IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received September 10, 2018, accepted October 26, 2018, date of publication November 19, 2018, date of current version December 19, 2018.

Digital Object Identifier 10.1109/ACCESS.2018.2882245

Gait Anomaly Detection of Subjects With Parkinson's Disease Using a Deep Time Series-Based Approach

GIOVANNI PARAGLIOLA AND ANTONIO CORONATO
Institute for High-Performance Computing and Networking, National Research Council, 80131 Naples, Italy
Corresponding author: Giovanni Paragliola (giovanni.paragliola@icar.cnr.it)

This work has been partially supported by both the eAsy inteLligent service Platform for Healthy Aging (ALPHA) Project and the project named "Realization of services and tools for Public Administrations for the implementation of the Electronic Health Record", a Convention between the Agency for Digital Italy of the Presidency of the Council of Ministers and the National Research Council of Italy.

ABSTRACT Parkinson's disease (PD) is a cognitive degenerative disorder of the central nervous system that mainly affects the motor system. The earliest symptoms evidence a general deficit of coordination and an unsteady gait. Current approaches for the evaluation and assessment of gait disturbances in PD have proved to be expensive, inconvenient and ineffective in the detection of anomalous walking patterns. In this paper, we address these issues by defining a deep time series-based approach for the detection of anomalous walking patterns in the gait dynamics of elderly people by analyzing the acceleration values of their movements. The results show a training accuracy and testing accuracy of over 90% with an accuracy improvement of 4.28% in comparison with related works.

INDEX TERMS Deep learning, convolutional neural network, long short-term memory, human behavior recognition, gait classification, neurodegenerative, diseases, deep neural network.

I. INTRODUCTION

Parkinson's disease (PD) is a cognitive degenerative disorder of the central nervous system that mainly affects the motor system. The earliest symptoms that arise consist in problems with mood or mental abilities. A general deficit of coordination and an unsteady gait are also involved [1].

Gait disorders are one of the main constituents contributing to a deterioration in the quality of life in patients with PD.

It has been proved that *changes* in gait dynamics can be interpreted as a warning of gait impairment above and beyond gait speed. These *changes* involve alterations or anomalies in the walking pattern, which does not present as consistent from one step to the next [2].

These *changes* are not always easily quantifiable in clinical procedure assessment but become apparent when the gait is evaluated quantitatively using a gait analysis system.

This observation means that the gait analysis is considered an important task for an *early detection* of developing symptoms of PD.

The sooner an anomalous walking pattern is detected, the sooner it is possible to predict the onset of an *early* stage of the disease.

Current approaches for the evaluation and assessment of gait disturbances in PD require that patients attend a medical center under a clinician's supervision. As a consequence, these approaches have proved to be expensive, inconvenient and generally ineffective in the detection of anomalies in walking.

In this paper, we aim at overcoming these issues by defining a deep time series-based approach for the detection of anomalous walking patterns in the gait dynamics of elderly people by analyzing the acceleration values of movements.

Currently, the gait dynamics of a subject are evaluated by means of analyzing the acceleration values along the (x,y,z) axes.

These data have the structure of a *temporal time series* (TTS) with the X axis measuring the time and the Y axis measuring the acceleration value. The *temporal time series* is, therefore, a sequence of values indexed in time order.

The novelty of this paper consists in the presentation of an innovative and more efficient way of analyzing the TTS for the detection of anomalous walking patterns in the gait dynamics.

In Table 1 we have reported the related works that address the issue of walking pattern identification in the gait dynamics using state-of-the-art methods (e.g. SVM, Decision Tree, etc. [3]) for the classification problems.

All these related works analyze the TTS by applying classic pattern recognition based approaches to the gait dynamics;

TABLE 1. Summary of the related works.

Solution	Technologies	Results	Classification Approach
[12]	- Wavelet - SVM	- PD Vs CO = 90.32%	- sample-by-sample Classification
[13]	- SVM	- PD Vs CO = 83.33%	- sample-by-sample Classification
[14]	- HMM	- PD Vs CO = 90.3%	- sample-by-sample Classification
[16]	- MLP - RF - NB	- PD Vs CO = 91.1% * *the results refer to a Random Forest based on the Maximum Area under the Curve evaluation method (AUC)	- sample-by-sample Classification
[18]	- MLP - SVM - NB	- PD Vs CO = 86.75% * *the results refer to the neural network	- sample-by-sample Classification
[17]	- SVM	- PD Vs CO = 86.17%	- sample-by-sample Classification
Our Approach	- LSTM - Convolution Neural Network - Deep Neural Network	- PD Vs CO = 95%	- Time Series Classification

the training of the machine-learning methods adopted is sample-based.

These approaches do not take into account the fact that the values of the TTS have the property of being ordered by time.

Our approach does not rely on a sample-based pattern recognition process; instead, we have analyzed the whole TTS in order to evaluate the ordering by time of the data.

This is the aspect that most significantly distinguishes our solution from other works concerning the gait analysis of subjects with a cognitive disease (e.g. PD or Huntington's disease).

In order to analyze the TTS, a hybrid deep neural network (NN) has been defined, where the term *hybrid* indicates that our NN is a combination of three different types of neural network algorithm.

The architecture of the hybrid NN consists in two components (Figure 7):

1) a *Reduction Layer* (RL) in charge of producing a coded version of the input TTS in order to improve the classification performance.
2) a *Classification Layer* (CL) in charge of classifying a gait as normal or anomalous on the basis of the coded representation of the TTS.

The RL is defined as a Convolutional Neural Network (CNN), while the CL is a hybrid neural network, built by combining a Long Short-Term Memory (LSTM) and a Deep Neural Network (DNN)

An general diagram representing our solution is shown in Figure 1.

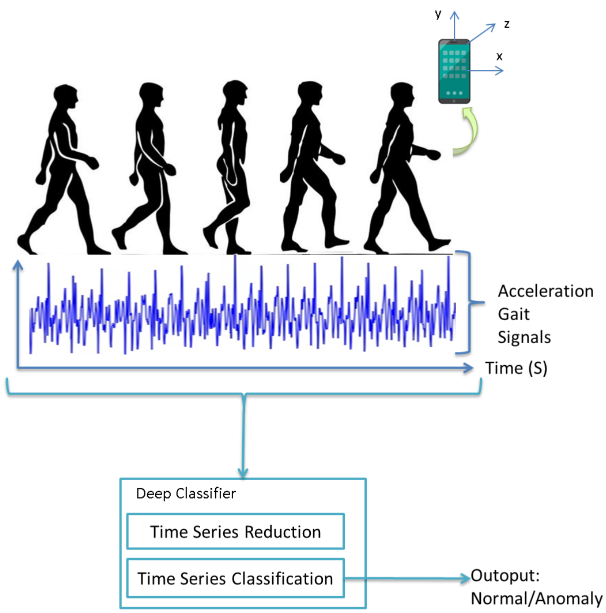

FIGURE 1. In-house self gait assessment.

The main contributions of this presented work consist in two aspects:

- The definition of a deep time series-based approach for the detection of a walking anomaly by analyzing the whole TTS of the gait dynamics.
  The results show an increase in the classification accuracy of up to 4.28% in comparison with related works

- A comparative analysis among different reduction approaches to address the coding of a TTS for classification.

The final objective of our research will be to overcome the problems of the traditional approaches by defining an easy in-house self-test mobile solution able to detect anomalies in the gait dynamics of elderly people.

In Section II a description of the state-of-the-art relating to the assessment and classification of the gait dynamics of a patient with PD is presented. In Section III we describe the proposed approach, highlighting the dataset and the definition of the hybrid neural network. In Section IV we present the results of our work, while in Section IV we provide a discussion of our research and indicate our future plans.

II. RELATED WORKS

This section includes the following contents:

- The most relevant ICT-based and clinical approaches for the assessment and classification of the gait dynamics of patients with PD
- The limitations of all ICT-based approaches
- - The advantages of our approach in comparison with the related works

A. CLINICAL APPROACHES

Currently, the state-of-the-art for the assessment of the gait dynamics and balance of PD patients relies on clinical approaches based on rating scales, questionnaires, and timed tests [4].

Timed Up and Go Test [5], Functional Gait Assessment [6], and Balance Evaluation Systems Test [7], are examples of gait assessment methods showing good results in the evaluation of the gait dynamics and risk analysis; all these methods are ambulation-based performed under the supervision of a clinician and expert.

Additionally, most clinical units adopt a personal gait assessment, such as the Unified Parkinson Disease Rating Scale (UPDRS) [8] or patient self-reports of function, such as the Parkinson's Disease Quality of Life Questionnaire (PDQ-39) [9].

These approaches are qualitative assessments with a possible large variability, where *variability* means that different supervisors may give a different assessment in relation to the same subject

These limitations make these methods expensive and less precise and, in some cases, require that the patients attend a medical center.

Furthermore, all these approaches are performed after the onset of the disease.

B. TRADITIONAL ICT-BASED APPROACHES

Gait Recognition is a type of classification problem where training examples currently consist of sequences of value which model the movements and the kinematics of a subject (Figure 2) [10].

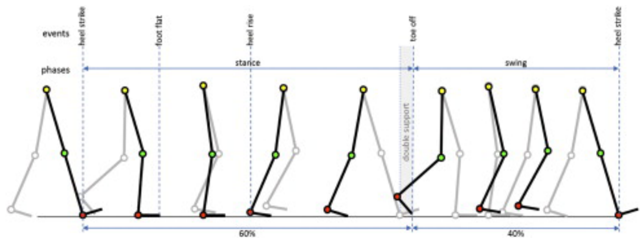

FIGURE 2. **Phases of the human gait [10].**

Figure 4 shows the main differences between our approach and the related works from the point of view of input data point.

At the top of the figure, generic (x,y,z) acceleration values of the gait dynamics are shown.

A features extraction process is performed over the values. A set of temporal windows is extracted from the signal. The choice of the length of the window and the overlapping rate between two consecutive windows is a typical engineering problem which, generally, makes such a process complex and time-consuming.

From each window a set of engineered features is extracted. This defines a generic sample as a tuple of values. A sample is a basic unit for the training of a generic machine learning algorithm.

All the extracted samples define a set of samples (A-Figure 3).

The ICT-based approaches adopt machine-learning-based algorithms in which a NN learns by casually batching the sample set.

This procedure is repeated many times during the learning, defining the so-called *training stage*.

In relation to the ICT-based approaches, we introduce the most relevant works (Table 1) aimed at classifying walking patterns using the Hausdorff et al dataset [11].

In Table 1, we report the following information:

- Technologies: the algorithms used for the analysis of the gait dynamics
- Results: a list of the results achieved
- Classification Model: a description of how the gait dynamics are analyzed

In [12], an approach based on the combination of Support Vector Machine (SVM) and wavelet analysis is explored to classify PD and healthy subjects using their gait cycle variability. Although the accuracy of the classification is 90.32%, the wavelet analysis increases the cost of the computations.

Shetty and Rao [13] apply a Gaussian-based kernel with an SVM classifier to successfully distinguish PD from healthy subjects. The SVM classifier achieves a good accuracy (83.33%).

Khorasani and Daliri [14] use the Hidden Markov Model (HMM) with Gaussian Mixtures to distinguish patients with PD from healthy subjects. They achieve an accuracy of over 90%. The authors use the raw gait data

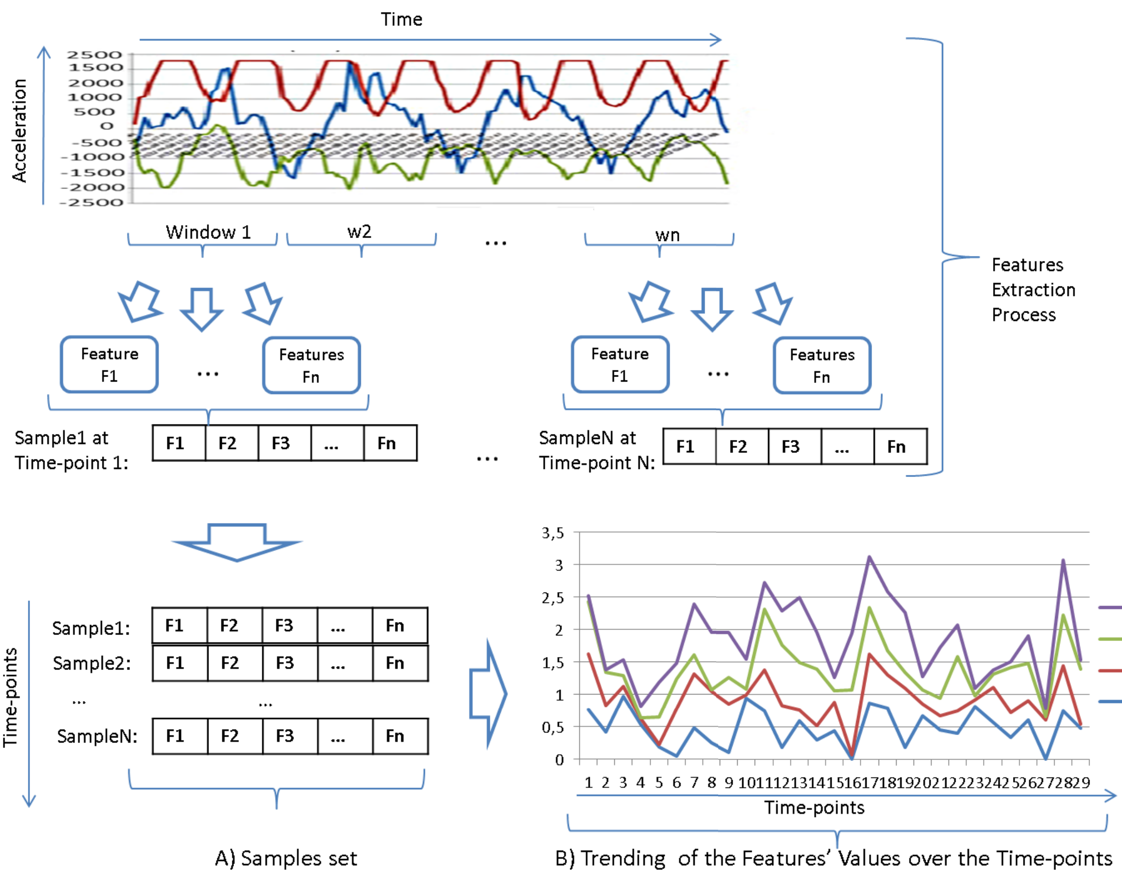

FIGURE 3. Overview of the classification model [10].

instead of the extracted features. Although the accuracy is good, the HMM involves certain factors which limit its application to the evaluation of a sequence of data, such as the assumption that successive observations are independent of each other whereas, in reality, they rarely are [15].

Ren *et al.* [16] adopt a machine-learning-based approach for further analysis. They use three types of classifier: multilayer perception (MLP), random forest (RF) and Naive Bayes (NB). The dimensionality of the data has been reduced using a PCA technique.

Yang *et al.* [17] investigate three feature selection techniques and one feature construction method for the analysis of a gait dataset of neurodegenerative disease patients and adopt an SVM-based classifier to assess the prediction.

Baby *et al.* [18] deal with the application of wavelet transform to the gait data of PD patients, with the aim of distinguishing PD patients from healthy controls. Statistical features extracted from wavelet coefficients are used to classify the subjects. The authors evaluate two types of classifier, namely an SVM classifier and a Naive Bayes' classifier.

Other time series classification methods include the series similarity ([19]) approaches. These are appropriate when there may be discriminatory features over the whole series [20]. However, this observation is not necessarily correct in respect of the gait dynamics signal of patients with a neurodegenerative disease [4]. For this reason we have not adopted this approach in our experiments.

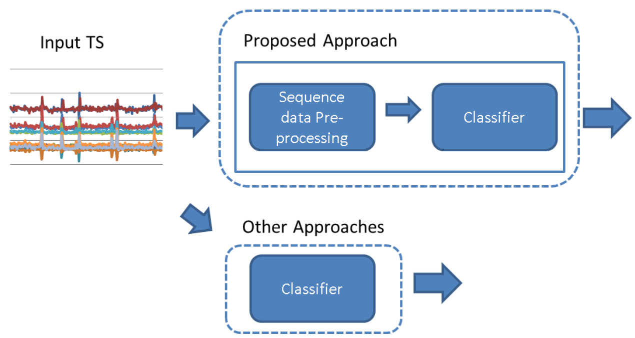

FIGURE 4. Logic view of the proposed approach.

It is worth performing an analysis of the strengths and weaknesses of our proposed solution in comparison with the other ICT-based approaches in order to highlight the advantages and limitations.

1) LIMITATIONS OF ALL ICT-BASED APPROACHES

Figure 4 shows a logic view of our approach compared to classic approaches.

The property of being a temporal sequence is quite difficult to model in classic supervised machine-learning algorithms.

The traditional approaches do not pre-elaborate the whole TS. Instead, they train a generic classifier (e.g. SVM or Decision Tree) by feeding the signal sample-by-sample (A - Figure 3). However, in relation to gait recognition, this kind of training has certain limitations due to the fact that it does not take into account the temporal-sequential ordering of the whole sequence during the training of the classifier.

On account of the casual batching of the sample set, the ordering by time of the values of the sequence is absent during the training stage.

On the other hand, in our solution the LSTM network performs the sequence data pre-processing step in which the entire TS is taken into account and analyzed as a whole. In this way, the temporal-sequential order is maintained unaltered.

The output of the LSTM is then used to feed a classifier that in our solution is a DNN.

The gait dynamics exhibit a significant sequential correlation which should be exploited to improve the prediction accuracy of the classifiers rather than requiring them to learn each sample independently [21].

In relation to the issue of gait classification, this characteristic, which also contributes to improving the accuracy of the classification process, represents the most significant difference between our solution and the state-of-the-art approaches

2) ADVANTAGE OF OUR APPROACH

Anomalous gait detection is a type of *classification* problem where the training examples currently consist of the whole sequence of values [10]

In order to address the issue of modeling the sequential correlation of the gait dynamics, we have designed our hybrid NN for the classification of the trend of the features values over the time-points (B - Figure 3).

Our hybrid NN learns the whole trend of the features' values over all time-points (B - Figure 3) in order to classify it as normal or, if an anomalous walking pattern is detected, anomalous.

The soundness of our approach is also highlighted in Figure 2 which shows that a gait is a sequence of temporal phases describing the movements of the subject.

In order to classify the gait dynamics, it is necessary to take into account all these temporal phases from the beginning, at time t=0, to the end.

Our approach performs a time-series classification of the whole sequence of samples ordered by time-points instead of making an independent sample-by-sample classification.

In the classification, our classifier recognizes as normal or anomalous the whole TTS of the gait dynamics, instead of classifying each sample one at a time as the state-of-the-art methods do.

III. PROPOSED APPROACH

In this section, we will present our proposed approach.

First of all, we will describe the Hausdorff's dataset [[11]– [22]] downloaded from PhysioNet (www.physionet.org) used for the training and testing stages of our NN. Next, we will present both a comparative analysis dealing with the problem of the dimensionality reduction of a time series and then, the classification process.

A. DATA DESCRIPTION

The dataset has been created by observing a set of more than 60 subjects, including patients with PD (age: 66.8 10.9 (SD) years), and healthy controls (age: 39.3 18.5 (SD) years).

In accordance with the experimental protocol, each subject was asked to walk at their normal pace along a 77-meter long hallway for 5 minutes.

Each subject performed one experiment.

Accelerometers were placed in the subject's shoes, the output of these sensors providing a raw TTS relating to the subject's movements.

In accordance with Figure 3, a feature extraction process was set and a sequence of temporal windows was applied over the TTS [22]. The length of each window was one second, each window defining a time-point.

From each window, a vector of engineering spatial-temporal features was extracted.

$$\forall x \in TTS, \quad x = [F_1, F_2, \ldots, F_n],\ n = 0, \ldots, 12$$

where F_n is characterized by the following:

- Left Stride Interval (sec)
- Right Stride Interval (sec)
- Left Swing Interval (sec)
- Right Swing Interval (sec)
- Left Swing Interval (% of stride)
- Right Swing Interval (% of stride)
- Left Stance Interval (sec)
- Right Stance Interval (sec)
- Left Stance Interval (% of stride)
- Right Stance Interval (% of stride)
- Double Support Interval (sec)
- Double Support Interval (% of stride)

Each vector defines a sample. The total dataset counts more than 15k samples extracted during the experiments.

Figure 5 shows a portion of the dataset where each row is a sample.

A machine learning-based approach is trained by learning sample-by-sample, aimed at classifying a generic sample as normal or anomalous at the testing time

In a different way, in our approach, we evaluate the set of samples of each patient as a sequence of pairs (time-point, sample) so defining a multivariate time series (MTS), where each series is related to a single feature. The total number of features is 12.

An MTS describes the trend of the features values over the time-points for a subject (B - Figure 3).

Figure 6 shows an example of two MTS, the one at the top relating to a healthy subject, the second one to a subject with early stage PD.

Each MTS had a different length and therefore we decided to pad each one to the maximum length of about 300 time-points.

Patient ID	Time-Point	Left Stride Interval	Right Stride Interval	Left Swing Interval	Right Swing Interval	Left Swing Interval (\% of stride)	Right Swing Interval (\% of stride)	Left Stance Interval	Right Stance Interval	Left Stance Interval (\% of stride)	Right Stance Interval (\% of stride)	Double Support Interval	Double Support Interval (\% of stride)
10	11	1,1833	1,2067	0,3833	0,3533	0,3239	0,2928	0,8	0,8533	0,6761	0,7072	0,4467	0,3775
10	12	1,1467	1,18	0,3867	0,37	0,3372	0,3136	0,76	0,81	0,6628	0,6864	0,39	0,3401
10	13	1,1767	1,14	0,3833	0,3467	0,3258	0,3041	0,7933	0,7933	0,6742	0,6959	0,4467	0,3796
10	14	1,1833	1,19	0,4033	0,3733	0,3408	0,3137	0,78	0,8167	0,6592	0,6863	0,4067	0,3437
10	15	1,1	1,1467	0,33	0,34	0,3	0,2965	0,77	0,8067	0,7	0,7035	0,43	0,3909
10	16	1,2367	1,1467	0,4133	0,3567	0,3342	0,311	0,8233	0,79	0,6658	0,689	0,4667	0,3774

FIGURE 5. Sample set.

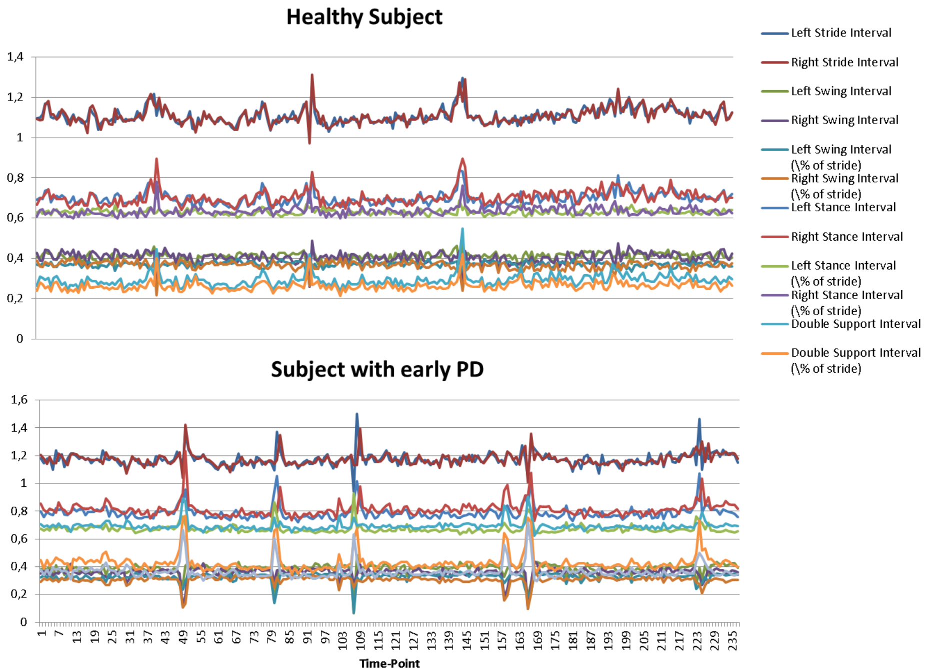

FIGURE 6. Trending of the Features values over time of a subject with early PD and a healthy subject.

Our approach relates to the learning of the whole MTS, aimed at classifying each MTS as normal, if it belongs to a healthy patient, or anomalous, if it belongs to a subject with PD.

B. PROPOSED APPROACH

Figure 7 shows an overview of the hybrid NN.

The hybrid NN is composed of two logical layers: a *Classification Layer* (CL) responsible for classifying a sequence of values as normal or anomalous and a *Reduction Layer* (RL), responsible for producing a coded version of the MTS to feed the CL.

1) CLASSIFICATION LAYER

The CL is a neural network, built by combining a Long Short Term Memory (LSTM) and a Deep Neural Network (DNN).

- the *Long Short-Term Memory* consists of NNs for the processing of sequential data x(1), . . . , x(t) with the time step index t ranging from 1 to T, with T being the length of the sequence [23].
  Each value x_i at a time-point i is handled by a unit cell. The information has to sequentially travel through all the cells before arriving at the cell at time T. This behavior enables the network to learn the information of the whole sequence of values.
- the *Deep Neural Network*, also called multilayer perception (MLPs) [24], is the quintessential deep learning model. The goal of an MLPs is to approximate some functions $f(x)$.
  For example, for a generic classifier, $y = f(x)$, it maps an input x to a category y. A Deep Neural Network defines a mapping $y = f(x; \theta)$ and learns the value of the parameters θ that result in the best function approximation [23].

The LSTM is fed with the sequence of values, providing as output a numerical vector. The LSTM's output has the important property of having been defined taking into account the property of the input values to be ordered by time. The LSTM output is used as input for the DNN.

Next, the DNN classifies the input as normal walking or anomalous walking, if an anomalous gait dynamics is detected

In our first experiments, the hybrid NN was composed of only the CL.

The MTS (Figure 6) submitted directly into the LSTM and, as a consequence, the network had to analyze a sequence of 300 time-points.

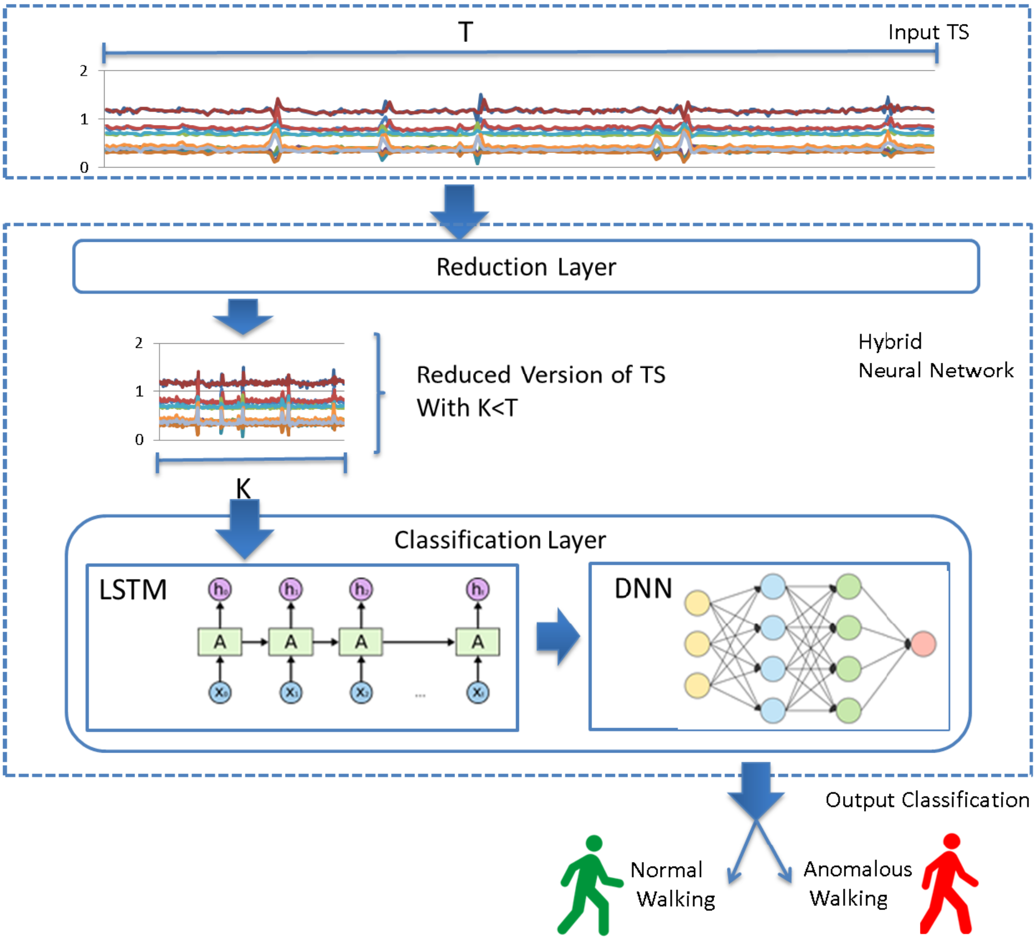

FIGURE 7. Overview of the proposed hybrid neural network.

TABLE 2. Performace classification without reduction layer.

Precision	Recall	Training Accuracy	Testing Accuracy
65%	70%	94%	70%

This approach did not produce satisfactory results.

These experiments achieved a classification accuracy lower than of the related works, an average training accuracy of 94% and an average testing accuracy of 70% (Table 2).

The gap between the training and testing accuracy also indicates a strong over-fitting.

In our opinion, these relatively poor results can be explained in terms of the excessive dimensionality of the MTS.

Feeding a LSTM layer with very long sequences makes the problem of vanishing and exploding gradients occur when back-propagating errors across the entire time-points sequence [25].

The training of a LSTM is based on the well-known gradient-based *back propagation algorithm* [26].

During the training stage, the gradients may vanish or explode *exponentially with respect to the number of time steps* [23].

This causes the NN to produce badly trained models.

This unwanted behavior brings to light the need to feed the CL with a smaller but loss-less representation of the MTS.

The RL component is responsible for addressing this issue.

2) REDUCTION LAYER

The necessity of analyzing the whole MTS makes the classification process more complex than a typical pattern-recognition problem in which a network tries to identify those portions of the signal characterized by a specific pattern ([27]–[30]).

The high dimensionality of a TTS may make many machine learning-based approaches inefficient and ineffective [31] when being applied to very large datasets.

This is a well-known issue sometimes referred to as the *curse of dimensionality* [32].

In most TTS problems, there is a requirement for a *dimensionality reduction* in order to define a new representation of the data series.

Dimensionality reduction (either of the number of time points or the number of features), can effectively reduce the computational overhead on condition that *the new representation safeguards sufficient information to solve the specific TTS problem correctly*, in our case a classification problem [33].

In general, a dimensionality reduction problem can be stated in terms of whether, given a TTS of an arbitrary length M (e.g. 300 time-points), it can be reduced to another representation of the data series of length K, with $K < M$.

In our case, with the maximum length of a MTS being 300 time-points; the aim of the RL is to provide a reduced version of the MTS to feed the CL.

We have carried out a comparative analysis of four approaches [34] for dimensionality reduction in order to improve the classification accuracy: *Under-sampling*, *Fourier Transformation*, *Autoencoder* and *Convolutional Neural network*.

For each approach, we have evaluated the improvement of the classification performance compared with the results achieved without applying a reduction process (Table 2).

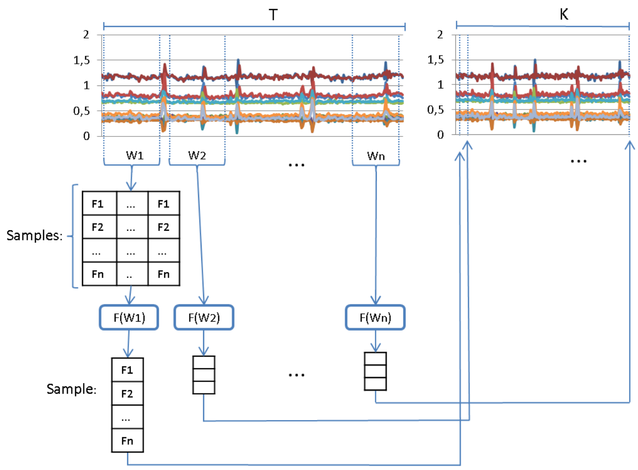

FIGURE 8. Overview of the under-sampling reduction process.

The input of each reduction approach is the MTS where each MTS has a length of 300 time points and each value is a sample described by 12 features (Figure 6).

- *Under-sampling*: This is the simplest method [35], in which a rate of M/K stands for the compression rate. However, the shape of the compressed TTS may be roughly similar.
  In our experiments, we defined a windows-based under-sampling reduction approach.
  Given the input MTS, we extracted a portion of the values as the length of the window; from each window, we applied an aggregation function to reduce the window of values to a single value (Figure 8).
  There is no overlapping between two consecutive windows.
  The sequence of these aggregated values defines the reduced version of the MTS (Figure 8) with a length at K time-points where each value is still described by 12 features (Figure 5)
  The parameters of interest of this approach are the *length of the window*, and the *aggregation function*.
  Figure 10-A, shows a list of the parameter configurations in relation to this approach, which we have tested to improve the classification performance. The *length of the window* was set to 10 or 5 time-points whereas the *aggregation function* was set as the average function or the random function.
  Each row describes the value of the parameters for a specific configuration.
  The column *Length of Reduced Signal* states the length of the reduced TTS. As an example, for the configuration *UCc4* we achieved a reduced version of 60 time-points.
- *Fourier Transformation*: Representing the data series in the transformed domain is another approach.
  One of the most frequently used transformation techniques is Discrete Fourier Transform (DFT) [36]. DFT maps a discrete periodic sequence f to a discrete sequence of coefficients F, representing the Fourier Transform of the sequence.
  In our experiment, the sequence f was the MTS. It was transformed by means of a Fourier Function $F(TS)- > TS_f$. Next, we filtered a subset of coefficients F from the transformed signal and used this for the classification.
  The parameter of interest of this approach is the *Fourier Coefficients*, F, that states the number of filtered coefficients.
  Figure 10-D shows the ranges of values of the Fourier coefficients F used during the experiments.
  The reduction process produces a code version with a length of F coefficients where each value is equal to the module of the coefficient.
  The time duration of the reduction process of both the DFT and Under-sampling approaches was in the order of seconds.
- *Autoencoder*: Autoencoder networks are a specific type of NNs where the input is the same as the output.
  These NN compress the input into a lower-dimensional code and then reconstruct the input from that representation to obtain an output signal equal to the input with a reconstruction error as small as possible. The code is a compact summary or compression of the input [23].
  We built our autoencoder network using a pair of MPLs, named the encoder and decoder, where the decoder is the mirror image of the encoder.
  We experimentally evaluated that reducing the whole MTS into a smaller version was time-consuming and that the level of the error between the MTS and the reconstructed MTS was considerable.
  In order to address this issue, we applied a windows-based approach. We extracted sequential windows of values from the MTS and reduced them using the autoencoder nets (Figure 9). There was no overlapping between two consecutive windows.
  The sequence of the reduced windows defines an output sequence with a length of K time-points, where each value is still described by 12 features (Figure 9)
  Figure 10-B shows the parameters of interest for the tuning of the autoencoder.

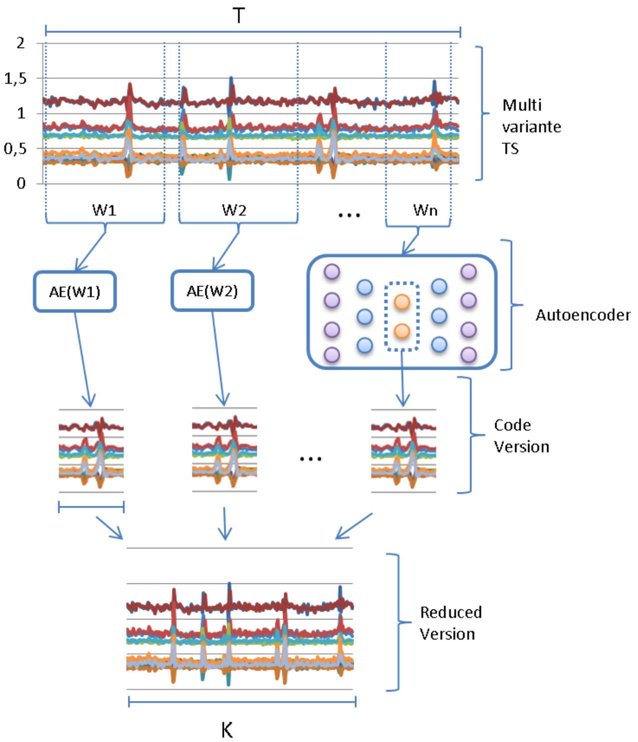

FIGURE 9. Overview of the autoencoder-based reduction process.

The *window size* states the length of the window. The *compression rate* is the compression rate to apply over the window (e.g. a reduction rate of 0.80 produces a window of 30 time-points to obtain a size of 6 time-points)

The *reconstruction error* assesses the error between the input MTS and the reconstructed MTS.

Concerning the structure of the autoencoder, the hyper-parameters of both the encoder and decoder are:

- Number of layers: 4
- Number of units per layer: 1024 units
- Learning Rate: 0.01
- Training Epochs: 1000
- Number of Iterations: 500

The time duration of the reduction process was in the order of days

A grid search approach was adopted setting the range of the values as:

- Number of the layers: {1;2;3;4;5}
- Number of units of the layers:{512;1024;2048;4096}
- Learning rate: {0.01;0.05;0.001;0.005;0.0001}
- Training Epochs: {500; 1000; 1500}
- Number of Iterations: {250; 400; 500;800}

A grid search approach is an exhaustive searching through a manually specified subset of the hyperparameter space.

- *Convolutional Neural network*: The CNN structure is formed of two basic layers: a *convolutional layer* (C), and a *polling layer* (P).

The convolutional layer is in charge of applying a convolutional operator over the input. The pooling layer performs a filtering operation from the results of the convolutional layer.

A convolutional layer uses a set of filters that process small local regions of the input where these filters are replicated along the whole input space.

A subsampling step (pooling) generates a lower resolution version of the convolution layer activations by taking the maximum (or average) filter activation from different positions within a specified window.

The CNN takes as its input a MTS of 300 time-points and returns a reduction version.

It is important to note that each time-point of both the input and output TTS is still modeled as a vector of features, and that the CNN does not change the dimensionality of the space of features.

Figure 10-C shows the parameters of interest for the convolutional network: the *kernel size* specifying the length of the convolution window; the *filter size* dimensionality of the output space; and the *pooling size* of the pooling windows.

We determined the parameters by using a grid search approach.

The range of values was:

- Window size: {10, 11, 12}
- Filter size: {4, 5, 6}
- Pooling size: {10, 5}

For each configuration we adopted an average pooling layer

IV. RESULTS

In this section, we present the results of the final configuration of the hybrid NN and the experiments.

The classification layer is composed of two neural networks, a LSTM and a DNN. The final configuration of the hyperparameters of both NNs is reported above. For each network, we have performed a search grid approach in order to find the best configuration of the parameters.

For each network also, we have adopted a search grid approach in order to find the best parameters' configuration.

- Number of the units of the Layer of the DNN: 2048
- Number of fully-connected layers of the DNN: 5
- Number of units for each cell of the LSTM: 10
- Learning Rate: 0.005
- Training Epochs: 1000
- Number of Iterations: 1000

We randomly partitioned the whole dataset into training, validation and testing datasets, respectively. For the training stage, we defined a training set as 70% of the original set whereas the validation set was defined as 10%. For the testing stage, we defined a testing set as 20% of the whole dataset.

First, we describe the overall results of the classification process and then we present the contribution of the reduction layer to the results.

Undersampling Configuration (US)			
Configuration	Windows Size	Agregation Function	Length of Reduced Signal (Time-points)
US_c1	10	Average	30
US_c2	5	Average	60
US_c3	10	Random	30
US_c4	5	Random	60

A)

Autoencoder Configuration (AE)				
Configuration	Windows Size	Compression Rate	Length of Reduced Signal (Time-points)	Reconstruction Error
AE_c1	10	0,5	150	0,02
AE_c2	20	0,5	150	0,41
AE_c3	30	0,2	240	0,31
AE_c4	30	0,5	150	0,5
AE_c5	30	0,8	60	1,75

B)

Convolution network Configuration (CNN)				
Configuration	Kernal Size	Filter Size	Pooling Size	Length of Reduced Signal (Time-points)
CNN_c1	10	4	10	30
CNN_c2	11	5	5	60
CNN_c3	12	6	10	30

C)

Fourier Trasformation Configuration (FFT)		
Configuration	Fourier Coefficients	Length of Reduced Signal (Time-points)
FTT_c1	60	60
FTT_c2	50	50
FTT_c3	40	40
FTT_c4	30	30

D)

FIGURE 10. Overview of the configuration parameters for the input reduction process.

For the evaluation of our model, three metrics have been used: precision, recall and accuracy [37].

These measurements are defined as follows:

$$Recall = \frac{TP}{TP + FN} \tag{1}$$

$$Precision = \frac{TP}{TP + FP} \tag{2}$$

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \tag{3}$$

where TP is the number of true positives, FN is the number of false negatives, TN is the number of true negatives and FP is the number of false positives.

Figure 12 shows the classification results related to all reduced approaches for each configuration.

Interesting results were achieved by the CNN and AE approaches.

The best results were achieved by the CNN-based reduction approach, where the classification accuracy at the testing stage ranged from 89% - 95%.

The configuration *CNNc3* achieved both a training accuracy and testing accuracy over 90%, respectively 94% and 95%, with an accuracy improvement of 4.28% compared with related works.

The quality of the solution is also confirmed by the small gap between the accuracy of the training stage and that of the testing stage which demonstrates a lack of over-fitting.

Moreover, the values of recall and precision, both up to 85%, also suggest that this is a highly effective method.

These results rate our model as the best solution among the presented ICT-based related works.

Additionally, the Autoencoder approach shows good results. The accuracy increases as the window size increases achieving the best results with a *window size* at 30 time-points and a compression rate of 0.8.

Reduction Approach	Configurations	Accuracy Improvement
Undersampling	US_c1	0,00%
Undersampling	US_c2	-12,86%
Undersampling	US_c3	-12,86%
Undersampling	US_c4	-10,00%
Fourier	FTT_c1	-5,71%
Fourier	FTT_c2	-10,00%
Fourier	FTT_c3	-25,71%
Fourier	FTT_c4	-12,86%
Autoencoder	AE_c1	-11,43%
Autoencoder	AE_c2	-2,86%
Autoencoder	AE_c3	0,00%
Autoencoder	AE_c4	2,86%
Autoencoder	AE_c5	7,14%
CNN	CNN_c1	34,29%
CNN	CNN_c2	27,14%
CNN	CNN_c3	35,71%

FIGURE 11. Improvement contribution of the reduction layer to the classification accuracy.

This configuration reaches a classification accuracy of 75%, with a small gap between this results and the accuracy of the training step. This demonstrates that the model is not significantly affected by over-fitting.

The Undersampling and Fourier transformation approaches do not achieve results comparable with related works. The accuracy classification is about 60% and also recall and precision obtain results lower than 70%.

The improvement contribution of the reduction layer is reported in Figure 11 and Table 3.

The accuracy achieved without reducing the MTS is reported in Table 2, the testing accuracy being about 70%.

TABLE 3. Summary of the improvement for each reduction approach.

	CNN	AE	FTT	US
Best	35.71%	7.14%	-5.7%	0.0%
Average	32.38%	-0.86%	-13.57%	-8.93%
Worst	27.14%	-11.43%	-25.71%	-12.86%

Reduction Approach	HyperParameters	Precision	Recall	Accuracy of Training Set	Accuracy of Testing Set
Undersampling	US_c1	0,75	0,65	1	0,7
Undersampling	US_c2	0,6	0,6	0,96	0,61
Undersampling	US_c3	0,68	0,55	0,83	0,61
Undersampling	US_c4	0,68	0,65	0,76	0,63
Fourier Traformation	FTT_c1	0,61	0,68	0,66	0,66
Fourier Traformation	FTT_c2	0,57	0,68	0,6	0,63
Fourier Traformation	FTT_c3	0,47	0,57	0,6	0,52
Fourier Traformation	FTT_c4	0,55	0,62	0,7	0,61
Autoencoder	AE_c1	0,63	0,75	0,76	0,62
Autoencoder	AE_c2	0,7	0,75	0,8	0,68
Autoencoder	AE_c3	0,66	0,85	0,93	0,72
Autoencoder	AE_c4	0,6	0,85	0,96	0,7
Autoencoder	AE_c5	0,66	0,75	0,83	0,75
CNN	CNN_c1	0,92	0,84	1	0,94
CNN	CNN_c2	0,9	0,9	0,95	0,89
CNN	CNN_c3	0,87	0,87	0,94	0,95

FIGURE 12. Overview of the classification results.

It is possible to note that by applying the CNN-based reduction approach, the classification process improves by an average of 32.38% (Table 3).

The best configuration is the CNN_c3 which brings an improvement of 35.75%, achieving an accuracy classification of 95% (Figure 11).

The Autoencoder-based approach achieves the best improvement with the configuration AE_c5 increased by 7.14%, achieving an accuracy classification of 75%. The average improvement, however, is -0.83%.

The Under-sampling and Fourier approaches do not show any improvement in accuracy.

V. DISCUSSION

In this section, we will present some points for discussion relating to our solution emerging from the training and testing processes.

The experiments concerning the gait classification have brought to light some interesting points.

The CNN-based approach achieved the best results with a testing accuracy ranging from 89% to 95%.

The classification of the gait dynamics without applying any reduction approach achieved an average training accuracy of 94% and an average testing accuracy of 70%. Comparing these results with those of the CNN-based approach the improvement ranges from 27% to 35% (Figure 12).

The use of a reduction layer has brought a definite benefit to the overall classification process.

The drawback of the usage of CNN is the training time which is in the order or days

Additionally, the Autoencoder-based approach shows a less positive trend of improvement ranging from -11.43% to 7.14%.

This approach provides better results as the window size increases, but the larger the window size is, the higher the compression rate needs to be in order to reduce the input MTS.

This aspect makes the Autoencoder-based approach time-consuming and computationally heavy for the achievement of a low value of reconstruction error.

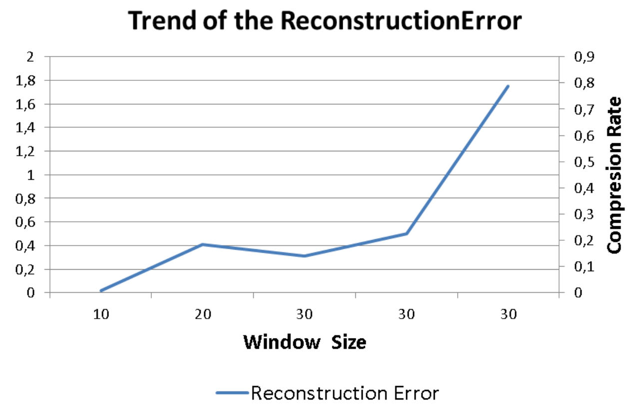

FIGURE 13. Tread of the reconstruction error.

Figure 13 shows the increasing trend of reconstruction error as the window size and compression rate increases.

Both the Under-sampling and Fourier approaches show negative improvements.

These results suggest an interesting consideration relating to the reduction process.

Although the effectiveness of the training of the LSTM is correlated to the number of time steps of the MTS [23], the results show that the property of a reduced MTS of being smaller than the original is not alone sufficient to guarantee a good classification performance

For example, all the approaches produce a reduced MTS with a length equal to 60 time-points but the improvement of the classification accuracy is quite different, ranging from -10% for the Under-sampling to $+27.14\%$ for CNN.

This evidence indicates that, notwithstanding the fact that the various reduced approaches are able to produce MTSs of comparable size, the classification accuracy varies according to the approach adopted.

A. LIMITATIONS

An honest discussion about the weaknesses of our solution is also presented.

1) The main issue in relation to our approach is the effect of the dimensionality of the TTS over the whole classification process.

The training time of the CNN-based reduction approach increases as the dimensionality increases. This behavior makes our solution time-consuming.

The total duration time of the experiments including both the reduction (using CNN or Autoencoder) and classification has been in the order of one week.

2) For the training and testing of our experiments we have used public data [11] produced by using accelerometers placed in the subject's shoes.

In order to realize our final objective, that of defining an in-house self-test mobile solution, we will set new experiments in which the data will be collected from the subject's smartphone.

This aspect is extremely challenging, since accelerometers embedded into a smart-phone do not have the same quality as those used in Hausdorff's dataset [11].

3) Another weak point is connected to the consideration that the classification process is performed off-line.

Our idea is to transfer the classification process onto a mobile device in order to make it available everywhere.

B. FUTURE WORK

The experiments performed have enabled us to learn a few lessons which will help us to define certain interesting research activities.

1) Defining a novel approach able to scale the training time of the CNN as the dimensionality of the TTS increases.
2) Defining a structured approach for the evaluation of the effectiveness of the reduction approaches.
3) Improving the dataset by collecting data from a larger population of patients. A larger set would enable us to improve the accuracy of our model.
4) Setting new experiments for the evaluation of our solution in a mobile environment in order to train and evaluate our model using the accelerometer values acquired by the sensors embedded in a smartphone.

VI. CONCLUSIONS

In this paper we have addressed the problem of the gait analysis of elderly subjects for the detection of the onset of a cognitive impairment disease such as PD.

In order to achieve this objective, we have defined a deep learning-based approach for the classification of the gait dynamics by using as input data a well-known public dataset.

The core of our solution is a hybrid Neural Network built by combining three types of NN algorithm: a Convolutional Neural Network (CNN) a Long Short Term Memory (LSTM) and a Deep Neural Network (DNN).

Our solution achieves a classification accuracy better than that of related works with an accuracy improvement of 3.9%.

The CNN proved to be the best algorithm for the reduction of the dimensionality of the input TTS, enabling the LSTM and DNN networks to produce a better classification of the input. The final objective of our research is to overcome the limitations of the traditional approaches by defining an easy in-house self-test mobile solution able to detect anomalies in the gait dynamics of elderly people.

In accordance with a specific test protocol, a patient will walk for a few meters inside his/her house, and a mobile app will acquire his/her movements by means of a smartphone's embedded accelerometers. These values will describe the gait dynamics of the patient (Figure 1) and will be analyzed to assess whether or not an anomalous walking pattern has been detected.

CONFLICT OF INTEREST

We confirm that we have no conflict of interest

REFERENCES

[1] S. Sveinbjornsdottir, "The clinical symptoms of Parkinson's disease," *J. Neurochem.*, vol. 139, pp. 318–324, Oct. 2016, doi: 10.1111/jnc.13691.

[2] J. M. Hausdorff, "Gait dynamics in Parkinson's disease: Common and distinct behavior among stride length, gait variability, and fractal-like scaling," *Chaos*, vol. 19, no. 2, p. 026113, May 2009, doi: 10.1063/1.3147408.

[3] G. Biagetti, P. Crippa, L. Falaschetti, G. Tanoni, and C. Turchetti, "A comparative study of machine learning algorithms for physiological signal classification," *Procedia Comput. Sci.*, vol. 126, pp. 1977–1984, 2018. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S1877050918312316, doi:10.1016/j.procs.2018.07.255

[4] A. F. Barbosa *et al.*, "Gait, posture and cognition in Parkinson's disease," *Dementia Neuropsychol.*, vol. 10, no. 4, pp. 280–286, Dec. 2016, doi: 10.1590/s1980-5764-2016dn1004005.

[5] D. Podsiadlo and S. Richardson, "The timed 'Up & Go': A test of basic functional mobility for frail elderly persons," *J. Amer. Geriatrics Soc.*, vol. 39, no. 2, pp. 142–148, Feb. 1991, doi: 10.1111/j.1532-5415.1991.tb01616.x.

[6] D. M. Wrisley, G. F. Marchetti, D. K. Kuharsky, and S. L. Whitney, "Reliability, internal consistency, and validity of data obtained with the functional gait assessment," *Phys. Therapy*, vol. 84, pp. 906–918, Oct. 2004, doi: 10.1093/ptj/84.10.906.

[7] F. B. Horak, D. M. Wrisley, and J. Frank, "The balance evaluation systems test (BESTest) to differentiate balance deficits," *Phys. Therapy*, vol. 89, no. 5, pp. 484–498, May 2009, doi: 10.2522/ptj.20080071.

[8] C. Ramaker, J. Marinus, A. M. Stiggelbout, and B. J. van Hilten, "Systematic evaluation of rating scales for impairment and disability in Parkinson's disease," *Movement Disorders*, vol. 17, no. 5, pp. 867–876, Sep./Oct. 2002, doi: 10.1002/mds.10248.

[9] J. A. Opara, W. Brola, M. Leonardi, and B. Blaszczyk, "Quality of life in Parkinsons disease," *J. Med. Life*, vol. 5, no. 4, pp. 375–381, 2012.

[10] I. Birch, W. Vernon, J. Walker, and M. Young, "Terminology and forensic gait analysis," *Sci. Justice*, vol. 55, no. 4, pp. 279–284, Jul. 2015, doi: 10.1016/j.scijus.2015.03.002.

[11] J. M. Hausdorff, A. Lertratanakul, M. E. Cudkowicz, A. L. Peterson, D. Kaliton, and A. L. Goldberger, "Dynamic markers of altered gait rhythm in amyotrophic lateral sclerosis," *J. Appl. Physiol.*, vol. 88, no. 6, pp. 2045–2053, 2000.

[12] D. Joshi, A. Khajuria, and P. Joshi, "An automatic non-invasive method for Parkinson's disease classification," *Comput. Methods Programs Biomed.*, vol. 145, pp. 135–145, Jul. 2017 doi: 10.1016/j.cmpb.2017.04.007.

[13] S. Shetty and Y. S. Rao, "SVM based machine learning approach to identify Parkinson's disease using gait analysis," in *Proc. Int. Conf. Inventive Comput. Technol. (ICICT)*, vol. 2, Aug. 2016, pp. 1–5.

[14] A. Khorasani and M. R. Daliri, "HMM for classification of parkinson's disease based on the raw gait data," *J. Med. Syst.*, vol. 38, no. 12, p. 147, Dec. 2014, doi: 10.1007/s10916-014-0147-5.

[15] C. Chakraborty and P. Talukdar, "Issues and limitations of HMM in speech processing: A survey," *Int. J. Comput. Appl.*, vol. 141, no. 7, pp. 13–17, May 2016, doi: 10.5120/ijca2016909693.

[16] P. Ren, W. Zhao, Z. Zhao, M. L. Bringas-Vega, P. A. Valdes-Sosa, and K. M. Kendrick, "Analysis of gait rhythm fluctuations for neurodegenerative diseases by phase synchronization and conditional entropy," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 24, no. 2, pp. 291–299, Feb. 2016, doi: 10.1109/TNSRE.2015.2477325.

[17] M. Yang, H. Zheng, H. Wang, and S. McClean, ''Feature selection and construction for the discrimination of neurodegenerative diseases based on gait analysis,'' in *Proc. 3rd Int. Conf. Pervasive Comput. Technol. Healthcare*, Apr. 2009, pp. 1–7.

[18] M. S. Baby, A. J. Saji, and C. S. Kumar, ''Parkinsons disease classification using wavelet transform based feature extraction of gait data,'' in *Proc. Int. Conf. Circuit, Power Comput. Technol. (ICCPCT)*, Apr. 2017, pp. 1–6.

[19] T. M. Rath and R. Manmatha, ''Word image matching using dynamic time warping,'' in *Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit.*, vol. 2, Jun. 2003, p. II.

[20] A. Bagnall, J. Lines, A. Bostrom, J. Large, and E. Keogh, ''The great time series classification bake off: A review and experimental evaluation of recent algorithmic advances,'' *Data Mining Knowl. Discovery*, vol. 31, no. 3, pp. 606–660, May 2017, doi: 10.1007/s10618-016-0483-9.

[21] T. G. Dietterich, *Machine Learning for Sequential Data: A Review*. Berlin, Germany: Springer, 2002, pp. 15–30, doi: 10.1007/3-540-70659-3_2.

[22] J. M. Hausdorff, C. K. Peng, Z. Ladin, J. Y. Wei, and A. L. Goldberger, ''Is walking a random walk? Evidence for long-range correlations in stride interval of human gait,'' *J. Appl. Physiol.*, vol. 78, no. 1, pp. 349–358, 1995. [Online]. Available: http://jap.physiology.org/content/78/1/349

[23] I. Goodfellow, Y. Bengio, A. Courville, and F. Bach, *Deep Learning*. Cambridge, MA, USA: MIT Press, 2016. [Online]. Available: http://www.deeplearningbook.org

[24] J. Orbach, ''Principles of neurodynamics. Perceptrons and the theory of brain mechanisms,'' *Arch. Gen. Psychiatry*, vol. 7, no. 3, pp. 218–219, 1962, doi: 10.1001/archpsyc.1962.01720030064010.

[25] Z. C. Lipton, J. Berkowitz, and C. Elkan, ''A critical review of recurrent neural networks for sequence learning,'' *CoRR*, vol. abs/1506.00019, Oct. 2015. [Online]. Available: http://arxiv.org/abs/1506.00019

[26] D. E. Rumelhart, R. Durbin, R. Golden, and Y. Chauvin, ''Backpropagation: The basic theory,'' in *Backpropagation*, Y. Chauvin and D. E. Rumelhart, Eds. Hillsdale, MI, USA: Lawrence Erlbaum Associates, 1995, pp. 1–34. [Online]. Available: http://dl.acm.org/citation.cfm?id=201784.201785

[27] A. Coronato, G. D. Pietro, and G. Paragliola, ''A situation-aware system for the detection of motion disorders of patients with Autism spectrum disorders,'' *Expert Syst. Appl.*, vol. 41, no. 17, pp. 7868–7877, 2014. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0957417414002917

[28] J. Park, D. Kim, C. Yang, and H. Ko, ''SVM based dynamic classifier for sleep disorder monitoring wearable device,'' in *Proc. IEEE Int. Conf. Consum. Electron. (ICCE)*, Jan. 2016, pp. 309–310.

[29] G. Paragliola and A. Coronato, ''A deep learning-based approach for the recognition of sleep disorders in patients with cognitive diseases: A case study,'' in *Proc. Federated Conf. Comput. Sci. Inf. Syst.*, vol. 12, Sep. 2017, pp. 43–48, doi: 10.15439/2017F532.

[30] G. Paragliola and A. Coronato, ''Intelligent monitoring of stereotyped motion disorders in case of children with Autism,'' in *Proc. 9th Int. Conf. Intell. Environ.*, Jul. 2013, pp. 258–261.

[31] K. Beyer, J. Goldstein, R. Ramakrishnan, and U. Shaft, ''When is 'nearest neighbor' meaningful?'' in *Database Theory—ICDT*, C. Beeri and P. Buneman, Eds. Berlin, Germany: Springer, 1999, pp. 217–235.

[32] J. F. Elder, IV, and D. Pregibon, ''A statistical perspective on knowledge discovery in databases,'' in *Advances in Knowledge Discovery and Data Mining*, U. M. Fayyad, G. Piatetsky-Shapiro, P. Smyth, and R. Uthurusamy, Eds. Menlo Park, CA, USA: American Association for Artificial Intelligence, 1996, pp. 83–113. [Online]. Available: http://dl.acm.org/citation.cfm?id=257938.257947

[33] M. Krawczak and G. Szkatuła, ''An approach to dimensionality reduction in time series,'' *Inf. Sci.*, vol. 260, pp. 15–36, Mar. 2014. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0020025513007743

[34] T.-C. Fu, ''A review on time series data mining,'' *Eng. Appl. Artif. Intell.*, vol. 24, no. 1, pp. 164–181, Feb. 2011, doi: 10.1016/j.engappai.2010.09.007.

[35] K. J. Åström, ''On the choice of sampling rates in parametric identification of time series,'' *Inf. Sci.*, vol. 1, no. 3, pp. 273–278, Jul. 1969, doi: 10.1016/S0020-0255(69)80013-7.

[36] C. Faloutsos, M. Ranganathan, and Y. Manolopoulos, ''Fast subsequence matching in time-series databases,'' in *Proc. ACM SIGMOD Int. Conf. Manage. Data (SIGMOD)*, New York, NY, USA, 1994, pp. 419–429, doi: 10.1145/191839.191925.

[37] ''Evaluation: From precision, recall and F-factor to ROC, informedness, markedness & correlation,'' *J. Mach. Learn. Technol.*, vol. 2, no. 1, pp. 37–63, 2011, doi: 10.9735/2229-3981.

GIOVANNI PARAGLIOLA received the M.Sc. degree in computer engineering from Federico II University in Naples. He is currently enrolled in the Ph. D. program in information engineering at the University Parthenope in Naples. He is also a Research Fellow at the Institute for High-Performance Computing and Networking, National Research Council. His research focuses on artificial intelligence, machine learning, and pervasive computing.

ANTONIO CORONATO received the M.Sc. degree in computer engineering from Federico II University in Naples. He is currently a Researcher at the Institute for High-Performance Computing and Networking, National Research Council, Italy. His research focuses on pervasive computing and component-based architectures. He is a member of the ACM.