A Hybrid Spatio-temporal Model for Detection and Severity Rating of Parkinson's Disease from Gait Data

Aite Zhao[a], Lin Qi[a,1], Jie Li[a], Junyu Dong[a,*], Hui Yu[b]

[a]*Department of Computer Science and Technology, Ocean University of China, Qingdao, China*
[b]*University of Portsmouth, Portsmouth, UK*

Abstract

When diagnosing Parkinson's disease (PD), medical specialists normally assess several clinical manifestations of the PD patient and rate a severity level according to established criteria. This rating process is highly depended by doctors' expertise, which is subjective and inefficient. In this paper, we propose a machine learning based method to automatically rate the PD severity from gait information, in particular, the sequential data of Vertical Ground Reaction Force (VGRF) recorded by foot sensors. We developed a two-channel model that combines Long Short-Term Memory (LSTM) and Convolutional Neural Network (CNN) to learn the spatio-temporal patterns behind the gait data. The model was trained and tested on three public VGRF datasets. Our proposed method outperforms existing ones in terms of prediction accuracy of PD severity levels. We believe the quantitative evaluation provided by our method will benefit clinical diagnosis of Parkinson's disease.

Keywords: Parkinson's disease, diagnosis, gait, temporal data, LSTM, CNN.

*Corresponding author
Email addresses: `zhaoaite@stu.ouc.edu.cn` (Aite Zhao), `qilin@ouc.edu.cn` (Lin Qi), `lijie@stu.ouc.edu.cn` (Jie Li), `dongjunyu@ouc.edu.cn` (Junyu Dong), `hui.yu@port.ac.uk` (Hui Yu)
[1]The author contributed equally to this work and should be considered co-first author

1. Introduction

Parkinson's disease (PD) is a degenerative brain disorder characterized by a loss of midbrain dopamine (DA) neurons [1]. It affects mainly elderly people, causing movement problems such as static tremors, rigidity, bradykinesia, gait disturbance, and postural instability [2]. As apart of these motor symptoms, gait disturbance occurs in early stages and shows obvious manifestations. Some gait disturbances such as festinating gait, short gait, and freezing gait have been used in literature to identify a prognosis of Parkinson's disease by characterizing gait analysis [3–5].

When assessing the severity level of Parkinson disease, numeric scales are preferred. The Hoehn & Yahr scale (H & Y scale) was widely adopted, which consisted of 5 stages originally and was further extended with additional stage 1.5 and 2.5 [6]. The Unified Parkinsons Disease Rating Scale (UPDRS) is more complex and consists of more levels [7]. When medical specialists employ these scales to rate the PD severity, subjectivity and low efficiency are inevitable as most of the diagnostic criteria use descriptive symptoms, which cannot provide a quantified diagnostic basis. Therefore, the development of computer-assisted diagnosis and computer-expert system is very important.

Modern computer-assisted diagnosis, such as automatic medical image processing and large scale medical data analysis, has been widely used in corresponding medical fields, where machine learning plays a core role in these systems.

For PD detection from gait data, machine learning methods, such as kernel Fisher discriminant, naïve Bayesian, and support vector machine, have been employed and achieved promising results [8–15]. However, these approaches only deal with it as a two-category classification problem, i.e. detecting PD from gait information, whereas the severity grading, which requires a finer investigation, has not been well studied.

In addition, the machine learning methods used in literature are not specifically designed to deal with temporal sequential data, whereas the gait data

2

captured by sensory devices (cameras, force sensors) contains important temporal information that is critical for PD diagnosis. As one of the popular deep learning model, Long Short-Term Memory (LSTM) has recently been used in various fields, including action recognition and gait recognition [16–22] as it is good at handling time series with long intervals and multi-classification problems. Besides, the Convolution Neural Networks (CNN) can automatically learn commendable features from the given gait data, which considers spatial proximity using a convolution operation, significantly improves the classification accuracy.

In this paper, we propose a hybrid model that fuses the Long Short-Term Memory and Convolutional Neural Network to estimate the PD severity level from gait data. The model was trained and tested using public datasets of vertical ground reaction force recorded by force sensors. The model well captures the dynamic characteristics of PD patients' gait and the experimental results show that the efficiency and accuracy of the model is better than existing approaches.

The paper is organized as follows. Section 2 describes related research on computer-assisted PD diagnosis. Section 3 introduces the proposed two-channel model, and Section 4 depicts the gait dataset. The experiments are described in Section 5. We conclude the paper and discuss possible future work in Section 6.

2. Related Work

Huge amounts of medical data have been being produced and collected, from which researchers have built modern computer-assisted medical diagnostic systems with machine learning methods.

Choi *et al.* developed a Doctor AI using the Recurrent Neural Networks (RNN), which could predict the diagnosis and medication categories for the subsequent visits of a patient [23]. The dataset contained a longitudinal time-stamped electronic health record data from more than 250,000 patients over 8 years. They showed that their model could be implemented in another insti-

tution with high prediction accuracy. Lipton *et al.* applied LSTM to analyze clinical medical data, including 13 observations (such as diastolic and systolic blood pressures, peripheral capillary refill rate, and heart rate) collected from the Intensive Care Unit (ICU) [24]. This was a multi-label task with 128 diagnostic (labels) classes in the experiment, which included a wide range of acute conditions, such as acute respiratory distress, congestive heart failure, and sepsis. With the same data, they also utilized RNN to deal with missing data in sequential inputs and improved multi-label classification of a given clinical time series [25].

In diagnosing PD, early studies used the PD gait sensor data for binary classification (PD or non-PD). Cho *et al.* presented a vision-based analysis system for PD detection from motion (walking) video clips of 7 PD participants and 7 normal participants [26]. They adopted principal component analysis (PCA) and linear discriminant analysis (LDA), and achieved 95.49% accuracy for the 2-class classification. However, the few number of participants limits the generalization of the method.

Das used a dataset that consists of speech signals from 31 people, 23 of which were PD patients, and compared different methods for PD detection, including neural networks, DMneural, regression, and decision tree [27]. Manap *et al.* proposed to use the feed-forward multilayer perceptron neural network in distinguishing PD patients from normal people with gait patterns extracted from Ground Reaction Force (GRF) recordings[28]. Sarbaz *et al.* designed a feed-forward artificial perceptron neural network with a hidden layer as the classifier that takes frequency features extracted from gait signals (time intervals collected by force sensor) [29]. Khorasani *et al.* adopted hidden Markov model (HMM) and least squares support vector machine (LS-SVM) to detect PD from 16 healthy subjects and 15 PD subjects[30]. Other works distinguish between PD patients and healthy people from GRF gait data by employing SVM and PCA.[31–34].

Existing approaches mainly considers PD detection, which is a two-category classification problem. Little work has been carried out in predicting PD severity

4

levels, which involves more complex multi-category information. Hammerla *et al.* proposed two consecutive restricted Boltzmann machines (RBMs) to detect PD states such as "on time" and "off time" (motor symptoms of patients were both well controlled or not well controlled) [35]. The dataset they used consists of accelerometer records from 34 PD patients including approx. 5,500 hours of movement data. Our work is most related with those proposed by Jane *et al.*, in which they employed Q-backpropagated time delay neural network to predict gait disturbances [36]. The gait data they used is the Vertical Ground Reaction Force (VGRF) collected by force sensors.

Among popular deep learning models, LSTM has shown its advantages in dealing with temporal data, which can be used for learning gait features of PD patients. Moreover, CNN shows superiority in extracting spatial features from multi data source (RGB image, Gait Energy Image) in gait recognition field, In this paper, we combine the two state-of-the-art models in the PD detection and severity prediction. In the future, we will also embed our model into robots, and use the hardware such as robotic arm and teleoperation system [37, 38] to build a complete robotic disease diagnosis system.

2.1. Brief Introduction to RNN and LSTM

To be self-explanatory, the following briefly introduces basic theory of RNN and LSTM, after which the architecture of our proposed model is described.

RNN is a special artificial neural network with node-oriented connections. This internal state of the network can represent dynamic timing behavior. Unlike feed-forward neural networks, the internal memory in RNN makes it naturally good at tasks involving sequential data, such as non-segmented handwriting recognition, speech recognition and action recognition. The RNN is a chainlike neural network comprised of repeating modules that allows for information retention by combining previous states with current input [39]. In a standard RNN, this repetitive module has a very simple structure, which is a "*tanh*" function. For a given input series x_t $(t = 1, 2..., T)$, the hidden state of a recurrent module h_t is calculated using Eq.1. The output of the module y_t is
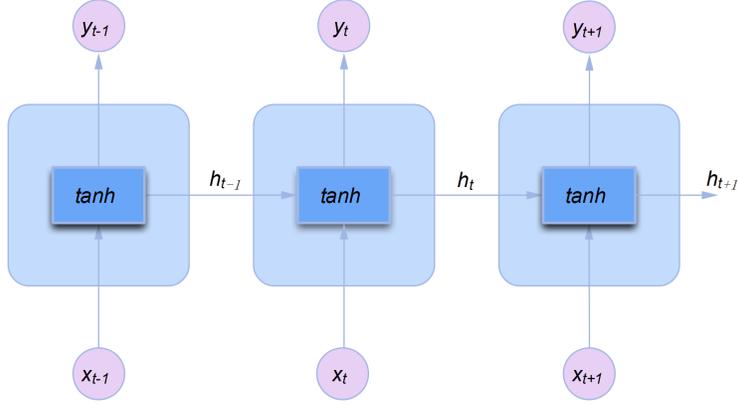
Figure 1: A RNN expand model. Every RNN block represents a RNN cell with two inputs, two outputs and an activation function.

calculated as in Eq.2. Fig. 1 shows a typical structure of RNN with 3 modules.

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

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

where W_{xh}, W_{hh}, W_{ho} denote 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 $softmax$ 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 simplicity. Hochreiter & Schmidhuber proposed a new model called LSTM (Long Short-Term Memory) to resolve this problem [40]. LSTM is an extension of RNN, which can learn long-term dependency information from the input data and has been successfully applied in various fields. The repetitive module in LSTM has a different structure from RNN in that there are four interactive operations (3 *sigmoid* and 1 *tanh*), which enables stronger memory ability. Like RNN, the structure of LSTM is flexible and can be modified ,leading to many variants [41, 42].

The basic cell of LSTM has three gates (input, forget and output). Every gate has a sigmoid activation function and a pointwise multiplication operation. Fig. 2 shows the structure of a basic cell and illustrates the operations of the gates. The basic cell of the LSTM is defined with the following equations:

$$\begin{pmatrix} i_t \\ f_t \\ o_t \\ \tilde{c}_t \end{pmatrix} = \begin{pmatrix} \sigma \\ \sigma \\ \sigma \\ \tanh \end{pmatrix} \left(M \begin{pmatrix} x_t \\ h_{t-1} \end{pmatrix} \right) \tag{3}$$

$$c_t = i_t \odot \tilde{c}_t + f_t \odot c_{t-1} \tag{4}$$

$$h_t = o_t \odot \tanh(c_t) \tag{5}$$

where x_t denotes the input at time t, $\tilde{c}_t$ is the processed intermediate state, σ denotes the sigmoid activation function, M is an affine transformation, i_t and f_t denote the input gate and forget gate respectively. The input gate determines the input information, and the forget gate is responsible for updating the memory at time t. Previous state of the memory (c_{t-1}) and the intermediate state $\tilde{c}_t$ jointly decide the current state (c_t) after the calculation of the two gates. i_t, c_t and the output o_t determine the hidden state (h_t) which provides information from previous time.

2.2. Two-Channel model for PD Detection and Severity Rating

In this paper, we propose a two-channel model which consists of a 5-layer Convolutional Neural Network (CNN) and a 2-layer LSTM for PD detection and severity rating to capture the spatio-temporal information of the gait data of PD patients.

The LSTM is two-layer, where each layer has the same number of cells. We observed that the gait cycle of most participants is about 1 *second*, whereas the gait data was sampled by every 10 *ms*. Therefore, the number of cells in each layer is set to 100, while each 100 successive pieces of gait data were fed to the
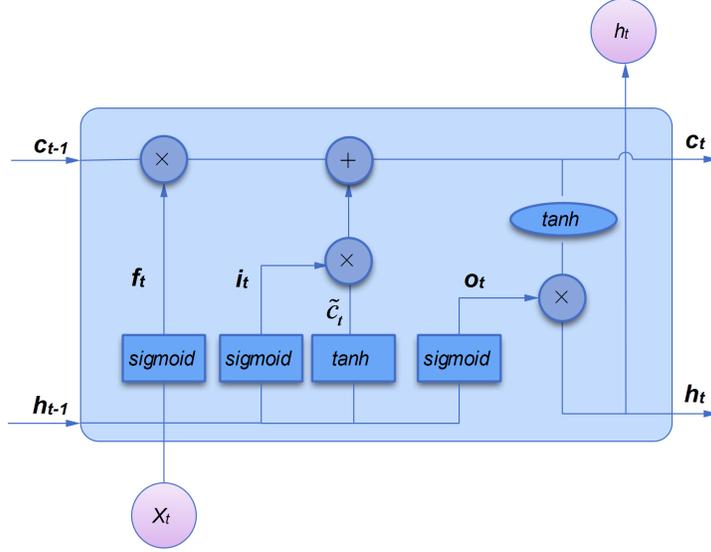
Figure 2: The illustration of one LSTM cell. Unlike RNN cell, it adds a state input (c_{t-1}) and a state output (c_t). c_t combines the state information from c_{t-1} and $\tilde{c}_t$, which affects the hidden layer output (h_t).

LSTM. The LSTM cell outputs a feature vector for predicting the probability to another softmax classifier. Fig. 3 depicts the structure of the two-layer LSTM.

The CNN channel includes 2 convolutional layers, 2 max pooling layers and a fully connected layer. The convolution kernel in the two convolutional layers are both 5×5, where the first convolutional layer outputs 32 feature maps and the second convolutional layer outputs 64 feature maps. The fully connected layer outputs a 1024-dimensional feature vector to a softmax classifier for PD detection and severity level prediction.

The above CNN and LSTM were combined for joint training and the weighted average of the softmax gives the final classification results.

We input the $N \times f$-dimensional feature vectors (every training sample) into the two channels of the model respectively. Every time M training samples can be sent into the model ($N \times M \times f$ dimensions).

The CNN model convolute and down-sample these training data to extract 1024-dimensional features for each sample. On the other hand, the LSTM model
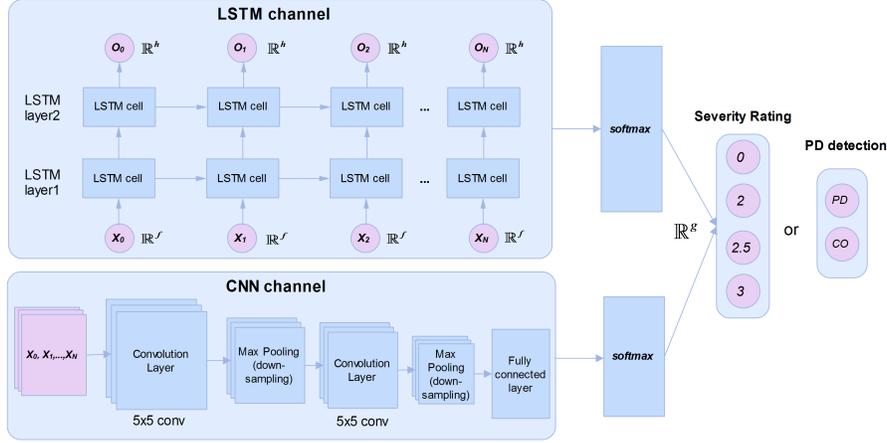
Figure 3: The structure of proposed two-layer LSTM and CNN model. The input of the LSTM is $X_N, X_N \in \mathbb{R}^f$ ($f = 19$), which from different time periods are entered into N ($N = 100$) LSTM cell. The output of the last time step (Nth LSTM cell) is used for feature classification.

can be expanded according to time steps and every sample in one training sample can be input to one LSTM cell respectively. Every LSTM cell takes $M \times f$ data as the input. The output of each LSTM cell is an h-dimensional feature vector (hidden layer output), it 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 from all of the above nodes, so we take only the last output ($O_{100} \in \mathbb{R}^h$) as the basis for the classification.

The outputs of the two channels will be transformed into g-dimensional(g classes) vector 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 detect whether participant with PD or rate the severity level of the PD patient.

3. Datasets

In this study, we have utilized gait signals from PhysioNet [43]. The database consists of three PD gait sub-datasets, which are contributed by three re-

searchers (Ga[44], Ju[45] and Si[46]).

The dataset includes gait information from 93 patients with the idiopathic PD and 73 healthy controls (average age 66.3, 55% male). Every participant was asked to walk in their usual, self-selected pace for about two minutes while wearing a pair of shoes with 8 force sensors located under each foot. The sensors measure Vertical Ground Reaction Force (VGRF, in Newton) as a function of time at 100 samples per second.

With information from VGRF, one can investigate the force record as a function of time and location, derive measures that reflect the center-of-pressure as a function of time, and determine stride time or swing time for each foot changing time. Accordingly, one can study the stride-to-stride dynamics and the variability of these time series.

There were 16 sensor outputs and two total outputs for each foot in each sample. When the subjects were standing still, the 16 sensors would have an initial value that changes when the subjects start walking. Every sensor had 19 parameters that corresponded to each row in the dataset file. It includes the time stamp, 16 VGRF foot parameters and 2 total VGRF data.

The dataset had also explained the specific situation of each participant including gender, age, height, weight, and severity level of PD. PD Severity level was graded according to two scales (H&Y, UPDRS). Table 1 describes the main contents about the participants and the specific number of them counted by H&Y rating scale has been recorded in Table 2.

4. Experiments

This section illustrates the main process of the experiment showing experimental results. Experiments were implemented using Tensorflow library and tested on an Intel Core i5 computer with 31.3 GB RAM. We compared our experimental results with those of other researchers who utilized the same data and other modern machine learning models. For the pre-process of the gait data, we took the whole data set reshaped as $100 \times 19 \times N$ (N="Ga:13592,Si:7744,Ju:11734"),

Table 1: The basic information of the subjects in the three datasets.

Study	Subjects	Total Subjects	Male	Female	Age (<50)	Age (50 ~70)	Age (>=70)
Ga[44]	PD patients	29	20	9	0	13	16
	Healthy controls	18	10	8	0	10	8
Ju[45]	PD patients	29	16	13	1	18	10
	Healthy controls	26	12	14	0	20	6
Si[46]	PD patients	35	22	13	0	28	7
	Healthy controls	29	18	11	1	26	2

Table 2: The number of people in three datasets divided by severity.

Study	Severity 0(Healthy)	Severity 2	Severity 2.5	Severity 3
Ga[44]	18	15	8	6
Ju[45]	26	12	13	4
Si[46]	29	29	6	0

each row had 100×19 dimensions. After an L2 normalization, the above mentioned raw gait data was fed into the prepared network. Some related experimental settings will be explained in detail.

We conducted three experiments: one used the dataset that was divided into two classes, PD patients and normal people for PD detection; the second used experimental subjects that were divided into multi-classes in accordance with the H & Y scale. According to the dataset information, the three datasets have been classified as 0 (healthy control), 2, 2.5, and 3 rating levels, and the above two types of diagnoses covered comparisons of various machine learning models. Finally, we transformed core parameters of LSTM to reach the best results of classification. The specific experimental process will be described adequately in the following part.

Furthermore, for the experimental settings, the dropout was set to 0.5 to prevent over-fitting with the learning rate of 0.001 in the combined models with a fixed batch size 100. We determined the output dimensions of the two models that were 128 in LSTM and 1024 in CNN. The training iterations were different

Table 3: Experimental results on 3 datasets for classification of PD patients and healthy controls using different models.

						Accuracy on different methods					
Dataset	Perumal *et al.*,[31]	NB	KNN	LR	DT	RF	SVM	GBDT	CNN	LSTM	CNN+LSTM
Ga[44]	92.25%	68.04%	94.19%	73.67%	88.53%	93.27%	67.78%	91.95%	93.44%	*98.63%*	*98.70%*
Ju[45]	92.50%	82.25%	95.15%	82.25%	90.29%	94.25%	82.25%	92.00%	92.34%	*98.38%*	*98.41%*
Si[46]	90.00%	72.82%	92.51%	78.05%	85.93%	91.48%	53.13%	89.80%	95.68%	*98.84%*	*98.88%*

Table 4: Experimental results on all data from the three datasets for classification of PD patients and healthy controls using different models.

		Total Accuracy(All data from the three datasets)					
Ertuğrul *et al.*,[34]	Wu *et al.*,[33]	Medeiros *et al.*,[32]	KNN	RF	CNN	LSTM	CNN+LSTM
88.89%	84.48%	81.00%	94.70%	93.85%	95.23%	*98.52%*	*98.61%*

in the three experiments in order to achieve the best result, which was set up to 300000,500000 and 500000 respectively. Finally, the accuracy rates of the two channels were summed up with the weight of 0.5.

In the first experiment, there were two classes can be classified that contained 93 PD patients and 73 healthy controls. Every subject contained between a thousands and more than 10 thousand for each time period (within two minutes). The gait data was fed into the LSTM with 100 time steps and 19-dimensional input features, with the number of training iterations and batch size can be set according to the practical demand of what was used when the model was trained. In this part, PD patients and healthy controls were well identified. The results from the first experiment are shown in Table 3 and Table 4.

Others' work on diagnosing PD employed different data from the three datasets. In Table 3, all of the methods were performed on the three data sets, respectively, with the best results 92.5% put forward by Perumal *et al.* Some researchers also merged the three data sets into a single set for their training, shown in Table 4. In the two conditions, our work has achieved a higher accuracy that proved the high-performance of LSTM in this research.

In addition, PD severity rating and detection were completed using different machine learning methods and the accuracy was evaluated and com-

pared with the proposed LSTM method. We adopted the open-source Python-sklearn library to implement these methods including naive Bayesian (NB), k-nearest neighbor (KNN), logistic regression (LR), random forest (RF), decision tree (DT), support vector machine (SVM), and gradient boosting decision tree (GBDT). By comparison, the performance of LSTM to the original data out-performed these methods. After verification, 100 samples (100 time steps) were considered as a training sample as well as the testing sample. 80% of the data can be regarded as the training data; the rest used for testing. It can be seen that KNN and RF achieved a higher accuracy. Thus, Table 4 has removed the method of poor performance. However, these machine learning models do not deal with time series well, and the classification results of these models cannot compare with those of LSTM+CNN.
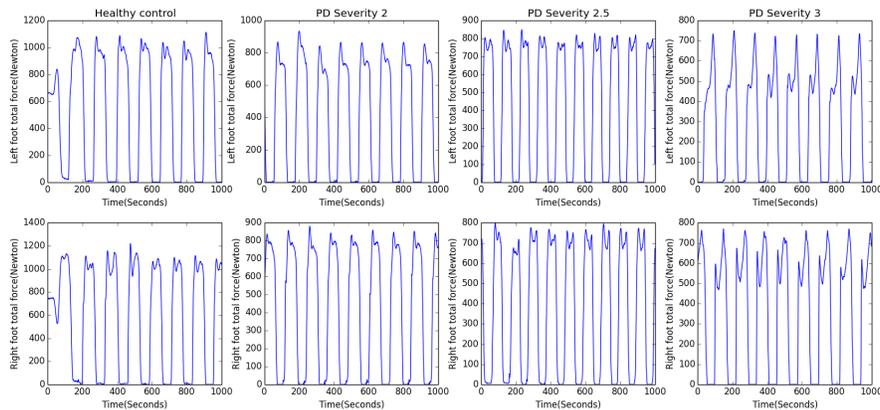

Figure 4: Comparison of the performance of subjects for different severity. And the total force data on the left and right feet for 10 seconds has been displayed. The fluctuation of data is very different.

In the second experiment, we divided the data from Ga and Ju's studies into 4 classes (data from Si's study only had 3 severity levels, leading to 3 classes) in accordance with PD severity levels. Zero to ten seconds of total VGRF data from the left and right feet are shown in Fig. 4 and illustrates that patients in each severity level have different performances. The first row represents the

13

left foot data and the second row represents the right, and the four columns denote PD severity levels. It can be seen that the total force of healthy people is slightly larger than the PD patients, and the data amplitude of PD patients gradually decreases with the aggravation of the disease, which means that the healthy control's foot VGRF is greater than that of patients with PD can be demonstrated. With this information, we have trained the two-channel model for 4 rating levels, showing the experimental results in Table 5.

Table 5: Experimental results on 3 datasets for classification of PD patients with different severity and healthy controls using different models.

	Accuracy on different methods										
Dataset	Jane *et al.*,[36]	NB	KNN	LR	DT	RF	SVM	GBDT	CNN	LSTM	CNN+LSTM
Ga[44]	93.10%	36.85%	89.48%	58.62%	77.64%	87.68%	35.23%	86.94%	96.36%	*97.46%*	*97.48%*
Ju[45]	89.66%	41.51%	89.36%	68.92%	81.78%	90.55%	41.29%	91.06%	96.67%	*97.83%*	*97.86%*
Si[46]	91.67%	65.78%	87.02%	69.40%	78.11%	85.73%	44.42%	84.31%	96.57%	*98.77%*	*98.80%*

According to the results, we made a a confusion matrix to visually show the accuracy of predicting rating levels. As we can see from Fig. 5, the classification accuracy for level 2 and level 2.5 is lower than other levels, indicating that there is little difference for the two severity levels in gait features and that our model is accurate for a PD severity rating. We exploited a 10-fold-cross-validation to evaluate the use 10% of data for testing, with the rest was for training. In this experiment, the two-channel model also performed well.

In the last experiment, we changed vital parameters of the LSTM model that was superior to the basic LSTM model. These parameters included model layer and feature dimension for every sample. We set up some variables in advance in order to get the best results. 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 can be set to 100, since every 100 samples are obtained in 1 second (100Hz). We showed the optimal values of the parameters in Table 6 and Table 7. For the two classification tasks above, the LSTM with two layers was preferable to that with other layers. We therefore chose the 16 VGRF sensor data as the training data, excluding the time stamp column and
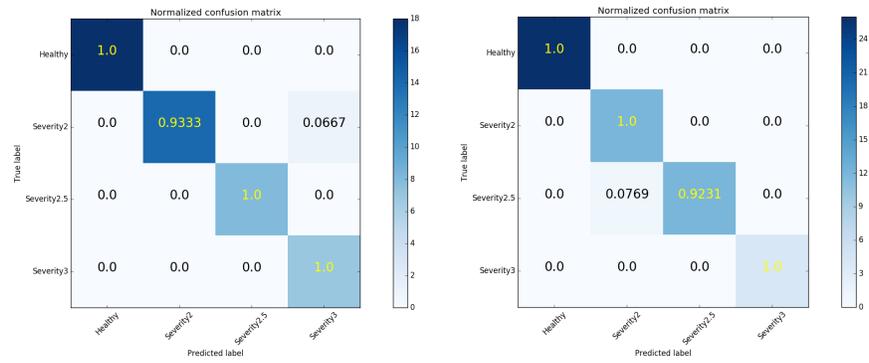
(a) The classification results in Ga's dataset. (b) The classification results in Ju's dataset.
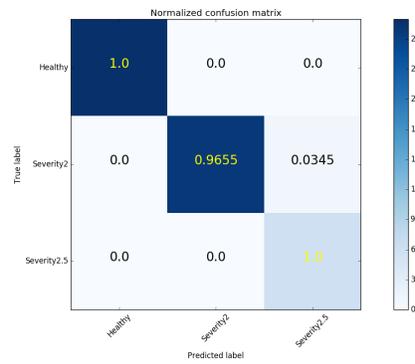
(c) The classification results in Si's dataset.

Figure 5: Diagnostic accuracy on 3 datasets for classification of PD patients with different severity and healthy controls.

Table 6: Experimental results on 3 datasets for classification of PD patients and healthy controls with different parameters.

Dataset	Dimension		Layer		
	16	19	1	2	3
Ga[44]	98.46%	**98.70%**	97.04%	**98.63%**	97.40%
Ju[45]	97.96%	**98.38%**	97.37%	**98.38%**	97.78%
Si[46]	98.69%	**98.84%**	97.33%	**98.84%**	98.25%

Table 7: Experimental results on 3 datasets for classification of PD patients with different severity and healthy controls with different parameters.

Dataset	Dimension		Layer		
	16	19	1	2	3
Ga[44]	97.45%	**97.46%**	95.67%	**97.46%**	96.87%
Ju[45]	97.23%	**97.83%**	95.87%	**97.83%**	96.88%
Si[46]	98.13%	**98.77%**	96.98%	**98.77%**	98.03%

two total VGRF columns, which got inadequate results when compared with 19-dimensional data.

5. Conclusion

In this paper, we reported our investigation of predicting the severity of Parkinson's disease from the gait data (Vertical Ground Reaction Force) recorded by foot sensors. We employed deep learning architectures and proposed a two-channel network including LSTM and CNN to model the gait data, which captures the spatio-temporal characteristics of gait dynamics over a period of time. In contrast to previous machine learning based approaches which only provide "binary" detection, our model can perform multi-category classification, i.e. predicting the severity level of PD. Once the model is well trained, the predicted severity can be directly obtained from sensor data, which resolves any subjectivity concerns. This benefits neurosurgeons as it simplifies the diagnosis

procedure and provides further personalized treatment.

Our work is a preliminary investigation of predicting PD severity. Limited by the current datasets, only four levels of severity were labeled and used to train the model. Collecting VGRF gait data from patients of high severity level is difficult because they can hardly walk. In the future work, more severity levels 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 LSTM model we proposed can also be improved to accommodate more data types. The experiments reported in this paper have shown the superiority of the LSTM in predicting PD severity. Our model will be a baseline and the proposed approach is expected to inspire more applications of machine learning techniques in computer-assisted medical diagnosis.

6. Acknowledge

References

[1] A. Yarnall, N. Archibald, D. Burn, Parkinson's disease, Medicine 40 (7) (2012) 529–535.

[2] A. Samii, J. G. Nutt, B. R. Ransom, Parkinson's disease., Lancet 363 (17) (2004) 896–912.

[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 huntingtons disease, Journal of Applied Physiology 82 (8) (1997) 262–269.

[4] J. M. Hausdorff, N. B. Alexander, Gait disorders : evaluation and management, Crc Press 1 (8) (2005) 228–236.

[5] A. Sam&Agrave, C. Angulo, D. Pardo, A. Catal&Agrave, J. Cabestany, Analyzing human gait and posture by combining feature selection and kernel methods, Neurocomputing 74 (16) (2014) 2665–2674.

[6] C. G. Goetz, W. Poewe, O. Rascol, C. Sampaio, G. T. Stebbins, C. Counsell, N. Giladi, R. G. Holloway, C. G. Moore, G. K. Wenning, Movement disorder society task force report on the hoehn and yahr staging scale: Status and recommendations, Movement Disorders 19 (9) (2004) 1020–1028.

[7] V. R. Preedy, R. R. Watson, Unified parkinsons disease rating scale, Movement Disorders 21 (8) (2006) 1200–1207.

[8] F. Wahid, R. K. Begg, C. J. Hass, S. Halgamuge, Classification of parkinson's disease gait using spatial-temporal gait features, IEEE Journal of Biomedical & Health Informatics 19 (7) (2015) 1794–1802.

[9] J. Hannink, T. Kautz, C. F. Pasluosta, J. Barth, S. Schülein, K.-G. Gaßmann, J. Klucken, B. M. Eskofier, Stride length estimation with deep learning, arXiv preprint arXiv:1609.03321 1 (9).

[10] T. Khan, J. Westin, Motion cues analysis for parkinson gait recognition, in: 15th International Congress of Parkinson's Disease and Movement Disorders, Toronto, Canada, 5-9 juni, 2011, 2011, pp. 1–2.

[11] J. D. A. Paredes, B. Muñoz, W. Agredo, Y. Ariza-Araújo, J. L. Orozco, A. Navarro, A reliability assessment software using kinect to complement the clinical evaluation of parkinson's disease, in: 2015 37th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), IEEE, 2015, pp. 6860–6863.

[12] A. P. Rocha, H. Choupina, J. M. Fernandes, M. J. Rosas, R. Vaz, J. P. S. Cunha, Kinect v2 based system for parkinson's disease assessment, in: 2015 37th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), IEEE, 2015, pp. 1279–1282.

[13] A. P. Rocha, H. Choupina, J. M. Fernandes, M. J. Rosas, R. Vaz, J. P. S. Cunha, Parkinson's disease assessment based on gait analysis using an innovative rgb-d camera system, in: 2014 36th Annual International Conference of the IEEE Engineering in Medicine and Biology Society, IEEE, 2014, pp. 3126–3129.

[14] H. L. Chen, G. Wang, C. Ma, Z. N. Cai, W. B. Liu, S. J. Wang, An efficient hybrid kernel extreme learning machine approach for early diagnosis of parkinson's disease, Neurocomputing 184 (C) (2016) 131–144.

[15] D. Rodríguez-Martín, A. Samà, C. Pérez-López, J. Cabestany, A. Català, A. Rodríguez-Molinero, Posture transition identification on pd patients through a svm-based technique and a single waist-worn accelerometer, Neurocomputing 164 (C) (2015) 144–153.

[16] A. Haque, A. Alahi, F. F. Li, Recurrent attention models for depth-based person identification, in: Computer Vision and Pattern Recognition, 2016, pp. 1229–1238.

[17] 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.

[18] L. Wu, C. Shen, A. v. d. Hengel, Convolutional lstm networks for video-based person re-identification, arXiv preprint arXiv:1606.01609 1 (11).

[19] 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, in: Computer Vision and Pattern Recognition, 2016, pp. 3697–3703.

[20] R. R. Varior, 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.

[21] S. Sharma, R. Kiros, R. Salakhutdinov, Action recognition using visual attention, in: International Conference on Learning Representations, 2016, pp. 1–11.

[22] N. Y. Hammerla, S. Halloran, T. Ploetz, Deep, convolutional, and recurrent models for human activity recognition using wearables, Journal of Scientific Computing 61 (23) (2016) 454–476.

[23] E. Choi, M. T. Bahadori, J. Sun, Doctor ai: Predicting clinical events via recurrent neural networks, Computer Science 56 (6) (2015) 1–6.

[24] Z. C. Lipton, D. C. Kale, C. Elkan, R. Wetzell, Learning to diagnose with lstm recurrent neural networks, in: International Conference on Learning Representations (ICLR), 2015, pp. 1–18.

[25] Z. C. Lipton, D. C. Kale, R. Wetzel, Directly modeling missing data in sequences with rnns: Improved classification of clinical time series, in: Proceedings of Machine Learning for Healthcare, 2016, pp. 1–17.

[26] C.-W. Cho, W.-H. Chao, S.-H. Lin, Y.-Y. Chen, A vision-based analysis system for gait recognition in patients with parkinsons disease, Expert Systems with applications 36 (7) (2009) 7033–7039.

[27] R. Das, A comparison of multiple classification methods for diagnosis of parkinson disease, Expert Systems with Applications 37 (5) (2010) 1568–1572.

[28] 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.

[29] Y. Sarbaz, F. Towhidkhah, S. Gharibzadeh, A. Jafari, Gait spectral analysis: an easy fast quantitative method for diagnosing parkinson's disease, Journal of Mechanics in Medicine and Biology 12 (3) (2012) 529–531.

[30] A. Khorasani, M. R. Daliri, Hmm for classification of parkinsons disease based on the raw gait data, Journal of medical systems 38 (6) (2014) 1–6.

[31] S. V. Perumal, R. Sankar, Gait and tremor assessment for patients with parkinsons disease using wearable sensors, ICT Express 2 (7) (2016) 168–174.

[32] L. Medeiros, H. Almeida, L. Dias, M. Perkusich, R. Fischer, A gait analysis approach to track parkinson's disease evolution using principal component analysis, in: IEEE International Symposium on Computer-Based Medical Systems, 2016, pp. 48–53.

[33] Y. Wu, P. Chen, X. Luo, M. Wu, L. Liao, S. Yang, R. M. Rangayyan, Measuring signal fluctuations in gait rhythm time series of patients with parkinson's disease using entropy parameters, Biomedical Signal Processing & Control 31 (7) (2017) 265–271.

[34] Ömer Faruk Ertuğrul, Y. Kaya, R. Tekin, M. N. Almal, Detection of parkinson's disease by shifted one dimensional local binary patterns from gait, Expert Systems with Applications 56 (8) (2016) 156–163.

[35] N. Y. Hammerla, J. Fisher, P. Andras, L. Rochester, R. Walker, T. Plötz, Pd disease state assessment in naturalistic environments using deep learning., in: AAAI2015, 2015, pp. 1742–1748.

[36] Y. N. Jane, H. K. Nehemiah, K. Arputharaj, A q-backpropagated time delay neural network for diagnosing severity of gait disturbances in parkinsons disease, Journal of biomedical informatics 60 (18) (2016) 169–176.

[37] C. Yang, X. Wang, Z. Li, Y. Li, C. Y. Su, Teleoperation control based on combination of wave variable and neural networks, IEEE Transactions on Systems Man & Cybernetics Systems PP (99) (2016) 1–12.

[38] C. Yang, K. Huang, H. Cheng, Y. Li, C. Y. Su, Haptic identification by elm-controlled uncertain manipulator, IEEE Transactions on Systems Man & Cybernetics Systems PP (99) (2017) 1–12.

[39] A. Graves, A. Mohamed, G. Hinton, Speech recognition with deep recurrent neural networks, ICASSP 1 (2003) (2013) 6645–6649.

[40] S. Hochreiter, J. Schmidhuber, Long short-term memory, Neural computation 9 (8) (1997) 1735–1780.

[41] 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, Computer Science 2 (3) (2015) 5.

[42] N. Kalchbrenner, I. Danihelka, A. Graves, Grid long short-term memory, in: International Conference on Learning Representations (ICLR), 2016, pp. 1–15.

[43] NIGMS, physionet database, `http://physionet.org/pn3/gaitpdb/`.

[44] G. Yogev, N. Giladi, C. Peretz, S. Springer, E. S. Simon, J. M. Hausdorff, Dual tasking, gait rhythmicity, and parkinson's disease: which aspects of gait are attention demanding?, European journal of neuroscience 22 (5) (2005) 1248–1256.

[45] J. M. Hausdorff, J. Lowenthal, T. Herman, L. Gruendlinger, C. Peretz, N. Giladi, Rhythmic auditory stimulation modulates gait variability in parkinson's disease, European Journal of Neuroscience 26 (8) (2007) 2369–2375.

[46] S. Frenkel-Toledo, N. Giladi, C. Peretz, T. Herman, L. Gruendlinger, J. M. Hausdorff, Treadmill walking as an external pacemaker to improve gait rhythm and stability in parkinson's disease., Movement Disorders 20 (9) (2005) 1109–1114.