An Interpretable Deep Learning Optimized Wearable Daily Monitoring System for Parkinson's Disease Patients

Fei Su (sufei@tju.edu.cn)
Shandong Agricultural University

Min Chen
Shandong Agricultural University

Zhanfang Sun
Provincial Hospital Affiliated to Shandong First Medical University

Tao Xin
The First Affiliated Hospital of Shandong First Medical University

Degang Bu
Shandong Agricultural University

Yan Chen
Shanghai Jiahui International Hospital

Article

Keywords: Parkinson's disease, wearable sensors, daily monitoring, deep learning, visual interpretation

Posted Date: January 9th, 2023

DOI: https://doi.org/10.21203/rs.3.rs-2411539/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Additional Declarations: (Not answered)
An Interpretable Deep Learning Optimized Wearable Daily Monitoring System for Parkinson's Disease Patients

Min Chen¹, Zhanfang Sun², Tao Xin³, Degang Bu¹, Yan Chen², ⁴ *, Fei Su¹ *

1. College of Mechanical and Electronic Engineering, Shandong Agricultural University, Taian 271018, China
2. Department of Neurology, Provincial Hospital Affiliated to Shandong First Medical University, Jinan 250021, China
3. Department of Neurosurgery, The First Affiliated Hospital of Shandong First Medical University, Jinan 250013, China
4. Department of Neurology, Shanghai Jiahui International Hospital, Shanghai 200233, China
* Authors to whom any correspondence should be addressed.
Email: sufei@sdau.edu.cn and cyrz001@163.com

Abstract

Walking monitoring in the daily life of patients with Parkinson's disease (PD) is of great significance for tracking the progress of the disease. This study aims to implement an accurate, objective, and long-term walking monitoring algorithm optimized by the interpretable deep learning architecture, and explore the most representative spatiotemporal motor features. Five inertial measurement units attached to the wrist, ankle, and waist are used to collect motion data from 100 subjects during a 10-meter walking test. The raw data of each sensor are subjected to the continuous wavelet transform before training in the constructed deep learning architecture based on a 6-channel convolutional neural network (CNN). The results show that the sensor located at the waist has the best classification performance with an accuracy of 98.01% ± 0.85% and the area under the receiver operating characteristic curve (AUC) of 0.9981 ± 0.0017 under ten-fold cross-validation. The gradient-weighted class activation mapping shows that the feature points with greater contribution to PD were concentrated in the lower frequency band (0.5~3Hz) compared with healthy controls. The visual maps of 3D CNN show that only three out of six time series have a greater contribution, which is used as a basis to further optimize the model input, and the raw data processing cost is reduced by 50% while the classification performance (AUC = 0.9929 ± 0.0019) is still maintained. As far as we know, this is the first study that considers visual interpretation-based optimization of an intelligent classification model in the objective diagnosis of PD.

Keywords: Parkinson's disease, wearable sensors, daily monitoring, deep learning, visual interpretation
1. INTRODUCTION

Parkinson's disease (PD), as a neurodegenerative disease, is affecting the health of more and more people, accompanied by movement symptoms such as bradykinesia, rigidity, tremor, and gait disorders. It has become the fastest developing neurological disease in the world, which poses a great challenge to the health of the elderly population due to the prevalence rate increased significantly with age. Objective assessment of the clinical progress of PD plays an important role in formulating treatment strategies for patients. However, there are some limitations in the way of evaluating the motor performance of patients through the Unified Parkinson’s Disease Rating Scale (UPDRS) in clinical practice: (1) the strained medical resources due to the need for specialized neurologists; (2) the subjectivity of clinical scoring and patient response; (3) the fluctuating patient symptoms; (4) the unsustainability of the assessment. The wearable sensor technology is becoming popular in the medical field, which can provide continuous objective longitudinal data, emerging as an effective approach for detecting and evaluating the severity of disease.

Almost all relevant studies using wearable sensors to objectively assess PD patients are implemented based on machine learning (ML) and deep learning (DL) algorithms. For ML algorithms, Talitckii et al. revealed 3 of the most effective exercises to assist doctors in diagnosing PD from 15 common motor tasks. Chen et al. constructed the genetic algorithm optimized random forest classification model by collecting 3 upper limb tasks in UPDRS-III, which can help clinicians identify subtle changes in the motor performance of PD patients. However, ML based on hand-crafted features is very time-consuming, and features determined by experience may not comprehensively describe motion details of complex disease progression. Comparatively, DL technology that can automatically learn the spatial hierarchy of features from low to high levels has gradually become the current development trend. Vásquez-Correa et al. proposed a DL architecture that integrates multimodal information from speech, handwriting, and gait to evaluate patients at different disease stages. The above studies significantly improved the objective diagnosis of PD, but at the same time require patients to perform a series of tedious tests on their initiative, which not only makes them bear a large burden but also fails to continuously monitor the exercise performance of patients.

The gait disorders are also important factors affecting the quality of life of PD patients, and walking occupies most of the exercise time in daily life. Therefore, for PD patients with gait disorders, passive detection of motor performance has been proposed in many studies to achieve the goal of long-term monitoring. Peraza et al. proposed an automatic gait analysis pipeline based on DL algorithms for 4 wearable sensors, and the results showed the potential feasibility of using wearable sensors to assess gait in free-living and home scenarios. However, multiple wearable sensors may cause burden and discomfort to patients, thus current research is also moving toward a minimum number or even external monitoring of sensors.
Liu et al. passively measured home gait speed by detecting and analyzing reflected radio waves from the human body using radio devices, and used it as a key metric for characterizing and monitoring PD, enabling an objective assessment of disease severity, progression, and medication response at home. However, the ensuing problem is that the uninterpretable nature of the DL model makes people unable to fully trust it, thus limiting further clinical application. Some interpretable methods have been proposed to explain the underlying decision mechanism of the DL model, such as gradient-weighted class activation mapping (Grad-CAM), Shapley additive explanations (SHAP), etc. However, there are few studies on DL model visual interpretation and further optimization in the field of intelligent PD assessment. In this work, we conduct research that will help to address these questions.

This work provides an objective assessment algorithm that promises to enable long-term monitoring during the daily walking of patients, track the progression of PD, and reduce the dependence on clinicians. A total of 100 subjects were included in the experiment, in which each subject repeatedly collected 10-meter walking data using wearable sensors. The raw data are subjected to continuous wavelet transform (CWT) to obtain more detailed time-frequency information before training in the constructed DL architecture based on a 6-channel convolutional neural network (CNN). The classification performance of each sensor data-driven model was compared and interpretable visualized, with the waist sensor having the best performance. Finally, this work further visualizes the contribution of data in each channel to the model performance and optimizes the model according to the analysis results.

2. METHODS

All data processing process was conducted using Python 3.7, PyTorch 1.10, and other relevant python libraries.

2.1 Data acquisition

A total of 50 patients with varying severity of PD symptoms and 50 age-matched healthy controls (HC) participated in this work, excluding psychological, other neurological diseases, and movement disorders. All PD subjects were scored by the same neurologist according to UPDRS-III, and other details of subject characteristics are shown in table 1. All subjects understood the purpose of data collection and the significance of the study and signed the informed consent. This work was approved by the Ethical Committee of the Shandong Provincial Hospital in Jinan, China (NO.2020-600).

The IMUs (BWT901CL, Wit-Motion Company) with the 3-axis accelerometer and 3-axis gyroscope were employed to collect motion data, which could be sent synchronously to the connected personal computer via Bluetooth. Five IMUs attached to the wrist, ankle, and waist by comfortable Velcro straps form a body sensor network, as shown in figure 1(a). This fixed manner ensured consistent sensor position and orientation throughout the acquisition. From the
perspective of sensors, the x, y, and z axes were defined as transverse direction (left is positive), anterior-posterior (backward is positive), and vertical direction (up is positive), respectively. Subjects deploying the body sensor network were instructed by the neurologist to perform the 10-meter walking test at their preferred speed, with each subject repeating the test three times, as shown in figure 1(b). All sensors collected gait data with a sampling frequency of 50Hz, and the built-in accelerometer calibration function of the sensor eliminated human error caused by different wearing angles of different subjects.

<table>
<thead>
<tr>
<th>Characteristics</th>
<th>HC (N=50)</th>
<th>PD patients (N=50)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age (Years, μ ± σ)</td>
<td>63.3 ± 8.6</td>
<td>63.6 ± 7.2</td>
</tr>
<tr>
<td>Gender (Male/Female)</td>
<td>22/28</td>
<td>22/28</td>
</tr>
<tr>
<td>Duration of disease (Years, μ ± σ)</td>
<td>N.A.</td>
<td>4.6 ± 3.2</td>
</tr>
<tr>
<td>Total score of UPDRS-III (μ ± σ)</td>
<td>N.A.</td>
<td>27.0 ± 13.3</td>
</tr>
</tbody>
</table>

μ = average, σ = standard deviation, N.A. = not applicable.

Figure 1. Data acquisition. (a) The placements of sensors. (b) The 10-meter walking test.

2.2 Data pre-processing

2.2.1 Data screening and filtering

All data were initially screened to discard data recorded due to substandard performance of the subject or non-compliant sensor placement. Also, at the beginning and end of each walking test, redundant data that did not belong to the actual test due to human error were removed. Aiming at the problem that a small part of the signal was lost due to unstable transmission during the data recording process, the linear interpolation method was used to supplement the data according to the position of the lost data determined by the data acquisition time interval.

Band-pass filtering is often performed on time series data to eliminate low-frequency and high-frequency noise to reduce external influences. However, DL models usually require minimal filtering or introduce noise into the input data, thereby preventing model overfitting.
and improving generalization performance and robustness. Therefore, the data was not filtered at all in this work.

2.2.2 Continuous wavelet transform

Compared with the fixed time-frequency resolution of the short-time Fourier transform (STFT) technique, CWT enables dynamic resolution analysis of non-stationary signals to capture more detailed time-frequency information. In this work, the Morlet wavelet was used to perform the continuous transform on each time series (3-axis acceleration and angular velocity) of each sensor, providing richer input for the CNN model. The specific form is as follows.

\[
CWT(\psi, a, b) = \frac{1}{\sqrt{a}} \int_{-\infty}^{+\infty} f(t)\psi^*(\frac{t-b}{a}) dt,
\]

(1)

where \(\psi\) represents the Morlet mother wavelet, \(\psi^*\) represents the complex conjugate, \(a\) and \(b\) represent the scale and translation parameters, respectively.

2.2.3 Dataset construction

To capture more features for better monitoring of walking, all data were segmented into 2 seconds windows (100 samples) with 50% overlap. The entire dataset consisting 1832 windows, including 975 for PD and 857 for HC. The training set and test set were randomly divided according to 8:2 with the same PD and HC ratio.

For each time series of the sensor, the modulus range of coefficients obtained from CWT was different, so the data was normalized to improve the convergence speed of the model. To prevent potential overfitting, only the training set was used to calculate the normalization parameters, and the entire dataset was processed.

2.3 CNN architecture and training strategy

The CNN was widely used in classification, target detection, image segmentation, and other fields due to its advantages of automatically learning complex features. However, training a deeper model requires a large number of computing resources and data. Considering the limitation of the input windows sizes, this work constructed a 6-channel CNN architecture with only 6 convolutional layers according to the data characteristics to avoid overfitting, in which the channel represents the modulus obtained by CWT on the time-series component of the sensor data.

As shown in figure 2, each convolution block included two convolution layers with the same number of filters, and the kernel size was 3×3. The number of filters for different blocks was 8, 16, and 32, respectively. To improve the convergence speed and robustness of the model, the batch normalization (BN) layer was added after each convolutional layer. The activation function adopted the linear rectification function (ReLU). Each block ends with a max pooling
layer of pool size $2\times2$ and stride $2\times2$ for downsampling to reduce information redundancy. The last convolutional block was followed by two fully connected layers with 512 and 64 neurons respectively.

The performance of the model was evaluated by 10-fold cross-validation that divides the training set equally into 10 subsets. In the training process, 9 subsets were used for training, and the remaining one was used as validation data to test the performance of the model. This process was repeated 10 times to ensure that each subset can participate in training and validation. The hyperparameters are optimized according to the performance of the model on each subset, and the test set is used to evaluate the final performance of the model. The following settings and hyperparameters were selected for the model:

- Adaptive Moment Estimation (Adam) optimizer with initial learning rate = 0.001.
- Cross entropy loss function.
- Batch size of 8.
- Max epochs equal 200.

Moreover, the problem of model overfitting was often seen in CNNs, which perform extremely well on the training set but fail to generalize to the test set. Therefore, this work employs L2-norm regularization techniques to decay weights, combined with early stopping (stopping training when the validation set loss does not drop significantly within 30 epochs) to overcome overfitting.

2.4 Model evaluation and visualization

Accuracy, defined as the percentage of subjects who were correctly classified, was involved throughout the training and testing phases as a measure of performance. Besides, the area under the receiver operating characteristic (ROC) curve (AUC) was introduced to comprehensively evaluate the performance of the proposed model. AUC represents the probability that the value of the positive sample is higher than the value of the negative sample in a random sampling of a positive sample and a negative sample, that is, the larger the AUC value (no more than 1), the better the performance.
To make the output more understandable, the Grad-CAM was used for the visual interpretation of the model. The Grad-CAM can make existing state-of-the-art deep models interpretable, avoiding the trade-off between interpretability and accuracy. Specifically, described as:

\[ \alpha_{k} = \frac{1}{Z} \sum_{i} \sum_{j} \frac{\partial y^{c}}{\partial A_{ij}} \]

\[ L_{Grad-CAM} = ReLU \left( \sum_{k} \alpha_{k} A^{k} \right) \]  

where \( \frac{\partial y^{c}}{\partial A_{ij}} \) represents the gradient of the score \( y^{c} \) (before the softmax) for class \( c \) with respect to the neuron \( A_{ij} \) of a convolutional layer. These gradients obtained by backpropagation are subjected to global average pooling to obtain the importance weights \( \alpha_{k} \) of feature map \( A^{k} \).

The Grad-CAM can visualize any layer of any structural CNN without modifying the network structure or retraining. In this work, the last layer of feature maps was visually interpreted considering rich high-level semantic information and detailed spatial information. The obtained activation feature maps and weights can be used to visualize the basis for model judgment, which in turn helps us to distinguish the frequency range and spatiotemporal features with the most significant differences between PD and HC.

3. RESULTS

3.1 Results of data pre-processing

Figure 3 shows the raw data and the time-frequency plot of each time-series component obtained by the CWT from waist sensor (sensor 3) of randomly selected HC and PD subjects. The green and red boxes represent HC and PD subjects, respectively, where the acceleration and angular velocity curves are shown on the left ((a) and (c)), (b) and (d) are the normalized time-frequency plots corresponding to each component of acceleration (upper) and angular velocity (lower). From the raw data, HC subjects exhibited greater motion amplitude and frequency than PD subjects, while exhibiting larger values on the time-frequency map.
Figure 3. The raw data and the time-frequency plot from waist sensor (sensor 3) of HC and PD. The green and red boxes represent HC and PD subjects, respectively. (a) and (c) are the acceleration and angular velocity curves. (b) and (d) are the normalized time-frequency plots corresponding to each component of acceleration (upper) and angular velocity (lower). The orange box represents the 2s time window.

3.2 Classification Results

To find the optimal sensor locations for daily walking monitoring, the classification performance of 5 different sensors data-driven CNN models were compared in figure 4, including accuracy and AUC. Figure 4(a) showed the average accuracy of different models with 10-fold cross-validation on the test set, and the solid black lines represent their respective standard deviations. The results showed that the performance of waist sensor (sensor 3) was the best, reaching 98.01% ± 0.85%. Figure 4(b-f) depicted the ROC and AUC of different models with 10-fold cross-validation on the test set, in which the blue solid line represents the average value of 10 solutions, the other lines in different colors represent the results of 10 calculations, and the gray shading represents standard deviation. The analysis results also showed that the waist sensor exhibited the best performance (AUC = 0.9981 ± 0.0017). One-way analysis of variance verified that there was a significant difference (p<0.05) between the AUC results of all sensors. To further analyze the specific differences between groups, Tukey's HSD post-hoc analysis method was used. The results showed that the performance of the CNN model driven by the waist sensor was significantly better than that of the other sensors, while there is no significant difference among the other models.

Figure 4. Accuracy and AUC comparison of CNN models driven by 5 different sensor data. (a) The average
accuracy of different models with 10-fold cross-validation on the test set, and the solid black lines represent their respective standard deviations. (b-f) The ROC and AUC of different models with 10-fold cross-validation on the test set. The blue solid line represents the average value of 10 solutions, the other lines in different colors represent the results of 10 calculations, and the gray shading represents the standard deviation.

3.3 Model Visualization

The last convolutional layers of the CNN models that can capture higher-level visual constructs were visually interpreted by Grad-CAM ²⁷, and the visual maps of the same size as the input data were obtained through upsampling, as shown in figure 5. Four subjects (2 HC and 2 PD) were randomly selected for presentation, where each column in the figure represents a subject, and each row represents a sensor in figure 1(a). As shown, the feature points that significantly contributed to the discrimination of PD subjects were more concentrated in lower frequency bands (0.5-3 Hz) relative to HC (4-6 Hz).

The best-performing waist sensor (sensor 3) has better discrimination between HC and PD because the waist is closer to the center of gravity of the body and can better reflect the overall motion condition than the unilateral limb ³⁶,³⁷. Moore et al. ³⁸ defined the locomotor band of walking as 0.5-3Hz through the IMU attached on the left shank, while the sensor mounted on the waist in this work could display the activities of both body segments, thus the identified frequency band of HC was 4-6 Hz. In addition, the decision basis of the left-hand sensor (sensor 1) is also evident between the two classes, which was also reflected in the AUC results. Motor symptoms were more serious in the dominant hand ³⁹, and the subjects included in this work were almost all right-handed, which was reflected in sensor 2 of PD with a wider range of frequency band (0.5-7 Hz, including tremor).

Although there was no significant difference in the performance of the four sensor-driven models other than sensor 3, the decision basis was worse when attached to the ankle (sensor 4 and 5) compared to the wrist (sensor 1 and 2). This was because the motor manifestations of PD patients started distally in one upper limb ⁴⁰, while this work involved a small number of early stage patients without motor disorders in the lower limb. The analysis results of sensors attached to the ankles also showed some differences, but most of the features were concentrated in the lower frequency band determined by the locomotor band (0.5-3 Hz) of unilateral limb ³⁸, resulting in invalid differentiation. Another reason may be caused by the heterogeneity of limb motor symptoms in PD patients. Due to the individual differences of patients, the sensors located in limbs cannot accurately measure the overall movement of patients, so the visual maps showed disorder.
3.4 Model Optimization

Although the human body posture can be solved by IMU data, the limited time series components (e.g., only accelerometer data) of a particular movement may be enough to describe its motion characteristics, thus could reduce the burden of data acquisition. The above model visualization method explained the decision basis of each sensor and proved the advantage of sensor 3 over the others, however it was unable to verify the contribution of its various time series components to classification. Therefore, the waist sensor-driven model with the best performance was further optimized to remove redundant data, reduce data processing pressure, improve the generalization performance, and explore the interpretability of the model.

The dataset structure was reconstructed from a 6-channel 2D input (6×64×100) to a 1-channel 3D input (1×6×64×100) to visualize the key features of each time series (3-axis acceleration and angular velocity). The proposed network was modified to the 3D CNN and
retrained while keeping the basic architecture unchanged, only performing the waist sensor with the best performance. The contribution of each time-series component for the waist sensor to HC and PD classification is visually interpreted in figure 6.

As shown, PD focused on lower frequency band features in $a_x$ and $\omega_z$ compared with HC, which meant that the lateral movement of patients becomes difficult. Moreover, PD exhibited higher frequency features on $a_z$ than HC due to the more stable gait of HC. The $a_y$ component was not obvious in HC but focused on the low-frequency band in PD. Interestingly, these two groups had the same interpretation of $\omega_y$ that represents the side-to-side shaking. The reason for the inability to distinguish effectively may be due to the stable gait of HC and the slow movement of PD, both of which show low-frequency features. Besides, due to individual differences in patient symptoms, $\omega_x$ was confused on both.
According to the above analysis, the components that could not effectively distinguish HC from PD were removed and only \( a_x \), \( a_z \), and \( \omega_z \) were retained to train the new CNN model. Even though the data component was reduced by half, the new model still showed good performance (AUC = 0.9929 ± 0.0019). The optimized model paid more attention to the important characteristics of walking in PD patients. Although the performance of the optimized model was slightly reduced compared with the original model, the raw data processing pressure was greatly reduced (50%), which provided the possibility for the development of a real-time monitoring system of daily gait.

4. DISCUSSION

In this work, we adopted DL technology and wearable sensors to monitor gait disorders in PD patients for a long time. The original data collected by IMU were one-dimensional time series. If it was directly input into the CNN model for one-dimensional convolution operation, only time-domain features could be obtained while the equally important frequency domain features were ignored, resulting in limited model performance. Some studies used the fast Fourier transform method to extract the frequency domain features of data \(^9,41,42\), but it could not guarantee the time-frequency information of data at the same time. The STFT solved this problem to some extent but had a fixed time-frequency resolution \(^43,44\). Therefore, CWT with the dynamic resolution was used in this work to capture more detailed time-frequency features before input to CNN. CNNs were designed to process data from multiple arrays, such as the color image consisting of three channels (RGB) \(^20\). According to the collected data format (3-axis acceleration and angular velocity), we regarded each time series component as a channel array to construct a customized 6-channel CNN. Considering the limited sample size and reduced computational cost, only six convolutional layers were designed for the network.

Although the body wearable sensor network can reflect the motor performance of PD patients more comprehensively, using multiple sensors increases patient burden and may not be feasible for long-term monitoring of patients or disease management \(^45\). Therefore, the selection of the minimum number of sensors has become the current research hotspot \(^25,46\). Lonini et al. \(^46\) used six sensors to record movement data from 20 participants during 13 common tasks and showed that a single sensor was sufficient for the classification of bradykinesia and tremor. Hence, in this work, we constructed 5 CNN models based on the data of each sensor and compared their classification performance. The reason why the waist sensor (sensor 3) has better performance (AUC = 0.9981 ± 0.0017) than other sensors may be that its position is closer to the center of gravity of the body, thus can reflect the overall movement during walking \(^36,37\).

However, the black-box nature of CNN makes it limited in terms of high-risk decision-
making in the medical field even with excellent performance \(^{47}\). To build trust in intelligent systems, we must construct “transparent” models that explain the decision-making mechanisms behind them \(^{27}\). Chu et al. \(^{35}\) obtained characteristic frequency bands (high-delta (3.5–4.5 Hz) and low-alpha (7.5–11 Hz)) and spatial distribution features for early PD identification using EEG based on the interpretability of the model and visualization of the decision-making process. In our previous work, SHAP was introduced to explore the contribution of different motor tasks to the model, and the results showed that the features with the greatest impact on the model output almost all come from the alternating movement of the hand \(^{9}\). Therefore, in order to further explore the decision-making mechanism behind the proposed model, we used the Grad-CAM method to visually interpret the feature maps of the last layer of each model without changing the structure. The CNN model driven by the waist sensor does show better discrimination ability, and the feature points that make a great contribution to PD are concentrated in the lower frequency band (0.5~3Hz) compared with HC.

In addition, not all windows were classified correctly, and a few wrong cases were used to further reveal the decision-making mechanism behind them. Figure 7 illustrates the visual maps of different sensor-driven models in the case of decision-making errors, where the first two columns show that HC is misclassified as PD, and the last two columns show that PD is misclassified as HC. As shown, the misclassification occurred because a small number of HC windows exhibit disordered visual maps, and PD windows had an important decision basis in higher frequency bands (4-6Hz). This was consistent with the above findings that the feature points that play an important role in the classification of PD were concentrated in lower frequency bands relative to HC. Some patients had mild motor symptoms, which meant that gait impairment was not always present. Similarly, HC subjects may have gait fluctuations due to the influence of the external environment, resulting in classification errors.
The ultimate goal of PD intelligent monitoring is to develop an integrated system that combines data collection sensor and classification algorithms to monitor gait performance for a long time in the daily life of patients. However, the large computation cost of the CNN model is a great challenge for the microprocessor, even if CNN with simple architecture was used in this work. Therefore, we further optimized the waist sensor-driven model to reduce data processing pressure. The decision-making basis of each time series component was visualized by the reconstructed 3D CNN, and a new 3-channel CNN model was trained after removing the 3 poorly performing components. Although the performance of the optimized model is slightly reduced, the computational cost of raw data is reduced by 50%, which makes it possible to develop a system for monitoring daily gait for a long time. Some researchers have also explored the hardware implementation of the DL algorithm. Mikos et al. demonstrated that the freezing of gait detection system built from neural networks capable of learning in real-time was integrated into a single sensor node based on field programmable gate arrays (FPGA),
and could operate for more than 9 hours while providing auditory biofeedback cues. Langer et al. \(^{51}\) reported the implementation of a temporal convolutional network trained to detect freezing of gait based on FPGA, which can ensure sufficient time to trigger the cue in less than a millisecond to prevent the patient from falling. These will facilitate the design of dedicated hardware for PD.

5. CONCLUSIONS

Due to impaired motor function, PD patients showed symptoms such as slow walking and unsteady gait, and showed lower movement amplitude and frequency in sensor data and CWT map. By comparing the performance of CNN models driven by different sensors, the optimal sensor position is determined to be the waist with an AUC of 0.9981 ± 0.0017. The Grad-CAM further demonstrated that the feature points with greater contribution to PD were concentrated in the lower frequency band (0.5–3Hz) compared with HC. Moreover, we combine Grad-CAM with 3D CNN to propose a model optimization scheme that visually interpreted the contribution of each time series component to the model to remove components with insufficient contribution, which reduces raw data processing pressure by 50% at a slight performance sacrifice. In the future, we plan to design an integrated dedicated hardware circuit that can be attached to the waist of patients to monitor gait performance for a long time and track the progress of PD, which is expected to guarantee the realization of personalized treatment and improve the treatment effect.

DATA AVAILABILITY STATEMENT

The datasets generated for this study are available on request to the corresponding author. The data generated and/or analyzed during the current study are not publicly available for legal/ethical reasons but are available from the corresponding author upon reasonable request.

CODE AVAILABILITY

The code used to train and generate results is available on GitHub (https://github.com/Warden-X/Walking)

REFERENCES


Badola, A., Nair, V. P. & Lal, R. P. in *2020 IEEE 17th India Council International Conference (INDICON)*. 1-6 (IEEE).


Ancona, S. *et al.* Wearables in the home-based assessment of abnormal movements in Parkinson’s disease: a


**ACKNOWLEDGEMENTS**

This work was supported in part by the Innovation Project of Clinical Science and Technology of Jinan City (202134047) and in part by the Academic promotion program of Shandong First Medical University (2019LJ005).

**AUTHOR CONTRIBUTIONS**

M.C., Z.S., T.X., Y.C. and F.S designed the study. M.C., D.B., Z.S. and Y.C. collected the raw data. M.C. processed the data and developed the models. M.C. and F.S. produced the figures and tables and wrote the first draft of the manuscript. All authors (M.C., Z.S., T.X., D.B., Y.C., F.S.) critically reviewed, contributed to the preparation of the manuscript, and approved the final version. All authors vouch for the data, analyses, and interpretations.

**COMPETING INTERESTS**

The authors declare no competing interests.