Artificial Intelligence In Medicine 94 (2019) 54–66

Contents lists available at ScienceDirect

Artificial Intelligence In Medicine

journal homepage: www.elsevier.com/locate/artmed

Normal and pathological gait classification LSTM model

Margarita Khokhlova[a,*], Cyrille Migniot[a], Alexey Morozov[b], Olga Sushkova[b], Albert Dipanda[a]

[a] Le2i, FRE CNRS 2005, Univ. Bourgogne Franche-Comté, France
[b] Kotelnikov Institute of Radio Engineering and Electronics of RAS, Moscow 125009, Russia

ARTICLE INFO

Keywords:
Gait assessment
Gait modeling
Low-limbs motion
LSTM
RGB-D
Kinect skeletons

ABSTRACT

Computer vision-based clinical gait analysis is the subject of permanent research. However, there are very few datasets publicly available; hence the comparison of existing methods between each other is not straightforward. Even if the test data are in an open access, existing databases contain very few test subjects and single modality measurements, which limit their usage. The contributions of this paper are three-fold. First, we propose a new open-access multi-modal database acquired with the Kinect v.2 camera for the task of gait analysis. Second, we adapt to use the skeleton joint orientation data to calculate kinematic gait parameters to match golden-standard MOCAP systems. We propose a new set of features based on 3D low-limbs flexion dynamics to analyze the symmetry of a gait. Third, we design a Long-Short Term Memory (LSTM) ensemble model to create an unsupervised gait classification tool. The results show that joint orientation data provided by Kinect can be successfully used in an inexpensive clinical gait monitoring system, with the results moderately better than reported state-of-the-art for three normal/pathological gait classes.

1. Introduction

Traditional measures used to analyze gait parameters in clinical settings are semi-subjective. Medical specialists monitor the quality of the patient's walk. Sometimes, a patient is also asked to perform a specific action (get up from the chair and walk, etc.). These methods provide subjective measurements and do not guarantee accuracy, repeatability, and reproducibility, which adversely affect the diagnosis, supervision, and treatment of gait pathologies.

To make gait analysis more quantitative, floor sensor-based (ground force platforms) and marker-based (Vicon, Optotrac) motion capture (MOCAP) solutions exist. These systems provide accurate and reliable data for gait assessment, but pose a significant cost burden for clinics. In addition, the setup of such systems for each patient is time costly.[1] Machine vision-based gait analysis approaches gained popularity as alternatives to subjective gait analysis and costly marker-based or floor sensor-based methods.

RGB-D sensors are often used in gait analysis. Such sensors provide not only a standard color image, but also the structural information of a scene. Earlier studies have shown that Kinect sensor can be exploited to examine gait. The second version of Microsoft Kinect was recognized as valid for important clinical gait parameters assessment [1–4]. Kinect was also successfully used for gait recognition [5–8]. The affordability and accessibility of modern RGB-D cameras makes them a potential tool for the automatic objective motion analysis. This work propose a Kinect v.2-based setup, a database, and a Long-Short Memory (LSTM) model for automatic pathological gait assessment, detecting a limp and knee rigidity. Kinect camera provides with several types of data: color image, depth image, near infrared image, audio, face features and skeleton data. Color image is the regular image with RGB values for each pixel. Depth image contains the real distances between the camera and objects on the scene. Audio is a simple sound recording. Face data are the set of points capturing face features. Skeletons are the simplified human body parameters calculated from the depth silhouette of a detected human. A skeleton contains 25 key positions called joints, joint orientations (in form of quaternions), describing the body parts' rotations, and joint states, which signify if the joint coordinates were calculated from a depth frame or approximated.

The objective of gait analysis is to quantify objectively the parameters of the gait and detect the presence of a pathology. The list of gait parameters used in clinical research is exhaustive. Gait parameters can be broadly divided into three groups: kinetic which deals with the forces acting on or exerted by the body, spatio-temporal which deals with distance and time related parameters, and kinematic which deals

* Corresponding author.
E-mail addresses: margarita.khokhlova@u-bourgogne.fr (M. Khokhlova), cyrille.migniot@u-bourgogne.fr (C. Migniot), albert.dipanda@u-bourgogne.fr (A. Dipanda).

[1] According to the clinical staff of the University Medical Center in Dijon, the time to place Vicon markers and calibrate the system takes at least 15 minutes per patient.

https://doi.org/10.1016/j.artmed.2018.12.007
Received 30 November 2018; Accepted 27 December 2018
0933-3657/

with joint angles, joint velocity and joint accelerations. The in-depth research on the subject was performed by Roberts and Mongeon [9] to identify the most used gait parameters. According to the authors, kinetic parameters were found to be the most often measured parameters, followed by spatio-temporal and kinematic parameters.

1.1. Contributions

The goal of this research is to propose a complete gait analysis and classification system based on the data from the Kinect v.2 sensor. We aim to build an acquisition protocol, a feature extraction method from the RGB-D sensor data, and a Machine Learning (ML) based classification tool.

Our gait databases study presented later in this section has shown that the number of existing pathological gait databases available for a reuse in research purposes is very small, and commonly they contain only the coordinates of skeleton joints. To overcome this issue, we collected a new multi-modal gait symmetry (MMGS) database in our laboratory. Proposed MMGS contains depth and skeletons data, including the skeleton joints orientations. A normal ground walking mode is used to make the acquisition condition natural and easily reproducible without additional equipment.

We demonstrate how the Kinect v.2 joint orientations data can be used in order to obtain flexion (a bending movement that decreases the angle between a body segment and its proximal segment), abduction (a motion that pulls apart away from the mid-line of the limb), and rotation (a rotating movement towards the axis of the body) angles of the low-limbs. Low-limbs dynamics are commonly used to diagnose gait diseases and pathologies. Usually, their measurements are performed by MOCAP systems, such as Vicon. Then a clinician can make a diagnosis of the patients gait based on the obtained low-limbs kinematic parameters. Research studies using Kinect for gait assessment, typically use the joints data directly, or use the flexion angles only. Meanwhile, the information about joints' orientations can be used to calculate the low-limbs kinematic parameters in an effective and simple way. Earlier it was shown that Kinect bone orientations data can be used for low-limbs dynamics assessment during a vertical jump [10]. However, the amplitude of flexion angles during walk is typically less than that of a jump, and the capturing volume is bigger. To validate kinematic parameters during gait, we compare the knee and hip dynamics between Kinect v.2 and Vicon system.

Finally, we propose an ML-based gait model to show how the kinematic parameters described in the previous section can be used for the automatic pathological gait assessment. Gait data is a multi-dimensional time-domain sequence. In gait sequences, the joint angles of current time are related with the previous frame angles, future frame angles, and also other joints angles. To capture these complex relations, we use a bi-directional LSTM gait model. We perform experiments using the collected MMGS database. An LSTM ensemble model is trained to detect the normal, limping, and knee rigid gait based on the hip and knee flexion dynamics.

1.2. Proposed algorithm

We propose to evaluate gait symmetry based on low-limbs flexion angles to detect deviations from normal gait pattern.

The proposed gait model is based on the data from the Kinect v.2 sensor. It is obtained in several steps. We use the human body skeletons extracted from each depth video frame coming from the sensor. First, we segment these skeleton sequences coming from the sensor into repetitive parts (gait cycles). Then we calculate low-limbs flexion features from Kinect orientation data for each gait cycle. These features are used in a ML-based gait model. A ensemble LSTM-based model is trained using data from a set of persons to learn the difference between normal and abnormal gait based on kinematic parameters. Finally, the model is validated on the testing data coming from the other subjects not used for training.

The main novelties of the proposed algorithm are the following. First, we choose to use only low-limbs joint data and excluded the ankle data as the least reliable, in contrast to the methods which use all the joints [11,12] and ankle, hip, and knee joint [13]. Second, we propose to use flexion and rotation dynamics calculated from orientation data, validating our features by a comparison with a golden-standard MOCAP system. Third, we use flexion and rotation angles and not just flexion data as it is commonly done [13,14]. Finally, we adapt a novel LSTM model architecture to classify the data and detect two gait pathologies. To our knowledge, this is the first work adopting an LSTM model for the normal/pathological gait classification task based on 3D joints data.

1.3. Normal/abnormal gait datasets

Latest gait assessment algorithms [11–13] report accurate results, however, the used benchmarks are either very small ([12], just 56 sequences), contain a lot of noise [15], or are taken into special condition ([16] the treadmill is used). The collection of patient data requires many formal permissions and a special laboratory setup. In the same moment, the research teams working on Computer Vision-based motion analysis do not commonly have an access to the clinical environment. This brings the situation when there are very few benchmark datasets publicly available in gait assessment field.

The absence of data is a big issue in gait clinical studies. Normal gait is more prevalent, since we can use data from the gait recognition domain [7,17]. The amount of publicly available gait data is small compared to the number of gait studies that have been performed over the years. The data that are available generally suffer from limitations such as few subjects, few gait cycles, highly clinical, no raw data, lack of metadata, non-standard formats, and restrictive licensing [18]. However, there are very few examples of the abnormal gait publicly available for a reuse. We collected the information about the gait datasets acquired with RGB-D camera devices and shared with the other users. It should be noted that more databases collected with the other devices, such as 2D cameras, golden-standard MOCAP systems, wearable sensors and others exist. Moore et al. [18] reviews many of these databases. Therefore, we concentrate here on the normal/pathological gait data acquired by RGB-D sensors, and the Kinect camera in particular. We also add two popular gait recognition databases [17,6], since samples from them can be used as normal gait examples. Information about the existing gait datasets is grouped in Table 1. It can be seen that commonly the datasets contain a small number of test subjects, and the gait-affecting diseases are simulated.

Table 1
Available pathological gait datasets.

Dataset	Year	Device	Subjects	Modalities	Comment
TUM Gaid [17]	2014	RGB-D	305	audio, image and depth	Person identification and assessment of the soft biometrics
SPHERE-Walking2015 [15]	2014	Kinect v.1	20	skeleton	Normal, PD and stroke simulating persons climbing stairs, 15 joints, noisy data
DAI [12]	2015	Kinect v.2	7	skeleton, orientations	Normal/abnormal (asymmetric) sequences performed by actors
UPCV Gait K2 [7]	2016	Kinect v.2	30	skeleton	Gait recognition database
Walking gait dataset [16]	2018	Kinect v.2	8	skeleton, point clouds	A treadmill is used, padding sole and attaching weight to simulate the pathological asymmetric gait

The biggest multi-modal database containing abnormal gait samples of all is the recent one by Nguyen et al. [16]. However, the authors do not store all the data about the skeletons (joint states and orientations are missing). In addition, a treadmill is used to acquire this dataset, which slightly affects kinematic parameters [19], so this data cannot be used in the more common solid surface walking scenario. To continue with the promising direction of the Computer Vision-based gait assessment, more data is required. This data should satisfy the multi-modal criteria, contain a bigger number of test subjects, and be captured in normal conditions.

The remainder of this article is organized as follows. State of the art gait analysis and classification CV-based methods are reviewed in Section 2. We present a MMGS database containing normal and pathological gait samples in Section 3. We present the procedure to calculate knee and hip flexion angles from Kinect v.2 orientations in 4. A ensemble LSTM gait model design and parameters are described in Section 5. Section 6 describes the performed experiments. We assess the reliability of the low-limbs flexion angles calculated from Kinect v.2 using intraclass correlation coefficient in Section 6.1. The gait model performance is validated on the MMGS dataset in 6.2. Finally, Section 7 concludes the paper, where we highlight the outcomes of this research and propose an outlook on future work.

2. State of the art

From the first days of their appearance, researchers actively used RGB-D sensors for gait research. A gait analysis system based on skeleton joints calculated from depth images was proposed by Gabel et al. [20] in 2012. Using regression models, the authors successfully estimated the stride duration and arm angular velocity from the joints data. Rocha et al. [14] evaluated the possibility to use skeleton joints data from the Kinect v.1 sensor for gait assessment. First, the acquired sequences were temporally segmented into cycles. Then for each cycle velocities, accelerations and intra-distances for the selected joints along with the flexion angles were obtained. Final feature set was calculated as mean, median, and variance of these values. Afterwards, the most characteristic features for Parkinson Disease (PD) patients were identified among calculated features. The testing data used by researchers was quite small – just 3 training and 3 testing subjects and the data is not available.

Abnormal or pathological gait deviates from normal gait patterns. There could be many different reasons for such a deviation, the most common are different neurological diseases. Paiement et al. [15] targeted abnormal gait detection by analyzing the skeleton data coming from the Kinect v.1 sensor. A non-linear manifold learning technique was employed to reduce the dimensionality of the noisy skeleton data. Then a statistical model was learned from the normal gait samples and the detection was performed based on matching the new observations to the model following Markov assumptions. This method was then applied to detect simulated gait anomalies on subjects climbing a flight of stairs/walking on the flat surface. The data collected for this experiment is shared with the research community.

Kastaniotis et al. [21] proposed a Kinect-based acquisition platform and classification algorithm for healthy and Multiple Sclerosis (MS) patients. The authors used selected eight joint angles as the features. Then each static pose from a gait sequence is represented by a 16-dimensional space of 2D Euler angles. This features are calculated for each frame and then transformed into dissimilarity space to get constant-size (independent on the number of frames) matrix representation for the gait sequence. Linear discriminant analysis is then used to perform the binary classification for two classes, containing in total 8 healthy and 9 MS patients.

Chaaraoui et al. [12] performed abnormal gait detection with RGB-D devices using so-called Joint Motion History (JMH) features. Skeletons were used to track motion over a segment of frames based on the three-dimensional location of their joints. Each skeleton (obtained by the Kinect v.2 sensor) was normalized for location, size, and rotation. The skeletons were then accumulated using a sliding window approach. For each window, all skeleton joints were shown as 3D coordinates of a volume divided in 3D cells called voxels. The JMH was encoding the information about the occupancy of voxels. To reduce the size of JMH, researchers used a dimensionality reduction method based on an axis projection. The BagOfKeyPoses (variation of the bag-of-words model) was used to detect abnormal behaviour based on JMH features. The temporal relation was obtained via learning templates of sequences of key poses, which indirectly captures the gait events. The method seems to be over complicated and does not identify the source of the gait pathology. The datasets used are not big enough to generalize the results, however, available for reuse.

Devanne et al. [22] analyzed the motion trajectory for each step separately. Full body posture described by joints positions is used for the analysis. The trajectory made by the skeleton joints during each step is obtained by concatenating the joints feature vectors for corresponding time intervals. A custom distance metric was then used to compare different gait sequences and remove the outliers. A statistical model of the normal distribution in a Riemannian shape space was built by analyzing shape variations among all samples belonging to left and right steps. New gait sequences were then evaluated using the two learned step models. Experimental results demonstrate the effectiveness of the proposed approach in the context of asymmetric gait detection. However, the resulting trajectory feature relies on prior segmentation and can be affected by the arm motions and erroneous estimations of the feet joints. Researchers used two earlier datasets [12,15].

Nguyen et al. [13] proposed a normal gait model build on skeleton joints to detect abnormal human gait. The authors decomposed each gait sequence into cycles. Each human instant posture was represented by a feature vector which described relationships between pairs of lower body bone joints. Selected features were: hip angles, knee angles, ankle angles, and the angle between two feet. The angles corresponding to flexion angles (one for each joint) were calculated via two planes computed with the three joints of each leg. The feature vectors were then converted into codewords using a clustering technique. The normal human gait model was created based on multiple sequences of codewords corresponding to different gait cycles. A Left-to-Right Markov chain was used to model the skeleton data. In the detection stage, a gait cycle with a normality likelihood below a selected threshold, which was determined automatically in the training step, was assumed to be an anomaly. The method has shown state-of-the-art results for binary gait assessment. However, the Kinect sensor fails to reliably detect low-limbs joints (particularly, ankles and feet joints [1]), so it is arguably to base a feature vector on them. Many researchers reported the inaccuracy of feet position estimations and corresponding angles [23] coming from Kinect v.2. Available gait datasets [16,7] also suffer from this issue. Researchers reported using a combined dataset from Vicon data and Kinect v.1 data. The dataset used is not available; however, later researchers evaluated this method on a newly shared dataset [16].

Li et al. [11] proposed a method to classify normal and abnormal gait based on skeleton data. They focused on the group of motion anomalies representative for Parkinson and Hemiplegia (tremor, partial paralysis, gestural rigidity, and postural instability). At the heart of their descriptor lie 2 covariance matrices. One for 24 mutual joint positions and another one for motion rate of the same 24 joints. A k-nearest neighbors based classifier and a custom covariance distance metric were used. Sequences were temporally segmented by a sliding window. A distinctive classifier was learned for the time windows to yield an importance weight assignment. The method gives good results with an average accuracy of 80%. However, the custom dataset used is not publicly available. Although the full list of body joints is commonly used to compose a covariance matrix, we believe that in many cases upper joint motion can be a source of ambiguity for the gait assessment task. In addition, the dataset exploited by the researchers in this work is

not available for reuse.

Sequential models are often used in gait research. Such models are popular because this technique is effective in describing the transition of human posture states during a gait cycle. Classically, Hidden Markov based models were exploited in gait research [13,15,24,25]. Lately LSTMs became popular and replaced HMMs in many applications. Both HMMs and LSTMs can capture transitional features along with the structural features extracted beforehand. LSTMs do not make the Markov assumption and so can, in theory, take into account long-term dependencies. HMMs are more applicable when the number of features is not excessive since they require less parameters to learn. That is why they are often applied in skeleton-based methods [26]. LSTM [27] recurrent neural networks can be used to learn the features for gait recognition or classification. Usually they are used in the same applications as HMMs, but lately, researchers reported a higher performance of LSTMs in traditional for HMM domains, such as handwriting recognition [28] and speech recognition. Similar to HMM, there are more examples of their usage with 2D data, or data coming from wearable sensors [29].

Only several recent methods adapted Recurrent Neural Networks for gait pathology detection. Liu et al. [30] proposed a Deep Rehabilitation Gait Learning (DRGL) for modeling the knee joints of a lower-limb exoskeleton, which leverage a simple one-layer LSTM to learn the inherent spatial-temporal correlations of gait features. Researchers sought to predict the abnormal knee joint trajectories based on the other joints. The method results were evaluated visually, by comparing the predicted and actual trajectory. The task addressed by the authors is different from the gait assessment targeted in our work, hence the model design and gait features are different. Feng et al. [31] trained an LSTM model on human joints data to characterize gait. An LSTM was used to model temporal parameters of the gait sequence. The hidden activation values of the Neural Network represented the final gait feature. Researchers used 2D data, extracting the joint features with a CNN, however the sequence part of the model can be easily used with skeleton data directly. The output of their LSTM model is the gait sequence representation, which can be used for gait recognition. The design of the model is not very well explained, but it seems to be a single-layer model. The closest to ours and the most recent is the work of Zhao et al. [32]. The authors adopted an LSTM net to diagnose Neurodegenerative diseases (ND). The two-layer model was trained and tested using temporal data that was recorded by force-sensitive resistors including time series, such as stride interval and swing interval. Researchers indicated the results allowing to diagnose particular ND.

3. Multi-modal gait symmetry database

A database of normal and pathological gait examples was collected in our laboratory. It is a multi-modal gait symmetry database, or MMGS. We used the single Kinect v.2 sensor placed in front of the test subject. Each person walked towards the camera. The setup is shown in Fig. 1 (left). The resulting acquisition volume is quite small and is limited by the Kinect v.2 field of view and distance range. It can be advantageous to use a treadmill for gait acquisition to be able to acquire more data, however, it will impose additional charges and difficulties,[2] so in this work we choose the simple setup stating its limitations.

27 persons participated in the experiment, among which 19 men and 8 women. The average age of participants is 30 years old (standard deviation is equal to 7 years). The information about test subjects can be found in Appendix A. All subjects were wearing shoes without heels. The acquisition protocol consisted of three steps. First, each person was instructed to walk normally with a self-assigned comfortable speed from a starting line towards the camera. Second, we added a padding sole into the right shoe of each patient. The height of the padding material is 7 cm. A similar sole padding to simulate a pathology was already performed in [16]. Padding allows us to simulate limping gait. Third, we asked the person to not bend the right knee during walking as it was done in previously by [12]. This test is designed to simulated gait problems which can be due to the wearing a prosthesis or characteristic for recovery after a fracture. Second and third parts of the experiment were performed after an initial time, when the person walked around and practiced the simulation. The first normal trial was executed immediately after a person was instructed. Each walk was executed between 5 and 7 times. In total, MMGS database contains 489 videos.

Examples of the depth image, silhouette, and a skeleton provided by Kinect v.2 and stored in our database are shown in Fig. 2. It should be noticed also that our system does not capture and store the color information, which guarantees the confidentiality for patients.

We stored the following type of data:

- Depth maps of the scene
- Binary masks with the silhouettes
- Skeleton joints 3D coordinates (X, Y, Z) and their state (i.e., *tracked, inferred, not tracked*)
- Joint orientations (q_w, q_x, q_y, q_z)

The data in the database are provided in two forms. First, we publish the data in the form they are delivered by Kinect v.2 sensor and corresponding SDK. Second, we added a pre-processed version, where skeleton sequences are segmented in cycles and filtered. Skeleton information consists of joints coordinates and bone orientations in form of quaternions. Orientation quaternion is a complex number with 4 dimensions, used to describe a rotation in 3d space. For our experiments, we pre-process the skeleton data from our database in the following way. First we exclude all the joint measurements where the joint status returned by the Kinect SDK is *not tracked* or *inferred*. This allows us to remove the data, where ML algorithms failed to track the joint. Next a low-pass filter is applied to the joints and quaternions data. A standard normalization is then done by subtracting the center of the spine joint from other joints for each frame. Then we segment the gait sequences into cycles. Such segmentation allows us to compare gait features in an easy fashion later. The sequences are segmented using the method proposed by [13], but in this work it is based on the right hip Z flexion angle peaks and not the intra-feet distance. After the segmentation, we obtain a total of 686 gait cycles. We normalize each cycle to have the same number of frames. For our tests we used the normalization to 30 frames. The normalization is not necessary, especially in the case when self-selected speed is an important characteristic; however, in our tests we wanted the speed factor not to affect the assessment procedure.

MMGS dataset skeletons part is available online: https://github.com/margokhokhlova/LSTM_gait_model. The data in txt format can be freely downloaded and used for research purposes.

4. Knee and hip flexion angles from the Kinect v.2

Low-limbs flexion angles (Fig. 3(a)) can be calculated from the skeleton data provided by the standard software of RGB-D cameras such as the Kinect v.2. Kinect provides the skeleton joint orientation values in the form of quaternions. A quaternion q is a set of 4 values: w and x, y, z (in the following, q_w, q_x, q_y, q_z). More details about quaternions can be found in [33] and in Appendix B.1. Guess et al. [10] compared the low-limbs flexion, extension, and abduction data obtained via Kinects' v.2 quaternions and Vicon. The comparison has shown that reasonable measurement of hip and knee flexion and rotation angles, as well as isolated hip abduction angles, can be obtained from the Kinect skeletal tracker for a person performing a jumping exercise. Inspired by [10], we adopted the proposed method to calculate low-limbs joint angles from the Kinect v.2 orientation information. Commonly, researches

[2] The price of the treadmill and the fact that the control panel commonly is in the front of the device, occluding the patient and making the joints estimation more error-prone.

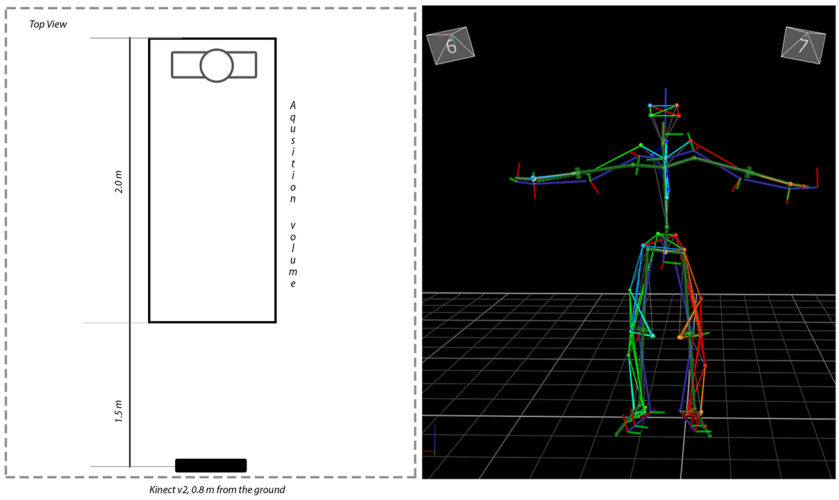

Fig. 1. (a) Acquisition setup with Kinect v.2 sensor: a subject walks towards the camera from a fixed starting point. (b) Skeleton obtained from Vicon markers using the custom software.

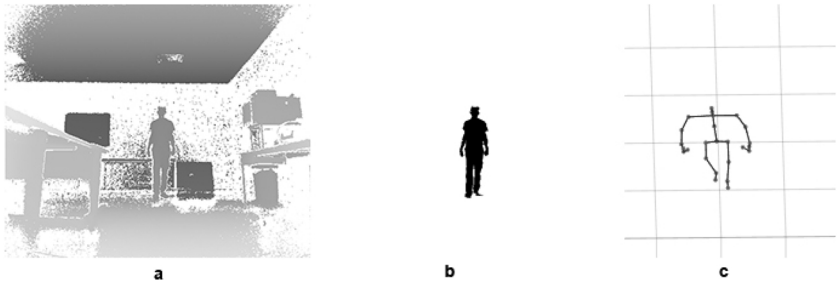

Fig. 2. (a) depth map (b) silhouette mask (c) skeleton (25 joints positions & orientations).

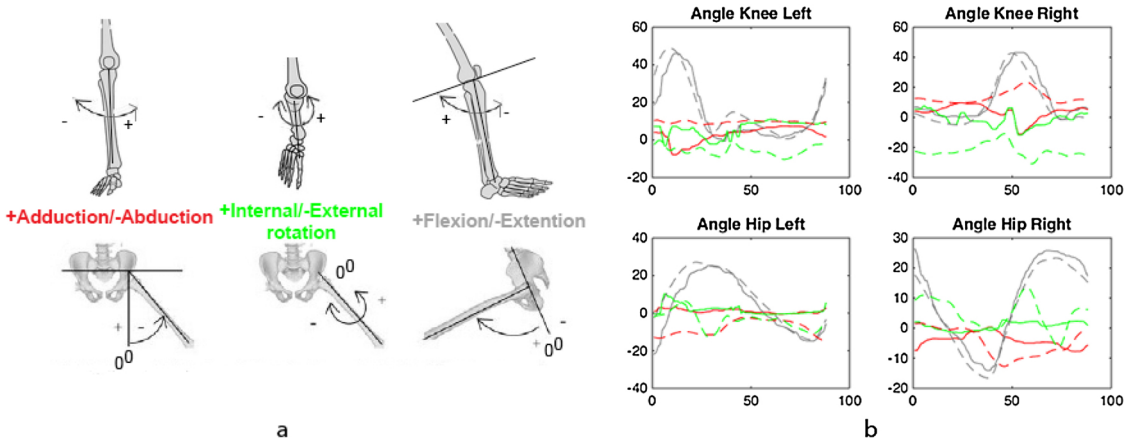

Fig. 3. (a) Knee (top) and hip (bottom) angles on a scheme. Adduction refers to a motion that pulls a part towards the midline of a limb. Abduction refers to a motion that pulls a part away from the mid-line of the limb. Internal rotation refers to rotation towards the axis of the body, external rotation refers to rotation away from the center of the body. Flexion describes a bending movement that decreases the angle between a segment and its proximal segment. (b) Angles calculated for one gait cycle by Kinect v.2 (solid) and Vicon (dashed). Visually, for a given cycle, the flexion angles correspond well, and so are rotation angles.

[13,11,15,14] only use skeleton joint data provided by Kinect and quaternions are mainly used for avatar animation. Quaternions provide with additional information and can be used to calculate flexion, abduction, and rotation limb angles, which are important characteristics for many gait deviations. Here we use quaternions to obtain the angles that match the Plug-in Gait kinematics data coming from the Vicon software. Using the orientation quaternion, we can calculate the rotation of the bone represented as a vector around the Z, X, and Y axis. These yaw (ϕ), roll (ψ), and pitch (χ) correspond to flexion/extension, adduction/abduction, and internal/external rotation when applied to the hip and knee (femur and tibia bones).

The Microsoft Kinect SDK version 2.0 provides us with the orientation of a joint in 3D camera global reference system of coordinates. The orientation of each joint is relative to the child joint (i.e., Knee to Ankle, Hip to Knee) while the Hip Center joint still contains the orientation of the person.

We calculate the knee and hip angles of a walking person via quaternions in the following way. We re-orient the hip quaternion to match the Vicon hip orientation as advised in [10]. The angle between a line connecting left and right hip joints and a line from either the left or

right hip to hip center joint was calculated for each frame. The left hip was then reoriented by rotating 270° about the *X*-axis, 270° about the *Z*-axis, and then rotating about the *Y*-axis by the previously calculated left hip angle. The right hip was then reoriented by rotating 270° about the *X*-axis, 90° about the *Z*-axis, and then rotating about the *Y*-axis by the previously calculated right hip angle.

To calculate the knee flexion angles, we need to use the knee joint and its child, the ankle joint orientations. These orientations are $q_k = (q_{w_k} + q_{x_k} i + q_{y_k} j + q_{z_k} k)$ for the knee quaternion and $q_a = (q_{w_a} + q_{x_a} i + q_{y_a} j + q_{z_a} k)$ for the ankle quaternion. The quaternions provided by Kinect 2.0 SDK satisfy the definition of a unit quaternion.

Each rotation, q_k and then q_a, is done in the absolute frame of reference so we use $q_a^{-1} \times q_k$ to obtain a quaternion q_{ka} $(q_{w_{\mathrm{ka}}}, q_{x_{\mathrm{ka}}}, q_{y_{\mathrm{ka}}}, q_{z_{\mathrm{ka}}})$ corresponding to the knee flexion angle.

Note that q_a^{-1} in our case is equal to the conjugate of a quaternion. Then q_{ka} is converted to Euler angles in predefined order, specified as '*ZXY*'. The conversion between quaternions and Euler angles is:

$$\begin{bmatrix} \phi \\ \chi \\ \psi \end{bmatrix} = \begin{bmatrix} \tan^{-1}\left(2\frac{q_w q_z + q_x q_y}{1 - 2(q_y^2 + q_z^2)}\right) \\ \sin^{-1}(2(q_w q_y - q_x q_z)) \\ \tan^{-1}\left(2\frac{q_w q_x + q_y q_z}{1 - 2(q_x^2 + q_y^2)}\right) \end{bmatrix} \quad (1)$$

A particular case arises when $|q_w q_y - q_x q_z|$ is equal to 1/2 [33].

In order to calculate the angle for the hip joint, we need to use the re-oriented hip joint and its child, the knee joint orientations. Let the q_h be the quaternion of a hip joint. Then q_{hk} is equal to:

$$q_{\mathrm{hk}} = q_k^{-1} \times q_h \quad (2)$$

Eq. (1) is used to obtain angles in radians, which are then converted to degrees. The knee-ankle and hip-knee angle visualizations for a gait cycle are shown in Fig. 3 (Right). We compared gait kinematic parameters between Kinect v.2 and Vicon systems and find the Kinect angles satisfactory except for the adduction/abduction angles.

5. Gait model

Flexion low-limb angles are widely used by clinicians to evaluate the gait of a person. Earlier, an experiment was performed to assess the reliability of these angles calculated from the Kinect v.2 camera. Here we describe the ML learning method, which classify the gait automatically based on the obtained low-limbs dynamics. We describe the proposed ensemble LSTM network, exploring several architectural variations. Long Short Term Memory networks are a special kind of Recurrent Neural networks, capable of learning long-term dependencies. Similar to the other ML-based techniques, an LSTM network should be trained on a significant amount of data. The optimal architecture and the choice for the hyper-parameters of the network depend on the specific problem.

Single LSTM net. Based on some initial tests, we adopted a 2-layers bi-directional LSTM model for the task of gait classification. Our test has shown the strong evidence that the use of both past (backwards) and future (forward) gait cycle states is beneficial for gait classification model. The exact architecture of a single model used in our work is shown in Fig. 4. More details on the LSTM cell structure are provided in Appendix B.3. We use a bidirectional LSTM model to capture the temporal information, and use a gait cycle features as an input for our model.

The model was implemented in Python using Keras framework 4 with Tensorflow as the backend. After some prior testing, we adopt the number of LSTM hidden units to be equal to 64. This gives us the model with 140, 163 parameters to train. For the classification task, we selected categorical cross-entropy loss function, improving the training and validation accuracy of the model. Adam optimizer [34] is used to control the learning rate, with the final optimal parameters learning rate = 0.0002, beta1 = 0.9, beta2 = 0.999, epsilon = 1e-08, no decay. We also use standard mini-batches technique for training with the size of samples equal to 21. Each of our architectures was trained end-to-end using a stopping criterion, reducing the learning rate (by a factor of 0.1) in case when a validation accuracy metric has stopped improving. 50 epochs were used, since we do not have a very big dataset and input features are low-dimensional.

Ensemble LSTM model. As it is shown in the next section, single bi-directional LSTM model described earlier gives good results, however, we discovered that the performance varies depending on the training/validation data partitioning. In order to achieve better and more consistent results, a ensemble model was designed, in which 5 separate LSMT models are trained on different training/validation partitioning, and the final label for the test set is assigned based on the elaborated weighting scheme. The scheme weights the assigned probability for each model based on its validation accuracy. Each single LSTM output is a 3×1 probability vector, we scale these probabilities by the validation accuracy of the corresponding model. The final label is then assigned based on the mean scaled probability by 5 single models. The described model architecture is shown in Fig. 5. The 5 models Softmax weighted output example is shown in Fig. 6(a).

6. Experiments

The experimental part is divided into two main tasks. First, we compare the low-limbs kinematic parameters calculated using the Kinect's v.2 data with the golden standard measurements of Vicon. Second, we demonstrate the LSTM ensemble model-based application which uses the obtained kinematic parameters to automatically classify normal/pathological gait.

6.1. Validation of kinematic parameters from Kinect v.2 in comparison with Vicon

Previous section describes how to calculate low-limbs dynamics using the Kinect v.2 orientation data. Here we validate our results by comparing the obtained angles with the ones from a MOCAP system. The second database capturing walking data simultaneously with Kinect v.2 and Vicon sensor was collected in clinical settings. The purpose of this data was to validate the method used to calculate the low-limbs 3D flexion angles from Kinect orientation data presented in the paper by Guess et al. [10]. The authors used Kinect to evaluate the flexion dynamics of a person performing a vertical jump. The collected dataset contains gait sequences of 2 healthy persons, each has performed 11 walking trials. Kinects skeleton joints, orientations, and joint states (i.e., *tracked*, *not tracked*, *inferred*) are saved. Marker coordinates and flexion angles from the Plug-in Gait Vicon module are saved for marker-based MOCAP system. The acquisition was performed in a platform of the University Medical Center (CHU) in Dijon, where the Vicon camera is employed to perform movement analysis. A single experienced researcher placed retro-reflective markers on every participant. All markers were placed directly on the skin for a higher accuracy. The acquisition setup is shown in Fig. 1. The person walked towards Kinect from a distance of 6 m and stopped in 1.5 in front of the sensor. Kinect detects a person approximately on a distance of 5 m, when he/she is walking normally. This gives us on average about 2–3 ,m of walking distance. All subjects walked bare-foot. They were instructed to perform 11 trials. The acquisition process was started simultaneously by the Kinect and Vicon. Marker trajectories from Vicon were saved with a sampling frequency of 120 Hz. Kinect data frequency is about 30 Hz. We later downsample the Vicon data to have the same frequency as Kinect.

We assess the reliability of the knee and hip flexion angles calculated from Kinect v.2 data. The data from Vicon and Kinect were

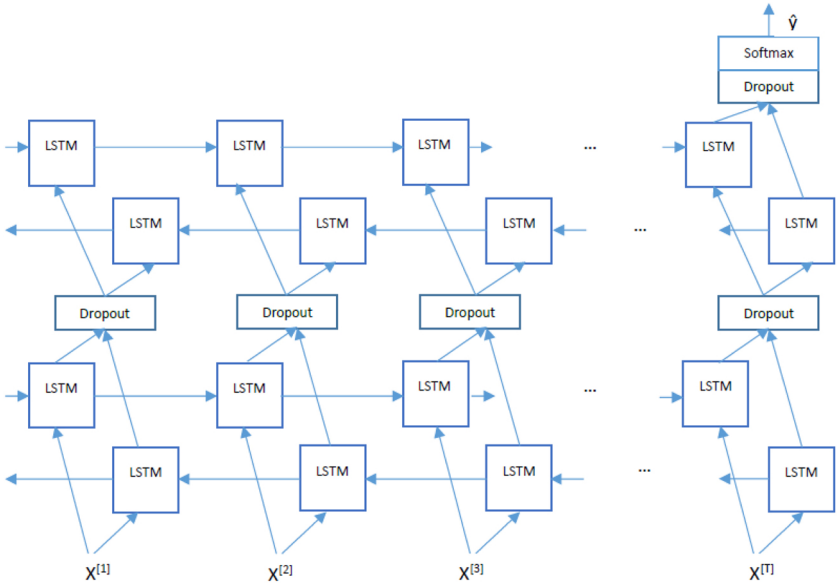

Fig. 4. Adopted bi-directional LSTM structure: input is the flexion angles, output is the assigned class.

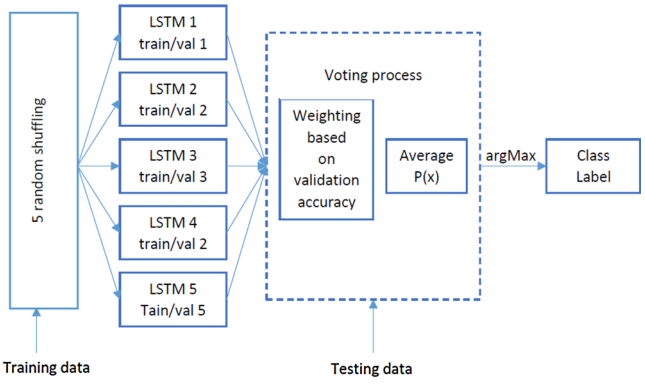

Fig. 5. Adopted ensemble model structure. Final label is assigned by voting of 5 models trained on randomly selected training/validation persons.

manually segmented into cycles. For this experiment, we segment the cycles by identifying corresponding intra-feet joints distance peaks for Kinect and Vicon as it was proposed in [13]. An example of calculated kinematic parameters for a cycle from Kinect v.2 and Vicon data can be found in Fig. 3(b).

We measure the correlation coefficients to estimate how strong a relationship between Kinect v.2 and Vicon kinematic data is. We choose to use correlation index coefficient for consistency for single rater/measurement ICC(C,1), also known as norm-referenced reliability. More details about ICC can be found in Appendix B.2. The index was calculated for 32 gait cycles acquired by Kinect and Vicon simultaneously. Each trial had just 1 or 2 complete cycles estimated by the right foot floor contact periodicity. Even knowing the Vicon and Kinect frame rate, we did not obtain the perfect cycles alignment with our segmentation algorithm (see Fig. 3(b)).

This study demonstrates that overall moderate measurement of hip and knee flexion angles during walking can be obtained from the Kinect v.2 skeletal tracker. Our results are shown in Table 2, fair to good reliability is shown in bold in agreement with the ICC evaluation recommendations [35]. Values less than 0.5 are indicative of poor reliability, values between 0.5 and 0.75 indicate moderate reliability, values between 0.75 and 0.9 indicate good reliability, and values greater than 0.90 indicate excellent reliability.

Our results signify worse overall agreement than the ones reported by [10], but follow the same trend: flexion/extension and internal/external rotation angles are correlated with the Vicon measurements while the abduction/abbuction angles are not. The smaller values of the ICC in our case can be due to the fact that Kinect provides fewer accurate measurements within specific distance range [2]. Guess et al. [10] used an optimal capture space for their experiment; however, we need more space to perform a gait acquisition. Nonetheless, we obtain moderate reliability for knee flexion and good reliability for hip flexion. The confidence intervals are small. Rotation angles signify poor

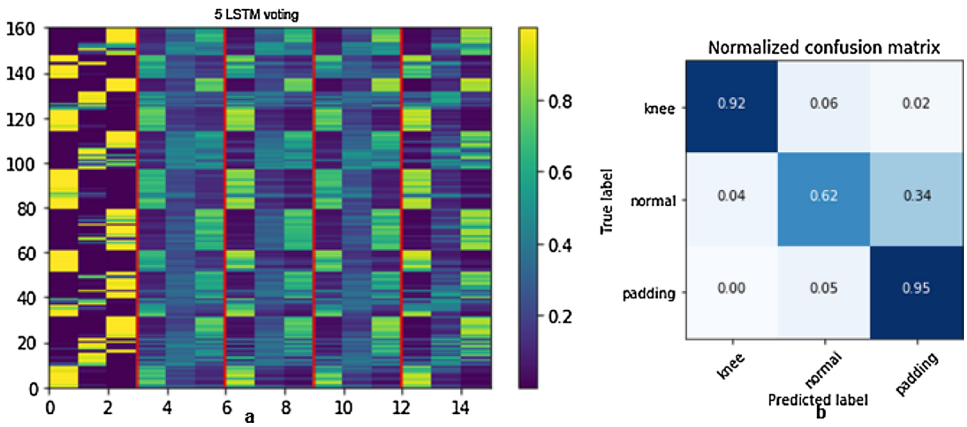

Fig. 6. Left: five LSTM models Softmax outputs: each LSTM output is a 3x1 probability vector, color codes the probability value. Right: final normalized confusion matrix for 3 classes. Two pathological gaits are detected by the system with a high accuracy. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

Table 2
ICC(C,1) correlation index and its 95% confidence interval for Vicon and Kinect angles.

Joint	Flexion/Extension ICC(C,1)	Abduction/Adduction ICC(C,1)	Rotation Internal/ External ICC(C,1)
Left Knee	**0.69** (0.67–0.71)	−0.12 (−0.15–0.09)	0.35 (0.31–0.38)
Right Knee	**0.73** (0.71–0.75)	−0.37 (−0.04–0.33)	0.05 (0.01–0.08)
Left Hip	**0.85** (0.84–0.86)	−0.18 (−0.22–0.15)	**0.45** (0.42–0.48)
Right Hip	**0.76** (0.74–0.78)	0.14 (0.10–0.17)	0.32 (0.28–0.36)

reliability, but can still be interesting for a coarse analysis. Abduction angles were judged as not correlated with the exception of the right hip data.

6.2. Validation of the gait classification model

Our MMGS dataset is used to train the gait model and validate the results. Flexion and rotation (ϕ and χ) angles calculated for a gait cycle by the method described in Section 4 are the model input. Each cycle is divided into 10 equal parts. The number 10 was found experimentally, after testing the division in 6, 10, 15 and 30 parts. The angles are averaged in each part. Final gait feature has the dimensionality 8×10. We also experimented with the option when all three angles (ϕ, χ and ψ) are used and different number of parts in cycle, but got slightly worse results. In order to train and evaluate the model, we need to divide dataset into training, validation, and testing data. This is done in the following way: we use 19 persons for training and validation (13 and 6 correspondingly) and 8 persons are left for testing. This guarantees the generalization of the algorithm for the data of new, unseen by the algorithm subjects. We perform the selection of the persons for training and testing data randomly and repeat the procedure several times, each time training a new LSTM model, in order to show it is independence of the particular data partitioning. For the final result, such random partitioning was performed 23 times.

Our results can be found in Table 3. A single LSTM model actually gives good results, however, we have noticed that, depending on the persons' distribution in training/validation data, possible deviations are about 3%. This can further increase significantly based on the training/testing subject assignment as can be seen by the accuracy distribution. The ensemble model architecture was proposed in order to decrease the variance of each model and its dependency on the test partitioning. Training/validation partitioning affects the results, so we generalize them by making 30 trials for the ensemble model, which means that each single LSTM was trained 150 times. Table 3 provides with the mean results value over 30 trials. We handle the problem of variance by using the ensemble model for pathological gait assignment. However, we also have high bias between training and validation results, so more training data will be beneficial for the model.

The confusion matrix is shown in Fig. 6(b).

We performed exhaustive testing to check if the resulting score of the algorithm does no depend on the data distribution. However, to make it easier to compare with our results in the future studies, Table 4 provides with the single ensemble model data, with training/validating/testing subjects specified. We specifically selected testing subjects so that there are both female and male subjects in the testing data. The corresponding confusion matrix can be found in Fig. 7(a). The results obtained with this set of persons are higher than average model accuracy demonstrated earlier, but fixed distribution allowed as to fine-tune all the parameters for the given model.

Table 3
LSTM model accuracy.

	Single LSTM			Ensemble Model		
	Mean	Min	Max	Mean	Min	Max
Test accuracy (%)	**77**	71	94	**82**	75	91

We also compared the proposed kinematic features with the use of 24 joint positions from Kinect used in [11,26,36]. The skeletons were normalized as described earlier in Section 3. We used the same model, same data partitioning and same parameters as in the previous test, but replaced the angles to 34 joints positions in XYZ, which gives us a 72×15 feature for each gait cycle. 15 represents the number of inputs from one cycle, maximizing the accuracy of the model with joints data. The final ensemble model accuracy is just 48%, which shows the advantage of our kinematic features over skeleton joints.

Finally, to justify the use of our LSTM model and not a simple classifier, we trained the SVM classifier [37] using the same data partitioning used in the previos test (Testing persons: 1, 9, 15, 18, 20, 21, 23, 25, Training and validation persons: 2, 3, 4, 5, 6, 7, 8, 10, 11, 12, 13, 14, 16, 17, 19, 22). For this test we used standard normalization towards mean, calculating μ and σ from the training part of the dataset.[3] We use the sklearn python library implementation of the SVM [38] and nested cross-validation to optimize the hyper parameters. We also tested different kernels (linear, RBF and polynomial), the polynomial kernel has demonstrated the best results in terms of validation accuracy. The final best parameters for the kernel are: penalty = 100, degree = 3. Cross-validation on 10 different validation partitioning has given an average accuracy of 0.785 with standard deviation of 0.058. The final test accuracy is 0.776. The results show that out kinematic features can be used for gait classification, however, temporal relation in a gait cycle captured by the LSTM model is important for the optimal results. The corresponding confusion matrix is shown in Fig. 7(b).

From these results, we can conclude the following. Overall, the proposed system is comparable to the accuracy reported by other studies in the domain [11,13,15]. The latest works by Li et al. [11] report 80% accuracy for calssification of Parkinson Disease and Hemiplegia on a small custom dataset, Nguyen et al. [13,16] obtained 81% accuracy with their database acquired by the Kinect device, and 86% using mixed Vicon/3D sensor data for the task of binary normal/pathological gait classification. The proposed features perform better for gait assessment than the skeleton joints positions. Moreover, we use more data than [11,15] and assign three classes, but not a binary pathology detection as [13]. Hip and flexion angles work extremely good for recognizing of the knee rigidity problem. A limp gait pathology simulated by padding is harder to separate from the normal walk data. We see a solution in a search for a reliable way to estimate ankle kinematics from point cloud data directly.

The system accuracy is not high enough for clinical studies, but the results are stimulating. Although the researchers are used to deal with confusion matrices provided in Figs. 6(b), 7, the clinicians usually look into other parameters to evaluate the performance of a diagnostic tool. Table 5 reports the sensitivity ($TP/(TP + FN)$) and specificity ($TN/(TN + FP)$) of the system for two pathological classes. To make it easier for future comparisons, we report the results for the Model from Table 4. Please note that our mean data for 23 trials from Fig. 6 shows similar trend, even more pronounced in case of two pathological classes.

7. Conclusion

It has been shown that Kinect v.2 sensor has potential for abnormal gait detection. However, the number of publicly available datasets is very small which makes it difficult to generalize the results and compare the existing algorithms. In this paper, we propose a gait assessment database for the use in clinical gait analysis research. The database is

[3] We did not obtain convergence for different sets of SVM kernel parameters without normalization.

Table 4
Accuracy on a particular data partitioning.

Persons		LSTM1	LSTM2	LSTM3	LSTM4	LSTM5	Ensemble model
Training + validation: 2,3,4,5,6,7,8,10,11,12,13, 14,16,17,19,22,24,26,27	Val	79	90	80	77	82	n/a
Test: 1,9,15,18,20,21,23,25	Test	85	89	82	83	88	91

Table 5
Sensitivity and specificity on a particular data partitioning.

Category (simulation)	TP	FP	FN	TN	Sensitivity, %	Specificity, %
Knee rigidity (no bending)	70	2	4	133	95	98
Limping (sole padding)	62	4	9	134	87	97

available for reuse in research purposes, without the need for individual applications, through open access. The proposed MMGS database capture more individuals than the existing databases, and provides the data of different modalities. We hope that this database will allow advancing the Computer Vision based clinical gait research. We propose an enhanced gait representation for abnormal gait detection. Our novel approach tackles the problem of low-limb kinematic symmetry description and modeling via an LSTM Network. We calculated the kinematic gait parameters related to the knee and hip using the Kinect v.2 data, compared them with the Vicon data, and used them in a gait assessment sequential model. Experimental results demonstrate the effectiveness of the proposed approach in the context of automatic abnormal gait detection and classification. We extensively evaluated our technique. Our results are comparable with the state-of-the-art results gait assessment methods [11,13]. The features we use correspond to the ones evaluated by clinicians and can be used on their own for visual evaluation, or as part of a complete gait assessment system. The proposed classification ensemble model architecture has shown itself as a powerful tool, and is available for reuse. We believe that this sets a ground for future works in automatic gait classification.

Based on the performed research work, it can be confirmed that Microsoft Kinect v.2 can be conveniently exploited for gait assessment. Low-limb kinematics calculated from the Kinect data can be used to classify gait deviations affecting gait symmetry: limping, rigidity, and possibly others. However, we envisage further work in order to increase the accuracy and reliability of the skeletonization algorithms.

Future work is to make a continuous measure of movement quality based on a normal gait model. We consider the possibility to estimate classical gait assessment indexes such as The Gillette Gait Index (GGI) [39] with the Kinect sensor data. We also seek the possibility to perform the data acquisition with real patients. We envisage more advanced research directions. In the last years, 3D shape completion methods from point clouds had shown good results [40] and even 2D images [41]. Using these algorithms, the shape and posture of a patient can be successfully restored and used for the analysis. Adaptation of such a technique can lead to very promising results in gait assessment studies. Ideally, a patient model can be acquired prior to the walking test, and then, when the shape is restricted, it will be easier to estimate the posture more accurately in a temporal sequence and evaluate the motion of different segments.

Conflict of interest

The authors declare no conflict of interest.

Funding information

This research was funded by Region of Burgundy. Alexei Morozov and Olga Sushkova express their gratitude to Russian Foundation for Basic Research (grant number 16-29-09626-ofi-m). A special gratitude is given to the Region of Burgundy, who financed the PhD research of Khokhlova Margarita.

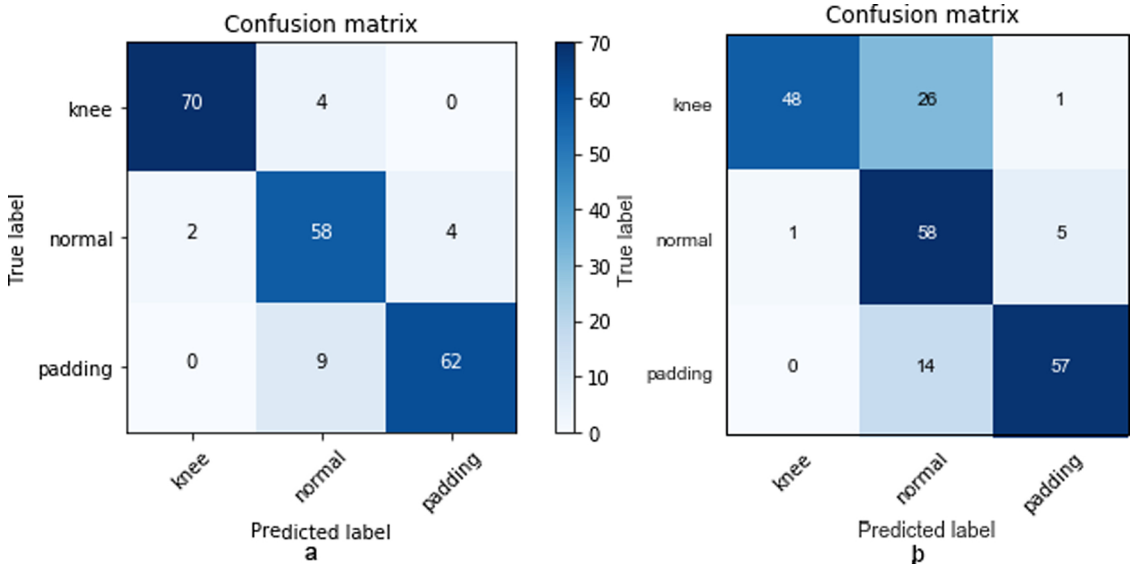

Fig. 7. Confusion matrix for a particular data partitioning. Train and Validation persons: 2,3,4,5,6,7,8,10,11,12,13,14,16,17, 19,22. Test persons: 1,9,15,18,20,21,23, 25. (a) ensemble LSTM model confusion matrix for testing data. (b) SVM confusion matrix for testing data.

Appendix A. Test subjects

Detailed information about the database participants is summarized here in Table A.6. This can be used in case of a need to normalize the data, or perform a gender recognition.

The sole we used to simulate limping gait in the database is shown in Fig. A.8.

Table A.6
Custom dataset composition. 19 male and 3 female subjects of different age participated in the data acquisition.

Num	Gender	Age	Height, cm	Weight, kg
1	M	23	182	89
2	M	26	167	67
3	M	34	188	106
4	M	30	180	65
5	M	28	170	73
6	M	31	176	65
7	M	23	170	71
8	F	55	156	58
9	M	42	175	85
10	M	27	170	75
11	M	28	175	70
12	F	28	176	58
13	F	25	171	53
14	M	47	178	70
15	M	25	175	84
16	M	24	181	86
17	M	35	176	65
18	M	29	183	70
19	M	28	180	80
20	M	28	172	65
21	M	34	175	63
22	M	24	180	65
23	F	30	160	47
24	F	24	172	52
25	F	26	172	62
26	F	27	162	45
27	F	34	160	73

Fig. A.8. Sole padding used in this work.

Appendix B. Tools and concepts used in this work

B.1 Quaternions

In mathematics, the quaternions are a number system that extends the complex numbers. A 6D pose can be described as a displacement in 3D plus a rotation defined by means of a specific case of Euler angles, which can be represented via quaternions. Quaternions find uses in applied mathematics, in particular for calculations involving three-dimensional rotations such as in three-dimensional computer graphics, and computer vision. A quaternion q is a set of 4 values: q_x, q_y, q_z, and q_w. Using the orientation quaternion, we can calculate the rotation of the bone around the Z, X, and Y axis. The quaternions provided by Kinect 2.0 SDK satisfy the definition of a unit quaternion:

$$N(q) = 1; \quad N(q) = N(q_{w0} + q_{x0}i + q_{y0}j + 1_{z0}k) = q_w^2 + q_{x0}^2 + q_{y0}^2 + q_{z0}^2 = 1 \tag{B.1}$$

The conjugate of a quaternion $q = q_w + q_x i + q_y j + q_z k$ is defined by:

$$q^- = q_w - q_x i - q_y j - q_z k \tag{B.2}$$

Multiplication of two quaternions $q_1 = q_{1w} + q_{1x}i + q_{1y}j + q_{1z}k$ and $q_2 = q_{2w} + q_{2x}i + q_{2y}j + q_{2z}k$ combines two rotations in the following way as a dot product. The new quaternion's $q_3 = q_{3w}a + q_{3x}i + q_{3y}j + q_{3z}k$ values a, i, j, k are then obtained as:

$$
\begin{aligned}
a &= q_{2_w} q_{1_w} - q_{2_x} q_{1_x} - q_{1_y} q_{1_y} - q_{2_z} q_{1_z} \\
i &= q_{2_w} q_{1_w} + q_{2_x} q_{1_x} - q_{1_y} q_{1_y} + q_{2_z} q_{1_z} \\
j &= q_{2_w} q_{1_w} + q_{2_x} q_{1_x} + q_{1_y} q_{1_y} - q_{2_z} q_{1_z} \\
k &= q_{2_w} q_{1_w} - q_{2_x} q_{1_x} + q_{1_y} q_{1_y} + q_{2_z} q_{1_z}
\end{aligned} \tag{B.3}
$$

More details about quaternions can be found in [33] and in the Appendix B.1.

B.2 Reliability via correlation coefficient

Reliability is defined as the extent to which measurements can be replicated. In other words, it reflects not only degree of correlation but also agreement between measurements. Historically, Pearson correlation coefficient, paired t test, and Bland-Altman plot have been used to evaluate reliability. However, paired t test and Bland-Altman plot are methods for analyzing agreement, and Pearson correlation coefficient is only a measure of correlation, and hence, they are non-ideal measures of reliability [42]. Intra-class correlation coefficient (ICC) is an index to measure a reliability which reflects both degree of correlation and agreement between measurements. McGraw [35] defined 10 forms of ICC based on the Model (1-way random effects, 2-way random effects, or 2-way fixed effects), the Type (single rater/measurement or the mean of k raters/measurements), and the Definition of relationship considered to be important (consistency or absolute agreement). Using the recommendations of [35], for our intra-systems comparisons we selected the correlation coefficient for consistency for single rater/measurement (ICC(C,1)). The confidence interval selected is 0.05 (95%). Confidence interval has upper and lower bound, which are constructed so that the designated proportion (confidence level) of such intervals will include the true population value.

Commonly, the ICC agreement is evaluated using guidelines by Cicchetti [43] for interpretation ICC inter-rater agreement measures:

- Less than 0.40 — poor.
- Between 0.40 and 0.59 — fair.
- Between 0.60 and 0.74 — good.
- Between 0.75 and 1.00 — excellent.

B.3 LSTM

LSTM is a specific recursive neural network with a set of memory blocks. Each block contains inputs, outputs, forget gate units, and one or more self-connected memory cells. There several slightly different LSTM architectures, so here we introduce the mostly used one. The block diagram of the LSTM memory block is shown in Fig. B.9. An LSTM cell c and the candidate value for updating it, $\tilde{c}^{\langle t \rangle}$ are design in the following way. The three gates can perform update (Γ_u), forget (Γ_f) operations for the cells:

$$
\begin{aligned}
\tilde{c}^{\langle t \rangle} &= \tanh(W_c[a^{\langle t-1 \rangle, x^{\langle t \rangle}}] + b_c) \\
\Gamma_u &= \sigma(W_u[a^{\langle t-1 \rangle}, x^{\langle t \rangle}] + b_u) \\
\Gamma_f &= \sigma(W_f[a^{\langle t-1 \rangle}, x^{\langle t \rangle}] + b_f) \\
\Gamma_o &= (W_o[a^{\langle t-1 \rangle}, x^{\langle t \rangle}] + b_o) \\
c^{\langle t \rangle} &= \Gamma_u * \tilde{c}^{\langle t \rangle} + \Gamma_f * c^{\langle t-1 \rangle} \\
a^{\langle t \rangle} &= \Gamma_o * c^{\langle t \rangle}
\end{aligned} \tag{B.4}
$$

Γ_o is the output gate, b – biases.

In practice, LSTM are considered to be more powerful and general than Gated Recurrent Unit (GRU) models. Because the state of LSTM is regulated much more delicately than that of the basic RNN, the LSTM cell can learn very long term relations in the data. Multiple LSTM cells can be stacked for more expressive power.

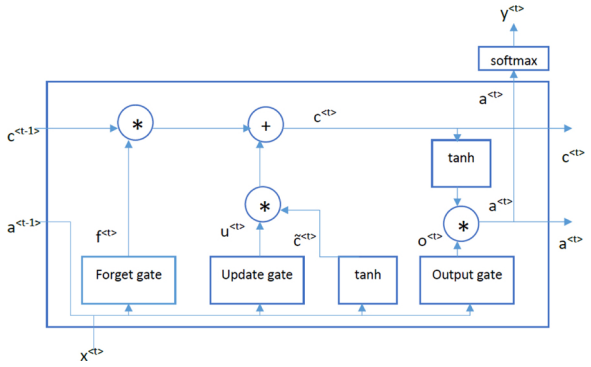

Fig. B.9. LSTM unit as introduced in Andrew Ng Sequential Models course.

Appendix C. Train, validation, and test results

Some examples of the results we obtained are grouped in Table B.7.

The database (joints data only, please contact us for depth images) and pre-trained LSTM model are available from: https://github.com/margokhokhlova/LSTM_gait_model.

Table B.7
Examples of the concrete single LSTMs and ensemble model accuracy on particular data partitioning (using 22 persons dataset).

Train\Test	Validation	Train acc, %	Val acc, %	Test acc, %	Ensemble model test acc, %
1,2,3,6,7,8,9,10,	1,6,10,17,18	55	96	90	80
11,13,14,15,17,	1,8,10,11,13	80	88	77	
18,19,21	1,3,6,10,11	83	84	77	
\	9,13,17,19,21	57	80	72	
4,5,12,16,20,22	2,8,14,17,18	87	75	76	
1,2,4,5,8,10,11,	1,4,17, 18,21	88	72	86	84
13,14,15,16,17,	10,11,16,18,20	87	78	82	
18,19,20,21	13,14,15,16,21	84	70	85	
\	10,11,15,17,19	77	83	82	
3,6,7,9,12,22	1. 5. 15. 16. 19.	84	84	75	
2,3,5,6, 8,9,10,11,	5,6, 9,11,21	83	72	80	**77**
12,13,16,17,	8,9,10,13,16	80	92	78	
\	2,5,11,18,21	78	78	75	
18,20,21,22	6,12,13,20,22	85	72	73	
1,4,7,14,15,19	3,5,6,8,20	86	81	75	

References

[1] Geerse DJ, Coolen BH, Roerdink M. Kinematic validation of a multi-kinect v2 instrumented 10-meter walkway for quantitative gait assessments. PLOS ONE 2015;10(10). https://doi.org/10.1371/journal.pone.0139913.

[2] Geerse D, Coolen B, Kolijn D, Roerdink M. Validation of foot placement locations from ankle data of a kinect v2 sensor. Sensors 2017;17(10):2301.

[3] Mentiplay BF, Perraton LG, Bower KJ, Pua Y-H, McGaw R, Heywood S, et al. Gait assessment using the microsoft xbox one kinect: concurrent validity and inter-day reliability of spatiotemporal and kinematic variables. J Biomech 2015;48(10):2166–70.

[4] Springer S, Yogev Seligmann G. Validity of the kinect for gait assessment: a focused review. Sensors 2016;16(2):194.

[5] Bobillo F, Dranca L, Bernad J. A fuzzy ontology-based system for gait recognition using kinect sensor. International conference on scalable uncertainty management. 2017. p. 397–404.

[6] Kastaniotis D, Theodorakopoulos I, Economou G, Fotopoulos S. Gait based recognition via fusing information from Euclidean and Riemannian manifolds. Pattern Recognit Lett 2016;84:245–51.

[7] Kastaniotis D, Theodorakopoulos I, Fotopoulos S. Pose-based gait recognition with local gradient descriptors and hierarchically aggregated residuals. J Electron Imaging 2016;25(6):063019.

[8] Ahmed F, Paul PP, Gavrilova ML. Dtw-based kernel and rank-level fusion for 3d gait recognition using kinect. Vis Comput 2015;31(6–8):915–24.

[9] Roberts M, Mongeon D, Prince F. Biomechanical parameters for gait analysis: a systematic review of healthy human gait. Phys Therapy Rehabil 2017;4(1):6.

[10] Guess TM, Razu S, Jahandar A, Skubic M, Huo Z. Comparison of 3d joint angles measured with the kinect 2.0 skeletal tracker versus a marker-based motion capture system. J Appl Biomech 2017;33(2):176–81. https://doi.org/10.1123/jab.2016-0107.

[11] Li Q, Wang Y, Sharf A, Cao Y, Tu C, Chen B, et al. Classification of gait anomalies from kinect. Vis Comput 2018;34(2):229–41. https://doi.org/10.1007/s00371-016-1330-0.

[12] Chaaraoui AA, Padilla-López JR, Flórez-Revuelta F. Abnormal gait detection with rgb-d devices using joint motion history features. 11th IEEE international conference and workshops on automatic face and gesture recognition (FG), vol. 7 2015. p. 1–6. https://doi.org/10.1109/FG.2015.7284881.

[13] Nguyen T-N, Huynh H-H, Meunier J. Skeleton-based abnormal gait detection. Sensors 2016;16(11):1792. https://doi.org/10.3390/s16111792.

[14] Rocha AP, Choupina H, Fernandes JM, Rosas MJ, Vaz R, Cunha JPS. Parkinson's disease assessment based on gait analysis using an innovative rgb-d camera system. 36th annual international conference of the IEEE engineering in medicine and biology society (EMBC) 2014:3126–9. https://doi.org/10.1109/EMBC.2014.6944285.

[15] Paiement A, Tao L, Hannuna S, Camplani M, Damen D, Mirmehdi M. Online quality assessment of human movement from skeleton data. British machine vision conference 2014:153–66.

[16] Nguyen T-N, Meunier J. Walking gait dataset: point clouds, skeletons and silhouettes, Tech. Rep. 1379. DIRO, University of Montreal; 2018.

[17] Hofmann M, Geiger J, Bachmann S, Schuller B, Rigoll G. The tum gait from audio, image and depth (gaid) database: multimodal recognition of subjects and traits. J Vis Commun Image Represent 2014;25(1):195–206.

[18] Moore JK, Hnat SK, van den Bogert AJ. An elaborate data set on human gait and the effect of mechanical perturbations. PeerJ 2015;3:e918. doi:10.7717.

[19] Gates DH, Darter BJ, Dingwell JB, Wilken JM. Comparison of walking overground and in a computer assisted rehabilitation environment (caren) in individuals with and without transtibial amputation. J Neuroeng Rehabil 2012;9(1):81.

[20] Gabel M, Gilad-Bachrach R, Renshaw E, Schuster A. Full body gait analysis with kinect. Engineering in medicine and biology society (EMBC), annual international conference of the IEEE 2012:1964–7.

[21] Kastaniotis D, Economou G, Fotopoulos S, Kartsakalis G, Papathanasopoulos P. Using kinect for assesing the state of multiple sclerosis patients. 2014 EAI 4th international conference on wireless mobile communication and healthcare (Mobihealth). 2014. p. 164–7.

[22] Devanne M, Wannous H, Daoudi M, Berretti S, Del Bimbo A, Pala P. Learning shape variations of motion trajectories for gait analysis. 23rd international conference on pattern recognition (ICPR) 2016. p. 895–900. https://doi.org/10.1109/ICPR.2016.7899749.

[23] Eltoukhy M, Oh J, Kuenze C, Signorile J. Improved kinect-based spatiotemporal and kinematic treadmill gait assessment. Gait Posture 2017;51:77–83.

[24] Papageorgiou XS, Chalvatzaki G, Tzafestas CS, Maragos P. Hidden Markov modeling of human pathological gait using laser range finder for an assisted living intelligent robotic walker. 2015 IEEE/RSJ international conference on intelligent robots and systems (IROS). 2015. p. 6342–7.

[25] Gu J, Ding X, Wang S, Wu Y. Action and gait recognition from recovered 3-d human joints. IEEE Trans Syst Man Cybern Part B (Cybern) 2010;40(4):1021–33.

[26] Drumond RR, Marques BAD, Vasconcelos CN, Clua E. Peek – an LSTM recurrent network for motion classification from sparse data. Proceedings of the 13th international joint conference on computer vision, imaging and computer graphics theory and applications – volume 1: GRAPP. 2018. p. 215–22.

[27] Hochreiter S, Schmidhuber J. Long short-term memory. Neural Comput 1997;9(8):1735–80.

[28] Graves A, Schmidhuber J. Offline handwriting recognition with multidimensional recurrent neural networks. Advances in neural information processing systems 2009:545–52.

[29] Ordóñez FJ, Roggen D. Deep convolutional and LSTM recurrent neural networks for multimodal wearable activity recognition. Sensors 2016;16(1):115. https://doi.org/10.3390/s16010115.

[30] Liu D-X, Du W, Wu X, Wang C, Qiao Y. Deep rehabilitation gait learning for modeling knee joints of lower-limb exoskeleton. 2016 IEEE international conference on robotics and biomimetics (ROBIO). 2016. p. 1058–63.

[31] Feng Y, Li Y, Luo J. Learning effective gait features using LSTM. 23rd international conference on pattern recognition (ICPR). 2016. p. 325–30.

[32] Zhao A, Qi L, Li J, Dong J, Yu H. LSTM for diagnosis of neurodegenerative diseases using gait data. Ninth international conference on graphic and image processing (ICGIP 2017), vol. 10615 2018. p. 106155B. https://doi.org/10.1117/12.2305277.

[33] Blanco J-L. A tutorial on Se (3) transformation parameterizations and on-manifold

optimization. University of Malaga; 2010. Tech. Rep 3.
[34] Kingma DP, Ba J. Adam: a method for stochastic optimization. 2014arXiv preprint arXiv:1412.6980.
[35] McGraw KO, Wong SP. Forming inferences about some intraclass correlation coefficients. Psychol Methods 1996;1(1):30.
[36] Zhu W, Lan C, Xing J, Zeng W, Li Y, Shen L, et al. Co-occurrence feature learning for skeleton based action recognition using regularized deep lstm networks. AAAI, vol. 2 2016:6.
[37] Vapnik V, Mukherjee S. Support vector method for multivariate density estimation. Advances in neural information processing systems 2000:659–65.
[38] Pedregosa F, Varoquaux G, Gramfort A, Michel V, Thirion B, Grisel O, et al. Scikit-learn: machine learning in python. J Mach Learn Res 2011;12(October):2825–30.
[39] Wren TA, Do KP, Hara R, Dorey FJ, Kay RM, Otsuka NY. Gillette gait index as a gait analysis summary measure: comparison with qualitative visual assessments of overall gait. J Pediatr Orthop 2007;27(7):765–8.
[40] Litany O, Bronstein A, Bronstein M, Makadia A. Deformable shape completion with graph convolutional autoencoders. 2017arXiv preprint arXiv:1712.00268.
[41] Bogo F, Kanazawa A, Lassner C, Gehler P, Romero J, Black MJ. Keep it SMPL: automatic estimation of 3d human pose and shape from a single image. European conference on computer vision. 2016. p. 561–78.
[42] Koo TK, Li MY. A guideline of selecting and reporting intraclass correlation coefficients for reliability research. J Chiropr Med 2016;15(2):155–63.
[43] Cicchetti DV. Guidelines, criteria, and rules of thumb for evaluating normed and standardized assessment instruments in psychology. Psychol Assess 1994;6(4):284.