; x_i is the i-th sample minus the average of the samples in the gait segment. After the autocorrelation coefficients have been found, a peak detection algorithm is used to find the dominant periods in the autocorrelation signal.

AC_C1, AC_DP2, median, and RMS are calculated on the acceleration magnitude of the gait segment samples. Mean and standard deviation are calculated on the horizontal acceleration. P2P and ZCR are calculated on the vertical acceleration. AAV is calculated on both vertical and horizontal acceleration. In total, a *gait instance* is a vector with eleven features. Since vertical acceleration and horizontal acceleration are estimated using an orientation-independent technique, the user can wear the device without caring about its orientation.

Hereafter we use the term *gait instance* or simply *instance* to refer to the vector of features obtained from a gait segment through the above described feature extraction process.

E. Anomaly detection

The proposed anomaly detection algorithm is a binary classifier based on k-Nearest Neighbors (k-NN) analysis. Gait instances are either classified as *abnormal* (positive) or *normal* (negative). In this study, the term *normal* denotes the user's typical gait patterns when he/she starts wearing the device. These patterns are learned by the system in a totally unsupervised fashion: a personalized training set is created with the gait segments detected in the first few days of use. This unsupervised approach for generating the training set relies on the implicit assumption made by unsupervised anomaly detection systems: normal instances are far more frequent than anomalies [35]. In our application scenario, the user starts wearing the device after that a clinical assessment has been performed by a geriatrician. If the user is not currently recovering from an injury, it is reasonable to expect that his/her gait ability will not change significantly in a few days following the assessment. Hence, the instances obtained in the training period represent a reliable picture of the gait ability assessed by the geriatrician.

At the end of the training phase, the instances in the training set are used as model to identify anomalies. Firstly, for each gait instance g_i in the training set, an anomaly score AS_i is found as

$$AS_i = \sum_{j=1}^{k} distance\ (g_i, n_j),$$

where n_j is the j-th nearest neighbor of g_i in the training set. The $distance$ between two gait instances is found using the Euclidean distance. Let f_{ij} be the j-th feature of gait instance g_i, the distance between two instances g_x and g_y is defined as

$$distance\ (g_x, g_y) = \sqrt{\sum_{j=1}^{11} (f_{xj} - f_{yj})^2}.$$

TABLE III
GAIT EXPERIMENTS CHARACTERISTICS (AVG. ± STD.DEV.)

Experiment	Duration [s]	Speed [m/s]	Stride [s]	Steps
Normal	106.9 ± 14.3	1.33 ± 0.19	1.12 ± 0.08	190 ± 21
Mild	113.9 ± 16.5	1.26 ± 0.19	1.15 ± 0.08	198 ± 23
Severe	117.2 ± 17.8	1.22 ± 0.19	1.16 ± 0.09	201 ± 24
Mild Both	117.7 ± 18.3	1.22 ± 0.20	1.16 ± 0.09	203 ± 24

Fig. 2. Each volunteer simulated three different types of impaired gait: mild knee condition, severe knee condition, and mild condition in both knees. Mild condition was simulated using a Neo-G thigh support strap wrapped around one knee. Severe condition was simulated wrapping two straps at the same knee. Finally, to simulate mild problems in both knees, a strap was wrapped around each knee.

Before this calculation, the features are normalized using the minimum and maximum values in the training set. This ensures that different features contribute to the distance with equal importance.

Afterwards, a threshold TH is defined to discriminate between normal and abnormal instances. Given a *coverage index* $c \in [0, 1]$, TH is chosen such that the proportion of instances in the training set having an anomaly score smaller than TH is equal to c. Whenever a new gait instance g_{new} is found, its corresponding anomaly score AS_{new} is calculated, and the new instance is classified as abnormal if and only if

$$AS_{new} > TH.$$

The behavior of the anomaly detection algorithm is determined by the choice of the parameters k (number of neighbors) and c (coverage index). In particular, c represents the specificity obtained by the classifier on the training set, and is used to define the threshold TH. For example, if c = 0.5, the threshold TH is set as the median among the AS values of the training instances, so that 50% of the instances in the training set are classified as normal. A higher choice for c is likely to produce a system less prone to false alarms (high specificity). On the other hand, increasing c may decrease the sensitivity of the classifier, leading to detect several abnormal instances as normal. The method used to select k and c is discussed in Section IV-C.

IV. Experimental setting and evaluation

In order to validate the proposed method, a gait dataset was created with the help of volunteers. Experiments were performed both in controlled and uncontrolled environment. During data collection, the volunteers carried a Shimmer device at their lower back using a belt. Acceleration samples, collected with 51.2 Hz frequency, were saved into the device's Micro SD card to ensure repeatable evaluation of the method. Then, an evaluation procedure was applied to the dataset to verify the performance of the method.

A. Controlled experiments

30 healthy volunteers (18 males, 12 females, age 32.9 $\pm$ 12.2, height 171.1 $\pm$ 9.6 cm, weight 69.8 $\pm$ 15.2 kg) performed gait-related experiments in a corridor measuring $\sim$ 35 m. The following four types of gait were collected: normal, mild knee condition, severe knee condition, and mild condition in both knees. Each experiment was repeated four times. In order to simulate gait changes due to limitation in knee flexion, the volunteers were asked to wrap *Neo-G thigh support* straps around their knees as shown in Figure 2. In the *mild* condition simulation, one strap was wrapped around one knee. In the *severe* condition simulation, instead, two straps were wrapped around the same knee to emulate a severe impairment. Finally, to simulate mild problems with *both* knees, one strap was wrapped around each knee. In the following we refer to these experiments using the terms *normal*, *mild*, *severe*, and *mild both*.

In total 2673 gait segments were collected. Additional statistics about these gait experiments are presented in Table III. For each experiment, the average and standard deviation values of the duration, speed, stride time, and number of steps are shown. It is interesting to note that the average speed variation between normal gait and mild knee condition gait was only 0.07 m/s. Such a difference cannot be easily detected relying only on a single accelerometer [36].

B. Uncontrolled experiments

Four users who participated in the controlled experiments were also involved in *uncontrolled experiments*, where about 48 hours of acceleration data was collected in total. During these experiments, the users performed their habitual activities and were allowed to change their footwear during the day. The collected traces include different activities, such as climbing stairs and walking outdoors on uneven terrain.

C. Evaluation procedure

The acceleration traces were transferred onto a PC for off-line analysis. The first two parts of the method, gait detection and feature extraction, were applied to the raw acceleration signals to obtain gait instances. To evaluate the *controlled experiments*, the following steps were taken for each user u: (i) the anomaly detection algorithm was trained using all the *normal* gait instances belonging to the user u except one instance x, which was left out for validation; (ii) specificity was estimated using the normal instance x; (iii) sensitivity was estimated with respect to each simulated condition (*mild*, *severe*, and *mild both*). This procedure was repeated N times, where N is the number of normal instances belonging to u: each time, a different normal instance was left out for validation (leave-one-out cross-validation). The overall specificity and sensitivity values for the user u were calculated averaging the results obtained for each different training set.

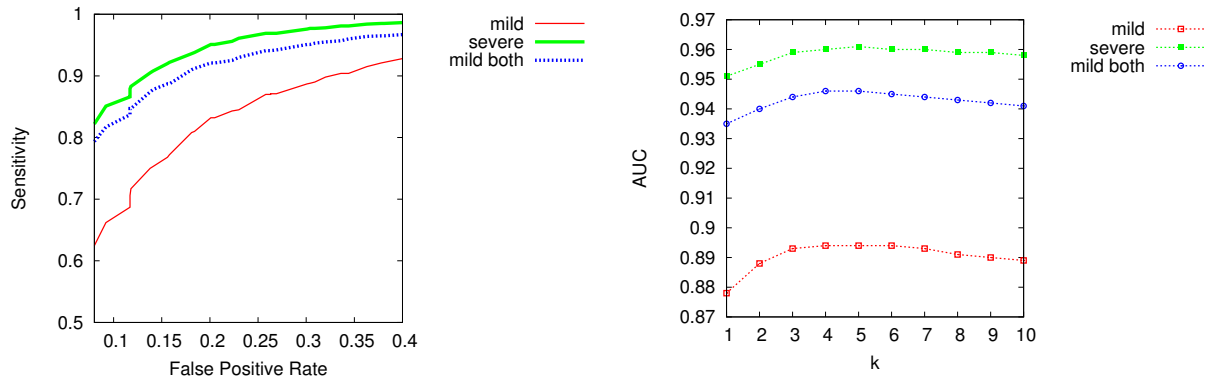
A similar analysis was carried out on the uncontrolled experiments data. Ten-fold cross-validation was used to train the classifier with a user's normal instances: specificity was estimated on the left-out normal instances, while sensitivity was estimated on the simulated conditions performed by the same user in the controlled experiments. In order to limit the growth of the training set, a filter based on autocorrelation was used. For each gait instance, the highest autocorrelation peak (AC_HP) was found: only the 100 most regular gait instances detected during the training phase were actually included in the training set. Also, the lowest AC_HP value in the training set AC_HP$_{min}$ was used to reduce the number of validation instances: gait instances having an AC_HP value lower than AC_HP$_{min}$ were discarded. These highly irregular gait segments, indeed, may have been produced while climbing stairs or on highly uneven terrain and are unlikely to offer a proper representation of the user's gait patterns.

As a corollary contribution, we verified the efficacy of the proposed feature set in representing gait patterns with supervised classification. In this case, some standard classifiers were used, such as Neural Networks and Random Forests. The classifiers were trained using the user's typical gait as well as instances belonging to the anomalies to be detected. The performance of the classifiers was evaluated using ten-fold cross-validation on the controlled experiments. This analysis allowed a partial comparison with the system presented in [27].

D. Selection of parameters

Before performing the procedures described above, the parameters of the anomaly detection algorithm, k and c, were set using the controlled experiments data. The value of k was selected using Receiver Operating Characteristic (ROC) plots and Area Under Curve (AUC) analysis. Then, c was selected in order to maximize the following performance index: the average between the mean and the worst case accuracy. The worst case accuracy was included with the aim to have a more conservative choice of c.

Figure 3a shows the ROC plot obtained by setting k = 3 and averaging the results obtained among different users. As expected, the algorithm is more accurate in detecting the *severe* and *mild both* conditions with respect to the anomaly produced wrapping just one strap around a knee (*mild* condition). ROC plots were generated for several k values (1-10), then the respective AUC values were calculated and used as the performance metric of the classifier (Figure 3b). The AUC value shows negligible variations in the [3,7] interval: k = 3 was selected as a trade-off between accuracy and processing requirements. According to that choice for k, the coverage index c maximizing the overall accuracy was 0.80.



(a) (b)

Fig. 3. (a) ROC plot of the anomaly detection algorithm for each gait anomaly, using k = 3; (b) AUC of the anomaly detection algorithm (based on k-NN sum) for different values of k and for each gait anomaly.

TABLE IV
CONTROLLED EXPERIMENTS AVERAGE AND WORST-CASE RESULTS

Class	**Anomalies (%)** average (worst)	**Accuracy (%)** average (worst)
Normal	23.0 (26.9)	N/A
Mild	84.5 (48.0)	80.8 (63.5)
Severe	96.1 (55.7)	86.5 (67.4)
Mild Both	93.1 (55.8)	85.1 (66.1)

TABLE V
UNCONTROLLED EXPERIMENTS AVERAGE AND WORST-CASE RESULTS PLUS COMPARISON WITH CONTROLLED EXPERIMENTS

Class	**Anomalies (%)**		**Accuracy (%)**	
	avg (worst)	**controlled**	**avg (worst)**	**controlled**
Normal	28.7 (35.4)	24.1	N/A	
Mild	87.8 (52.7)	86.1	79.6 (63.5)	81.0
Severe	99.3 (97.1)	99.0	85.3 (67.4)	87.4
Mild Both	100 (100)	100.0	85.7 (66.1)	87.9

V. RESULTS AND DISCUSSION

This section reports and discusses the results obtained in the controlled and uncontrolled experiments.

A. Results of controlled experiments

Table IV shows the results related to the controlled experiments, which were performed by 30 users in a corridor. The first row shows the rate of normal instances that were mistakenly detected as anomalies. The remaining rows of Table IV show the results related to the simulated impairments: the second column shows the rate of abnormal instances which were correctly detected (i.e. the sensitivity of the method), while the third column shows the accuracy obtained for each impairment. The accuracy was calculated as the average between specificity and sensitivity.

On average, 23% of the gait instances are classified as abnormal when the user is walking normally. Instead, when the user's gait pattern is abnormal, the rate of detected anomalies raises up to 84.5% in the mild condition, 96.1% in the severe condition, and 93.1% in the mild both condition. These results meet the requirements of the novel application that we are proposing, since the aim is to detect long-term changes in gait patterns. Indeed, a predefined threshold on the detection rate of anomalies could be used to classify a set of gait instances as normal or abnormal, for example the instances collected during a day. Only if an abnormal condition is detected for multiple consecutive days an alarm should be raised, alerting the user and his/her carer to take preemptive actions. For example, according to the results presented in Table IV, in the controlled experiments it is possible to successfully distinguish the whole set of normal instances belonging to a single user from the set of abnormal instances by using a threshold equal to 45%.

B. Results of uncontrolled experiments

Table V presents the results obtained in the uncontrolled environments, where four users collected acceleration data without supervision and during their habitual activities. For each condition, it is shown the average rate of anomalies detected and the average classification accuracy. Table V also shows the average result obtained by the same four users in the controlled experiments performed in the corridor. It is thus possible to compare the results obtained in controlled and uncontrolled environments by the same group of volunteers.

As expected, due to the higher variability of the gait pattern during unsupervised activities, the classifier detected a higher rate of anomalies during normal walking (+4.6%). Average sensitivity, instead, was not affected significantly by the use of a larger training set. In terms of average accuracy, there was a 1.4% reduction in the mild condition, a 2.1% reduction in the severe condition, and a 2.2% reduction in the mild both condition experiment. However, a threshold equal to 45% can be still used to successfully discriminate a user's set of normal gait instances from each set of abnormal instances.

TABLE VI
SUPERVISED CLASSIFICATION ACCURACY (%)

Classifier	Mild	Severe	Mild Both
Neural Network	91.2	97.2	96.5
Random Forest	88.6	93.8	94.2
Rotation Forest	91.1	96.7	94.9
M. Logistic	89.3	95.7	93.7
3-NN	89.9	96.6	94.6

C. Discussion

In all of the experiments described above, the anomaly detection classifier was trained using only the user's typical gait. This is key for the application we are proposing, since in practice abnormal gait examples (anomalies) will not be available when the user starts wearing the device. In the proposed approach, unsupervised and user-specific training of the classifier can be obtained using the gait instances detected in the first few days of use.

In these experiments, the number of instances used for the training phase is not dynamically set. In a real deployment, the system may terminate the training phase automatically when the classifier's specificity proves to be relatively stable, despite new instances being detected and added to the training set.

Long term users complained that wearing a Shimmer at the lower back for long periods is uncomfortable, mostly because of its thickness. To overcome this issue, we envisage that devices that can be attached to the skin as patches will be commonly available to allow continuous and unobtrusive monitoring of gait [37]–[40]. In addition, we plan to investigate the use of the proposed technique on an ear-worn sensor.

To further verify the efficacy of the feature set, we computed the classification accuracy also in a supervised fashion. Five standard classifiers were used, and the average accuracy among the users is shown in Table VI. Each column of the table indicates the average accuracy achieved in distinguishing *normal* gait from a specific gait anomaly. The classifiers were trained using normal gait instances as well as examples of the anomaly to be detected, thus achieving higher accuracy with respect to the unsupervised anomaly detection algorithm. As previously explained, this approach cannot be used in the proposed application, since anomalies are totally unknown when the user starts using the system.

Nevertheless, these results allow a partial comparison with the system presented in [27]. In that work, an ear-mounted sensor has been used to distinguish moderate knee impairment, simulated using a knee brace, from normal gait. The average accuracy is similar to the results presented in Table VI, but the method presented here offers two main advantages: (i) the use of a reduced number of features with low computational requirements, enabling on-node feature extraction, (ii) successful recognition of a mild condition, in which the reduction of knee flexion capability and average speed were minimal for most of the users in the dataset.

VI. ESTIMATION OF ENERGY CONSUMPTION

Being characterized by low computational requirements, the proposed method can be implemented on a wearable device like the Shimmer 2r (8 MHz CPU, 10 Kbyte RAM). This aspect is key for continuous monitoring applications, as it makes the wearable device totally independent from external devices. On the other hand, it must be verified that on-node processing is feasible and does not exhaust the device's battery. We thus estimated how energy consumption is affected if the method is partially or entirely distributed. More precisely, we evaluated four different implementation strategies: (i) the wearable device only performs acceleration sampling and streams data to a base station, (ii) the wearable device, beside acceleration sampling, also performs gait detection and transmits acceleration samples only during gait periods, (iii) the wearable device samples acceleration, performs gait detection, extracts features and sends the latter ones to the base station, (iv) all processing is executed on-node (sampling, gait detection, feature extraction, and classification). In practice, we identified two extreme strategies (complete off-loading of computation vs all computation on-node) and two hybrid strategies (with increasing amount of on-node computation).

The four implementation strategies can be built by combining the six tasks described in Table VII. Each strategy was first implemented and executed on a Shimmer node in order to verify that the wearable device is capable of performing the required tasks in real-time. After, it was estimated the power consumption determined by the strategy as well as the expected battery duration on the Shimmer.

A. Tasks

The first task, *Acceleration Sampling (AS)*, consists in sampling acceleration with 51.2 Hz frequency: raw ADC readings are saved into a ring buffer without further processing. According to our experiments, the power consumption generated by AS was $\sim$ 3.0 mW.

The second task, *Radio Streaming (RS)*, consists in transmitting the raw acceleration samples to a base station. RS was implemented using the CC2420 IEEE 802.15.4 radio module of the Shimmer. A simple communication protocol was defined to ensure reliability. The radio is switched on every 5 s to send the acceleration samples collected in the last 5 s interval. The receiver is supposed to be always listening, and must acknowledge the beginning as well as the end of each communication. When no retransmissions are required, the transmission of the last 5 s of samples is executed in $\sim$ 280 ms. The average power required was estimated using CC2420 specifications, and thus considering a 52.2 mW consumption while transmitting. Taking into account idle intervals, the average consumption was $\sim$ 2.9 mW.

The third task is *Gait Detection (GD)*, which requires $\sim$ 0.6 ms of processing for each acceleration sample. The average power consumption was estimated considering a consumption equal to 15.0 mW while processing (worst case indicated in the Shimmer 2r manual). The result was $\sim$ 0.5 mW.

The fourth and fifth tasks are *Feature extraction (FE)* and *Feature Streaming (FS)* respectively. FE, applied to a gait

TABLE VII
POWER REQUIRED FOR EACH TASK

Task	Description	Consumption
Acceleration Sampling (AS)	Sampling acceleration with 51.2 Hz freq.	3.0 mW
Radio Streaming (RS)	Streaming samples via radio (802.15.4)	2.9 mW
Gait Detection (GD)	Real-time detection of gait segments	0.5 mW
Feature Extraction (FE)	Feature extraction applied to a gait segment	2.7 mW
Feature Streaming (FS)	Features are sent via radio to the base	0.5 mW
Classification (CL)	Classification of a gait segment (anomaly detection)	$<$ 0.1 mW

TABLE VIII
POWER CONSUMPTION ACCORDING TO DIFFERENT IMPLEMENTATION STRATEGIES

Implementation strategy	Idle user		Walking user		Average consumption	Battery duration (days)
	Tasks	Power	Tasks	Power		
Streaming samples continuously	AS+RS	5.9 mW	AS+RS	5.9 mW	5.90 mW	9.5
Streaming only during gait	AS+GD	3.5 mW	AS+GD+RS	6.4 mW	3.74 mW	15.0
On-node feature extraction	AS+GD	3.5 mW	AS+GD+FE+FS	6.7 mW	3.77 mW	14.9
Everything is done on-node	AS+GD	3.5 mW	AS+GD+FE+CL	6.2 mW	3.73 mW	15.1

segment, is executed in about 18% of the duration of the gait segment itself. The average consumption, in this case, was calculated with respect to a gait segment lasting 5 s. The result was an average consumption of $\sim$ 2.7 mW. FS, instead, consists in transmitting the calculated features via radio (11 values, 22 bytes). The result was $\sim$ 0.5 mW.

The last task is *Classification (CL)*. For a training set made of 100 gait instances, the time required to classify a gait segment was $\sim$ 13 ms. The impact on overall consumption was negligible.

B. Energy consumption of the four implementation strategies

The evaluation of the different implementation strategies is described in Table VIII. The second column reports which tasks are executed when the user is idle (not walking) according to each strategy. The third column indicates which tasks are executed during walking activity. In both situations (idle and walking), the power required was found by summing up the consumptions due to each task. The fourth column reports the average consumption according to the assumption that the users walks two hours per day. Finally, the last column reports the expected battery duration when a 450 mA h battery is used (the one available on Shimmer 2r).

In the first implementation strategy (first row of Table VIII), the device does not perform any processing on sampled data: samples are streamed via radio to an external device (AS+RS tasks). Consequently, the average consumption is $\sim$ 5.9 mW regardless of the user's activity. In all of the remaining strategies GD is executed on-node. This allows to identify idle intervals, which do not require further processing: when the user is not walking the consumption is $\sim$ 3.5 mW. During gait, instead, samples can be alternatively sent to a base (RS), partially analyzed on-node (FE+FS), or completely analyzed on-node (FE+CL). If gait samples are transmitted via radio, the power required is $\sim$ 6.4 mW (AS+GD+RS). If feature extraction is done on-node and only the features are sent via radio, the consumption is $\sim$ 6.7 mW (AS+GD+FE+FS). Finally, if everything is done on-node, there are no radio transmissions and the power required is $\sim$ 6.2 mW (AS+GD+FE+CL).

These results confirmed that executing the whole method on the Shimmer device is feasible in terms of computational power. At the same time the battery duration, when the whole method is executed on-board, is much better than the solution where computing is executed completely on the base station. Hybrid solutions have approximately the same energy requirements with respect to the completely on-board solution, as in these cases the battery duration is always $\sim$ 15 days. However, the completely on-board solution is much more flexible than the others and it is thus the one to be preferred: it does not require to be in the communication range of the base station and the device can operate autonomously. It should be also noted that we considered a basic radio protocol, in which the receiver is always listening and does not establish a secure connection with the wearable device. Therefore, in practice the consumption due to radio transmissions is likely to be higher than the estimation we have just presented.

VII. CONCLUSIONS

A novel method to detect deviation in gait using a wearable accelerometer has been presented. The proposed approach is based on a specific set of acceleration features and unsupervised anomaly detection. The results of the experiments with 30 volunteers have demonstrated the robustness and accuracy of the proposed method. An evaluation in uncontrolled environments was performed with four volunteers, and the results obtained in the respective controlled experiments were confirmed. It was proved that all the required processing can be executed on the wearable sensor node with limited resources, enabling low-cost and continuous monitoring of gait. Early

detection of gait changes could be used to reduce the risk of falls, capture early signs of deterioration, and monitor the progression of degenerative diseases.

In future work, we plan to investigate the use of clustering to improve the performance of the anomaly detection process: firstly, clusters may be used to merge similar instances and make classification more efficient; secondly, clusters' density may be used as an additional indicator for detecting changes in gait and suggest when a new clinical assessment would be advisable.

REFERENCES

[1] R. Camicioli, D. Howieson, B. Oken, G. Sexton, and J. Kaye, "Motor slowing precedes cognitive impairment in the oldest old," *Neurology*, vol. 50, no. 5, pp. 1496–1498, 1998.

[2] J. Verghese, C. Wang, R. B. Lipton, R. Holtzer, and X. Xue, "Quantitative gait dysfunction and risk of cognitive decline and dementia," *Journal of Neurology, Neurosurgery & Psychiatry*, vol. 78, no. 9, pp. 929–935, 2007.

[3] T. Buracchio, H. Dodge, D. Howieson, D. Wasserman, and J. Kaye, "The trajectory of gait speed preceding mild cognitive impairment," *Archives of Neurology*, vol. 67, no. 8, pp. 980–986, 2010.

[4] J. Verghese, R. Holtzer, R. B. Lipton, and C. Wang, "Quantitative gait markers and incident fall risk in older adults," *The Journals of Gerontology Series A: Biological Sciences and Medical Sciences*, vol. 64A, no. 8, pp. 896–901, 2009.

[5] J. M. VanSwearingen, K. A. Paschal, P. Bonino, and J.-F. Yang, "The modified gait abnormality rating scale for recognizing the risk of recurrent falls in community-dwelling elderly adults," *Physical Therapy*, vol. 76, no. 9, pp. 994–1002, 1996.

[6] J. M. Hausdorff, D. A. Rios, and H. K. Edelberg, "Gait variability and fall risk in community-living older adults: A 1-year prospective study," *Archives of Physical Medicine and Rehabilitation*, vol. 82, no. 8, pp. 1050 – 1056, 2001.

[7] D. Giansanti, G. Maccioni, S. Cesinaro, F. Benvenuti, and V. Macellari, "Assessment of fall-risk by means of a neural network based on parameters assessed by a wearable device during posturography," *Med Eng and Physics*, vol. 30, pp. 367–372, 2008.

[8] S. Mellone, C. Tacconi, and L. Chiari, "Validity of a smartphone-based instrumented timed up and go," *Gait & Posture*, vol. 36, no. 1, pp. 163 – 165, 2012.

[9] R. Senden, H. Savelberg, B. Grimm, I. Heyligers, and K. Meijer, "Accelerometry-based gait analysis, an additional objective approach to screen subjects at risk for falling," *Gait & Posture*, vol. 36, no. 2, pp. 296 – 300, 2012.

[10] K. Sheehan, B. Greene, C. Cunningham, L. Crosby, and R. Kenny, "Early identification of declining balance in higher functioning older adults, an inertial sensor based method," *Gait & Posture*, vol. 39, no. 4, pp. 1034 – 1039, 2014.

[11] E. Stone and M. Skubic, "Unobtrusive, continuous, in-home gait measurement using the microsoft kinect," *Biomedical Engineering, IEEE Transactions on*, vol. 60, no. 10, pp. 2925–2932, Oct 2013.

[12] L. Wang, T. Tan, W. Hu, and H. Ning, "Automatic gait recognition based on statistical shape analysis," *Image Processing, IEEE Transactions on*, vol. 12, no. 9, pp. 1120–1131, Sept 2003.

[13] S. Hagler, D. Austin, T. Hayes, J. Kaye, and M. Pavel, "Unobtrusive and ubiquitous in-home monitoring: A methodology for continuous assessment of gait velocity in elders," *Biomedical Engineering, IEEE Transactions on*, vol. 57, no. 4, pp. 813–820, April 2010.

[14] K. M. Culhane, M. OConnor, D. Lyons, and G. M. Lyons, "Accelerometers in rehabilitation medicine for older adults," *Age and Ageing*, vol. 34, no. 6, pp. 556–560, 2005.

[15] M. J. Mathie, A. C. F. Coster, N. H. Lovell, and B. G. Celler, "Accelerometry: providing an integrated, practical method for long-term, ambulatory monitoring of human movement," *Physiological Measurement*, vol. 25, no. 2, p. R1, 2004.

[16] K. Aminian, B. Najafi, C. Bla, P.-F. Leyvraz, and P. Robert, "Spatio-temporal parameters of gait measured by an ambulatory system using miniature gyroscopes," *Journal of Biomechanics*, vol. 35, no. 5, pp. 689 – 699, 2002.

[17] K. Tong and M. H. Granat, "A practical gait analysis system using gyroscopes," *Medical Engineering & Physics*, vol. 21, no. 2, pp. 87 – 94, 1999.

[18] W. Zijlstra and A. L. Hof, "Assessment of spatio-temporal gait parameters from trunk accelerations during human walking," *Gait & Posture*, vol. 18, no. 2, pp. 1 – 10, 2003.

[19] S. Mukhopadhyay, "Wearable sensors for human activity monitoring: A review," *Sensors Journal, IEEE*, vol. 15, no. 3, pp. 1321–1330, 2015.

[20] T. Shany, S. Redmond, M. Narayanan, and N. Lovell, "Sensors-based wearable systems for monitoring of human movement and falls," *Sensors Journal, IEEE*, vol. 12, no. 3, pp. 658–670, March 2012.

[21] B. Caby, S. Kieffer, M. de Saint Hubert, G. Cremer, and B. Macq, "Feature extraction and selection for objective gait analysis and fall risk assessment by accelerometry," *BioMedical Engineering OnLine*, vol. 10, no. 1, p. 1, 2011.

[22] M. Gietzelt, G. Nemitz, K.-H. Wolf, H. Meyer Zu Schwabedissen, R. Haux, and M. Marschollek, "A clinical study to assess fall risk using a single waist accelerometer," *Inform Health Soc Care*, vol. 34, pp. 181–188, 2009.

[23] A. Dalton, H. Khalil, M. Busse, A. Rosser, R. van Deursen, and G. Laighin, "Analysis of gait and balance through a single triaxial accelerometer in presymptomatic and symptomatic huntington's disease," *Gait & Posture*, vol. 37, no. 1, pp. 49 – 54, 2013.

[24] M. Marschollek, A. Rehwald, K.-H. Wolf, M. Gietzelt, G. Nemitz, H. Meyer Zu Schwabedissen, and R. Haux, "Sensor-based fall risk assessment-an expert 'to go'," *Methods Inf Med*, vol. 50, pp. 420–426, 2011.

[25] J. Howcroft, J. Kofman, and E. Lemaire, "Review of fall risk assessment in geriatric populations using inertial sensors," *Journal of NeuroEngineering and Rehabilitation*, vol. 10, no. 1, p. 91, 2013.

[26] D. Giansanti, S. Morelli, G. Maccioni, and G. Constantini, "Toward the design of a wearable system for fall-risk detection in telerehabilitation," *Telemed e-Health*, vol. 15, pp. 296–299, 2009.

[27] L. Atallah, O. Aziz, B. Lo, and G.-Z. Yang, "Detecting walking gait impairment with an ear-worn sensor," in *Wearable and Implantable Body Sensor Networks (BSN), Sixth Int Workshop on*, June 2009, pp. 175–180.

[28] S. Jiang, B. Zhang, and D. Wei, "The elderly fall risk assessment and prediction based on gait analysis," in *Computer and Information Technology (CIT), 2011 IEEE 11th International Conference on*, Aug 2011, pp. 176–180.

[29] Shimmer manual (rev. 2Rk), "http://www.shimmersensing.com/support," 2015.

[30] G. Cola, A. Vecchio, and M. Avvenuti, "Improving the performance of fall detection systems through walk recognition," *Journal of Ambient Intelligence and Humanized Computing*, pp. 1–13, 2014.

[31] D. Mizell, "Using gravity to estimate accelerometer orientation," in *Proc IEEE Int Symp Wearable Comput*, 2003, pp. 252–253.

[32] M. Zhang and A. A. Sawchuk, "A feature selection-based framework for human activity recognition using wearable multimodal sensors," in *Proc of the 6th Int Conf on Body Area Networks*, ser. BodyNets '11, ICST, Brussels, Belgium, 2011, pp. 92–98.

[33] S. Abbate, M. Avvenuti, G. Cola, P. Corsini, J. V. Light, and A. Vecchio, "Recognition of false alarms in fall detection systems," in *Proc IEEE Int Workshop Consum eHealth Platf Serv Appl*, Las Vegas, NV, USA, Jan. 2011, pp. 538–543.

[34] S. Abbate, M. Avvenuti, F. Bonatesta, G. Cola, P. Corsini, and A. Vecchio, "A smartphone-based fall detection system," *Pervasive Mob Comput*, vol. 8, no. 6, pp. 883 – 899, 2012.

[35] V. Chandola, A. Banerjee, and V. Kumar, "Anomaly detection: A survey," *ACM Comput. Surv.*, vol. 41, no. 3, pp. 15:1–15:58, Jul. 2009.

[36] W.-S. Yeoh, I. Pek, Y.-H. Yong, X. Chen, and A. Waluyo, "Ambulatory monitoring of human posture and walking speed using wearable accelerometer sensors," in *Engineering in Medicine and Biology Society(EMBS), 30th Annu Int Conf of the IEEE*, Aug 2008, pp. 5184–5187.

[37] A. Chan, N. Selvaraj, N. Ferdosi, and R. Narasimhan, "Wireless patch sensor for remote monitoring of heart rate, respiration, activity, and falls," in *Engineering in Medicine and Biology Society (EMBC), 35th Annual International Conference of the IEEE*, July 2013, pp. 6115–6118.

[38] Meria IH1^{TM}, "https://www.metriaih1.com/," 2015.

[39] F. Martinez-Tabares, N. Gaviria-Gomez, and G. Castellanos-Dominguez, "Very long-term ecg monitoring patch with improved functionality and wearability," in *Engineering in Medicine and Biology Society (EMBC), 36th Annu Int Conf of the IEEE*, Aug 2014, pp. 5964–5967.

[40] H.-J. Yoo, J. Yoo, and L. Yan, "Wireless fabric patch sensors for wearable healthcare," in *Engineering in Medicine and Biology Society (EMBC), Annu Int Conf of the IEEE*, Aug 2010, pp. 5254–5257.