Parkinson's Disease Freezing of Gait (FOG) Symptom Detection Using Machine Learning from Wearable Sensor Data

by

Mahmudul Hasan
24341117

A thesis submitted to the Department of Computer Science and Engineering
in partial fulfillment of the requirements for the degree of
B.Sc. in Computer Science

Department of Computer Science and Engineering
Brac University
November 2024

Declaration

It is hereby declared that

1. The thesis submitted is my/our own original work while completing degree at Brac University.
2. The thesis does not contain material previously published or written by a third party, except where this is appropriately cited through full and accurate referencing.
3. The thesis does not contain material which has been accepted, or submitted, for any other degree or diploma at a university or other institution.
4. We have acknowledged all main sources of help.

Student's Full Name & Signature:

Mahmudul Hasan
24341117

Approval

The thesis titled "Parkinson's Disease Freezing of Gait (FOG) Symptom Detection Using Machine Learning from Wearable Sensor Data" submitted by

1. Mahmudul Hasan(24341117)

Of Summer, 2024 has been accepted as satisfactory in partial fulfillment of the requirement for the degree of B.Sc. in Computer Science on November 30, 2024.

Examining Committee:

Supervisor:
(Member)

Dr. Md. Golam Robiul Alam
Professsor
Department of Computer Science and Engineering
Brac University

Thesis Coordinator:

Dr. Md. Golam Robiul Alam
Professsor
Department of Computer Science and Engineering
Brac University

Head of Department:
(Chair)

Sadia Hamid Kazi, PhD
Chairperson and Associate Professor
Department of Computer Science and Engineering
Brac University

Abstract

Freezing of gait (FoG) is a special symptom found in patients with Parkinson's disease (PD). Patients who have FoG abruptly lose the capacity to walk as they normally would. Accelerometers worn by patients can record movement data during these episodes, and machine learning algorithms may be able to categorize this information. Thus, the combination may be able to identify FoG in real time. In order to identify FoG events in accelerometer data, we introduce transformer encoder-Bi-LSTM fusion and transformer encoder-GRU fusion models in this study. The model's capability to differentiate between FoG episodes and normal movement was used to evaluate its performance, and on the Kaggle freezing of gait dataset, the proposed transformer encoder-Bi-LSTM fusion model produced better results with 52.06% compared to combination of transformer encoder and GRU with 49.38% in respect of mean average precision. The findings highlight how Deep Learning-based approaches may progress the field of FoG identification and help PD patients receive better treatments and management plans.

Keywords: Deep Learning, Time Series Analysis, Parkinson's Disease

Table of Contents

List of Figures

List of Tables

Nomenclature

Several symbols and abbreviations that will be used later in the document's body are described in the following list.

AP Average Precision

$BERT$ Bidirectional Encoder Representations from Transformers

EMG Electromyography

FI Freezing Index

FN False Negative

FoG Freezing of Gait

FP False Positive

GRU Gated Recurrent Unit

IoU Intersection over Union

$LSTM$ Long Short Term Memory

mAP Mean Average Precision

ML Machine Learning

PD PD

RNN Recurrent Neural Network

SVM Support Vector Machine

TN True Negative

TP True Positive

Chapter 1

Introduction

The second most widespread neurodegenerative illness and the one with the greatest rate of increase in frequency, associated disability, and mortality is Parkinson's disease (PD) [37]. 1-2% of people over 65 have PD, and as the people ages, its prevalence is rising quickly [4]. It affects the nervous system that causes physical discomfort, disruptions in sleep patterns, mental health issues, and various other health complications. A survey of 6620 PD patients revealed that 28% of them reported FoG every day, and 47% reported regular FoG [14]. FoG is an abnormal gait pattern characterized by unforeseen, episodic, transient episodes of gait inhibition marked by the incapability to propel the feet ahead in spite of the individual's desire to ambulate. It usually occurs suddenly and is brief, with the motor system being halted for a few seconds to a few minutes [24]. When experiencing FOG, PD patients frequently report feeling despite that their legs are glued to the floor for no apparent reason [7]. FOG is less common in patients whose primary symptom is tremor and more regular in men than in women [5]. Evaluating FoG in the hospital is tough, as the occurrence and patterns of freezing episodes differ from where sufferers perform their daily activities for living. In most situations, patients exhibit a minimal amount of freezing in the hospital, whereas their caretakers claim they freeze extensively at Residence [3]. The reason for this is that the activations for FoG episodes (for example, certain locations at home such as walk-in kitchens, crossing the road) are difficult to locate in the clinic's small area. Schaafsma et al. have identified five varieties of freezing: open space hesitation, destination hesitation, start hesitation, turn hesitation, and hesitation in confined spaces [7]. For patients, FOG has significant social and medical repercussions. It seriously lowers quality of life [1], disrupts everyday activities, and is a frequent cause of falls [8]. Therefore, developing methods that can aid in lowering the prevalence of FoG is crucial.

Treatment with pharmaceuticals does not work effectively for FoG. Levodopa (LD) is the most often prescribed medication for PD patients' motor symptoms. The duration of LD's effects on parkinsonism symptoms ranges from two to six hours, and they gradually fade off. Some individuals experience a slow decline in motor function as a result of this wearing off effect, whereas others experience a sudden and rather severe decline. It is possible to identify distinct ON and OFF phases for these people, where ON periods denote when the drug is working and OFF periods denote when it isn't. The right duration of each medication dose decreases with the progression of the disease, necessitating more frequent LD administration

[5]. Gait abnormalities in PD patients are commonly unresponsive to medication, despite the fact that FOG episodes typically occur more frequently during the OFF state [8]. Therefore, in order to alleviate symptoms and enhance mobility, effective non-pharmacologic treatments must be created as an adjuvant therapy. Clinical research, however, indicates that rhythmic cueing that is in sync with the gait—such as metronome ticking sounds or periodic lines projected on the floor—can assist patients in breaking out of the freezing condition and starting to walk again ([10], [15], [23]). Body-worn accelerometers ([19], [29]) can be used by wearable devices to detect FoG and, when detected, provide a rhythmic cue. In their thorough analysis of how external rhythmical cueing effects in PD patients' gait, Lim et al. [11] discovered compelling evidence that using auditory signals can increase walking speed. The usefulness of visual and somatosensory cueing was not sufficiently supported by the findings. Likewise, Nieuwboer et al. demonstrated the benefits of auditory cueing over visual and tactile cueing [20]. It has been demonstrated that RAS is especially successful in helping PD patients walk more easily [13]. However, compared to PD patients without FOG, there was no discernible benefit to using this technique to enhance gait in PD patients who also had FOG [20]. It's interesting to note that a research in which PD and FOG patients cued at home using metronome recordings revealed no reduction in freezing symptoms [9].

Mental health issues may have a significant influence in the pathophysiology of FoG, which is a known fact that has not yet been fully utilized [12]. Stress and anxiety are associated with and probably contribute to the incidence of FoG [17]. When a person feels a strong need to act quickly, such as when they have little time to get on an elevator before the doors shut, the likelihood of freezing may become more noticeable. Therefore, by utilizing wearable sensory device technology, it is important to track patient gait and obtain an accurate estimate of the occurrence and intensity of freezing episodes experienced by a patient in their usual environment. Having a precise figure for the number and intensity of these episodes could help clinicians accurately select the quantity of levodopa or adjust the settings of brain stimulation therapy. Machine perception of freezing events could also be utilized to enable and alter real-time therapeutic techniques.

1.1 Research Objective

Utilizing ML techniques, the aim of this research is to identify the beginning and end of each freezing episode, besides the incidence of the following three categories of freezing gait events-walking, turning, and start hesitation. The aims of this study include:
1. Detect the start and stop of each freezing episode and the occurrence in series of three types of FoG events.
2. Improve the ability of medical professionals to optimally evaluate, monitor, and ultimately, prevent FoG events.
3. Participate in the creation of novel techniques and resources for the early identification and monitoring of cognitive and motor decline.
4. Help researchers better understand when and why FoG episodes occur.

1.2 Research Contribution

In this thesis, we design and develop a machine learning (ML) model to detect and classify FoG in patients with PD using wearable sensor data. When wearable sensor devices are used with ML methods, the accuracy of detecting FOG from a lower back accelerometer is relatively high. The key contributions of this research are as follows:

- We proposed a Transformer Encoder-Bi-LSTM fusion model to detect freezing of gait where Transformer Encoder was utilized for feature extraction and Bi-LSTM for classifying the type of FoG event.
- We have tested our model on two benchmark datasets. The proposed model outperformed the state-of-the-art models without extensive data preprocessing and ensemble multiple models.

1.3 Report Organization

The remainder of this report is structured as follows: Chapter 2 provides a comprehensive review of the literature relevant to the field. Chapter 3 outlines the background study and details the architecture of the proposed model. Chapter 4 describes the methodology employed, including a thorough discussion of the dataset and the proposed model. Chapter 5 presents the performance evaluation metrics and the experimental results. Finally, Chapter 6 concludes the paper with a summary of findings and potential future directions.

Chapter 2

Related Work

Early attempts at identifying FoG episodes relied on the concept of acceleration signals. Progressive work by Han et al. in 2003 utilized a 4-level Daubechies wavelet transform on acceleration data from PD patients' ankles [6]. They discovered a distinct shift in the main frequency of the signal during freezing events, rising from 2 Hz during normal movement to a range of 6 to 8 Hz. Building on this, Moore et al. in 2008 introduced the freezing index (FI) [16]. This index calculated the energy distribution within specific frequency bands of a single vertical acceleration signal from the shank. The FI employed a threshold classifier to distinguish between normal movement and freezing based on the energy distribution. Bächlin et al. further refined detection in 2010 by incorporating an additional criterion that differentiated standing from freezing based on the overall energy disparity in the signal [18].

However, these accelerometer-based methods had limitations. They primarily identified freezing patterns resembling tremors, which while common, don't encompass all FoG presentations. Recognizing this, researchers explored incorporating additional sensors and more sophisticated classifiers. Coste et al. in 2014 demonstrated the limitations of relying solely on acceleration energy [28]. They proposed a new benchmark by utilizing a shank-mounted sensor to measure stride length and cadence alongside acceleration. These additional gait characteristics, combined with a threshold classifier, improved FoG detection accuracy. Another approach involved combining different sensor modalities. Cole et al. in 2011 employed accelerometers and EMG sensors positioned on the shank, thigh, and arm to detect FoG [22]. They utilized a two-stage classification system. First, a linear classifier distinguishes between sitting/lying and standing. Then, a Dynamic Neural Network classifier identified freezing during movement. Mazilu et al. further emphasized the value of multimodal approaches, suggesting that incorporating more information could enhance detection speed and accuracy [25]. Cappeci et al. [31] echoed this sentiment, demonstrating improved results when integrating gait factors into deep learning models for FoG detection. For a more comprehensive exploration of FoG identification research, refer to [34].

Several studies explored effective machine learning techniques for FoG detection in real-world settings. Ahlrichs et al [30]. introduced a method for monitoring FoG occurrences in home environments. They extracted various frequency-based features from 3.2-second data segments. Analysis using a SVM classifier with 10-fold

cross-validation yielded impressive results, achieving a sensitivity of 92.3% and a perfect specificity (100%). Rodríguez-Martín et al. [26] explored FoG detection in daily activities using an SVM classifier with 55 features encompassing statistical and spectral characteristics. Their LOSO evaluation yielded a sensitivity of 0.7903 and a specificity of 0.7467. Samà et al. [38] further optimized this approach by lessening the number of attributes while maintaining high performance (sensitivity of 91.81% and specificity of 87.45%).

Recent research has successfully employed deep learning models for FoG detection using sensor data. CNNs and fully connected neural networks have shown promising results. Camps et al. [36] proposed a deep learning architecture with eight levels incorporating 1D convolutional layers for FoG identification. Their model achieved a precision rate of 90.6%, an AUC of 88%, a sensitivity of 0.9190, and a specificity of 0.8950. San-Segundo et al. [42] compared various classification methods, including Random Forest, Multilayer Perceptron, and deep CNN, for FoG detection on the Daphnet dataset. The deep convolutional neural network emerged as the most effective method, achieving an Area Under the Curve (AUC) of 0.931 and an EER of 12.5%. Sigcha et al. [43] explored using Recurrent Neural Networks (RNNs) with 3D accelerometers for FoG detection. Their evaluation, using a LOSO cross-validation approach, revealed that the CNN-LSTM model achieved the finest performance with a sensitivity and specificity of 87.1% each, using a three-step look-back window. This review highlights the ongoing advancements in FoG detection techniques. From early explorations with acceleration signals to sophisticated deep learning models, researchers are continuously refining methods to achieve more accurate and reliable identification of this debilitating symptom.

The development of accurate FoG detection models is becoming more and more popular with the introduction of wearable sensors and the progress made in ML. Wearable sensors that track gait objectively and continuously include magnetometers, gyroscopes, and accelerometers. They offer a non-invasive method to capture the intricate details of movement, making them ideal for detecting FoG episodes. Numerous studies have leveraged data from these sensors to develop algorithms for FoG detection. Mancini et al. (2012) demonstrated the feasibility of using accelerometer data to distinguish FoG episodes from normal walking in PD patients [33]. Similarly, Moore et al. (2008) highlighted the potential of wearable sensors to monitor gait and predict FoG, underscoring the need for accurate and real-time detection methods [16]. Traditionally used machine learning techniques including SVM, KNN, and Decision Trees were used in the early research on FoG identification. For instance, Rodríguez-Martín et al. (2017) used SVM to classify FoG episodes based on features extracted from accelerometer data, achieving promising results. However, these models often require extensive feature engineering and may not capture the temporal dependencies in gait data effectively [26]. Recurrent Neural Networks (RNNs) and their derivatives are a class of deep learning models that have demonstrated tremendous promise in processing sequential data, such as gait signals. RNNs are useful for FoG detection because they can record temporal dependencies. An extension of LSTM called Bi-LSTM analyzes data both forward and backward to give a thorough grasp of the temporal context. For example, Hannink et al. (2017) utilized an LSTM network to classify different gait patterns, including

FoG, using wearable sensor data [32]. Their results demonstrated the superiority of LSTM over traditional methods in capturing the temporal dynamics of gait.

Machine vision is an evaluative technology that uses a machine to see instead of a human. Machine vision-based approaches are more objective than sensor wearables because they don't require the patient to wear a gadget and don't interfere with their movement or cause discomfort. A few studies have used the Kinect depth camera, which is a 3D motion capture system, to extract motion data for the purpose of identifying gait abnormalities [39]. However, these cameras need pricey specialized equipment like the Microsoft Kinect. Using RGB technology for 2D key point recognition, as exemplified by OpenPose [41], is an additional technique that may estimate the joint coordinates of individuals in films captured with a monocular camera without the need for external scales or markers. In order to identify FOG using 2D keypoint estimation, a study suggested a unique design for a graph convolutional neural network and obtained good detection performance [40].

Chapter 3

Background Study

3.1 Transformer

Vaswani et al., Google researchers, initially presented the transformer design in their 2017 publication "Attention Is All You Need" [35]. An encoder and a decoder, each consisting of several layers of feed forward and self-attention neural networks, make up the transformer architecture. The self-attention mechanism, at the core of the transformer, allows the model to evaluate the relative relevance of many words in a sentence by taking into account their mutual affinities. This is comparable to how a human could read a phrase, concentrating on the most important details rather than reading it in a straight line from start to finish. To help the model remember the comparative positions of words in a sentence, the transformer adds positional bias in addition to self-attention. This is essential because a sentence's meaning can be greatly influenced by the order in which its words are used.

3.1.1 Transformer Encoder

For applications like text classification, where the model must assign a text passage to one of several predetermined categories, such sentiment analysis, subject classification, or spam detection, the transformer encoder architecture is employed. A whole series of tokens is fed into the encoder, which creates a fixed-size vector representation of the sequence that can be used for categorization. In 2018, Google unveiled BERT, one of the most widely used transformer encoder models. Large volumes of text data have already been used to train BERT, which may be adjusted for a variety of NLP tasks. The transformer encoder does not produce any output sequence; instead, it just cares about the input sequence, in contrast to the encoder-decoder architecture. By applying a self-attention mechanism to the input tokens, it is able to focus on the portions of the input that are most relevant to the recent task. Examples of the transformer encoder architecture in action include email spam detection, where the model must determine if an email is spam or not, and sentiment analysis, where the model must determine whether a given review is good or negative.

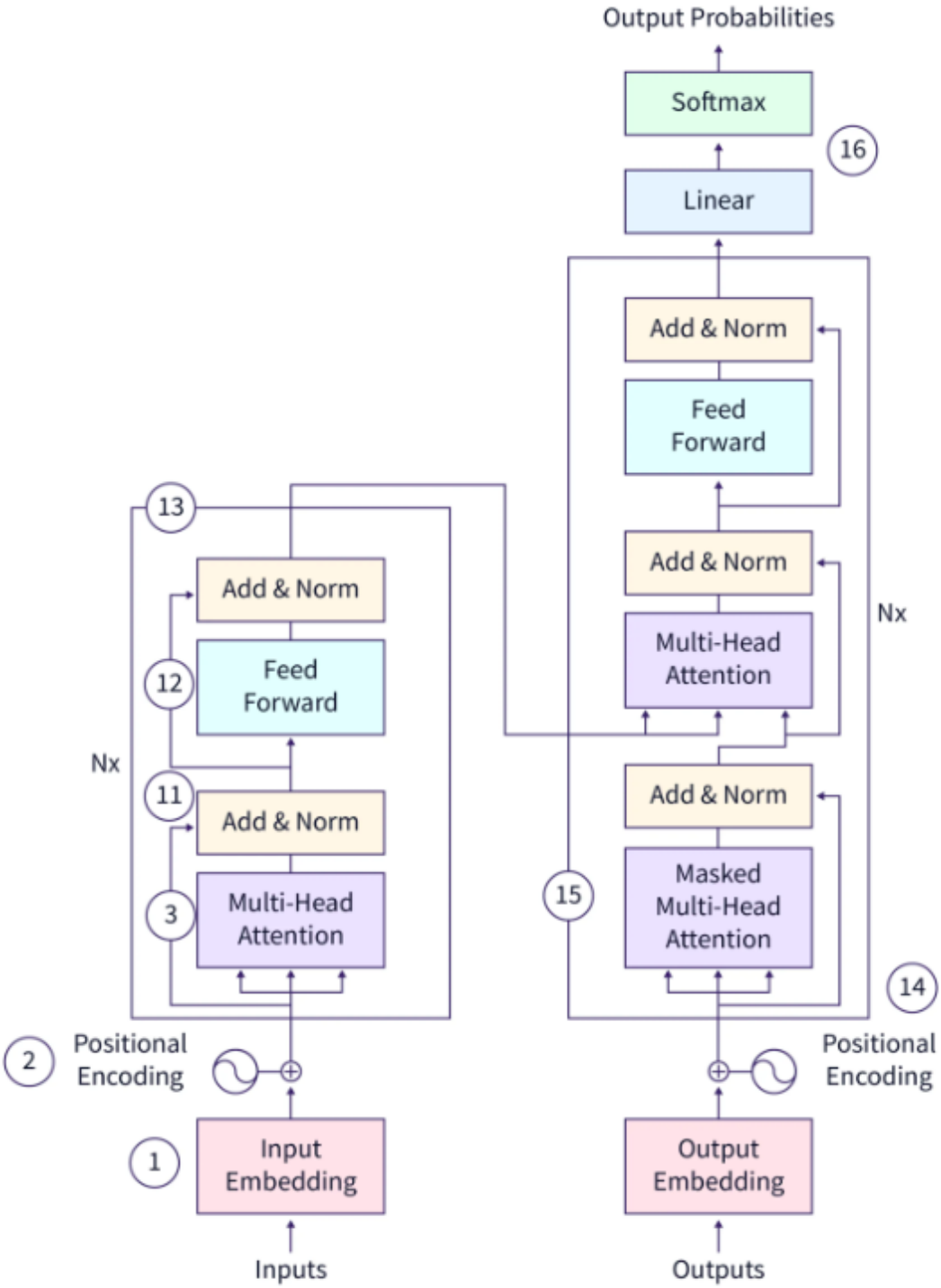

Figure 3.1: Architechture of Transformer [35]

3.2 Bi-LSTM

Recurrent neural networks of the Bi-LSTM type process sequential data both forward and backward [2]. The model is able to consider the input sequence's past and future context by fusing bidirectional processing with LSTM capabilities. To enable the progressive selective retention and forgetting of information, memory cells and gating mechanisms are incorporated. Long-term dependencies spanning several time steps can be captured by LSTMs due to their persistent internal memory state. Bi-LSTM architecture is made up of two LSTM layers that process the sequence in forward and backward. Each layer maintains its own memory cells and secret states. Beginning with the initial time step and continuing until the forward pass is complete, the input sequence is transmitted to the forward LSTM layer. The forward LSTM improves its memory cell depending on the recent input, the preceding hidden state, and the memory cell in each time step. Furthermore, the input sequence is delivered to the backward LSTM layer concurrently in backward order, from the last time step to the beginning. Alike the forward pass, the reverse LSTM determine its hidden state and improve its memory cell. The hidden states from both LSTM layers are integrated at each time step when the forward and backward passes are finished. Applying an extra transformation or combining the hidden states will work for this combo.

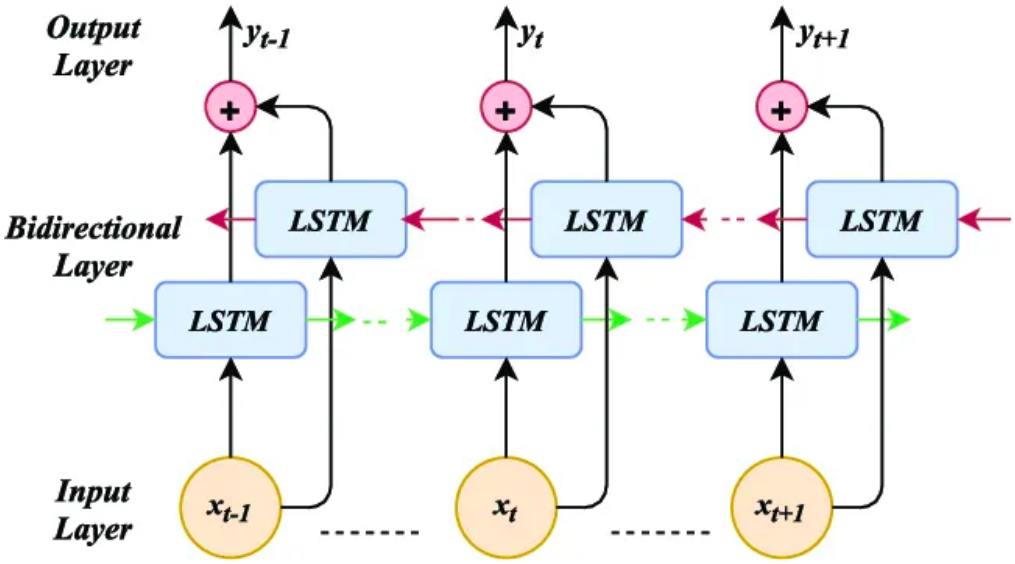

Figure 3.2: Architechture of Bi-LSTM [2]

3.3 GRU

To address the vanishing gradient problem with conventional RNNs, a type of RNN architecture known as the Gated Recurrent Unit was developed [27]. Its smaller structure and fewer gates make it computationally more efficient than the LSTM model. The two main gates that comprise the GRU are the update gate and the reset gate. By regulating the information flow and choosing what to remember and what to keep, the update gate, determines what quantity of the past data should

be transferred forward. In contrast, the reset gate determines what quantity of the previous data should be discarded in order to determine the recent concealed state. The GRU combines these features, making it less complicated and quicker to train than the LSTM, which employs input, forget, and output gates and a separate memory cell to efficiently capture temporal dependencies in time-series data.

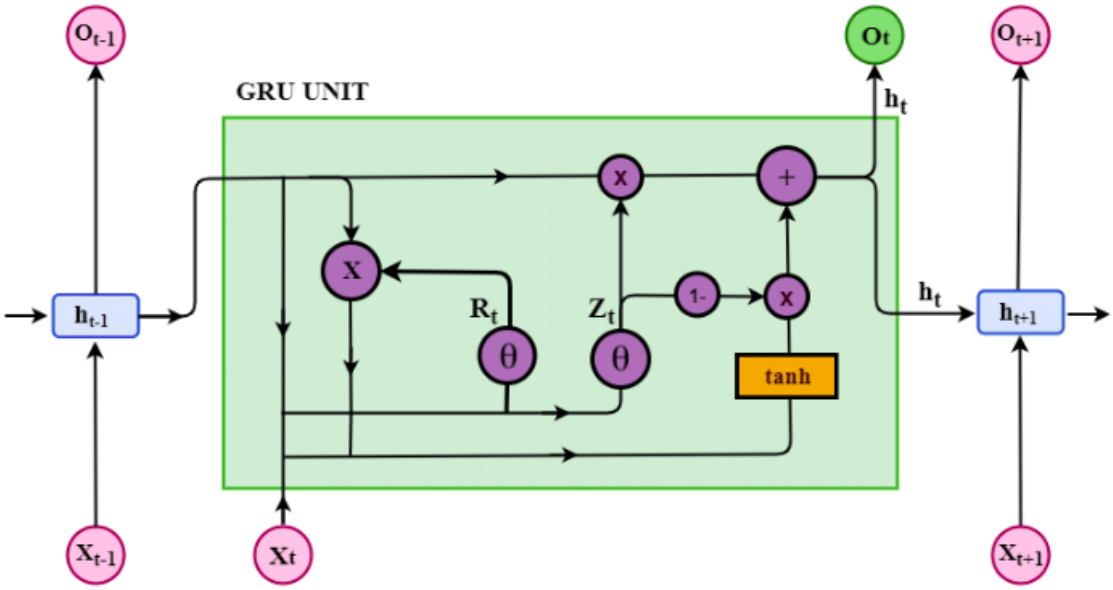

Figure 3.3: Architecture of GRU [27]

Chapter 4

Methodology

Finding the beginning and end of each freezing episode also the recurrence of these three types of FoG events are the goals of the proposed Parkinson's FoG detection model: walking, turning, and start hesitation. To do this, the model has to be designed with a procedure that receives as input lower-back 3D accelerometer data from participants, processes the data in a methodical manner, detects the beginning and end of each FoG episode, and generates predictions of three folds: start hesitation, turn, or walking. The model uses the input data to detect freezing of gait events after it has undergone pre-processing. Since the distribution of the two datasets differs, it is required to create two distinct models: one for tdcsfog and one for defog. The top level overview of the proposed FoG detection system is shown in Fig.4.1.

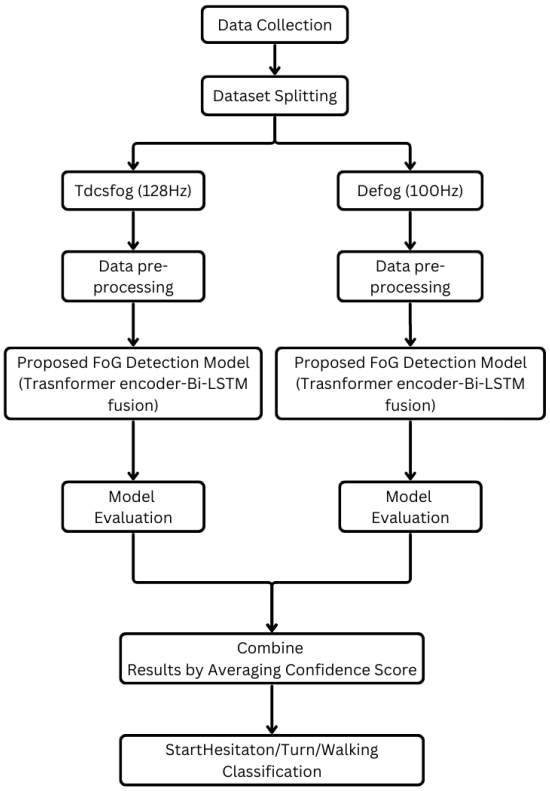

Figure 4.1: Top Level Overview of the Proposed FoG Detection System

Defog data is not utilized to train tdcsfog models, and tdcsfog models are not trained using defog data. The proposed FoG detection system is composed of several phases: (1) Collection of data, (2) Processing the input data, (3) Propose FoG detection Model, (4) Classification of FoG into three types of events. In the proposed FoG detection model, two data sets (explained in Section 4.1) have been used. The next phase is the data processing steps described in Section 4.2. After that, we proposed a Transformer Encoder-Bi-LSTM fusion model (explained in Section 4.3) to detect freezing episodes from the data and categorize them into three discrete classes: Start Hesitation, Turn and Walking. Finally, the model was evaluated by the mean average precision of predictions for each event class.

4.1 Dataset

Machine learning based FoG detection has been conducted on two freezing of gait datasets, namely tdcsfog and defog [45]. These data sets have been acquired from publicly accessible database. These data sets were annotated by expert reviewers documented the freezing of gait episodes.

The tDCS FOG Dataset (tdcsfog): Consisting of data series collected in the lab. During each visit, individuals wore a 3D accelerometer on their lower back (sampling rate 128 Hz). Every trial that caused FoG was captured on camera and examined offline. Before the test protocol begins, there is a brief (2–3 s) interval of silent standing while data recordings. The correlation matrix for tdcsfog data is shown in Fig.4.2.

Figure 4.2: Correlation matrix of tdcsfog

The DeFOG Dataset (defog): Consisting of data series gathered while subjects were completing a program designed to induce FOG in their homes. Two trips to the subject's home surroundings were part of this investigation. The subjects were assessed at both the off and on medication states at each visit. The subjects wore a 3D accelerometer on their lower backs during the motor evaluation, which gathered

data at a sample rate of 100 Hz. The correlation matrix for tdcsfog data is shown in Fig.4.3.

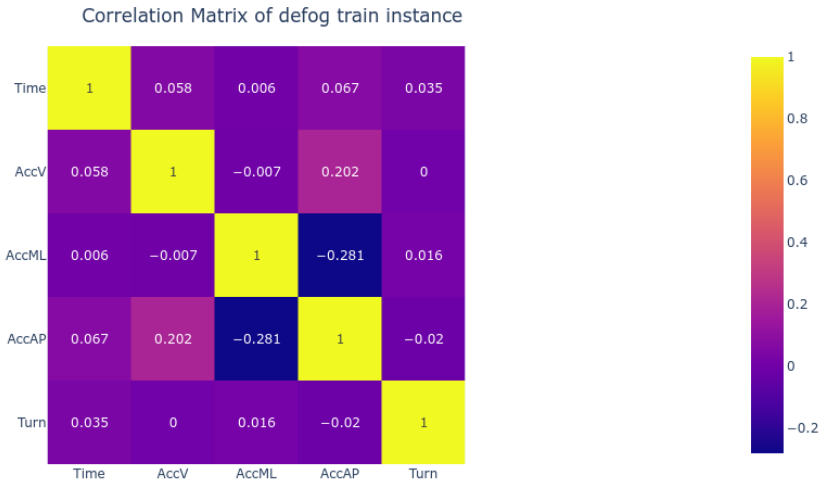

Figure 4.3: Correlation matrix of defog

The following protocols described by Ziegler et al. 2010 [21] shown in Fig.4.4 were performed each time for collecting data:

1. The 4-Meter Walk Test: Involves the participant walking a distance of 4 meters at their usual, comfortable pace. The time needed to complete the walk is collected using a stopwatch, and the test may be repeated multiple times, with the average time used for analysis.

2. The Timed Up & Go (TUG) Single Task: It requires the participant to begin seated in a standard chair. Upon the command "go", the participant stands up, walks a distance of 3 meters, turns around, walks back, and sits down again. The total time needed to complete the task is recorded, and the test assesses mobility, balance, and functional ability.

3. The Timed Up & Go (TUG) Dual Task: The participant performs the standard TUG test while simultaneously being asked to subtract a specified number (e.g., subtract 3s from 100) out loud while walking. The total time needed to complete the task and the accuracy of the subtraction are recorded to evaluate the effect of cognitive load on mobility.

4. The Turning - Single Task: Involves the participant performing four 360-degree turns, alternating the direction with each turn (clockwise and counterclockwise). The participant's speed, stability, and turning technique are observed and recorded.

5. The Turning – Dual Task: The participant performs the same turning task as before, but with an added cognitive challenge of subtracting a specified number (e.g., subtract 3s from 100) while turning. The time, stability, and accuracy of the cognitive task (subtraction) are recorded.

6. The Hotspot Door Test: Involves the participant walking to a designated door, opening it, entering an adjacent room, turning around, and returning to the starting point. This test simulates real-life mobility challenges, such as navigating through doorways and making directional changes. The total time needed and any hesitations or difficulties, such as freezing of gait, are recorded.

7. The Personalized Hotspot Test: The participant is asked to identify an area within their home, that typically triggers freezing of gait (FoG). The participant then walks through this area under observation, with the occurrence and duration of FoG episodes noted, along with any compensatory strategies used.

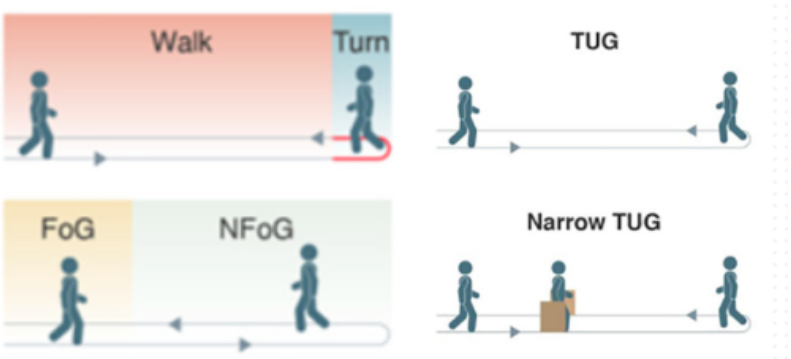

Figure 4.4: Data collection protocols

Field description of tdcsfog and defog dataset:.

Time: An integer timestep. The datasets differ in their sampling rates: tdcsfog data is collected at 128Hz, whereas defog data is recorded at 100Hz.

AccV, AccML, and AccAP: Figure 4.5 illustrates the three axes of acceleration from a lower-back sensor: vertical (V), mediolateral (ML), and anteroposterior (AP). For tdcsfog and defog, the data is expressed in m/s^2 and g, respectively.

Start Hesitation, Turn, Walking: Indicator variables for the occurrence of each of the event types.

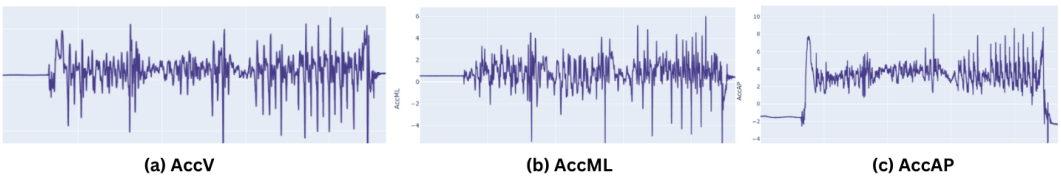

Figure 4.5: Time series features from dataset

4.2 Data Pre-processing

4.2.1 Normalization

Mean-std normalization is applied to the acceleration columns : AccV, AccML, and AccAP for both tdcsfog and defog data series. By calculating the mean of each data point in the dataset and dividing it by the standard deviation, mean-std or Z-score normalization is applied in data preparation to standardize characteristics so that the distribution has a standard deviation of one and a mean of zero. The tdcsfog dataset records acceleration in m/s^2, while the defog dataset uses 'g' (gravitational units). These differences in measurement units can lead to inconsistencies when the data is processed. Mean-std normalization standardizes these values, ensuring that the different units do not affect the modeling process. Moreover, the datasets have different sampling frequencies: tdcsfog at 128 Hz and defog at 100 Hz. Normalizing the features ensure that the variance due to different sampling rates is minimized, providing a uniform scale across the datasets. Without normalization, features with larger magnitudes (acceleration values in certain axes) could dominate others in the modeling process, potentially biasing the model. By applying mean-std normalization, all features are brought to the same scale, ensuring that no single feature dominates the training process.

4.2.2 Reducing the Sampling Rate of the Data

In this project, the frequency of the target data is reduced to improve the model's performance. The original target data has a high resolution with a frequency of 128Hz, which means it captures detailed, high-frequency information over time. While such high-resolution data can be useful, deep learning models often struggle to handle targets with a complex structure or excessive detail. In particular, high-frequency data can introduce noise or unnecessary complexity that creates it harder for the system to understand relevant patterns effectively.
To address this, the target data, initially in a high-resolution format, is reshaped and reduced through a series of transformations. The first step reshapes the target data into smaller patches, reducing the temporal resolution. The data is then transposed to facilitate operations across specific dimensions. Lastly, by choosing the most important value within each patch, the TensorFlow reduction max operation is applied throughout the patch dimension, reducing the complexity of the target data.
This reduction in resolution simplifies the structure of the target data, enabling the model to better focus on the core signals and features that are relevant for learning. It also removes high-frequency noise and variability that could impair the model's ability to generalize, leading to better performance on lower-resolution, smoothed-out target data. This preprocessing step helps the model capture important trends and behaviors more effectively, resulting in improved overall performance.

4.2.3 Partitioning Data into Fixed-length Blocks

By copying and selecting specific columns, and then converting these into a standardized numerical format, the method prepares the data for further analysis using patches like Vision Transformers. Padding the series to make its length a multiple of

the block size ensures that all blocks are uniformly sized, which is crucial for model compatibility. The use of overlapping blocks, determined by the block stride, allows the method to capture sequential dependencies and transitional features within the series, enhancing the model's ability to detect patterns like freezing of gait (FOG) events. By extracting these blocks as individual units containing start and end indices along with the values, the processing facilitates the efficient transformation of continuous time series data into a format suitable for training and testing predictive models. This structured segmentation approach is essential for ensuring that the model can learn temporal patterns effectively, increasing its accuracy and robustness in identifying events. Moreover, the overlapping blocks provide a mechanism for minimizing information loss at the boundaries of each segment, ensuring that critical transitions are preserved and analyzed. This design also enables the model to handle variability in signal dynamics, capturing subtle patterns that may indicate early signs of FOG events. The standardized block-based format further simplifies batching and parallel processing during training, significantly improving computational efficiency. By organizing the data in this manner, the method not only ensures compatibility with advanced architectures like Vision Transformers but also establishes a foundation for scalability to larger datasets and more complex scenarios, such as multi-sensor fusion and real-time event detection.

4.3 Proposed Transformer encoder-Bi-LSTM Fusion Model

In this thesis, a Transformer encoder and two Bi-LSTM layers were combined to create a hybrid deep learning model that can identify FoG episodes shown in Fig.4.6. To find long-range relationships in the data, the model first splits the accelerometer time-series data into blocks, which are then processed by a Transformer encoder that employs multi-head self-attention. Each encoder layer applies multi-head attention with residual connections and layer normalization, followed by a feed-forward network with dropout to enhance generalization. To incorporate temporal positional information, a trainable position encoding is added to the input, which allows the Transformer to capture the sequential structure of the data. After passing through multiple Transformer encoder layers, the features are further processed by two Bi-LSTM layers, which model both forward and backward dependencies in the time-series data, crucial for identifying transitions between FOG episodes. Finally, the model outputs the confidence score of three FOG event types (Start Hesitation, Turn, and Walking) using a dense layer followed by a sigmoid activation function. This hybrid architecture leverages the strengths of both Transformers and Bi-LSTMs, making it highly effective in capturing both long-term dependencies and local temporal patterns for accurate FOG detection. Additionally, the model's use of multi-head self-attention enhances its ability to focus on relevant features across the input blocks, while the Bi-LSTM layers ensure that fine-grained temporal relationships are preserved. By combining these two approaches, the architecture is particularly well-suited for handling the complex dynamics of accelerometer data during FOG episodes. Furthermore, dropout regularization throughout the network helps mitigate overfitting, making the model robust across varying patient datasets. This design allows the model to generalize effectively, demonstrating significant po-

tential for real-world clinical applications.

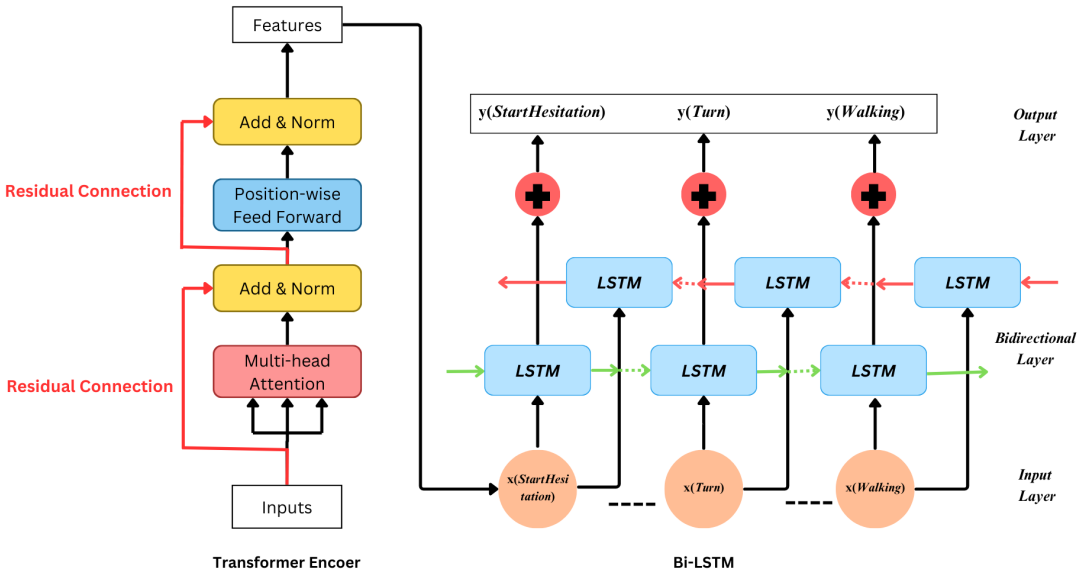

Figure 4.6: Proposed Transformer encoder-Bi-LSTM Fusion Model

4.3.1 Model Training

In this study, we fine-tune several key hyperparameters to optimize the detection and classification of FoG episodes using a Transformer Encoder and Bi-LSTM model. All the training was conducted with TensorFlow capabilities and Kaggle Notebook resources. The input data is processed in blocks, with each block divided into smaller patches, allowing the model to handle large sequences while retaining essential temporal information. The model architecture integrates a Transformer Encoder and Bi-LSTM layers. The Transformer Encoder applies multiple attention heads, enabling it to focus on different aspects of the input concurrently, while several encoder layers help capture complex patterns in the sequential data. The LSTM layers capture long-term dependencies crucial for detecting FoG episodes. To prevent overfitting and improve generalization, dropout is applied at multiple stages, including the first layer and the Transformer Encoder. A key aspect of the training process is the learning rate schedule, which starts with a warm-up phase to stabilize initial training and then maintains a gradual learning rate for fine-tuning. Batch size and the number of steps per epoch are optimized to balance computational efficiency and model stability, while the model also benefits from accelerated training, allowing it to process the large dataset effectively. Each of these parameters contributes to the model's capacity to precisely detect and distinguish FoG events in real-time.

The loss function used in this research is based on binary cross-entropy that computes the loss between the real values and the model's predicted output without reducing the dimensionality immediately, enabling flexibility in further processing. The function first expands the dimensions of both the real values and the output to

match a shape suitable for element-wise comparison. Next, two particular columns of the actual tensor are multiplied to create a mask that represents circumstances or occurrences that should be taken into account when calculating the loss. This mask is expanded and tiled across the target dimensions to match the shape of the loss tensor, ensuring that only relevant parts of the input contribute to the overall loss calculation. After that, all the masked loss values are added, and they are normalized using the total mask sum to provide the average loss value that only represents the valid segments. This approach effectively handles the sparsity of relevant events and ensures the model focuses on learning from meaningful parts of the data, improving the accuracy and robustness of the detection model. The optimizer used in this model is a custom implementation of the Adam optimizer from TensorFlow.

Chapter 5

Results and Discussion

5.1 Performance Evaluation Metrics

The models were evaluated using the mean average precision (mAP) of forecasts for each type of occurrence. We independently compute the average precision on predicted scores for the three event classes (Start Hesitation, Turn, Walking). These three numbers are then averaged to determine the total score. The trade-off between precision and recall is seen in mAP. Since they take into account both false positive (FP) and false negative (FN), mAP is a good statistic for the majority of detection applications. The mean average precision is shown in the equation (5.1) where AP_k is the average precision of class k and n is the number of classes.

$$mAP = \frac{1}{n} \sum_{k=1}^{k=n} AP_k \tag{5.1}$$

Average Precision: Average Precision is calculated as the weighted mean of precisions at each threshold, where the weight corresponds to the increase in recall from the preceding threshold. It provides a single metric summarizing the precision-recall curve. The formulation of Average Precision is provided in Equation (5.2).

$$Confidence = AP = \frac{1}{|Thresholds|} \sum_{T} \frac{TP}{TP + FP + FN} \tag{5.2}$$

It is composed of four submetrics: confusion matrix, Intersection over Union (IoU), Recall and Precision.

Confusion Matrix: A confusion matrix, depicted in Fig.5.1, is constructed using four key components.

True Positives: Instances where the model correctly predicts a label, matching the ground truth.
True Negatives: Instances where the model correctly identifies that a label is absent, aligning with the ground truth.
False Positives: Instances where the model predicts a label that is not present in the ground truth.
False Negatives: Instances where the model fails to predict a label that is present

in the ground truth.

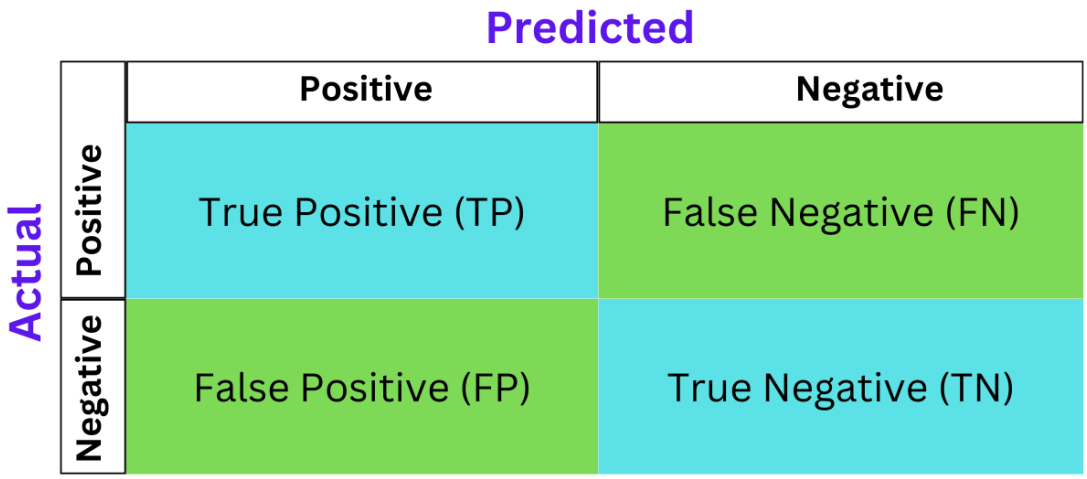

Figure 5.1: Confusion Matrix

IoU: The Intersection over Union, denoted in Equation (5.3), quantifies the overlap between the predicted bounding box and the ground truth bounding box. It is a standard metric for evaluating object detection models.

$$IoU = \frac{Area\ of\ Overlap}{Area\ of\ Union} \tag{5.3}$$

Precision: Precision, as defined in Equation (5.4), measures the proportion of true positive predictions relative to the total number of positive predictions made by the model. It evaluates the model's accuracy in identifying true positives among all predicted positives.

$$Precision = \frac{TP}{TP + FP} \tag{5.4}$$

Recall: Recall, as defined in Equation (5.5), measures the model's ability to identify true positives (TP) among all actual positives. It is a key metric for evaluating how effectively the model captures relevant instances.

$$Recall = \frac{TP}{TP + FN} \tag{5.5}$$

Accuracy: Accuracy, shown in equation (5.6), is the percentage of correct classifications that a trained model achieves.

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \tag{5.6}$$

F1 Score: The F1 score, described in Equation (5.7), represents the harmonic mean of precision and recall. It balances the trade-off between these two metrics, providing a single score that reflects both the model's accuracy in identifying true positives and its ability to retrieve all relevant instances.

$$F1 - score = \frac{2 * Precision * Recall}{Precision\ +\ Recall} \tag{5.7}$$

5.2 Result Analysis

The proposed fusion model combining a Transformer Encoder with Bi-LSTM layers achieved a mean Average Precision (mAP) score of 0.5206, surpassing the performance of the Transformer Encoder-GRU fusion model, which achieved a lower mAP score of 0.4938. This improvement underscores the effectiveness of Bi-LSTM layers in capturing bidirectional temporal dependencies, which are critical for accurately modeling the transitions and complexities inherent in Freezing of Gait (FoG) episodes.

In this study, the performance of various machine learning architectures was compared for the detection and classification of FoG episodes, with mAP as the primary evaluation metric. The proposed Transformer Encoder-Bi-LSTM model outperformed several baseline architectures, including the Multilayer Bidirectional GRU (mAP: 0.4729) [46], LSTM (mAP: 0.4445) [47], and 1D ResNet (mAP: 0.3568) [44]. While another model leveraging an ensemble GRU [48] achieved a slightly higher mAP score of 0.5367, that approach involved extensive data preprocessing and an ensemble of multiple networks, resulting in a significantly higher computational cost.

The hybrid architecture of the Transformer Encoder-Bi-LSTM combines the global feature extraction capabilities of attention mechanisms with the temporal modeling strengths of recurrent networks. Additionally, the use of Bi-LSTM layers enhances the ability to model both forward and backward temporal dependencies, ensuring a more comprehensive understanding of sequential data. The lightweight nature of the architecture, compared to ensemble-based models, also makes it more suitable for deployment in real-time and resource-constrained environments. The comparative analysis of mAP scores, illustrated in Figure 5.2, highlights the potential of the proposed architecture to serve as a competitive and computationally efficient alternative for FoG detection and classification.

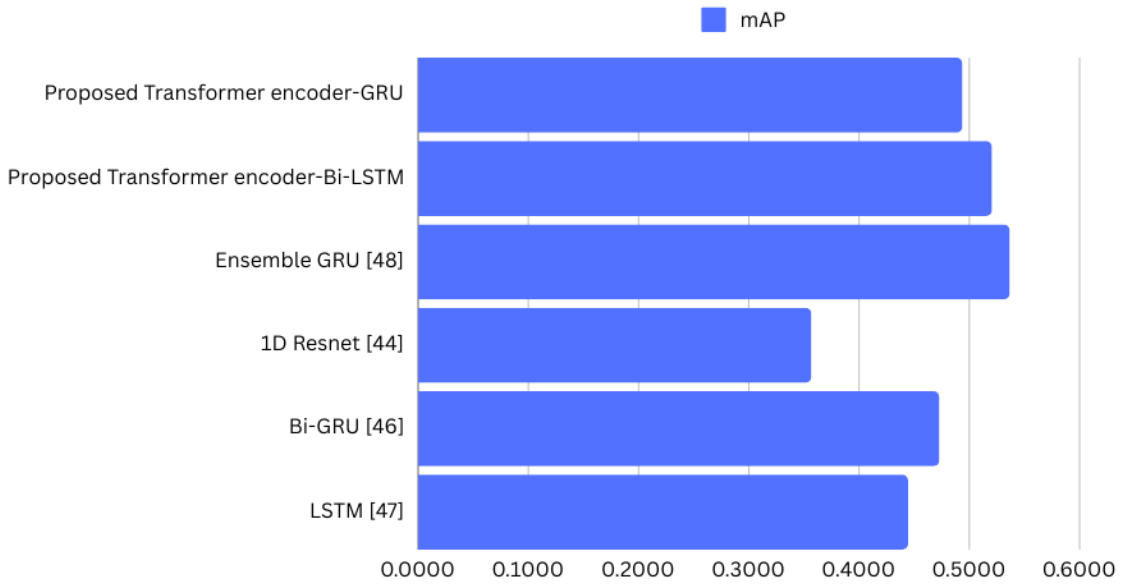

Figure 5.2: Comparison of Proposed Models with respect to Mean Average Precision

The results of the local cross-validation for the Tdcsfog training models are summarized in Table 5.1, where the mean Average Precision (mAP) scores range from 0.473 to 0.602. The individual Average Precision (AP) for each of the three outcomes across different folds varies as follows: StartHesitation achieves AP values between 0.360 and 0.594, Turn shows AP values ranging from 0.850 to 0.889, and Walking records AP values between 0.187 and 0.463.

Table 5.1: Training Results of Tdcsfog Models using Transformer endcoder-Bi-LSTM Fusion

Configuration	AP			mAP
	StartHesitation	Turn	Walking	
1	0.360	0.872	0.187	0.473
2	0.594	0.850	0.282	0.575
3	0.474	0.879	0.430	0.594
4	0.455	0.889	0.463	0.602

These results were achieved by tuning the following parameters of the model configuration:

- Block Size: The size of each input segment processed by the model.
- Patch Size: The size of smaller chunks into which each block is divided.
- Fog Model Dimension: The dimensionality of the feature embeddings.
- Attention Head: The number of attention heads in the multi-head self-attention mechanism.
- Encoder Layer: The number of encoder layers in the Transformer.
- First Dropout: Dropout rate applied to the input or initial layers to prevent overfitting.
- Encoder Dropout: Dropout rate applied to the Transformer encoder, enhancing generalization by randomly deactivating neurons during training.
- MHA Dropout: Dropout rate applied within the multi-head attention mechanism.

Similarly, the mAP scores for the Defog models, as presented in Table 5.2, range from 0.432 to 0.489. The individual AP for the three outcomes across the folds is as follows: StartHesitation achieves AP values ranging from 0.215 to 0.485, Turn records AP values between 0.410 and 0.492, and Walking shows AP values between 0.238 and 0.490.

Table 5.2: Training Results of Defog Models using Transformer endcoder-Bi-LSTM Fusion

Configuration	AP			mAP
	StartHesitation	Turn	Walking	
1	0.215	0.410	0.238	0.432
2	0.440	0.452	0.446	0.446
3	0.460	0.468	0.461	0.463
4	0.485	0.492	0.490	0.489

The change of mAP score for tdcsfog and defog models using combination of Transformer Encoder and GRU is shown in Table.5.3 and Table 5.4. The individual AP for each of the three outcomes in the individual folds ranges from 0.460 to 0.519 for StartHesitation, 0.520 to 0.696 for Turn, and 0.320 to 0.600 for Walking as shown in Table 4.3. The Transformer encoder employs 320-dimensional embeddings, 6 attention heads, and 5 encoder layers, which are designed to capture global patterns and relationships in the input data. Additionally, the GRU component contributes to modeling sequential dependencies while maintaining computational efficiency compared to more complex recurrent layers. The dropout set to 0.1 for the initial, encoder, and multi-head attention layers—provide a balanced approach to regularization, helping mitigate overfitting while retaining critical signal information.

Table 5.3: Training Results of Tdcsfog Models using Transformer endcoder-GRU Fusion

Configuration	AP			mAP
	StartHesitation	Turn	Walking	
1	0.519	0.520	0.320	0.453
2	0.430	0.620	0.579	0.549
3	0.450	0.672	0.600	0.574
4	0.460	0.696	0.590	0.582

Table 5.4: Training Results of Defog Models using Transformer endcoder-GRU Fusion

Configuration	AP			mAP
	StartHesitation	Turn	Walking	
1	0.300	0.410	0.505	0.405
2	0.310	0.420	0.527	0.419
3	0.320	0.436	0.552	0.436
4	0.345	0.462	0.579	0.462

The performance of our proposed models using other evaluation metrics like precision, recall, F1 score and accuracy are given in Table.5.5. The results highlight the performance metrics of two proposed models: the Transformer Encoder-Bi-LSTM fusion and the Transformer Encoder-GRU fusion. For the Transformer encoder with Bi-LSTM, the metrics achieved are an F1 score of 0.783, accuracy of 0.897, precision of 0.798, and recall of 0.768. For the Transformer encoder with GRU, the results are an F1 score of 0.771, accuracy of 0.844, precision of 0.697, and recall of 0.747. These metrics demonstrate effective recall performance, capturing a high proportion of true positive instances, while maintaining a reasonable level of precision.

Table 5.5: Test Results of the Proposed Models

Model	F1 score	Accuracy	Precision	Recall
Proposed Transformer Encoder-Bi-LSTM	0.783	0.897	0.798	0.768
Proposed Transformer Encoder-GRU	0.771	0.844	0.697	0.747

5.3 Discussion

The best-performing models on the same dataset [45] performed slightly better than our solution, even though the proposed Transformer Encoder-Bi-LSTM model has a strong mAP score. It suggests that doing intensive data preprocessing and ensemble models might produce better overall FoG detection outcomes. Although there are some drawbacks, the claimed solution's simplicity makes it stand out. It is shown how to achieve competitive performance in FOG detection without the need for excessively complex models, pre-processing, or substantial CPU resources by integrating two Deep Learning architectures, such as the transformer encoder and Bi-LSTM. This simplicity not only makes implementation easier, but it also makes it relatively easy for other researchers who are interested to adopt and expand upon the published methodology.

In order to improve the standard of life of PD patients, reliable automatic FOG detection is an important aspect to consider. The present study reported a highly ranked solution to the Parkinson's FoG detection. Results show that using a combination of transformer encoder and two Bidirectional LSTM layers can be successfully applied to detect FOG related events in accelerometer data from different settings without applying sophisticated pre-processing. The procedure did not make use of the daily living dataset. To support the detection modeling, one might add semi-supervised or unsupervised techniques to the daily dataset's series.

Chapter 6

Conclusion

One of the main motor signs of PD that can lead to falls is freezing of gait. The majority of PD patients are older, therefore falls can cause fractures or even death. Finding ways to diagnose FOG and stop patients from falling is therefore crucial. An effective FOG recognition model based on deep learning was created in this study. This diagnostic examination service lessens the financial burden on PD patients in addition to mitigating the effects of individual subjectivity. Although there are alternative methods, FOG-provoking protocols are used in the majority of FOG evaluation methodologies. Individuals with FOG are captured on camera while engaging in activities that are likely to exacerbate the condition. After watching the footage, experts score each frame to determine when FOG happened. This method of scoring is quite time-consuming and necessitates specialized knowledge, but it is also reasonably sensitive and dependable. Another approach is to use wearable technology to enhance FOG-provoking testing. FOG detection is made easier with additional sensors, but usability and compliance may suffer. Thus, the optimal strategy might be a mix of these two techniques. The accuracy of identifying FOG from a lower back accelerometer is comparatively high when paired with machine learning techniques. This approach can truly provide full-process management and timely intervention in the PD-FOG population.

6.1 Limitations and Future Work

One key limitation is the computational complexity of this architecture. The Transformer encoder, with its multi-head attention mechanism, can be resource-intensiv. This makes it challenging to deploy the model on edge devices such as wearable sensors, where computational power and battery life are constrained. Another limitation is related to the real-time performance and latency. Although the model has strong predictive capabilities, ensuring it operates in real-time with minimal lag is crucial for practical applications, such as triggering therapeutic responses for Parkinson's patients during a FoG episode.

Future work could focus on optimizing the model for real-time deployment. The model could benefit from incorporating data from multiple sensors, such as combining accelerometer data with gyroscope or electromyography inputs, to enhance detection capabilities.

Bibliography

[1] A. De Boer, W. Wijker, J. Speelman, and J. De Haes, "Quality of life in patients with parkinson's disease: Development of a questionnaire.," *Journal of Neurology, Neurosurgery & Psychiatry*, vol. 61, no. 1, pp. 70–74, 1996.

[2] M. Schuster and K. K. Paliwal, "Bidirectional recurrent neural networks," *IEEE transactions on Signal Processing*, vol. 45, no. 11, pp. 2673–2681, 1997.

[3] A. Nieuwboer, W. d. Weerdt, R. Dom, and E. Lesaffre, "A frequency and correlation analysis of motor deficits in parkinson patients," *Disability and rehabilitation*, vol. 20, no. 4, pp. 142–150, 1998.

[4] M. d. De Rijk, L. Launer, K. Berger, M. Breteler, J. Dartigues, M. Baldereschi, L. Fratiglioni, A. Lobo, J. Martinez-Lage, C. Trenkwalder, *et al.*, "Prevalence of parkinson's disease in europe: A collaborative study of population-based cohorts. neurologic diseases in the elderly research group.," *Neurology*, vol. 54, no. 11 Suppl 5, S21–3, 2000.

[5] N. Giladi, M. McDermott, S. Fahn, S. Przedborski, J. Jankovic, M. Stern, C. Tanner, and P. S. Group, "Freezing of gait in pd: Prospective assessment in the datatop cohort," *Neurology*, vol. 56, no. 12, pp. 1712–1721, 2001.

[6] J. H. Han, W. J. Lee, T. B. Ahn, B. S. Jeon, and K. S. Park, "Gait analysis for freezing detection in patients with movement disorder using three dimensional acceleration system," in *Proceedings of the 25th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (IEEE Cat. No. 03CH37439)*, IEEE, vol. 2, 2003, pp. 1863–1865.

[7] J. Schaafsma, Y. Balash, T. Gurevich, A. Bartels, J. M. Hausdorff, and N. Giladi, "Characterization of freezing of gait subtypes and the response of each to levodopa in parkinson's disease," *European journal of neurology*, vol. 10, no. 4, pp. 391–398, 2003.

[8] B. R. Bloem, J. M. Hausdorff, J. E. Visser, and N. Giladi, "Falls and freezing of gait in parkinson's disease: A review of two interconnected, episodic phenomena," *Movement disorders: official journal of the Movement Disorder Society*, vol. 19, no. 8, pp. 871–884, 2004.

[9] E. Cubo, S. Leurgans, and C. G. Goetz, "Short-term and practice effects of metronome pacing in parkinson's disease patients with gait freezing while in the 'on'state: Randomized single blind evaluation," *Parkinsonism & related disorders*, vol. 10, no. 8, pp. 507–510, 2004.

[10] M. Suteerawattananon, G. Morris, B. Etnyre, J. Jankovic, and E. Protas, "Effects of visual and auditory cues on gait in individuals with parkinson's disease," *Journal of the neurological sciences*, vol. 219, no. 1-2, pp. 63–69, 2004.

[11] I. Lim, E. van Wegen, C. de Goede, M. Deutekom, A. Nieuwboer, A. Willems, D. Jones, L. Rochester, and G. Kwakkel, "Effects of external rhythmical cueing on gait in patients with parkinson's disease: A systematic review," *Clinical rehabilitation*, vol. 19, no. 7, pp. 695–713, 2005.

[12] N. Giladi and J. M. Hausdorff, "The role of mental function in the pathogenesis of freezing of gait in parkinson's disease," *Journal of the neurological sciences*, vol. 248, no. 1-2, pp. 173–176, 2006.

[13] J. M. Hausdorff, J. Lowenthal, T. Herman, L. Gruendlinger, C. Peretz, and N. Giladi, "Rhythmic auditory stimulation modulates gait variability in parkinson's disease," *European Journal of Neuroscience*, vol. 26, no. 8, pp. 2369–2375, 2007.

[14] M. Macht, Y. Kaussner, J. C. Möller, K. Stiasny-Kolster, K. M. Eggert, H.-P. Krüger, and H. Ellgring, "Predictors of freezing in parkinson's disease: A survey of 6,620 patients," *Movement disorders*, vol. 22, no. 7, pp. 953–956, 2007.

[15] A. Nieuwboer, G. Kwakkel, L. Rochester, D. Jones, E. van Wegen, A. M. Willems, F. Chavret, V. Hetherington, K. Baker, and I. Lim, "Cueing training in the home improves gait-related mobility in parkinson's disease: The rescue trial," *Journal of Neurology, Neurosurgery & Psychiatry*, vol. 78, no. 2, pp. 134–140, 2007.

[16] S. T. Moore, H. G. MacDougall, and W. G. Ondo, "Ambulatory monitoring of freezing of gait in parkinson's disease," *Journal of neuroscience methods*, vol. 167, no. 2, pp. 340–348, 2008.

[17] S. Rahman, H. Griffin, N. Quinn, and M. Jahanshahi, "The factors that induce or overcome freezing of gait in parkinson's disease," *Behavioural neurology*, vol. 19, no. 3, pp. 127–136, 2008.

[18] M. Bachlin, M. Plotnik, D. Roggen, I. Maidan, J. M. Hausdorff, N. Giladi, and G. Troster, "Wearable assistant for parkinson's disease patients with the freezing of gait symptom," *IEEE Transactions on Information Technology in Biomedicine*, vol. 14, no. 2, pp. 436–446, 2009.

[19] E. Jovanov, E. Wang, L. Verhagen, M. Fredrickson, and R. Fratangelo, "Defog—a real time system for detection and unfreezing of gait of parkinson's patients," in *2009 annual international conference of the IEEE engineering in medicine and biology society*, IEEE, 2009, pp. 5151–5154.

[20] A. Nieuwboer, K. Baker, A.-M. Willems, D. Jones, J. Spildooren, I. Lim, G. Kwakkel, E. Van Wegen, and L. Rochester, "The short-term effects of different cueing modalities on turn speed in people with parkinson's disease," *Neurorehabilitation and neural repair*, vol. 23, no. 8, pp. 831–836, 2009.

[21] K. Ziegler, F. Schroeteler, A. O. Ceballos-Baumann, and U. M. Fietzek, "A new rating instrument to assess festination and freezing gait in parkinsonian patients," *Movement Disorders*, vol. 25, no. 8, pp. 1012–1018, 2010.

[22] B. T. Cole, S. H. Roy, and S. H. Nawab, "Detecting freezing-of-gait during unscripted and unconstrained activity," in *2011 Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, IEEE, 2011, pp. 5649–5652.

[23] S. Donovan, C. Lim, N. Diaz, N. Browner, P. Rose, L. Sudarsky, D. Tarsy, S. Fahn, and D. Simon, "Laserlight cues for gait freezing in parkinson's disease: An open-label study," *Parkinsonism & related disorders*, vol. 17, no. 4, pp. 240–245, 2011.

[24] J. G. Nutt, B. R. Bloem, N. Giladi, M. Hallett, F. B. Horak, and A. Nieuwboer, "Freezing of gait: Moving forward on a mysterious clinical phenomenon," *The Lancet Neurology*, vol. 10, no. 8, pp. 734–744, 2011.

[25] S. Mazilu, M. Hardegger, Z. Zhu, D. Roggen, G. Tröster, M. Plotnik, and J. M. Hausdorff, "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, pp. 123–130.

[26] D. Rodriguez-Martin, A. Sama, C. Perez-Lopez, A. Catala, J. Cabestany, and A. Rodriguez-Molinero, "Svm-based posture identification with a single waist-located triaxial accelerometer," *Expert Systems with Applications*, vol. 40, no. 18, pp. 7203–7211, 2013.

[27] J. Chung, C. Gulcehre, K. Cho, and Y. Bengio, "Empirical evaluation of gated recurrent neural networks on sequence modeling," *arXiv preprint arXiv:1412.3555*, 2014.

[28] C. A. Coste, B. Sijobert, R. Pissard-Gibollet, M. Pasquier, B. Espiau, and C. Geny, "Detection of freezing of gait in parkinson disease: Preliminary results," *Sensors*, vol. 14, no. 4, pp. 6819–6827, 2014.

[29] S. Mazilu, U. Blanke, M. Hardegger, G. Tröster, E. Gazit, and J. 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*, 2014, pp. 2531–2540.

[30] C. Ahlrichs, A. Samà, M. Lawo, J. Cabestany, D. Rodríguez-Martín, C. Pérez-López, D. Sweeney, L. R. Quinlan, G. Ò. Laighin, T. Counihan, *et al.*, "Detecting freezing of gait with a tri-axial accelerometer in parkinson's disease patients," *Medical & biological engineering & computing*, vol. 54, pp. 223–233, 2016.

[31] M. Capecci, L. Pepa, F. Verdini, and M. G. Ceravolo, "A smartphone-based architecture to detect and quantify freezing of gait in parkinson's disease," *Gait & posture*, vol. 50, pp. 28–33, 2016.

[32] J. Hannink, T. Kautz, C. F. Pasluosta, K.-G. Gaßmann, J. Klucken, and B. M. Eskofier, "Sensor-based gait parameter extraction with deep convolutional neural networks," *IEEE journal of biomedical and health informatics*, vol. 21, no. 1, pp. 85–93, 2016.

[33] M. Mancini and F. B. Horak, "Potential of apdm mobility lab for the monitoring of the progression of parkinson's disease," *Expert review of medical devices*, vol. 13, no. 5, pp. 455–462, 2016.

[34] A. L. Silva de Lima, L. J. Evers, T. Hahn, L. Bataille, J. L. Hamilton, M. A. Little, Y. Okuma, B. R. Bloem, and M. J. Faber, "Freezing of gait and fall detection in parkinson's disease using wearable sensors: A systematic review," *Journal of neurology*, vol. 264, pp. 1642–1654, 2017.

[35] A. Vaswani, "Attention is all you need," *Advances in Neural Information Processing Systems*, 2017.

[36] J. Camps, A. Sama, M. Martín, D. Rodriguez-Martin, C. Pérez-López, J. M. M. Arostegui, J. Cabestany, A. Catala, S. Alcaine, B. Mestre, *et al.*, "Deep learning for freezing of gait detection in parkinson's disease patients in their homes using a waist-worn inertial measurement unit," *Knowledge-Based Systems*, vol. 139, pp. 119–131, 2018.

[37] E. R. Dorsey, A. Elbaz, E. Nichols, N. Abbasi, F. Abd-Allah, A. Abdelalim, J. C. Adsuar, M. G. Ansha, C. Brayne, J.-Y. J. Choi, *et al.*, "Global, regional, and national burden of parkinson's disease, 1990–2016: A systematic analysis for the global burden of disease study 2016," *The Lancet Neurology*, vol. 17, no. 11, pp. 939–953, 2018.

[38] A. Samà, D. Rodríguez-Martín, C. Pérez-López, A. Català, S. Alcaine, B. Mestre, A. Prats, M. C. Crespo, and À. Bayés, "Determining the optimal features in freezing of gait detection through a single waist accelerometer in home environments," *Pattern Recognition Letters*, vol. 105, pp. 135–143, 2018.

[39] A. Amini, K. Banitsas, and W. R. Young, "Kinect4fog: Monitoring and improving mobility in people with parkinson's using a novel system incorporating the microsoft kinect v2," *Disability and Rehabilitation: Assistive Technology*, vol. 14, no. 6, pp. 566–573, 2019.

[40] K. Hu, Z. Wang, S. Mei, K. A. E. Martens, T. Yao, S. J. Lewis, and D. D. Feng, "Vision-based freezing of gait detection with anatomic directed graph representation," *IEEE journal of biomedical and health informatics*, vol. 24, no. 4, pp. 1215–1225, 2019.

[41] K. Hu, Z. Wang, W. Wang, K. A. E. Martens, L. Wang, T. Tan, S. J. Lewis, and D. D. Feng, "Graph sequence recurrent neural network for vision-based freezing of gait detection," *IEEE Transactions on Image Processing*, vol. 29, pp. 1890–1901, 2019.

[42] R. San-Segundo, H. Navarro-Hellín, R. Torres-Sánchez, J. Hodgins, and F. De la Torre, "Increasing robustness in the detection of freezing of gait in parkinson's disease," *Electronics*, vol. 8, no. 2, p. 119, 2019.

[43] L. Sigcha, N. Costa, I. Pavón, S. Costa, P. Arezes, J. M. López, and G. De Arcas, "Deep learning approaches for detecting freezing of gait in parkinson's disease patients through on-body acceleration sensors," *Sensors*, vol. 20, no. 7, p. 1895, 2020.

[44] J. Brederecke, "Freezing of gait prediction from accelerometer data using a simple 1d-convolutional neural network–8th place solution for kaggle's parkinson's freezing of gait prediction competition," *arXiv preprint arXiv:2307.03475*, 2023.

[45] A. Howard, amit salomon, eran gazit, J. Hausdorff, L. Kirsch, Maggie, P. Ginis, R. Holbrook, and Y. F. Karim, *Parkinson's freezing of gait prediction*, https://kaggle.com/competitions/tlvmc-parkinsons-freezing-gait-prediction, Kaggle, 2023.

[46] *Parkinson's freezing of Gait prediction*. [Online]. Available: https://www.kaggle.com/competitions/tlvmc-parkinsons-freezing-gait-prediction/discussion/416410.

[47] *Parkinson's freezing of Gait prediction*. [Online]. Available: https://www.kaggle.com/competitions/tlvmc-parkinsons-freezing-gait-prediction/discussion/416248.

[48] *Parkinson's freezing of Gait prediction*. [Online]. Available: https://www.kaggle.com/competitions/tlvmc-parkinsons-freezing-gait-prediction/discussion/416057.