

# LSTM for Diagnosis of Neurodegenerative Diseases Using Gait Data

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

<sup>a</sup> College of Information Science and Engineering, Ocean University of China, Qingdao, China  
266100;

<sup>b</sup> University of Portsmouth, Portsmouth, UK  
zhaoaite@stu.ouc.edu.cn, qilin@ouc.edu.cn, lijie@stu.ouc.edu.cn, dongjunyu@ouc.edu.cn,  
hui.yu@port.ac.uk

## ABSTRACT

Neurodegenerative diseases (NDs) usually cause gait disorders and postural disorders, which provides an important basis for NDs diagnosis. By observing and analyzing these clinical manifestations, medical specialists finally give diagnostic results to the patient, which is inefficient and can be easily affected by doctors' subjectivity. In this paper, we propose a two-layer Long Short-Term Memory (LSTM) model to learn the gait patterns exhibited in the three NDs. The model was trained and tested using temporal data that was recorded by force-sensitive resistors including time series, such as stride interval and swing interval. Our proposed method outperforms other methods in literature in accordance with accuracy of the predicted diagnostic result. Our approach aims at providing the quantitative assessment so that to indicate the diagnosis and treatment of these neurodegenerative diseases in clinic.

**Keywords:** Neurodegenerative diseases, diagnosis, gait disorders, temporal data, LSTM.

## 1. INTRODUCTION

Neurodegenerative diseases (NDs) are diseases of the brain and spinal cord neuron loss, which are divided into two groups according to their phenotype: one that affects movement, such as cerebellar ataxia, one that affects memory, such as related dementia [1]. We embarked on the former in accordance with the gait characteristics to diagnose three NDs including Huntington's disease (HD), Parkinson's disease (PD) and Amyotrophic Lateral Sclerosis (ALS). The three diseases have different clinical manifestations in gait, which can be used as a basis for the diagnosis of NDs. Much literature also diagnoses these three NDs by gait analysis [2–4]. The following will describe the features of these three diseases in gait.

Huntington's disease (HD) is a dominant genetic disease. The patient's nervous system gradually degenerated, resulting in abnormal body movements. In the middle stage of this disease, there will be involuntary movements; walking and balance disorders; dance like movements, twisting, rolling, and unstable gait; unable to control the speed and power of the movement [5, 6].

Parkinson's disease (PD) is a degenerative brain disorder characterized by a loss of midbrain dopamine (DA) neurons [7–10]. Gait disturbance occurs in the early stages and is used as a principal manifestation for diagnosis. Some gait disturbances such as festinating gait, short gait and freezing gait may make PD diagnosis easier, much literature has identified PD by analyzing gait [6, 11].

Amyotrophic Lateral Sclerosis (ALS) is a motor neuron disease that causes the muscles of the limbs, trunk, chest and abdomen to gradually become weak and atrophied. Due to motor neuron lesions, the movement of the limbs tends to disorder. The complications can also occur in spastic paralysis of lower limbs, scissors gait, etc. [12, 13].

Based on these motor symptoms from ND patients, we are committed to optimizing the existing diagnostic system which is one of the more labor-intensive tasks in medical procedure. Medical experts make a decision by these gait features that cannot provide quantified diagnostic basis and the diagnostic process will be subjectivity and low efficiency. The general health care system does not always maintain accurate and rapid diagnosis, but machine learning methods can not only substantially reduce costs, but its diagnosis can be achieved almost at real time. In many cases, these methods can provide more accurate diagnosis than experienced doctors so that it can benefit the auxiliary diagnosis.

There are many approaches being successfully applied in diagnosing NDs which include the kernel Fisher discriminant (KFD), the naive Bayesian approach (NB), support vector machine (SVM) and nearest neighbor (NN) [3, 14–16]. In addition, neural networks have also been adopted in this field [4, 17]. However, these methods 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.

As the well known deep learning method, Long Short-Term Memory (LSTM) has recently covered various fields including action recognition and gait recognition [18–24]. Different from the above methods, LSTM is characterized by the ability to capture subtle changes in long continuous temporal data and memorize important data information, which is propitious to classify the temporal data. Inspired by these studies, we propose to use a variant of LSTM, Gated Recurrent Unit (GRU) for NDs diagnosis employing the gait information. A two-layer GRU model with a softmax classifier has been constructed which was trained and tested on public datasets that are time records of subjects measured by force sensitive resistors. These data quantify the dynamic characteristics of gait in the three NDs according to the time series. The experimental results show that our method improves the accuracy of the diagnosis and stay ahead of previous approaches in literature.

This paper is organized based on the following sections. Section II describes technologies on diagnosis of the three NDs and existing research about this study. Section III and Section IV present our methods and experiment on NDs diagnosis. Finally, Section V concludes this paper.

## 2. RELATED WORKS

The medical industry generates a large amount of data every day, which is an important application area for big data that is the source of new technology. As the most popular new method, machine learning has been applied to aid diagnosis, which is not merely a collection of data to draw statistical conclusions, but rather a certain artificial intelligence algorithm in which inferences can be derived.

Cho et al. designed a vision-based analysis system for gait recognition among patients with PD employing principal component analysis (PCA) and linear discriminant analysis (LDA) in PD diagnosis, which processed motion video data from 7 PD patients and 7 normal people captured by a camcorder with a 95.49% classification result [25]. Das compared different classification methods for detecting PD in a speech data set gathered from 31 people, 23 of which were PD patients [26]. Xia et al. feed the extracted features from Teager energy operator into the SVM classifier to diagnose pathological progression of ALS using time records [27]. Li et al. adopted the SVM and NB methods to support the classification of different NDs (ALS, PD, HD), which made a concrete analysis of the gait data measured by accelerometers [3].

For the past few years, neural networks are gradually attracting the attention of researchers. Zeng et al. demonstrated that the proposed RBF neural networks model can effectively separate the gait time series between the groups of healthy controls and neurodegenerative patients [17]. In addition, an artificial neural network was constructed to model healthy behavior and train gait time interval series data of patients with HD in 2016 [4].

There is a growing awareness that the gait data is based on the time interval data and the above methods are not designed for this condition. For problems with temporal data, Khorasani et al. adopted Hidden Markov Model (HMM) and Least Squares Support Vector Machine (LS-SVM) to diagnose PD from 16 healthy subjects and 15 PD subjects, which used the temporal data collected by force sensors [28].

This paper proposed a two layer GRU method to model the gait dynamics of patients in NDs, which can improve the diagnosis for objectively analyzing the complexity and change diversity in motor behaviors. Inspired by these existing advanced work, this paper applies GRU in the diagnosis of NDs achieving the experimental results that is superior to the popular machine learning approaches.

### 3. METHODS

#### 3.1 Brief Introduction to RNN and LSTM

To be self-explanatory, this paper will briefly introduce the basic theory of RNN and LSTM before describing the structure of our proposed two-layer GRU model in NDs diagnosis.

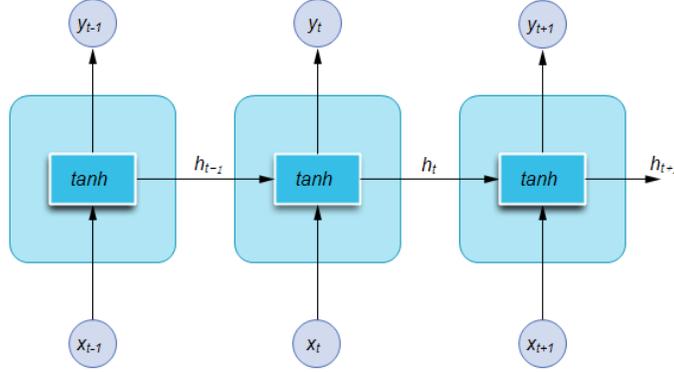


Figure 1. A RNN expand model.

RNN (Recurrent Neural Networks) is a kind of artificial neural network with node oriented connection and has a chainlike repeating neural network module, which can exhibit dynamic timing behavior. The RNN structure is shown in Fig. 1. LSTM is an extension of RNN, which can learn long term dependency information of the input data and has been successfully applied in various fields. [29]. The repetitive module in LSTM has more complex structure that are four interactive operations (3 *sigmoid* and 1 *tanh*) instead the *tanh* operation in RNN, which enables stronger memory ability. Like RNN, the structure of LSTM is flexible to be modified, which leads to many variants [30, 31].

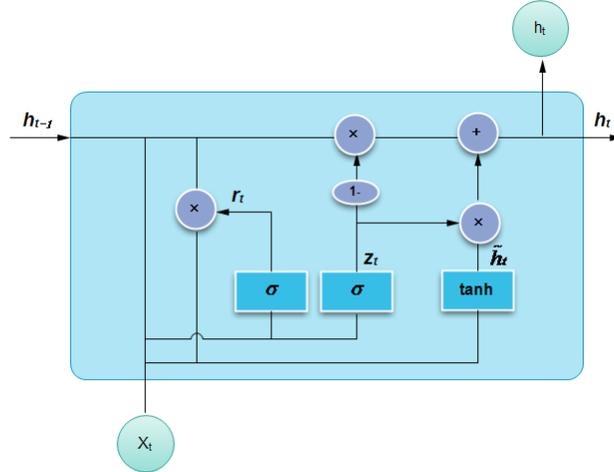


Figure 2. The illustration of one GRU cell.

The basic cell of LSTM has three gates (input, forget and output). Every gate has a *sigmoid* activation function and a pointwise multiplication operation. We choose the GRU as learning model, which combines the input gate and forget 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:

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

$$r_t = \sigma(W_r \cdot [h_{t-1}, x_t]) \quad (2)$$

$$\tilde{h}_t = \tanh(W \cdot [r_t \odot h_{t-1}, x_t]) \quad (3)$$

$$h_t = (1 - z_t) \odot h_{t-1} + z_t \odot \tilde{h}_t \quad (4)$$

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. (4). The update gate  $z$  selects whether the hidden state is to be updated with a new hidden state  $\tilde{h}_t$ . The reset gate  $r$  decides whether the previous hidden state is ignored.

### 3.2 The Proposed Two-Layer GRU for NDs Diagnosis

In this paper, we design a two-layer GRU for NDs Diagnosis, which first detects the NDs from healthy people and then identifies the three NDs according to their features. Each layer of GRU has the same number of cells receiving the temporal data by gait changing time steps, and we designs two layers make the feature more deeply. The temporal gait data is feed to GRU after a simple preprocess, which made the data fixed length in terms of the time steps. Gait cycle of the participant is nearly 20 records in the time series file with enough foot changing information thus we choose 20 as time steps. After calculating the data in the GRU cell there is a feature vector obtained to predict the probability of every class, to take the maximum value, the model can eventually achieve the accuracy of diagnosis class. By repeated training this model, connection weights between each node reaches optimal. Fig. 3 depicts the structure of the used two-layer GRU here. Experiments settings will be explained in Section V.

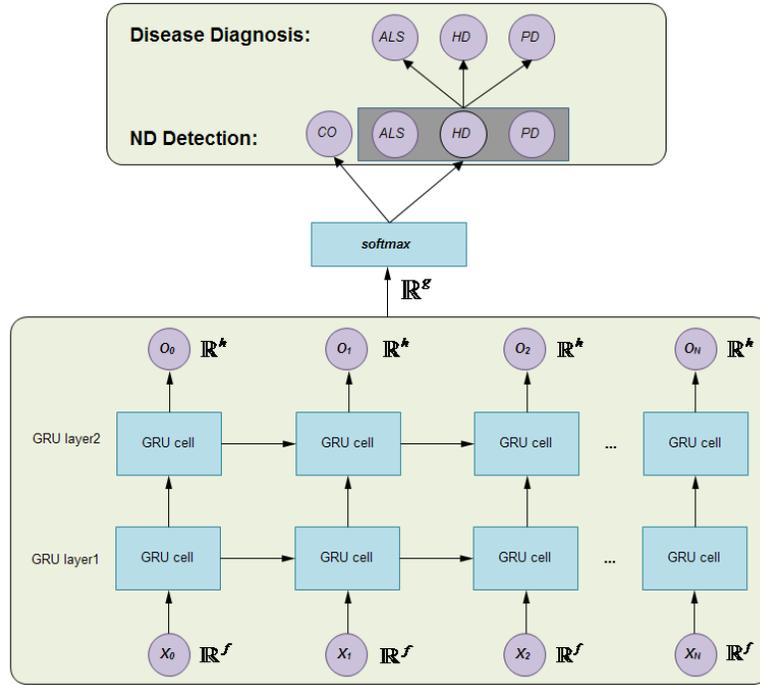


Figure 3. The illustration of one GRU cell.

Each time sequence has an  $f$  dimensional feature vector, we use  $N$  feature vectors as one training sample.  $M$  training samples can be entered into the first layer, which represents that the amount of data entered each time is  $N \times M \times f$ . GRU will be expanded according to time steps and every record in one training sample can be input to one GRU cell respectively. Every GRU cell takes  $M \times f$  data as the input. The output of each GRU cell is a specified  $h$ -dimensional feature vector (hidden layer output), it can be adjusted to an appropriate value. These settings also apply to the second layer of GRU. The output at time  $N$  contains information from all of the above points, so we take only the last  $h$ -

dimensional output ( $O_{20} \in \mathbb{R}_h$ ) as the basis for the classification. This output is transformed into a  $g$ -dimensional ( $g$  classes) vector based on weights and biases, which are classified using a multiclass classifier softmax to map the output of GRU to a probability distribution and derive which disease the patient has.

## 4. EXPERIMENT

### 4.1 Datasets Description

The gait data can be downloaded from web page of PhysioNet [32] including gait information from 48 patients with different NDs, and 16 healthy controls. Participants were walking at their usual pace along a 77-m-long hallway for 5 minutes. The raw data were obtained using force-sensitive resistors that were placed in the subject’s shoe, which produce a measure of the force applied to the ground during ambulation with a sampling frequency 300 Hz. Stride-to-stride measures of footfall contact times were derived from these signals, which have been applied in our experiment. There were 13 parameters that corresponded to each row in the dataset file including stride interval, swing interval, stance interval and double support interval of the left and right foot.

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

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          |

We randomly selected the time stride interval data of 4 subjects as the example, the difference between NDs is illustrated in Fig. 4. Gait data of the three diseases differ in amplitude.

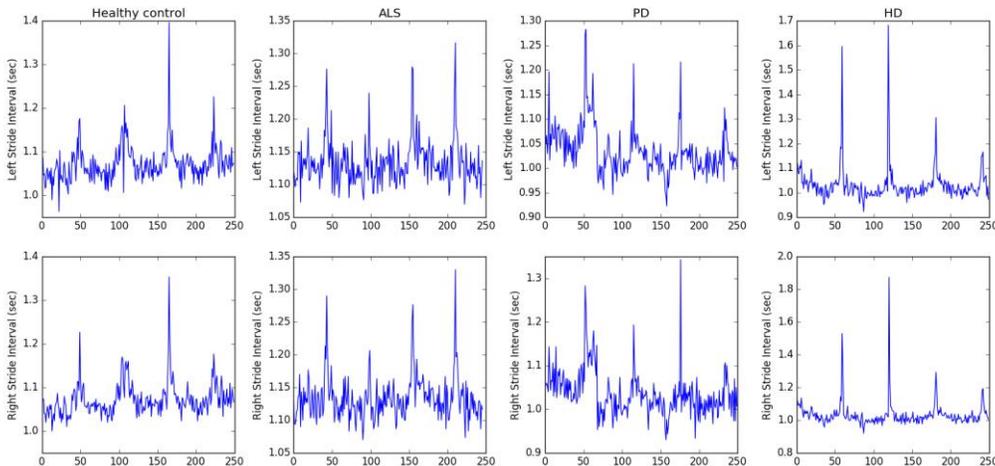


Figure 4. Different gait performance of participants.

## 4.2 Experimental Details

This section illustrates the details of the experiment comparing our experimental results with those of other researchers who utilized the same dataset and showing results of some machine learning models. Some related experimental settings can be also explained.

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, transformed the parameters of GRU to reach the optimum results of classification. We also showed the classification results of the two different GRU variants proving that the GRU performed better than the basic LSTM. The specific experimental process will be explained detailed in the following part. Before the experiment, we removed some bad data that had abnormal fluctuations impacting on the implementation of the experiment.

In the first experiment, there were two classes can be classified including 16 healthy controls, 13 ALS patients, 15 PD patients or 20 HD patients. Performance of the proposed classification approach was evaluated by the all-training-all testing and leave-one-out cross-validation methods. The results have been shown in Table 2 and “-” denotes that the study did not deal with. The raw data can be feed to the two-layer GRU, the time step was 20 and dimension of input features was 13, the number of training iterations and batch size can be changed according to the practical demand. Table 3 illustrates the classification results compared with other studies according to their experiment settings. The average result can be displayed if the research had several experimental results in one classification task. We took 80% percent data as the training set, 20% as testing sets.

Table 2. Performance of diagnosing NDS (two groups).

| Evaluation methods       | Subjects   | Zeng et al.,[17] | Original LSTM | Our Method    |
|--------------------------|------------|------------------|---------------|---------------|
| All-training-all-testing | ALS vs. CO | 93.10%           | 96.36%        | <b>99.49%</b> |
|                          | PD vs. CO  | 100%             | 95.56%        | <b>100%</b>   |
|                          | HD vs. CO  | 100%             | 96.58%        | <b>100%</b>   |
|                          | NDS vs. CO | 93.75%           | 95.48%        | <b>97.37%</b> |
| Leave-one-out            | ALS vs. CO | 89.66%           | 93.28%        | <b>97.25%</b> |
|                          | PD vs. CO  | 87.10%           | 90.69%        | <b>96.80%</b> |
|                          | HD vs. CO  | 83.33%           | 87.45%        | <b>92.25%</b> |
|                          | NDS vs. CO | -                | 90.77%        | <b>95.04%</b> |

Table 3. Performance of diagnosing nds (two groups).

| Studies                      | ALS vs. CO    | PD vs. CO     | HD vs. CO     |
|------------------------------|---------------|---------------|---------------|
| Athisakthi et al.,[33]       | 56.66%        | 51.00%        | 54.28%        |
| Our Method                   | <b>90.48%</b> | <b>90.91%</b> | <b>100%</b>   |
| Banaie et al.,[16]           | 100%          | 80.00%        | 71.43%        |
| Our Method                   | <b>100%</b>   | <b>90.28%</b> | <b>91.05%</b> |
| Sánchez-Delacruz et al.,[14] | 96.13%        | 90.36%        | 88.67%        |
| Our Method                   | <b>97.25%</b> | <b>96.80%</b> | <b>92.25%</b> |
| khorasani et al.,[28]        | -             | 90.32%        | -             |
| Our Method                   | -             | <b>96.80%</b> | -             |
| Xia et al.,[34]              | -             | 96.77%        | -             |
| Our Method                   | -             | <b>96.80%</b> | -             |
| Xia et al.,[27]              | 92.86%        | -             | -             |
| Our Method                   | <b>97.25%</b> | -             | -             |

In the second experiment, there were 3 classes should be identified. The difference between three diseases has been displayed in Fig. 4. It can be seen that ALS and PD data fluctuation range was similar. We used the confusion matrix to represent the classification accuracy of the three ND groups and healthy controls. We choose the best test result that has been shown in Fig. 5. The CO class and HD class both achieved 100% accuracy which was better than other diseases.

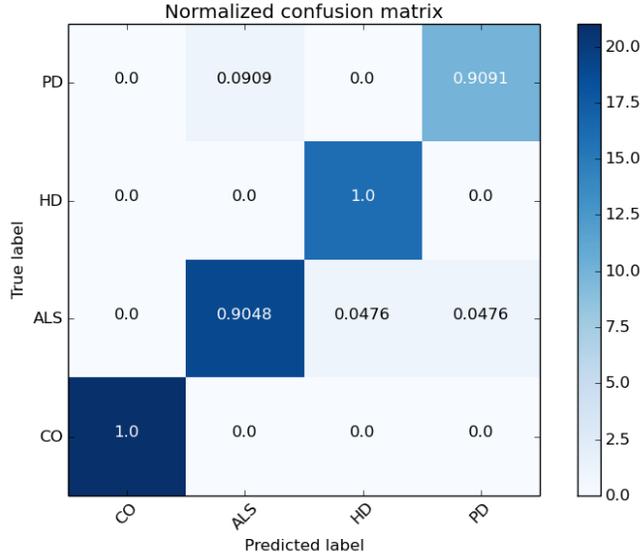


Figure 5. Confusion Matrix of Classification Accuracy.

The last experiment changed vital parameters of the GRU model that led to excellent effect superior to the ordinary LSTM model. These parameters included model layer and feature dimension in every sample. We set up some variables in advance, dropout was set to 0.8 to prevent over-fitting, the learning rate was 0.001, and the hidden layer output was 128. The number of time steps can be set up to 20. 80% of the data can be regarded as the training data, the rest was for testing. The last experiment shows the optimal values of the parameters in GRU. In every classification task, the two-layer GRU is better than the single-layer one. We extracted 12 sensor data as the training data that excluding elapsed time column, which got worse results than the data of 13 dimensions (all dimensions including stride interval, swing interval, stance interval and double support interval). The experimental results have been shown respectively in Table 4.

Table 4. Results for classification of NDS patients and healthy controls with different parameters.

| Dataset    | Dimension |               | Layer  |               |        |
|------------|-----------|---------------|--------|---------------|--------|
|            | 12        | 13            | 1      | 2             | 3      |
| ALS vs. CO | 93.45%    | <b>97.25%</b> | 78.29% | <b>97.25%</b> | 94.23% |
| PD vs. CO  | 92.96%    | <b>96.80%</b> | 82.37% | <b>96.80%</b> | 95.33% |
| HD vs. CO  | 88.72%    | <b>92.25%</b> | 79.33% | <b>92.25%</b> | 89.02% |
| NDS vs. CO | 90.20%    | <b>95.04%</b> | 85.56% | <b>95.04%</b> | 91.35% |

## 5. CONCLUSION

In this paper, we reported our investigation of predicting the neurodegenerative diseases using state-of-the-art machine learning techniques. We proposed a two-layer LSTM to model the gait data (time interval) recorded by force-sensitive sensors. The proposed model can capture and preserve features of gait dynamics over a long period. Compared with previous machine learning based approaches for NDs diagnosis, our proposed model is better in describing the

temporal sequential gait data. Once the model is well trained, the diagnostic result can be directly reflected from sensor data, which resolves the subjectivity concerns. This benefits neurosurgeons in simplifying the diagnostic procedure and providing further personalized treatment.

Our work is a preliminary research in diagnosing NDs. Limited by the current datasets, only 3 diseases were labeled and used to train the model. In the future work, more kinds of diseases and higher prediction accuracy can be achieved by using fused features from multiple data sources, such as skeleton data, fMRI image, and biochemical data. The GRU model we proposed can also be improved to accommodate more various data types. The experiments reported in this paper have proved the superiority of the GRU in diagnosis of NDs. Our model will be a baseline and the proposed approach is expected to inspire more applications of machine learning techniques in computer assisted diagnostic systems.

## ACKNOWLEDGMENT

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] K. Gammon, Neurodegenerative disease: brain windfall., *Nature* **515** (7526) (2014) 299–300
- [2] W. Aziz, M. Arif, Complexity analysis of stride interval time series by threshold dependent symbolic entropy., *European Journal of Applied Physiology* **98** (1) (2006) 30
- [3] 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
- [4] A. Pourhedayat, Y. Sarbaz, A grey box neural network model of basal ganglia for gait signal of patients with Huntington disease., *Basic & Clinical Neuroscience* **7** (2) (2016) 107
- [5] W.C. Koller, J. Trimble, The gait abnormality of Huntington’s disease, *Neurology* **35** (10) (1985) 1450–4
- [6] 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., *Journal of Applied Physiology* **82** (1) (1997) 262
- [7] A. Yarnall, N. Archibald, D. Burn, Parkinson’s disease, *Medicine* **40** (10) (2012) 529–535
- [8] B.R.B. M. PhD, J.M. Hausdorff, J.E. Visser, M.D. Nir Giladi, Falls and freezing of gait in Parkinson’s disease: A review of two interconnected, episodic phenomena, *Movement Disorders* **19** (8) (2004) 871–84
- [9] M. Panisset, Freezing of gait in Parkinson’s disease, *Journal of Neurology* **253** Suppl 7 (7) (2006) 255-260
- [10] A. Amini, Maghsoud Bigy, K. Banitsas, A. Badii, J. Cosmas, Recognition of postures and freezing of gait in Parkinson’s disease patients using microsoft kinect sensor, in: *International IEEE/EMBS Conference on Neural Engineering*, (2015), pp. 731-734
- [11] J.M. Hausdorff, N.B. Alexander, *Gait disorders: evaluation and management*, CRC Press
- [12] E.R. Kandel, J.H. Schwartz, T.M. Jessell, *Principles of neural science*, fourth edition, Medical
- [13] J.M. Hausdorff, A. Lertratanakul, M.E. Cudkowicz, A.L. Peterson, D. Kaliton, A.L. Goldberger, Dynamic markers of altered gait rhythm in amyotrophic lateral sclerosis., *Journal of Applied Physiology* **88** (6) (2000) 2045-2053
- [14] Sánchez-Delacruz, F. Acosta-Escalante, M.A. Wister, J.A. Hernandez-Nolasco, P. Pancardo, J.J. MndezCastillo, Gait recognition in the classification of neurodegenerative diseases **8867** (2) (2014) 128–135
- [15] L. Sugavanewaran, K. Umopathy, S. Krishnan, Ambiguity domain-based identification of altered gait pattern in als disorder., *J Neural Eng.* **9** (4) (2012) 046004
- [16] 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 Systems with Applications* **38** (6) (2011) 7359–7363

- [17] W. Zeng, C. Wang, Classification of neurodegenerative diseases using gait dynamics via deterministic learning, *Information Sciences An International Journal* **317** (C) (2015) 246–258
- [18] A. Haque, A. Alahi, L. Fei-Fei, Recurrent attention models for depth-based person identification
- [19] 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
- [20] L. Wu, C. Shen, A.V.D. Hengel, Convolutional lstm networks for video-based person re-identification, arXiv preprint arXiv:1606.01609.
- [21] 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, arXiv preprint arXiv:1603.07772
- [22] R.R. Variator, 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
- [23] S. Sharma, R. Kiros, R. Salakhutdinov, Action recognition using visual attention, arXiv preprint arXiv:1511.04119
- [24] N.Y. Hammerla, S. Halloran, T. Ploetz, Deep, convolutional, and recurrent models for human activity recognition using wearables, arXiv preprint arXiv:1604.08880
- [25] C.-W. Cho, W.-H. Chao, S.-H. Lin, Y.-Y. Chen, A vision-based analysis system for gait recognition in patients with Parkinson's disease, *Expert Systems with applications* **36** (3) (2009) 7033–7039
- [26] R. Das, A comparison of multiple classification methods for diagnosis of Parkinson disease, *Expert Systems with Applications* **37** (2) (2010) 1568–1572
- [27] Y. Xia, Q. Gao, Y. Lu, Q. Ye, A novel approach for analysis of altered gait variability in amyotrophic lateral sclerosis, *Medical & Biological Engineering & Computing* **54** (9) (2016) 1–10
- [28] A. Khorasani, M.R. Daliri, Hmm for classification of Parkinson's disease based on the raw gait data, *Journal of medical systems* **38** (12) (2014) 1–6
- [29] S. Hochreiter, J. Schmidhuber, Long short-term memory, *Neural computation* **9** (8) (1997) 1735–1780
- [30] K. Xu, J. Ba, R. Kiros, K. Cho, A. Courville, R. Salakhutdinov, R. S. Zemel, Y. Bengio, Show, attend and tell: Neural image caption generation with visual attention, arXiv preprint arXiv:1502.03044 2 (3) (2015) 5
- [31] N. Kalchbrenner, I. Danihelka, A. Graves, Grid long short-term memory, arXiv preprint arXiv:1507.01526
- [32] NIGMS, physionet database, <https://physionet.org/physiobank/database/gaitnidd/>
- [33] A. Athisakthi, D.M. Pushparani, Detection of Movement Disorders Using Multi SVM, *Society of Photo-Optical Instrumentation Engineers*, (2013)
- [34] X. Yi, Y. Qiang, Q. Gao, Y. Lu, D. Zhang, Symmetry analysis of gait between left and right limb using cross fuzzy entropy, *Computational & Mathematical Methods in Medicine* **2016** (1) (2016) 1737953