Automatic classification of gait patterns using a smart rollator and the BOSS model

Maribel Ojeda
Tecnológico de Monterrey - Campus Puebla
Puebla, México
a01097403@itesm.mx

Atia Cortés
Universitat Politècnica de Catalunya - BarcelonaTech
Barcelona, Catalunya, Spain
acortes@cs.upc.edu

Javier Béjar
Universitat Politècnica de Catalunya - BarcelonaTech
Barcelona, Catalunya, Spain
bejar@cs.upc.edu

Ulises Cortés
Universitat Politècnica de Catalunya - BarcelonaTech
Barcelona, Catalunya, Spain
ia@cs.upc.edu

ABSTRACT

Nowadays, the risk of falling in older adults is a major concern due to the severe consequences it brings to socio-economic and public health systems. Some pathologies cause mobility problems in the aged population, leading them to fall and, thus, reduce their autonomy. Other implications of ageing involve having different gait patterns and walking speed. In this paper, a non-invasive framework is proposed to study gait in elder people using data collected by a smart rollator, the *i*-Walker. The analysis presented in this article uses a feature extraction method and a spectral embedding to represent the information and Bayesian clustering for the knowledge discovery. The algorithm considers raw data from the *i*-Walker sensors along with the calculated walking speed of each individual, which has been already used in clinical studies to assess physical and cognitive status of older adults. The results obtained demonstrate that the proposed analysis has the potential to separate in clusters the people of the two groups of interest: young people and geriatric.

CCS CONCEPTS

• **Computing methodologies** → **Cluster analysis**; • **Applied computing** → *Health informatics*;

KEYWORDS

gait analysis, assistive technologies, machine learning, time series clustering

ACM Reference Format:
Maribel Ojeda, Atia Cortés, Javier Béjar, and Ulises Cortés. 2018. Automatic classification of gait patterns using a smart rollator and the BOSS model. In *PETRA '18: The 11th PErvasive Technologies Related to Assistive Environments Conference, June 26–29, 2018, Corfu, Greece.* ACM, New York, NY, USA, 7 pages. https://doi.org/10.1145/3197768.3201575

PETRA '18, June 26–29, 2018, Corfu, Greece

ACM ISBN 978-1-4503-6390-7/18/06…$15.00
https://doi.org/10.1145/3197768.3201575

1 INTRODUCTION

Older individuals often exhibit greater impairments in posture and gait. These structural and functional abnormalities were associated with changes in walk strategy and abnormal gait, eventually leading to an increased risk of falls. To improve the quality of life of affected people and avoid falls, assistive devices, as rollator, have been used to provide mobility support with the aim of helping to perform those compensation strategies [10].

According to the Center for Disease Control and Prevention [3], falls are the leading cause of mortality in this segment; they can also lead to injury, causing loss of independence and critical health problems. Today, many countries are facing a rapidly aging population. Worldwide, the number of this group was estimated to grow to more than two billion by 2050 [14]. The increased number of persons over 65 years will potentially lead to expensive healthcare costs at the same time that a higher Quality of Life (QoL) is a significant concern for people.

Gait analysis is used to assess and treat individuals with conditions affecting their ability to walk. To put-on sight walking deficiencies, clinicians have developed several walking and cognitive tests to detect problems in older adults with injuries and diseases. One of the better-known assessments is the Six Minute Walking Test (6mWT from now on). The former has been previously used as a traditional method to predict the risk of falling or other health morbidities by measuring the walking speed [14]. A growing body of research suggests that gait velocity is a good indicator and predictor of health issues like cognitive decline, falls, and even certain cardiac or pulmonary diseases [20] [11]. Another metric to clinically measure gait performance is Tinetti scale; it aims to assess the gait and balance in older adults; it is also an indicator of the risk of falling.

Even though traditional methods have been efficient and useful, in recent years, the Information and Communication Technologies (ICT) had presented an enormous potential to promote the development of a series of devices, sensors and techniques that promote better evaluation measures and, thus, contribute to provide a more reliable and complete information about a patient. It is expected that ICT will we able to provide new tools for early detection and prevention of diseases in the elder population.

The assistive device presented here (see Section 2.1) has been designed with the aim of helping physicians detecting certain kind of physical and cognitive conditions of a person by combining its sensors data along with the clinical information of an individual. It consists in a smart rollator which main novelty is to capture and provide information about the force applied by the user to the handlers of the rollaor while walking collected by sensors located in the handlers, among other data. This study hypothesizes that the force sensors of the presented assistive device can identify whether individuals perform compensation strategies while walking to avoid falls and that these strategies can be associated with their physical conditions. The analysis of force sensor has been poorly analyzed in comparison with other sensors as accelerometers and gyroscopes, but we believe that they enclosure critical information about person′s health, especially in the case of older adults.

In this work, we present an unsupervised approach to categorize elder individual′s walking patterns into multiple groups among a population of volunteers with diverse ages based on their interaction with the i-Walker, which has been used as a measuring tool. The methodology developed is an efficient system for determining a person′s walking-age.

The results of the presented work cluster the population into four categories in relation to their gait characteristics. The first two mainly separates the young people with higher and lower speed. The other two identify the geriatric gait and split the sample also considering the speed. The paper is organized as follows: Section 2 introduces related work and the i-Walker; Section 3 presents the approach used in this work, and Section 4 shows the experimental results. Lastly, Section 5 concludes this paper.

2 RELATED WORK

Several ICT tools and machine learning procedures had been used for the study walking patterns, each of them with different advantages and disadvantages mainly related to the costs and the reliability of the results. This chapter present some of these proposals as well as the previous work made with the *i*-Walker.

2.1 The i-Walker

The *i*-Walker is an assistive device based on a standard 4-wheeled rollator with embedded sensors and actuators. Fig 1 shows the *i*-Walker components, which have already been described in [4].

The most innovative feature of the *i*-Walker is the design of the handlebars, which include embedded force sensors able to measure the longitudinal, lateral and vertical forces (X, Y and Z respectively) applied by the user when walking.

In [2], authors developed a system to assess the estimation error of the force sensors. The results showed an error lower than 10%. This study also confirms the relationship between the force sensor and the support of the individual.

In [12], authors proposed a methodology for detecting common walking patterns in the elderly according to their physical and cognitive conditions by using the Symbolic Aggregation Approximation (SAX) algorithm for time series representation. They used a sample of 200 older adults where some of them have had suffered one or more falls during the last year, while the others did not. The algorithm was able to separate between fallers and non fallers using hierarchical clustering.

Figure 1: *i*-Walker components. The handlers contain the force sensors. The rear wheels have embed motors that work as actuators. The central box contains the computing power on- board as well a gyroscope and an accelerometer. Blocking brakes provide a braking help in down-hills.

2.2 Related Work using other systems

This section presents a description of previous research work in the field of gait analysis highlighting the device used to gather the information, their proposed approach and main contribution.

In 2012, [13] used a video camera they extracted different events of the gait cycle. They used a modification of the RANSAC algorithm to identify the motion of the person′s body. The result show that they could identify the phases of double-support, mid-swing, toe-off and heel-strike events of the gait cycle and estimate the gait speed.

In 2014, [8] also used a 3-D accelerometer and a gyroscope to measure speed, acceleration, and tilting of the legs. Their proposed approach was to use intrinsic mode functions (IMF) as a representation method and K-means as the clustering technique. The parameters used were the acceleration, and tilting of the legs and the results were compared with the walking speed. They made a walking-age clustering of three groups, children of ages 10 and below, adults between 20s and 60s, and elders in 70s and 80s.

In 2017, [22] recorded walking tasks using wearable inertial measurement units (IMU). They presented gait assessment system is based on Support Vector Machine (SVM) classifier. The parameters collected were the difference between the joint angle trajectories of the gait cycles and then they created a metric to estimate gait variability. Their main contribution that the could be able to differentiate a pathological gait from a normal one using a proposed gait variability index (GVI) as an indicator of the "healthiness" degree of a gait pattern.

In 2017, [6] used a system called GAITRite to characterize the pathological gait pattern. They used the allocated indexing membership (AIM) as a representation method and the Bayesian Gaussian Mixture as the clustering technique. They recorded spatio-temporal features such as gait symmetry ratios in swing time, stance time, and step length. The results found three clusters: the first group was formed by the healthy people, group 2 had the worse condition and group 3 contained people with conditions but with improvements in their recovery.

Figure 2: Flow diagram containing the stpes to perform the clustering for the walking analysis

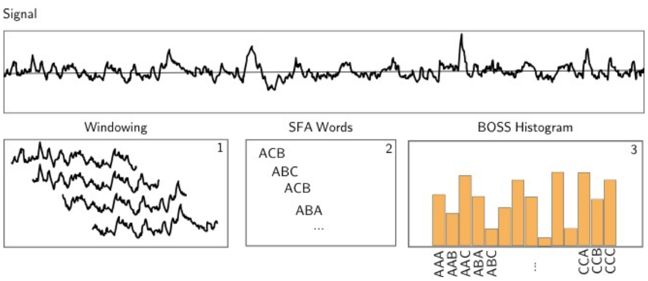

Figure 3: BOSS Model is used for indexing and representation and transforms the time series into BOSS histograms

3 APPROACH

This section introduces the data collection and methodology used to develop this work.

3.1 Data collection

Walking tests were conducted to record the data of the walking patterns of a group of volunteers. In this work, participants performed an adaptation of 6mWT, but reducing the time by half due to clinical recommendations so this study uses a Three Minute Walking Test (3mWT).

To perform the test, a total of 42 individuals whose ages range from 22 to 94 years were selected for this study. Each of them performed the 3mWT on a flat hallway moving back and forth. The variables are collected every 10ms while the user is walking to form a time series. The clinicians which nurse the older people of this sample provided their Tinetti scores for further analysis.

Several measures were obtained from the *i*-Walker sensors, but the most important are *(i)* the vertical, longitudinal and transversal forces of both hands which were used as the input of the model on which the clustering performs the pattern recognition and *(ii)* the estimated pose in X and Y to compute the distance travelled using the Euclidean metric and then to calculate the walking speed measuring the rate of change of the travelled distance.

3.2 Methodology

The methodology comprises the steps shown in in 2. The Bag-of-SFA-Symbols (BOSS) model is used for the representation and indexing. This method transforms the numeric time-series into a bag of words representation to later create a k-dimensional matrix.

The BOSS model describes time series as an unordered set of substructures using Symbolic Fourier Approximation (SFA) words. The BOSS flow model is represented as follows (see Fig 3 [16]): parameters definition, windowing, SFA and histograms aggregation and reduction.

3.2.1 Representation and indexing. The goal of this step is to transform each time series to a new representation based on a vocabulary extracted from the behaviour of the data in the frequency domain.

This transformation has a set of parameters that need to be tuned in order to find the optimal configuration: length of the windows, the number of Fourier coefficients for a window and the number of letters for the word representation.

For each time series, a set of fixed-size windows are generated using a sliding window of length w. The first window begins at 0 and ends in the position w, the second one will offset one position ending at position $w + 1$ and so on until the end of the series is reached. The result will be transformed into a vocabulary where each window represents a *word*. The length of the words are defined by the quantization of the Fourier coefficients of the window as explained in the following paragraphs.

The transformation of the windows into a vocabulary uses the SFA method, which aims to simplify information by removing the unnecessary data and keeping only the most representative characteristics. The SFA performs three steps to achieve its goal: approximation, quantization and words computation.

The SFA uses a discretization method based on the Momentary Fourier Transform (MFT), for the approximation process. The MFT targets to keep the critical data extracting the Fourier coefficients of the signal [1]. This method allows to incrementally compute the first f Fourier coefficients of a sliding window in a series in efficient manner.

The idea behind this discretization is to decompose the time series into two basic type of functions [17]. The former consists in identifying *slow* changes in the data, while the latter identifies *rapid* changes. For the approximation, those with slow changes are enough for a fair description of the signal. These types of functions also provide a smoother signal with low pass filtering.

The decomposition represents a time series by its Fourier coefficient. The magnitude of the coefficient represents the amplitude of the signal. The proposed approximation uses only the first f coefficients. The first Fourier coefficient can be optionally discarded because it stands for the mean value of the signal, obtaining this way offset invariance.

The quantization step reduces the granularity of the data by dividing the values of the Fourier coefficients into a histogram of equal frequency bins [7] and mapping each coefficient to its range. To define these histograms, a number of bins is defined as a parameters, representing the letters of the discretization alphabet that will be used for computing the words representing the windows extracted from the time series. Each position of the Fourier coefficients is discretized separately. This is done by computing Multiple Coefficient Binning (MCB), and it aims to minimize the lost information when performing the discretization process.

The discretization works as a map that contains intervals of numeric values per each coefficient. The outcome of the SFA is a word of f letters per window and a set of words per time series. In order to avoid a bias due to large periods of stable signal, the BOSS model reduces the length of the vocabulary by numerosity reduction, removing identical consecutive words.

The final stage of the BOSS model is to transform the vocabulary for a time series into a histogram of relative frequencies of words.

This transforms the series into a vector, so we can compare different series using different similarity measures. The main advantage of this final transformation is to allow the comparison of time series of different lengths.

3.2.2 Similarity Measure and Space Embedding. The next step is to evaluate *how* similar are the sets of time series. From the possible similarity functions that can be defined, the one that better fits the model is the cosine similarity measure because this metric considers only the orientation of a vector and not the magnitude [18]; this characteristic is useful when working with symbols instead of numbers. The output of this process is an affinity matrix that will be further transformed before clustering the data.

For this model, a Spectral Embedding [19] is applied to the affinity matrix to embed the data in a metric space and to enhance the relevant characteristics by using a non linear transformation. This embedding transforms the affinity matrix into a k-dimensional matrix. The number of dimensions k is part of the configuration parameters. For this paper, the transformation was limited at most to three dimensions.

3.2.3 Clustering. A Bayesian Gaussian Mixture Model (BGMM) with Dirichlet priors [9, 15] is used for partitioning the data represented by the k-dimensional data matrix. A Gaussian mixture model is a probabilistic model that assumes all the data points are generated from a mixture of a finite number of Gaussian distributions with unknown parameters. Each cluster is formed with a set of points that shape a Gaussian distribution using a Expectation Maximization (EM) algorithm. The use of a Dirichlet prior includes the determination of the number of clusters in the optimization process.

The EM for mixture models consists of two steps. The first step calculates the expectation of the component for each data point given by the model parameters. The second phase maximizes the expectations calculated in the previous step concerning the model parameters. Then these two steps are repeated until the result converges.

3.2.4 Evaluation. To assess the quality of the clustering, it is necessary to apply some evaluation techniques. This approach considers two techniques: *(i)* the adjusted rand index (ARI) [5] to measure the stability of the clustering to random initialization and *(ii)* the Silhouette index (SI) [21].

3.3 Scenarios

The work is based on the recognition of the patterns of the forces applied by the users to the *i*-Walker while walking. From the forces recorded by the *i*-Walker, the vertical ones are the most related to the individuals′ compensation strategies. Therefore, this study will examine two scenarios. The first model will be provided with all the forces to find if the transversal and longitudinal forces add relevant information. The second one will consider only the vertical forces to study whether the transversal and longitudinal forces add noise instead of contributing to the results.

4 EXPERIMENTS

This section presents the results obtained by applying the model clustering approach on both scenarios. The goal of the model is to find the set of parameters that provide the best clustering. Thus, the results are presented in two parts: the performance comparison of the different sets of parameters and the knowledge discovery.

Table 1: Results of the best two configurations of initial parameters per scenario and their evaluation.

Set of variables and evaluation of the best results per scenario				
Description	Set 1	Set 2	Set 3	Set 4
Scenario	1	1	2	2
Window size	135	128	131	131
Word length	7	4	6	4
No. of coeff.	2	2	2	8
Dimension	3	3	2	3
Adj. Rand index	0.90	0.93	0.93	0.93
Silhouette index	0.52	0.43	0.50	0.47
No. clusters	4	4	3	5

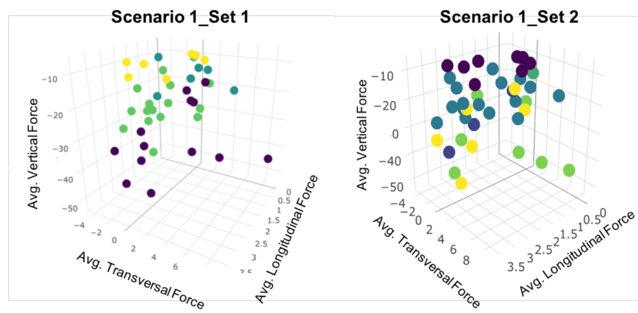

Figure 4: 3D plot of the vertical, longitudinal and transversal forces per cluster. The clustering is identified mainly over the axis of the vertical force.

Table 1 lists the best two sets of parameters per scenario with the most relevant findings along with their validation indexes and their number of clusters, all of them worthy to analyse for pattern recognition.

The knowledge discovery was carried out by visualizing the behaviour of patterns through several plots. It was necessary to select statistical variables that better represented the problem. A finding is identified if the comparison of those variables enhances the visualization of the groups of clusters. The selected variables are *(i)* the cluster, *(ii)* age, *(iii, iv, v)* the mean and standard deviation of the considered forces,*(vi)* the mean of the linear speed and the *(vii)* Tinetti score.

The next step is to identify if clustering visualization is preserved by comparing the selected variables. Fig 4 depicts the relationship between the X, Y and Z forces of the first scenario's set of parameters and it is noticeable that the vertical forces have a critical influence over the clustering.

Fig 5 shows that, if the standard deviation of the vertical force is added, the segments of clusters are more noticeable, which means that the variability measure is a factor that influences the clustering. Therefore, the proposed model meets its goal at considering the changes of the forces in the feature extraction.

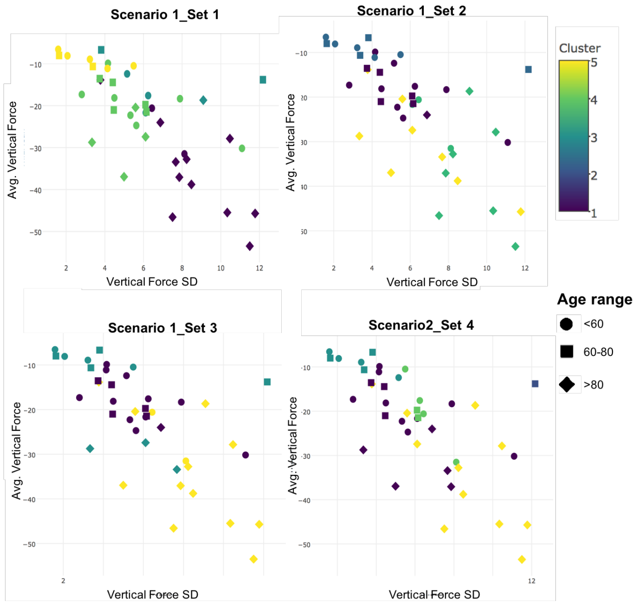

Figure 5: Comparison of the vertical force and its variability per cluster and age. It can be seen that the variability is directly related with the magnitude of the vertical force.

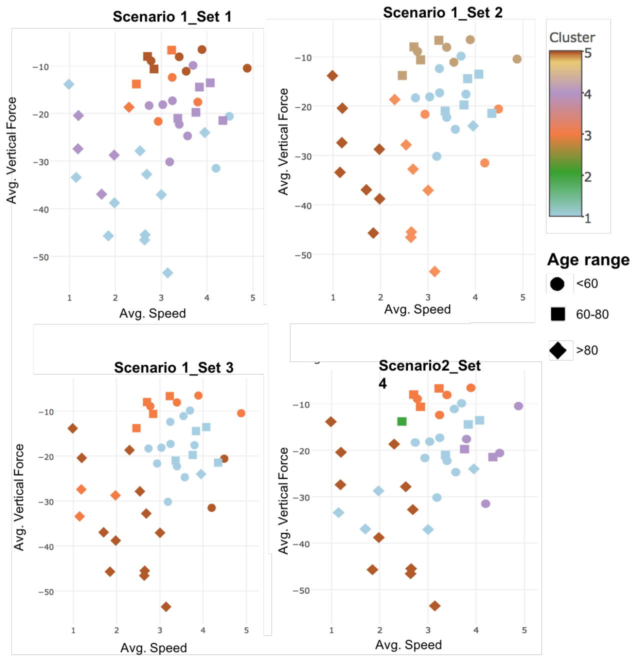

Figure 6: Comparison of the vertical force and the linear speed per cluster and age. The axis of the linear speed presents a tendency in the scenario 1.2 and 2.2.

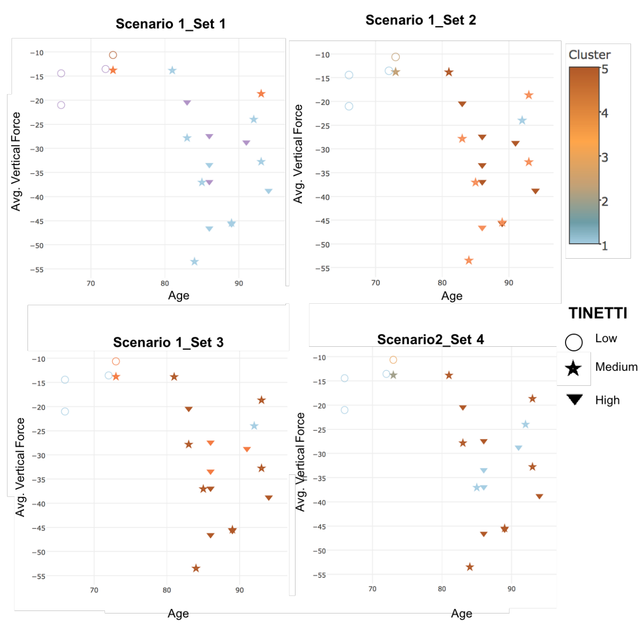

Figure 7: Comparison of the vertical force, age and Tinetti score.

When comparing the mean of the vertical force *vs*. the linear speed, it can be observed in Fig 6 that this variable is essential for the scenario 1.2. Moreover, if the age of the person is observed as well, there is a critical pattern that splits the clusters by age and speed as shown in Fig 6.

Fig 7 shows the relationship between the clustering formed and the Tinetti score considering only the group of elder people. The results from scenario 1 and test 2 depicts a strong correlation between the cluster and the risk of falling.

4.1 Clustering description

This section presents a comparison of the clustering results among scenarios. The characteristics considered are *(i)* the number of persons contained in each cluster per range of age, *(ii)* the average vertical force, *(iii)* the standard deviation of the vertical force variability and *(iv)* the average linear speed.

Table 2 describes the scenario 1.1. From the four clusters, clusters one and four have the most noticeable distribution of people per groups, older adults and young people respectively. The average forces are by far the most critical difference between each cluster as well as its variability.

The scenario 1.2, described in Table 3, has a more radical distribution of people: clusters one and three mainly contain the younger individuals and cluster two and four the older ones. These clustering characteristics are remarkably interesting. The clusters that contain the younger participants have similar average speed but a noticeable difference in their average vertical force. On the contrary, the groups including the older adults have similar average Z-force and variability, but they present a vast difference in the average speed. This contrast is enhanced because the primary descriptor of

Table 2: Scenario 1.1 clustering description. The vertical force and its variability are the two major differentiators

Clustering of Scenario 1 and Set 1				
Description	Clust.1	Clust.2	Clust.3	Clust.4
Under 65 years	2	3	7	5
65-80 years	0	2	5	2
Above 80 years	11	1	4	0
Avg. z-force(N)	34.7	15.3	21.5	9.1
Z-force sd.(N)	8.2	7.9	5.7	2.7
Avg.speed(m/s)	0.75	0.83	0.84	0.95

Table 3: Scenario 1.2 clustering description. All groups have a visible difference in the exerted forces, the variability and the speed.

Clustering of Scenario 1 and Set 2				
Description	Clust.1	Clust.2	Clust.3	Clust.4
Under 65 years	9	3	5	0
65-80 years	5	0	4	0
Above 80 years	1	7	0	8
Avg. z-force(N)	19.0	33.6	9.4	30.7
Z-force sd.(N)	5.6	8.8	4.1	7.0
Avg.speed(m/s)	0.99	0.85	0.92	0.42

Table 4: Scenario 2.3 clustering description.The clusters have differences mainly in the vertical force parameters

Clustering of Scenario 2 and Set 3			
Description	Clust.1	Clust.2	Clust.3
Under 65 years	11	4	2
65-80 years	5	4	0
Above 80 years	1	3	12
Avg. z-force(N)	18.7	14.8	8.5
Z-force sd.(N)	5.6	8.8	4.1
Avg.speed(m/s)	0.98	0.77	0.70

Table 5: Scenario 2.4 clustering description. The clusters have differences in the three variables presented.

Clustering of Scenario 2 and Set 4					
Description	Cl.1	Cl.2	Cl.3	Cl.4	Cl.5
Under 65 yrs	9	4	4	0	0
65-80 yrs	3	2	3	0	1
Above 80 yrs	5	0	0	11	0
Av.z-force(N)	22.5	20.2	8.8	33.7	13.8
z-force sd(N).	5.7	5.8	3.0	8.9	14.3
Av.speed(m/s)	1.18	0.88	0.58	0.68	0.80

cluster four is indeed the velocity, it contains the eight individuals with the lowest average speed, while the primary descriptor of the group two is the average Z-force. An observation about the clusters of young people is that even though all of them have similar characteristics, due to the lack of experience of using a rollator they tend to change their walking behaviour. Thus, the algorithm classifies them into two categories.

In Table 4, it can be observed that the scenario 2.1 has three clusters. In two of them, the distribution between young and elderly people is visible, but cluster two has a mix of both. The last scenario is presented in Table 5, which has been divided into five clusters, but the last one contains only one person. Therefore, it can be said that it has four relevant clusters and an outlier individual. From those four, again two clusters contain the young and healthy people, while the fourth cluster represents the elderly group. In fact, the distribution of clusters is quite similar to the one obtained in scenario 1.2 in terms of the applied vertical forces, however the average speed is different between scenarios, as well as the distribution of people aged 80+ years old.

The configuration of the scenario 1.2 has the bests results because it presents a better categorization per groups of age and because both variables, the speed (supported by the literature) and the forces applied (proposed by the presented approach) were noticeable differentiators between the clusters. This scenario also leads to thinking that the longitudinal and transversal forces do not add noise to the results, on the contrary, they enrich them.

5 CONCLUSION AND FUTURE WORK

This paper proposes an analysis methodology that can classify an individual's walking-age based on the 3 minutes Walking Test using the *i*-Walker. A database of signal recordings was built to test the methodology applying data preprocessing, clustering, and classification as an automated process. The experimental results determine a suitable configuration of initial parameters that distinct walking-age groups according to the forces exerted by the individual on the *i*-Walker and their walking speed. Therefore, there is an evidence that the compensation strategies for walking can influence the person's gait and that it is also a useful indicator metric to measure health. Another conclusion is that the *i*-Walker is a practical, safe and efficient tool that can be easily implemented to collect high-quality data successfully. In addition, results regarding the relationship between age and walking speed are coherent with the literature [11].

Although results are promising, the dataset used in this work fits for an exploratory analysis, however, statistically speaking, it falls into the realm of small datasets. As future work, it is expected to implement more walking tests to obtain a more extensive database to reinforce the analysis and findings. Also, it is recommended to include more elderly people to improve the quality of the results. The young group served as a baseline for this analysis, but according to [2], rollators are useful for gait characterization as long as users need the device for ambulation.

It would be helpful to work with a specialist to provide additional information about the gait features that can be critical for the analysis and match them with other *i*-Walker descriptors. Aforementioned, it is also essential to develop an efficient analyzer of age-related problems, such as early detection of Alzheimer disease and other physical or mental issues.

ACKNOWLEDGMENTS

The authors are partially supported by the project *Rehabilitación personalizada y adaptativa en tratamientos post-ictus: el i*-Walker: (TEC2014-56256-C2-2-P).

Prof. Ulises Cortés is a member of Sistema Nacional de Investigadores (SNI-III), México.

REFERENCES

[1] S. Albrecht, I. Cumming, and J. Dudis. 1997. The Momentary Fourier Transformation Derived from Recursive Matrix Transformations. *13th International Conference on Digital Signal Processing Proceedings* (7 1997). https://doi.org/10.1109/ICDSP.1997.628089

[2] J. Ballesteros, C. Urdiales, A.B Marinez, and J.H. van Dieen. 2016. On gait analysis estimation errors using force sensors on a smart rollator. *Sensors* 16, 11 (11 2016). https://doi.org/10.3390/s16111896

[3] Center of Disease Control and Prevention CDC. 2016. Falls are leading cause of injury and death in older Americans. (2016). https://www.cdc.gov/media/releases/2016/p0922-older-adult-falls.html.

[4] U. Cortés, A. Martínez-Velasco, C. Barrué, and R. Annicchiarico. 2012. AI Based Fall Management Services - The Role of the i-Walker in I-DONTFALL. In *Advances in Artificial Intelligence - 11th Mexican Int. Conf. on Artificial Intelligence, MICAI 2012, San Luis Potosí, Mexico.* 395–406. https://doi.org/10.1007/978-3-642-37807-2_34

[5] R. de Vargas and B. Bedrega. 2013. A Way to Obtain the Quality of a Partition by Adjusted Rand Index. *Workshop-School on Theoretical Computer Science (WEIT)* (10 2013). https://doi.org/10.1109/WEIT.2013.33

[6] E. Dolatabadi, A. Mansfield, K. Patterson, B. Taati, and A. Mihailidis. 2017. Mixture-Model Clustering of Pathological Gait Patterns. *IEEE Journal of Biomedical and Health Informatics* 21, 5 (9 2017). https://doi.org/10.1109/JBHI.2016.2633000

[7] A.P. Gurajada and J. Srivastava. 1991. Equidepth Partitioning of a Data Set based on Finding its Median. *Symposium on Applied Computing* (2 1991). https://doi.org/10.1109/SOAC.1991.143854

[8] B. Jin, T. Hoai Thu, E. Baek, S. Sakong, J. Xiao, T. Mondal, and M. Deen. 2014. Walking-Age Analyzer for Healthcare Application. *IEEE Journal of Biomedical and Health Informatics* 18, 3 (5 2014). https://doi.org/10.1109/JBHI.2013.2296873

[9] Z. Li and L.S. Mihaylova. 2017. Autonomous Flame Detection in Videos with a Dirichlet Process Gaussian Mixture Color Model. *IEEE Transactions on Industrial Informatics* PP, 99 (11 2017). https://doi.org/10.1109/TII.2017.2768530

[10] V. Monaco, A. Ghionzoli, P. Dario, and S. Micera. 2008. Muscle synergies during walking: Comparison between young and elderly people. Preliminary results. In *2008 30th Annual International Conference of the IEEE Engineering in Medicine and Biology Society.* 5370–5373. https://doi.org/10.1109/IEMBS.2008.4650428

[11] M. Montero-Odasso, M. Schapira, E.R. Soriano, M. Varela, R. Kaplan, L. A. Camera, and L. M. Mayorga. 2005. Gait Velocity as a Single Predictor of Adverse Events in Healthy Seniors Aged 75 Years and Older. *The Journals of Gerontology Series A: Biological Sciences and Medical Sciences* 60, 10 (2005), 1304–1309. https://doi.org/10.1093/gerona/60.10.1304 arXiv:http://biomedgerontology.oxfordjournals.org/content/60/10/1304.full.pdf+html

[12] J. Moreno, A. Cortés, C. Barrué, and U. Cortés. 2015. A Methodology for Pattern Detection in Gait Behaviour using the i-Walker. In *Artificial Intelligence Research and Development - Proc. of the 18th Int. Conf. of the Catalan Association for Artificial Intelligence, CCIA 2015 (Frontiers in Artificial Intelligence and Applications)*, Vol. 277. IOS Press, 287 – 290.

[13] Z. Duric N. Vishnoi and N.Gerber. 2012. Markerless identification of key events in gait cycle using image flow. *2012 Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)* (August 2012). https://doi.org/10.1109/EMBC.2012.6347077

[14] United Nations. 2017. World Population Ageing 2017 - Highlights. *World Population Ageing* (2017), 5.

[15] S. Sanjay-Gopal and T.J. Hebert. 1998. Bayesian pixel classification using spatially variant finite mixtures and the generalized EM algorithm. *IEEE Transactions on Image Processing* 7, 7 (7 1998). https://doi.org/10.1109/83.701161

[16] P. SchÃďfer. 2015. The BOSS is concerned with time series classification in the presence of noise. *Data Mining and Knowledge Discovery* 29, 6 (11 2015).

[17] P. SchÃďfer and M. HÃűgqvist. 2012. SFA: A Symbolic Fourier Approximation and Index for Similarity Search in High Dimensional Datasets. *Proceedings of the 15th International Conference on Extending Database Technology* (3 2012). https://doi.org/10.1145/2247596.2247656

[18] Michael Steinbach, George Karypis, and Vipin Kumar. 2000. A comparison of document clustering techniques. In *In KDD Workshop on Text Mining*.

[19] H. Strange and R. Zwiggelaar. 2014. *Open Problems in Spectral Dimensionality Reduction.* Springer, Alberystwyth. 7–20 pages.

[20] Perera S. Patel K. Rosano C. Faulkner K. Inzitari M. Guralnik Studenski, S. 2011. Gait Speed and Survival in Older Adults. *The Journal of the American Medical Association* (2011), 50–58. https://doi.org/10.1001/jama.2010.1923

[21] N. Verma, E. Dutta, and Y. Cui. 2015. Hausdorff Distance and Global Silhouette Index as Novel Measures for Estimating Quality of Biclusters. *Int. Conf. on Bioinformatics and Biomedicine, BIBM* (2015). https://doi.org/10.1109/BIBM.2015.7359691

[22] Durrant Matthias Spranger Xingchen Wang, Danijela RistiÄĞ and Axel GrÃďser. 2017. Gait assessment system based on novel gait variability measures. *2017 International Conference on Rehabilitation Robotics (ICORR)* (August 2017). https://doi.org/10.1109/ICORR.2017.8009292