Artificial Intelligence In Medicine 154 (2024) 102932

Contents lists available at ScienceDirect

Artificial Intelligence In Medicine

journal homepage: www.elsevier.com/locate/artmed

Research paper

Prediction of Freezing of Gait in Parkinson's disease based on multi-channel time-series neural network

Boyan Wang [a], Xuegang Hu [b], Rongjun Ge [c], Chenchu Xu [d,e,*], Jinglin Zhang [f], Zhifan Gao [g], Shu Zhao [e], Kemal Polat [h]

[a] Tsinghua University, Beijing, China
[b] Hefei University of Technology, Hefei, China
[c] Nanjing University of Aeronautics and Astronautics, Nanjing, China
[d] Institute of Artificial Intelligence, Hefei, China
[e] Anhui University, Hefei, China
[f] Shandong University, Jinan, China
[g] Sun Yat-sen University, Shenzhen, China
[h] Bolu Abant Izzet Baysal University, Bolu, Turkey

ARTICLE INFO

Keywords:
Freezing of gait
Multi-channel time-series neural network
Wearable sensors
Transformer
Parkinson's disease

ABSTRACT

Freezing of Gait (FOG) is a noticeable symptom of Parkinson's disease, like being stuck in place and increasing the risk of falls. The wearable multi-channel sensor system is an efficient method to predict and monitor the FOG, thus warning the wearer to avoid falls and improving the quality of life. However, the existing approaches for the prediction of FOG mainly focus on a single sensor system and cannot handle the interference between multi-channel wearable sensors. Hence, we propose a novel multi-channel time-series neural network (MCT-Net) approach to merge multi-channel gait features into a comprehensive prediction framework, alerting patients to FOG symptoms in advance. Owing to the causal distributed convolution, MCT-Net is a real-time method available to give optimal prediction earlier and implemented in remote devices. Moreover, intra-channel and inter-channel transformers of MCT-Net extract and integrate different sensor position features into a unified deep learning model. Compared with four other state-of-the-art FOG prediction baselines, the proposed MCT-Net obtains 96.21% in accuracy and 80.46% in F1-score on average 2 s before FOG occurrence, demonstrating the superiority of MCT-Net.

1. Introduction

The prevalence of Parkinson's disease continues to rise in adults [1], and an estimated 7 to 10 million people globally have Parkinson's disease right now [2]. Freezing of gait (FOG) is a debilitating symptom and an abnormal gait pattern that accompanies Parkinson's disease [3]. FOG manifests the patient's inability to take any continuous steps in a short time, thereby significantly obstructing their mobility and independence [4]. Even worse, FOG increases the risk of falls. The unpredicted fall causes patient injuries, accelerating the deterioration of the patient's conditions [5]. To avoid injuries, FOG detection and FOG prediction are proposed to protect patients. FOG detection is available to reduce the risk of falls by identifying FOG once it has occurred [6]. Unlike FOG detection, FOG prediction provides the potential to forecast FOG episodes before they happen [7], allowing for timely interventions and preventive measures before the fall. Hence, credible prediction of FOG is urgent for patients with Parkinson's disease.

Nowadays, wearable sensors can detect an imminent freeze, playing an important role in long-term symptom monitoring and deploying preemptive mitigation via multiple sensory cues. In the last 15 years, wearable sensor systems have been widely implemented for the daily care of FOG episodes [8]. Owing to the convenient wearable sensors, neurologists can observe the FOG symptom through the learned behavior patterns at any time [9]. Based on the objective data, an efficient FOG wearable sensor system significantly improves quality of life and increases life expectancy.

Even with mature wearable sensor techniques, it is still hard to predict FOG accurately and timely. Firstly, the FOG occurrences are difficult to record comprehensively because the mechanism behind FOG remains obscure in the current medical standard [10]. Besides, the goal

* Corresponding author at: Anhui University, Hefei, China.
E-mail addresses: wby000000@mail.tsinghua.edu.cn (B. Wang), jsjxhuxg@hfut.edu.cn (X. Hu), rongjun.ge@nuaa.edu.cn (R. Ge), cxu332@gmail.com (C. Xu), jinglin.zhang@sdu.edu.cn (J. Zhang), gaozhifan@gmail.com (Z. Gao), zhaoshuzs2002@hotmail.com (S. Zhao), kpolat@ibu.edu.tr (K. Polat).

https://doi.org/10.1016/j.artmed.2024.102932
Received 2 January 2024; Received in revised form 30 May 2024; Accepted 4 July 2024
Available online 6 July 2024

of the clinical treatment is often to ease and control the symptoms of FOG by using medication adjustment, physical therapy, or walking aids [11]. Such symptom control therapy and conservative treatment cannot heal the FOG thoroughly, and worsening of the condition with age and falling will inevitably occur [7]. Moreover, due to the sporadic nature of FOG at the early stage, the clinical assessment of FOG needs to consider various spatial–temporal gait parameters, which is time-consuming and incompatible with everyday clinical practice [12]. The combination of machine learning and wearable sensor systems is a sound way to solve the dilemma without professional knowledge [13].

By classical machine learning approaches, wearable sensor systems can successfully analyze behavior patterns using gait features from temporal data based on a single sensor [14,15]. Most machine learning methods attempt to extract the gait features and detect abnormal movements based on statistical computing [16,17]. Among these machine learning approaches, deep learning with its powerful representation learning capability, can automatically extract and learn the high-level features from raw time-series data [18–20]. Due to the outperforming representation results without expert complex settings, deep learning methods have become the best choice for the wearable sensor system.

However, existing deep learning approaches still face two challenges in predicting FOG. The first challenge is how to predict accurately the occurrence of FOG in advance. Considering the delay between wearable sensors and the remote device, the patient might have fallen before the alarm. Recognizing the gait pattern in different time slices is the key to providing the real-time monitor [21]. The feature distribution is not significantly distinct in different time slices, which causes embedding errors and is hard to warn early in the prediction task. The other challenge is the interference in the feature representations between multiple wearable sensors. The multi-channel wearable sensors have intra and inter interferences during representation learning [22], corresponding to the correlation and diversity in the sensor system. The interference covers the high-quality representation and needs to be reduced in a comprehensive multi-channel system.

To predict FOG in Parkinson's disease using a practical wearable sensor system, we propose a novel multi-channel time-series neural network (MCT-Net) to meet the challenges. The proposed MCT-Net creatively designs a feature learning layer that leverages causal distributed convolutions to learn gait features in behavioral patterns effectively. These convolutions can integrate the confusing multivariate feature distributions of the embeddings from different time slices into a unified framework, thereby improving the predicting accuracy. Besides, our MCT-Net inventively introduces an intra-channel and inter-channel transformer, which applies multi-head self-attention to optimize the multi-channel gait features by potential correlations and diversities among multiple wearable sensors. The intra-channel and inter-channel transformer can weaken interference between and within wearable sensors, thus enhancing the feature representations. By integrating the proposed convolution and transformer, MCT-Net enables an alarm to prevent falls before the FOG occurrence, as shown in Fig. 1. At last, the proposed approach aids in progressing our understanding of Parkinson's disease and underlying neuropsychiatric mechanisms that cause FOG.

The main contributions of this work include:

- A novel deep learning method, i.e., MCT-Net, is proposed to predict FOG for advanced multi-channel wearable sensor systems. MCT-Net enables an alarm about 2 s before the FOG to prevent falls and to be implemented in remote monitoring devices.
- To form a unified temporal feature embedding from different time slices, causal distributed convolutions integrate various gait features in behavioral patterns for the efficiency of prediction in advance.
- An intra-channel and inter-channel transformer is proposed to receive the correlations and diversities by the multi-head self-attention mechanism, thus overcoming the interference among wearable sensors and achieving outstanding performance on the real-world FOG dataset.

2. Related work

2.1. Deep learning for FOG

The approaches based on deep learning in handling FOG in Parkinson's disease patients have recently emerged as a pivotal area of gait pattern recognition [12,23,24]. Distinguished from traditional machine learning methods such as decision trees [25], random forests [26], Naive Bayes [25], nearest neighbors [6], and support vector machines [27], deep learning leverages multi-layered neural networks to significantly enhance the capability of modeling complex gait patterns. Hence, it performs superior in detecting, predicting, and observing FOG episodes [12]. Recent studies in FOG detection and prediction based on deep learning [28,29] can be categorized into two primary methods: CNNs [30–32] and RNNs [33,34]. CNN-based methods [35] are adept at processing spatial features from wearable sensors, identifying crucial representations of impending FOG episodes. RNN-based methods, focusing on the long short-term memory (LSTM) network, have advantages in predictive modeling with causality [36]. They are good at analyzing time-series data and extracting temporal gait patterns to predict FOG occurrences [37]. However, due to model analysis limitations and multi-source data integration, current deep learning methods for FOG prediction merely provide warnings about 1 s before an episode [38]. This small window of time significantly hampers effective intervention. Patients have inadequate response time, and automated systems cannot offer assistance [39]. Therefore, it is urgent to explore advanced neural network architecture that can extend the prediction window of time while keeping accurate feature representations based on deep learning.

2.2. Feature embedding by wearable sensors

The wearable sensors in healthcare have revolutionized patient data collection and analysis, making feature embedding increasingly important [40]. Accelerometers and gyroscopes, as traditional sensors, have high-dimensional and noisy data [41]. Therefore, a practical feature embedding framework, such as transforming raw sensor data into a comprehensive representation for FOG prediction [42], is essential. Traditional feature embedding methods are predominantly hand-crafted [43]. The feature representations of these methods are often coarse and inaccurate [44]. For instance, temporal feature embeddings are based on statistical measures using mean, standard deviation, and range values from sensor records [45]. They capture basic movement features but may overlook intricate patterns in specific conditions like FOG. The other example is time–frequency features, utilizing fast Fourier transform (FFT) to analyze frequency components of sensor signal [46]. FFT is adept at identifying rhythmic patterns in gait but ineffectively captures sudden, transient changes. Time–frequency features, employing Wavelet transform for a combined analysis [47], offer a more detailed perspective but still face limitations in the dynamic and complex features of gait patterns, especially in pathological conditions. The recent deep learning technology uses its representation learning capability to automatically extract and select the suitable feature embedding without the hand-crafted process, significantly improving the prediction performance [48]. However, finding the optimal feature embedding by representation learning is still a well-known open challenge in FOG prediction.

2.3. Attention mechanism for temporal data

The attention mechanism for weighting the features has a close relationship to the downstream task [49,50], including time-series forecasting. The attention mechanism is an effort to simulate the human mind, where a relative part of the task pays more attention than the others with the global picture [51]. This mechanism utilizes the neural network to calculate different correlations corresponding to all feature

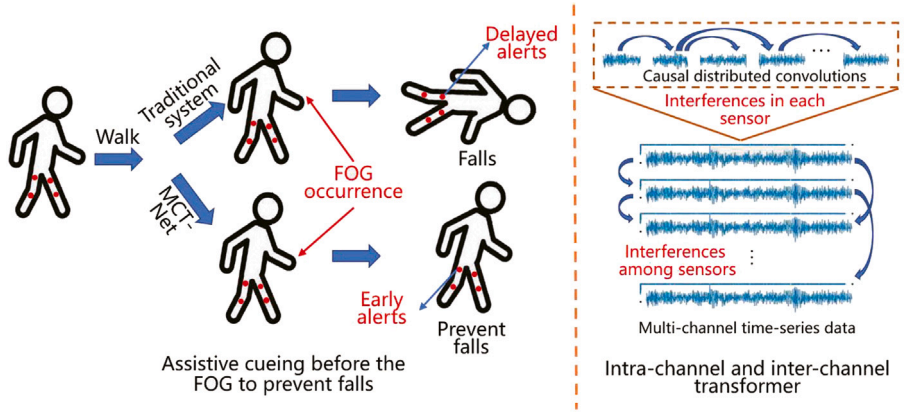

Fig. 1. Owing to two innovative submodules, causal distributed convolutions are designed for unified feature distribution, and an intra-channel and inter-channel transformer attempts to weaken the interference. MCT-Net can predict FOG before the occurrence to prevent falls and improve the patient's quality of life.

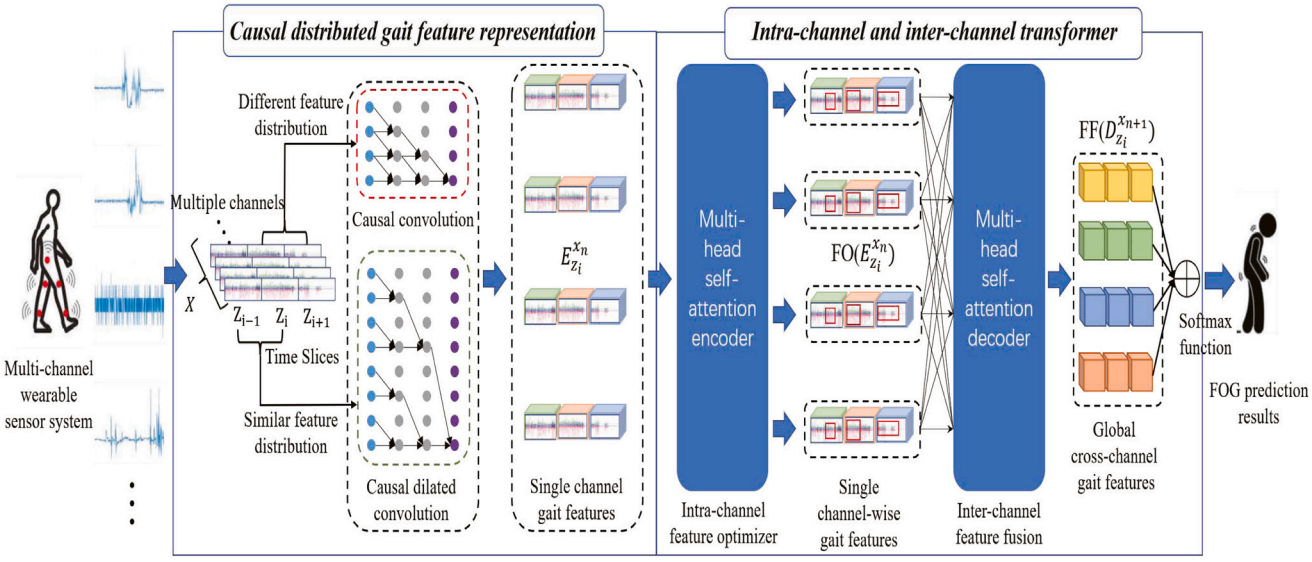

Fig. 2. The main framework of our proposed MCT-Net. Causal distributed gait feature representation is responsible for a unified temporal feature embedding by adaptive learning strategies. The intra-channel and inter-channel transformation is responsible for the correlations and diversities by encoding optimizer and decoding fusion.

positions of temporal data [52]. Otherwise, the attention framework can find the pivotal information of the features, which has a high sensitivity in learning the long-range dependencies [53]. Recent works on the attention mechanism of the transformer predominantly employ the multi-head self-attention framework to execute natural language processing tasks by establishing connections with all positions at the sequence level [54]. This methodology has also been adapted for image classification and segmentation to find the long-range dependency within the image's inner correlations [55]. Furthermore, the attention model has been applied in the time-series analysis of FOG detection [56,57] to utilize temporal contextual information. So far, the attention mechanism has yet to be investigated for time-series data from multi-channel wearable sensors in FOG prediction thoroughly.

3. Methodology

The proposed MCT-Net takes the multi-channel temporal data $X \in \{x_1, \ldots, x_n, \ldots, x_N\}$ as input, and the FOG prediction result $Y \in \{y_1, y_2, y_3, y_4\}$ as output, where each x_n is a single wearable sensor that has m time steps ($x_n = \{t_1, \ldots, t_m, \ldots, t_M\}$). y_1, y_2, y_3, and y_4 indicate four prediction types, i.e., conditions before FOG occurrence, during FOG, without FOG, and FOG cessation. The MCT-Net contains two submodules: the first is the causal distributed gait feature representation, and the other is the intra-channel and inter-channel transformer, as shown in Fig. 2. The transformer can be further divided into the intra-channel feature optimizer and the inter-channel feature fusion. Specifically, the gait feature representation (Section 3.1) leverages causal distributed convolution layers as an extraction to learn the behavior features from different time slices. The intra-channel feature optimizer (Section 3.2) builds a single-channel attention mechanism to highlight the key embedding of gait features corresponding to an optimal encoder in each channel. The inter-channel feature fusion (Section 3.3) integrates a cross-channel attention mechanism to handle the correlations and diversities among all channels, thus obtaining the comprehensive prediction by decoding multi-channel gait features.

3.1. Causal distributed gait feature representation

The causal distributed gait feature representation applies 3 residual block stacks (for each sensor's front, vertical, and lateral direction) based on the proposed causal distributed convolution as the main framework. The submodule takes temporal data X as input and learns each channel independently, corresponding to the output of the gait embedding E^{x_n}. Specifically, each block stack has a 1D causal distributed convolution layer, a LayerNorm layer [58], and a ReLU layer [59]. Owing to causal distributed convolution, our gait feature

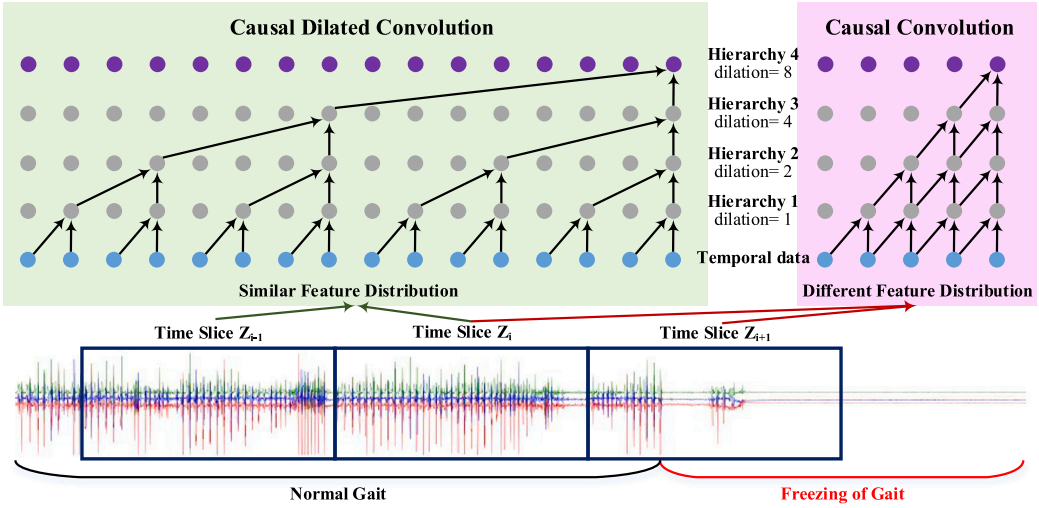

Fig. 3. Visualization of causal distributed gait feature representation. Based on the feature distributions of the gait pattern, our proposed causal distributed convolutions adaptively apply sampling area by the Kullback–Leibler divergence.

representation automatically adjusts the embedding strategies, accelerating sampling during normal gait periods and maintaining the original causal convolution layer whenever a freeze symptom occurs.

Compared with the classical causal dilated convolutions [60], the proposed causal distributed convolution is a time-series feature sampler based on time slices. Each time slice $Z_i = \{t_m, \ldots, t_{m+32}\}$ consists of 32 time steps (about 0.5 s) to calculate feature distribution. Based on the Kullback–Leibler divergence $D(Z_{i+1} \| Z_i)$ [61] of adjacent time slices, MCT-Net keeps the dilation rate increases by a factor of 2 in each layer until reaching the maximum dilation of 128 when adjacent time slices are similar ($D(Z_{i+1} \| Z_i) \leq 0.4$). Otherwise, MCT-Net resumes the original causal convolution to deal with the significant difference in gait feature distribution, as shown in Fig. 3. For each functioning wearable sensor, the freeze event is the main factor that causes the changes in the feature distribution. With adaptive strategies for gait feature embedding, the causal distributed gait feature representation is sensitive to the FOG episode starts and ends.

In advance, MCT-Net can effectively predict the FOG occurrence several seconds ahead based on the combination of multiple time slices. The gait feature representation increases and skips the dilated time steps while the feature distribution is similar. Therefore, causal distributed gait feature representation obtains a feature pyramid containing time slices from the past until now. We leverage 4 time slices (about 2 s) as a training batch in each block stack so our proposed representation can realize the changes in the feature distribution in the last 2 s. In the following intra-channel feature optimizer, MCT-Net will further emphasize the feature changes during the period that consisted of 4 time slices. Accordingly, MCT-Net can recognize the potential sign of freeze events via the feature distribution in a single sensor.

To sum up, for a random time slice Z_i, the gait embedding $E_{Z_i}^{x_n}$ of each block stack's output follows a logistic distribution $\mathcal{L}(x_n | \mu(Z_i), \delta(Z_i))$ parameterized by μ and δ:

$$E_{Z_i}^{x_n} = Conv(x_n | \mathcal{L}) \tag{1}$$

where $Conv$ represents a convolution operator of the causal distributed convolution. The convolution operator is the standard causal convolution when the adjacent logistic distribution $\mathcal{L}(x_n | \mu(Z_{i-3}, Z_{i-2}, Z_{i-1}, Z_i), \delta(Z_{i-3}, Z_{i-2}, Z_{i-1}, Z_i))$ is different. The convolution operator is the causal dilation convolution when $\mathcal{L}(x_n)$ is similar. We use the adaptive causal distributed convolutions as the base gait features embedding to recognize the behavioral pattern and strengthen the prediction before freeze events.

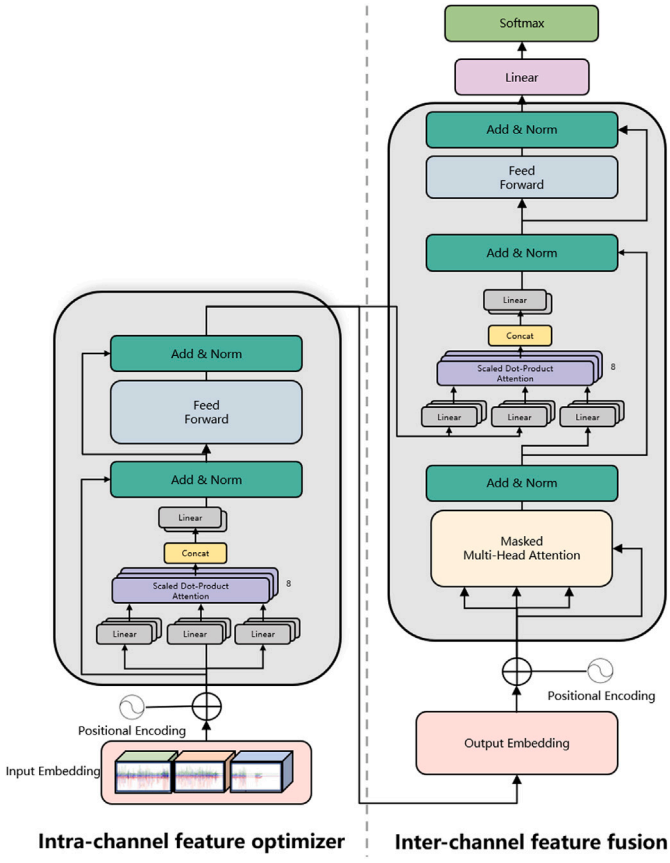

Fig. 4. Visualization of the intra-channel and inter-channel transformer. We employ the intra-channel feature optimizer as an encoder and the inter-channel feature fusion as a decoder.

3.2. Intra-channel feature optimizer

As shown in Fig. 4, we import the multi-head self-attention mechanism as the intra-channel and inter-channel transformer to highlight the crucial gait. As a single-channel encoder of the transformer, the intra-channel feature optimizer utilizes a self-attention block after each channel's gait feature representation to optimize single channel-wise

features, allowing our model to focus on the key behavior features. Using single channel gait embedding $E^{x_n}_{Z_i}$ for each time slice as the input, MCT-Net merges the input with the relative time step order encoding into the multi-head attention. The relative time step order encoding $Te()$ provides the relative position of a gait embedding of time step in each time slice. By the multi-head self-attention block, MCT-Net jointly attends to the embedding from different time slices at different time steps. With a single attention head, averaging inhibits this. We compute the output of multi-head attention $MA(E^{x_n})$ as:

$$\begin{aligned} &MA(E^{x_n}) = concat(Att_1(E^{x_n}_{Z_1}), \dots, Att_i(E^{x_n}_{Z_i}), \dots)W^{MA} \\ &where\ Att(E^{x_n}_{Z_i}) = \mathrm{softmax}(\frac{Te(E^{x_n}_{Z_i})Te(E^{x_n}_{Z_j})^T}{\sqrt{d_E}})Te(E^{x_n}_{Z_i}) \end{aligned} \tag{2}$$

where j is the index of any time slice, W^{MA} is the parameter matrix, and d_E is the dimension of the gait embedding $E^{x_n}_{Z_i}$. In this paper, we employ 8 parallel attention layers or heads. The self-attention is a dot-product attention that is much faster and more space-efficient in practice than the traditional additive attention [51], so it can effectively highlight similar gait embedding using optimized matrix multiplication code.

In order to reduce interferences in each wearable sensor, the intra-channel feature optimizer leverages a fully connected feed-forward network after the self-attention block, consisting of two linear transformations with a ReLU activation.

$$FO(E^{x_n}) = \max(0, W_1 MA(E^{x_n}) + b_1)W_2 + b_2 \tag{3}$$

where W_1 and W_2 are parameter matrices, and b_1 and b_2 are biases. As the final output of the proposed feature optimizer $FO()$, we obtain the key single channel-wise gait features by self-attention mechanism, which is applied to each time slice separately and identically. We employ the intra-channel feature optimizer as an optimal encoder of a channel for our transformer. Moreover, $FO(E^{x_n})$ will be the input of the next feature fusion to keep the correlations among time slices in each wearable sensor.

3.3. Inter-channel feature fusion

Algorithm 1 MCT-Net algorithm

Input: multi-channel temporal data $X \in \{x_1, \dots, x_n, \dots, x_N\}$.
Output: FOG prediction result $Y \in \{y_1, y_2, y_3, y_4\}$, where indicates pre-FOG, FOG, None-FOG, and post-FOG.

1: Parameter initialization of the model.
2: Divide each sensor temporal data x_n into time slice $Z_i^{x_n}$.
3: **if** $D(Z_{i+1}||Z_i) \le 0.4$ **then** Generate the gait embedding $E^{x_n}_{Z_i}$ of each similar-distributed time slice by causal dilated convolutions.
4: **else** Generate the gait embedding $E^{x_n}_{Z_i}$ of each different-distributed time slice by causal convolutions.
5: **for** each epoch **do**
6: Generate the intra-channel multi-head attention $MA(E^{x_n})$ using Formula (2).
7: Optimize the intra-channel gait encoder $FO(E^{x_n})$ using Formula (3).
8: Generate the other channel multi-head attention $D^{x_{n+1}}$ using Formula (2).
9: Integrate inter-channel multi-head attention $MA(D^{x_{n+1}})$ using Formula (4).
10: Generate the inter-channel gait decoder $FF(D^{x_{n+1}})$ by the feed-forward network.
11: **end for**
12: Predict the gait situation result Y by the softmax function to convert $FF(D^{x_{n+1}})$.

The inter-channel feature fusion integrates single channel-wise feature $FO(E^{x_n})$ into a unified decoder and uses the self-attention mechanism to learn the correlations and diversities among wearable sensors.

Table 1
Statistics of Daphnet FOG dataset.

Gender (male)	70% (7/10)	Age (years)	66.4 ± 4.8
Number of FOG events		237	
FOG/None-FOG distribution		11%	
Disease duration		13.7 ± 9.67	
Number of channels		9	

Similarly, we obtain the previous multi-head attention results $D^{x_{n+1}}_{Z_i} = MA(E^{x_{n+1}}_{Z_i})$ from the other channel sensors using the Formula (2). Compared with the multi-head self-attention of the feature optimizer, we need to mask rightward predicted time steps in the decoder to preserve the auto-regressive property. With the $FO(E^{x_n})$ of dimension d_{FO} comes from the output of the encoder, the encoder–decoder self-attention layer establishes the transformation between multi-channel sensors.

$$\begin{aligned} &MA(D^{x_{n+1}}) = concat(Att_1(D^{x_{n+1}}_{Z_1}), \dots, Att_i(D^{x_{n+1}}_{Z_i}), \dots)W^{MA} \\ &where\ Att(D^{x_{n+1}}_{Z_i}) = \mathrm{softmax}(\frac{D^{x_{n+1}}_{Z_i} FO(E^{x_n})^T}{\sqrt{d_{FO}}})FO(E^{x_n}) \end{aligned} \tag{4}$$

where W^{MA} is the parameter matrix for the encoder–decoder multi-head layer, different from Formula (2). Similarly, we employ a fully connected feed-forward network after the attention block. The output of the feed-forward network $FF(D^{x_{n+1}})$ is the final result of the decoding feature fusion corresponding to encoder input E^{x_n}. The feature fusion effectively reduces the interference by drawing global dependencies between the intra-channel and the other channels.

To obtain the prediction from the transformer embedding $FF(D^{x_{n+1}})$, MCT-Net uses a fully connected layer with a dropout layer to classify the multi-channel self-attention features into the default prediction categories. We convert the category output to the FOG label by the softmax function. The detailed pseudo-code of MCT-Net is shown in Algorithm 1. In Algorithm 1, from Line 2 to Line 4, describe the causal distributed gait feature representation submodule. This submodule focuses on the feature embedding of each time slice in behavioral patterns. Line 5 to Line 11 denote the intra-channel and inter-channel transformer submodule, utilizing the multi-head self-attention mechanism for multi-channel comprehensive prediction.

4. Experiments

4.1. Experimental setup

4.1.1. Experimental dataset

We evaluate the proposed MCT-Net on the Daphnet FOG dataset [62], which is based on real-world wearable sensors. The dataset has nine channels corresponding to three wearable sensors. Three wearable locations, ankle, thigh, and trunk, have been used for FOG detection, so there are interferences in prediction performance among them. Each sensor comprises the temporal data to describe the acceleration in the horizontal front, lateral, and vertical directions. Each sample in millisecond has nine acceleration values with a tag for a time step. More details of the Daphnet FOG dataset are shown in Table 1.

Following the labeling strategy [38] to predict FOG, we import new labels, pre-FOG and post-FOG, with the original FOG and None-FOG. In detail, pre-FOG denotes the patient's condition 2 s before the FOG episode. FOG is the FOG occurrence, and None-FOG is without FOG occurrence. The post-FOG label denotes the patient's condition 2 s after the FOG episode. These two new labels make the FOG prediction model focus on the start and the end of FOG events. Four prediction labels are mutually exclusive for each time step. Eight out of ten patients experienced FOG on the Daphnet FOG dataset, and professional physiotherapists identified 237 freeze events. On the one hand, FOG events have various distributions among ten patients, which are collected from the real-time record. On the other hand, using different portions of the

data on different iterations is the direct way to validate the efficiency of our MCT-Net in random data distribution. Hence, we perform five-fold cross-validation to train and test a model for the Daphnet FOG dataset, and all the results are shown as the mean.

4.1.2. Baseline methods and implementation details

We compare our MCT-Net with the following classical baselines:

- **LDA** [17]: Linear discriminant analysis (LDA), as a traditional machine learning method, to catch FOG and pre-FOG episodes.
- **PCA** [38]: Principal component analysis (PCA) is a clustering algorithm that reduces high-dimension features and applies KNN to predict the classification results for FOG.
- **DAE** [16]: A probabilistic denoising autoencoder (DAE) learns the distribution of normal human movements, thus quantifying the deviation of each test sample for abnormal movement detection.
- **LSTM** [18]: Long short-term memory (LSTM) is a neural network method to predict freeze events in short time durations before FOG occurrence.

To compare the baselines fairly, we follow the optimal setting in the corresponding reference. For LDA, we set Gamma in the range of [0.01, 1] and Delta in the range of [0.01, 100] to search for the optimal model. For PCA, we reduce the dimensionality to 45 and then use KNN for classification, following the original paper. For DAE, we set the learning rate to 0.0001, the batch size to 16, and ran the training for 200 epochs to reach the optimal performance. For LSTM, we set the learning rate to 0.0002, the batch size to 8, and ran the training for 200 epochs to achieve the best results. For our proposed MCT-Net, we conduct Adam optimizer ($\beta_1 = 0.9$, $\beta_2 = 0.999$, and $\varepsilon = 10^{-9}$) [63] using the learning rate = 0.0004, decay factor = 0.1, and batch size = 32. The dimensionality of the input and output of feed-forward networks is 128. All experiments were run on the Intel i9 10900KF CPU with NVIDIA GeForce RTX3090 and 64 GB RAM. The operating system and software platforms are Linux Ubuntu 20.04, CUDA 11.7, TensorFlow 2.6.0, Pytorch 1.10.0, and Python 3.9.12.

4.1.3. Evaluation metrics

To compare the prediction results with baselines, we judge the results of experiments with four quantitative evaluation metrics: accuracy, precision, recall, and F1 score. The accuracy, precision, and recall are formulated as:

$$\text{Accuracy} = \frac{TP + TN}{TP + TN + FP + FN} \tag{5}$$

$$\text{Precision} = \frac{\text{TP}}{\text{TP} + \text{FP}} \quad \text{Recall} = \frac{\text{TP}}{\text{TP} + \text{FN}} \tag{6}$$

where TP is the true positive value, indicating the number of present labels predicted correctly. FP is the false positive value, indicating the number of absent labels predicted correctly. TN is the true negative value, indicating the number of present labels predicted wrongly. FN is the false negative value, indicating the number of absent labels predicted wrongly.

The F1-score is the reciprocal metric of the harmonic mean of precision and recall to measure the prediction performance. The F1-score is defined as:

$$\text{F-measure} = \frac{(1 + \alpha^2) \times R \times P}{\text{R} + \alpha^2 \times \text{P}} \tag{7}$$

where R represents recall and P represents precision. α is a hyper-parameter and is default to 1 [64].

4.2. FOG prediction analysis

As shown in Fig. 5, MCT-Net obtains the best performances corresponding to four prediction types compared to the other baselines. The internal areas of MCT-Net are the largest among all the FOG prediction methods. Specifically, MCT-Net improves F1-score by about 2%, 1%, 3%, and 1% in terms of pre-FOG, FOG, post-FOG, and None-FOG. Focusing on two new labels, MCT-Net achieves efficient performance with 96.21% of accuracy, 78.64% of precision, 72.31% of recall, and 80.46% of F1-score corresponding to pre-FOG. For post-FOG prediction results, MCT-Net obtains 96.38% of accuracy, 75.99% of precision, 70.24% of recall, and 78.33% of F1-score. Compared with the classical baselines, the proposed MCT-Net successfully extends the warning time before freeze events to 2 s. The outstanding performance validates that potential freeze signals are embedded in the gait pattern before and after the FOG episode, which deserves further study. Furthermore, we provide ROC and PR curves for the complete temporal test data to prove our MCT-Net effectiveness in Fig. 6. For the rest of the experimental representations, we provide the overall performance, i.e., the average results for four prediction labels. In ROC curves, the x-axis represents 1-specificity, while the y-axis represents the overall sensitivity. MCT-Net (blue dashed line) is more inclined to the point (0, 1) in the upper left corner than the other baselines, illustrating the optimal results in real-time FOG prediction. In PR curves, the x-axis represents the overall recall, while the y-axis represents the overall precision. Compared with four baselines, MCT-Net (blue dashed line) is closer to the coordinate (1, 1) in the upper right corner, validating the credible classification results based on our comprehensive framework. The above experimental results conclude that MCT-Net can assist physiotherapists in automatically analyzing the gait patterns of patients and detecting different conditions before, during, and after FOG events.

4.3. Function of the causal distribution convolution

To validate the advantage of the proposed causal distributed convolutions, we compare it with the traditional time-series feature representation ways in overall F1-score, i.e., PCA, recurrent neural network (RNN), LSTM, causal convolutions, and causal dilated convolutions. Notice that PCA and LSTM are classical feature representations in this section rather than the baselines. We employ the intra-channel and inter-channel transformer as the downstream classifier to fairly compare with the other feature embedding results. As shown in Fig. 7(a), causal distributed convolutions defeat the other classical baselines and are the best temporal feature embedding among deep learning methods. Specifically, ours performs better than the original causal convolutions and causal dilated convolutions. It validates that the adaptive strategy of causal distributed convolutions is beneficial to forming a unified feature representation.

4.4. Function of the intra-channel and inter-channel transformer

Fig. 7(b) also shows the function of the intra-channel feature optimizer and the inter-channel feature fusion in the transformer submodule. We compare the overall F1-score results among three frameworks: without attention, only with the intra-channel feature optimizer, and the proposed complete transformer (with both the intra-channel feature optimizer and the inter-channel feature fusion). Three frameworks use causal distributed gait feature representation as feature embedding and a fully connected layer with a dropout layer to predict the FOG condition. The framework with the intra-channel feature optimizer and our proposed transformer outperform the framework without attention. It proves that the multi-head self-attention mechanism highlights the key gait features in a single channel or among multiple channels. The best performance based on the proposed transformer concludes that the correlations and diversities in the gait pattern can overcome the interference among wearable sensors.

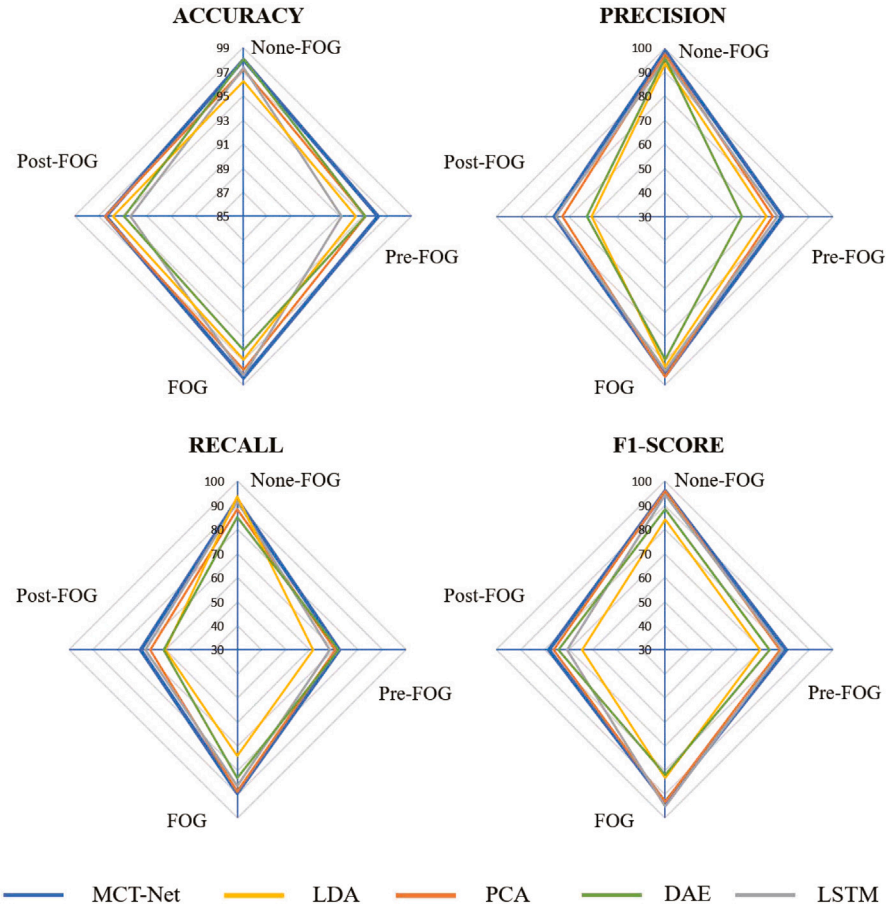

Fig. 5. Radar charts of FOG prediction results in Accuracy, Precision, Recall, and F1-score. The four subfigures show four evaluation metrics. Each radar chart has four axes corresponding to four prediction types (FOG, None-FOG, Pre-FOG, and Post-FOG). Each radar chart has five methods, and each method plots a curve with a unique color. The color legend is at the bottom of this figure. The largest internal area of the curve represents the best performance. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

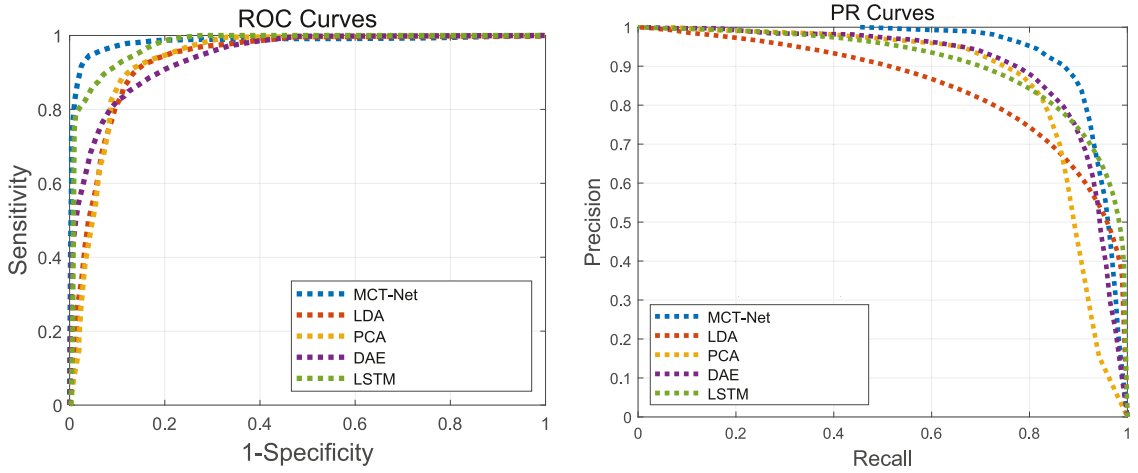

Fig. 6. ROC curves and PR curves for FOG prediction. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

4.5. Discussion and limitations

MCT-Net is an efficient method with an advanced multi-channel wearable sensor system. The proposed method is not only feasible for real-time FOG detection but also FOG prediction about 2 s before the episode. The overall performances are comprehensive indices corresponding to four prediction types, reflecting real-time detection and prediction ahead of time. In Table 2, MCT-Net outperforms the other four baselines in accuracy, specificity, precision, recall, and F1-score among five evaluation measures. The ultimate goal of MCT-Net is to be implemented in remote monitoring devices, such as the fitness tracker. Despite the considerable training phase, MCT-Net can give accurate FOG prediction timely during the test phase. The attention mechanism of the proposed transformer costs most computation resources in the training process, but the intra-channel and inter-channel attentions remain unchanged during the testing process. Compared with the other

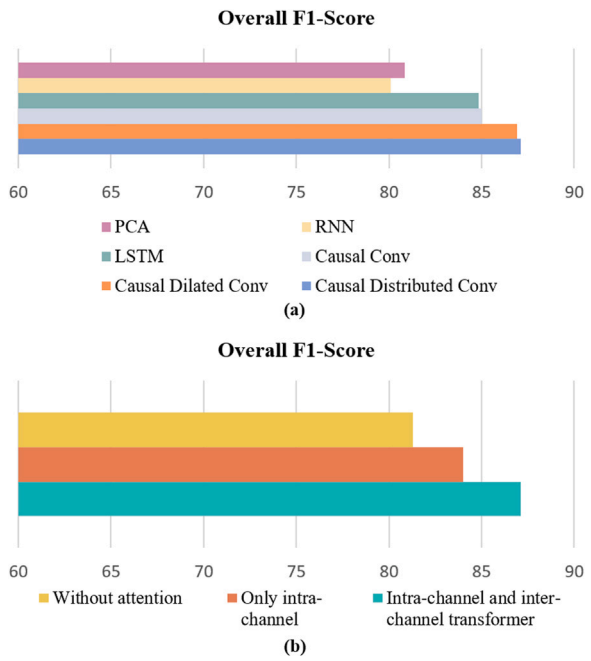

Fig. 7. Ablation study of MCT-Net. (a) illustrates the function of the causal distributed gait feature representation. (b) illustrates the function of the intra-channel and inter-channel transformer. By combining these two components, MCT-Net achieves the best overall F1-score.

Table 2
Experimental results of overall prediction on Daphnet FOG (%). The best ones are in **red&bold**.

Method	Accuracy	Specificity	Precision	Recall	F1-Score
LDA	95.81	98.32	79.64	72.32	75.33
PCA	96.64	98.72	85.81	78.53	85.87
DAE	96.07	98.41	77.53	75.30	79.63
LSTM	95.78	99.01	85.36	78.68	84.65
MCT-Net	**97.26**	**99.13**	**87.34**	**81.09**	**87.12**

deep learning models, the causal distributed gait feature representation is a real-time embedding method. Hence, MCT-Net can handle the real-time gait feature.

Each component of MCT-Net is beneficial to FOG prediction, which boosts overall performance submodule by submodule with a consistent rise. As shown in Fig. 7, the ablation studies have reflected the outstanding performance of MCT-Net, 87.12% of comprehensive F1-scores, in both temporal feature embedding and transformer fields. Compared with classical time-series feature representation methods, our causal distributed distribution convolution pipeline has a similar operating time to casual dilated convolutions with 1% improvements. Fig. 7(b) can conclude that the encoder and decoder are necessary for the multi-head self-attention mechanism.

FOG symptom acts on the entire body, and each wearable sensor contributes to the final prediction. As shown in Table 3, we list all the on-body wearable locations, i.e., ankle, thigh, and trunk. Each location corresponds to a wearable sensor. For each location, we treat its time-series samples as input and retain the rest of MCT-Net. Among three wearable locations, the sensor (ankle) above the ankle and at the shank has the closest relationship to the FOG symptom, which obtains 84.46% in the F1-score. However, MCT-Net with all the multi-channel sensor locations beats the result using a single sensor, expressing FOG causes movement disorder in the whole body. The optimal results of MCT-Net validate the comprehensive multi-channel system will perform better with more relative wearable sensors for FOG prediction. The gait representation includes more detailed and diverse temporal inputs, which enhance the proposed model's ability to learn and predict complex gait

Table 3
Experimental results of overall prediction with different wearable locations (%). The best ones are in **red&bold**.

Location	Accuracy	Specificity	Precision	Recall	F1-Score
Ankle	95.24	96.67	85.01	80.31	84.46
Thigh	94.84	96.08	84.15	79.39	81.76
Truck	94.53	95.93	85.23	79.75	83.31
MCT-Net	**97.26**	**99.13**	**87.34**	**81.09**	**87.12**

patterns. Hence, MCT-Net improves the robustness and accuracy of FOG prediction with multiple sensors. Considering the physical burden on patients and the limitation of on-body sensors, a single wearable sensor above the ankle is the best way to prevent Parkinson's patients from falling.

Despite a leap of accuracy in terms of multi-channel FOG prediction in this study, there remain several limitations in the current work. The first concern with MCT-Net is the limited high-quality samples. We train and test our model using offline and real-world FOG data, i.e., the Daphnet FOG dataset. Real-world credible data is scarce, especially for the multi-channel system. They are collected by Parkinson's patients equipped with wearable devices performing random walks. Besides, expert therapists must be present to instruct the patient and care for safety. Hence, it is hard to collect and simulate to enrich the existing samples. The scarcity of credible data will cause an insufficient training process, which restricts the adaptiveness of MCT-Net. The second limitation of MCT-Net is its complexity, especially considering its deployment with limited wearable sensors and computational resources. Our proposed transformer requires quite extra computational cost compared with the single-channel framework. Although Table 3 has validated that MCT-Net is also available with a single sensor, the advantage of MCT-Net cannot reach its full potential without an expensive multi-channel system. The final limitation of MCT-Net is that the current framework is only for temporal data. Multimodal learning of FOG is an urgent area for development, which has also been noted and discussed in recent literature [24].

5. Conclusion

An efficient and credible prediction of FOG will prevent patient who has Parkinson's Disease from falling and improve the quality of life. Based on the wearable sensor system, we deploy a novel multi-channel time-series neural network approach, MCT-Net, to realize freeze events prediction. MCT-Net has two components: causal distributed gait feature representation and the intra-channel and inter-channel transformer. The former employs causal distributed convolutions to integrate various gait features into a unified temporal feature embedding. By the encoder–decoder framework, the latter handles the correlations and diversities by a multi-head self-attention mechanism. By qualitative and quantitative evaluation with multiple metrics on real-world records, our MCT-Net achieves good performance with 96.21% in accuracy and 80.46% in F1-score about 2 s before FOG occurrence (for pre-FOG). It demonstrates that MCT-Net is a significant framework for real-time FOG prediction and to be implemented in remote monitoring devices.

Now, multimodal temporal data is trending in gait pattern recognition. In the future, we want to apply the proposed framework to complicated multimodal datasets to improve motion prediction in Parkinson's Disease.

CRediT authorship contribution statement

Boyan Wang: Writing – original draft. **Xuegang Hu:** Writing – review & editing. **Rongjun Ge:** Validation. **Chenchu Xu:** Writing – review & editing. **Jinglin Zhang:** Methodology. **Zhifan Gao:** Data curation. **Shu Zhao:** Supervision. **Kemal Polat:** Supervision.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

This work was supported in part by The National Natural Science Foundation of China (62106001, U1908211), The University Synergy Innovation Program of Anhui Province, China (GXXT-2021-007), and The Anhui Provincial Natural Science Foundation, China (2208085Y19).

References

[1] Han JW, Ahn YD, Kim W-S, Shin CM, Jeong SJ, Song YS, Bae YJ, Kim J-M. Psychiatric manifestation in patients with Parkinson's disease. J Korean Med Sci 2018;33(47).
[2] Vos T, Allen C, Arora M, Barber RM, Bhutta ZA, Brown A, Carter A, Casey DC, Charlson FJ, Chen AZ, et al. Global, regional, and national incidence, prevalence, and years lived with disability for 310 diseases and injuries, 1990–2015: a systematic analysis for the global burden of disease study 2015. Lancet 2016;388(10053):1545–602.
[3] Jankovic J. Parkinson's disease: clinical features and diagnosis. J Neurol Neurosurg Psychiatry 2008;79(4):368–76.
[4] Bloem BR, Okun MS, Klein C. Parkinson's disease. Lancet 2021;397(10291):2284–303.
[5] Anikwe CV, Nweke HF, Ikegwu AC, Egwuonwu CA, Onu FU, Alo UR, Teh YW. Mobile and wearable sensors for data-driven health monitoring system: State-of-the-art and future prospect. Expert Syst Appl 2022;202:117362.
[6] Tahafchi P, Judy JW. Freezing-of-gait detection using wearable sensor technology and possibilistic k-nearest-neighbor algorithm. In: 2019 41st annual international conference of the IEEE engineering in medicine and biology society. EMBC, IEEE; 2019, p. 4246–9.
[7] Zhang W, Sun H, Huang D, Zhang Z, Li J, Wu C, Sun Y, Gong M, Wang Z, Sun C, et al. Detection and prediction of freezing of gait with wearable sensors in Parkinson's disease. Neurol Sci 2023;1–23.
[8] Sveinbjornsdottir S. The clinical symptoms of Parkinson's disease. J Neurochem 2016;139:318–24.
[9] Thomas I, Westin J, Alam M, Bergquist F, Nyholm D, Senek M, Memedi M. A treatment-response index from wearable sensors for quantifying Parkinson's disease motor states. IEEE J Biomed Health Inform 2017;22(5):1341–9.
[10] Dickson DW, Braak H, Duda JE, Duyckaerts C, Gasser T, Halliday GM, Hardy J, Leverenz JB, Del Tredici K, Wszolek ZK, et al. Neuropathological assessment of Parkinson's disease: refining the diagnostic criteria. Lancet Neurol 2009;8(12):1150–7.
[11] Armstrong MJ, Okun MS. Diagnosis and treatment of Parkinson disease: a review. Jama 2020;323(6):548–60.
[12] Borzì L, Sigcha L, Rodríguez-Martín D, Olmo G. Real-time detection of freezing of gait in Parkinson's disease using multi-head convolutional neural networks and a single inertial sensor. Artif Intell Med 2023;135:102459.
[13] Irrera F, Cabestany J, Suppa A. New advanced wireless technologies for objective monitoring of motor symptoms in Parkinson's disease. Front Neurol 2018;9:216.
[14] Arami A, Poulakakis-Daktylidis A, Tai YF, Burdet E. Prediction of gait freezing in Parkinsonian patients: a binary classification augmented with time series prediction. IEEE Trans Neural Syst Rehabil Eng 2019;27(9):1909–19.
[15] Palmerini L, Rocchi L, Mazilu S, Gazit E, Hausdorff JM, Chiari L. Identification of characteristic motor patterns preceding freezing of gait in Parkinson's disease using wearable sensors. Front Neurol 2017;8:394.
[16] Mohammadian Rad N, Van Laarhoven T, Furlanello C, Marchiori E. Novelty detection using deep normative modeling for imu-based abnormal movement monitoring in Parkinson's disease and autism spectrum disorders. Sensors 2018;18(10):3533.
[17] Borzì L, Mazzetta I, Zampogna A, Suppa A, Olmo G, Irrera F. Prediction of freezing of gait in Parkinson's disease using wearables and machine learning. Sensors 2021;21(2):614.
[18] Torvi VG, Bhattacharya A, Chakraborty S. Deep domain adaptation to predict freezing of gait in patients with Parkinson's disease. In: 2018 17th IEEE international conference on machine learning and applications. ICMLA, IEEE; 2018, p. 1001–6.
[19] Xu C, Xu L, Gao Z, Zhao S, Zhang H, Zhang Y, Du X, Zhao S, Ghista D, Liu H, et al. Direct delineation of myocardial infarction without contrast agents using a joint motion feature learning architecture. Med Image Anal 2018;50:82–94.
[20] Mikos V, Heng C-H, Tay A, Chia NSY, Koh KML, Tan DML, Au WL. Real-time patient adaptivity for freezing of gait classification through semi-supervised neural networks. In: 2017 16th IEEE international conference on machine learning and applications. ICMLA, IEEE; 2017, p. 871–6.
[21] Unni M, Menon P, Livi L, Wilson M, Young W, Bronte-Stewart H, Tsaneva-Atanasova K. Data-driven prediction of freezing of gait events from stepping data. 2020.
[22] Pham TT, Moore ST, Lewis SJG, Nguyen DN, Dutkiewicz E, Fuglevand AJ, McEwan AL, Leong PH. Freezing of gait detection in Parkinson's disease: a subject-independent detector using anomaly scores. IEEE Trans Biomed Eng 2017;64(11):2719–28.
[23] Borzì L, Mazzetta I, Zampogna A, Suppa A, Irrera F, Olmo G. Predicting axial impairment in Parkinson's disease through a single inertial sensor. Sensors 2022;22(2):412.
[24] Mesin L, Porcu P, Russu D, Farina G, Borzì L, Zhang W, Guo Y, Olmo G. A multi-modal analysis of the freezing of gait phenomenon in Parkinson's disease. Sensors 2022;22(7):2613.
[25] Orphanidou NK, Hussain A, Keight R, Lishoa P, Hind J, Al-Askar H. Predicting freezing of gait in Parkinsons disease patients using machine learning. In: 2018 IEEE congress on evolutionary computation. CEC, IEEE; 2018, p. 1–8.
[26] Chen Z, Li G, Gao C, Tan Y, Liu J, Zhao J, Ling Y, Yu X, Ren K, Chen S. Prediction of freezing of gait in Parkinson's disease using a random forest model based on an orthogonal experimental design: a pilot study. Front Hum Neurosci 2021;15:636414.
[27] Kleanthous N, Hussain AJ, Khan W, Liatsis P. A new machine learning based approach to predict Freezing of Gait. Pattern Recognit Lett 2020;140:119–26.
[28] Tanveer M, Rashid AH, Kumar R, Balasubramanian R. Parkinson's disease diagnosis using neural networks: Survey and comprehensive evaluation. Inf Process Manage 2022;59(3):102909.
[29] Sigcha L, Borzì L, Amato F, Rechichi I, Ramos-Romero C, Cárdenas A, Gascó L, Olmo G. Deep learning and wearable sensors for the diagnosis and monitoring of Parkinson's disease: a systematic review. Expert Syst Appl 2023;120541.
[30] Kim HB, Lee HJ, Lee WW, Kim SK, Jeon HS, Park HY, Shin CW, Yi WJ, Jeon B, Park KS. Validation of freezing-of-gait monitoring using smartphone. Telemed e-Health 2018;24(11):899–907.
[31] Filtjens B, Ginis P, Nieuwboer A, Afzal MR, Spildooren J, Vanrumste B, Slaets P. Modelling and identification of characteristic kinematic features preceding freezing of gait with convolutional neural networks and layer-wise relevance propagation. BMC Med Inform Decis Mak 2021;21(1):1–11.
[32] Shi B, Tay A, Au WL, Tan DM, Chia NS, Yen S-C. Detection of freezing of gait using convolutional neural networks and data from lower limb motion sensors. IEEE Trans Biomed Eng 2022;69(7):2256–67.
[33] Hu K, Wang Z, Wang W, Martens KAE, Wang L, Tan T, Lewis SJ, Feng DD. Graph sequence recurrent neural network for vision-based freezing of gait detection. IEEE Trans Image Process 2019;29:1890–901.
[34] Esfahani AH, Dyka Z, Ortmann S, Langendörfer P. Impact of data preparation in freezing of gait detection using feature-less recurrent neural network. IEEE Access 2021;9:138120–31.
[35] Kaur S, Aggarwal H, Rani R. Diagnosis of Parkinson's disease using deep CNN with transfer learning and data augmentation. Multimedia Tools Appl 2021;80(7):10113–39.
[36] Thummikarat H, Chongstitvatana P. An implementation of machine learning for Parkinson's disease diagnosis. In: 2021 18th international conference on electrical engineering/electronics, computer, telecommunications and information technology. ECTI-CON, IEEE; 2021, p. 258–61.
[37] Naghavi N, Wade E. Towards real-time prediction of freezing of gait in patients with Parkinson's disease: a novel deep one-class classifier. IEEE J Biomed Health Inf 2021;26(4):1726–36.
[38] Halder A, Singh R, Suri A, Joshi D. Predicting state-transition in freezing of gait via acceleration measurements for controlled cueing in Parkinson's disease. IEEE Trans Instrum Meas 2021.
[39] Deng K, Li Y, Zhang H, Wang J, Albin RL, Guan Y. Heterogeneous digital biomarker integration out-performs patient self-reports in predicting Parkinson's disease. Commun Biol 2022;5(1):58.
[40] Mazilu S, Hardegger M, Zhu Z, Roggen D, Tröster G, Plotnik M, Hausdorff JM. Online detection of freezing of gait with smartphones and machine learning techniques. In: 2012 6th international conference on pervasive computing technologies for healthcare (PervasiveHealth) and workshops. IEEE; 2012, p. 123–30.
[41] Giannakopoulou K-M, Roussaki I, Demestichas K. Internet of things technologies and machine learning methods for Parkinson's disease diagnosis, monitoring and management: a systematic review. Sensors 2022;22(5):1799.
[42] Kubota KJ, Chen JA, Little MA. Machine learning for large-scale wearable sensor data in Parkinson's disease: Concepts, promises, pitfalls, and futures. Mov Disorders 2016;31(9):1314–26.
[43] Akbar S, Hayat M, Tahir M, Khan S, Alarfaj FK. cACP-DeepGram: classification of anticancer peptides via deep neural network and skip-gram-based word embedding model. Artif Intell Med 2022;131:102349.
[44] Samà A, Rodríguez-Martín D, Pérez-López C, Català A, Alcaine S, Mestre B, Prats A, Crespo MC, Bayés À. Determining the optimal features in freezing of gait detection through a single waist accelerometer in home environments. Pattern Recognit Lett 2018;105:135–43.

[45] Mazilu S, Blanke U, Hardegger M, Tröster G, Gazit E, Hausdorff JM. GaitAssist: a daily-life support and training system for Parkinson's disease patients with freezing of gait. In: Proceedings of the SIGCHI conference on human factors in computing systems. 2014, p. 2531–40.
[46] Polat K. Freezing of gait (fog) detection using logistic regression in Parkinson's disease from acceleration signals. In: 2019 scientific meeting on electrical-electronics & biomedical engineering and computer science. EBBT, Ieee; 2019, p. 1–4.
[47] Silik A, Noori M, Altabey WA, Ghiasi R, Wu Z. Comparative analysis of wavelet transform for time-frequency analysis and transient localization in structural health monitoring. Struct Durab Health Monit 2021;15(1):1.
[48] Zia J, Tadayon A, McDaniel T, Panchanathan S. Utilizing neural networks to predict freezing of gait in Parkinson's patients. In: Proceedings of the 18th international ACM SIGACCESS conference on computers and accessibility. 2016, p. 333–4.
[49] Bahdanau D, Cho KH, Bengio Y. Neural machine translation by jointly learning to align and translate. In: 3rd international conference on learning representations. ICLR 2015, 2015.
[50] Xu C, Xu L, Ohorodnyk P, Roth M, Chen B, Li S. Contrast agent-free synthesis and segmentation of ischemic heart disease images using progressive sequential causal GANs. Med Image Anal 2020;62:101668.
[51] Vaswani A, Shazeer N, Parmar N, Uszkoreit J, Jones L, Gomez AN, Kaiser Ł, Polosukhin I. Attention is all you need. In: Advances in neural information processing systems. 2017, p. 5998–6008.
[52] Shaw P, Uszkoreit J, Vaswani A. Self-attention with relative position representations. In: Proceedings of the 2018 conference of the North American chapter of the association for computational linguistics: human language technologies, volume 2 (short papers). 2018, p. 464–8.
[53] Beltagy I, Peters ME, Cohan A. Longformer: The long-document transformer. 2020, arXiv preprint arXiv:2004.05150.
[54] Han K, Xiao A, Wu E, Guo J, Xu C, Wang Y. Transformer in transformer. Adv Neural Inf Process Syst 2021;34:15908–19.
[55] Dosovitskiy A, Beyer L, Kolesnikov A, Weissenborn D, Zhai X, Unterthiner T, Dehghani M, Minderer M, Heigold G, Gelly S, et al. An image is worth 16x16 words: Transformers for image recognition at scale. In: International conference on learning representations. 2020.
[56] Li B, Yao Z, Wang J, Wang S, Yang X, Sun Y. Improved deep learning technique to detect freezing of gait in Parkinson's disease based on wearable sensors. Electronics 2020;9(11):1919.
[57] Sun R, Hu K, Martens KAE, Hagenbuchner M, Tsoi AC, Bennamoun M, Lewis SJ, Wang Z. Higher order polynomial transformer for fine-grained freezing of gait detection. IEEE Trans Neural Netw Learn Syst 2023.
[58] Xu J, Sun X, Zhang Z, Zhao G, Lin J. Understanding and improving layer normalization. 2019, arXiv preprint arXiv:1911.07013.
[59] Agarap AF. Deep learning using rectified linear units (relu). 2018, arXiv preprint arXiv:1803.08375.
[60] Hamad RA, Kimura M, Yang L, Woo WL, Wei B. Dilated causal convolution with multi-head self attention for sensor human activity recognition. Neural Comput Appl 2021;1–18.
[61] Hershey JR, Olsen PA. Approximating the Kullback Leibler divergence between Gaussian mixture models. In: 2007 IEEE international conference on acoustics, speech and signal processing-iCASSP'07. Vol. 4, IEEE; 2007, p. IV–317.
[62] Bachlin M, Plotnik M, Roggen D, Maidan I, Hausdorff JM, Giladi N, Troster G. Wearable assistant for Parkinson's disease patients with the freezing of gait symptom. IEEE Trans Inf Technol Biomed 2009;14(2):436–46.
[63] Zhang Z. Improved adam optimizer for deep neural networks. In: 2018 IEEE/ACM 26th international symposium on quality of service. IWQoS, IEEE; 2018, p. 1–2.
[64] Fisher NI. Statistical analysis of circular data. cambridge University Press; 1995.