

# Dual channel LSTM based multi-feature extraction in gait for diagnosis of Neurodegenerative diseases

Aite Zhao<sup>a</sup>, Lin Qi<sup>a</sup>, Junyu Dong<sup>a,\*</sup>, Hui Yu<sup>b</sup>

<sup>a</sup> Department of Computer Science and Technology, Ocean University of China, Qingdao, China

<sup>b</sup> University of Portsmouth, Portsmouth, UK

## ARTICLE INFO

### Article history:

Received 9 August 2017

Revised 29 December 2017

Accepted 2 January 2018

Available online 6 January 2018

### Keywords:

Neurodegenerative diseases

Diagnosis

Gait disorders

Time series

Force series

Dual channel LSTM

## ABSTRACT

The performance of gait disturbances differ in various Neurodegenerative diseases (NDs), which is an important basis for the diagnosis of NDs. In the diagnosis, doctors can judge disease state by observing patients' gait features without quantification, such a subjective diagnosis has been seen as a problem because diagnostic results may differ among doctors. Moreover, there are some irresistible factors such as fatigue may effects diagnostic procedure. To make use of these observations, we build an automatic deep model based on Long Short-Term Memory (LSTM) for the gait recognition problem. In our model, a dual channel LSTM model is designed to combine time series and force series recorded from NDs patients for whole gait understanding. Experimental results demonstrate that our proposed model improves gait recognition performance compared to baseline methods. We believe the quantitative evaluation provided by our method will assist clinical diagnosis of Neurodegenerative diseases.

© 2018 Elsevier B.V. All rights reserved.

## 1. Introduction

Neurodegenerative diseases (NDs), including Parkinson's disease (PD), Huntingtons disease (HD), and amyotrophic lateral sclerosis (ALS), produce lesions in the central nervous system that controls the motions of two lower limbs causing gait disorders. Studies assert that every individual has unique gait pattern [1] indicating that gait disorder is recognized as a contributing diagnostic criterion for NDs.

Additionally, gait performance varies among the three diseases. HD is an dominant genetic disease with nervous system gradually degenerated resulting in abnormal body movements, which may present as involuntary movements, walking and balance disorders, dance like movements, twisting, rolling, and unstable gait in middle stage [2]. Gait disturbance occurs in PD patients' early stage, such as festinating gait, short gait and freezing gait may make PD diagnosis easier [3]. The clinical features of ALS are indicative of the loss of neurons at all levels of the motor system - from the cortex to the anterior horn of the spinal cord. Physical signs of this disorder thus encompass both upper motor neuron and lower motor neuron findings; this causes the movement disorder of the limbs, for example, scissors gait and spastic gait [4].

By employing these motor symptoms, we are committed to optimizing the existing disease diagnosis system which is one of the more labor-intensive tasks in medical procedure. Despite advances in medical care, gait disturbances are known to worsen as the disease progresses, contributing to loss of independence, falls, and poor quality of life. Moreover, gait disorder based diagnosis can not provide quantitative data of diseases leading to a subjective and inefficient diagnosis process, besides, general health care systems seem not always maintain accurate and rapid diagnosis. Machine learning methods have gained popularity as they offer an objective approach to identifying or differentiating subgroups of individuals with movement disorders and quantifying outcomes of gait classification in low cost. In many cases, these methods can provide more accurate diagnosis than experienced nerve physicians so that it facilitates auxiliary diagnosis.

Recently, machine learning technologies have been effectively applied to study the gait variability in neurological diseases including the kernel Fisher discriminant (KFD), the naive Bayesian approach (NB), support vector machine (SVM) and nearest neighbor (NN) [5–8], neural networks [9,10]. However, these researches are not specifically designed to cope with temporal data, whereas the gait recorded by devices (cameras, force sensors) contains important temporal information which is significant for NDs diagnosis.

Long Short-Term Memory (LSTM) has done a good job on this issue in various fields including action recognition and gait recognition owing to the ability of processing and forecasting the time series with very long interval and delay [11–17]. Although these

\* Corresponding author.

E-mail address: [dongjunyu@ouc.edu.cn](mailto:dongjunyu@ouc.edu.cn) (J. Dong).

works can effectively extract the single gait feature, it is insufficient for representing the details of gait changes. We propose a dual channel LSTM model to fuse two gait features including time series and force series to learn gait pattern of the patient. By experimenting on a public NDs gait data set, our method is testified to outperform other popular algorithms in solving this problem.

In addition, the two types of gait data were measured by a foot-switch system simultaneously that provides accurate estimates of the start and end of stance phase for sequential steps based on a commercially available transducer and can be readily reproduced for use in a laboratory setting. Specifically, it contained two  $1.5 \text{ in}^2$  force sensitive resistors and a  $390\Omega$  measuring resistor, which can obtain the stride time intervals according to the variation tendency of force changing in gait. In order to comprehensively capture the gait changes in time and space, we combine these two features to obtain better diagnostic results [18].

The rest of this paper is organized as follows. Section 2 explores the related work. Section 3 introduces the proposed dual channel LSTM network. Section 4 describes the used public data set and provides our experimental results, and Section 5 is a discussion for NDs diagnosis. Section 6 concludes the paper.

## 2. Related work

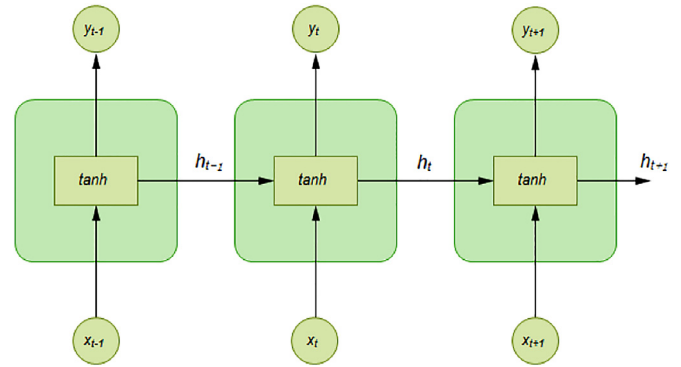
It is generally known that the medical institutions produce a large number of data, different features are extracted from these data, which can help doctors and patients understand the state of illness and select the proper treatment. This section introduces the related work for NDs diagnosis by using data with different features, such as skeleton feature, time feature, and force feature.

Skeleton feature is an influential reference of gait recognition, which eliminates some interference characteristics and retains only the joint point coordinate of human bones to identify gait difference. Torres et al. presented an approach of recording the posture of PD with a Kinect sensor, which can capture the changes of joint points in order to assist physicians in PD diagnosis [19]. Galna et al. established a system to measure clinically relevant movements in people with PD using skeleton data, which calculated temporal and spatial features of skeleton joints [20]. Oskarsson et al. proposed a method to determine the spectrum of 3-dimensional reachable workspace encountered in a cross-sectional cohort of individuals with ALS applying skeleton data aiming to assess the difference between healthy controls and ALS patients [21].

Gait is a time-dependent process, in which the time features of lower limb movements are also essential for gait recognition. Sarbaz et al. designed a feed-forward artificial perceptron neural network with a hidden layer as the classifier that took frequency features extracted from time intervals with a power spectrum as the input [22]. Moreover, Zeng et al. demonstrated that the proposed Radial Basis Function (RBF) neural networks model and deterministic learning fusion method can effectively separate the gait time series between the groups of healthy controls and Neurodegenerative patients (ALS, HD and PD) [9]. In addition, an artificial neural network were constructed to model healthy behavior and train gait time interval series data of patients with HD [10].

Another feature obtained by force sensors (such as force-sensitive resistors with the output roughly proportional to the force under the foot) can represent more subtle changes in lower limb movements. Manap et al. proposed the feed-forward multilayer perceptron neural network to identify PD patients from normal people with gait patterns extracted from Ground Reaction Force (GRF, in Newton) recordings [23]. Jane et al. built a Q-backpropagated time delay neural network in predicting PD severity of gait disorders using GRF [24].

Based on these temporal data in gait, finer pattern of dynamic gait details can be effectively distinguished by LSTM. We proposed



**Fig. 1.** A typical RNN model. Every RNN module represents a RNN cell at a time, there is an activation function “tanh” in it, and each module has two inputs and two outputs.

a dual channel LSTM model that combining two types of features including time data and force data to model the gait dynamics of patients in NDs, which can improve the diagnostic process for objectively analyzing the changes in motor behaviors. Our experiment demonstrates that using fusion feature is more effective than that of single feature.

## 3. Methods

### 3.1. Introduction to RNN and LSTM

To be self-explanatory, the basic theory of RNN and LSTM will be briefly introduced before describing the structure of our proposed dual channel LSTM model in NDs classification.

RNN is a special artificial neural network where connections between units form a directed cycle, which can exhibit dynamic timing behavior. Unlike the feed-forward neural networks, the internal memory cell in RNN makes it naturally adept in handling sequential data such as connected handwriting recognition, speech recognition, and activity recognition. The RNN is a chainlike structure comprised of repeating modules that allows for information retention by combining previous states with current input [25]. This repetitive module has a simple formation (“tanh” function) in a basic RNN. For a given input series  $x_t$  ( $t = 1, 2, \dots, T$ ), the hidden state of a recurrent module  $h_t$  is calculated using Eq. (1). The output of the module  $y_t$  is calculated as in Eq. (2). Fig. 1 shows a typical structure of RNN with 3 modules.

$$h_t = \tanh(W_{xh}x_t + W_{hh}h_{t-1} + b_h) \quad (1)$$

$$y_t = \text{softmax}(W_{ho}h_t + b_o) \quad (2)$$

where  $W_{xh}$ ,  $W_{hh}$ ,  $W_{ho}$  indicate the connection weights from the input  $x$  to the hidden state  $h$ , the hidden state  $h$  to itself and the hidden state to the output  $y$  respectively.  $b_h$  and  $b_o$  are bias vectors,  $\tanh$  and  $\text{softmax}$  are the activation functions in the hidden layer and the output layer.

However, the memory ability of RNN is weak for long time steps because of its limited function, to solve this problem, LSTM has complicated dynamics that allow it to easily memorize information for an extended number of timesteps, which has similar repetitive module to RNN and can learn long-term dependency information from the input data [26]. The internal organization of repetitive module in LSTM has four interactive operations (3 sigmoid and 1 tanh), which enables it by extracting valid information from dynamic data.

There are three gates (input, forget and output) in the basic cell of LSTM, each gate has a sigmoid activation function and a point-wise multiplication operation. We choose a variant of LSTM called

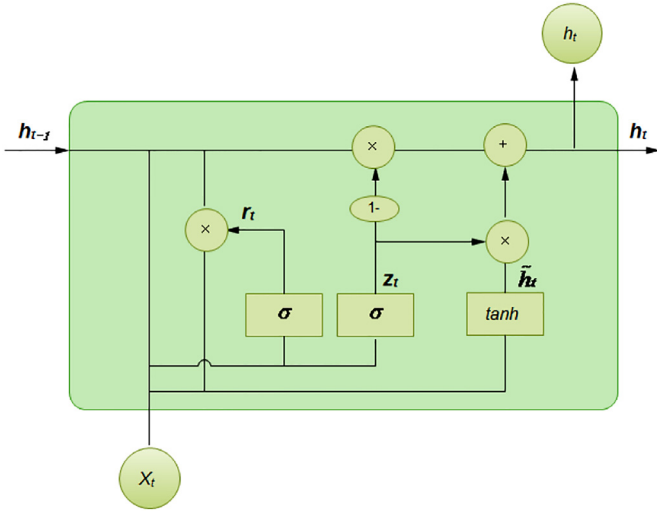


Fig. 2. The GRU model.

the Gated Recurrent Unit (GRU) as learning model, which combines the input gate and forgot gate into an update gate and mixes cell state and hidden state. Fig. 2 shows the structure of a GRU cell and illustrates the operations of the gates. The basic cell of the GRU is defined as the following equations:

$$z_t = \sigma(W_z \cdot [h_{t-1}, x_t]) \quad (3)$$

$$r_t = \sigma(W_r \cdot [h_{t-1}, x_t]) \quad (4)$$

$$\tilde{h}_t = \tanh(W \cdot [r_t \odot h_{t-1}, x_t]) \quad (5)$$

$$h_t = (1 - z_t) \odot h_{t-1} + z_t \odot \tilde{h}_t \quad (6)$$

where  $z_t$  denotes the output of update gate to the network at time step  $t$ , where  $\sigma$  is the logistic sigmoid function.  $x_t$  and  $h_{t-1}$  are the input and the previous hidden state, respectively.  $W_r$  and  $W_z$  are weight matrices which are learned.  $r_t$  denotes the reset gate and the actual activation of the proposed unit  $h_t$  is then computed by Eq. (6). The update gate  $z$  selects whether the hidden state is to be updated with a new hidden state  $\tilde{h}$ . The reset gate  $r$  decides whether the previous hidden state is ignored.

### 3.2. The dual channel LSTM

In this paper, we propose a dual channel LSTM for diagnosis of NDs, each channel is composed of two layers of LSTM each of which consists of the same number of expanded nodes (time steps, cells) according to the input data. After calculation of the LSTM memory cells, outputs of the two channels are concatenated into a feature vector (256D) with a 128-dimensional vector generated in one channel, as the fusion feature that is input into the *softmax* classifier for NDs classification. According to each training accuracy, the dual channel LSTM model minimizes the loss to obtain the optimum weights and biases. The proposed dual channel LSTM model is illustrated in Fig. 3.

The dual channel LSTM model is divided into two channels taking time data and force data as input in batches, each channel determines different number of LSTM cells in terms of the number and time variation of the input ( $M$  and  $N$ ). By synchronous training, the output feature of the last cell from the second layer of LSTM ( $O_M$ ,  $O_N$ ) is extracted to connect into a feature vector that can predict the probability of each class. Taking the maximum probability value, the model can eventually achieve the accuracy of disease

classification. Taking force channel for example, the internal composition of two-layer LSTM is introduced in detail.

In the force channel, we design a two-layer LSTM for obtaining deeper feature, each LSTM layer has  $N$  ( $N = 100$ ) cells with the last output  $O_N$ . The binary force data will be fed into the LSTM after converted into decimal and normalized, which makes the data fixed in length in terms of the time steps. The gait cycle of the participant is nearly 1 s with enough foot changing information for time steps (10 ms each time step). After calculating the data in the LSTM cell, a feature vector is obtained to represent the changing force information. Through repeatedly training and learning, connection weights between each cell can be optimized. Fig. 4 depicts the structure of the used two-layer LSTM.

The only difference between the time channel and the force channel is that it has  $M$  ( $M = 20$ ) cells in each LSTM layer due to a small amount of data that excludes erroneous data.

Each training sample is a reshaped sequence of  $N$   $f$ -dimensional force feature vector. Suppose that  $K$  (batch size) training samples are fed into the first LSTM layer, the size of the input is  $N \times K \times f$ . LSTM will be expanded by time steps and every sample in one training sample can be input to one LSTM cell respectively. One LSTM cell takes  $K \times f$  data as the input, the output of which is an  $h$ -dimensional feature vector (hidden layer output) that can be adjusted to an appropriate value. These settings are also applicable to the second layer of LSTM. The output at time  $N$  contains information of all the previous nodes, so we just choose the last output ( $O_{100} \in \mathbb{R}^h$ ) as the basis for the classification. After merging with the output of time channel, a  $g$ -dimensional ( $g$  classes) vector is generated based on the weights and biases, which are classified using a multi-class classifier *softmax* to map the output of the LSTM to a probability distribution and diagnose whether the patient suffer from certain NDs.

## 4. Experiment

### 4.1. Datasets description

In this study, we used predetermined data taken from PhysioNet [27]. The dataset contained gait signal of 48 patients with NDs and 16 healthy persons. Participants were walking at their usual pace along a 77-m-long hallway for 5 min in each sample including recorded signal of stride, swing, and stand times for each leg and double support signal for both legs, the specific attributes we used are shown in Table 2. An expert physician labeled patients states from 0 to 13 (0 equal to the most severe state and 13 for a healthy one). To measure time intervals, a 12-bit on-board analog-to-digital converter sampled the output of foot switches at 300 Hz. First 20 s of records were excluded to reduce initial oversight and median filter was used to remove data points (outliers) that were far away from the median value. These outliers were mainly due to the turns at the end of the hallway [28].

The raw data were obtained using force-sensitive resistors, with the output roughly proportional to the force under the foot. Stride-to-stride measures of footfall contact times were derived from these signals. In addition, the force data was placed in binary files distinguishing left and right feet. Each of the 64 records was identified by the name of the subject group (hunt, park, als, or control) followed by an arbitrary ID number.

The dataset had also explained clinical information for each subject including age, gender, height, weight, walking speed, and a measure of disease severity or duration. Table 1 describes the information of subjects.

We randomly selected the time stride interval data of 4 subjects ('control1', 'als10', 'park6', 'hunt9') as the example, the difference between NDs can be illustrated in Fig. 5, and the corresponding

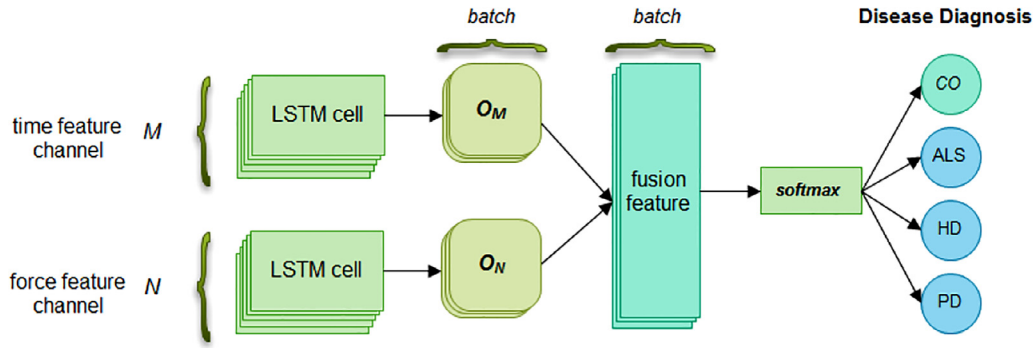
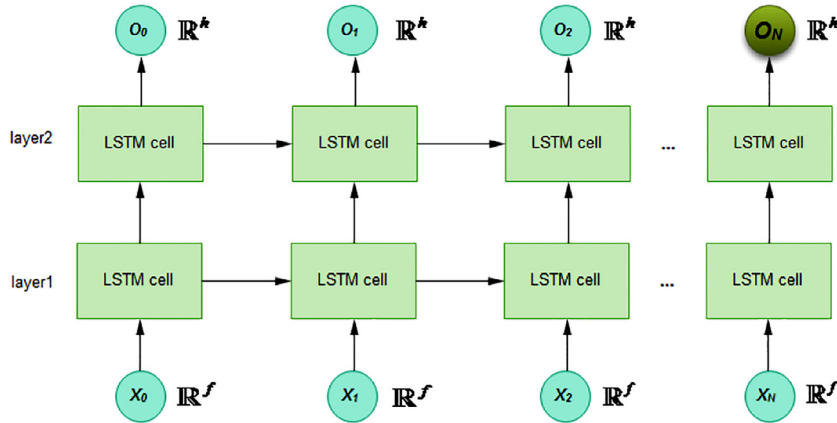


Fig. 3. The dual channel LSTM model.

Fig. 4. The structure of the two-layer LSTM. The input of the LSTM is  $x_N, x_N \in \mathbb{R}^f$  ( $f = 12$ ), which from different time periods are entered into  $N$  ( $N = 100$ ) LSTM cell. The output of  $N$ th time point is used for feature fusion.

**Table 1**  
The basic information of the subjects in the dataset.

Subjects	Total number	Male	Female	Age ( < 50)	Age (50–70)	Age ( > =70)
Healthy controls	16	2	14	11	4	1
ALS	13	10	3	4	7	2
PD	15	10	5	1	7	7
HD	20	6	14	13	5	2

**Table 2**  
The stride-to-stride measures of the subjects.

Column	Contents
1	Elapsed Time (s)
2	Left Stride Interval (s)
3	Right Stride Interval (s)
4	Left Swing Interval (s)
5	Right Swing Interval (s)
6	Left Swing Interval (% of stride)
7	Right Swing Interval (% of stride)
8	Left Stance Interval (s)
9	Right Stance Interval (s)
10	Left Stance Interval (% of stride)
11	Right Stance Interval (% of stride)
12	Double Support Interval (s)
13	Double Support Interval (% of stride)
14	Left Force Signal (extracted from binary file)
15	Right Force Signal (extracted from binary file)

#### 4.2. Experimental results

Experiments were implemented using Tensorflow library and tested on an Intel Core i5 computer with 31.3GB RAM. The computation time was related to the number of the LSTM cells and the size of the feature data employed. We assigned feature vector sequences for all the 64 healthy controls and NDs subjects. The time series and force series for each subject were divided into two sub-series, one for training and the other for test. Based on the method described in Section 3, we extracted all the 64 subjects gait features through time and force data, which meant the input of the LSTM networks is  $x=[\text{left swing interval, right swing interval, ..., left stance interval, right stance interval}]^T$  and  $y=[\text{force signal}]^P$ . In order to eliminate the data difference between different gait features, all the gait feature data are normalized to  $[0,1]$ .

We conducted three experiments. First, we detected ND patients from the healthy one, the dataset was divided into two classes, ND patients and normal people; second, the participants was classified into 3 classes in accordance with different diseases (ALS, PD, and HD); and the above two diagnostic approaches covered comparison of various machine learning methods; finally, the parameters of dual channel LSTM were transformed to reach the optimum results of classification. We also showed the classification results of the original LSTM and our proposed model. Before the experiment, we removed some bad data that had abnormal fluctuations impacting on the implementation of the experiment. The classification results of the NDs in the all-training-all-test and leave-one-out training-testing styles are presented in the following tables.

force data is shown in Fig. 6. We can see that the range of fluctuation of time is smaller than that of fluctuation.



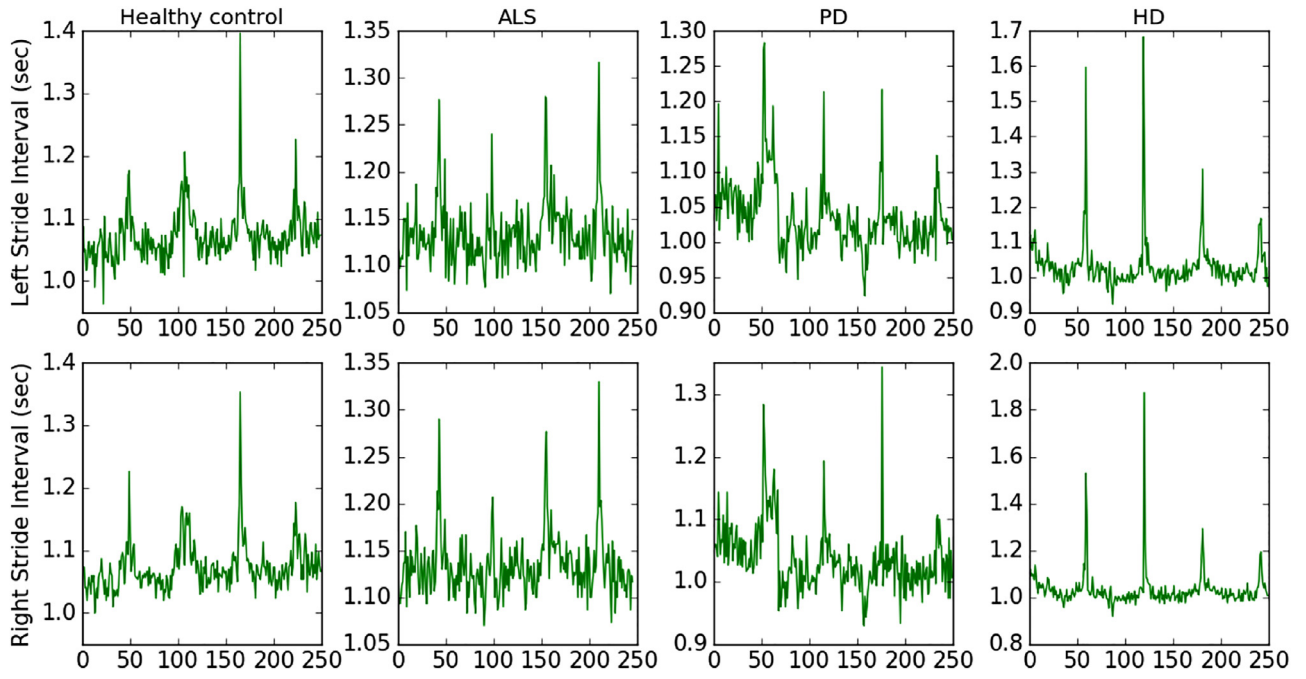


Fig. 5. Comparison of the performance of subjects for different NDs. And the stride interval data of the left and right foot has been displayed.

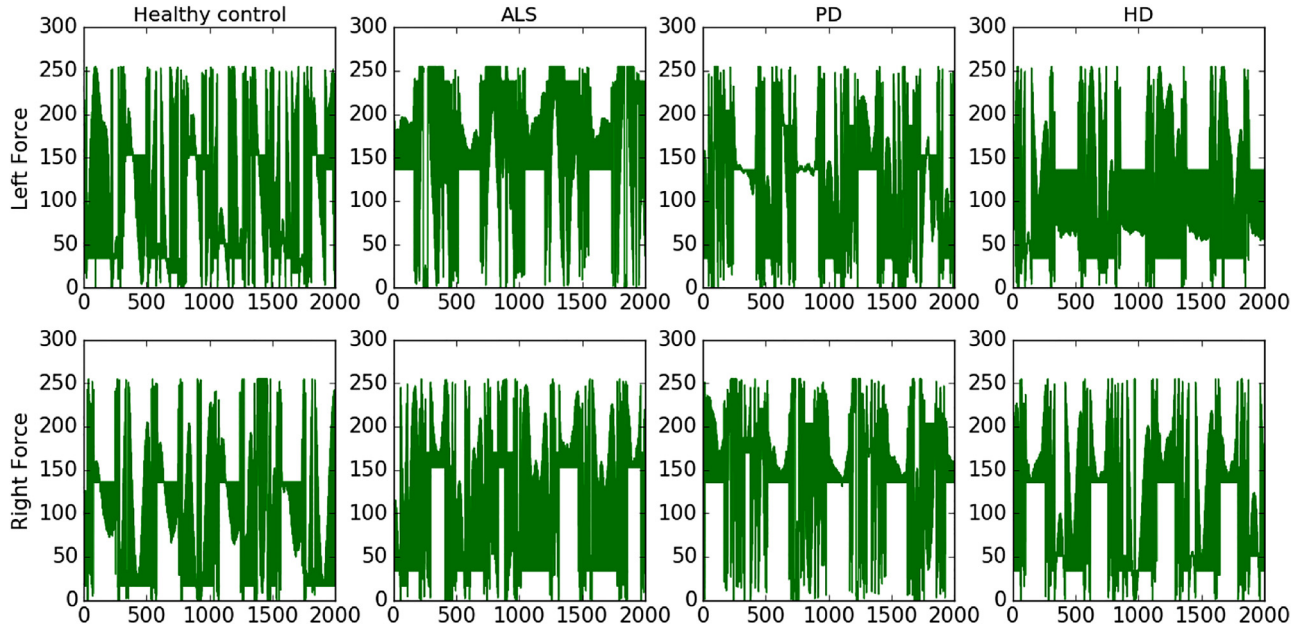


Fig. 6. Comparison of the performance of subjects for different NDs. And the force data of the left and right foot has been displayed.

#### 4.2.1. Classification for NDs and healthy controls

In the first experiment, NDs patients and healthy persons were classified including 16 healthy controls, 13 ALS patients, 15 PD patients and 20 HD patients. Performance of the proposed classification approach was evaluated by the all-training-all testing and leave-one-out cross-validation methods using the strategy explained above. The results have been shown in Table 3.

We respectively fed the force and time series into the two LSTM channels with different settings of time steps and input dimension. The time channel got 20 time steps and 13 dimensional input, and the force channel obtained 100 and 2 (every force sample had only 2 decimal digits, left and right feet force data), the number of training iterations and batch size can be changed according to the practical demand. Table 4 illustrates the classification results

Table 3

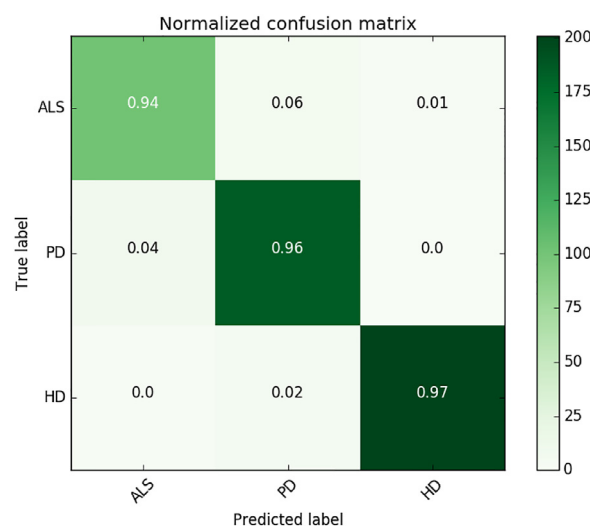
Performance of diagnosing NDs (two groups) evaluated by the all-training-all-testing and leave-one-out cross-validation methods. “-” denotes that the study did not deal with.

Evaluation methods	Subjects	Zeng et al., [9]	Original LSTM	The proposed method
All-training-all-testing	ALS vs. CO	93.10%	98.36%	<b>99.57%</b>
	PD vs. CO	100%	100%	<b>100%</b>
	HD vs. CO	100%	100%	<b>100%</b>
	NDs vs. CO	93.75%	95.48%	<b>97.88%</b>
Leave-one-out	ALS vs. CO	89.66%	93.28%	<b>97.43%</b>
	PD vs. CO	87.10%	90.69%	<b>97.33%</b>
	HD vs. CO	83.33%	87.45%	<b>94.96%</b>
	NDs vs. CO	-	-	<b>96.42%</b>

**Table 4**

Performance of diagnosing NDs (two groups) compared with other studies. “–” denotes that the study did not deal with.

Studies	ALS vs. CO	PD vs. CO	HD vs. CO
Athisakthi et al., [29]	56.66%	51.00%	54.28%
Our Method	<b>91.23%</b>	<b>90.34%</b>	<b>99.88%</b>
Banaie et al., [8]	100%	80.00%	71.43%
Our Method	<b>100%</b>	<b>92.45%</b>	<b>92.10%</b>
Sánchez-Delacruz et al., [6]	96.13%	90.36%	88.67%
Our Method	<b>97.43%</b>	<b>97.33%</b>	<b>94.96%</b>
khorasani et al., [30]	–	90.32%	–
Our Method	–	<b>97.33%</b>	–
Xia et al., [31]	–	96.77%	–
Our Method	–	<b>97.33%</b>	–
Xia et al., [32]	92.86%	–	–
Our Method	<b>97.43%</b>	–	–

**Fig. 7.** Performance of diagnosing NDs.**Table 5**

Performance of diagnosing NDs (3 groups).

Banaie et al., [8]	Original LSTM	The proposed method
86.96%	89.58%	<b>95.67%</b>

compared with other studies according to their different training settings. The average result would be used if the research had several experimental results in one classification task. In this work, we took 80 percent data as the training set, 20% as testing sets.

#### 4.2.2. Classification of NDs

In the second experiment, NDs (ALS, HD and PD) were identified. The difference between three diseases has been displayed in Figs. 5 and 6. It can be seen that fluctuation ranges of ALS and PD data were similar. We used the confusion matrix to represent the classification accuracy of the three NDs. We choose the best test result that has been shown in Fig. 7. We can see that the HD obtains highest identifying rate employing our experiment setting (20% testing, because the work we want to contrast adopted it). And the results were compared with one research in Table 4. (2 groups) and Table 5 (3 groups).

#### 4.2.3. Parameter transformation

In the last experiment, we changed vital parameters of the LSTM model that was superior to the basic LSTM model including model layer and feature dimension for every sample. We set up some variables in advance in order to reach the best results.

**Table 6**

Experimental results for classification of NDs patients and healthy controls with different parameters using leave-one-out cross-validation methods.

Data	Time:20 steps Force:100 steps 2 layers LSTM	Time:20 steps Force:50 steps 2 layers LSTM	Time:20 steps Force:100 steps 1 layer LSTM	Time:20 steps Force:50 steps 1 layer LSTM
ALS vs. CO	<b>97.43%</b>	97.25%	94.03%	93.58%
PD vs. CO	<b>97.33%</b>	96.47%	95.43%	93.45%
HD vs. CO	<b>94.96%</b>	92.11%	92.10%	92.66%
NDs vs. CO	<b>96.42%</b>	96.14%	95.04%	95.23%

**Table 7**

Experimental results for classification of 3 groups of NDs patients with different parameters using leave-one-out cross-validation methods.

Time:20 steps Force:100 steps 2 layers LSTM	Time:20 steps Force:50 steps 2 layers LSTM	Time:20 steps Force:100 steps 1 layer LSTM	Time:20 steps Force:50 steps 1 layer LSTM
<b>95.67%</b>	95.38%	93.44%	93.25%

The dropout was set to 0.5 to prevent over-fitting with the learning rate of 0.001 and 128-dimensional hidden layer output.

The number of time steps in time channel can be set to 20, since every 20 samples contained enough gait information. With this setting we adjusted the steps of force channel and LSTM layer of the two channels to achieve the best classification accuracy. We showed the optimal values of the parameters in Tables 6 and 7. For the two classification tasks above, our design was preferable to other model transformations.

## 5. Discussion

Automated methods for classifying gait disorders prior to further assessment and diagnosis of the severity of diseases may help nerve physicians to focus on the correct information of diseases and to choose better health treatment. We suggest to use the proposed dual channel LSTM model to diagnose different NDs achieving better results than the popular classification methods, which can not only simplify complicated pre-diagnosis process, but also assist doctors to diagnose diseases effectively. It can also quantify gait data characteristics and provide a detailed objective description of diseases. We implement our architecture on the computer with NVIDIA Tesla K40c GPU, it takes 23.45 min for training and 2.66 ms for testing of one testing sample on average. Moreover, our model achieves higher accuracy among the state-of-the-art studies (Zeng et al. [9], Banaie et al. [8], Sánchez-Delacruz et al., [6] etc.) as we mentioned above. On the basis of these methods, we can present new viewpoints to help solve the problem of NDs diagnosis.

According to the survey, there are most researches related to NDs is based on data processing and feature recognition before disease identification, only a few focus on treatment or rehabilitation period after the diagnosis. In addition, patient data is not public which is difficult to acquisition limiting the application of excellent algorithms in medical diagnosis. Equipments for monitoring and collecting patient's action clinical signs and symptoms should be developed to obtain more disease information. We will request for a large amount of daily data from NDs patients to produce data sets utilizing more features to describe and analyze NDs diseases.

## 6. Conclusion

In this paper, we reported our study of predicting the Neurodegenerative diseases using state-of-the-art machine learning techniques. A dual channel LSTM was proposed to model the gait data (time intervals and force binary data) recorded by force-sensitive

sensors representing features of gait dynamics over a long period. Once the model is well trained, it can directly diagnose disease from sensor data to avoid the subjectivity concerns. The provided experimental results can show the availability of gait data enables the data-driven learning frameworks to outperform hand-crafted features. They also show the effectiveness of the proposed dual channel LSTM model over traditional recurrent models.

Our work is a preliminary research in diagnosing NDs. In future work, we intend to provide more comprehensive gait disorder diagnostic tools for more complex gait disorders that are difficult for the clinicians to detect. We plan to assist their assessment process in the clinic, evaluate these analytic systems with properly designed clinical studies, and design new methods for diagnosis, rehabilitation evaluation and treatment plan development. Moreover, more kinds of diseases and higher prediction accuracy can be achieved by using fused features from multiple data sources, such as skeleton data, fMRI and CT image, RGB image and biochemical data. Our model will be also applied in the mobile devices for routine monitoring of patients, which can be a baseline and the proposed approach is expected to inspire more applications of machine learning techniques in computer-aided diagnosis system.

## Acknowledgments

This paper is supported by the National Natural Science Foundation of China (NSFC) (no. 61501417 & no. 61271405), the International Science & Technology Cooperation Program of China (ISTCP) (no. 2014DFA10410).

## References

- [1] J.E. Cutting, L.T. Kozlowski, Recognizing friends by their walk: gait perception without familiarity cues., *Bull. Psychon. Soc.* 9 (5) (1977) 353–356.
- [2] W.C. Koller, J. Trimble, The gait abnormality of huntington's disease., *Neurology* 35 (10) (1985) 1450–1454.
- [3] J.M. Hausdorff, S.L. Mitchell, R. Firtion, C.K. Peng, M.E. Cudkowicz, J.Y. Wei, A.L. Goldberger, Altered fractal dynamics of gait: reduced stride-interval correlations with aging and huntington's disease., *J. Appl. Physiol.* 82 (1) (1997) 262.
- [4] E.R. Kandel, J.H. Schwartz, T.M. Jessell, *Principles of Neural Scienc*, Fourth edition, Medical, 2000.
- [5] Z. Li, W. Chen, J. Wang, J. Liu, An automatic recognition system for patients with movement disorders based on wearable sensors, in: *Industrial Electronics and Applications*, 2014, pp. 1948–1953.
- [6] E. Sánchez-Delacruz, F. Acosta-Escalante, M.A. Wister, J.A. Hernández-Nolasco, P. Pancardo, J.J. Méndez-Castillo, Gait recognition in the classification of neurodegenerative diseases, in: *International Conference on Ubiquitous Computing and Ambient Intelligence* 2014, 8867, 2014, pp. 128–135.
- [7] L. Sugavanewaran, K. Umapathy, S. Krishnan, Ambiguity domain-based identification of altered gait pattern in als disorder., *J. Neural Eng.* 9 (4) (2012) 046004.
- [8] M. Banaie, M. Pooyan, M. Mikaili, Introduction and application of an automatic gait recognition method to diagnose movement disorders that arose of similar causes, *Expert Syst. Appl.* 38 (6) (2011) 7359–7363.
- [9] W. Zeng, C. Wang, Classification of neurodegenerative diseases using gait dynamics via deterministic learning, *Inf. Sci.* 317 (C) (2015) 246–258.
- [10] A. Pourhedayat, Y. Sarbaz, A grey box neural network model of basal ganglia for gait signal of patients with huntington disease., *Basic Clin. Neurosci.* 7 (2) (2016) 107.
- [11] A. Haque, A. Alahi, L. Fei-Fei, Recurrent attention models for depth-based person identification, *IEEE Conference on Computer Vision & Pattern Recognition CVPR2016*.
- [12] J. Liu, A. Shahroudy, D. Xu, G. Wang, Spatio-temporal lstm with trust gates for 3d human action recognition, in: *European Conference on Computer Vision*, Springer, 2016, pp. 816–833.
- [13] L. Wu, C. Shen, A.v.d. Hengel, Convolutional lstm networks for video-based person re-identification, *Comput. Vis. Pattern Recogn.* (2016).
- [14] W. Zhu, C. Lan, J. Xing, W. Zeng, Y. Li, L. Shen, X. Xie, Co-occurrence feature learning for skeleton based action recognition using regularized deep lstm networks, Co-occurrence feature learning for skeleton based action recognition using regularized deep lstm networks(2016).
- [15] R.R. Viorio, B. Shuai, J. Lu, D. Xu, G. Wang, A siamese long short-term memory architecture for human re-identification, in: *European Conference on Computer Vision*, Springer, 2016, pp. 135–153.
- [16] S. Sharma, R. Kiro, R. Salakhutdinov, Action recognition using visual attention, *Comput. Sci.* (2015).
- [17] N.Y. Hammerla, S. Halloran, T. Ploetz, Deep, convolutional, and recurrent models for human activity recognition using wearables, *J. Sci. Comput.* (2016).
- [18] J.M. Hausdorff, Z. Ladin, J.Y. Wei, Footswitch system for measurement of the temporal parameters of gait, *J. Biomech.* 28 (3) (1995) 347–351.
- [19] R. Torres, M. Huerta, R. Clotet, R. González, G. Sagbay, M. Erazo, J. Pirrone, Diagnosis of the corporal movement in parkinson's disease using kinect sensors., in: *World Congress on Medical Physics and Biomedical Engineering*, 2015, pp. 1445–1448.
- [20] G. Brook, B. Gillian, J. Dan, M. Dadirayi, O. Patrick, R. Lynn, Accuracy of the microsoft kinect sensor for measuring movement in people with parkinson's disease, *Gait Posture* 39 (4) (2014) 1062–1068.
- [21] B. Oskarsson, N.C. Joyce, E. De Bie, A. Nicorici, R. Bajcsy, G. Kurillo, J.J. Han, Upper extremity 3-dimensional reachable workspace assessment in amyotrophic lateral sclerosis by kinect sensor, *Muscle Nerve* 53 (2) (2016) 234–241.
- [22] Y. Sarbaz, F. Towhidkhah, S. Gharibzadeh, A. Jafari, Gait spectral analysis: an easy fast quantitative method for diagnosing parkinson's disease, *J. Mech. Med. Biol.* 12 (03) (2012) 1250041.
- [23] H.H. Manap, N.M. Tahir, A.I.M. Yassin, R. Abdullah, Anomaly gait classification of parkinson disease based on ann, in: *System Engineering and Technology (ICSET)*, 2011 IEEE International Conference on, IEEE, 2011, pp. 5–9.
- [24] Y.N. Jane, H.K. Nehemiah, K. Arputharaj, A q-backpropagated time delay neural network for diagnosing severity of gait disturbances in parkinson's disease, *J. Biomed. Inf.* 60 (2016) 169–176.
- [25] A. Graves, A. Mohamed, G. Hinton, Speech recognition with deep recurrent neural networks," *icassp* 1(2003) (2013) 6645–6649.
- [26] S. Hochreiter, J. Schmidhuber, Long short-term memory, *Neural Comput.* 9 (8) (1997) 1735–1780.
- [27] NIGMS, physionet database., (<https://physionet.org/physiobank/database/gaitndd/>).
- [28] J.M. Hausdorff, M.E. Cudkowicz, R. Firtion, J.Y. Wei, A.L. Goldberger, Gait variability and basal ganglia disorders: stride-to-stride variations of gait cycle timing in parkinson's disease and huntington's disease., *Movement Disorders* 13 (3) (1998) 428–437.
- [29] A. Athiakhithi, D.M. Pushparani, Detection of Movement Disorders Using Multi SVM, *Society of Photo-Optical Instrumentation Engineers*, 2013.
- [30] A. Khorasani, M.R. Daliri, Hmm for classification of parkinson's disease based on the raw gait data, *J. Med. Syst.* 38 (12) (2014) 1–6.
- [31] X. Yi, Y. Qiang, Q. Gao, Y. Lu, D. Zhang, Symmetry analysis of gait between left and right limb using cross-fuzzy entropy, *Comput. Math. Methods Med.* 2016 (1) (2016) 1737953.
- [32] Y. Xia, Q. Gao, Y. Lu, Q. Ye, A novel approach for analysis of altered gait variability in amyotrophic lateral sclerosis, *Med. Biol. Eng. Comput.* 54 (9) (2016) 1–10.