

Gait acquisition and analysis system for osteoarthritis based on hybrid prediction model

Fang Chen^{a,*}, Xiwen Cui^b, Zhe Zhao^c, Daoqiang Zhang^a, Cong Ma^b, Xinran Zhang^b, Hongen Liao^{b,**}

^a Department of Computer Science and Engineering, Nanjing University of Aeronautics and Astronautics, MIIT Key Laboratory of Pattern Analysis and Machine Intelligence, China

^b Department of Biomedical Engineering, School of Medicine, Tsinghua University, China

^c Department of Orthopaedics, Beijing Tsinghua Changgung Hospital, China



ARTICLE INFO

Keywords:

Gait analysis
RGB-D camera
Hybrid model
Osteoarthritis

ABSTRACT

Osteoarthritis (OA) is the most common type of joint-related diseases, which affects millions of people worldwide. Expensive and time-consuming medical imaging can provide precise structural description of knee joints, but lacks the functional descriptions. Gait analysis can provide functional descriptions of knee joints. However, orthopedic surgeons always observe the patient's gait qualitatively and perform subjective assessments through rating scales at present due to the lack of a quantitative gait analysis system. To solve these problems, a gait acquisition and analysis system is developed to provide a cheap, easy-to-use solution for quantitative recording and functional description of OA patients. Firstly, an automatic gait acquisition platform is designed for the clinical setting based on the RGB-D camera and the developed software of gait data recording. In addition, the effective working space of gait acquisition platform is evaluated for clinical applications by comparing with the ground-truth from infrared optical trackers. Secondly, the acquired gait data is analyzed with a novel hybrid prediction model to assess the gait anomalies quantitatively and objectively. In the hybrid model, the extracted features of gait data contain the manually-extracted features and the automatically-extracted features from Long Short-Term Memory network. Experimental results on real patients demonstrate that the proposed gait analysis system can quantitatively predict gait anomalies with a high accuracy of 98.77 %. Therefore, this gait acquisition and analysis system achieves quantitative recording and objective assessment of gait anomalies for clinical OA treatments.

1. Introduction

Osteoarthritis (OA) is the most common type of orthopedic joint-related diseases in the elderly, and affects millions of people worldwide (Goekoop et al., 2011). The functional characters of bone joints commonly change for OA patients, because OA can damage joints and cause pain, stiffness of knee joints (Lonner et al., 2003). Gait pattern is useful to provide the functional description of knee joints and judge the status of joint-related diseases (Kaufman et al., 2001; Messier et al., 1992). However, the quantitative gait analysis framework hasn't been fully utilized due to the lack of cheap and easy-to-use systems. On the one hand, orthopedic surgeons always observe a patient's gait qualitatively and perform subjective assessments through rating scales. This

method is subjective. On the other hand, expensive and time-consuming medical imaging like X-ray and magnetic resonance images can provide the precise structural descriptions of knee joints, but lacks of the functional descriptions. Therefore, a cheap and easy-to-use system of gait acquisition and analysis is helpful not only for subjective and quantitative analysis of gait anomalies, but also for functional descriptions of joints that medical imaging cannot provide.

1.1. Related work

Over the last few years, several studies are proposed for gait data collections. Optical infrared tracker is used for the detailed motion capture of the gait (Chen et al., 2016a); however, complex operations of

* Corresponding author at: Department of Computer Science and Engineering, Nanjing University of Aeronautics and Astronautics, Nanjing, China.

** Corresponding author at: Department of Biomedical Engineering, School of Medicine, Tsinghua University, Beijing, China.

E-mail addresses: chenfang@nuaa.edu.cn (F. Chen), liao@tsinghua.edu.cn (H. Liao).

attaching infrared markers bring inconvenience in clinical usages. In addition, wearable sensor based gait analysis by using inertial sensors, pressure insole sensors, and electromyography (EMG) have been used as an alternative method to optical motion tracking systems (Tadano et al., 2016; Chen et al., 2016b). However, the challenge in using wearable sensor systems lies in the fact that sensors do not directly measure positional data of gait, but the data such as acceleration, angular velocity and magnetic field vector. Although a number of the custom force plates and the EMG systems are valuable in research studies of gait collections, there are obstacles from laboratory settings to real clinical applications (Sutherland, 2005). Moreover, video-based gait recognition has been a traditional way for human gait acquisitions and explorations. For example, silhouette information plays a fundamental role in video-based gait recognition approaches (Wang et al., 2003; Choudhury et al., 2012). Lee et al. proposed to extract features from silhouettes of human walking motion for the purposes of gait identification (Lee et al., 2014). Similarly, Lam et al. introduced an algorithm that utilizes spatio-temporal silhouette templates for gait recognition (Lam et al., 2007). Akae et al. proposed a temporal super-resolution method for gait recognition in quite low frame-rate videos (Akae et al., 2012). However, video-based approaches mainly use two-dimensional information and rely on high-quality videos and coherent silhouette extractions.

RGB-D cameras based on human stereo vision can provide an easy-to-use solution for 3D gait tracking and recording (Zhang et al., 2014). Kinect is one of the popular RGB-D cameras due to its low cost and high collection speed. Some studies have validated the Kinect system for gait tracking (Clark et al., 2013; Baldewijns et al., 2014). For example, (Clark et al., 2013) assessed the validity of the anatomical gait parameters such as step length, velocity etc. from the Kinect system by comparing with the gold standard from marker-based 3D motion analysis system. In addition, Baldewijns et al. (Baldewijns et al., 2014) measured the step length and step time using the Kinect, and the measured data showed excellent agreement with the ground-truth from GAITRite Electronic Walkway system. The gait data collection based on Kinect is robust (Geerse et al., 2015), and the current studies mainly focus on the neurological gait disorders for neuro-degenerative diseases.

Remarkably, gait analysis based OA recognition is a prominent area that requires developing automated systems that not only have reliability but are also affordable for real clinical applications (Kour et al., 2020; Iijima et al., 2019). About gait data collection for OA diagnosis, (Zeng et al. (2019)) used the marker-based motion analysis system to collect knee kinematic data, but this system needs attaching two rigid plates that composed of infrared light-reflecting markers. In addition, the ground reaction force (GRF) measurements were used for describing the knee gait patterns through sensors such as insole shoes (Mezghani et al., 2008; Pamukoff et al., 2019). Elbaz et al. performed the measurements of spatiotemporal gait parameters by a computerized walking mat (Elbaz et al., 2014). Furthermore, Higa et al. (Higa et al., 2015) compared the gait patterns of hip OA from the Kinect sensor and the ground-truth from force plate based gait-tracking system, but the comparison experiments were finished on common subjects instead of real patients. In all, easy-to-use and complete framework of OA gait measurement by using the cheap RGB-D camera has not been developed for real clinical applications (Ishikawa et al., 2017).

About analyzing the time sequence of gait data collected from RGB-D camera, the existing studies mainly implemented machine-learning methods to finish the disease prediction. For example, Li et al. used the covariance matrices and K-Nearest Neighbors (KNN) classifier to finish gait assessment, achieving the average accuracy about 80 % (Li et al., 2018). Jiang et al. manually set the static features of length and the dynamic features of joint angle, and use the nearest neighbor classifier to finish the gait assessment with the accuracy of 82 % (Jiang et al., 2014). In addition, Ahmed et al. combined the manually-defined dynamic features of horizontal and vertical distance with the KNN classifier to realize the gait assessment rate of 92 % (Ahmed, 2014). Dolatabadi et al. proposed to combine the KNN and dynamic time

warping (DTW) distance for pathological gait classification and the mean accuracy was about 90 % (Dolatabadi et al., 2017). Li et al. used three types of gait features and the KNN classifier for high predictive sensitivity of 94 % (Li et al., 2019). Without considering the usage habits of clinicians, these traditional analysis methods have not been used in real clinic applications.

Deep neural networks have become increasingly popular for analyzing the time sequence data, specifically, the long short-term memory (LSTM) architecture (Liu et al., 2018). Several LSTM recurrent neural network models have been proposed to solve the sequence labeling tasks like speech recognition and entity recognition, and achieved competitive performance against traditional machine learning methods (Sutskever et al., 2014). For example, Chung et al. used the LSTM recurrent neural network to finish the sequence data modeling, aiming at learning a probability distribution over sequences (Chung et al., 2014). Ma et al. proposed the bi-directional LSTM recurrent neural network to finish sequence-labeling tasks and obtain state-of-the-art performance with 97.55 % accuracy (Ma et al., 2016). In addition, Baccouche et al. trained a LSTM recurrent neural network to classify the time sequences of human actions by learning the spatio-temporal features of each sequence (Baccouche et al., 2011). Gao et al. proposed an algorithm based on LSTM and convolution neural network to extract the features automatically without manual interventions, and the highest accuracy was 93.1 % (Gao et al., 2019). Yin et al. also used the neural network to finish automatic gait classification on a database of 15 participants (Yin et al., 2019). Zhao et al. proposed to use a dual channel LSTM to extract gait features in order to diagnosis the neurodegenerative diseases, and the acquired accuracy was 95.67 % (Zhao et al., 2018). Inspired by the LSTM network, our paper implements the automatically optimized features from the LSTM network to improve the accuracy of gait analysis system.

1.2. Our contributions

This paper proposes a cheap and easy-to-use gait acquisition and analysis system for OA to finish quantitative recording and objective functional description of knee joints. The work described here expands on our previous study (Cui et al., 2019), and the contributions of the paper can be summarized as follows:

- 1) The gait acquisition platform is built for OA's clinical setting by using the RGB-D camera and the developed software of gait data recording. In addition, the effective working space of the gait acquisition platform is evaluated for clinical application by comparing with the ground-truth from infrared optical trackers.
- 2) We propose a novel hybrid prediction model to objectively assess the gait anomalies. Combining the LSTM network and traditional support vector machine classifier, the hybrid prediction model implements both the manually-extracted features and automatically-extracted features from LSTM network to achieve a high assessment accuracy.
- 3) Experimental results validate that, the gait analyzing system using RGB-D camera and the hybrid prediction model can quantitatively record joint's functional status for knee OA assessment. Gait function analysis is a good supplement to the joint's structure information from the medical imaging.

The remainder of this paper is organized as follows. In Section 2, we introduce the gait data acquisition platform (Sect. 2.1), the pipeline of hybrid prediction model (Sect. 2.2), the manually-defined features according to the clinical experience (Sect. 2.3), and the automatically-extracted features from the convolutional LSTM network (Sect. 2.4). The experimental setup and evaluation results are illustrated in detail in Sect. 3. We further discuss and analyze our study in Sect. 4. Conclusions are finally drawn in Sect. 5.

2. Methods

In this section, we describe the proposed gait acquisition and analysis system for OA patients. This system consists of two parts, including the gait data acquisition and pre-processing; and a novel hybrid prediction model for objectively assessing the gait anomalies. The hybrid prediction model combines the clinic-related features from manual extraction, and the automatically extracted features from deep neural network.

2.1. Gait data acquisition and pre-processing

The portable gait data acquisition platform mainly included a Kinect V2 depth camera (Microsoft Corp., Redmond, WA, United States) and a laptop that incorporated a graphics card (NVIDIA 1080 Ti) and an Intel multicore CPU (i7-4790 CPU), as shown in Fig. 1(a). About the location of depth camera, the Kinect v2 was placed horizontally at a 90-cm height and fixed on a tripod, and it was set approximately 100 cm away from the working range so as to capture the subjects' entire body in the cameras' fields of view. About the working range, we used the optical tracking device as the golden standard to determine this proper working range, as shown in Fig. 1(b). Through experiments, a proper working range of a rectangle with the width of 120 cm and length of 230 cm was determined in our study. In such a proper working range, each person was asked to walk back and forth, and position of the 25 main joints were obtained from the Kinect v2 at 30 Hz. The gait collection and recording software was developed based on the Kinect API with C++ programming toolkit. For OA patients, the gait pattern is reflected with joints, and the major 13 joints of a person were shown in Fig. 1(c).

After collecting the raw joint data, the pre-processing procedure is essential for normalizing and excluding the noisy skeletons. The raw joint data is a motion sequence in time $t \in [1 \dots T]$, and it consists of a set of skeletons, each of which is defined by 25 joints (the major 13 joints of a person are shown in Fig. 1(c)). We numerically denote 3D position of the i^{th} joint at time t , as $p_i^t = [x_i^t \ y_i^t \ z_i^t]$. The joint of SpineBase is chosen as the center, thus, the new position can be written as: $p_i^t - p_{\text{SpineBase}}^t$. In order to gain a scale invariant joint skeleton, we follow the observation that the Euclidean distance between SpineShoulder and SpineMid should be fixed when the people walks, and finish zooming the position of other joints accordingly, as shown in Fig. 2 (a). Then we put the center of the skeleton at the same location, which is the re-center process, as shown in Fig. 2 (b). Furthermore, a low-pass filter is used for noise reductions, because the noise has higher frequency comparing to human movement during acquisition. Finally, the linear interpolation is finished to ensure all the gait sequences have the same lengths (Fig. 2 (c)). The gait data after pre-processing is recorded by a time sequence

$$\hat{p}_i^t = [\hat{x}_i^t \ \hat{y}_i^t \ \hat{z}_i^t] \quad (i = 1 \dots 25; t = 1 \dots T).$$

2.2. Pipeline of the hybrid prediction model

After the gait data acquisition and pre-processing steps, the time-sequence of gait data is input into the hybrid prediction model for analyzing gait anomalies quantitatively. The pipeline of the hybrid prediction model is shown in Fig. 3. On the one hand, according to the clinical experience of orthopedic surgeons, manually-defined features of time-sequence gait data are extracted by using gait parameters and the dynamic time warping technique. On the other hand, the deep neural networks with convolution, LSTM and fully connected layers are implemented to automatically extract the contextual information of time-sequence gait data. Finally, the combined features are inputted into the support vector machine (SVM) classifier for assessing gait anomalies objectively (Xue et al., 2016; Niu and Suen, 2012).

2.3. Manually-defined features

According to the clinical experience of orthopedic surgeons, we manually define some clinic-related features to describe the knee joint's functions of OA patients. These gait features mainly include the static and dynamic features. In total, fourteen manually extracted gait features are calculated from the pre-processed gait data.

Static gait parameters are calculated based on the gait cycle, and a gait cycle includes the swing stage and stance stage, as shown in Fig. 4 (a). Stance stage means that two feet are both supported on the ground, and swing stage means that one leg is swinging and moving forwards with the other one supported on the ground. These two stages can be defined by the periodically changed distance between two ankles (David et al., 2012; Hunter et al., 2006). According to the clinical experience, static parameters such as the stride length (the moving distance of the same side foot during a gait cycle) and step length (the distance between two feet during a gait cycle) can provide the functional descriptions of the joint. Fig. 4 (b) shows the definitions of stride length and step length. Dynamic parameters include the DTW distance of different sequences of gait data. Through the DTW algorithm, different sequences of gait data are warped by using a nonlinear transformation in the time dimension to determine the similarities of different sequences. Therefore, DTW measures a distance-like quantity between two given sequences. DTW is firstly used in comparing two voice series (Rath et al., 2003). Here, we compute the DTW distances between k^{th} gait sample θ_k and the all healthy gait samples $\{\varphi_j\}_{j=1 \dots N_h}$ in the training dataset, and the mean value is the final DTW result D_k of gait sample θ_k , as follows:

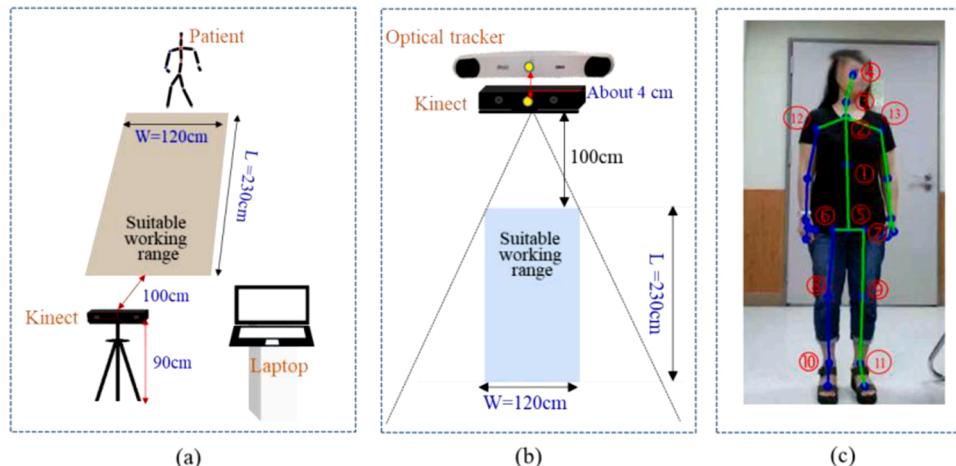


Fig. 1. (a) the portable acquisition system consisting of a Kinect and a laptop; (b) the proper rectangular working range, determined by the golden standard of optical tracker; (c) an example of the major 13 joints of a person collected by the RGB-D camera.

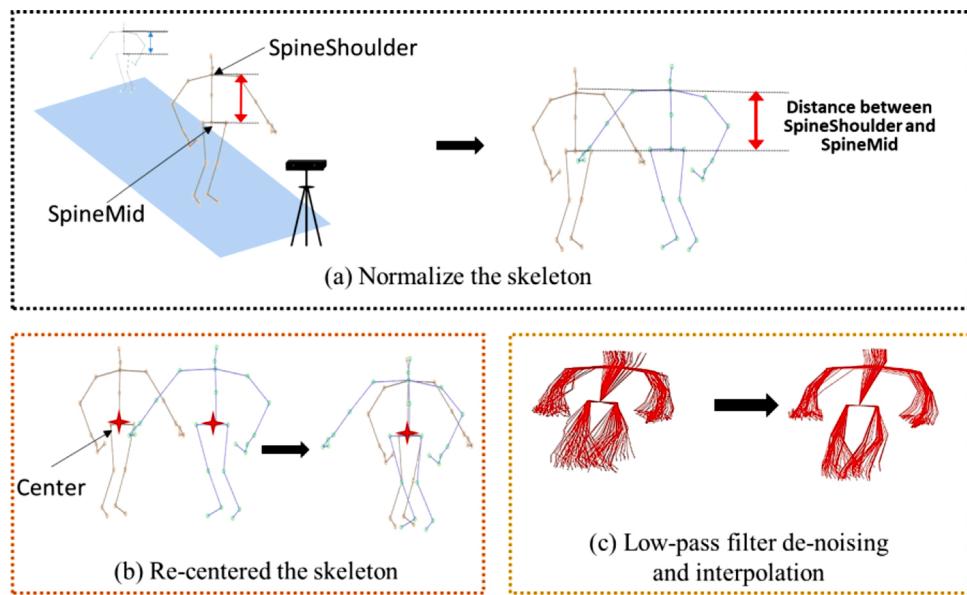


Fig. 2. The pre- processing procedure of the joints' data; (a) Normalize the skeleton; (b) Re-center the skeleton; (c) De-noising and interpolation.

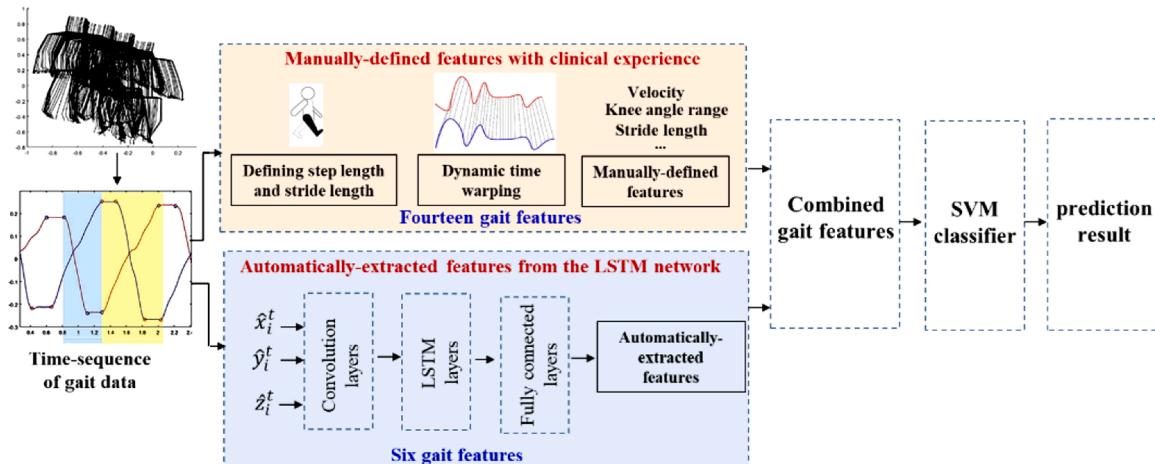


Fig. 3. The pipeline of hybrid prediction model.

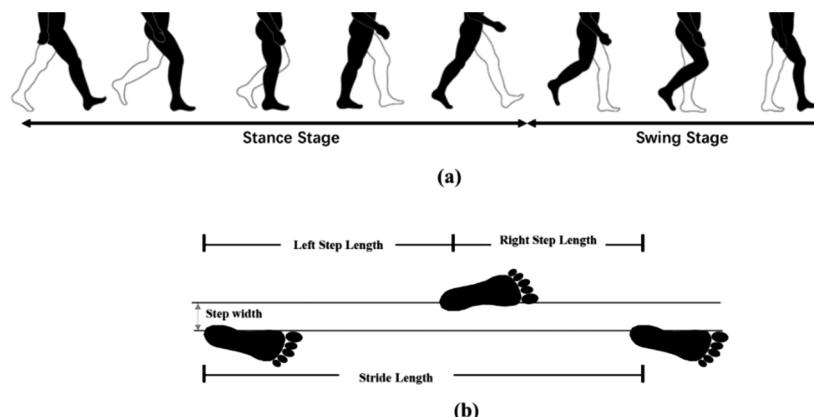


Fig. 4. (a) Stance stage and swing stage during a gait cycle; (b) the definition of gait parameters (step length and stride length).

$$D_k = \frac{1}{N_h} \left(\sum_{j=1}^{N_h} DTW(\theta_k, \varphi_j) \right)$$

(1) , where N_h is the subject number of the healthy group, and $DTW(\cdot)$ denotes the DTW distance.

In order to select the proper features for gait anomalies' assessment,

T-test result of each parameter is used as a rule. It represents the statistic differences between healthy people and OA patient with gait anomalies. The features with t-test's p-value less than 0.05 are chosen. All the t-test results are illustrated in Table 1, and the selected features are input into the classifier afterwards. Abbreviation of the manual features are also listed in Table 1.

2.4. Automatically-extracted features

We use deep neural network architecture (convolutional LSTM network) to automatically extract the contextual information of time-

We use 2 LSTM layers and each has 256 hidden units. LSTM part allows us to easily memorize the context information for long periods of time in sequence data. Finally, we pass the output of the LSTM to a few fully connected layers. As shown in (Krizhevsky et al., 2012), these higher layers can produce a higher-order feature representation. The higher-order feature is more separable for the task of gait data discriminating. The fully connected part in our network contains two fully connected layers of 60 and 6 nodes. The output of the final fully connected layer represents the automatically-extracted features.

<i>Algorithm 1: Gait processing and analysis method for osteoarthritis</i>
Input: Gait position sequences for training
$P_1 = \{[p_i^t]^1, [p_i^t]^2, \dots, [p_i^t]^{K1}\}_{i=1:25; t=1:T}$ (K1 is the number of training walking sequences)
Label: $L_1 = \{[l]^1, [l]^2, \dots, [l]^{K1}\}$
Gait position sequences for testing
$P_2 = \{[p_i^t]^1, [p_i^t]^2, \dots, [p_i^t]^{K2}\}_{i=1:25; t=1:T}$ (K2 is the number of testing walking sequences)
Output: The predict result of the testing data $\{[l']^1, [l']^2, \dots, [l']^{K2}\}$
(a) pre-processing of the gait position sequences
1: Normalize and re-centered the skeleton;
2: low-pass filter de-noising and interpolating;
(b) calculate the manually defined features for training and testing gait position sequences
1: calculate the manually defined features shown in Tab. 1.
(c) calculate the automatically defined features for training and testing gait position sequences
Training process to acquire the parameters ω_l of the convolutional LSTM network
1: setting the learning rate η , number of iterations n, batch size N
2: initialize ω_l
3: Repeat
for i=1 to n do
sample a mini-batch from the training data
$\omega_l \leftarrow \omega_l + \eta \text{ADAM}(\omega_l, \text{object function})$
end for
Until ω_l converges
Calculating the automatically defined features using the network with trained ω_l
1: Input the gait sequences to the convolutional LSTM network with parameters ω_l ;
2: Output the result of the final fully connected layer as the automatically defined features;
(d) using the manually and automatically defined features of the training data P_1 and the labels L_1 to train the SVM classifier
(e) Input the features of testing data into the trained SVM, and output the prediction result $\{[l']^1, [l']^2, \dots, [l']^{K2}\}$

Algorithm implementation

sequence gait data. An overview of our proposed network is shown in Fig. 5. The input of the network is the gait data after pre-processing $\hat{p}_i^t = [\hat{x}_i^t \ \hat{y}_i^t \ \hat{z}_i^t]$ ($i = 1 \dots 25; t = 1 \dots T$). The convolutional LSTM network is trained as an extractor of the automatic features, which will be combined with manually-defined features and inputted into a SVM classifier. SVM is used to replace the original sigmoid classifier because some researches have shown that the neural network + SVM model can obtain better performance than the neural network + sigmoid model (Xue et al., 2016; Niu and Suen, 2012).

The convolutional LSTM network of automatic feature extraction mainly contains the convolutional part, LSTM part and fully connected part (Fig. 5). First, the convolutional part can capture the information of the input gait data at different scales. The good convolutional operations have the capability to analyze the feature representations for capturing all intermediate scales. The gradual subsampling is used to achieve different scales. In our network, we use the pooling layers to perform gradual subsampling. Specifically, we use 4 convolutional layers, each with different feature maps of 64, 32, 16 and 8. A pooling size of 2 is used for the first three convolution layers, and no pooling is done in the final convolutional layer. Then, we pass the convolutional part output to LSTM part, which is appropriate for temporally modeling the gait data.

2.4.1. Algorithm implementation

In this task of gait analysis, the status of the input gait sequence are labelled into two classes (normal person and OA patients with gait anomalies), which serve as the ground-truth. For model training, we use the cross-entropy-loss function \mathcal{L} , which can compare the similarity between the ground-truth and the output of convolutional LSTM network. The cross-entropy loss is defined as follows:

$$\mathcal{L} = -\sum_{m=1}^{M=2} t^m \log o^m + (1 - t^m) \log(1 - o^m) \quad (2)$$

, where M denotes the number of output neurons, t^m is the target probability of the ground-truth, and the network output o^m represents an estimated posterior probability at which the status is classified as gait anomalies. The loss function is optimized by using the Adaptive Moment Estimation method (Adam) (Kingma et al., 2014). The proposed algorithm of gait processing and analysis is illustrated in Algorithm 1.

3. Experiments and results

Some evaluation experiments were conducted to validate our gait analysis system based on RGB-D camera and the hybrid prediction

Table 1

Manually defined features of the gait data.

Parameter (Unit)	Definition	Parameter abbreviation (Patient/ Healthy)	Correlation coefficient	p-value
Velocity(s^{-1})	Walking velocity	p-V/h-V	-0.9216	4.0784e-55
Step length(m)	Step length of both side	p-Lsl/h-Lsl(left) p-Rsl/h-Rsl(Right)	-0.3029 0.0652	0.0039 0.752
Stride length(m)	Length of a single stride	p-Str/h-Str	-0.9140	3.1970e-55
Step difference(m)	The difference between left and right step	p-Sdiff/h-Sdiff	0.4173	1.0361e-06
Step width(m)	Distance between the two ankles in double support phase	p-sw/h-sw	0.2230	0.0053
Knee angle range(deg)	Knee angle in a gait cycle on both sides	p-krr/h-krr(Left) p-krr/h-krr(Right)	0.2656 0.3407	0.2205 0.9788
DTW for knee angle(deg)	DTW distance for knee angle sequence of both sides	p-Dkl/h-Dkl(Left) p-Dkl/h-Dkl(Right)	0.1480 0.2314	0.1052 9.4845e-5
DTW for knee angle in yz plane/ xy plane(deg)	DTW value for abduction-adduction (x-y plane), extension- flexion angle(y-z plane)	p-Dkrz/h-Dkrz(Left) p-Dkrz/h-Dkrz(Right) p-Dkxz/h-Dkxz(Left) p-Dkxz/h-Dkxz(Right)	-0.1107 0.0908 -0.0104 0.1260	0.2269 0.0415 0.9097 0.1686

model. The first experiment validated the feasibility of the gait acquisition platform based on RGB-D camera. The effective working space of gait acquisition platform was evaluated for clinical applications by comparing with the ground-truth from infrared optical trackers. The second experiment finished the evaluations of manually-defined features. The third experiment evaluated the automatically extracted features. The fourth experiment validated the proposed hybrid prediction model with manually and automatically extracted features. Finally, the usage pipelines of the gait acquisition and analysis system in clinic was presented.

3.1. Validating the reliability of gait acquisition platform

3.1.1. System setup and working space evaluation

With the gait acquisition software and hardware, the joints' positions can be collected and saved. The correctness of the whole system relies on the RGB-D camera and a suitable working space. According to the user's operation manual of Kinect, the practical working range is a sector with

the maximum distance of 4.0 m and the minimum distance of 0.8 m. In our system, we reshaped the working range into a rectangle, because the rectangle range is convenient when patients walk alone a straight line. In order to determine the suitable working space, we compared the collected joint points' positions of the same person from Kinect system and the golden standard: optical tracking system. When the person walks into the improper space, the Kinect will collect the gait data inaccurately. Comparatively, the golden standard of optical tracking system has higher tracking accuracy and can collect the gait data correctly. Therefore, the collected positions from the Kinect and optical tracking system will have a big difference, and the position difference will have a sudden change. According to this rule, a proper working range could be determined in our study.

The optical tracking system can track the position of infrared markers, so in our experiments, we adhered the markers on the position of lower limb joints, as shown in Fig. 6 (a). We asked the experiment participant adhered with the infrared markers to move in a measuring range from -500 mm to +500 mm around the raw working boundaries

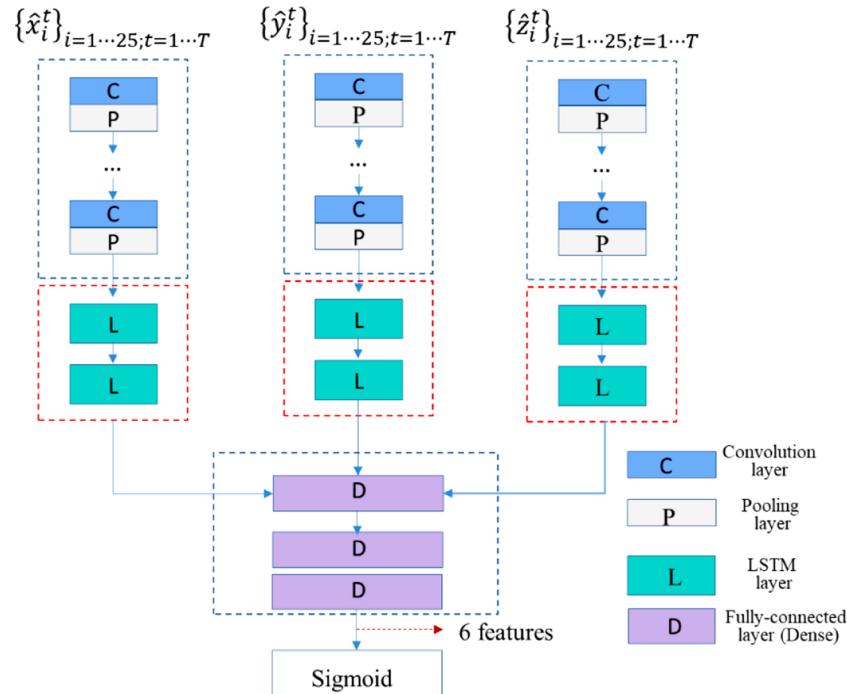


Fig. 5. The deep neural network architecture (convolutional LSTM network) for automatic feature extraction of the gait data.

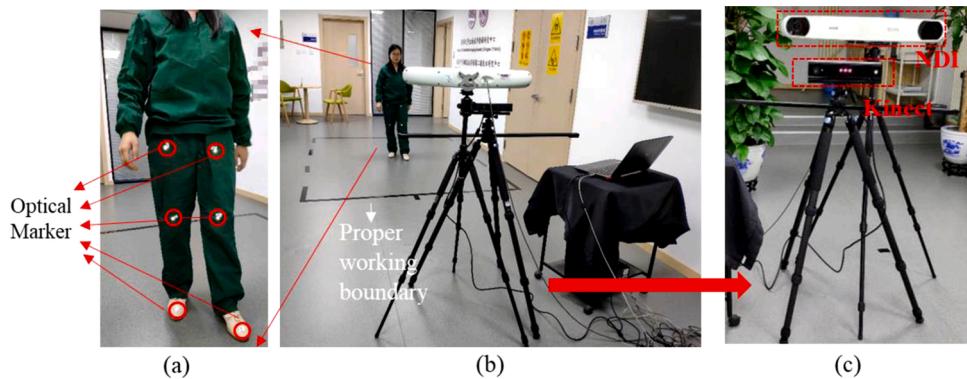


Fig. 6. (a) Adhering optical markers on the position of lower limb joints; (b) measuring joints' position with Kinect and optical tracker simultaneously; (c) enlarged view of the Kinect and optical tracking system.

([Fig. 6 \(b\)](#)). During the evaluation procedure, we measured the lower limb joints' position with Kinect and optical tracker simultaneously, and calculated the measured position differences between Kinect and optical tracking system ([Fig. 6 \(b\) \(c\)](#)).

A raw working range was a rectangle with width of 140 cm and length of 230 cm (shown in [Fig. 7 \(a\)](#)). [Fig. 7 \(b\)](#) showed an example of the measured position differences within the measuring range of $-500 \text{ mm} \sim +500 \text{ mm}$ around the raw working boundary, and we could see the sudden change of the position differences happened at the location of 100 mm location. Therefore, we adjusted the raw working boundary to the sudden point of the position differences ([Fig. 7\(a\)](#)). In this way, a proper working range of a rectangle with the width of 120 cm and length of 230 cm was determined.

3.1.2. Walking sequences' collection

This study was approved by the Institutional Review Board of Beijing Tsinghua Changgung Hospital. The gait data acquisition was finished in the clinic room of Beijing Tsinghua Changgung Hospital. Each person was asked to walk back and forth for several times to get multiple walking sequences ([Fig. 8 \(a\)](#)), and the collected walking sequences were used as the raw joint data. The schematic of the whole system and software interface could be seen in [Fig. 8 \(b\)](#).

In total, we collected 121 walking sequences, including 65 walking sequences of experiment group from 19 OA patients with gait anomalies, and 56 walking sequences of control group from 19 healthy volunteers. We could see that there were 121 walking sequences from 38 persons in total.

3.2. Evaluating the manually-extracted features

The manually-extracted features include the static and dynamic parameters. Some examples of the box plots for the manually-defined features of static and dynamic parameters between healthy persons and patients with gait anomalies were illustrated in [Fig. 9](#). Static parameters such as the gait velocity and stride length had significant differences between healthy persons and patients. Orthopedic surgeons found that OA patients with gait anomalies walked slower than normal people with shorter stride length ([Kaufman et al., 2001](#)). Therefore, the conclusion of static parameter was agreement with the clinical finding of patients with gait anomalies. In addition, the dynamic parameters like DTW values of the right knee angle (Dkr: DTW value for knee angle; and Dkrz: DTW value for abduction-adduction (x-y plane)) also showed separability between patients and healthy persons.

3.3. Evaluating the convolutional LSTM network for automatically extracted features

The convolutional LSTM network was trained and tested with keras (it was a high-level neural networks API, written in Python and capable of running on top of TensorFlow, CNTK, or Theano) on two NVIDIA GTX 1080Ti. We used a batch size of 10, a learning rate of 0.0001 and epochs of 80. To validate the effectiveness of the automatically extracted features from the convolutional LSTM network, the *t*-test's p-value and Pearson correlation coefficient of each extracted feature was illustrated in [Table 2](#). The results confirmed that the automatically extracted features had obvious relevance and significances.

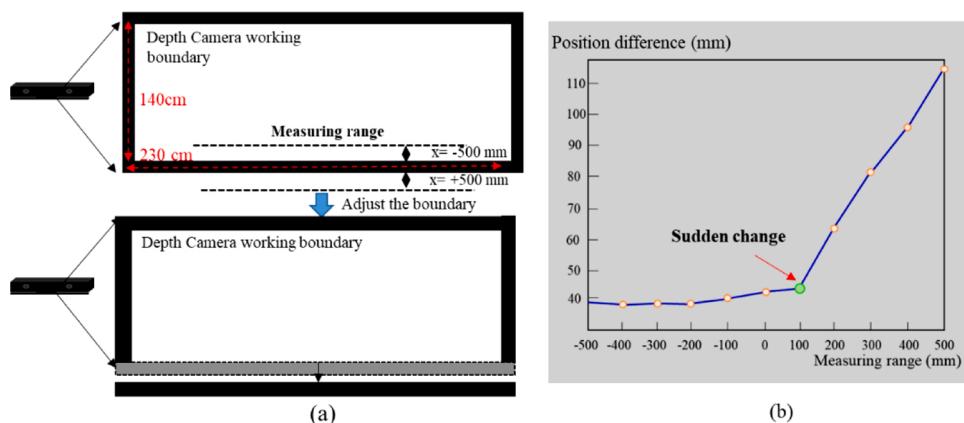


Fig. 7. (a) The original working range and the working range is adjusted by evaluating the measuring error around the raw working boundary; (b) an example of the position differences within the measuring range.

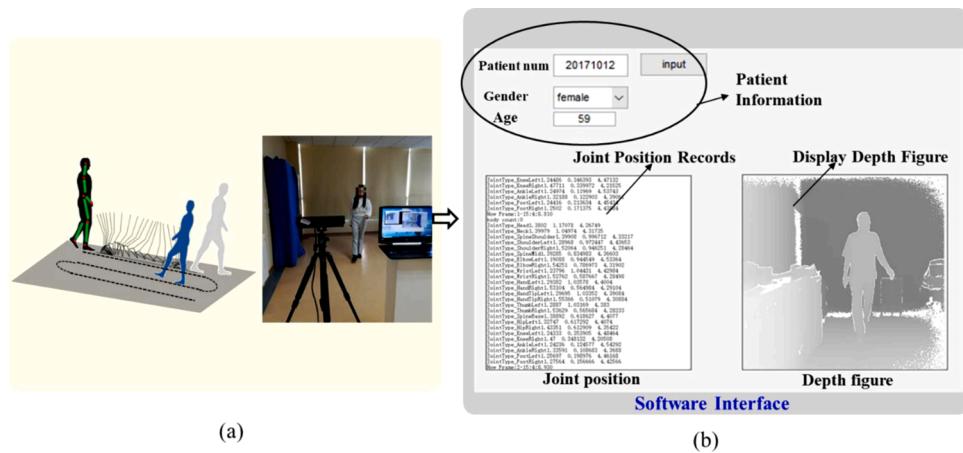


Fig. 8. (a) The person is asked to walk back and forth several times, and the middle parts of the whole walking procedure is collected; (b) Software interface of the data acquisition platform.

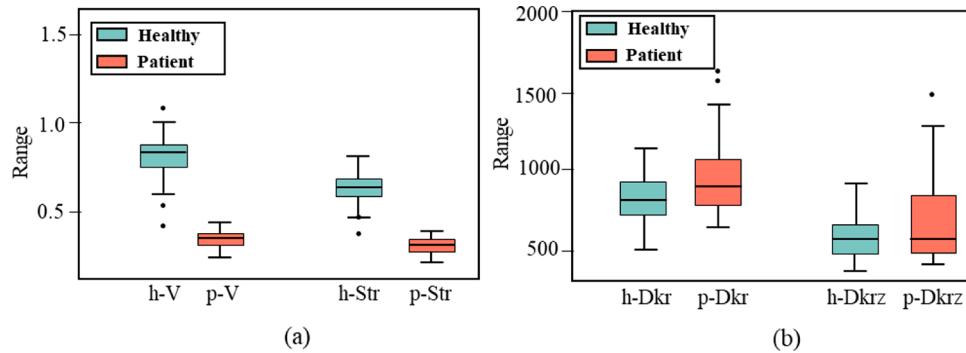


Fig. 9. Some examples of the box plots of the manually-defined features of static and dynamic parameters between healthy persons and patients (Parameter abbreviation was shown in Table 1); (a) Box plots of static parameters; (b) Box plots of dynamic parameters.

Table 2

T-test's P-value and Pearson correlation coefficient of the automatically extracted features.

Features	correlation coefficient	p-value
Feature1	0.87	6.47e-39
Feature2	-0.94	9.70e-59
Feature3	0.89	1.52e-42
Feature4	0.84	1.20e-33
Feature5	0.87	3.32e-39
Feature6	0.87	4.87e-39

Table 3

Average prediction results of 5-fold cross validation by using different features of gait data.

Tasks	Features	Prediction accuracy (%)
Assessing the gait anomalies	Manually extracted features	97.03 ± 2.75
	Automatically extracted features	97.35 ± 1.65
	Hybrid features	98.77 ± 1.01

3.4. Comparison of hybrid features with single extracted features

This proposed hybrid prediction model combined the manually and automatically extracted features of the gait data from RGB-D camera. Therefore, we compared the results of analyzing the gait anomalies by using the SVM classifier on three feature sets including: 1) Manually

Table 4

the used time for each step of the gait analysis method.

Gait data transmission	Pre-processing	Calculation of the manually-defined (clinic-related) features	Calculation of the features from LSTM	Model prediction	Total time
0.04 s	0.87 s	2.43 s	0.61 s	0.96 s	4.91 s

extracted features; 2) Automatically extracted features; and 3) Hybrid features. During the experiment, we adopted the 5-fold cross validation to evaluate the gait prediction results for 121 walking sequences from 38 persons. Please note that, all sequences from the same person should be either in the training or the testing fold, and never on both. Table 3 represented the average 5-fold cross validation results by using different feature sets.

On the one hand, the average prediction accuracy of the traditional manually-extracted-feature based SVM method was 97.03 % whereas that of the hybrid-feature based method was 98.77 %, demonstrating the effectiveness of hybrid features. We further calculated the statistically significance of the accuracy improvement between these two methods and the results demonstrated the improvement of classification accuracy was statistically significant with a p-value of 0.039. On the other hand, the average prediction accuracy of the hybrid-feature based method outperformed the single automatically-extracted-feature based method by at least 1.42 % (98.77 % Versus 97.35 %), as shown in Table 3. In

Table 5

Average prediction results of 5-fold cross validation by using different classifiers.

Task	Used classifier	Features	Prediction accuracy (%)
Assessing the gait anomalies	SVM	Manually extracted features	97.03 ± 2.75
		Automatically extracted features	97.35 ± 1.65
		Hybrid features	98.77 ± 1.01
	KNN	Manually extracted features	96.42 ± 3.41
		Automatically extracted features	96.75 ± 2.42
		Hybrid features	97.37 ± 2.16
Random Forest	KNN	Manually extracted features	96.78 ± 3.02
		Automatically extracted features	97.14 ± 2.13
	Random Forest	Hybrid features	98.03 ± 1.97

addition, paired *t*-test showed that this accuracy improvement was statistically significant with a *p*-value of 0.028. Therefore, the proposed hybrid features achieved better performance with statistically significance.

There are two common rules for combining features. The first rule is the forward combination rule that means different types of features are concatenated and then input into a SVM classifier to finish prediction. The second rule is the backward combination rule that means inputting the different types of features into two classifiers, and then averaging the results to acquire the final prediction. We compared the prediction accuracy by using these two rules, respectively. The experimental results validated that using the forward combination rule could acquire higher accuracy than using backward combination rule (98.77 ± 1.01 % Versus 98.43 ± 1.24 %). Therefore, the forward combination rule was chosen to combine the manually-defined features and automatically extracted feature.

As shown in Table 4, we illustrated the used time for each step of the gait analysis method, and the total time was about 4.91 s to analyze one person. A user speeded specifically about 2.43 s and 0.61 s for calculating the manually-defined (clinic-related) and automatically-extracted features respectively.

3.5. Comparison of using different classifiers

As shown in Table 5, we represented the prediction results by using different classifiers including SVM, KNN and random forest. According to the experimental results, we can found two conclusions. Firstly, no matter which classifier is used, the proposed hybrid features achieved better performance than using the single type of feature (SVM: 97.35 % Versus. 98.77 %; KNN: 96.75 % Versus 97.37 %; Random forest: 97.14 %

Versus 98.03 %). Therefore, hybrid features were significant for recognizing the gait anomalies. Secondly, using SVM classifier could acquire the best prediction results compared with using KNN and random forest (SVM: 98.77 %; KNN: 97.37 %; Random forest: 98.03 %).

3.6. Real clinical use of the gait analysis system

To confirm that the proposed system is clinically practical, the usage pipeline of the gait acquisition and analysis system in clinic was presented by the collaborative medical team, as follows:

- 1) Ask the person to enter the room of orthopedic clinic;
- 2) In the determined working range, the gait acquisition platform collects the gait data of the person during walking;
- 3) Analyze the function character of the gait data by using the gait processing and analysis algorithm, and judge whether this patient has OA with gait function anomalies;
- 4) The patients suspected to have gait function anomalies are supposed to finish the further structural diagnosis by using medical imaging. For these suspected patients, structural information from medical image and the functional characters from gait analyzing are used together to describe the knee status (Sowers et al., 2003).

In our experiment of the real clinical use of the gait analysis system, totally five people were finished the gait analysis, and two people suspected to have gait function anomalies. Then, these two people further finished the medical imaging for structural diagnosis (see Fig. 10). Therefore, through the above gait-analyzing step, the orthopedist could directly finish function pre-diagnosis in clinic rooms quantitatively, and only the people suspected to have gait function anomalies, instead of all the people were need to perform further medical imaging.

4. Discussion

The quantitative description and recording of patient's gait are very important for the joint-related disease, such as OA. In this paper, we propose a gait acquisition and analysis system for objective functional descriptions of OA patients. Traditional clinical subjective assessment observes patient's gait qualitatively and documents assessments through rating scales. Comparatively, our gait analysis system can provide quantitative recording and objective functional descriptions of OA patients. In addition, the knee's function characters from the gait analysis system can be a good supplement to the knee's structure information from the medical imaging, and can avoid some unnecessary and expensive medical imaging in some cases.

The proposed gait analysis system finishes the hardware and software platforms. The hardware platform contains a RGB-D camera and a computer with the developed joint data acquisition interface. This

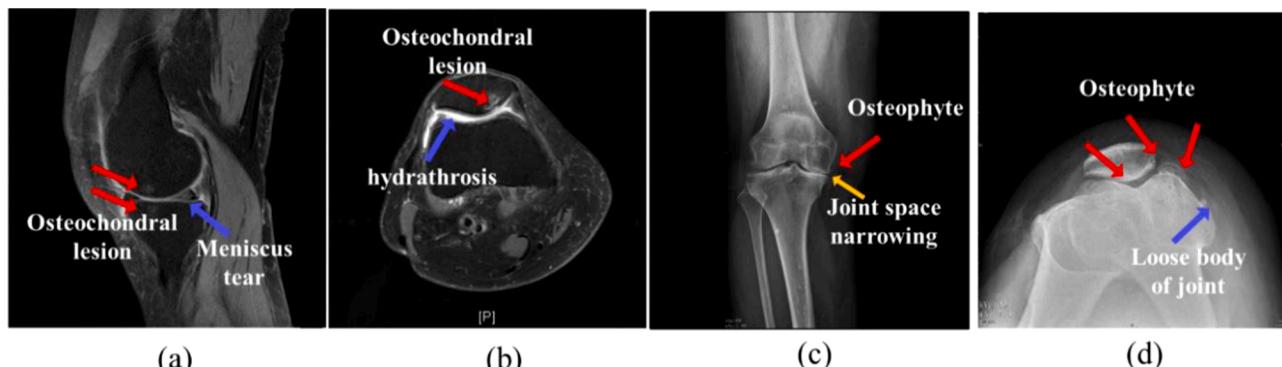


Fig. 10. (a) (b) structural information from magnetic resonance images for patient one; (c) (d) structural information from X-ray images for patient two; Different arrows showed structural signs of osteoarthritis.

hardware platform is portable and convenient for clinical usages. The software platform mainly contains the novel hybrid prediction model for subjectively assessing the gait anomalies. Our experiments on real patients have validated the key aspects of the hybrid prediction model. We first confirm that the manually-defined features of gait data can offer useful information for analyzing gait data. Further, the results of the additional automatically-extracted features show that convolutional LSTM network is effective for descriptions of gait data. The hybrid model that combines manual and automatic features achieves a clear increase in OA prediction accuracy compared with the model based on single feature set. Finally, the usage pipeline of the gait acquisition and analysis system in clinic is presented.

4.1. Contribution and novelty of the proposed system and method

4.1.1. The contribution of the proposed method

In the field of gait analysis system for OA recognition, this is the first study that tries the ideas of combining the describable features from clinicians' experiences and indescribable implicit features from the deep networks. About the proposed gait analysis method, firstly considering that clinicians have some experience in gait discriminations, we translated the clinician's experience into quantifiable gait features such as step difference, knee angle range, etc., and these features were further analyzed and selected to acquire valuable clinical parameters. Secondly, considering that some deep features are implicit in data and are difficult to describe artificially, we implemented the convolutional LSTM network to find implicit features automatically. Therefore, our method used not only the describable features from clinicians' experiences, but also the implicit and indescribable features from convolutional LSTM network. Experimental results also validated that combining the describable features from clinical prior and implicit features from deep model could achieve higher accuracy than using one type of feature (98.77 % Versus 97.03 % and 98.77 % Versus 97.35 %). Most of the existing studies just used a single type of feature (Mezghani et al., 2008; Ishikawa et al., 2017; Gao et al., 2019; Yin et al., 2019). Therefore, our analysis framework can provide some reference ideas for the researches in this field.

4.1.2. The contribution of the proposed system

The final goal of this study is developing a clinically available gait collection and analysis system for OA. In order to ensure that the system can meet the usage habits of orthopedic clinicians and be more easily accepted by clinicians, our system mainly finished the following aspects of design pertinently. Firstly, considering the portability and low cost requirements in clinical use, a portable gait data collection platform was set up by using the inexpensive Kinect depth camera and a common laptop. This platform was portable, inexpensive and suitable for different environments, such a narrow clinic room, primary medical units, and so on. Thus, it has the potential of wide applications. Secondly, considering the narrow and square area of the clinic room, we defined and evaluated the perfect working range of a rectangle for gait collections, and the rectangle range was convenient when patients walk alone a straight line. Thirdly, considering that using clinical prior information for gait recognition is easier to be accepted by the doctors, we combined the clinical features accumulated by doctors' clinical experience and the automatic implicit features to finish gait recognition in this study.

In all, because of these special technologies from the perspective of clinical applications, our system was finally applied in real clinical research. In the field of gait based OA analysis, most of the existing studies (Higa et al., 2015; Liu et al., 2018; Gao et al., 2019) only focus on the design of complex gait analysis methods, while ignoring the overall system construction and clinical evaluation. Different from the existing studies, our system is applied in real clinic and has the potential of wide applications.

4.2. Significance of the gait recognition for OA

OA can damage joints. Painful joints undoubtedly cause people lose the ability to walk normally and have gait abnormality. Thus, gait recognition can be used to judge whether the joints have the function of normal walking. This functional descriptions of joints based on gait recognition is of great significance for OA diagnosis and treatment. On the one hand, medical imaging like X-ray and magnetic resonance image can provide the structural description of knee joints for OA, but lacks of the functional descriptions. Gait recognition can achieve functional descriptions of joints that medical imaging cannot provide. On the other hand, the developed gait recognition system is cheap and easy-to-use in a narrow clinic environment. Thus, it can be easy to use in primary medical units to help the diagnosis and rehabilitation analysis for a larger amount of OA patients. In summary, gait recognition provides an important and significant tool for OA diagnosis and treatment.

We also acknowledge several limitations of this work. First, the proposed gait analysis system of OA is created based on the data from a single institution; therefore, its generalizability needs to be further validated with the data from multi-institutions. Second, a future study will focus on more experiments of clinical cases to prove the practical importance of the system in clinical treatment. Third, considering that the Kinect sensor has been discontinued by Microsoft, we will replace the Kinect by another RGB-D camera to acquire the same gait collection and analysis system in the future.

5. Conclusion

In this paper, we present a cheap and easy-to-use gait acquisition and analysis system for OA by using the RGB-D camera and the hybrid prediction model. Our approach utilizes two type of features including the manually-defined features from clinical experience and automatically extracted features from deep neural network to improve the prediction accuracy. The experimental results further demonstrate that our system achieves objective gait assessment with a high accuracy of 98.77 %. Our system can describe the knee's function and quantitatively record joint status. In addition, it can be a good supplement to the joint's structure information from the medical imaging.

CRediT authorship contribution statement

Fang Chen: Conceptualization, Methodology, Writing - original draft. **Xiwen Cui:** Visualization. **Zhe Zhao:** Software, Validation. **Daqiang Zhang:** Conceptualization. **Cong Ma:** Data curation. **Xinran Zhang:** Project administration. **Hongen Liao:** Project administration.

Declaration of Competing Interest

The authors report no declarations of interest.

Acknowledgements

Fang Chen, Xiwen Cui and Zhe Zhao are contributed equally to this work. The authors acknowledge supports from National Nature Science Foundation of China grants (61901214, 81771940), China Postdoctoral Science Foundation (2020M671484), High-level Innovation and Entrepreneurship Talents Introduction Program of Jiangsu Province of China, National Key Research and Development Program of China (2017YFC0108000, 2018YFC2001600, 2018YFC2001602).

References

- Ahmed, M., 2014. Gait recognition based on Kinect sensor. *Real-Time Image and Video Processing*, pp. 32–39.
- Akae, N., et al., 2012. Video from nearly still: an application to low frame-rate gait recognition". *IEEE Computer Vision and Pattern Recognition* 1, 1537–1543.

- Baccouche, M., et al., 2011. Sequential deep learning for human action recognition. In: International Workshop on Human Behavior Understanding, Berlin, Heidelberg, 4, pp. 29–39.
- Baldewijns, G., et al., 2014. Validation of the kinect for gait analysis using the GAITRite walkway, in IEEE. Conf. Proc. IEEE Eng. Med. Biol. Soc. 2, 5920–5923.
- Chen, S., et al., 2016a. Toward pervasive gait analysis with wearable sensors: a systematic review. *IEEE J. Bio. Health Inform.* 20, 1521–1537.
- Chen, S., Lach, J., Lo, B., Yang, G.Z., 2016b. Toward pervasive gait analysis with wearable sensors: a systematic review. *IEEE J. Biomed. Health Inform.* 20 (6), 1521–1537.
- Choudhury, S.D., et al., 2012. Silhouette-based gait recognition using Procrustes shape analysis and elliptic Fourier descriptors. *Pattern Recognit.* 45, 3414–3426.
- Chung, J., et al., 2014. Empirical evaluation of gated recurrent neural networks on sequence modeling. arXiv preprint 5 arXiv:1412.3555.
- Clark, R.A., et al., 2013. Concurrent validity of the Microsoft Kinect for assessment of spatiotemporal gait variables. *J. Biomech.* 46, 2722–2725.
- Cui, X., Zhao, Z., Ma, C., et al., 2019. A Gait Character Analyzing System for Osteoarthritis Pre-diagnosis Using RGB-D Camera and Supervised Classifier. *World Congress on Medical Physics and Biomedical Engineering* 13, 297–301.
- David, L., et al., 2012. Whittle's Gait Analysis, 5th ed., Volume 10. Elsevier LTD., Oxford, UK, pp. 32–37.
- Dolatabadi, E., Taati, B., Mihailidis, A., 2017. An automated classification of pathological gait using unobtrusive sensing technology. *Ieee Trans. Neural Syst. Rehabil. Eng.* 25, 2336–2346.
- Elbaz, A., et al., 2014. Novel classification of knee osteoarthritis severity based on spatiotemporal gait analysis. *Osteoarthr. Cartil.* 22, 457–463.
- Gao, J., Gu, P., Ren, Q., et al., 2019. Abnormal gait recognition algorithm based on LSTM-CNN fusion network. *IEEE Access* 7, 163180–163190.
- Geerse, D.J., et al., 2015. Kinematic validation of a multi-kinect v2 instrumented 10-Meter walkway for quantitative gait assessments. *PLoS One* 10, e0139913.
- Goekoop, R.J., et al., 2011. Determinants of absence of osteoarthritis in old age. *Scand. J. Rheumatol.* 40, 68–73.
- Higa, M., et al., 2015. Impact of gait modifications on hip joint loads during level walking. In: *World Congress on Medical Physics and Biomedical Engineering*, 12, pp. 346–349.
- Hunter, D.J., et al., 2006. Increases in bone marrow lesions associated with cartilage loss: a longitudinal magnetic resonance imaging study of knee osteoarthritis. *Arthritis & Rheumatism: Off. J. Am. Coll. Rheumatol.* 54, 1529–1535.
- Iijima, H., Shimoura, K., Ono, T., et al., 2019. Proximal gait adaptations in individuals with knee osteoarthritis: a systematic review and meta-analysis. *J. Biomech.* 87, 127–141.
- Ishikawa, Y., et al., 2017. Gait analysis of patients with knee osteoarthritis by using elevation angle: confirmation of the planar law and analysis of angular difference in the approximate plane. *Adv. Robot.* 31, 68–79.
- Jiang, S., et al., 2014. Real time gait recognition system based on kinect skeleton feature. *Asian Conference on Computer Vision* 21, 46–57.
- Kaufman, K.R., et al., 2001. Gait characteristics of patients with knee osteoarthritis. *J. Biomech.* 34, 907–915.
- Kingma, D.P., et al., 2014. Adam: a method for stochastic optimization. arXiv preprint arXiv:1412.6980.
- Kour, N., Gupta, S., Arora, S., 2020. A survey of knee osteoarthritis assessment based on gait. *Arch. Comput. Methods Eng.* 11, 1–41.
- Krizhevsky, A., et al., 2012. Imagenet classification with deep convolutional neural networks. *Adv. Neural Inf. Process. Syst.* 21, 1097–1105.
- Lam, T.W., et al., 2007. Human gait recognition by the fusion of motion and static spatio-temporal templates. *Pattern Recognit.* 40, 2563–2576.
- Lee, T.M., et al., 2014. A comprehensive review of past and present vision-based techniques for gait recognition. *Multimed. Tools Appl.* 72, 2833–2869.
- Li, Q., et al., 2018. Classification of gait anomalies from Kinect. *Vis. Comput.* 34, 229–241.
- Li, M., Tian, S., Sun, L., et al., 2019. Gait analysis for post-stroke hemiparetic patient by multi-features fusion method. *Sensors* 19, 1737–1746.
- Liu, J., et al., 2018. Skeleton-based action recognition using spatio-temporal LSTM network with trust gates. *IEEE Trans. Pattern Anal. Mach. Intell.* 40, 3007–3021.
- Lonner, J.H., et al., 2003. A 57-year-old man with osteoarthritis of the knee. *Jama* 298, 1016–1025.
- Ma, X., et al., 2016. End-to-end sequence labeling via bi-directional lstm-cnns-crf. arXiv preprint arXiv:1603.01354.
- Messier, S.P., et al., 1992. Osteoarthritis of the knee: effects on gait, strength, and flexibility. *Arch. Phys. Med. Rehabil.* 73, 29–36.
- Mezghani, N., Husse, S., Boivin, K., et al., 2008. Automatic classification of asymptomatic and osteoarthritis knee gait patterns using kinematic data features and the nearest neighbor classifier. *IEEE Trans. Biomed. Eng.* 55, 1230–1232.
- Niu, X.X., Suen, C.Y., 2012. A novel hybrid CNN-SVM classifier for recognizing handwritten digits. *Pattern Recognit.* 45, 1318–1325.
- Pamukoff, D.N., Garcia, S.A., Holmes, S.C., et al., 2019. Association between ground reaction force characteristics during gait and knee injury and osteoarthritis outcome scores in young adults with obesity. *Osteoarthr. Cartil.* 27, S133–S134.
- Rath, T.M., et al., 2003. Word image matching using dynamic time warping, in IEEE Computer Vision and Pattern Recognition. Madison 23, 13–21.
- Sowers, M.F., et al., 2003. Magnetic resonance-detected subchondral bone marrow and cartilage defect characteristics associated with pain and X-ray-defined knee osteoarthritis. *Osteoarthr. Cartil.* 11, 387–393.
- Sutherland, D., 2005. The evolution of clinical gait analysis part III—kinetics and energy assessment. *Gait Posture* 21, 447–461.
- Sutskever, I., et al., 2014. Sequence to sequence learning with neural networks. *Montreal International Conference on Neural Information Processing Systems* 14, 3104–3112.
- Tadano, S., Takeda, R., et al., 2016. Gait characterization for osteoarthritis patients using wearable gait sensors (H-Gait systems). *J. Biomech.* 49, 684–690.
- Wang, L., et al., 2003. Silhouette analysis-based gait recognition for human identification. *IEEE Trans. Pattern Anal. Mach. Intell.* 25, 1505–1518.
- Xue, D.X., Zhang, R., Feng, H., Wang, Y.-L., 2016. CNN-SVM for microvascular morphological type recognition with data augmentation. *J. Med. Biol. Eng.* 36, 755–764.
- Yin, S., Chen, C., Zhu, H., et al., 2019. Neural networks for pathological gait classification using wearable motion Sensors//2019. In: *IEEE Biomedical Circuits and Systems Conference (BioCAS)*. IEEE, pp. 1–4.
- Zeng, W., Ma, L., Yuan, C., et al., 2019. Classification of asymptomatic and osteoarthritic knee gait patterns using gait analysis via deterministic learning. *Artif. Intell. Rev.* 52, 449–467.
- Zhang, Q., et al., 2014. Quality dynamic human body modeling using a single low-cost depth camera IEEE. Conference on Computer Vision and Pattern Recognition 2, 676–683.
- Zhao, A., Qi, L., Dong, J., et al., 2018. Dual channel LSTM based multi-feature extraction in gait for diagnosis of Neurodegenerative diseases. *Knowledge Based Syst.* 145, 91–97.