Real-Time Freezing of Gait Detection: Harnessing Advanced AI for Better Mobility

Luyao Yang, Osama Amin, Basem Shihada
Networking Lab, Computer, Electrical and Mathematical Sciences and Engineering
King Abdullah University of Science and Technology
Thuwal, Saudi Arabia
{luyao.yang, osama.amin, basem.shihada}@kaust.edu.sa

***Abstract*—Freezing of gait (FOG) represents a critical and debilitating symptom of Parkinson's disease, posing significant challenges in patient mobility and safety. Numerous research efforts have focused on predicting the onset of FOG using wearable sensors and computational aids. However, the unpredictability and brief nature of FOG episodes complicate the ability to predict their onset in real-time, rendering timely detection a complex goal. In our study, we employed an Autoencoder-Wavenet network architecture designed to effectively utilize data from tri-axial accelerometers and tri-axial gyroscopes positioned at the ankle. This approach allowed for a more nuanced analysis of movement patterns associated with FOG. Our findings indicated that this model successfully achieved a high level of accuracy in detecting FOG, with a specificity of 0.81, and sensitivity of 0.87. Furthermore, the model demonstrated the capability to provide real-time warnings of FOG onset, achieving a specificity of 0.68 and a sensitivity of 0.86. And an accuracy of 0.67 within a 1-second timeframe on the test set. Consequently, these results underscore the potential of our model in contributing to the real-time detection and management of FOG in Parkinson's disease patients.**

***Index Terms*—Parkinson's disease; Freezing of Gait; wearables; accelerometer; gyroscope**

I. Introduction

A. Background

Freezing of gait (FOG) is recognized as one of the most debilitating symptoms of Parkinson's disease. It manifests as a sudden and temporary inability to move the feet forward despite the intention to walk, often described metaphorically as the feeling of one's feet being "glued" to the floor. This phenomenon leads to a significant disruption in the rhythmic and smooth execution of gait, markedly impacting the patient's mobility and substantially increasing the risk of falls, which are a major cause of injury in Parkinson's patients [1].

The implications of FOG are profound as they not only compromise physical safety but also contribute to psychological stress, social isolation, and overall diminished quality of life. Patients may become apprehensive about engaging in physical activities, thereby exacerbating other Parkinson's symptoms due to reduced activity levels. Consequently, the ability to provide advanced warnings before the actual onset of a FOG episode represents a critical enhancement in patient care. Such proactive detection allows for timely interventions, such as cueing strategies that can help in overcoming the freeze [2], or engaging support systems that can assist the patient in managing the episode safely [3].

Technological advancements in wearable sensors coupled with progress in machine learning (ML) and deep learning (DL) have facilitated the widespread adoption of AI-based wearable devices in the domains of sports and healthcare. AI-driven analytics, derived from data collected via wearable sensors, can be smoothly incorporated into the daily routines of patients, offering uninterrupted support independent of clinical environments. This integration not only aids in consistent monitoring but also facilitates the collection of extensive data sets, which are invaluable for enhancing the understanding and management of Parkinson's disease [4].

As a result, an increasing number of AI-based applications have been developed to detect and predict the onset of FOG, with remarkable results. These techniques leverage ML and DL models to analyze signal patterns, such as shortened stride length, increased stride variability, and altered cadence, which are evident in motion data from wearable devices [5]–[11]. This analysis primarily employs ML and DL techniques to identify features indicative of the onset of FOG. A significant limitation of these approaches is their inability to provide accurate real-time warnings for FOG. The primary challenge is precisely predicting the occurrence of FOG within the critical seconds preceding the event, which is essential for effective real-time detection. Consequently, an effective wearable system should encompass two primary functionalities. First, it should be capable of detecting FOG warnings in advance, thereby facilitating the implementation of fall prevention measures. Second, in cases where early warnings are not predicted in time, the system should promptly detect ongoing FOG occurrences to minimize the risk of falls.

In pursuit of achieving these functionalities, our study makes the following principal contributions:

- We propose an Autoencoder-WaveNet architecture that is highly effective for processing FOG signals.
- Our method achieves a sensitivity of 0.68 and a specificity of 0.86 on an independent test set of real-time FOG warnings, surpassing the performance metrics of existing approaches.
- Our method also accurately detects discrete FOG occurrences with a sensitivity of 0.81 and a specificity of 0.87.

II. Related Works

Numerous studies have explored sensor-based prediction of FOG using various methodologies. These approaches include modeling based on signal features extracted from the time and frequency domains [12], applications of wavelet transforms [10], [11], and classifications conducted via ML and DL algorithms such as Convolutional Neural Network (CNN), Support Vector Machine (SVM), Graph Convolutional Network (GCN), Decision Trees (DT) [5]–[9]. Current methodologies in the field predominantly concentrate on the classification or detection of FOG by analyzing signal patterns. These methods have reached a relative maturity in detecting and classifying FOG within controlled datasets, often achieving high levels of sensitivity and specificity. However, there is a noticeable deficiency in the development of systems capable of providing short-term predictive warnings prior to the occurrence of a FOG event. This shortfall underscores a significant opportunity for future research dedicated to advancing real-time predictive capabilities. Enhancing these capabilities could substantially benefit patients by preemptively alerting them to imminent FOG episodes, thereby allowing preventative measures to be taken.

Among these works, two provide the possibility of early warning of FOG. For instance, Reches et al. [5] employed a Radial Basis Function Support Vector Machine (RBF-SVM) alongside leave-one-out cross-validation (LOOV) on a dataset comprising 71 FOG episodes, utilizing inertial measurement unit (IMU) data from both the back and feet. This approach achieved a sensitivity of 84.1% and a specificity of 83.4% of detecting FOG. However, the accuracy of providing real-time warnings for FOG was only 54%, and they didn't have an independent test dataset. In a related study, Borz et al. [6] utilized a CNN combined with differential data methods to detect FOG in real time. This study focused on an independent dataset, analyzing IMU data from the waist and back, and achieved an accuracy rate of 50%. These studies underscore the challenges in achieving high accuracy in the real-time prediction and warning of FOG events, highlighting an area for potential improvement in predictive methodologies.

III. Dataset

This research utilizes a public dataset distinct from other public datasets, which primarily captured FOG episodes during continuous walking states [13], [14]. This unique dataset encompasses data from 35 patients diagnosed with Parkinson's disease, specifically focusing on gait transitions, a critical aspect that has been underrepresented in previous research [15]. It includes recordings of 173 FOG episodes, accumulating a total duration of 1161 seconds of FOG. This dataset resulted in a comprehensive dataset representing three medication phases. To refine our analysis, we excluded phases from the dataset that did not exhibit any FOG attacks. This filtering process resulted in a focused dataset comprising 42 phases.

To prepare the dataset for further analysis, we employed a standard min-max normalization technique to scale the features, which had a significant impact on the distribution

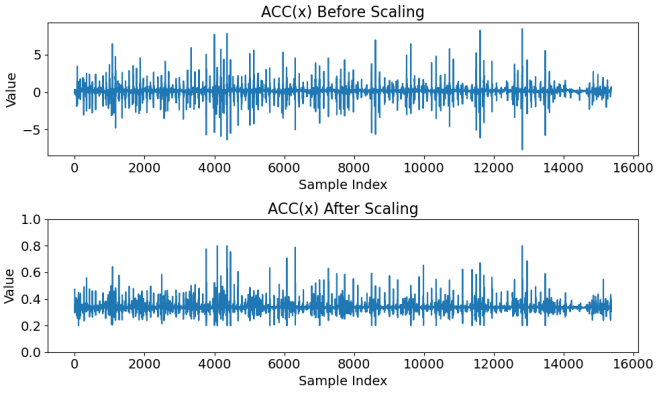

Fig. 1. The data preprocessing

of the accelerometer x-axis features, as illustrated in Figure 1. By applying the min-max scaler, we were able to achieve a more stable and symmetrical feature distribution, effectively minimizing the distortion caused by the presence of outliers in the original data. Furthermore, to address the issue of particularly large noise values, we applied a clipping operation to the normalized data, limiting the values within the 0.1-0.9 range. This additional step helped to reduce the impact of extreme noise on the overall feature distribution, resulting in a more concentrated and evenly distributed dataset that facilitated the subsequent gradient descent-based learning of the predictive models, as the models were less susceptible to the detrimental effects of outliers and noise.

Fig. 2. The definition for different segments

As illustrated in Figure 2, the dataset is partitioned into three primary segments: pre-FOG, FOG, and non-FOG. The duration of the pre-FOG segment ranges from 1 to 3 seconds. Excluding the pre-FOG and FOG segments, the remainder constitutes the non-FOG portion. Given that the non-FOG portion predominantly constitutes the dataset, random down-sampling techniques are employed to achieve a balanced representation among the different categories.

Subsequently, we adopted variable sliding window sizes (1s, 2s, and 3s) to standardize the segmentation of all data into uniform window sizes. Proportional sampling was then employed, based on the quantity and ratio of each class within the dataset, to distribute 80% of the windows to the training set and 20% to an independent test set. This stratagem ensures that both the training and testing sets accurately reflect the overall distribution of classes.

IV. Materials and Methods

A. Methods

In this study, we proposed the AutoEncoderWavenet, an architecture tailored for efficient signal processing and temporal detection in FOG warnings, leveraging autoencoder [16]

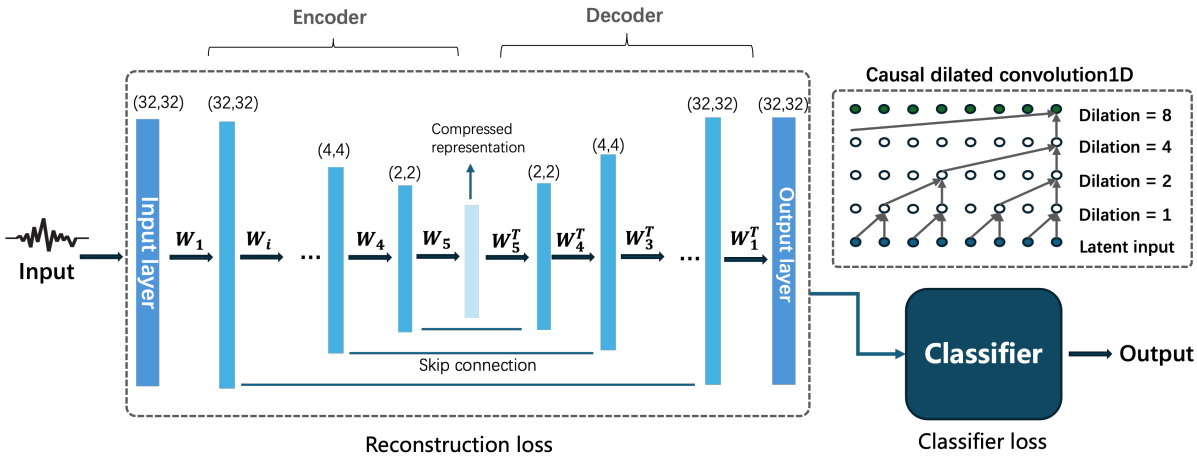

Fig. 3. The proposed autoencoder structure

and dilated causal convolution mechanisms [17] to handle sequence data processing.

As shown in Figure 3, the encoder begins with a input convolutional layer converting six input channels to the hidden channels. Subsequent layers include multiple blocks of causal dilated convolutional layers with increasing dilation rates (2^i). Specifically, the dilation pattern allows the network to exponentially increase the receptive field without a loss in resolution or coverage, enabling the model to capture temporal dependencies across scales. Causal dilated convolutional layers, as depicted in Figure 3, were selected for their ability to preserve the chronological order of temporal data. This characteristic is crucial for the analysis of time-series data, where the sequence and timing of inputs directly influence the accuracy and relevance of the output. Each convolution is followed by ReLU activation and dropout for non-linearity and regularization, respectively.

The decoder layers module mirrors the encoder structure but in reverse order, which aids in reconstructing the original input signal from the condensed feature representation. Additionally, skip layers are incorporated between the encoder and decoder layers to enhance the model's ability to comprehend and capture important information from different levels of the feature hierarchy. These skip connections enable the model to leverage both low-level and high-level features, facilitating more effective and accurate signal reconstruction. The final transformation involves a simple output convolutional layer that maps back to the original six input channels.

The AutoEncoderWaveNet, with its deep causal convolutions and systematic dilation setup, is optimized for signals reconstruction and effective learning of temporal dynamics of different windows of FOG detection.

Following the signal reconstruction through the autoencoder, the classifier was then utilized to further categorize the signal and accurately identify the FOG patterns within the data. This two-step approach enabled a comprehensive analysis of the data. The autoencoder's ability to capture the underlying structure and features of the signal provided a solid foundation, which was then leveraged by the classifier to make the final determination of the FOG episodes.

B. Loss function

The training process of the proposed framework is governed by two distinct loss functions, which work in tandem to guide the model's optimization. The first loss function employed in the training process is the reconstruction loss, which serves as a measure of the model's ability to accurately reconstruct the input signals. For this purpose, we utilize the Mean Squared Error (MSE) as the reconstruction loss metric:

$$L_{\text{reconstruction}} = \frac{1}{N}\sum_{i=1}^{N}(\hat{y}_i - y_i)^2 \tag{1}$$

where $\hat{y}_i$ represents the predicted values, y_i denotes the true values, and N is the number of samples in the dataset. This loss function penalizes the square of the difference between the predicted and actual values, thus heavily penalizing larger errors which is suitable for the windows regression.

In the classifier of the model, the second loss function employed is the cross-entropy loss, which serves to facilitate the accurate classification of the different patterns across the various temporal windows of the input data:

$$L_{\text{entropy}} = -\sum_{i=1}^{N}\sum_{j=1}^{C} p_{ij}\log(p_{ij}) \tag{2}$$

In the given context, N denotes the total number of windows present in the dataset, while C represents the number of categories or classes in the window classification problem. For this particular scenario, C is equal to 3. The training target for the final model was established by computing the sum of two losses, as depicted in Equation 3. The initial values of the coefficients α and β were set to 0.5.

$$L_{\text{total}} = \alpha L_{\text{entropy}} + \beta L_{\text{reconstruction}} \tag{3}$$

C. Criteria for FOG Detection

To evaluate the precision of early warnings, we employ the metric of accuracy. This is quantified by the following equation:

$$\text{Accuracy}_t = \frac{P_t}{T} \tag{4}$$

Where The subscript t indicates the system's capability to make predictions t seconds in advance. P_t represents the number of FOG occurrences accurately predicted i seconds before they happen. And T is the total number of FOG occurrences in the testing dataset.

Sensitivity, or the true positive rate, measures the ability of the FOG detection system to correctly identify actual episodes of FOG:

$$\text{Sensitivity} = \frac{TP}{TP + FN} \tag{5}$$

Where TP represents the number of instances where the model correctly identified a FOG episode. FN represents the number of instances where the model failed to detect an actual FOG episode.

Specificity assesses the ability of the FOG detection system to correctly identify when FOG is not occurring:

$$\text{Specificity} = \frac{TN}{TN + FP} \tag{6}$$

Where TN represents the number of instances where the model correctly identified a non-FOG event. FP represents the number of instances where the model incorrectly identified a non-FOG event as a FOG event.

D. Model Optimization

The optimization process is pivotal for enhancing the performance of the model, which, in turn, significantly influences the outcomes of the experiments. In this study, the AdamW optimizer was employed, incorporated weight decay to prevent overfitting. The initial learning rate was set at 0.001. This choice of optimizer and learning rate is grounded in their proven effectiveness in similar deep learning tasks involving autoencoders and sequence data To optimize the learning process, a ReduceLROnPlateau learning rate scheduler was employed. This scheduler dynamically adjusts the learning rate based on the validation loss performance. When the validation loss does not improve for a specified number of epochs, the learning rate is reduced by a factor of 0.5. This adjustment aims to fine-tune the learning rate and facilitate convergence during training.

E. Environment

All experiments were performed using the pytorch library on a computer equipped with a Tesla V100 32 GB GPU.

V. Results

A. Results

In this section, a quantitative evaluation is conducted to gauge the model's effectiveness in issuing warnings at varying time intervals preceding an emergency event. Specifically, the assessment focuses on the model's performance 1, 2, and 3 seconds prior to the anticipated emergency occurrence. Additionally, experiments are carried out using different window lengths, utilizing the CNN model employed by Borz et al. [6], in order to validate the model's performance. The evaluation comprises an examination of the sensitivity and specificity metrics for each category, alongside the accuracy of pre-FOG warnings. These metrics are computed based on an independent test dataset and are presented in Table I.

We observed that the model performed best when it issued a FOG warning 2 seconds before the FOG actually occurred. This performance was achieved using a window size of 256 and a step size of 128. The sensitivity of the model, which represents its ability to correctly detect positive cases, was measured at 0.68. Similarly, the specificity of the model, which reflects its accuracy in identifying negative cases, was determined to be 0.86. It is noteworthy that any deviation in the prediction time of the other two models, either 1 second ahead or 1 second behind, increased the sensitivity by 0.11, the specificity by 0.07, and the accuracy by 0.11. Interestingly, the warning capabilities of the models 1 second ahead and 3 seconds ahead were statistically comparable in terms of performance, while the CNN model outperformed our model in 1 second and 3 seconds ahead, while our model showed a certain advantage when the window size was increased.

TABLE I
PERFORMANCE OF DIFFERENT PRE_FOG DETECTION MODELS

Models	Sensitivity	Specificity	Accuracy
Model($\leq 1s$)	0.56	0.78	0.56
Model($\leq 3s$)	0.57	0.79	0.55
CNN($\leq 1s$)	0.6	0.79	0.6
CNN($\leq 2s$)	0.62	0.8	0.61
CNN($\leq 3s$)	0.64	0.82	0.63
Model($\leq 2s$)	**0.68**	**0.86**	**0.67**

The performance evaluation of the optimal model, with a lead time of less than 2 seconds, was further conducted by assessing its sensitivity and specificity across different segments. The detailed performance results can be found in Table II. The findings demonstrate that the model exhibits a strong capability to accurately predict the pre_FOG, as evidenced by a sensitivity of 0.64 and a specificity of 0.89. Out of the 264 windows generated in the test set, the model successfully made accurate judgments for 177 windows, resulting in an overall accuracy rate of 0.67.

B. Ablation study

In this section, we investigate the influence of different window sizes and step sizes on the performance of the model. For the sensor operating at a frequency of 128Hz, a window size of 128 corresponds to a duration of 1 second. The moving

TABLE II
PERFORMANCE OF MODEL($\leq 2s$) ON INDEPENDENT TESTING DATASET

Model	Sensitivity	Specificity
Pre_FOG	0.64	0.89
FOG	0.81	0.87
Non_FOG	0.6	0.78
Balanced	**0.68**	**0.86**

step, commonly set to half the window size, is therefore 64, equivalent to 0.5 seconds. In the case of the model with a lead time of 1 second, we employed windows of 1 second and 0.5 seconds, respectively. Likewise, for the model with a lead time of 2 seconds, windows of 1 second and 2 seconds were utilized. Similarly, for the model with a lead time of 3 seconds, windows of 1 second, 2 seconds, and 3 seconds in advance were employed. The balanced sensitivity and specificity for all categories, along with the accuracy of pre-FOG warnings, were computed and are presented in Table III.

TABLE III
PERFORMANCE FOR DIFFERENT WINDOW SIZE AND STEP SIZE

Models	Step	Window	Sensitivity	Specificity	Accuracy
Model($\leq 1s$)	32	64	0.43	0.61	0.42
Model($\leq 1s$)	64	128	0.58	0.78	0.57
Model($\leq 2s$)	64	128	0.58	0.79	0.61
Model($\leq 2s$)	128	256	0.68	0.86	0.67
Model($\leq 3s$)	64	128	0.55	0.77	0.53
Model($\leq 3s$)	128	256	0.57	0.79	0.58
Model($\leq 3s$)	256	512	0.62	0.81	0.59

Based on the findings presented in the Table III, it is evident that the performance of each model is optimized when the window size is set to its maximum feasible value. For instance, in the case of the model with a lead time of 2 seconds, the most effective outcome is achieved when the window size is set to 2 seconds, corresponding to a value of 256. This trend is also observed for the 1-second and 3-second-ahead models, where the best performance is attained when the window size aligns with the respective lead time. Conversely, when the window size is exceptionally small, the model exhibits notably poor performance. For instance, in the 1-second-ahead model, setting the window size to half a second results in prediction accuracy that falls below the level of random probability. This observation leads us to speculate that this phenomenon may be attributed to the inherent characteristics of the model itself. The utilization of dilated convolution enables the model to capture information from a broader temporal context, thus implying that a larger window size facilitates the convolution process in acquiring more comprehensive contextual information.

We also carried out ablation experiments to study the influence of different sensor data sources on the model's performance. Specifically, we evaluated the models using accelerometer data, gyroscope data, and the combination of both accelerometer and gyroscope data as input features. These experiments allowed us to gain insights into the relative importance and complementary nature of the sensor modalities.

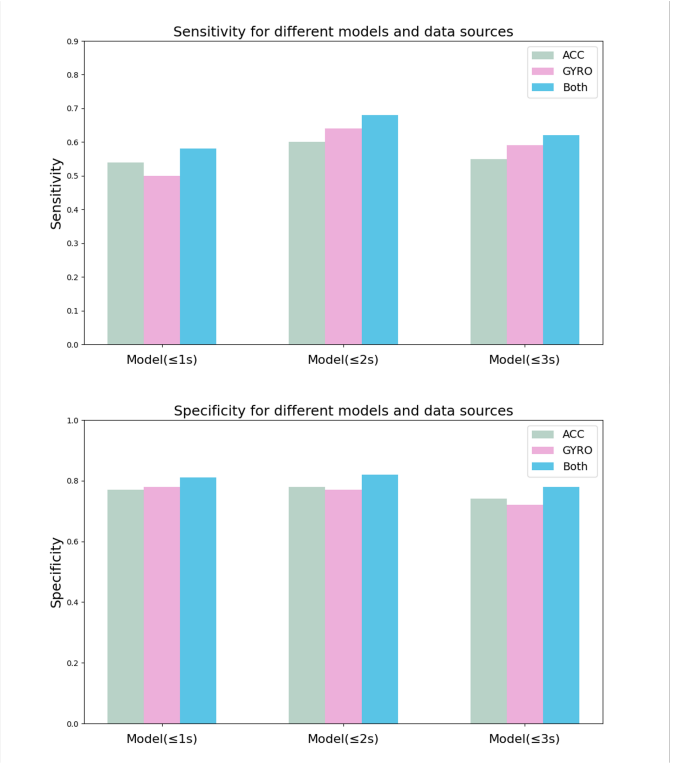

Fig. 4. Sensitivity and specificity under different kinematic data sources

As shown in Figure 4, the model trained on the multimodal data, which combines the accelerometer and gyroscope signals, achieved the highest sensitivity across the three models. This finding is reasonable, as incorporating higher-dimensional information from both motion and angular measurements can facilitate the model's acquisition of more comprehensive knowledge about the target phenomena. We also observed that the gyroscope data had a greater influence on model sensitivity compared to the accelerometer data. We attribute this to the nature of our dataset, which consists of transition recordings of patients. These ablation study results highlight the importance of considering multimodal sensor data, as the fusion of complementary signal sources can lead to improved model performance, especially for tasks involving complex kinematic patterns.

VI. DISCUSSION & CONCLUSIONS

A. Discussion

This study proposes a comprehensive network architecture that can directly process data from wearable sensors of users and generate appropriate prompts and warnings. Given that the data source includes daily wearable sensor measurements, not just FOG detection, the model has the potential to be further optimized and applied to the field of fall detection and prevention in the elderly, as well as monitoring and detecting physical recovery after surgery. Therefore, in the future, as the available datasets expand, it is expected that the model can be

further generalized to adapt to a wider range of personal health management and disease rehabilitation applications.

This paper first attempts to use autoencoders to perform unsupervised reconstruction and denoising of motion signal sequences, as well as extract relevant feature information. However, in order to more accurately detect the precursors and occurrence of FOG, we subsequently adopted a supervised classification training method based on the output of the autoencoder model. While this approach is effective, it is limited by the need for labeled data, as existing methods. As wearable sensor technology becomes more and more common, the amount of patient data generated will grow exponentially. Domain experts need to observe patients' motion data and manually label them, which is resource-intensive. Therefore, developing unsupervised or semi-supervised modeling techniques is essential to fully realize the potential of these large-scale unlabeled datasets.

Designing effective FOG detection models requires a careful balance between sensitivity and specificity. High sensitivity, which ensures reliable warning of impending FOG events, often comes at the cost of reduced specificity. To address this trade-off, future research protocols or performance metrics may be proposed to measure and optimize the balance between these two crucial metrics.

B. Conclusion

In this study, we employed an end-to-end autoencoder network to predict the warnings and onset of FOG episodes. This approach produced significant results when evaluated on an independent test set. Compared to prior methods, the present study not only aimed to predict the occurrence of FOG episodes, but also incorporated regression-based training focused on the pre-FOG phases. This expanded approach sought to provide more granular insights into the precursory indicators of FOG, rather than solely relying on binary classification of the onset itself. Our results showed that during testing, the best warning effect was achieved about two seconds before the event. With the expected advancement of technology, patients are expected to increasingly use smart wearable devices that can predict and alert them to impending FOG episodes in real-time, thereby mitigating significant risks.

The future research agenda of this work will concentrate on further refining the model architecture to enable high-dimensional, multimodal analysis. Specifically, upcoming efforts will seek to integrate patient-specific data, such as medication schedules, clinical records, and other relevant parameters, with the aim of enhancing the predictive accuracy and broader applicability of the model across diverse patient populations. It is anticipated that these advancements will significantly improve the development of personalized management and intervention strategies for individuals suffering from Parkinson's disease. Furthermore, the sensor data collected are not limited to FOG warnings and detection, but can also be leveraged for gait analysis and rehabilitation training purposes, thereby expanding the scope and utility of the proposed approach.

REFERENCES

[1] John G Nutt, Bastiaan R Bloem, Nir Giladi, Mark Hallett, Fay B Horak, and Alice Nieuwboer. Freezing of gait: moving forward on a mysterious clinical phenomenon. *The Lancet Neurology*, 10(8):734–744, 2011.

[2] Alice Nieuwboer. Cueing for freezing of gait in patients with parkinson's disease: a rehabilitation perspective. *Movement disorders: official journal of the Movement Disorder Society*, 23(S2):S475–S481, 2008.

[3] Sinziana Mazilu, Ulf Blanke, Michael Hardegger, Gerhard Tröster, Eran Gazit, and Jeffrey M Hausdorff. 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*, pages 2531–2540, 2014.

[4] Luyao Yang, Osama Amin, and Basem Shihada. Intelligent wearable systems: Opportunities and challenges in health and sports. *ACM Computing Surveys*, 56(7):1–42, 2024.

[5] Tal Reches, Moria Dagan, Talia Herman, Eran Gazit, Natalia A Gouskova, Nir Giladi, Brad Manor, and Jeffrey M Hausdorff. Using wearable sensors and machine learning to automatically detect freezing of gait during a fog-provoking test. *Sensors*, 20(16):4474, 2020.

[6] Luigi Borzì, Luis Sigcha, Daniel Rodríguez-Martín, and Gabriella Olmo. Real-time detection of freezing of gait in parkinson's disease using multi-head convolutional neural networks and a single inertial sensor. *Artificial intelligence in medicine*, 135:102459, 2023.

[7] Florenc Demrozi, Ruggero Bacchin, Stefano Tamburin, Marco Cristani, and Graziano Pravadelli. Toward a wearable system for predicting freezing of gait in people affected by parkinson's disease. *IEEE journal of biomedical and health informatics*, 24(9):2444–2451, 2019.

[8] Helena Cockx, Jorik Nonnekes, Bastiaan R Bloem, Richard van Wezel, Ian Cameron, and Ying Wang. Dealing with the heterogeneous presentations of freezing of gait: how reliable are the freezing index and heart rate for freezing detection? *Journal of neuroengineering and rehabilitation*, 20(1):53, 2023.

[9] Scott Pardoel, Gaurav Shalin, Julie Nantel, Edward D Lemaire, and Jonathan Kofman. Early detection of freezing of gait during walking using inertial measurement unit and plantar pressure distribution data. *Sensors*, 21(6):2246, 2021.

[10] Catalina Punin, Boris Barzallo, Roger Clotet, Alexander Bermeo, Marco Bravo, Juan Pablo Bermeo, and Carlos Llumiguano. A non-invasive medical device for parkinson's patients with episodes of freezing of gait. *Sensors*, 19(3):737, 2019.

[11] Saba Rezvanian and Thurmon E Lockhart. Towards real-time detection of freezing of gait using wavelet transform on wireless accelerometer data. *Sensors*, 16(4):475, 2016.

[12] Y Zhao, Karin Tonn, Khalil Niazmand, Urban M Fietzek, Lorenzo T D'Angelo, A Ceballos-Baumann, and Tim C Lueth. Online fog identification in parkinson's disease with a time-frequency combined algorithm. In *Proceedings of 2012 IEEE-EMBS International Conference on Biomedical and Health Informatics*, pages 192–195. IEEE, 2012.

[13] Marc Bachlin, Meir Plotnik, Daniel Roggen, Inbal Maidan, Jeffrey M Hausdorff, Nir Giladi, and Gerhard Troster. Wearable assistant for parkinson's disease patients with the freezing of gait symptom. *IEEE Transactions on Information Technology in Biomedicine*, 14(2):436–446, 2009.

[14] Sinziana Mazilu, Ulf Blanke, Daniel Roggen, Gerhard Tröster, Eran Gazit, and Jeffrey M Hausdorff. Engineers meet clinicians: augmenting parkinson's disease patients to gather information for gait rehabilitation. In *Proceedings of the 4th augmented human international conference*, pages 124–127, 2013.

[15] Caroline Ribeiro De Souza, Runfeng Miao, Júlia Ávila De Oliveira, Andrea Cristina De Lima-Pardini, Débora Fragoso De Campos, Carla Silva-Batista, Luis Teixeira, Solaiman Shokur, Bouri Mohamed, and Daniel Boari Coelho. A public data set of videos, inertial measurement unit, and clinical scales of freezing of gait in individuals with parkinson's disease during a turning-in-place task. *Frontiers in Neuroscience*, 16:832463, 2022.

[16] Cheng-Yuan Liou, Wei-Chen Cheng, Jiun-Wei Liou, and Daw-Ran Liou. Autoencoder for words. *Neurocomputing*, 139:84–96, 2014.

[17] Aaron van den Oord, Sander Dieleman, Heiga Zen, Karen Simonyan, Oriol Vinyals, Alex Graves, Nal Kalchbrenner, Andrew Senior, and Koray Kavukcuoglu. Wavenet: A generative model for raw audio. *arXiv preprint arXiv:1609.03499*, 2016.