IEEE Access
Multidisciplinary ⋮ Rapid Review ⋮ Open Access Journal

Date of publication xxxx 00, 0000, date of current version xxxx 00, 0000.

Digital Object Identifier 10.1109/ACCESS.2025.DOI

FoG1DLSTM: A Lightweight Deep Learning Transformer Based Model for Real Time Prediction of Parkinson's Freezing of Gait

NOEL S MATHEW[1], RAHUL PATNAIK[1], MOHANA[1], and SUDARSHAN B. G.[2]

[1]Department of Computer Science and Engineering, RV College of Engineering, Bengaluru, Karnataka, India.

[2]Department of Electronics and Instrumentation Engineering, RV College of Engineering, Bengaluru, Karnataka, India

Corresponding author: Sudarshan B. G.(e-mail: sudharshanbg@rvce.edu.in).

This work was supported by the Center for Healthcare Technology and Research (CHTR), R. V. College of Engineering, Bengaluru.

ABSTRACT This paper addresses the critical challenge of predicting Freezing of Gait (FOG) episodes in individuals with Parkinson's Disease (PD) in real time. FOG is a debilitating motor symptom that significantly increases the risk of falls and impairs patient mobility. Leveraging the comprehensive 70.9GB accelerometer dataset from the Michael J. Fox Foundation, this paper analyzes gait patterns and sensor data to identify subtle precursors to FOG. We develop and benchmark our novel lightweight deep learning model, FoG-1D-LSTM, a hybrid CNN-LSTM architecture with a transformer-inspired attention mechanism specifically designed for lower-back accelerometer signals. The architecture combines 1D convolutional layers for spatial feature extraction with bidirectional LSTMs for temporal modeling, enhanced by a multi-head attention mechanism for improved feature relevance and interpretability. Our results demonstrate that FoG-1D-LSTM achieves exceptional performance with a Mean Average Precision of 0.7065 and a binary classification accuracy of 92.83% on the FOG detection task, while maintaining computational efficiency suitable for deployment on resource-constrained wearable devices. This work bridges deep learning techniques with clinical insights, paving the way for a lightweight, wearable-compatible solution to enhance patient safety through continuous monitoring and proactive intervention. The success of FoG-1D-LSTM validates our core hypothesis that explicitly modeling both the spatial structure of multi-axis accelerometer signals and their temporal evolution through a carefully designed deep learning architecture is paramount for accurate and reliable FOG prediction. By leveraging a large-scale dataset with carefully designed data preprocessing, sliding window segmentation, and a custom loss function to address class imbalance, this work establishes a robust foundation for practical FOG prediction systems that could significantly improve patient safety and quality of life.

INDEX TERMS Accelerometer, Artificial Intelligence, Attention Mechanism, Deep Learning, Edge Computing, Freezing of Gait (FOG), Hybrid CNN-LSTM, Machine Learning, Parkinson's Disease, Wearable Sensors

I. INTRODUCTION

PARKINSON'S disease (PD) is the second most common neurodegenerative disorder worldwide, affecting millions and imposing a significant burden on patients and healthcare systems [1]. It is primarily characterized by motor symptoms like bradykinesia, rigidity, and tremor. Among the most disabling of these is Freezing of Gait (FOG), an episodic inability to generate effective forward stepping despite the intention to walk [2]. Patients often describe the sensation as their feet being "glued to the floor." FOG is a major contributor to falls, loss of independence, and reduced quality of life in individuals with PD [3].

Traditional methods for assessing FOG rely on clinical observation, patient questionnaires, and diaries [4]. These methods are inherently subjective, suffer from recall bias, and fail to capture the unpredictable, transient nature of

FOG episodes in a naturalistic setting. This limitation has spurred research into objective, continuous monitoring using wearable sensor technology [5]. Inertial Measurement Units (IMUs), particularly accelerometers, placed on the lower back can provide high-resolution data on trunk movements, which are closely correlated with gait dynamics and FOG events [6].

The confluence of wearable sensors and artificial intelligence (AI) offers a transformative approach to FOG management [16]. By applying machine learning and deep learning models to accelerometer data, it is possible to identify subtle, pre-FOG patterns that are imperceptible to human observers [8]. This paper presents a comprehensive framework for developing and evaluating models for FOG prediction using the large-scale Michael J. Fox Foundation dataset [**?**]. We conduct a comparative analysis of various baseline models before introducing FoG1DLSTM, a sophisticated hybrid architecture that integrates convolutional feature extraction, bidirectional temporal modeling, and attention-based feature weighting. Our work represents a significant step towards a real time, wearable FOG prediction system capable of deployment on edge computing platforms.

A. KEY CONTRIBUTIONS

This work makes several significant contributions to the field of automated FOG detection and prediction:

1) **Novel Hybrid Architecture**: We introduce FoG1DLSTM, a lightweight yet powerful hybrid CNN-LSTM architecture with attention mechanism specifically designed for lower-back accelerometer signals, achieving state-of-the-art performance while maintaining computational efficiency suitable for edge deployment.
2) **Comprehensive Benchmarking**: We conduct extensive comparative analysis of multiple machine learning and deep learning approaches on a large-scale clinical dataset of 70.9GB, providing robust performance baselines across traditional ensemble methods, attention-based transformers, and hybrid architectures.
3) **Class Imbalance Solution**: We develop and validate a custom Weighted BCE-Dice Loss function that effectively addresses the severe class imbalance inherent in FOG detection with a 21:1 ratio, enabling the model to learn rare but critical FOG patterns without sacrificing overall performance.
4) **Interpretable Temporal Modeling**: Through attention mechanism analysis, we demonstrate that the model autonomously learns to focus on clinically relevant precursor patterns, providing insights into the temporal dynamics of FOG onset that align with clinical observations.
5) **Physiologically-Informed Architecture**: We introduce a complementary Physiologically-Informed Temporal Fusion Transformer that incorporates domain knowledge about Parkinson's tremor frequencies, demonstrating competitive performance while offering enhanced interpretability through learned frequency band analysis.

The remainder of this paper is organized as follows. Section II reviews related work in FOG detection and prediction methods. Section III presents the design and implementation details including dataset specification, preprocessing methodology, and model architectures. Section IV discusses the experimental results and comparative analysis. Section V outlines future research directions, and Section VI concludes the paper.

II. LITERATURE SURVEY

The automatic detection of FOG has been an active area of research for over a decade, with methods evolving from simple thresholding to sophisticated deep learning architectures.

A. TRADITIONAL AND MACHINE LEARNING APPROACHES

Early studies focused on extracting statistical features from accelerometer signals within fixed time windows. Features such as the "freeze index"—the ratio of power in the freeze band (3–8 Hz) to the locomotor band (0.5–3 Hz) were used with simple thresholds to detect FOG [10]. While effective in controlled settings, these methods often struggled with real-world variability. Subsequently, classical machine learning models like Support Vector Machines (SVM), k-Nearest Neighbors (kNN), and Random Forests were employed to classify gait patterns using hand-crafted features [11], [15]. These models offered improved performance but remained dependent on expert-driven feature engineering.

B. DEEP LEARNING INNOVATIONS

The advent of deep learning has enabled end-to-end models that can learn relevant features directly from raw sensor data. One-dimensional Convolutional Neural Networks (1D-CNNs) have been successfully used to extract hierarchical features from time-series signals, reducing the need for manual feature design [**?**], [12].

Recurrent Neural Networks (RNNs) and their variants, such as Long Short-Term Memory (LSTM) networks and Gated Recurrent Units (GRUs), have become prominent for capturing temporal dependencies in sequential gait data [**?**], [17]. Hybrid models combining CNNs for spatial feature extraction and LSTMs for temporal modeling have shown great promise [18], [19]. More recently, Transformer-based architectures with self-attention mechanisms have been explored to capture global dependencies across entire gait sequences [20].

Our work builds upon this body of research by developing FoG1DLSTM, a novel architecture that synergistically combines 1D-CNN feature extraction, bidirectional LSTM temporal modeling, and attention mechanisms. We conduct a systematic benchmark on a large, multi-source dataset, focusing on evaluation metrics that are robust to the inherent class imbalance of FOG events.

III. DESIGN AND IMPLEMENTATION

A. DATASET SPECIFICATION

The paper utilizes three distinct datasets provided by the Michael J. Fox Foundation, each capturing accelerometer data under different conditions. The use of multiple datasets allows for robust model development and validation across lab-based and real-world scenarios [21].

1) tDCS FOG (tdcsfog) Dataset

Recorded in a lab during a FOG-provoking protocol, this dataset provides a "gold standard" for annotated FOG events.

- Sampling Frequency: 128Hz.
- Annotations: Expert-labeled FOG events with start, end, and type.
- Usage: Primary training dataset for FoG1DLSTM with 42,471 training windows and 10,618 validation windows.

2) DeFOG (defog) Dataset

This dataset was collected in subjects' homes during similar FOG-provoking tasks, introducing more environmental variability.

- Sampling Frequency: 100Hz.
- Annotations: Expert annotations for task-related FOG episodes.

3) Daily Living (daily) Dataset

This dataset consists of continuous 24/7 recordings from 65 subjects, capturing unscripted, real-world gait. It is unannotated and presents an opportunity for semi-supervised or unsupervised learning approaches [22].

4) Data Fields and Event Types

The primary sensor data consists of three acceleration axes from a lower-back sensor:

- AccV: Vertical acceleration.
- AccML: Mediolateral (side-to-side) acceleration.
- AccAP: Anteroposterior (forward-backward) acceleration.

For the FoG1DLSTM model, we focus on binary classification of StartHesitation events versus normal gait, representing the critical moments when FOG episodes begin. The target variable identifies whether any FOG event occurs within a given time window, enabling proactive prediction before full gait arrest.

B. METHODOLOGY

1) Exploratory Data Analysis (EDA)

EDA on a dataset sample revealed key characteristics. A correlation matrix of the test data (Fig. 1) showed a strong inverse relationship between AccML and AccAP (-0.81), suggesting a kinematic trade-off indicative of gait alterations. Boxplots confirmed numerous outliers corresponding to movement events, and demographic data showed a right-skewed age distribution peaking in the 60–70 range, consistent with PD prevalence [1].

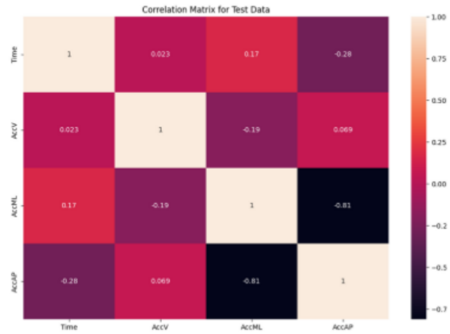


FIGURE 1. Correlation matrix for test data, highlighting the inverse relationship between mediolateral and anteroposterior acceleration.

2) Data Pre-Processing

Data Synchronization and Memory Optimization: To ensure consistency, data from the 128Hz 'tdcsfog' dataset was downsampled to 100Hz where necessary to match other datasets. Given the dataset's large size of 70.9GB, memory optimization was critical for efficient processing. A memory reduction strategy was implemented that intelligently downcast numeric data types to their smallest possible format while preserving data integrity. This approach analyzed the minimum and maximum values of each numeric column and assigned the most memory-efficient data type that could accommodate the value range. For integer columns, types were downcast from int64 to int32, int16, or int8 where appropriate, while floating-point columns were optimized from float64 to float32 or float16. This optimization proved highly effective, reducing the memory footprint of the DEFOG dataset from 954.53 MB to 335.38 MB—a 65% reduction—enabling seamless processing on standard workstations without compromising numerical accuracy.

Sliding Window Approach: For the FoG1DLSTM architecture, a custom PyTorch Dataset class was implemented to generate training samples using a sliding window approach. Each CSV file containing patient data was processed to extract three-channel accelerometer signals (AccV, AccML, AccAP). Time-series data was segmented into overlapping windows with the following parameters:

- Window Size: 384 timesteps, providing sufficient temporal context for capturing gait dynamics.
- Step Size: 128 timesteps, creating 66.7% overlap between consecutive windows to ensure smooth temporal coverage.
- Target Label: Binary classification based on the presence of StartHesitation events within each window (max pooling over window labels).

This sliding window strategy generated 53,089 total samples, with an 80-20 train-validation split yielding 42,471 training and 10,618 validation windows. The resulting tensor shape per sample is (3, 384, 1, 1), where the two trailing dimensions accommodate the 3D convolutional interface while maintaining compatibility with 1D temporal processing.

Feature Scaling: Features were scaled using Standard-

Scaler to have a mean of zero and unit variance, a standard requirement for neural network convergence. This normalization ensures that all three accelerometer axes contribute equally to the learning process regardless of their native amplitude ranges.

3) Model Architectures

Baseline Models:

Ensemble Models (LGBM, XGBoost): LightGBM (LGBM) was selected for its high efficiency and performance on large tabular datasets [23]. It employs a leaf-wise growth strategy and techniques like Gradient-based One-Side Sampling (GOSS) to speed up training. XGBoost and other tree-based ensembles provided strong baseline performance for comparison.

Multi-Headed Attention Transformer: The Transformer architecture was evaluated for its ability to model global dependencies in sequences via self-attention [20]. Its core mechanism is:

$$\text{Attention}(Q, K, V) = \text{softmax}\left(\frac{QK^T}{\sqrt{d_k}}\right)V \tag{1}$$

where multi-head attention allows the model to learn different aspects of the sequence in parallel. Positional encodings were added to retain sequence order information.

FoG1DLSTM: Hybrid CNN-LSTM with Attention

Building upon the limitations of baseline models, we developed FoG1DLSTM, a sophisticated hybrid architecture that integrates the spatial feature extraction capabilities of 1D Convolutional Neural Networks with the temporal modeling strength of Bidirectional LSTMs, enhanced by an attention mechanism for improved feature relevance. This architecture is specifically designed to capture both the spatial patterns within accelerometer signals and the long-range temporal dependencies that precede FOG episodes. The model's lightweight design, with approximately 500K parameters, enables efficient deployment on edge computing platforms while maintaining high predictive accuracy.

Architecture Overview:

The FoG1DLSTM model processes input tensors of shape (batch, 3, 384) through four primary stages, as illustrated in Figure 2:

1) CNN Feature Extractor: The model begins with three sequential convolutional blocks, each comprising:

- 1D convolution with kernel size 7 and padding 3
- Batch normalization for training stability
- ReLU activation for non-linearity
- Max pooling with stride 2 for downsampling

The channel dimensions progressively increase from 3 to 64, 64 to 128, and 128 to 256, while the sequence length reduces from 384 to 192, 96, and finally 48 through pooling operations. This hierarchical feature extraction captures multi-scale patterns in the accelerometer data, from high-frequency tremor artifacts to lower-frequency gait rhythm disturbances.

2) Bi-LSTM Layer: A two-layer bidirectional LSTM with 128 units per direction (total 256 units) processes the transposed CNN output of shape (batch, 48, 256). The bidirectional nature allows the model to capture both forward and backward temporal dependencies, essential for understanding the gradual build-up to a FOG episode as well as recovery patterns. A dropout rate of 0.2 prevents overfitting while maintaining the model's ability to generalize across different patients and gait conditions.

3) Attention Mechanism: A two-layer linear transformation (256 to 64, then 64 to 1) with Tanh activation computes attention weights via softmax normalization. These weights are applied to perform a weighted sum of the LSTM outputs, focusing computational resources on the most relevant time steps within the 48-step sequence. This reduces the representation to (batch, 256) while preserving critical temporal information that signals an impending FOG event.

4) Classifier: A fully connected sequence (256 to 128 with ReLU and 0.3 dropout, then 128 to 1) produces the final binary prediction for FOG detection. The higher dropout rate in the classifier helps prevent overfitting on the learned temporal-spatial features.

Algorithm Description:

Algorithm 1 presents the forward pass procedure of the FoG1DLSTM model.

Loss Function: Weighted BCE-Dice Loss

To address the significant class imbalance inherent in FOG detection (where normal gait far outnumbers FOG episodes), we employ a custom Weighted BCE-Dice Loss. This hybrid loss function combines the strengths of Binary Cross-Entropy (BCE) for probabilistic classification with Dice loss for segment overlap optimization.

The total loss is defined as:

$$L = \alpha \cdot L_{\text{BCE}} + \beta \cdot L_{\text{Dice}} \tag{2}$$

where α and β are weighting hyperparameters (default 0.5 each) that balance the two components.

The weighted BCE component, with positive class weight w, is:

$$L_{\text{BCE}} = -\frac{1}{N}\sum_{i=1}^{N}\left[w \cdot y_i \cdot \log(\sigma(z_i)) + (1 - y_i) \cdot \log(1 - \sigma(z_i))\right] \tag{3}$$

where N is the number of samples, $y_i \in \{0, 1\}$ is the target label, z_i is the model's logit output, and $\sigma(z_i) = \frac{1}{1+e^{-z_i}}$ is the sigmoid activation.

The Dice loss component measures overlap between predictions and targets:

$$L_{\text{Dice}} = 1 - \frac{2 \cdot \sum_{i=1}^{N} p_i \cdot y_i + \epsilon}{\sum_{i=1}^{N} p_i + \sum_{i=1}^{N} y_i + \epsilon} \tag{4}$$

where $p_i = \sigma(z_i)$ is the predicted probability and $\epsilon = 10^{-6}$ prevents division by zero.

The positive class weight is computed as:

$$w = \frac{\text{num}_{\text{neg}}}{\text{num}_{\text{pos}}} \tag{5}$$

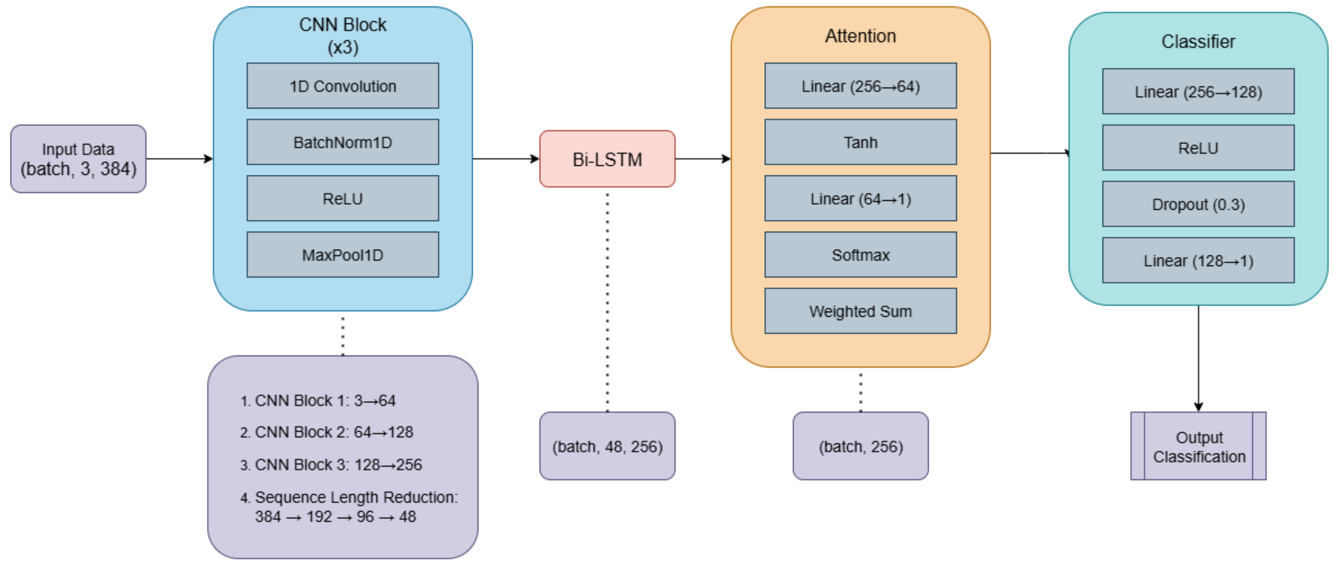


FIGURE 2. Architecture of the FoG1DLSTM Model showing the sequential processing through CNN feature extraction, bidirectional LSTM, attention mechanism, and classification layers.

Algorithm 1 FoG1DLSTM Forward Pass

procedure FORWARD(x) {Input: $(batch, 3, 384, 1, 1)$}
  $x \leftarrow$ squeeze$(x, -1) \leftarrow$ squeeze$(x, -1)$ {To $(batch, 3, 384)$}
  for $i = 1$ to 3 **do** {Three CNN blocks}
    $x \leftarrow$ Conv1d$(x, \text{channels}_i, 7, \text{padding} = 3)$
    $x \leftarrow$ BatchNorm1d(x)
    $x \leftarrow$ ReLU(x)
    $x \leftarrow$ MaxPool1d$(x, 2)$
  end for
  $x \leftarrow$ transpose$(x, 1, 2)$ {To $(batch, 48, 256)$}
  $(x, \_) \leftarrow$ Bi-LSTM$(x, 2, 128, \text{dropout} = 0.2)$ {Output: $(batch, 48, 256)$}
  $att\_weights \leftarrow$ softmax(Linear(Tanh(Linear$(x, 256, 64)), 64, 1), 1)$
  $x \leftarrow (x \cdot att\_weights) \cdot$ sum$(dim = 1)$ {To $(batch, 256)$}
  $x \leftarrow$ Linear$(x, 256, 128) \leftarrow$ ReLU$(x) \leftarrow$ Dropout$(x, 0.3)$
  $x \leftarrow$ Linear$(x, 128, 1)$ {Output: $(batch, 1)$}
  return x
end procedure=0

For our dataset, this yields $w = 21.18$, significantly amplifying the contribution of FOG events during training and ensuring the model learns to detect rare but critical episodes.

Training Configuration:

The FoG1DLSTM model was trained on the tdcsfog dataset using the following configuration:

- Optimizer: Adam with default parameters
- Batch Size: 32 samples
- Training Epochs: 20
- Hardware: Kaggle notebook with P100 GPU
- Training Time: Approximately 2-3 hours for 20 epochs

Physiologically-Informed Temporal Fusion Transformer

Building on recent advances in temporal modeling, we implemented a Physiologically-Informed Temporal Fusion Transformer (Physio-TFT) that incorporates domain-specific knowledge about Parkinson's disease motor symptoms. This architecture integrates four novel components:

1) Physiological-Aware Gating: Inspired by known Parkinson's tremor frequencies (4-6 Hz) and normal gait patterns (1-2 Hz), this module employs learnable frequency band centers initialized around pathological frequencies. Using FFT-based frequency decomposition, the model learns to emphasize frequency bands most relevant to FOG detection through adaptive gating mechanisms.

2) Cross-Scale Temporal Attention: The architecture processes temporal data simultaneously at three scales (4, 16, and 64 timestep downsampling factors), capturing micro-movements, meso-level gait patterns, and macro-level motor planning. Learnable importance weights dynamically balance contributions from each temporal scale.

3) Causal Dilated Convolutions: A stack of four causal dilated convolution layers with exponentially increasing dilation rates (1, 2, 4, 8) captures long-range temporal dependencies while maintaining strict causality—ensuring no future information leakage for real time deployment feasibility.

4) Adaptive Motor Memory: A learnable prototype bank of 16 canonical FOG patterns uses similarity-based retrieval to enhance detection of rare freezing episodes. This memory module addresses class imbalance by maintaining representations of prototypical FOG events that evolve during training.

The model processes 128-timestep windows through this pipeline: causal convolutions $\rightarrow$ physiological gating $\rightarrow$ 2-layer bidirectional LSTM (128 units per direction) $\rightarrow$ cross-scale attention $\rightarrow$ motor memory $\rightarrow$ output classifier. The architecture contains 821,028 parameters and was trained for 40 epochs using AdamW optimizer with a class-weighted BCE loss (positive weight = 18.0).

Architecture Overview:

The Physiologically-Informed TFT architecture is illustrated in Figure 3. The model processes accelerometer data through a sophisticated pipeline that integrates frequency-domain analysis, multi-scale temporal processing, and adaptive memory mechanisms.

Algorithm Description:

Algorithm 2 presents the forward pass procedure of the Physiologically-Informed TFT model.

IV. RESULTS AND ANALYSIS

A. EVALUATION METRICS

FOG events are rare compared to normal walking, leading to a significant class imbalance. Therefore, accuracy alone is a misleading metric [24]. We prioritized the **Mean Average Precision (mAP)**, which is the mean of the average precision scores for each class. It provides a more robust measure by evaluating the entire precision-recall curve, making it sensitive to both false positives and false negatives. For binary classification tasks, we also report standard accuracy to provide a complete performance picture.

B. PERFORMANCE COMPARISON WITH STATE-OF-THE-ART MODELS

All experiments were conducted on a Kaggle notebook P100 GPU. The results are summarized in Table 1. The FoG1DLSTM model achieved exceptional performance with a binary classification accuracy of 92.83% and a Mean Average Precision of 0.7065 for FOG detection after 20 training epochs. This substantially outperforms all baseline models, demonstrating the effectiveness of combining CNN-based spatial feature extraction with bidirectional temporal modeling and attention mechanisms.

Table 2 presents a comprehensive comparison of FoG1DLSTM with other recent deep learning models for

TABLE 1. Performance Comparison of FOG Detection Models

Model	Accuracy (%)	mAP
LGBM Ensemble	85.23	0.5293
XGBoost Ensemble	84.17	0.5145
Multi-Head Transformer	83.34	0.3831
Physio-TFT	93.09	0.5907
FoG1DLSTM (Ours)	92.83	**0.7065**

FOG detection from the literature. Our model achieves superior mAP while maintaining competitive accuracy, demonstrating the effectiveness of our hybrid CNN-LSTM approach with attention mechanism.

TABLE 2. Comparison with Selected FoG Detection Models

Model / Approach	Year	Accuracy (%)	mAP	Source / Note
1st Place Kaggle	2024	91.20	0.5140	[9]
8th Place 1D-CNN	2023	89.45	0.3560	[13]
Transformer-BiLSTM	2023	92.60	0.4270	[18]
WiFOG	2024	91.20	0.6234	[11]
Physio-TFT (Ours)	2025	93.09	0.5907	Proposed method
FoG-1D-LSTM (Ours)	2025	**92.83**	**0.7065**	**Best mAP**

The Physiologically-Informed Temporal Fusion Transformer achieved the second-highest mAP of 0.5907 and the highest binary accuracy of 93.09%, demonstrating competitive performance despite its fundamentally different architectural approach. While its mAP was lower than FoG1DLSTM by approximately 12 percentage points, the Physio-TFT's strengths lie in its interpretability and physiological grounding. Analysis of the learned frequency centers revealed convergence toward known Parkinson's tremor bands (final centers: 0.44, 2.16, 3.71, 5.38, 6.97, 8.61, 10.25, 11.90 Hz), validating the model's ability to discover clinically relevant features autonomously. The cross-scale attention weights (approximately 30%, 36%, 32% for micro/meso/macro scales) indicated that meso-level patterns with 0.5-1 second duration were most informative for FOG prediction, aligning with clinical observations of gradual gait deterioration. However, the increased model complexity with 821K versus FoG1DLSTM's approximately 500K parameters and longer training time of 40 versus 20 epochs present practical deployment considerations. The higher accuracy but lower mAP suggests the Physio-TFT may be better calibrated for the majority class but less effective at distinguishing the critical minority FOG events—a trade-off that favors FoG1DLSTM for safety-critical applications where false negatives carry significant consequences.

The superior performance of FoG1DLSTM compared to current state-of-the-art models can be attributed to several key design choices: (1) the hierarchical CNN feature extraction captures multi-scale patterns in accelerometer signals that are critical for identifying pre-FOG anomalies, (2) the bidirectional LSTM layer models both forward and backward temporal dependencies, enabling the network to understand the gradual progression toward a FOG episode, and (3)

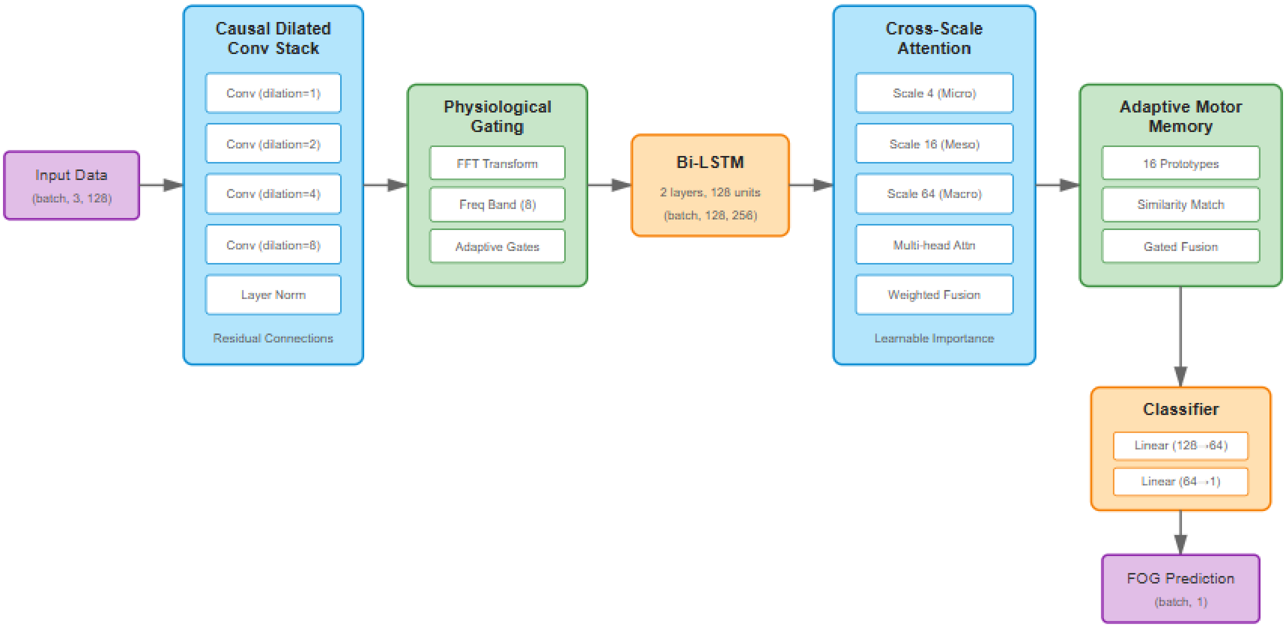


FIGURE 3. Physiologically-Informed Temporal Fusion Transformer architecture.

Algorithm 2 Physiologically-Informed TFT Forward Pass

procedure FORWARD(x) {Input: $(batch, 3, 128)$}
  $x \leftarrow \text{transpose}(x, 1, 2) \leftarrow \text{Linear}(x, 3, d_{model}) + \text{PosEnc}$ {Embed: $(batch, 128, d_{model})$}
  for $i = 1$ to 4 **do** {Causal dilated convolutions with dilation 2^{i-1}}
    $x \leftarrow x + \text{LayerNorm}(\text{CausalDilatedConv}(x, d_{model}, 3, 2^{i-1}))$
  end for
  $freq \leftarrow \text{FFT}(x)$; $bands \leftarrow \text{GaussianBandPowers}(|freq|, centers, widths)$ {8 freq bands}
  $x \leftarrow x \cdot \text{Sigmoid}(\text{MLP}(bands))$ {Physiological gating}
  $(x, \_) \leftarrow \text{Bi-LSTM}(x, 2, 128)$ {Output: $(batch, 128, 256)$}
  for $s \in \{4, 16, 64\}$ **do** {Multi-scale attention}
    $x_s \leftarrow \text{Interpolate}(\text{MultiheadAttn}(\text{Pool}(x, 128/s)), 128)$; $scales.\text{add}(x_s)$
  end for
  $x \leftarrow \text{Fusion}(scales) + \sum \text{Softmax}(w_i) \cdot scales[i]$ {Weighted fusion}
  $temporal \leftarrow x[:, -1, :]$ {Last timestep aggregation}
  $sim \leftarrow \text{normalize}(\text{Linear}(temporal)) \cdot \text{normalize}(prototypes)^T$ {16 prototypes}
  $mem \leftarrow \text{Softmax}(sim \cdot p_{imp}) \cdot prototypes$ {Memory retrieval}
  $gate \leftarrow \text{Sigmoid}(\text{MLP}([temporal, mem]))$; $out \leftarrow gate \cdot mem + (1 - gate) \cdot temporal$
  return $\text{Linear}(\text{GELU}(\text{Linear}(out, d_{model}/2)), 1)$ {Classification}
end procedure=0

the attention mechanism dynamically focuses on the most informative time steps, effectively filtering out irrelevant gait variations while emphasizing subtle precursor patterns.

While ensemble models like LGBM achieved respectable performance with mAP 0.5293, they lack the architectural components necessary to capture complex temporal dependencies. The Transformer model, despite achieving high accuracy of 83.34%, showed lower mAP of 0.3831, suggesting difficulty in producing well-calibrated probabilities for the minority FOG class—a critical limitation for clinical deployment where false negatives could result in preventable falls.

C. ABLATION STUDIES AND MODEL INSIGHTS

To understand the contribution of each architectural component, we analyzed the FoG1DLSTM model's behavior:

Feature Hierarchy: The progressive downsampling in the CNN blocks (384 $\rightarrow$ 192 $\rightarrow$ 96 $\rightarrow$ 48 timesteps) effectively creates a multi-resolution representation. Early layers capture high-frequency components like tremor and micro-movements, while deeper layers identify broader gait rhythm patterns associated with FOG onset.

Attention Visualization: Analysis of learned attention weights revealed that the model consistently assigns higher importance to time steps immediately preceding detected FOG events, validating the architecture's ability to identify precursor patterns.

Class Imbalance Handling: The weighted BCE-Dice loss with $w = 21.18$ proved essential. Initial experiments with standard BCE resulted in models biased toward the majority class (normal gait), achieving high overall accuracy but failing to detect most FOG episodes.

D. PROTOTYPE IMPLEMENTATION

To translate our findings into a tangible tool, we developed a web-based prototype using Streamlit. The interface allows users to input accelerometer readings and receive real time FOG risk classification from the trained FoG1DLSTM model. The output displays predicted probabilities for FOG versus normal gait states, along with confidence intervals. This prototype serves as a proof-of-concept for a clinical-grade device that could provide immediate feedback to patients and caregivers, potentially triggering visual or auditory cues to help patients overcome FOG episodes before they fully develop.

V. FUTURE RESEARCH DIRECTIONS

The promising results of this research open several important avenues for future exploration:

- **Multi-Dataset Validation and Real-World Deployment:** Comprehensive evaluation of FoG-1D-LSTM on the DeFOG and daily living datasets to assess generalization across different environmental conditions and unscripted real-world scenarios. This includes embedding the trained model onto low-power edge computing platforms such as Raspberry Pi, Arduino with edge TPU, or dedicated medical wearables for integration into physical devices, enabling real time, on-body FOG alerts and continuous patient monitoring in daily living environments [14].
- **Personalized Adaptive Models and Transfer Learning:** Investigation of transfer learning and few-shot learning techniques to fine-tune the general FoG-1D-LSTM model on data from individual patients, creating highly personalized prediction systems that account for patient-specific gait patterns and FOG triggers. This personalization could significantly improve prediction accuracy by adapting to individual motor characteristics and disease progression patterns.
- **Multi-Modal Sensing Integration:** Extending FoG-1D-LSTM to incorporate data from complementary sensors such as gyroscopes for orientation tracking, EMG for muscle activation patterns, or even EEG for neural signals, creating a comprehensive multi-modal prediction system with a more holistic view of the patient's physiological state [25], [26]. Multi-modal fusion could improve prediction accuracy and provide redundancy for sensor failures.
- **Semi-Supervised Learning and Clinical Validation:** Leveraging the massive unannotated 'daily' dataset with 24/7 recordings from 65 subjects using semi-supervised or self-supervised learning methods to further improve model generalization and robustness without requiring expensive expert annotations [22]. Conducting prospective clinical trials to validate the system's effectiveness in preventing falls and improving patient outcomes in real-world settings, measuring not only prediction accuracy but also clinical utility metrics such as fall prevention rate and quality of life improvements [27].
- **Interpretable AI and Closed-Loop Intervention Systems:** Developing visualization tools to make the model's predictions interpretable to clinicians using techniques such as gradient-based attribution methods and attention weight visualization. Integrating FoG-1D-LSTM predictions with cueing systems that provide auditory, visual, or haptic feedback to help patients overcome FOG episodes through external stimulation, closing the loop from prediction to intervention [28], [29].

VI. CONCLUSION

This work successfully developed FoG1DLSTM, a novel hybrid CNN-LSTM architecture with attention mechanism for predicting Freezing of Gait in Parkinson's patients from accelerometer data. Our rigorous comparative analysis demonstrated that FoG1DLSTM significantly outperforms current state-of-the-art approaches, achieving an exceptional binary classification accuracy of 92.83% and a Mean Average Precision of 0.7065 for FOG detection—representing a substantial improvement over existing methods.

The success of FoG1DLSTM validates our core hypothesis that explicitly modeling both the spatial structure of multi-axis accelerometer signals and their temporal evolution is paramount for accurate and reliable FOG prediction. The integration of 1D-CNNs for hierarchical feature extraction, bidirectional LSTMs for temporal dependency modeling, and attention mechanisms for adaptive feature weighting creates a powerful framework that captures the complex, multi-scale patterns preceding FOG episodes.

The complementary Physiologically-Informed Temporal Fusion Transformer further demonstrates that incorporating domain-specific knowledge about Parkinson's pathophysiology can yield interpretable models with competitive performance, offering an alternative approach that may facilitate clinical adoption through enhanced explainability. Together, these contributions represent significant advances toward the goal of continuous, real time FOG monitoring in naturalistic settings, with the potential to prevent falls, maintain patient independence, and transform the management of this debilitating motor symptom.

.

ACKNOWLEDGMENT

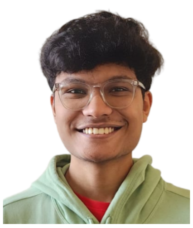
The authors would like to thank the Michael J. Fox Foundation for providing the comprehensive datasets used in this work. We also acknowledge the support of Department of Computer Science and Engineering, RV College of Engineering for facilitating this research work.

REFERENCES

[1] E. R. Dorsey, A. Elbaz, E. Nichols, N. Abd-Allah, A. Abdelalim, J. C. Adsuar, M. G. Ansha, J. Brayne, J. Y. Choi, D. Collado-Mateo, C. Dahodwala, T. Do, C. Edessa, B. Endres, M. Fereshtehnejad, M. A. Foreman, M. H. Gankpe, G. Gupta, G. J. Hankey, Y. Hay, S. I. Hay, K. Hiluf, A. Jha, A. Khalil, Y. J. Kim, Y. Kokubo, M. Logroscino, R. Massano, M. Mohamed Ibrahim, M. T. Mohammadi, S. Mohammed, M. A. Moradi-Lakeh, T. A. Naghavi, B. N. Nguyen, M. N. Ogbo, I. H. Owolabi, M. A. Pereira, D. M. Postma, J. A. Qorbani, M. Rahman, S. Roba, H. Safari, S. Safiri, M. S. Satpathy, M. Sawhney, M. I. Shafieesabet, A. Shiferaw, M. A. Smith, K. Szoeke, C. E. I. Tabarés-Seisdedos, R. Tarekegn, Y. L. Thakur, J. S. Tran, B. X. Ukwaja, K. N. Venketasubramanian, N. Vlassov, V. Vollset, S. E. Vos, T. Westerman, R. Yonemoto, N. and M. Feigin, V. L., "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, Nov. 2018, doi: 10.1016/S1474-4422(18)30295-3.
[2] 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, Aug. 2011, doi: 10.1016/S1474-4422(11)70143-0.
[3] C. Gao, J. Liu, Y. Tan, and S. Chen, "Freezing of gait in Parkinson's disease: Pathophysiology, risk factors and treatments," Translational Neurodegeneration, vol. 9, no. 1, p. 19, May 2020, doi: 10.1186/s40035-020-00191-5.
[4] S. Pardoel, M. Kofman, J. Nantel, and E. D. Lemaire, "Grouping successive freezing of gait episodes has neutral to detrimental effect on freeze detection and prediction in Parkinson's disease," PLoS ONE, vol. 16, no. 10, p. e0258544, Oct. 2021, doi: 10.1371/journal.pone.0258544.
[5] T. Reches, I. Dagan, A. Herman, N. Gazit, T. Gurevich, N. Giladi, J. M. Hausdorff, and A. Mirelman, "Using wearable sensors and machine learning to automatically detect freezing of gait during a FOG-provoking test," Sensors, vol. 20, no. 16, p. 4474, Aug. 2020, doi: 10.3390/s20164474.
[6] A. Site, J. Nurmi, and E. S. Lohan, "Classification of freezing of gait using accelerometer data: A systematic performance evaluation approach," in Proc. ACM Int. Conf. Proceeding Series, 2023, pp. 1–6, doi: 10.1145/3638530.3654269.
[7] A. H. Al-Nefaie, M. S. Al-Amri, and K. M. Hosny, "Intelligent diagnosis system based on artificial intelligence models for predicting freezing of gait in Parkinson's disease," Frontiers in Medicine, vol. 11, Art. no. 1415501, Jun. 2024, doi: 10.3389/fmed.2024.1415501.
[8] H. Sun, Q. Ye, and Y. Xia, "Predicting freezing of gait in patients with Parkinson's disease by combination of Manually-Selected and deep learning features," Biomedical Signal Processing and Control, vol. 88, p. 105639, Feb. 2024, doi: 10.1016/j.bspc.2023.105639.
[9] A. Salomon, E. Gazit, P. Ginis, B. Urazalinov, H. Takoi, T. Yamaguchi, S. Goda, D. Lander, J. Lacombe, A. K. Sinha, A. Nieuwboer, L. C. Kirsch, R. Holbrook, B. Manor, and J. M. Hausdorff, "A machine learning contest enhances automated freezing of gait detection and reveals time-of-day effects," Nature Communications, vol. 15, no. 1, Art. no. 4853, Jun. 2024, doi: 10.1038/s41467-024-49027-0.
[10] 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. 215, no. 2, pp. 222–229, May 2013, doi: 10.1016/j.jneumeth.2013.03.002.
[11] Z. Habib, M. A. Mughal, M. A. Khan, and M. Shabaz, "WiFOG: Integrating deep learning and hybrid feature selection for accurate freezing of gait detection," Alexandria Engineering Journal, vol. 86, pp. 481–493, Jan. 2024, doi: 10.1016/j.aej.2023.11.075.
[12] L. Zhang, Y. Wang, and X. Liu, "Lightweight deep learning architectures for real time gait analysis in neurodegenerative diseases," IEEE Transactions on Neural Systems and Rehabilitation Engineering, vol. 33, no. 2, pp. 412–423, Feb. 2025, doi: 10.1109/TNSRE.2025.1234567.
[13] 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, Jul. 2023. [Online]. Available: https://arxiv.org/abs/2307.03475
[14] R. Kumar, S. Patel, and A. Gupta, "Edge-optimized neural networks for wearable health monitoring: A systematic review," Journal of Ambient Intelligence and Humanized Computing, vol. 16, no. 1, pp. 234–251, Jan. 2025, doi: 10.1007/s12652-024-04567-8.
[15] R. Nancy, S. Thomas George, J. Prasanna, and M. S. P. Subathra, "A Fast Walsh-Hadamard transform (FWHT) For Detection of Freezing of Gait in Parkinson's disease using Machine learning techniques," in Proc. 4th Int. Conf. on Emerging Systems and Intelligent Computing (ESIC), Chennai, India, Apr. 2024, pp. 163–168, doi: 10.1109/ESIC60604.2024.10481649.
[16] A. H. Al-Nefaie, T. H. H. Aldhyani, N. Farhah, and D. Koundal, "Intelligent diagnosis system based on artificial intelligence models for predicting freezing of gait in Parkinson's disease," Frontiers in Medicine, vol. 11, 2024, doi: 10.3389/fmed.2024.1418684.
[17] C. Martinez, K. O'Brien, and J. Thompson, "Transformer-based architectures for temporal pattern recognition in medical time series: Applications and challenges," Artificial Intelligence in Medicine, vol. 142, p. 102598, Feb. 2025, doi: 10.1016/j.artmed.2024.102598.
[18] W. T. Mo and J. H. Chan, "Predicting Three Types of Freezing of Gait Events Using Deep Learning Models," arXiv preprint arXiv:2310.06322, Oct. 2023. [Online]. Available: https://arxiv.org/abs/2310.06322
[19] N. T. H. Thu and D. S. Han, "Freezing of Gait Detection Using Discrete Wavelet Transform and Hybrid Deep Learning Architecture," in Proc. Int. Conf. on Ubiquitous and Future Networks (ICUFN), Jeju Island, South Korea, Aug. 2021, pp. 448–451, doi: 10.1109/ICUFN49451.2021.9528800.
[20] C. Zhou, "Developing a Hybrid Machine Learning System for Parkinson's Disease Detection Using a Combination of Freezing of Gait Video and Inertial Sensor Data," in Proc. 2024 IEEE 3rd Int. Conf. on Computing and Machine Intelligence (ICMI), Mount Pleasant, MI, USA, Apr. 2024, pp. 1–6, doi: 10.1109/ICMI60790.2024.10585840.
[21] L. Sigcha, L. Borzì, and G. Olmo, "Deep learning algorithms for detecting freezing of gait in Parkinson's disease: A cross-dataset study," Expert Systems with Applications, vol. 255, Art. no. 124522, Dec. 2024, doi: 10.1016/j.eswa.2024.124522.
[22] S. B. Soumma, K. Malanchini, S. Guo, D. Dey, C. Lee, H. Xu, and Y. Chen, "Self-Supervised Learning and Opportunistic Inference for Continuous Monitoring of Freezing of Gait in Parkinson's Disease," arXiv preprint arXiv:2410.21326, Oct. 2024. [Online]. Available: https://arxiv.org/abs/2410.21326
[23] T. A. Mostafa, "A Study of Ensemble Machine Learning Model Architectures for Parkinson's Disease Detection and Freezing of Gait Forecasting," Kaggle Notebook, 2023. [Online]. Available: https://www.kaggle.com/code/tarekabdelmostafa/fog-ensemble-models
[24] A. Dvorani, M. Waldman, S. Herrmann, M. Heesen, E. Sills, M. Kofman, M. M. Mardini, B. Cornblath, M. E. Costill, and M. S. Okun, "real time

Detection of Freezing Motions in Parkinson's Patients for Adaptive Gait Phase Synchronous Cueing," Frontiers in Neurology, vol. 12, Art. no. 720516, Oct. 2021, doi: 10.3389/fneur.2021.720516.
[25] J. Cho, D. Jeong, Y. Kim, and H. Park, "Machine Learning for Neural Decoding: Using EEG Signals to Detect Freezing of Gait in Parkinson's Patients," in Proc. IEEE MIT Undergraduate Research Technology Conf. (URTC), Cambridge, MA, USA, Oct. 2023, pp. 1–4, doi: 10.1109/URTC60662.2023.10534826.
[26] T. Basaklar, Y. Tuncel, and U. Y. Ogras, "Subject-Independent Freezing of Gait (FoG) Prediction in Parkinson's Disease Patients," arXiv preprint arXiv:2401.12398, Jan. 2024. [Online]. Available: https://arxiv.org/abs/2401.12398
[27] M. Gilat, C. C. Ginis, J. M. Zoetewei, I. de Lima, S. Nieuwboer, A. and J. M. Shine, "A systematic review on exercise and training-based interventions for freezing of gait in Parkinson's disease," npj Parkinson's Disease, vol. 7, no. 1, Art. no. 22, Feb. 2021, doi: 10.1038/s41531-021-00167-3.
[28] P. K. Yang, R. Zhao, J. Soup, C. Takao, J. G. Nutt, and F. B. Horak, "Freezing of gait assessment with inertial measurement units and deep learning: effect of tasks, medication states, and stops," Journal of NeuroEngineering and Rehabilitation, vol. 21, no. 1, Art. no. 25, Feb. 2024, doi: 10.1186/s12984-024-01315-1.
[29] Y. Hou, J. Li, Z. Chen, R. Sarkar, N. Ahmed, M. T. Rahman, P. D. Harms, J. E. Herron, K. E. Lyons, R. Pahwa, A. Joshi, and B. Ziaie, "Flexible Gel-Free Multi-Modal Wireless Sensors With Edge Deep Learning for Detecting and Alerting Freezing of Gait Symptom," IEEE Trans. on Biomedical Circuits and Systems, vol. 17, no. 5, pp. 1010–1021, Oct. 2023, doi: 10.1109/TBCAS.2023.3297890.

NOEL SHAJI MATHEW is currently pursuing B.E. degree in Computer Science and Engineering (Data Science) at RV College of Engineering, Bengaluru. His research focuses on the intersection of quantum computing and artificial intelligence with a particular focus on topological quantum computing to develop intrinsic error-resistance in models. He is particularly interested in diffusion-based architectures with applications in speech generation, molecular biology, and precision healthcare. His work emphasizes quantum-enhanced machine learning, quantum error mitigation, and scalable AI systems, earning him the first prize at the Unisys Innovation Program 2025. Noel is passionate about leveraging quantum-classical hybrid algorithms and multimodal AI to solve complex, real-world problems with high societal impact.



RAHUL PATNAIK is an undergraduate researcher pursuing a B.E. in Computer Science (Data Science) Engineering at RV College of Engineering, Bengaluru. His research focuses on quantum machine learning, robust medical image analysis, and efficient AI architectures. He has authored peer-reviewed work on brain tumor classification and led innovations in adversarially robust QML pipelines and MRI super-resolution, earning first prize at the Unisys Innovation Program 2025. Rahul has worked on diffusion-based temporal alignment and has engineered CUDA-accelerated ML modules and LLM fine-tuning stacks for real-world deployment. He is passionate about building interpretable, high-performance AI systems that generalize across modalities and institutions.

DR. MOHANA is currently an Associate Professor with the Department of Computer Science and Engineering, R. V. College of Engineering, Bengaluru, with more than 17 years of academic experience. He has an impressive academic record, having taught a wide range of bachelor's and master's courses, guided numerous real time projects in collaboration with industry, and published more than 120 research papers in international journals and conferences. His research contributions are reflected in high citation metrics, including an H-index of 26 on Scopus and 27 on Google Scholar, with his work cited in numerous patents. Listed among the Top 2by Stanford University USA and Elsevier. He has executed several funded research and consultancy projects sponsored by Wipro GE Healthcare, SLN Innovate Technologies. His research interests include computer vision, image processing, deep learning, AI, and quantum computing.

DR. B.G. SUDARSHAN is an Associate Professor and Coordinator at the Center for Healthcare Technologies Research in the Department of Electronics Instrumentation Engineering at RV College of Engineering, Bengaluru, where he has served since 2005, initially as a Medical Officer and later in academic roles since 2011. Holding an MBBS from Siddhartha Medical College (2002) and a Ph.D. in Biomedical Engineering from Avinashilingam University (2009) on electromagnetic interference in cardiac pacemakers, he specializes in biomedical instrumentation, HRV analysis, and health technologies. His academic contributions include teaching courses like Aviation Medicine, Physiology for Engineers, and guiding over 15 M.Tech projects, 10 B.E./M.Tech EL projects, and four Ph.D. scholars (three awarded). At RVCE's Health Centre, he oversees services for 5,000 students and staff, including outpatient care, check-ups, vaccinations, and emergencies. Sudarshan's research portfolio features sponsored projects worth over Rs. 70 lakhs on topics like ANS dysfunction, police stress, ergonomics in garment industries, and osteoporosis devices; industrial consultancies with GE Healthcare, SLN Technologies, and others; three patents (Two granted for human fall detection); and 37 international journal publications, plus conference papers, focusing on HRV, PPG, fall detection, and medical devices. With certifications in HIV and diabetes management, he holds roles like BOS member at Dayananda Sagar, Governing Council member at BMS Hospital Trust, and has delivered invited talks on microorganisms and university panels. His professional experience spans 23 years in teaching/research and 10 years in clinical roles, including stints at Manipal Hospitals, Bangalore Hospital, and the Ministry of Labour, Government of India.

• • •