IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Date of publication xxxx 00, 0000, date of current version xxxx 00, 0000.

Digital Object Identifier 10.1109/ACCESS.2017.Doi Number

Gait analysis for early neurodegenerative diseases classification through the Kinematic Theory of Rapid Human Movements

Vincenzo Dentamaro[1] Member IEEE, Donato Impedovo[1], Senior Member, IEEE, and Giuseppe Pirlo[1], Senior Member, IEEE

[1]Computer Science department, University of Bari, Italy

Corresponding author: Vincenzo Dentamaro, vincenzo.dentamaro@uniba.it

This work is within the BESIDE project (no. YJTGRA7) funded by the Regione Puglia POR Puglia FESR - FSE 2014-2020. Fondo Europeo Sviluppo Regionale. Azione 1.6 - Avviso pubblico "InnoNetwork".

ABSTRACT Neurodegenerative diseases are particular diseases whose decline can partially or completely compromise the normal course of life of a human being. In order to increase the quality of patient's life, a timely diagnosis plays a major role. The analysis of neurodegenerative diseases, and their stage, is also carried out by means of gait analysis. Performing early stage neurodegenerative disease assessment is still an open problem. In this paper, the focus is on modeling the human gait movement pattern by using the kinematic theory of rapid human movements and its sigma-lognormal model. The hypothesis is that the kinematic theory of rapid human movements, originally developed to describe handwriting patterns, and used in conjunction with other spatio-temporal features, can discriminate neurodegenerative diseases patterns, especially in early stages, while analyzing human gait with 2D cameras. The thesis empirically demonstrates its effectiveness in describing neurodegenerative patterns, when used in conjunction with state-of-the-art pose estimation and feature extraction techniques. The solution developed achieved 99.1% of accuracy using velocity-based, angle-based and sigma-lognormal features and left walk orientation.

INDEX TERMS pose-estimation, computer vision, computer aided diagnosis, gait analysis, machine learning, early neurodegenerative diseases assessment, kinematic theory of rapid human movements, sigma-lognormal

I. INTRODUCTION

Neurodegenerative diseases are incurable diseases whose decline can partially or completely compromise the normal course of life of a human being.

In order to increase the quality of patient's life, a timely diagnosis plays a major role. Usually physicians asses the disease by letting the patient performing several tasks, such as Timed Up And Go (TUG) Test, Tandem Test, Sit to Stand, ADL (Activities of Daily Living), IADL (Instrumental Activities of Daily Living), Romberg and many more, in which, usually, the patient is asked to walk on a straight line or sit and standup, try to remain immobile in balance and so on.

These tests are important, because neurodegenerative diseases act by destroying several parts of the nervous system and showing various disorders in also other parts of human body such as speaking, posture, coordination, handwriting and many more.[1][6]

Gait is, thus, a complex activity which involves cognitive, kinesthetic and perceptual-motor components [6]. Gait is known to be one of the foundation movements of human being and it is carried out from the first years of life. It requires the use of an ensemble of resources of the nervous system to bring it to completion.

The accuracy of the analysis carried out depends greatly on the quality of the tools used. In past centuries, the analysis of the gait was carried out mainly through the careful observation of experts, who were able to extrapolate qualitative and quantitative data from it (such as the cadence, the speed of the pace and the distance traveled). The data obtained were, certainly, valid for recognizing serious gait disturbances, but proved inadequate to recognize small variations in the progression of diseases. This prevented early and precise diagnoses, identifying the actual severity of a given disease and carefully planning the next treatments.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/ACCESS.2020.3032202, IEEE Access

Up to now, some Computer Aided Diagnosis (CAD) tools for neurodegenerative assessment, have shown to be effective. Those tools are based on behavioral biometrics [1]. One of main examples of CAD tools for neurodegenerative disease assessment is handwriting [2], [3] analyzing both neuromuscular system parameters [4], and spatio-temporal models [5]. In particular, as reported in [1],[2],[4],[5] the sigma-lognormal (ΣΛ) model [8] of the kinematic theory of rapid human movements [9],[10] was employed, with successful results, in the early prediction of Parkinson Disease through handwriting [5].

In the review presented in [2] authors subdivide the state of art in neurodegenerative disease classification problem with handwriting into two main categories: computational and cognitive. Our focus is mainly on computational models. Computational models try to model or reconstruct the final result of movements, in terms of velocity and acceleration profiles or stroke shapes by means of mathematics, physics and computer science. One example of this computational model is the kinematic theory of rapid human movements. This theory is defined in terms of two elements: the agonist (acting in the direction of the movement) and the antagonist (acting in the opposite direction) neuromuscular systems involved in the production of rapid human movements.[9][10]. Plamondon showed that these kinds of system have a log-normal impulse response that results from the limiting behavior of a large number of interdependent neuromuscular networks controlling the velocity of a movement [1],[2],[8],[9],[10]. An evolution of this theory uses the sigma-lognormal model (ΣΛ). The effectiveness of this theory in predicting Parkinson Diseases through handwriting has been already demonstrated [1],[2],[4],[5].

The hypothesis at the very basis of this work is that transposing the Plamondon's kinematic theory of rapid human movements and its sigma-lognormal model (ΣΛ) from handwriting to gait analysis, it will still be possible to almost perfectly reconstruct and thus, model, the movement patterns of various body joints in terms of acceleration and velocity profiles. The intuition is that, during gait, several body parts share similar motor functions and neuromuscular parameters according to the sigma-lognormal model. Thus, these movements are dictated by the brain and are not involuntary.

These built models will be used, in conjunction with other synthetized features, to classify if a person 'gait is affected by some neurodegenerative disease or is healthy. As it will be shown later, the sigma-lognormal model will be used specifically for capturing small variations in acceleration profiles able to discriminate borderline subjects in early stages of the disease with the consequent increase of the overall accuracy of the developed decision support system.

Pose estimation technique in conjunction with Barrel distortion removal algorithm and Kalman filter will be used to fix various problems discovered and thus, with the best possible fidelity, extract the coordinates of body joints.

The main contributions of this work are the following:

1. The first application of kinematic theory of rapid human movements to neurodegenerative disease assessment through gait and commercial 2D cameras
2. Innovative pipeline comprehensive of gait phases segmentation
3. Outstanding state of the art results and, thanks to the interpretability capabilities of the developed system, important findings on the body parts that play a major role in neurodegenerative diseases assessment through gait and computer vision

The use of off the shelf 2D cameras for pose estimation is twofold: firstly, its cheap compared to infrared cameras, depth cameras or 3D tracking wearable sensors, in second instance, the system can easily be deployed in every hospital's room or at patient's home. This is important when it comes to remote monitoring, because there is no need from the patient side to wear anything, avoiding problems such as forgetfulness and rejection, but also for hospitals to reduce hospitalization costs.

The paper is organized as follows. Section II describes neurodegenerative diseases and motor patterns in gait analysis from a pure pattern recognition perspective. Section III shows a synthetic review on the techniques used in human gait analysis. Section IV depicts the kinematic theory of rapid human movements and its sigma-lognormal model (ΣΛ). In addition, the intuition behind the use of sigma-lognormal model in gait analysis is also provided. Section V describes the dataset used. Section VI sketches the experiment, implementation details and results. Section VII provides a comprehensive discussion from a purely pattern recognition perspective. Conclusions and future work are presented in section VIII.

II. NEURODEGENERATIVE DISEASES MOTOR PATTERNS

Neurodegenerative diseases affect long neuronal cells. The affected cells create physical disorder during the walk. Some neurodegenerative diseases are Parkinson's disease (PD), Amyotrophic lateral sclerosis (ALS), Alzheimer disease (AD), the Huntington Korea (HD) and various forms of Dementias (DD) [11]. PD patients show slow automatic movements and slow balance. The symptoms are body rigidity with hypertonia, bradykinesia plus akinesia and lack of balance, especially when the disease is severe [12].

ALS shows an evolutionary muscular atrophy with decrease in strength, with phonation and chewing disorders [13].

AD neuronal damages affect the short-term memory comprehension and thus patients usually forget important information. Also, the walk pattern is affected: AD patients show uncoordinated movements with erect standing and walking, dysarthria, alteration of breathing, facial grimaces, dysphagia and hyperkinesia [14].

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/ACCESS.2020.3032202, IEEE Access

DD represents a group of typical neurodegenerative diseases of old age, with irreversible loss or reduction of intellectual abilities. The HUN is a hereditary disease in which cognitive and motor skills are particularly compromised. The first clues are mood changes, memory loss, dementia, difficulty in walking, language and swallowing, strong depression in advanced stages and all its consequences.[14]

The Gait Analysis studies the ways human Gait cycle is a succession of physical actions involved during walking. Formally, gait cycle is defined as the interval between two successive occurrences of the same event. Literature reports two ways of measuring gait cycle from a computational perspective. These methods make use of Temporal/Spatial features or Pressure measurements. In this paper gait cycle is analyzed with respect to the first method (temporal) along with its gait patterns [15], [16], [17]. This method is, by far, the most common and scientifically validated by neurologists and for which there are well-defined protocols (e.g. TUG test).

Phases of a gait cycle include:

- Initial contact (IC): when the foot touches the ground;
- Loading response (LR): when the other foot is lifted for the swinging;
- Mid Stance (MS): the swinging foot exceeds the foot that acts as a lever;
- Terminal stance (TS): the right foot's heel moves vertically until the left foot touches the ground;
- Pre-swing (PS): now the left foot acts as a lever allowing the right foot to walk in
- Initial Swing: the hip, knee, and ankle are flexed to begin advancement of the limb forward and create clearance of the foot over the ground.
- Mid-swing (MS): the left leg's tibia is vertical so that right leg can overcome it;
- Terminal swing (TS): the progress of the limbs is completed when the right leg moves in front of the left thigh and the right foot touches the ground, going back to the IC phase.

This gait cycle will be used in section VI to describe the segmentation phase.

Gait abnormalities show deviations from normal gait patterns and are necessary to assess a specific neurodegenerative disease [18].

The gaits abnormal patterns presented in [18] are important to understand, from a pattern recognition perspective, the features to be employed in the analysis as well as their connection with the disease. From this perspective, it is clear that spatio-temporal features, as well as kinematic features play an important role in discriminating healthy subjects from subjects with some form of neurodegenerative disease. Of no less importance is the analysis of which parts of the body to observe with greater importance. It seems natural to impute the lower parts of the body as legs, ankles and feet as the most important parts for gait analysis. As it will be shown later in the discussion provided in Section VII, nonlinear correlation among features, hardly observable with the naked eye, results in a rise of importance of other analysed body parts.

III. LITERATURE REVIEW ON TECHNIQUES USED IN HUMAN GAIT ANALYSIS

In this section, literature review is performed with respect to data acquisition, preprocessing, tasks and mostly related works.

A. DATA ACQUISITION

Several authors have used different types of sensors for real-time data acquisition of human gaits. It is possible to divide the sensors in three main categories [19]:

1. Wearable sensors on the patient's body.
2. Floor Sensors deployed on the floor, usually a matrix of pressure sensors.
3. Cameras able to capture video information.
4. Optical Motion Capture with model based optoelectronic motion capture system [57].

Wearable sensors

Accelerometers are used for measuring the acceleration of the body [20]. The gyroscope measures the angular velocity, thus is used for measuring human posture by analyzing patient's feet, leg and torso angulation [19]. Magneto-resistive sensors estimate the change of orientation in relation to magnetic North. Inertial sensors measure body's parts velocity, acceleration, orientation, and gravitational forces using what is called sensor fusion techniques based on accelerometers, gyroscopes and magnetometers. Flexible goniometers are used for measuring angles of ankles, knees, hips and metatarsal.

Sensitive tissues are a mix of detection technology and electronic devices made by fabric materials. This type of sensor ensures a good level of comfortability because patients wear them without particular problems. Authors in [20] have integrated force sensors into shoes in order to detect clinical ground forces (GRF) measures during gait. Electromyography is used for detecting the activation of one or more muscles and it is important to evaluate patients with lower lib problems.

Floor Sensors

Force or pressure sensors deployed on the floor create the "force platforms". These platforms collect, in a matrix format, differentiated pressure measurements of each part of the foot separately [21].

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/ACCESS.2020.3032202, IEEE Access

Video-Image Sensors

One or more cameras are used in conjunction with various image processing techniques for removing noise and preserving relevant information. Techniques such as threshold filtering and background segmentation allow for the separation of background from the silhouette of the human body [19], [21]. Among the most successful techniques, the Infrared Thermography, which make use of temperatures of human body, is able to reach accuracy rates of about 91% [22]. In recent years, human pose estimation from video clips has been another emerging field of gait analysis. These techniques use deep learning techniques to extract information about body parts from video frames. [23]

Optical Motion Capture

The typical optical motion capture system uses, in general, about six calibrated infra-red (IR) cameras which tracks about 44 reflective markers to be attached to the human body. These markers are necessary to extrapolate kinematic features from the body joints and thus provide spatio-temporal features for gait assessment [58]. Such systems can be used only in laboratories. This is due to the setup, but also to the equipment, the need for calibration and trained operators.

B. PREPROCESSING

The pre-processing phase usually involves trimming of non-relevant information as well as normalizing the environment. This last aspect is relevant when doing 3D image reconstruction because there is no zero-point in a three-dimensional space. In order to overcome this problem, authors in [27] and [31] performed a series of "manual" calibrations in collaboration with patients. In [31] the gait analysis starts when the patient touches the force platform with the foot. In [41] measurements start after the patient has walked on to a starting platform located before the force platform used for real data capturing, while in [37] pre-processing cuts noise. Especially in Silhouette analysis, pre-processing techniques are used to remove noise, background information and in some image processing cases, overlapping obstacles [21], [29], [30], [32], [33], [38], [39], [40], [42].

C. WALKING TASKS

Most part of literature considers only healthy subjects. Only a small part of works, such as [24], [25] and [26] kept into consideration both healthy and un-healthy people (patients affected by PD).
Another common factor among publications is that authors performed tests with non-conventional protocols [24], [27], [28], [29], [30]: for example, authors in [30] trained some subjects to perform tests holding a ball as well as wearing a coat; authors in [28], instead, asked patient to move the leg up and down 5 times while a 3D ultra-wideband receiver estimated the position of the leg during time. A very interesting task was done by authors in [31] where 5 males and 5 females subjects were asked to perform overground walking (30 trials), stair ascent (30 trials) and stair descent (30 trials), while their body parameters were collected. In a similar fashion, authors of work [26] asked students and professors as well as patients affected by PD to perform tests such as: random walking for 2-4 times, lifting a stair starting from being sitting. Another interesting work is presented in [32] where authors used Microsoft Kinect to extract patients' skeleton by performing 25 meters on walkways and then stand in front of the Kinect with 3.5 meters distance and then performing swings on one side of the body. Similar approach was used by authors in [33]. In general, a common practice is to ask subjects to perform walk for different meters: 12 meters in [27], 3 to 7 meters in [24], alternate walk and run [34] and [25], perform walk for only 20 seconds [35] or perform random walk [26]. Depending on the capturing sensor, the subjects were asked to walk on custom force platform [36], or run in front of a camera [37], [38], walk while being captured by 24 cameras all positioned with different angles in order to reconstruct the 3D movement of the person [39], perform random walk in front of the camera [29] and walk straight in front of a camera [40].

D. RELATED WORKS

In [43] authors used a pose estimation algorithm called Convolutional Pose Estimation, which used Deep Learning and Convolutional Neural Networks. This algorithm was used to extract individual joints movement trajectories and synthetize several features (e.g. kinematic, frequency). These features were then used to train random forests to detect and estimate the severity of parkinsonism and LID. Binary F1 score with an interpatient [6] separation scheme was 0.906.

Ye et al. [44] observed that patients' gait dynamic is non-linear, for this reason, they proposed an Adaptive Neuro-Fuzzy Inference System (ANFIS) which combined neural network adaptive capabilities and the fuzzy logic approach. A reduced set of features were adopted: left stride interval, right stride interval, left stance interval, right stance interval, and double support interval. An interpatient separation scheme was used for carrying out the binary classification. Accuracy ranges from 90 to 94%. At the same time, standard classifiers were adopted observing comparable results to those obtained by authors. In [45] authors used Gaussian radial basis function and SVM in conjunction with a similar separation scheme proposed in [44] to predict Parkinson's disease patterns from human gait with accuracy of 83.3%.
In [47] authors built a decision support system (DSS) with the aim of helping medical doctors in diagnosing the PD influenced patients. Authors made use of grid search optimization to develop an optimized deep learning model to

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/ACCESS.2020.3032202, IEEE Access

predict the early onset of Parkinson's disease by tuning several hyperparameters of the model. The grid search optimization consisted in optimization of the deep learning model topology, the hyperparameters, and the overall performance. The resulting binary classification accuracy was of 91.69%.

In [48] authors used 3D body pose estimation technique using Deep Neural Network. The resulting 3D coordinates of joints time series are then fed in a classifier, with the aim of classifying healthy subjects, Parkinson's disease patients, post stroke patients, and patients with orthopedic problems. By using deep learning, no feature engineering was used, differently from the solution proposed in this work, but also the interpretability of their model was limited. Their classification accuracy ranged from 56% to 96% for different groups. The average accuracy was 71.25%.

Since neurodegenerative disease recognition from human gait is within the broader area of activity recognition in healthcare, it has been decided to compare results with state of the art shallow learning activity recognition techniques such as motion blobs [59] and optical flow [60]. In [59] authors analyzed human activities in sequence of frames. Firstly, they performed background subtraction in the video stream. Then, authors computed motion blobs of the current frame and a series of frames before the current frame to form a new feature image in certain rules. Finally, authors combined the non-zero pixels in the feature image into blobs using the connected component method. At the end, authors used Gaussian Mixture to model features and used standard classifiers to ensure the accuracy. The second technique used for comparison is the optical flow. Authors in [60] analyzed the use of local descriptor built by optical flow vectors along the edges of the human silhouette. As before, they first removed the background, then computed the optical flow on the silhouette along the boundary lines. From the centroid, radial lines are drawn at 5k degrees to intersect the boundary lines. The radial distances of the boundary points that lie on these radial lines along with the optical flow vectors computed at these points and are later used for classification through rbf-svm. For a more in depth analysis deep learning has been investigated. The 3D ResNet, a ResNet that captures 3D spatio-temporal relation has been used for comparison purposes as shown in [61].

IV. THE KINEMATIC THEORY OF RAPID HUMAN MOVEMENTS

The Kinematic Theory of Rapid Human Movements can be defined as an instrument to analyze handwriting movements as a statistical process that leverages on neuromuscular parameters of human body and brain [8], [9].

At the basis of this theory there is the intuition that any movement (movements of elbow, wrist, but also arms, legs, and so on), is the combination of primitives, called strokes, whose velocity and acceleration profile is lognormal [49].

Following the Sigma-lognormal model in equation (1), the velocity profile of each stroke j has a lognormal shape Λ which is scaled by an input command D and time-shifted by the instant the command t_0 starts:

$$\left|\vec{v}_j(t)\right| = \frac{D_j}{\sigma_j(t-t_{0j})\sqrt{2\pi}} exp\left(-\frac{[ln(t-t_{0j})-\mu_j]^2}{2\sigma_j^2}\right) \quad (1)$$

μ_j is the log-time delay generated by the action plan of the neuromuscular system, while σ_j is its log-response time. The model assumption is built around the fact that each stroke happens around a pivot making use of a starting angle θ_{sj} and an ending angle θ_{ej} as shown in equation (2)

$$\theta(t) = \theta_s + \frac{\theta_e-\theta_s}{D}\int_0^t \left|\vec{v}_j(t)\right| d_t \quad (2)$$

Equation (2) reveals that each movement could be composed by several lognormal strokes in continuous domain.

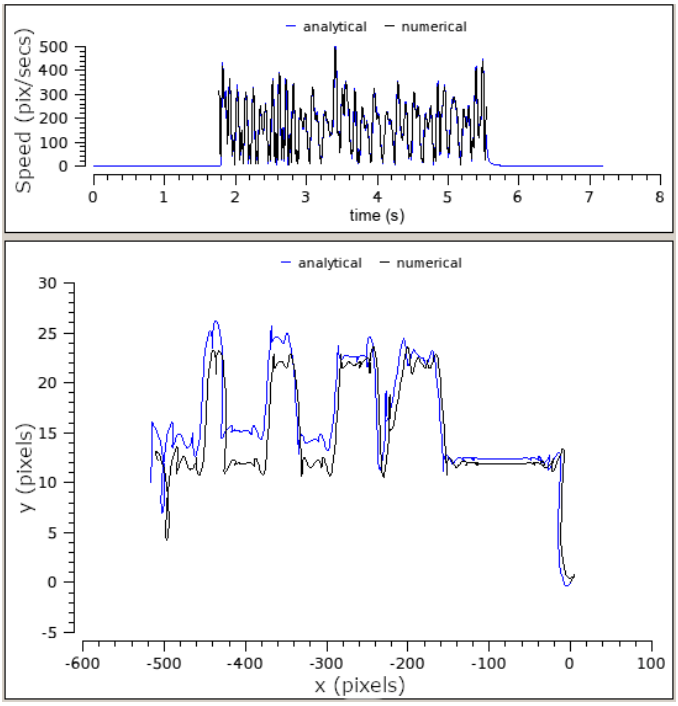

FIGURE 1. Sigma-lognormal velocity profile of healthy subject' nose. The blue line is the reconstructed sigma-lognormal profile, the black line is the original signal' coordinates of subject 'nose while walking.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/ACCESS.2020.3032202, IEEE Access

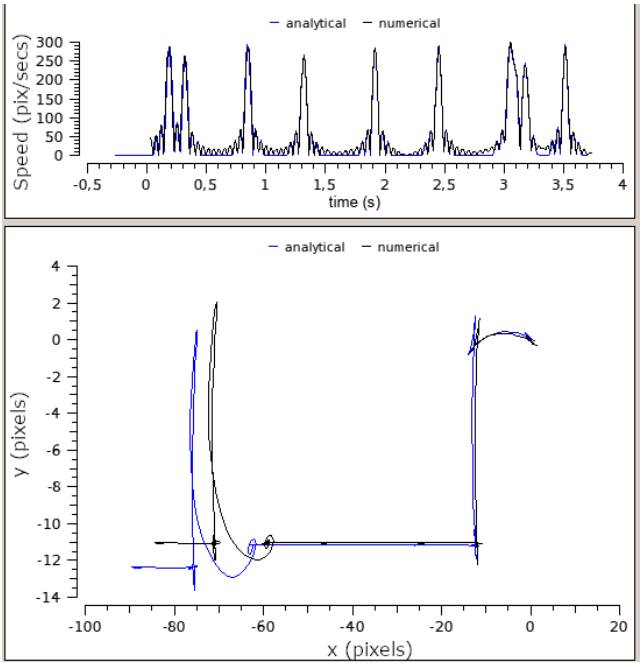

FIGURE 2. Sigma-lognormal velocity profile of Parkinsonian patient' nose. The blue line is the reconstructed sigma-lognormal profile, the black line is the original signal' coordinates of patient while walking.

Comparing the velocity profile (the upper side of the image with time as abscissa) in Figure 1 and Figure 2 it is possible to assert that:

- The sigma-lognormal almost perfectly reconstruct the signal from the velocity profile. Its reconstruction signal to noise ratio (SNR) is of about 17-35 dB in average.
- The patient's nose walk (Fig.2) is composed of several small oscillations after a high spike, instead intensity and numerosity of spikes in normal subject (Fig. 1) are almost always in a predefined range without such high variations.

The intuition behind this work is that, using the Kinematic Theory of Rapid Human Movements for capturing such small variations, it is possible to especially discriminate borderline subjects in early stages of the disease and thus add important information to the plethora of features used in reviewed works. Without loss of generality, an example of this intuition is that, features such as mean acceleration or mean displacement would smooth out such small oscillations, instead the sigma-lognormal model, if used correctly, would put them back in the game.

The Sigma-Lognormal model has been already successfully used for detecting neurodegenerative diseases thorough handwriting [1], [2], [4], [5].

V. DATASET DESCRIPTION

The reference dataset contains 115 videos regarding total 40 subjects divided in patients and controls as follows:

- 20 healthy control subjects
- 20 patients with neurodegenerative diseases (with different severity levels)

21 subjects are female, 19 have male sex.
Overall, the dataset has:

- 61 videos related to healthy subjects
- 54 videos related to patients.

The videos are variable in length and were made at different times and structures. Each video presents a patient's walk, which follows a linear path in both directions, from left to right and from right to left. All videos were fixed at 25 fps.
The videos were made using the following guidelines reassumed in Figure 3:

- The subjects traveled a fixed distance of 4 meters following a traced route, represented by a straight line highlighted on the floor.
- The videos were recorded with a camera placed perpendicularly 4 meters away from the straight line highlighted on the floor and at a height of 2 meters (or higher where needed).

Some video presented a fish-eye effect that was mitigated as reported in section VI.
The age of the patients is unbalanced: it was difficult to find subjects in their 80s who were perfectly healthy. Healthy subjects aging ranged [30-75] while patients aging ranged [65-90].
The dataset is also accompanied with the following metadata:

- The ID of the patient
- The sex of the patient
- The neurodegenerative disease stage which could be one of the following:
    1. Mild
    2. Found
    3. Moderate
    4. Severe

In this work, the classification performed is binary (healthy/unhealthy), but such disease' stages, as will be discussed in section VIII, will be used to evaluate the importance of sigma-lognormal features in discriminating borderline subjects, especially subjects with a neurodegenerative disease in mild or found stage from healthy control subjects. The neurodegenerative disease stage assessment was performed with several cognitive tests performed by trained and experienced psychologists and neurologists.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/ACCESS.2020.3032202, IEEE Access

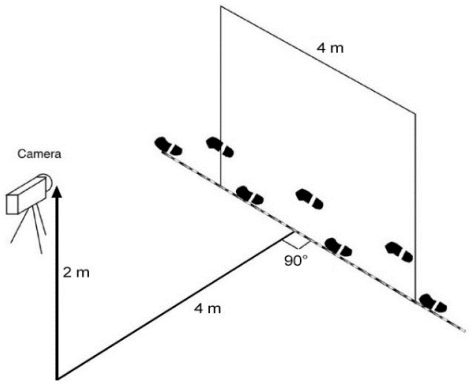

FIGURE 3. **The camera setup for video acquisition**

Some video presented mirrors and various kind of obstacles (chairs, body of the physiotherapist and so on) which added lots of complexity to the overall computation pipeline depicted in Section VI.

VI. EXPERIMENT AND IMPLEMENTATION DETAILS

A. PIPELINE & PREPROCESSING

The classification pipeline is depicted in Figure 4. The necessary phases are, in order, the lens-distortion algorithm to remove the fish-eye effect (also known as Barrel distortion) [50] present in some videos followed by the subject selection phase. This second phase is necessary when there are more people in the video. In almost all patient' videos, physiotherapist was present to supervise the patient during his/her performance. The third phase is the pose estimation, where for each frame of each video, joints coordinates of the body parts of the analyzed subject were extracted.

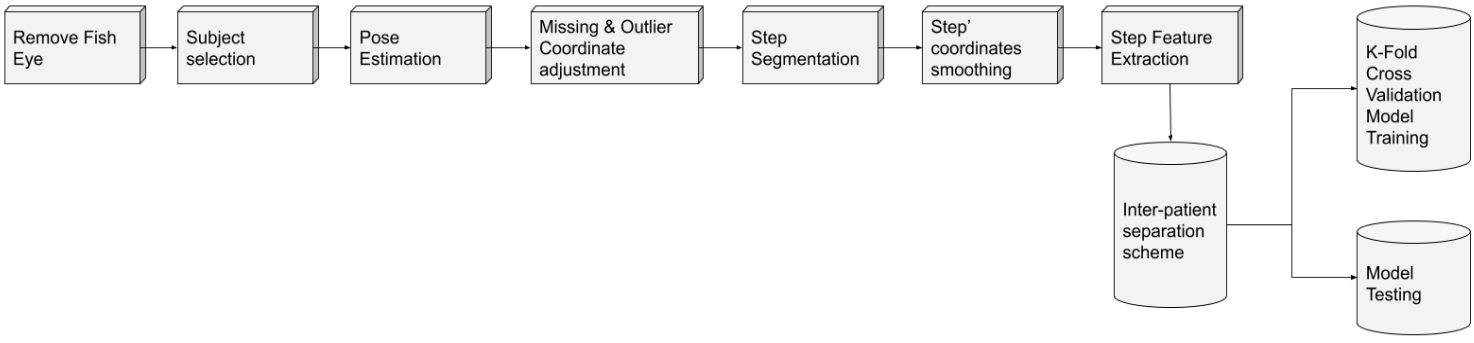

FIGURE 4. **The classification pipeline**

The fourth phase is necessary to fix some missing data and outlier present in extracted coordinates and caused by occlusion, pose estimation errors due to different light conditions and variations and colors' overlay. The fifth phase is of paramount importance: the step segmentation phase. Segmented steps will be smoothed in the sixth phase to remove unnatural and implausible coordinates. In the eight phase, for each segmented step, spatio-temporal features, sigma-lognormal features as well as angles features will be extracted. Various pattern recognition algorithms were trained with 10-Fold cross validation on an inter-patient separation scheme. In order to extract the relevant patterns that are able to discern subjects with some kind of neurodegenerative disease and healthy subjects, it is necessary to apply an inter-patient separation scheme as presented in [6]. The inter-patient scheme uses extracted features of some people 'steps for training and extracted features of completely other people 'steps for testing. This separation scheme is better suited for medical purposes, this is because, firstly, the model learnt must answer the question "what is the particular pattern of people with some form of dementia?" correctly, second because otherwise there would be bleeding of information from training to testing and third, but not of less importance, because without an inter-patient separation scheme, the i.i.d. (independently and identically distributed) assumption among instances, in this specific case, is not achievable.

Fish-Eye effect removal

For the fish-eye removal (Barrel distortion) the algorithm presented in [50] was used. Empirical parameters were estimated by fixing field of view (FoV) at 75.0. Results are shown in Figure 5.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/ACCESS.2020.3032202, IEEE Access

FIGURE 5. **Before and after application of fish-eye removal algorithm**

Subject Selection

In phase two, the user is asked to select with a click of the mouse the central point of the hips, which can be approximated to the center of gravity, of the subject protagonists of the analysis within the video. This procedure is necessary for both the start frame of the right walk and the start frame of the left walk. For all other frames, the algorithm seeks the center of gravity, which is closest, in Euclidean sense, to the position of the previous frame center of gravity. The subject selection window is presented in Figure 6.

FIGURE 6. **The first frame where the operator is asked to select the subject to analyze.**

Pose Estimation

For this task, OpenPose 1.6.1 was used. This version brought an improvement with respect to the previous especially in false positive detection rate. In general, the system takes as input a colored image and returns the 2D pixels coordinates of the anatomical key points of the people with respect to the image frame. Initially, a feedforward network forecasts 2D confidence maps of the positions of body parts and a set of 2D vectors for the affinity parts (which essentially describe the degree of association between the parts). At the end, each confidence map is converted to "greedy inference", returning the key points of the body parts within the image. [23]
There is no need of a calibration phase for performing pose estimation through OpenPose[23].
Microsoft COCO [51] image dataset training was used. The used body parts are listed in Table I.

TABLE I
MICROSOFT COCO BODY PARTS

Body Part	MS COCO ID
Head	0
Nose	1
Neck	2
Shoulder Right	3
Right Elbow	4
Right Wrist	5
Left Shoulder	6
Left Elbow	7
Left Wrist	8
Right Hip	9
Right Knee	10
Right Ankle	11
Left Hip	12
Left Knee	13
Left Ankle	14

Coordinates missing and adjusting

The system proceeds with the elimination of peaks from the data sequences: since walking is a continuous and harmonious movement, it is unthinkable that the extracted data will report sharp jumps forward or backward. Therefore, having established a threshold equal to 50 pixels and considering the position of each part of the body in correspondence with the first frame of each walk, all those measurements whose ordinate component y go out of the interval defined by the threshold [y - 50, y + 50] were removed. In this way, any errors made during the acquisition or isolation of the subject are eliminated. Figure 7 presents the result of this operation. Once the peak removal operation is complete, the system takes care of estimating the missing coordinates of each part of the body by linearly interpolating them both in cases where these are not visible within the video, and in cases where they were eliminated following the peak elimination process. The result is depicted in Figure 8. This operation is essential because it allows to create a database of homogeneous information: each part of the body will have the same number of points identified.

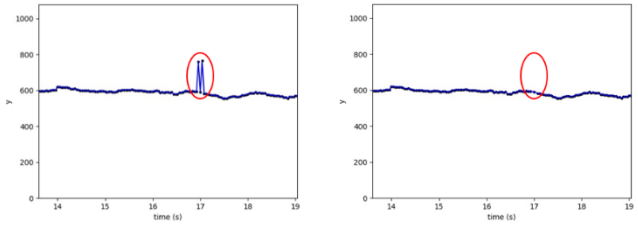

FIGURE 7. **On the left, the hip pixels coordinates before jump removal algorithm, on the right after the application of the removal algorithm.**

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/ACCESS.2020.3032202, IEEE Access

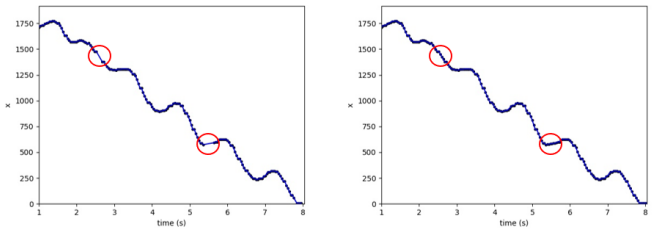

FIGURE 8. On the left, nose pixels coordinates before linear interpolation on the right after the application of the linear interpolation

Step Segmentation

Once the data modification phase is completed, the system proceeds with a new phase: recognition of the individual steps. To divide the entire walk into single steps, the system identifies stance phases discussed in Section II. In these phases the ankle remains still: these phases represent the initial part of a step and are followed by a swing phase, in which the ankle moves. To proceed with the identification of the stance phases, the system extrapolates from the entire database the x and y coordinates of the right and left ankles. Before being able to isolate the stance phases with enough accuracy, it is necessary to subject the ankle coordinates to further changes: the elimination of burrs. Despite the fact that the accuracy of data acquisition by OpenPose is quite high, it is inevitable that small smudges will be generated in the data, which would make it impossible to recognize the stance phases. The system then proceeds to analyze blocks of frames corresponding to a duration of 0.15 seconds (empirically estimated time as the minimum necessary to identify a stance phase), to check if there are any burrs, or if the central elements of the block have small variations compared to the elements that act as extremes for the considered block. If the system recognizes a burr, it proceeds to level the differences, replacing the values of the central elements of the analyzed block with the value present at the ends. Once the burr elimination procedure is over, the system can finally take care of recognizing the stance phases, saving the starting frame and the duration in the number of frames.
Once all the stance phases were identified, the system proceeds with further analyzes aimed at eliminating any of the following situations:

- The system combines two successive stance phases, if they are not interspersed with a swing phase: it could in fact be a slight movement of the foot during the stance phase that pushed the system to recognize two different phases instead of one.
- The system combines two phases of successive stances, if the starting point of the second stance coincides with the end point of the first stance: the fake swing phase recognized by the system could, in fact, derive from a data acquisition error that pushed the system to recognize a definitely wrong swing phase as it does not produce displacement.
- The system eliminates the first stance if it coincides with the start of the right or left walk, in order to avoid considering an incomplete stance phase.

Once the analysis of the identified stance phases was completed, the system proceeds with the identification of the individual steps: a step is identified as the interval between the beginning of a stance phase and the beginning of the next.

Step's coordinate smoothing

The sixth phase is the application of the Kalman filter [52] to the timeseries of data, of which every step is taken. This is necessary because consecutive frames, may suffer of high coordinate variance: key point of same body part in two consecutive frames, could have high variance due to small inaccuracies of the pose estimation algorithm.
Kalman Filter is an important and widely used estimation algorithm. The Kalman Filter works by computing estimates of hidden variables with the hypotheses that these variables are inaccurate and under uncertain measurements. The aim of the algorithm is to predict the future state of the system by using the past estimations. The Kalman filter is used in control system, tracking, space navigation and so on. [52], [53]
Given the entire sequence of joint coordinates of a segmented step, the Kalman filter, through 5 iterations of the estimation operation, outputs the smoothed sequence of coordinates. This task is performed for all key points. Results are shown in Figure 9.

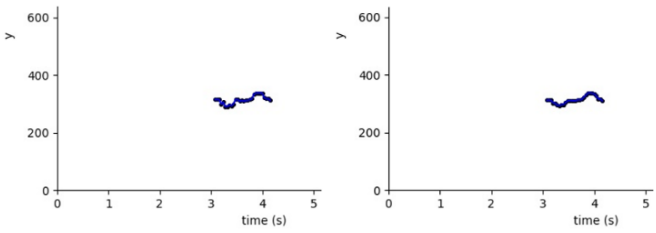

FIGURE 9. On the left, step's coordinate before Kalman smoothing and after on the right.

B. FEATURE EXTRACTION

The features are calculated on each analyzed step, distinguishing between right and left steps. For each step, all the features are calculated for each frame that makes up the video fragment relating to the analyzed step. Finally, after all the features were calculated, the timeseries containing the measurements relating to each feature are subjected to 5 statistical measurements shown in Table II (mean, median, standard deviation, first percentile and ninety-ninth percentile), which are then used by the system for recognition and classification phase. This is because, as said previously, features are computed on timeseries of joints coordinates of each segmented step.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/ACCESS.2020.3032202, IEEE Access

Features can be divided into three major macro categories: spatio-temporal features, angles, sigma-lognormal features.

TABLE II
STATISTICAL MEASURES

Feature name	Formulation
Mean	$\bar{x} = \frac{x_1 + x_2 + .. + x_n}{n}$
Median	given $x = [x_1, x_2, .., x_n]$ thus $\mu = x[\frac{n}{2}]$
Standard Deviation	$\sigma_X = \sqrt{\frac{\sum_{i=1}^{N}(x_i - \bar{x})^2}{N}}$ where $\bar{x} = \frac{1}{N}\sum_{i=1}^{N} x_i$
1 and 99 percentile	$n = [\frac{P}{100} \times N]$

TABLE III
SPATIO-TEMPORAL FEATURES

Feature name	Formulation
Displacement	$d_i = \sqrt{\Delta x_i^2 + \Delta y_i^2}$
Displacement x	$\Delta x_i = x_{i+1} - x_i$
Displacement y	$\Delta y_i = y_{i+1} - y_i$
Velocity	$v_i = d_i/\Delta t_i$
Velocity x	$v_{x,i} = \Delta x_i/\Delta t_i$
Velocity y	$v_{y,i} = \Delta y_i/\Delta t_i$
Acceleration	$a = v_i/\Delta t_i$
Acceleration x	$a_{x,i} = v_{x,i}/\Delta t_i$
Acceleration y	$a_{y,i} = v_{y,i}/\Delta t_i$
Tangent angle	$\rho_i = \tan^{-1}(\Delta y_i/\Delta x_i)$

Spatio-temporal features are shown in Table III. These features are synthetized from the timeseries of coordinates of each segmented step. These features are reassumed using the statistical measures used in Table II.

By carefully viewing the walking of patients with very serious neurodegenerative pathologies and referring to the literature reviewed regarding abnormal gait and reported in Section II, the following angles were computed:

- nose, neck, hip (right or left), this because patients tend to walk with their heads more inclined downwards;
- neck, hip (right or left), knee (right or left), this because patients tend to walk with their torso more inclined forward;
- shoulder (right or left), elbow (right or left), wrist (right or left), this because patients tend to walk with their arms much more curled up;
- hip (right or left), knee (right or left), ankle (right or left), this other because patients tend to walk with their knees more bent;
- knee (right or left), hip (right or left), knee (left or right) this also because patients tend to take smaller steps, with a much smaller opening of the legs.

For angles computation, the formula provided in equation (3) was used.

$$\gamma = deg(arc\cos\left(\frac{\overrightarrow{AB} * \overrightarrow{BC}}{|AB| * |BC|}\right)) \qquad (3)$$

In equation (3), *deg* converts radians to degrees. $\vec{A}$, $\vec{B}$ and $\vec{C}$ are, in order, the coordinates of the joints discussed in the previous bulleted list. The timeseries containing the measurements of these angles are subjected to 6 statistical measurements, instead of 5: the statistical measure shown in Table II plus the maximum amplitude reached of each angle.

TABLE IV
SIGMA-LOGNORMAL FEATURES

Feature name	Description
Lognormal stroke number	Number of lognormal strokes
D parameter	*D* parameter for all lognormal strokes
μ parameter	μ parameter for all lognormal strokes
σ parameter	σ parameter for all lognormal strokes
θ_s parameter	θ_s parameter for all lognormal strokes
θ_e parameter	θ_e parameter for all lognormal strokes

The sigma-lognormal features reported in Table IV, with the exception of the first, the number of lognormals found (which is calculated only once), are calculated, based on the type of step, for each part of the visible body: nose, neck, shoulder (right or left), elbow (right or left), wrist (right or left), hip (right or left), knee (right or left), ankle (right or left). For all those, apart the first, the statistical measures in Table II are applied.

The dataset which contains all the features (velocity, angles and sigma-lognormal) is composed by 675 synthetized segmented steps (the rows) and for a total of 679 features (columns).

In addition to this dataset, two additional datasets were created as the combination of one or more different features. Specifically, the first is composed only by velocity-based features and angles. The third is composed by only sigma-lognormal features. These datasets are fed into the classification models explained in section C.

C. CLASSIFIERS AND EXPERIMENTAL SETUP

The classification is performed with 10-Fold cross validation with 70-30 ratio maintaining the inter-patient separation scheme: 28 subjects in training and 12 in test set randomly kept 10 different times. The produced train-test sets within 10-Fold cross validation may be imbalanced [54]. For this

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/ACCESS.2020.3032202, IEEE Access

reason and only when there was relevant imbalance between healthy and sick, it was decided to perform, before feature selection, a novel oversampling technique called LICIC [54]. This oversampling technique creates new instances balancing the minority classes by preserving nonlinearities and the particular pattern present in each specific class.

It works by copying most important components and permutating less important components among instances of the same class, and thus, create new offspring. It was avoided to use LICIC at the beginning of the pipeline for the dataset as a whole, because it would bleed specific patterns of samples used for testing, also in training, thus making results, less reliable. This kind of bleeding is not considered cheating, but it is not suitable for medical application research.

Before doing the final classification, feature selection was performed using Extra Trees [55] with ordered feature importance. Extra Trees is an ensemble technique which uses the results of multiple de-correlated trees, aggregating them and outputting the classification result. The feature importance metric used was the Gini Index. At the end of the procedure, the features were ordered with respect to their Gini index, from higher to lower and the top 100 features in descending order were kept.

For the binary classification, the following 5 classifiers were used:

1. K-Nearest Neighbors
2. Random Forest
3. AdaBoost with Decision Tree as base learner
4. Linear Support Vector Machines
5. Radial Basis Function (rbf) Support Vector Machines

K-Nearest Neighbor was configured with 5 nearest neighbors. Support Vector Machine with linear kernel and C = 1. The *rbf* version of Support Vector Machine used automatic gamma adjusting and C=1. Random Forest classifier was used with 50 trees and maximum depth of 5 for conquering overfitting. Finally, AdaBoost with 10 decision trees as week learner each with 10 as max depth.

For comparison purposes, two widely known activity recognition algorithm were used for shallow learning. These two algorithms are Motion Blobs [59] and Optical Flow [60]. Both techniques extracted their respective features who were tested against AdaBoost with 50 decision trees and SVM with linear classifier and C=1.

For deep learning, the ResNet deep learning architecture making use of 3D CNN [61] for encoding spatio-temporal information was used. In this work it has been used the 34-layer 3D ResNet with Adam [62] as optimizer. The training time took over 8 days on GPU enabled server with Nvidia Tesla V100 with 16GB video RAM. For additional comparison purposes, well known ResNet-50 [64] and Inception-V3 [65] deep neural network architectures were tested in exactly same conditions. These two last networks were trained end-to-end on a frame by frame basis.

TABLE V
RESULTS ON COMPLETE DATASET

Algorithm	F1	Sensitivity	Specificity	Precision	Mean Acc	AUC
KNN	0.949	0.930	0.968	0.968	0.949	0.951
Random Forest	0.942	0.945	0.939	0.943	0.942	0.945
Ada Boost	0.932	0.936	0.939	0.943	0.942	0.943
Linear SVM	**0.955**	0.953	0.957	0.959	**0.955**	**0.961**
RBF SVM	0.954	0.959	0.949	0.952	0.954	0.955

TABLE VI
RESULTS WITH AND WITHOUT SIGMA-LOGNORMAL FEATURES

Algorithm	F1	Sensitivity	Specificity	Precision	Mean Acc	AUC
SVM All Features	0.955	0.953	0.957	0.959	**0.955**	**0.961**
SVM without Sigma-lognormal Features	0.950	0.947	0.952	0.955	0.950	0.951
SVM only Sigma-lognormal features	0.948	**0.982**	0.929	9.951	0.948	0.947

TABLE VII
RESULTS COMPARISON BETWEEN LEFT STEPS, RIGHT STEPS AND BOTH

Algorithm	F1	Sensitivity	Specificity	Precision	Mean Acc	AUC
SVM Right + Left steps	0.955	0.953	0.957	0.959	0.955	0.961
SVM Right steps	0.975	0.981	0.969	0.971	0.975	0.976
SVM Left Steps	**0.990**	**0.981**	**0.999**	**0.987**	**0.991**	**0.990**

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/ACCESS.2020.3032202, IEEE Access

TABLE VIII
RESULTS COMPARISON BETWEEN OTHER WORKS ON SAME DATASET

Work	F1	Recall	Precision	Accuracy
Motion Blobs + SVM [59]	0.9074	0.8964	0.9074	0.9152
Motion Blobs + AdaBoost [59]	0.9254	0.9232	0.9129	0.9254
Optical Flows + SVM [60]	0.9238	0.881	0.9628	0.8823
Optical Flows + AdaBoost [60]	0.9032	0.8235	0.9634	0.8353
3D ResNet [61]	0.84	0.88	0.83	0.84
2D ResNet 50	0.82	0.86	0.79	0.82
Inception V3	0.82	0.86	0.80	0.82
This Work	**0.990**	**0.976**	**0.987**	**0.991**

D. RESULTS

The averaged results of 10-Fold cross validation performed on the whole dataset by the 5 different algorithms is shown in Table V. The best performing classifier is the linear SVM. Table VI, instead, shows results of Linear SVM, with and without sigma-lognormal features. This comparison is important and will be discussed in more details in Section VII. To evaluate the possibility that the right and left steps can generate different features due to their intrinsic nature, it was decided to try to classify the right and left steps also separately. The results are presented in Table VII.
Table VIII contains comparison results with other state of the art techniques on the same dataset.

VII. DISCUSSION

The results presented in Table V show a mean accuracy of the system of 95.5% when linear SVM is used. This result is higher with respect to other reviewed works, but a rigorous comparison among other systems cannot be done because also other reviewed works, make use of private datasets to remain compliant with the privacy laws, in this case the GDPR (European).
The hypotheses at the basis of this work is that sigma-lognormal features would help in discriminating healthy from unhealthy subjects.

Table VI shows that mean accuracy is increased when sigma-lognormal features are used by 0.5%. This sensible increase in accuracy is important because, thanks to the inter-patient separation scheme, it was therefore possible to make an analysis to understand to whom, the badly classified steps, belonged to when no sigma-lognormal features were kept into consideration.
As results, all badly classified instances belonged to mild and found stages. In this case, adding sigma-lognormal features, increased the chance of recognizing borderline subjects, especially subjects with an early stage disease.
This finding confirms our hypotheses and empirically shows that sigma-lognormal features and, more in general, the Plamondon's Kinematic Theory of Rapid Human Movements [8], [9], [10] can be effectively transposed from the handwriting domain to a more general scenario. That is, these sigma-lognormal features can capture small grain movement details, otherwise impossible to be captured by effective but coarse features such as velocity, displacement and angles.

Sigma-lognormal features, at the end, contribute in classifying the most difficult, borderline, instances, contributing, in an important way, in the realization of the big picture of this work: early stage neurodegenerative disease classification. From Table VI is also possible to note that when sigma-lognormal features are considered in isolation, the sensitivity, which is the number of sick patients correctly identified as sick, is higher with respect to other, hitting 0.982. This finding contributes in confirming the hypothesis that the Kinematic Theory of Rapid Human Movement and its sigma-lognormal model captures fine grained variabilities present in borderline subjects.

Another interesting finding is that the separate classification of the right and left steps produces better results than the joint classification. This is because, the classification algorithm would learn more discriminated and correlated patterns when features are synthetized on exactly same walking direction and orientation. In this specific scenario, the highest accuracy of 99.1% was achieved by using all types of features but analyzing only the left walk orientation, as shown in Table VII.

The comparison results in Table VIII shows that our technique is at least 7 percentual points more accurate than all others. Motion Blobs and Optical Flows perform better than the 3D (spatio-temporal) Deep Learning technique called 3D ResNet and also all other DL architectures tested. This result contradicts the common opinion in the community that deep learning is on par or outperforms other techniques. The hypothesis is that DL architectures were biased toward several scene's details: algorithms may learn several scene parameters, such as the surrounding environment, the presence of another person (nurse), the camera orientation and the focus with respect to the subject acting. This is particularly visible when interpatient

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/ACCESS.2020.3032202, IEEE Access

separation scheme is applied, in fact few videos of patients walking were recorded at the exact same room, some, different, healthy subject videos were recorded in another place with wooden surroundings.

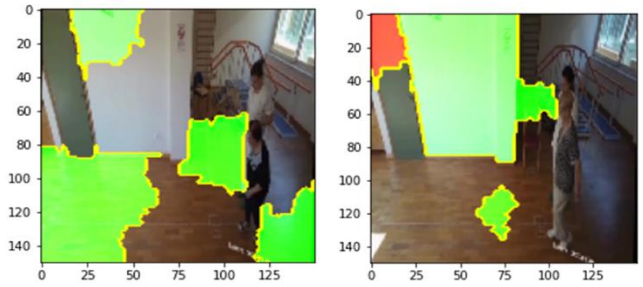

FIGURE 10. Lime result of patient performing gait (a) and healthy subject performing gait (b). In red the areas that contributed less in the classification of healthy/unhealthy subject, in green areas that contributed most.

For understanding what the DL models have learnt, it has been used the local interpretable model-agnostic explanations (LIME) [62] technique. This explainable artificial intelligence [56] technique is one of the most widely used implementation of local surrogate models. Surrogate models are trained to approximate the predictions of the underlaying neural network model. In this specific case, LIME focused on training local surrogate models to explain individual predictions of the 3D ResNet.

Figure 10 shows that LIME [62] algorithm found the environment, as most important parts for discriminating healthy from unhealthy subjects. This confirms the hypotheses of scene bias of some deep learning techniques applied to this kind of tasks.

Returning to the original technique developed in this work, thanks to the interpretability brought by Decision Trees integrated within Extra Trees algorithm used in feature selection, it is possible to rank features from the most to the least important. Figure 11 shows the top 13 features.

As it is possible to observe, velocity, acceleration and displacement of nose, hip and neck are the most important. By important, it is meant, the characteristics that allow to better distinguish the pattern of a healthy person from the pattern of a sick person.

Interesting is also the presence of the input command D of the sigma-lognormal model for the hip in 13th position (the green bar in Figure 11).

The ability for such a system to be interpretable is of paramount importance in computer aided diagnosis tools. Explainable Artificial Intelligence (briefly XAI) [56] allows the human being to understand, but also trust, what the system has predicted.

In this case, the high interpretability of this system allowed to draw different conclusions with respect to what other authors in [6], [15], [18], [19], [25], [28], [34] reported. From a pure computer vision and pattern recognition perspective, differently from the global medical assumption of measuring legs muscle activities, there is important evidence that the upper side of the body, such as nose, hip, neck and shoulders, play a major role in discriminating healthy from unhealthy subjects.

VIII. CONCLUSIONS AND FUTURE WORK

In this work, an early neurodegenerative disease assessment computer aided tool was developed. The high accuracies

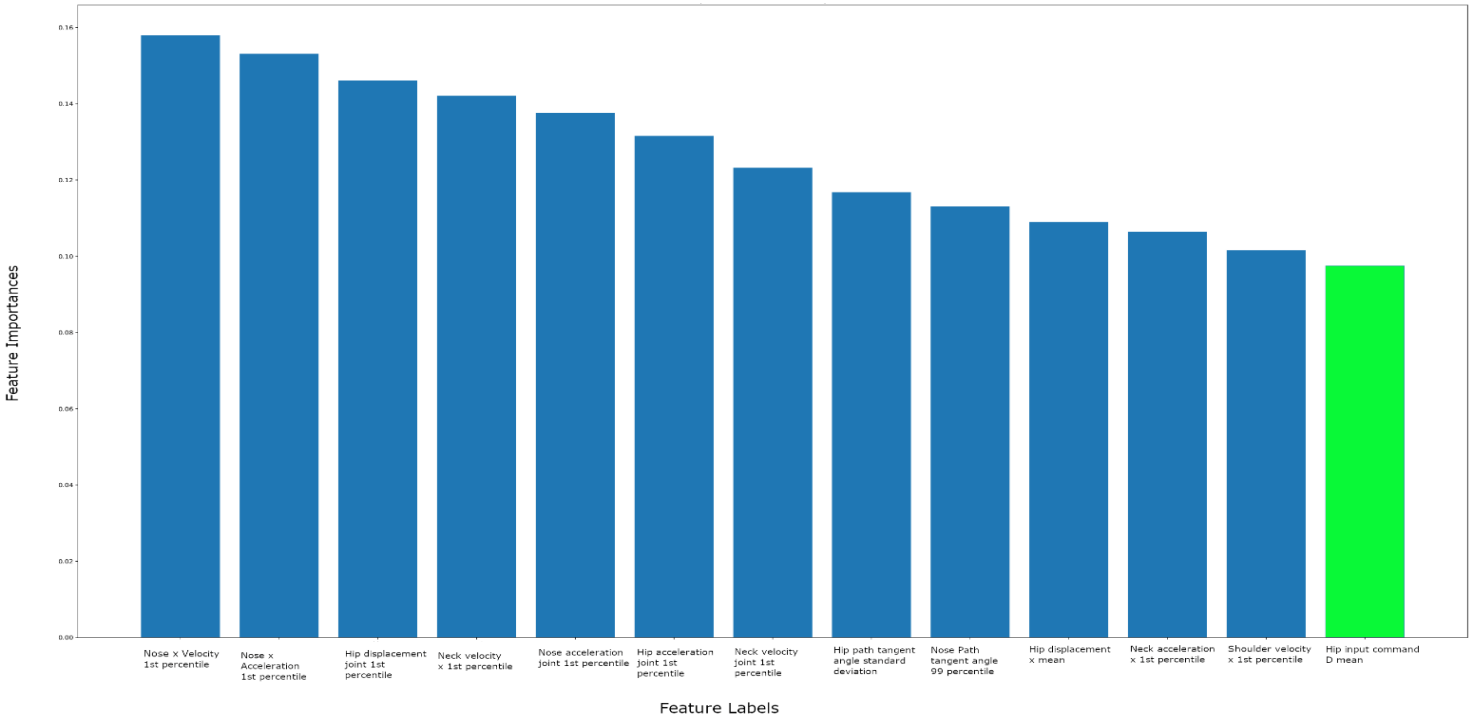

FIGURE 11. Top 13 ranked features used for classification

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/ACCESS.2020.3032202, IEEE Access

achieved by this system suggests that the proposed pipeline is effective.

It is important to assert that there is enough evidence that the Kinematic Theory of Rapid Human Movements can be used also in this specific domain. In this case, this theory and its sigma-lognormal model were capable of capturing fine grained movements that allowed to correctly classify borderline people affected by some kind of neurodegenerative disease in early stage, thus achieving the goal of this work.

Though this is a preliminary work in the field, thanks to the interpretability of the resulting model, it is possible to state that, from a pure computer vision perspective, the most important body parts when it comes to neurodegenerative disease classification trough 2D cameras are in the upper side of the body, in particular nose, neck, hip and shoulders.

Despite the numerosity of the dataset, results are encouraging and higher with respect to the reviewed works.

In a future work, 3D pose estimation will be performed and these features plus other new features from signal processing domain will be used on the 3D estimated pose. In addition, multi-class classification will be performed with the aim of predicting the severity of the disease.

With the help of trained psychologists and neurologists, more data will be collected. Of extreme importance is balancing the dataset finding healthy control subjects in their 80s.

The final aim of this research is to early predict the nature of different neurodegenerative diseases by means of computer vision. This is only our first step towards this big picture.

COMPLIANCE WITH ETHICAL STANDARDS

Ethical approval: All procedures performed in studies involving human participants were in accordance with the ethical standards of the institutional and/or national research committee and with the 1964 Helsinki declaration and its later amendments or comparable ethical standards.

Informed consent: Informed consent was obtained from all individual participants included in the study.

Conflict of interest: The authors declare no conflict of interest.

REFERENCES

[1] Impedovo, D., Pirlo, G.: Dynamic handwriting analysis for the assessment of neurodegenerative diseases: A pattern recognition perspective. IEEE reviews in biomedical engineering 12, 209-220 (2019).

[2] De Stefano, C., Fontanella, F., Impedovo, D., Pirlo, G., & di Freca, A. S. Handwriting analysis to support neurodegenerative diseases diagnosis: A review. Pattern Recognition Letters, 121, 37-45 (2019).

[3] Ubul, K., Tursun, Ga, Aysa, Ab, Impedovo, D., Pirlo, G., Yibulayin, T.: Script Identification of Multi-Script Documents: A Survey. IEEE Access 5, 7890400, pp. 6546-6559 (2017).

[4] Impedovo, D., Pirlo, G., Mangini, F.M., Barbuzzi, D., Rollo, A., Balestrucci, A., Impedovo, S., Sarcinella, L., O'Reilly, C., Plamondon, R.: Writing generation model for health care neuromuscular system investigation. In: CIBB 2013, LNBI, vol. 8452, pp. 137-148 (2014).

[5] Impedovo, D.: Velocity-based signal features for the assessment of Parkinsonian handwriting. IEEE Signal Processing Letters 26(4), pp. 632-636 (2019).

[6] Dentamaro Vincenzo, Donato Impedovo, and Giuseppe Pirlo. "Real-Time Neurodegenerative Disease Video Classification with Severity Prediction." International Conference on Image Analysis and Processing. Springer, Cham (2019).

[7] Chen, S., Lach, J., Lo, B., & Yang, G. Z. Toward pervasive gait analysis with wearable sensors: a systematic review. IEEE journal of biomedical and health informatics, 20(6), 1521-1537. (2016)

[8] O'Reilly, C., & Plamondon, R. Development of a Sigma–Lognormal representation for on-line signatures. Pattern Recognition, 42(12), 3324-3337. (2009)

[9] Plamondon, R. A kinematic theory of rapid human movements. Biological cybernetics, 72(4), 295-307. (1995a)

[10] Plamondon, R. A kinematic theory of rapid human movements. Biological cybernetics, 72(4), 309-320. (1995b)

[11] Emard, J. F., Thouez, J. P., & Gauvreau, D. Neurodegenerative diseases and risk factors: a literature review. Social science & medicine, 40(6), 847-858. (1995)

[12] Rastegari, E., Azizian, S., & Ali, H. Machine Learning and Similarity Network Approaches to Support Automatic Classification of Parkinson's Diseases Using Accelerometer-based Gait Analysis. In Proceedings of the 52nd Hawaii International Conference on System Sciences. (2019)

[13] Prabhu, P., Karunakar, A. K., Anitha, H., & Pradhan, N. Classification of gait signals into different neurodegenerative diseases using statistical analysis and recurrence quantification analysis. Pattern Recognition Letters. (2018)

[14] Najafabadian, B., Jalali, H., Sheibani, A., & Maghooli, K. Neurodegenerative Disease Classification Using Nonlinear Gait Signal Analysis, Genetic Algorithm and Ensemble Classifier. In Electrical Engineering (ICEE), Iranian Conference on (pp. 1482-1486). IEEE. (2018)

[15] Tao, W., Liu, T., Zheng, R., & Feng, H. Gait analysis using wearable sensors. Sensors, 12(2), 2255-2283. (2012)

[16] Moraes, C. R. L., Junior, E. A., Lucena, R. J. R. S., Cavalcanti, É. L., & Rodrigues, M. A. B. Human Gait Cycle Analysis Using an Adapted Mechanical Prosthesis. In XXVI Brazilian Congress on Biomedical Engineering (pp. 241-248). Springer, Singapore. (2019)

[17] Kharb, A., Saini, V., Jain, Y.K., Dhiman, S.: A review of gait cycle and its parameters. IJCEM International Journal of Computational Engineering & Management 13, pp. 78-83 (2011).

[18] Chen, S., Lach, J., Lo, B., & Yang, G. Z. Toward pervasive gait analysis with wearable sensors: A systematic review. IEEE journal of biomedical and health informatics, 20(6), 1521-1537. (2016)

[19] Muro-De-La-Herran, A., Garcia-Zapirain, B., & Mendez-Zorrilla, A.. Gait analysis methods: An overview of wearable and non-wearable systems, highlighting clinical applications. Sensors, 14(2), 3362-3394. (2014)

[20] Gafurov, D., Helkala, K., & Søndrol, T.. Biometric Gait Authentication Using Accelerometer Sensor. JCP, 1(7), 51-59. (2006)

[21] Dhiman, C., & Vishwakarma, D. K. A review of state-of-the-art techniques for abnormal human activity recognition. Engineering Applications of Artificial Intelligence, 77, 21-45. (2019)

[22] Sanchez, V., Prince, G., Clarkson, J. P., & Rajpoot, N. M. Registration of thermal and visible light images of diseased plants using silhouette extraction in the wavelet domain. Pattern Recognition, 48(7), 2119-2128. (2015)

[23] Cao, Z., Simon, T., Wei, S. E., & Sheikh, Y. Realtime multi-person 2D pose estimation using part affinity fields. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (pp. 7291-7299). (2017).

[24] Tay, A.; Yen, S.C.; Li, J.Z.; Lee, W.W.; Yogaprakash, K.; Chung, C.; Liew, S.; David, B.; Au, W.L. Real-Time Gait Monitoring for Parkinson Disease. Proceedings of 2013 10th IEEE International

Conference on Control and Automation (ICCA), Hangzhou, China, 12–14 June 2013; pp. 1796–1801 (2013)

[25] Salarian, A.; Burkhard, P.R.; Vingerhoets, F.J.G.; Jolles, B.M.; Aminian, K.A. Novel approach to reducing number of sensing units for wearable gait analysis systems. IEEE Trans. Biomed. Eng, 60, 72–77. (2013)

[26] Bamberg, S.; Benbasat, A.Y.; Scarborough, D.M.; Krebs, D.E.; Paradiso, J.A. Gait analysis using a shoe- integrated wireless sensor system. Trans. Inf. Tech. Biomed, 12, 413–423. (2008)

[27] Anna, A.S.; Wickström, N.; Eklund, H.; Zügner, R.; Tranberg, R. Assessment of Gait Symmetry and Gait Normality Using Inertial Sensors: In-Lab and In-Situ Evaluation. In Biomedical Engineering Systems and Technologies; Gabriel, J., Schier, J., Huffel, S.V., Eds.; ConchonE.CorreiaC.FredA.GamboaH.Springer: Berlin, Germany; pp. 239–254 (2013)

[28] Qi, Y.; Soh, C.B.; Gunawan, E.; Low, K.S.; Maskooki, A. Using Wearable UWB Radios to Measure Foot Clearance During Walking. Proceedings of 2013 35th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), Osaka, Japan, 3–7 July 2013; pp. 5199–5202. (2013)

[29] Samson, W.; van Hamme, A.; Sanchez, S.; Chèze, L.; van Sint Jan, S.; Feipel, V. Dynamic footprint analysis by time-of-flight camera. Comput. Methods Biomech. Biomed. Engin. 15, 180–182. (2012)

[30] Xue, Z.; Ming, D.; Song, W.; Wan, B.; Jin, S. Infrared gait recognition based on wavelet transform and support vector machine. Pattern Recognit. 43, 2904–2910. (2010)

[31] Zhang, J.T.; Novak, A.C.; Brouwer, B.; Li, Q. Concurrent validation of Xsens MVN measurement of lower limb joint angular kinematics. Physiolog. Meas. 34, N63–69 (2013)

[32] Clark, R.A.; Pua, Y.H.; Bryant, A.L.; Hunt, M.A. Validity of the Microsoft Kinect for providing lateral trunk lean feedback during gait retraining. Gait Posture 38, 1064–1066. (2013)

[33] Gabel, M.; Gilad-Bachrach, R.; Renshaw, E.; Schuster, A. Full Body Gait Analysis with Kinect. Proceedings of 2012 Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), San Diego, CA, USA; pp. 1964–1967. (2012)

[34] Novak, D.; Reberšek, P.; de Rossi, S.M.M.; Donati, M.; Podobnik, J.; Beravs, T.; Lenzi, T.; Vitiello, N.; Carrozza, M.C.; Munih, M. Automated detection of gait initiation and termination using wearable sensors. Med. Eng. Phys. 35, 1713–1720. (2013)

[35] Dominguez, G.; Cardiel, E.; Arias, S.; Rogeli, P. A Digital Goniometer Based on Encoders for Measuring Knee-Joint Position in an Orthosis. Proceedings of 2013 World Congress on Nature and Biologically Inspired Computing (NaBIC), Fargo, ND, USA pp. 1–4. (2013)

[36] Middleton, L.; Buss, A.A.; Bazin, A.; Nixon, M.S. A Floor Sensor System for Gait Recognition. Proceedings of 2005 4th IEEE Workshop on Automatic Identification Advanced Technologies, Buffalo, NY, USA; pp. 171–176. (2005)

[37] Leusmann, P.; Mollering, C.; Klack, L.; Kasugai, K.; Ziefle, M.; Rumpe, B. Your Floor Knows Where You Are: Sensing and Acquisition of Movement Data. Proceedings of 2011 12th IEEE International Conference on Mobile Data Management (MDM), Luleå, Sweden; pp. 61–66. (2011)

[38] Derawi, M., Ali, H., & Alaya Cheikh, F. Gait recognition using time-of-flight sensor. (2011)

[39] Muramatsu, D.; Shiraishi, A.; Makihara, Y.; Yagi, Y. Arbitrary View Transformation Model for Gait Person Authentication. Proceedings of 2012 IEEE 5th International Conference on Biometrics: Theory, Applications and Systems (BTAS), Arlington, VA, USA; pp. 85–90 (2012)

[40] Liu, H.; Cao, Y.; Wang, Z. Automatic Gait Recognition from a Distance. Proceedings of Control and Decision Conference (CCDC), Xuzhou, China; pp. 2777–2782. (2010)

[41] Vera-Rodriguez, R.; Mason, J.S.D.; Fierrez, J.; Ortega-Garcia, J. Comparative analysis and fusion of spatiotemporal information for footstep recognition. IEEE Trans. Pattern Anal. Mach. Intell. 35, 823–834. (2013)

[42] Maity, S., Chakrabarti, A., & Bhattacharjee, D. Robust Human Action Recognition Using AREI Features and Trajectory Analysis from Silhouette Image Sequence. IETE Journal of Research, 65(2), 236-249. (2019)

[43] Li, M. H., Mestre, T. A., Fox, S. H., & Taati, B. Vision-based assessment of parkinsonism and levodopa-induced dyskinesia with pose estimation. Journal of neuroengineering and rehabilitation, 15(1), 97. (2018)

[44] Ye, Q., Yi, X., Yao, Z: Classification of gait patterns in patients with neurodegenerative disease using adaptive neuro-fuzzy inference system. Computational and mathematical methods in medicine, 9831252 (2018).

[45] Yang, M., Zheng, H., Wang, H., McClean, S.: Feature selection and construction for the discrimination of neurodegenerative diseases based on gait analysis. In: 2009 3rd International Conference on Pervasive Computing Technologies for Healthcare, p. 1-7 (2009).

[46] Shetty, S., Rao, Y.S.: SVM based machine learning approach to identify Parkinson's disease using gait analysis. 2016 International Conference on Inventive Computation Technologies (ICICT), pp. 1-5 (2016).

[47] Kaur, S., Aggarwal, H., & Rani, R. Hyper-parameter optimization of deep learning model for prediction of Parkinson's disease. Machine Vision and Applications, 31, 1-15. (2020)

[48] Mehrizi, R., Peng, X., Zhang, S., Liao, R., & Li, K. Automatic Health Problem Detection from Gait Videos Using Deep Neural Networks. arXiv preprint arXiv:1906.01480. (2019)

[49] O'Reilly, C., & Plamondon, R. Design of a neuromuscular disorders diagnostic system using human movement analysis. In 2012 11th International Conference on Information Science, Signal Processing and their Applications (ISSPA) (pp. 787-792). IEEE. (2012)

[50] Drap, P., & Lefèvre, J. An exact formula for calculating inverse radial lens distortions. Sensors, 16(6), 807. (2016)

[51] Lin, T. Y., Maire, M., Belongie, S., Hays, J., Perona, P., Ramanan, D., & Zitnick, C. L. Microsoft coco: Common objects in context. In European conference on computer vision (pp.740-755).Springer, Cham. (2014)

[52] Welch, Greg, and Gary Bishop. "An introduction to the Kalman filter.": 41-95. (1995)

[53] Auger, F., Hilairet, M., Guerrero, J. M., Monmasson, E., Orlowska-Kowalska, T., & Katsura, S. Industrial applications of the Kalman filter: A review. IEEE Transactions on Industrial Electronics, 60(12), 5458-5471. (2013)

[54] Dentamaro, V., Impedovo, D., & Pirlo, G. LICIC: Less Important Components for Imbalanced Multiclass Classification. *Information*, *9*(12), 317. (2018)

[55] Geurts, P., Ernst, D., & Wehenkel, L. Extremely randomized trees. *Machine learning*, *63*(1), 3-42. (2006)

[56] Samek, W. Explainable AI: interpreting, explaining and visualizing deep learning (Vol. 11700). *Springer Nature*. (2019)

[57] Robinson, M. A., & Vanrenterghem, J. An evaluation of anatomical and functional knee axis definition in the context of side-cutting. *Journal of biomechanics*, 45(11), 1941-1946. (2012)

[58] Verlekar, T. T., De Vroey, H., Claeys, K., Hallez, H., Soares, L. D., & Correia, P. L. Estimation and validation of temporal gait features using a markerless 2D video system. *Computer methods and programs in biomedicine*, 175, 45-51. (2019).

[59] Yang, J., Cheng, J., & Lu, H. (2009, June). Human activity recognition based on the blob features. *In 2009 IEEE International Conference on Multimedia and Expo* (pp. 358-361). IEEE.

[60] Kumar, S. S., & John, M. (2016, October). Human activity recognition using optical flow based feature set. *In 2016 IEEE international Carnahan conference on security technology (ICCST)* (pp. 1-5). IEEE.

[61] Hara, K., Kataoka, H., & Satoh, Y. (2017). Learning spatio-temporal features with 3D residual networks for action recognition. *In Proceedings of the IEEE International Conference on Computer Vision Workshops (*pp. 3154-3160).

[62] Kingma, D. P., & Ba, J. (2014). Adam: A method for stochastic optimization. *arXiv preprint arXiv*:1412.6980.

[63] Ribeiro, M. T., Singh, S., & Guestrin, C. (2016, August). " Why should I trust you?" Explaining the predictions of any classifier. *In Proceedings of the 22nd ACM SIGKDD international conference on knowledge discovery and data mining* (pp. 1135-1144).

[64] He, K., Zhang, X., Ren, S., & Sun, J. (2016). Deep residual learning for image recognition. *In Proceedings of the IEEE conference on computer vision and pattern recognition* (pp. 770-778).

[65] Szegedy, C., Vanhoucke, V., Ioffe, S., Shlens, J., & Wojna, Z. (2016). Rethinking the inception architecture for computer vision. *In Proceedings of the IEEE conference on computer vision and pattern recognition* (pp. 2818-2826).