

## Article

# Explainable Gait Multi-Anchor Space-Aware Temporal Convolutional Networks for Gait Recognition in Neurological, Orthopedic, and Healthy Cohorts

Abdullah Alharthi 

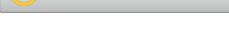
Department of Electrical Engineering, King Khalid University, Abha 62521, Saudi Arabia; asalharthi@kku.edu.sa

## Abstract

Gait recognition using wearable sensor data is crucial for healthcare, rehabilitation, and monitoring neurological and musculoskeletal disorders. This study proposes a deep learning framework for gait classification using inertial measurements from four body-mounted IMU sensors (head, lower back, and both feet). The data were collected from a publicly available, clinically annotated dataset comprising 1356 gait trials from 260 individuals with diverse pathologies. The framework, G-MASA-TCN (Gait Multi-Anchor, Space-Aware Temporal Convolutional Network), integrates multi-scale temporal fusion, graph-informed spatial modeling, and residual dilated convolutions to extract discriminative gait signatures. To ensure both high performance and interpretability, Integrated Gradients is incorporated as an explainable AI (XAI) method, providing sensor-level and temporal attributes that reveal the features driving model decisions. The framework is evaluated via repeated cross-validation experiments, reporting detailed metrics with cross-run statistical analysis (mean  $\pm$  standard deviation) to assess robustness. Results show that G-MASA-TCN achieves 98% classification accuracy for neurological, orthopedic, and healthy cohorts, demonstrating superior stability and resilience compared to baseline architectures, including Gated Recurrent Unit (GRU), Transformer neural networks, and standard TCNs, and 98.4% accuracy in identifying individual subjects based on gait. Furthermore, the model offers clinically meaningful insights into which sensors and gait phases contribute most to its predictions. This work presents an accurate, interpretable, and reliable tool for gait pathology recognition, with potential for translation to real-world clinical settings.

**Keywords:** Multi-Anchor Space-Aware Temporal Convolutional Neural Network (MASA-TCN); Clinical Gait; Multimodal Data; Integrated Gradients; Gated Recurrent Unit (GRU); Transformer Neural Networks; Neurological; Orthopedic

**MSC:** 37M10



Academic Editor: Cheng Tang

Received: 22 November 2025

Revised: 20 December 2025

Accepted: 5 January 2026

Published: 8 January 2026

**Copyright:** © 2026 by the author.

Licensee MDPI, Basel, Switzerland.

This article is an open access article distributed under the terms and conditions of the [Creative Commons Attribution \(CC BY\) license](#).

## 1. Introduction

Human locomotion, particularly walking, represents a cornerstone of daily human function, requiring seamless integration among musculoskeletal, neuromotor, and sensory systems; alterations in gait biomechanics are highly sensitive biomarkers for underlying physiological and neurological pathologies [1,2]. The accelerating global demographic shift toward aging has precipitated a surge in gait-associated morbidities, encompassing osteoarthritis, post-stroke hemiparesis, Parkinsonian gait freezing, and sarcopenic decline [3,4]. Such disorders profoundly impair mobility, erode autonomy, and elevate

susceptibility to falls, institutionalization, and premature death [5]. Consequently, proactive gait surveillance and early anomaly detection emerge as pivotal strategies in precision preventive medicine and neurorehabilitation [2,6].

Gait assessment traditionally relies on laboratory-based motion capture systems, force platforms, and instrumented treadmills [7,8]. While these tools provide precise biomechanical measurements such as joint angles, ground reaction forces, and step timing, their deployment is limited by high cost, technical complexity, and the inability to capture naturalistic, daily-life walking patterns [9,10]. Consequently, many clinically relevant gait abnormalities remain undetected until substantial functional decline occurs [11]. Wearable inertial measurement units (IMUs) have emerged as practical alternatives, capable of capturing continuous tri-axial acceleration and angular velocity data from key body segments such as the feet, shanks, or trunk [12,13]. These devices allow high-frequency motion data collection in real-world environments, enabling the detection of subtle gait deviations that may not manifest in laboratory assessments [2]. Recent studies validate the accuracy of IMUs against optical motion capture systems, demonstrating reliable measurement of stride length, cadence, and variability [14].

Despite their advantages, IMU datasets are high-dimensional, temporally complex, and often noisy, necessitating advanced computational approaches for meaningful interpretation [14,15]. Classical machine learning approaches such as support vector machines, random forests, and k-nearest neighbors rely heavily on handcrafted features, limiting their capacity to model complex, nonlinear temporal dependencies across multiple sensor axes [14]. Deep learning approaches, including convolutional neural networks (CNNs), recurrent neural networks (RNNs), long short-term memory (LSTM) networks, and Transformer architectures, have demonstrated superior capability in learning hierarchical features from raw time-series data [7,16,17]. CNNs are particularly effective for learning local temporal and spatial features, while LSTMs excel in modeling long-term temporal dependencies [16,17]. Transformers, utilizing self-attention mechanisms, can capture global dependencies across sequences but demand larger datasets and substantial computational resources [18,19].

Temporal Convolutional Networks (TCNs) offer an efficient alternative, employing dilated causal convolutions to capture long-range temporal dependencies while maintaining manageable computational cost [20,21]. However, standard TCNs often collapse multi-axis sensor data, losing critical spatial relationships such as inter-limb coordination and cross-joint interactions [14,22]. Preserving both spatial and temporal correlations is essential for accurate classification of gait conditions, particularly when subtle deviations reflect early-stage neurological or musculoskeletal impairments [11]. Insights from human-in-the-loop exoskeleton studies further underscore the importance of detailed spatiotemporal modeling. Exoskeleton-assisted walking has demonstrated that optimizing torque profiles based on real-time feedback can improve walking economy, reduce metabolic cost, and enhance gait symmetry [23,24]. Notably, adaptive multi-joint assistance based on continuous monitoring of joint angles and muscle activation has enabled personalized intervention strategies that improve mobility outcomes [25]. These findings emphasize the necessity of capturing multi-scale spatial and temporal features when designing AI-based gait assessment models.

The Temporal Convolutional Network (TCN) represents a recent advancement in deep learning for multivariate time-series classification, particularly in EEG for emotion recognition [26,27], and for action segmentation and detection [20]. The TCN extends conventional TCNs by incorporating multiple temporal anchors with distinct receptive fields, preserving both local and global temporal information while maintaining spatial correlations across sensor channels [28,29]. The model further integrates attention mechanisms to

dynamically weigh contributions from each temporal and spatial anchor, enhancing feature interpretability and robustness [30,31]. This architecture has demonstrated superior performance in classifying complex spatial and temporal datasets compared to conventional CNN, LSTM, and standard TCN models [14,20].

In addition to spatial and temporal datasets, TCN and related architectures are increasingly applied in brain–computer interface (BCI) systems for emotion recognition, cognitive workload estimation, and neural decoding [28,29,32]. These applications highlight the versatility of multi-scale, space-aware temporal models in extracting meaningful features from high-dimensional, temporally structured signals [28,33]. For instance, deep convolutional and graph convolutional neural networks have been successfully applied to EEG datasets, capturing local–global interactions essential for reliable decoding of neural signals [32,34]. Transformer-based architectures have further improved performance by leveraging hierarchical attention mechanisms, enabling simultaneous modeling of spatial and temporal dependencies [27,29,35,36].

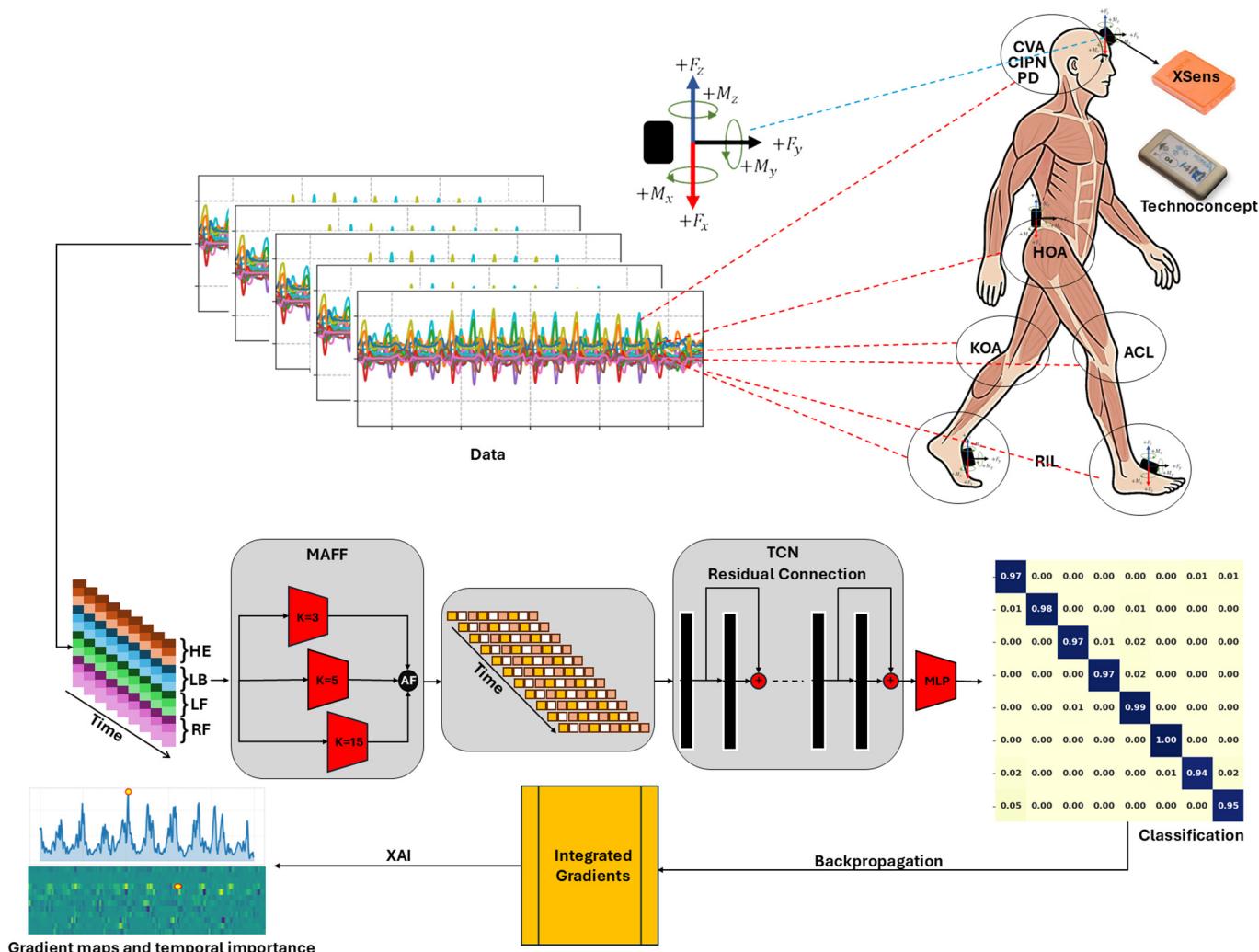
Despite its strong performance in modeling complex neural dynamics, Multi-Anchor Space-Aware Temporal Convolutional Network (MASA-TCN) has been used exclusively in EEG-based emotion recognition and has never been applied to human gait analysis [26]. This represents a notable gap, particularly given the critical role gait plays as an indicator of health, aging, fall risk, frailty, and neurological decline [1–6]. Wearable IMU-based gait analysis has become central to mobility assessment, enabling the extraction of stride parameters, joint coordination patterns, and dynamic balance metrics that are clinically meaningful and responsive to disease progression [7–15,37].

Introducing MASA-TCN into this domain offers several key advantages. First, its multi-anchor temporal convolutions are well suited to capture the fine-grained temporal structure of gait cycles, including stride timing, swing–stance transitions, and stride-to-stride variability features that are strongly linked to gait stability and fall risk [1,2,9]. Second, the model’s spatial-attention fusion provides a principled way to identify the most informative IMU channels and motion axes, supporting interpretable analysis of how specific body segments contribute to normal or pathological gait patterns, consistent with findings across sensor-based gait studies [10–15]. Third, MASA-TCN’s lightweight, TCN-driven architecture enables efficient real-time inference, aligning with advances in wearable-sensor gait tracking and activity recognition systems designed for continuous monitoring in real-world settings [21,38–40].

Taken together, these factors position MASA-TCN as a novel and technically well-justified architecture for gait analysis. Its proven ability to model multi-scale temporal-spatial patterns in EEG, combined with the clinical importance of gait and the growing maturity of wearable sensing, provides a strong foundation for applying MASA-TCN to gait datasets for the first time. This also paves the way for integrating MASA-TCN with XAI methods to deliver transparent, clinically interpretable gait assessments, an essential requirement for future healthcare and mobility-monitoring applications.

Wearable exoskeleton studies have further informed the design and evaluation of AI-based gait models. For instance, adaptive hip exosuits have been shown to reduce metabolic cost during walking and running, while ankle and knee exoskeletons improve stability and joint-specific assistance [24,25,41]. Human-in-the-loop optimization enables dynamic adjustment of torque profiles based on user-specific biomechanics, emphasizing the importance of accurate multi-channel sensor modeling [42]. These findings reinforce the notion that effective AI-based gait analysis must account for both local joint-level dynamics and global locomotor coordination, which MASA-TCN accomplishes through its multi-anchor, space-aware design [26,43]. This study’s contributions, as shown in Figure 1, are the following:

- Development of G-MASA-TCN, a unified deep learning model for multi-condition gait classification and individual identification.
- Space-aware temporal layers for spatial–spectral learning across multiple IMU channels.
- Multi-anchor attentive fusion for capturing dynamic gait events at multiple temporal scales.
- Evaluation on 260 real patients with eight clinically confirmed conditions, demonstrating high classification and identification accuracy.
- Comparative analysis with standard TCN, Gated Recurrent Unit (GRU), and Transformer neural network-based architectures.



**Figure 1.** Overview of the proposed G-MASA-TCN framework, consisting of three components: the MAAF block, the TCN block, and the classification block. IMU sensors are positioned using Velcro bands on the lower back (L5), forehead, and both feet (upper sensors: XSens; lower sensors: Technoconcept). Example gait sub-segments are shown for multiple conditions: HOA, KOA, ACL, CVA, PD, CIPN, and RIL. The parameter  $k$  in the SAT module denotes the temporal kernel length. Backpropagation is included for XAI Gradient maps and temporal importance. Where red dashed lines collected signal locations, black arrows data flow in the framework.

G-MASA-TCN offers a unified approach for analyzing human gait, capturing complex spatiotemporal patterns, and enabling accurate, clinically relevant classification and identification. It advances automated gait analysis by overcoming limitations of prior deep learning methods, providing both rich spatial–temporal modeling and high recognition accuracy.

The article is organized as follows. Section 2 describes the experimental design, participants, gait signal data, and preprocessing methods. It also presents the G-MASA-

TCN with the deep neural network architecture. Sections 3 and 4 present and analyze the results obtained. Finally, Section 5 offers conclusions and suggests potential directions for future research.

## 2. Materials and Methods

### 2.1. Data

This study employs a publicly accessible dataset titled Clinical Gait Signals with Wearable Sensors from Healthy, Neurological, and Orthopedic Cohorts [37]. The dataset includes multimodal gait recordings captured using four inertial measurement units (IMUs) to analyze human locomotion across multiple clinical populations.

Participants were recruited from hospitals affiliated with the Assistance Publique des Hôpitaux de Paris, France and the Hôpital d’Instruction des Armées, France. The sample comprised both healthy volunteers and patients from neurological and orthopedic departments. The control group, healthy subjects (HSs), included 73 adults (41 males, 32 females; ages 18–87) with no medical conditions confirmed through clinical examination. The patient population was divided into two major cohorts:

- The orthopedic cohort (44 participants) with diagnoses of hip osteoarthritis (HOA), knee osteoarthritis (KOA), or anterior cruciate ligament injury (ACL).
- The neurological cohort (143 participants) encompassing cerebrovascular accident (CVA), Parkinson’s disease (PD), chemotherapy-induced peripheral neuropathy (CIPN), and radiation-induced leukoencephalopathy (RIL).

Each participant underwent a standardized clinical assessment using established scoring systems corresponding to their condition (see Table 1). Differences in age distributions across cohorts reflect disease-specific characteristics rather than sampling bias. All procedures complied with ethical standards approved by the Comité de Protection des Personnes Ile de France II (CPP 2014-10-04 RNI). Written informed consent was obtained from all participants for data collection and sharing [37].

**Table 1.** Summary of clinical or radioclinical scores calculated by the physician for each pathology. HOA: hip osteoarthritis; KOA: knee osteoarthritis; ACL: anterior cruciate ligament injury; CVA: cerebrovascular accident; PD: Parkinson’s disease; CIPN: chemotherapy-induced peripheral neuropathy; RIL: radiation-induced leukoencephalopathy, HSs: healthy subjects [37].

Cohort	Score Name	Short Scale	Scale Range	Description
HOA/KOA	Western Ontario and McMaster Universities Osteoarthritis Index	WOMAC	100	Reference index for assessing symptoms (pain and stiffness) and functional limitations in patients with osteoarthritis, particularly for the hip and knee joints. The higher the score, the more severe the symptoms and limitations.
ACL	International Knee Documentation Committee Evaluation Form	IKDC	100	Clinical score assessing patients’ subjective perceptions, commonly used for various knee pathologies, especially after surgery or during rehabilitation. The higher the score, the less severe the symptoms.
CVA	Fugl-Meyer Assessment—Lower Extremity Subscore	FMA-LE	34	Stroke-specific disability index for the lower limb that evaluates movement, coordination, and reflex actions at the hip, knee, and ankle. The higher the score, the greater the preserved motor function.

**Table 1.** Cont.

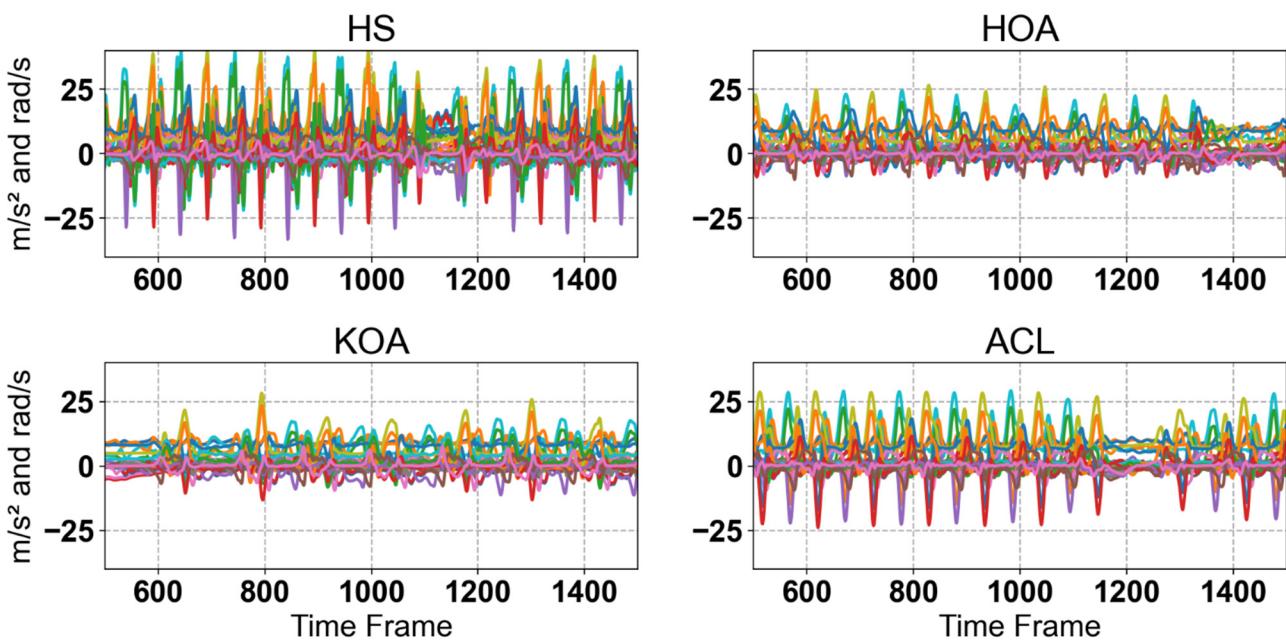
Cohort	Score Name	Short Scale	Scale Range	Description
PD	Unified Parkinson's Disease Rating Scale—Part III	UPDRS III	108	Parkinson's disease-specific clinical index that assesses disease progression and its impact on daily life. Part III focuses on motor symptoms. The higher the score, the more severe the disease and its effects.
CIPN	Total Neuropathy Score—Clinical Version	TNSc	28	Clinical form of the Total Neuropathy Score used to quantify motor and sensory impairment in peripheral neuropathy, without requiring electrophysiological testing. The higher the score, the more severe the neuropathy.
RIL	Original Clinical Score	OCS	12	Radioclinical score based on four assessments: dysexecutive disorder, gait apraxia, urinary incontinence, and white matter MRI involvement. Each dimension has five possible ratings (0 = None; 1 = Mild; 2 = Moderate; 3 = Severe). The higher the score, the more severe the symptoms and radiological damage.
HSs	-	-	-	Healthy subjects who entered the hospitals for treatment or accompaniment

### Sensor Configuration

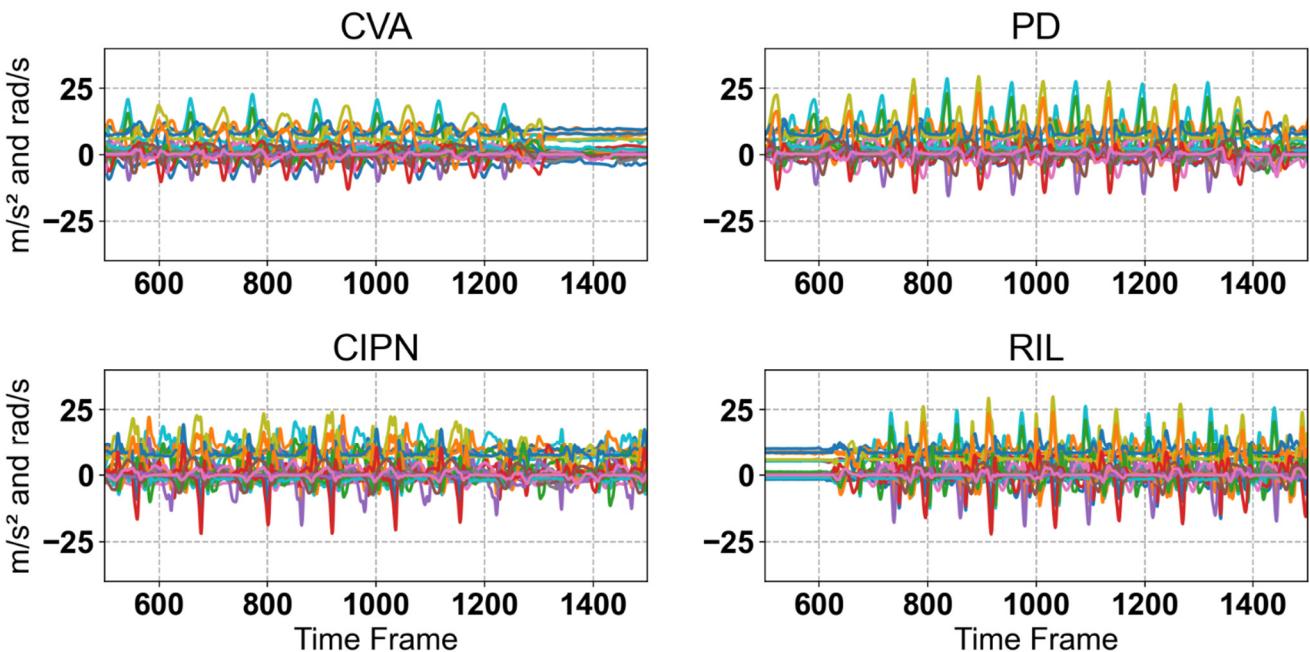
Four IMU sensors were positioned on the head (HE), lower back (LB), left foot (LF), and right foot (RF) (see Table 2) using adhesive straps as shown in Figure 1. Two hardware types were used: XSens and Technoconcept I4 Motion (see Figure 1), both functionally equivalent in capturing time-series acceleration and angular velocity signals. Sensors operated at a sampling frequency of 100 Hz, with full synchronization before each session [37]. The recorded data are presented for each class in this dataset in Figures 2 and 3.

**Table 2.** IMUs Sensor distribution and recorded data for reading for head, back, right foot, and left foot.

#	Feature Name	Description	Unit	#	Feature Name	Description	Unit
1	HE_Acc	Acceleration of the head sensor	m/s <sup>2</sup>	7	LF_Acc	Acceleration of the left foot sensor	m/s <sup>2</sup>
2	HE_FreeAcc	Acceleration minus gravity of the head sensor	m/s <sup>2</sup>	8	LF_FreeAcc	Acceleration minus gravity of the left foot sensor	m/s <sup>2</sup>
3	HE_Gyr	Angular velocity of the head sensor	rad/s	9	LF_Gyr	Angular velocity of the left foot sensor	rad/s
4	LB_Acc	Acceleration of the trunk (lower-back) sensor	m/s <sup>2</sup>	10	RF_Acc	Acceleration of the right foot sensor	m/s <sup>2</sup>
5	LB_FreeAcc	Acceleration minus gravity of the trunk (lower-back) sensor	m/s <sup>2</sup>	11	RF_FreeAcc	Acceleration minus gravity of the right foot sensor	m/s <sup>2</sup>
6	LB_Gyr	Angular velocity of the trunk (lower-back) sensor	rad/s	12	RF_Gyr	Angular velocity of the right foot sensor	rad/s



**Figure 2.** Raw IMUs Sensor singles for *Hs*, *HOA*, *KOA*, and *ACL*. Each plot shows multiple gait cycles for the corresponding class, illustrating the variability across samples. As can be seen, there are noticeable differences in signal patterns between the classes, reflecting variations in gait dynamics associated with different musculoskeletal conditions.



**Figure 3.** Raw IMU sensor singles for *CVA*, *PD*, *CIPN*, and *RIL*. Each plot shows multiple gait cycles for the corresponding class, illustrating the variability across samples. As can be seen, there are noticeable differences in signal patterns between the classes, reflecting variations in gait dynamics associated with different musculoskeletal conditions.

## 2.2. Data Processing Pipeline

The raw multivariate gait signals were processed through a structured pipeline to ensure consistency across subjects, robustness against inter-trial variability, and suitability for temporal modeling using the G-MASA-TCN architecture. Each trial in the dataset initially consisted of a variable-length sequence of T time steps sampled at 100 Hz, with

36 synchronized sensor channels capturing triaxial acceleration, free acceleration, and gyroscope readings from multiple body locations. Because gait speed differed significantly across clinical cohorts including HSs, HOA, KOA, ACL, PD, CVA, CIPN, and RIL, the resulting signal lengths also varied substantially. Empirically, the longest trial contained 18,639 frames, whereas the shortest contained 2490 frames. To manage this heterogeneity, two complementary preprocessing strategies were adopted for model training and evaluation.

### 2.2.1. Experiment 1: Zero-Padding Strategy (Full-Length Alignment)

In the first strategy, each trial was temporally aligned by zero-padding shorter sequences up to the maximum observed length  $T_{\max}$ . Formally, for a trial  $X \in \mathbb{R}^{T \times C}$ , the padded sequence  $\tilde{X} \in \mathbb{R}^{T_{\max} \times C}$  is defined as

$$\tilde{X}(t, c) = \begin{cases} X(t, c), & t \leq T \\ 0, & t > T \end{cases} \quad (1)$$

This approach preserved the full temporal structure of each trial and enabled the G-MASA-TCN to model complete gait cycles without discarding information. The padded sequences maintained the two-dimensional temporal–spatial organization of the data, consistent with the network’s convolutional design

### 2.2.2. Experiment 2: Fixed-Length Segmentation Strategy (500-Frame Windows)

In the second strategy, no temporal padding was applied. Instead, each variable-length trial was divided into consecutive non-overlapping sub-segments of frames with the fixed length  $L = 500$ :

$$X^{(k)} = X[kL : (k+1)L, :] \quad k = 0, 1, \dots, \lfloor \frac{T}{L} \rfloor - 1 \quad (2)$$

Because the sampling rate was 100 Hz, each 500-frame window corresponded to a 5 s segment of gait motion. Each segment retained its original spatial–temporal structure as a matrix of size  $500 \times 36$ . This method provided the G-MASA-TCN with uniformly sized windows while preserving the natural variability of gait patterns without introducing padding artifacts. Both preprocessing strategies were used independently for model training, allowing direct comparison of classification accuracy across full-length padded sequences and fixed-length unpadded segments.

### 2.2.3. Experiment 3: Subject-Level Identification via Gait Patterns

To ensure accurate prediction of the eight gait classes, encompassing neurological, orthopedic, and healthy cohorts, the model was evaluated on its ability to differentiate the gait patterns of 260 individual subjects as a biometric task. Gait analysis was approached as a subject identification problem, analogous to biometric fingerprint recognition. Each sample in the dataset is associated with a unique subject ID, and the model was trained to classify gait sequences according to their originating subject. Formally, let  $Y \in \mathbb{R}^{N \times 4}$  denote the label matrix, where the first column encodes the subject ID ( $0 \leq y \leq 259$ ), and the additional columns may encode metadata such as gait class. The dataset consists of 260 distinct subjects. Unlike traditional class-based partitioning, the G-MASA-TCN model is trained to distinguish among all 260 subjects simultaneously, learning unique gait signatures for each individual. Each subject is treated as a separate class:  $\mathcal{C} = \{0, 1, \dots, 259\}$ . The training and evaluation process uses the full set of subjects, enabling the model to associate specific spatiotemporal gait patterns with their corresponding subject IDs. By framing gait sequences as subject-specific biometric identifiers, the model can predict the subject from previously unseen gait sequences, akin to a forensic system linking latent

fingerprints to a known individual. This subject-level identification framework ensures that the model's performance reflects its ability to capture individual-specific gait characteristics, rather than only generic class differences. Such an approach provides a strong foundation for extending the model in healthcare applications, enabling reliable classification across multiple gait-related conditions and cohorts.

#### 2.2.4. Standardization of Training and Testing Data

Before feeding the signals into the G-MASA-TCN, each feature channel was standardized using statistics computed only from the training data. For training data  $X_{\text{train}}$ , the mean vector  $\mu$  and standard deviation vector  $\sigma$  across the temporal and sample dimensions were computed as

$$\mu = \frac{1}{N_{\text{train}}} \sum_{i=1}^{N_{\text{train}}} X^{(i)}, \quad \sigma = \sqrt{\frac{1}{N_{\text{train}}} \sum_{i=1}^{N_{\text{train}}} (X^{(i)} - \mu)^2} \quad (3)$$

To avoid division by zero, channels with  $\sigma_j = 0$  were replaced with 1. The standardized data was computed as

$$\hat{X}_{\text{train}} = \frac{X_{\text{train}} - \mu}{\sigma}, \quad \hat{X}_{\text{test}} = \frac{X_{\text{test}} - \mu}{\sigma} \quad (4)$$

Standardization ensures that each sensory channel contributes proportionally during convolution, stabilizes gradient flow, and improves training convergence by rescaling all features to comparable ranges.

#### 2.2.5. Final Data Representation for G-MASA-TCN

After segmentation, subject-level splitting, and standardization, each sample fed into the G-MASA-TCN has the shape

$$X \in \mathbb{R}^{T' \times 36} \quad (5)$$

where

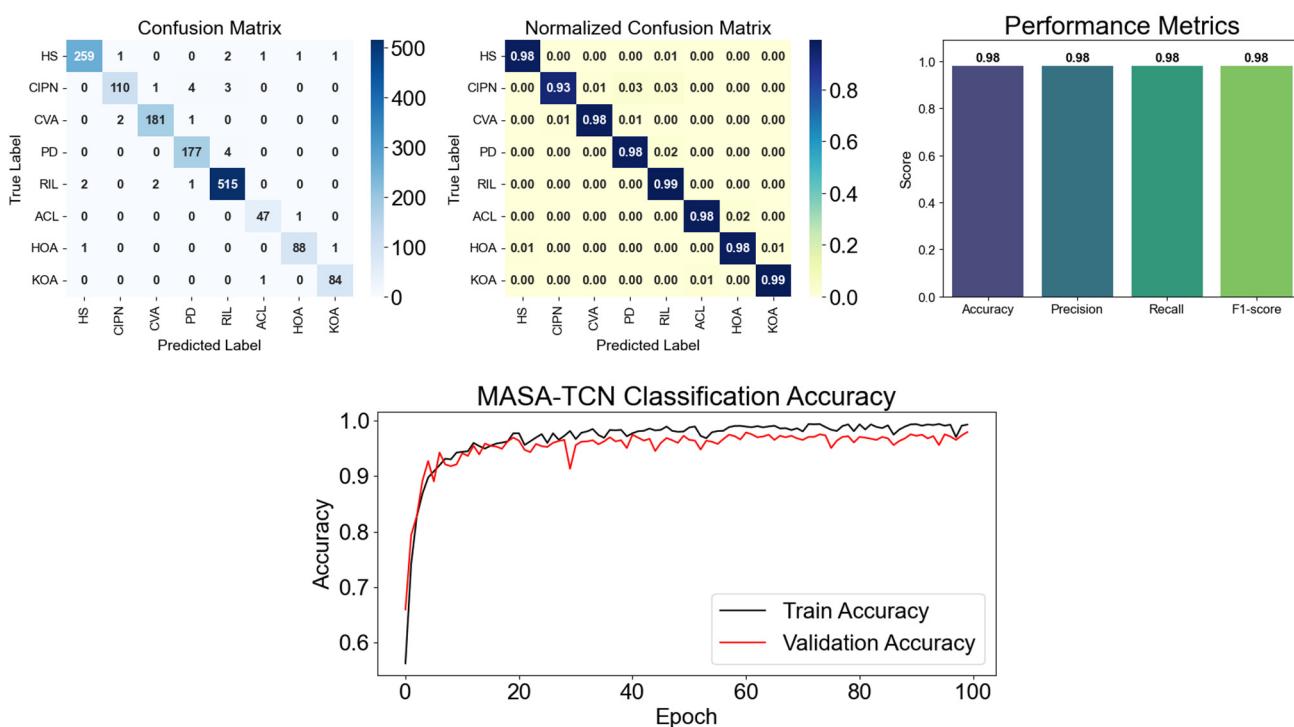
- $T' = T_{\max} = 18,639$  for zero-padded full-length trials, for a total of 1356 samples;
- $T' = 500$  for fixed-length segmented windows, for a total number of 7455 samples.

This consistent representation aligns with the G-MASA-TCN's multi-anchor spatial-temporal convolutional layers, which operate across the temporal dimension while jointly capturing inter-channel spatial dependencies.

#### 2.2.6. Real-World IMU Data for Compression

Although the G-MASA-TCN model is primarily developed for clinical gait analysis, specifically to differentiate between neurological, orthopedic, and healthy cohorts, it still requires validation under more generalized data and real-world scenarios. Therefore, without additional experimentation, we evaluate the applicability of the proposed model to human gait performance on irregular and uneven surfaces and during natural everyday walking in an urban environment as shown in Figure 4. For the real-world irregular surface dataset [44], the data were collected from 30 healthy participants (15 males and 15 females) who wore six IMU sensors placed on the forearms, thighs, ankles, and lower back while walking across nine different outdoor surface conditions (e.g., paved, grass, uneven stone, ramps, stairs) with multiple trials per participant covering a range of terrains and subtle gait pattern variations [44]. All subjects in the dataset were healthy; subject identity was used as the classification label in our experiments. We applied the same preprocessing procedure to these data: the raw IMU signals were segmented into fixed-length sequences of 500 steps, as described in Section 2.2.2, followed by assigning unique subject IDs as

class labels according to Section 2.2.3. Finally, the segmented sequences were standardized as outlined in Section 2.2.4, resulting in a structured dataset suitable for evaluating the generalizability of the G-MASA-TCN model on non-clinical gait data.



**Figure 4.** G-MASA-TCN classification results for Experiment 2 using the fixed-length segmentation strategy with 500-frame windows. The figure shows the confusion matrix for all classes, highlighting the normalized accuracy of predicted samples compared to true positives. Performance metrics including accuracy, precision, recall, and F1-score are also presented. Additionally, the training and validation learning curves illustrate the model’s performance across epochs, demonstrating proper alignment and indicating the absence of overfitting.

### 2.3. G-MASA-TCN: A Gait-Multi-Scale Space-Aware Temporal Convolutional Network for Gait Pathology Classification

This study classifies eight gait conditions—HSs, HOA, KOA, ACL, PD, CVA, CIPN, and RIL—using multivariate time series collected from four IMUs positioned on the head (HE), lower back (LB), left foot (LF), and right foot (RF). Each IMU recorded tri-axial acceleration, gravity-compensated (free) acceleration, and angular velocity, resulting in  $C = 36$  input channels per time step. The input tensor is defined as

$$X \in \mathbb{R}^{B \times T \times C}, T = 500, C = 36 \quad (6)$$

where  $B$  is the batch size and  $T$  corresponds to 5 s of data at 100 Hz sampling rate. To address this challenging multi-class classification task, an adapted version of the Multi-Anchor Space-Aware Temporal Convolutional Network (MASA-TCN), originally proposed in [26] for EEG-based emotion recognition, was employed. The model is specifically tailored to capture multi-scale temporal dynamics, inter-limb coordination, and causal real-time inference all critical for clinical gait assessment.

#### 2.3.1. SAT Layer: Space-Aware Temporal Feature Extraction

The foundational building block of G-MASA-TCN is the Space-Aware Temporal (SAT) layer, designed to extract local temporal patterns while preserving spatial relationships

across sensor channels, as shown in Figure 1. Unlike traditional 1D convolutions applied independently per channel, the SAT layer treats the input as a 2D feature map, enabling joint modeling of sensor location and signal dynamics. Given an input  $X \in \mathbb{R}^{B \times T \times C}$ , the SAT layer proceeds as follows:

1. Reshaping for 2D Convolution:

$$X' = X.\text{permute}(0, 2, 1).\text{unsqueeze}(1) \in \mathbb{R}^{B \times 1 \times C \times T} \quad (7)$$

2. Context Convolution (Multi-Channel Temporal Modeling): A 2D convolution with the kernel size  $(C, k)$  and dilation  $(1, d_t = 2)$  is applied:

$$H_c = \text{Conv}_{2D}\left(X'; W_c \in \mathbb{R}^{s \times 1 \times C \times k}, \text{dilation} = (1, d_t), \text{padding} = (0, (k - 1)d_t)\right) \quad (8)$$

3. This operation extracts temporal patterns of length  $k$  across all 36 channels simultaneously, with causal padding ensuring no future information leakage. The output  $H_c \in \mathbb{R}^{B \times s \times 1 \times (T + (k - 1)d_t)}$  is cropped to the original length:

$$H_c \leftarrow H_c[:, :, :, :T] \quad (9)$$

4. Spatial Mixing via  $1 \times 1$  Convolution: A pointwise convolution mixes information across the  $s$  feature maps:

$$H_s = \text{Conv}_{2D}\left(H_c; W_s \in \mathbb{R}^{s \times s \times 1 \times 1}\right) \quad (10)$$

5. Normalization and Regularization:

$$H = \text{Dropout}(\text{ReLU}(\text{BN}(H_s))) \in \mathbb{R}^{B \times s \times 1 \times T} \quad (11)$$

Traditional 1D convolutions treat each IMU channel independently, ignoring anatomical and functional relationships (e.g., head–trunk coupling in PD, asymmetric foot loading in KOA). The SAT layer’s 2D kernel spanning all channels enables the network to learn cross-sensor dependencies, for instance, compensatory head motion in response to lower limb instability. The use of causal dilated convolution ensures the model can be deployed in real-time clinical settings, where future frames are unavailable.

### 2.3.2. MAAF Block: Multi-Scale Adaptive Feature Fusion

Gait cycles exhibit hierarchical temporal structures: micro-events (heel-strike, toe-off), step-level patterns ( $\sim 1$  s), and stride-level symmetry ( $\sim 2$  s). To capture multi-scale dynamics, the Multi-Anchor Adaptive Fusion (MAAF) block, illustrated in Figure 1, is employed, which executes three SAT layers in parallel with kernel sizes  $\mathcal{K} = \{3, 5, 15\}$ :

$$H^{(i)} = \text{SAT}_k(X), k \in \{3, 5, 15\}, i = 1, 2, 3 \quad (12)$$

The outputs are concatenated:

$$H_{\text{cat}} = \text{Concat}\left(H^{(1)}, H^{(2)}, H^{(3)}\right) \in \mathbb{R}^{B \times 3s \times 1 \times T} \quad (13)$$

A  $1 \times 1$  fusion convolution learns to adaptively weight and combines multi-scale representations:

$$H_{\text{fuse}} = \text{Conv}_{2D}\left(H_{\text{cat}}; W_f \in \mathbb{R}^{s \times 3s \times 1 \times 1}\right) \quad (14)$$

This is followed by batch normalization and ReLU:

$$Z = \text{ReLU}(\text{BN}(H_{\text{fuse}})) \in \mathbb{R}^{B \times s \times 1 \times T} \quad (15)$$

#### Rationale for Multi-Scale Design

- Kernel = 3 (~30 ms): Captures rapid acceleration spikes during foot contact.
- Kernel = 5 (~50 ms): Models mid-stance and push-off phases.
- Kernel = 15 (~150 ms): Encodes step-level transitions and inter-limb coordination.

This inception-style parallelism allows the model to represent gait at multiple temporal resolutions without increasing depth, improving robustness to stride variability seen in pathological gait (e.g., shuffling in PD, limping in HOA).

#### 2.3.3. Residual Temporal Convolutional Blocks

The MAAF (See Figure 1) output is squeezed to  $Z \in \mathbb{R}^{B \times s \times T}$  and processed by three residual dilated 1D convolutional blocks with increasing dilations:  $d_l \in \{1, 2, 4\}$ . For the  $l$ -th block,

$$O_l = \text{Dropout}(PReLU(\text{BN}(\text{WeightNorm}(\text{Conv}_{1D}(Z_{l-1}; k=3, d=d_l)))))) \quad (16)$$

$$R_l = \begin{cases} \text{Conv}_{1D}(Z_{l-1}; k=1) & \text{if } c_{in} \neq c_{out} \\ Z_{l-1} & \text{otherwise} \end{cases} \quad (17)$$

$$Z_l = O_l + R_l \quad (18)$$

In this formulation,

- $O_l$  represents the nonlinear dilated convolutional transformation;
- $R_l$  denotes the residual connection, which either performs a  $1 \times 1$  projection (when channel dimensions differ) or passes the input directly;
- The block output  $Z_l$  is obtained by summing the transformed output and the residual pathway.

With  $c_{out} = 128$  for all blocks, the effective receptive field after three layers is

$$\text{RF} = 1 + 2(3 - 1)(2^0 + 2^1 + 2^2) = 1 + 4(1 + 2 + 4) = 29 \text{ frames} \approx 290 \text{ ms (at 100 Hz)} \quad (19)$$

However, when combined with SAT dilation ( $d_t = 2$ ) and multi-scale kernels, the total receptive field exceeds 1.2 s, sufficient to model full gait cycles. Why is the residual TCN used? Recurrent networks (LSTM, GRU) suffer from vanishing gradients and slow inference. Dilated TCNs offer exponential receptive field growth, parallel training, and constant-time inference ideal for embedded clinical devices. Weight normalization and residual connections stabilize training of deep temporal models.

#### 2.3.4. Classification Head and Training

After the final residual TCN block, the feature map  $Z_3 \in \mathbb{R}^{B \times 128 \times T}$  contains rich temporal representations that encode multi-scale gait dynamics across all sensor channels. To produce a single prediction per input sequence, global average pooling is applied along the temporal dimension, summarizing the entire 5 s gait segment into a compact fixed-length vector:

$$g = \frac{1}{T} \sum_{t=1}^T Z_3[:, :, t] \in \mathbb{R}^{B \times 128} \quad (20)$$

This operation not only reduces dimensionality but also enforces translation invariance across the gait cycle, making the model robust to variations in stride timing a common

challenge in pathological gait where cadence and phase durations differ significantly between conditions. The pooled representation  $\mathbf{g}$  is then passed through a fully connected layer with the weight matrix  $\mathbf{W}_{\text{fc}} \in \mathbb{R}^{K \times 128}$  and bias  $\mathbf{b}_{\text{fc}}$ , producing class logits for the eight target pathologies:

$$\mathbf{y} = \mathbf{W}_{\text{fc}}\mathbf{g} + \mathbf{b}_{\text{fc}} \quad (21)$$

During training, the model is optimized end-to-end using the cross-entropy loss, which measures the discrepancy between predicted probabilities and ground-truth labels:

$$\mathcal{L} = -\frac{1}{B} \sum_{i=1}^B \sum_{k=1}^K y_{i,k}^{\text{true}} \log(\text{softmax}(\mathbf{y}_i)_k) \quad (22)$$

The Adam optimizer is employed with an initial learning rate of  $10^{-3}$ , enabling stable convergence even with the high-dimensional 36-channel input. Dropout with a probability of 0.1 is consistently applied throughout the network to prevent overfitting, particularly important given the relatively small sample sizes typical in clinical gait studies. This training paradigm ensures that the model learns discriminative features directly from raw IMU signals without requiring hand-crafted gait parameters such as stride length, stance time, or symmetry indices.

#### 2.4. Why G-MASA-TCN for Gait Pathology Classification?

Pathological gait is characterized by complex, multi-scale temporal alterations. For instance, PD often manifests as reduced step length and increased cadence variability at the sub-step level, while hemiparetic gait following CVA introduces pronounced asymmetry over full stride cycles. The MAAF block, with its parallel SAT layers using kernel sizes of 3, 5, and 15, captures hierarchical dynamics: short kernels detect rapid acceleration transients during foot contact, medium kernels model mid-stance and propulsion phases, and longer kernels encode inter-limb coordination and stride-level symmetry. The effective receptive field of a single SAT layer with kernel size  $k$  and dilation  $d_t = 2$  is

$$RF_{\text{SAT}} = 1 + (k - 1)d_t \quad (23)$$

yielding 5, 9, and 29 frames, respectively, for  $k \in \{3, 5, 15\}$ . After adaptive fusion and the subsequent residual TCN stack with the dilations  $\{1, 2, 4\}$  and kernel size 3, the total receptive field expands exponentially:

$$\begin{aligned} RF_{\text{total}} &= RF_{\text{SAT}} + \sum_{l=1}^3 (3 - 1) \cdot d_l \\ &= RF_{\text{SAT}} + 2(1 + 2 + 4) \\ &= RF_{\text{SAT}} + 14 \end{aligned} \quad (24)$$

This covers  $> 1.2$  s at 100 Hz sampling sufficient for multiple gait cycles without explicit event detection pipelines [26]. Second, the spatial arrangement of IMU sensors on the head, lower back, and bilateral feet forms a natural anatomical hierarchy that influences gait stability and compensation strategies. In healthy walking, trunk and head motion are tightly coupled with lower limb kinematics; in contrast, patients with HOA may exhibit increased pelvic tilt to reduce joint loading, while those CIPN display reduced distal coordination. The SAT layer's 2D convolution with kernel height equal to all 36 channels jointly models cross-sensor dependencies, outperforming channel-independent 1D convolutions that miss inter-segment coordination [38].

Third, clinical deployment demands real-time, causal inference. Wearable gait monitoring systems must provide immediate feedback during ambulatory assessments or fall risk evaluation. The entire G-MASA-TCN architecture—causal padding in SAT layers,

dilated convolutions with increasing rates, and absence of recurrent units—ensures that predictions at the time  $t$  depend only on inputs  $\leq t$ :

$$y_t = f(x_{\leq t}; \theta) \quad (25)$$

Equation (25) enables low-latency inference on edge devices [39]. Furthermore, the model strikes an effective balance between expressiveness and efficiency. With approximately 0.92 million parameters, it is substantially lighter than hybrid ConvLSTM architectures, which often exceed 1.2 million parameters due to recurrent gating mechanisms. The residual TCN blocks, enhanced with weight normalization and PReLU activations, facilitate stable training of deep temporal hierarchies without gradient vanishing, a common issue in long-sequence gait data. Finally, the proven generalization of MASA-TCN from EEG emotion recognition to IMU-based movement analysis underscores its robustness to noisy, high-dimensional, multivariate time series, a hallmark of real-world sensor data affected by motion artifacts, sensor drift, and environmental interference [26].

### 2.5. Baseline Model Architectures

To establish a fair performance comparison, the proposed G-MASA-TCN was evaluated against three widely used sequence-modeling architectures: a Temporal Convolutional Network (TCN), Gated Recurrent Unit (GRU), and Transformer encoder. The first baseline, a TCN model, follows the design introduced in [45], which employs stacked dilated causal convolutions to efficiently expand the receptive field and capture long-range temporal dependencies without recurrence. While TCNs provide strong temporal modeling capabilities, they do not capture spatial correlations across IMU channels. The second baseline is a deep bidirectional GRU network, inspired by [46], which models temporal sequences through gated recurrence and hidden-state transitions. GRUs are effective for sequential data but struggle with long sequences and lack the ability to model multi-scale gait variations, especially when compared to convolutional approaches. The third baseline uses a Transformer encoder architecture introduced in [47], leveraging multi-head self-attention to model global dependencies across all time steps. Although Transformers excel at long-range modeling, they require larger datasets to avoid overfitting and do not inherently encode temporal priors or spatial sensor interactions. In contrast to these baselines, the proposed G-MASA-TCN integrates multi-anchor spatial-temporal filtering, graph-aware feature learning, and residual dilated convolutions, enabling it to capture fine-to-coarse gait dynamics and inter-sensor dependencies beyond those modeled by the TCN, GRU, or Transformer baselines.

### 2.6. Explainable Artificial Intelligence (XAI) for G-MASA-TCN

Although G-MASA-TCN provides state-of-the-art performance in classifying gait disorders, interpretability remains essential, especially in clinical and biomechanical applications where understanding model makes a particular decision is as critical as the prediction itself. To enhance transparency, the model integrates an explainable AI (XAI) framework centered on Integrated Gradients (IG), a mathematically principled attribution method introduced in [48]. IG was selected over alternative interpretation techniques such as Saliency Maps, SmoothGrad [49], Layer-wise Relevance Propagation (LRP), and SHAP because it satisfies key axioms required for reliable medical decision support: sensitivity, implementation invariance, linearity, and completeness. Most gradient-based methods fail to satisfy these axioms [50], while perturbation-based methods such as SHAP or LIME are computationally expensive for high-dimensional time-series models and often produce unstable results. IG, in contrast, provides smooth, theoretically grounded attributions that

align uniquely well with G-MASA-TCN's multi-anchor temporal representation, making it the most suitable XAI choice for clinical gait modeling.

Integrated Gradients measures how much each input feature contributes to the model's output by integrating the gradients along a straight-line path between an input,  $x$ , and a baseline,  $x'$  (typically a zero or mean sensor sequence). For an input feature  $x_i$ , the IG attribution is defined as

$$\text{IG}_i(x) = (x_i - x'_i) \int_0^1 \frac{\partial F(x' + \alpha(x - x'))}{\partial x_i} d\alpha \quad (26)$$

where  $F(\cdot)$  is the G-MASA-TCN output logit for the predicted class. In practical implementations, the integral is approximated using  $m$  interpolation steps:

$$\text{IG}_i(x) \approx (x_i - x'_i) \frac{1}{m} \sum_{k=1}^m \frac{\partial F(x' + \frac{k}{m}(x - x'))}{\partial x_i} \quad (27)$$

For G-MASA-TCN, each input is a multichannel IMU time series:

$$x \in \mathbb{R}^{T \times C} \quad (28)$$

where  $T$  denotes the temporal window and  $C$  indicates the number of sensor channels. IG backpropagates its gradients through every component of the architecture including the Multi-Anchor Attention Fusion (MAAF), the residual dilated TCN blocks, and the temporal global pooling stage, allowing the attribution map to reflect the true contribution of early spatial-temporal convolutional filters as well as deeper receptive-field expansions. This enables G-MASA-TCN to reveal class-specific reliance on short-range vs. long-range temporal filters, biomechanically meaningful gait events such as heel-strike or toe-off, and fine-grained cross-sensor dependencies. The resulting IG attribution tensor

$$A \in \mathbb{R}^{T \times C} \quad (29)$$

is used to construct two complementary interpretability visualizations. Feature importance is computed by averaging the absolute attributions over time:

$$\text{FI}(c) = \frac{1}{T} \sum_{t=1}^T |A_{t,c}| \quad (30)$$

This allows identification of the dominant sensor channels for each gait pathology. Temporal importance is obtained by averaging the absolute attributions across selected sensor channels:

$$\text{TI}(t) = \frac{1}{C'} \sum_{c \in \mathcal{C}'} |A_{t,c}| \quad (31)$$

This highlights which gait-cycle segments (e.g., stance, swing, transition) exert the strongest influence on the model's classification decision. Together, these visualizations form clinically meaningful sensor-time heatmaps that reveal how MASA-TCN grounds its predictions in biomechanically relevant features rather than noise or spurious correlations. Integrated Gradients therefore ensures both *local* (sample-level) and *global* (class-level) interpretability, enabling clinicians and researchers to validate, audit, and trust the G-MASA-TCN model. Its mathematical guarantees, stability, and compatibility with high-resolution temporal-convolutional models make it superior to alternative XAI methods for gait-pathology analysis.

### 3. Results

To ensure that the evaluation of all models was statistically reliable, reproducible, and free from artifacts caused by dataset shuffling, each experiment was repeated five times using different random seeds and a subject-level 5-fold cross-validation, with 70% of the data used for training, 10% for validation, and 20% for testing in each fold. This combined procedure allows assessment of model variance, robustness, and sensitivity to initialization, while maintaining rigorous, non-overlapping test partitions for unbiased evaluation. Model training was conducted on Google Colab cloud infrastructure using an NVIDIA GPU, with an average training time of approximately 5 min, while testing and inference evaluations were performed on a local workstation equipped with an Intel Core i7 (10th generation) CPU to assess deployment feasibility on standard hardware. For each experiment, two complementary sets of results are reported: (1) the best-performing run in terms of validation accuracy, including the confusion matrix, precision, recall, F1-scores, overall accuracy, and full training/validation accuracy curves; (2) aggregate statistics across all five runs, represented by the mean confusion matrix and corresponding standard deviations. To further visualize the model's ability to cluster gait dynamics across the eight classes (HSs, HOA, KOA, ACL, PD, CVA, CIPN, and RIL), a t-SNE projection of the learned embeddings is provided for each experiment.

Because diagnostic applications require not only strong predictive performance but also transparent decision-making, these quantitative results are complemented by a comprehensive explainable AI (XAI) analysis. This interpretability framework reveals the internal spatiotemporal representations learned by each model and identifies the most influential sensor channels and gait-cycle segments contributing to classification. Together, the statistical evaluation and XAI visualizations provide a complete and trustworthy assessment of both model accuracy and interpretability in biomechanical and clinical contexts. In the following sections, the results of the training and evaluation strategies introduced, followed by an evaluation using the Integrated Gradients (IG) XAI method.

#### 3.1. Experiment 1: Zero-Padding Strategy

This experiment investigated the effect of zero padding on modeling variability in gait duration among participants. Since real-world gait recordings vary in length due to differences in stride cycles, cadence, mobility level, and pathological conditions, all sequences were standardized to the global maximum length  $T_{\max}$  by padding shorter sequences with zeros. This ensured that architectures such as the TCN, Transformer, and the GRU could process all trials in a unified tensor format while preserving complete temporal dynamics. However, the impact of padding on learning stability varied substantially across models. Across the five independent seeds (see Table 3), the baseline TCN reached  $88\% \pm 2\%$ , confirming that conventional dilated convolutions can capture long-range temporal dependencies but remain sensitive to variation in gait rhythm. The Transformer, however, performed considerably worse at  $73\% \pm 5.6\%$ , likely due to its positional-encoding dependence and limited ability to regulate noisy or padded temporal segments. The GRU exhibited the weakest performance at  $52\% \pm 6.2\%$ , which suggests that recurrent architectures face difficulty maintaining stable hidden-state transitions when padding disrupts stride periodicity.

In sharp contrast, the proposed G-MASA-TCN achieved  $95.2\% \pm 0.8\%$ , not only outperforming all baseline models but also exhibiting extremely low variance across all five seeds. This demonstrates the architecture's stability under variable-length gait conditions and its ability to extract discriminative temporal cues from complete, natural gait cycles. The multi-anchor temporal filters of MASA-TCN allow the model to analyze gait at multiple temporal scales simultaneously, enabling it to capture subtle abnormalities such

as reduced stride length in HOA and KOA, elevated gait asymmetry in ACL and CVA, and tremor-induced high-frequency fluctuations in PD. Moreover, the spatial-attention fusion mechanism highlights the most informative IMU channels and axes when distinguishing neuropathic gait alterations in CIPN and central motor disruption patterns in RIL. The results of this experiment include the confusion matrix for the best seed run, the mean confusion matrix across all runs, accuracy/precision/recall/F1 metrics, and the complete training and validation accuracy curves are summarized in Figure 4. The data show that G-MASA-TCN manages temporal irregularities caused by pathology-specific gait disruptions, supporting its advantage over architectures with limited spatial adaptivity or weaker temporal stability.

**Table 3.** Mean and standard deviation with different random seed for each experiment for the classification of gait with 70% for training, 10% for validation, and 20% for testing.

Model	Experiment 1	Experiment 2	Experiment 3	Experiment 4	Experiment 5
TCN	$88\% \pm 2\%$	$89\% \pm 2.7\%$	$79\% \pm 3.5\%$	$62\% \pm 6.7\%$	$83\% \pm 4.2\%$
Transformer	$73\% \pm 5.6\%$	$87\% \pm 4.4\%$	$62\% \pm 4.1\%$	$77\% \pm 3.6\%$	$71\% \pm 3.7\%$
GRU	$52\% \pm 6.2\%$	$82\% \pm 3.2\%$	$76\% \pm 3.3\%$	$63\% \pm 5.4\%$	$66\% \pm 2.9\%$
G-MASA-TCN	<b><math>95.2\% \pm 0.8\%</math></b>	<b><math>96.8\% \pm 1.2\%</math></b>	<b><math>96.4\% \pm 2\%</math></b>	<b><math>94.3\% \pm 1.8\%</math></b>	<b><math>91\% \pm 2\%</math></b>

Bold text highlights the best-performing results.

### 3.2. Experiment 2: Fixed-Length Segmentation (500-Frame Windows)

This experiment evaluated the impact of fixed-length segmentation, which eliminated zero padding altogether. Gait signals were windowed into non-overlapping 500-frame segments, corresponding to approximately 5 s of motion. This segmentation reduced temporal variability between samples, thereby stabilizing the learning process by presenting the models with consistent temporal structures. Under this standardized alignment, the baseline models improved significantly across all five seeds, as shown in Table 3. The traditional TCN reached  $89\% \pm 2.7\%$ , the Transformer achieved  $87\% \pm 4.4\%$ , and the GRU improved markedly to  $82\% \pm 3.2\%$ . These gains reflect the benefit of ensuring that each input segment contained a comparable number of gait cycles, a crucial factor given the distinct stride dynamics across HSs, HOA, KOA, and PD and the irregular timing patterns in ACL, CVA, CIPN, and RIL.

The G-MASA-TCN once again delivered the strongest performance, achieving  $96.8\% \pm 1.2\%$  across the five seeds. This marks an improvement over Experiment 1 and indicates that MASA-TCN benefited from clean, uniformly segmented gait windows in which stride-level micro-patterns were preserved without padding artifacts. The multi-scale anchoring mechanism was especially advantageous here: it captured the interaction between low-frequency gait-cycle components (e.g., stance–swing transitions) and high-frequency pathological markers (e.g., PD tremor- or CIPN-related proprioceptive instability). The spatial-attention module also became more effective in this setting, consistently highlighting hip-mounted and shank-mounted IMU axes that strongly differentiate HOA, KOA, and ACL conditions. The results of this experiment, including detailed confusion matrices for the best run, the aggregated mean matrices, and the learning curves, are displayed in Figure 4. The improvement across all architectures confirms that fixed-length segmentation fostered more stable temporal learning, but G-MASA-TCN maintained a decisive performance advantage, approaching near-perfect discrimination across all eight gait classes.

The models were initially trained and validated using 70% of the available dataset, with 60% allocated for training and 10% reserved for validation, while the remaining 30% was held out for testing. This initial split enabled optimization of model parameters

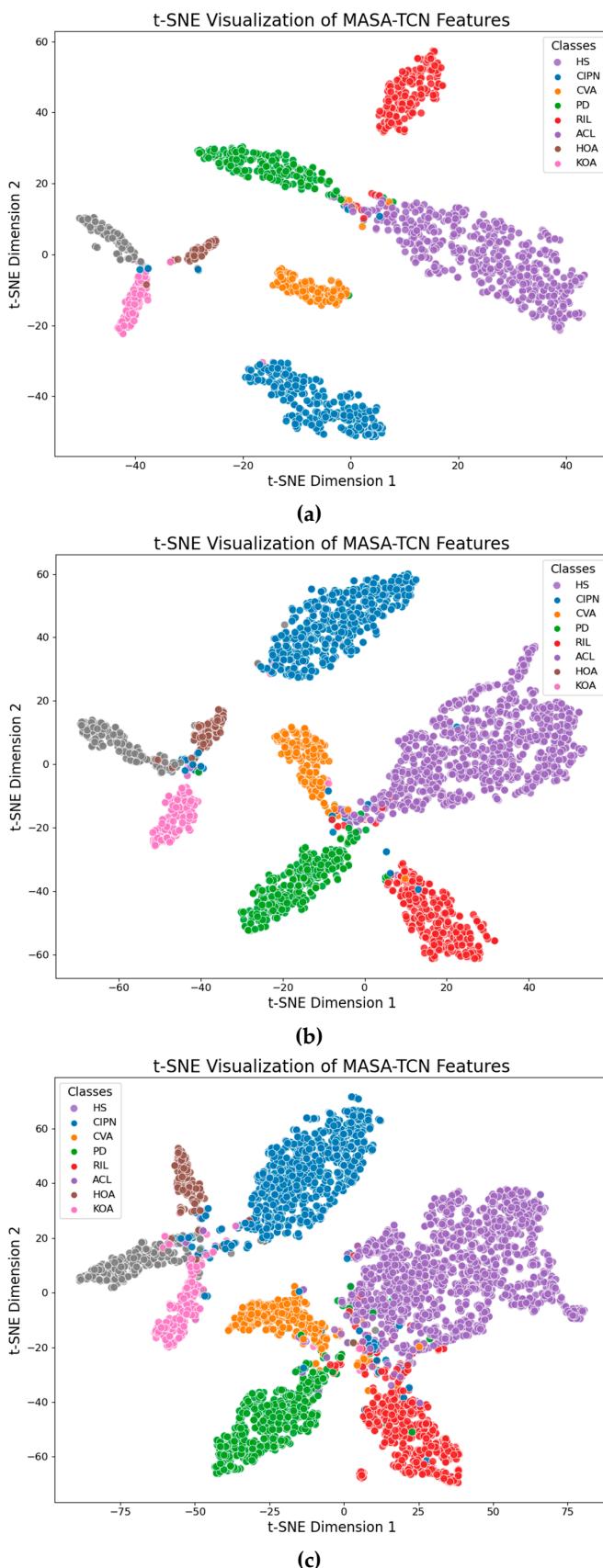
and hyperparameters while providing an unbiased evaluation of performance on unseen data. Given that G-MASA-TCN achieved the highest predictive performance among all architectures, further experiments were conducted to evaluate its robustness across different training-to-testing ratios. Specifically, the model was retrained on 60% of the data and tested on the remaining 40%, then trained on 40% and tested on 60%, and finally trained on 30% with testing on 70% of the dataset. Across all these splits, the observed changes in model accuracy were minimal and statistically insignificant, which can be attributed to the large size of the dataset comprising 7455 samples and the use of fixed-length segmentation of 500 samples per segment, which ensured consistent input representations for training and testing. The results for each split are presented not only in terms of standard performance metrics but also visually through t-SNE clustering of the predicted eight gait classes (HSs, HOA, KOA, ACL, CVA, PD, CIPN, RIL), demonstrating clear separation and alignment with clinical groupings as shown in Figure 5. These analyses collectively confirm the stability, robustness, and generalization capability of G-MASA-TCN across varying training data sizes.

### 3.3. Experiment 3: Subject-Level Identification via Gait Signatures

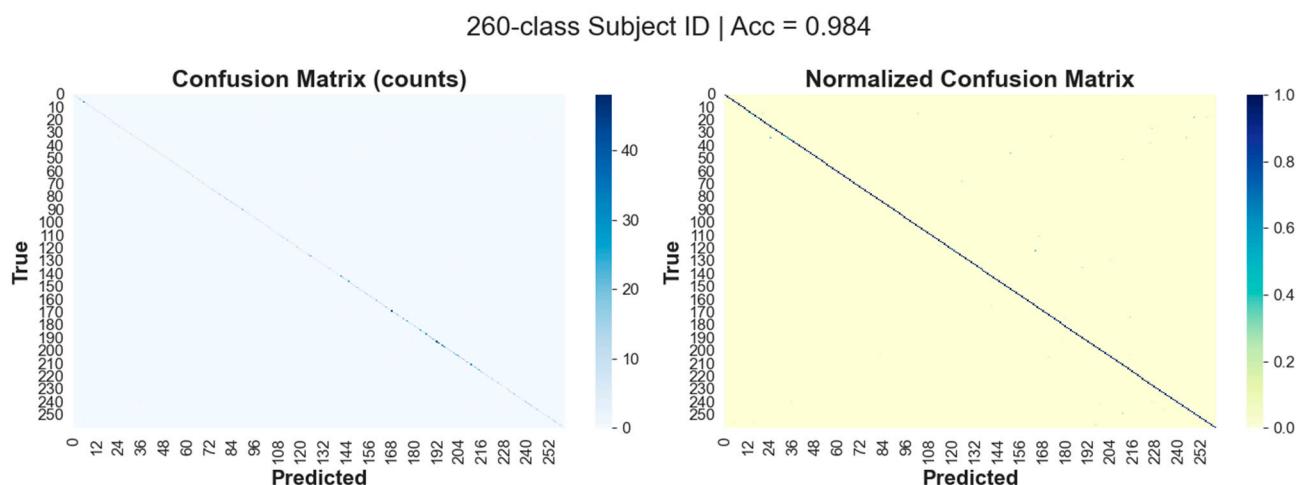
This experiment explored a fundamentally different problem setup: subject identification from gait signals. Unlike pathology classification, the models were tasked with identifying 260 individual subjects, requiring each architecture to learn unique, person-specific gait signatures. This scenario emphasized subtle spatiotemporal features that remained consistent across trials, despite variations in pathology or sensor noise. Using 500-frame segments, as shown in Table 3, the GRU achieved  $76\% \pm 3.3\%$ , outperforming the Transformer ( $62\% \pm 4.1\%$ ) but slightly below the TCN ( $79\% \pm 3.5\%$ ). The GRU's performance reflects its suitability for capturing individualized rhythmic patterns. However, the G-MASA-TCN again achieved the highest performance, reaching  $96.4\% \pm 2.1\%$  across the five seeds.

To further evaluate generalization and mimic biometric identification scenarios, an additional test was conducted in which 10 subjects were completely removed from the training set and treated as unknown “imposters” during testing. G-MASA-TCN returned zero predictions for these unseen subjects, indicating that it did not falsely classify unknown individuals as any of the enrolled subjects. This result mirrors real-world biometric systems, where unknown individuals are correctly recognized as imposters, demonstrating that G-MASA-TCN reliably captures subject-specific gait signatures without overfitting to the training population.

This result highlights architecture's exceptional capability to learn high-resolution gait signatures, capturing discriminative temporal cues that differ across individuals even within the same pathology class. Such signatures include personalized cadence, foot-ground contact profiles, habitual asymmetries, and joint-coordination rhythms. The multi-anchor mechanism, by analyzing the gait signals at varying temporal scales, enables the model to discover nuanced personal stride features that conventional TCN, Transformer, and GRU architectures fail to isolate consistently. The spatial attention module complements this by selecting the most identity-relevant IMU axes, often focusing on lateral acceleration and angular velocity channels associated with personalized stride lateralization. The results for this experiment, including the best-run confusion matrix, mean confusion matrix, and complete evaluation metrics, are presented in Figure 6. The t-SNE plots reveal well-separated clusters, demonstrating that G-MASA-TCN learns highly discriminative representations of individual gait signatures, confirming its promise for gait biometrics.



**Figure 5.** G-MASA-TCN t-SNE plots for different data split. (a) Training used 70% and remaining 30% was held out for testing. (b) Training used 60% and remaining 40% was held out for testing. (c) Training used 30% and remaining 70% was held out for testing.



**Figure 6.** G-MASA-TCN classification results for Experiment 3 using the fixed-length segmentation strategy (500-frame windows) for ID classification, showing the confusion matrix.

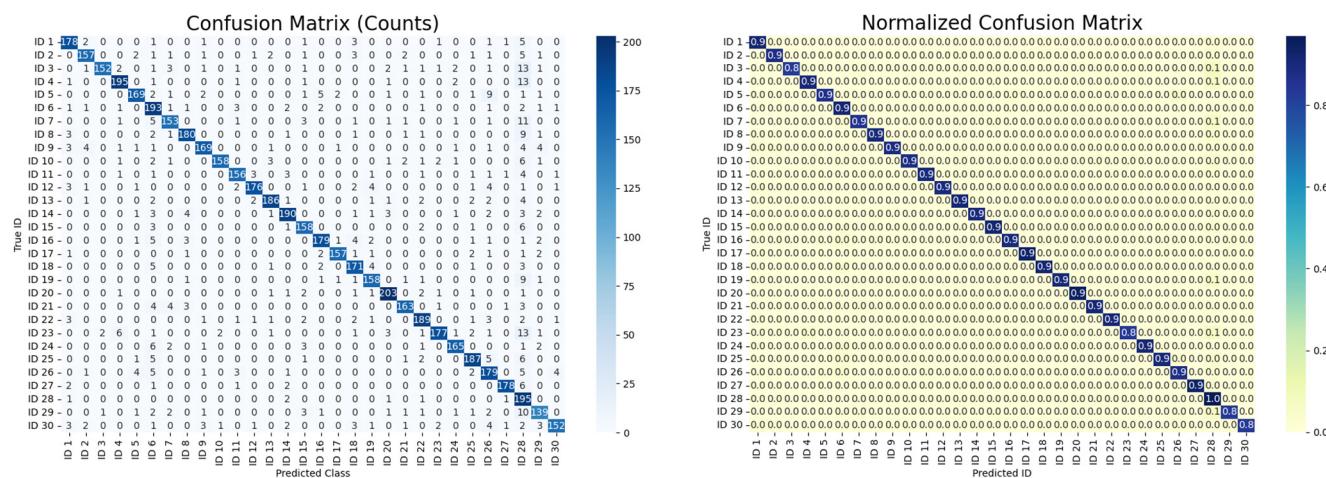
### 3.4. Experiment 4: 5-Fold Cross Validation

In this experiment, all models were trained and evaluated using 5-fold cross-validation. In each fold, two subjects from each class were held out for testing, while the remaining subjects were used for training. The test subjects were randomly selected under the constraint that no subject was included in the testing set more than once across all five folds.

The choice of using two test subjects per class was motivated by the limited number of available subjects in the HO, KOA, and ACL classes, each containing fewer than 20 subjects. Selecting two subjects per fold provided a more reliable and balanced evaluation under these constraints. The testing results, reported in Table 3, indicate that the proposed G-MASA-TCN with fixed-length inputs achieved the best overall performance. Specifically, it attained a peak accuracy of 96.1%, with a mean accuracy of  $94.3\% \pm 1.8\%$  across the 5-fold cross-validation. This outcome further confirms the consistency of the results obtained in Experiments 1 and 2.

### 3.5. Experiment 5: Subject-Level Identification via Real-World Gait Signature Data

To evaluate the generalization capability of the proposed G-MASA-TCN model, experiments were conducted on two real-world IMU-based gait datasets (see Section 2.2.5): irregular and uneven surface walking 30 subjects and natural everyday walking in an urban environment. For this dataset, the models were trained twice under two evaluation protocols. In the first experiment, the model was trained to classify all subjects, where each subject was treated as an independent class. This setup evaluated the model's ability to capture subject-specific gait signatures under non-clinical, real-world walking conditions. The obtained classification performance demonstrates that the proposed architecture can effectively learn discriminative gait representations from IMU data recorded on uneven and irregular surfaces as well as during natural urban walking. The results, as reported in the confusion matrix in Figure 7, show that the best-performing model achieved 93% accuracy in predicting the classes of each subject. Across the five random-seed repetitions, the model obtained a mean accuracy of  $91\% \pm 2\%$ , demonstrating consistent performance and reliability across different data splits.



**Figure 7.** G-MASA-TCN classification results for Experiment 5 using the fixed-length segmentation strategy (500-frame windows) for ID classification of *real-world gait signature data*, showing the confusion matrix.

In the second experiment, an imposter detection protocol was employed to further assess G-MASA-TCN model robustness. Five subjects were randomly selected and completely excluded from the training set. During testing, the model was required to identify these unseen subjects as imposters, i.e., subjects not belonging to any of the known classes learned during training and it return zero F1 score. This experiment evaluated the model's ability to generalize beyond trained identities and to detect unfamiliar gait patterns encountered in real-world scenarios.

The primary objective of this validation was twofold: to assess the ability of G-MASA-TCN to capture human gait characteristics outside the clinical environment and to evaluate its robustness in modeling gait patterns during walking on uneven, irregular, and uncontrolled surfaces. The quantitative results for the dataset experimental protocol are summarized in Table 3. Overall, the results confirm that the proposed G-MASA-TCN model maintains strong performance in real-world settings, highlighting its potential applicability for robust gait modeling beyond controlled clinical conditions.

### 3.6. XAI Evaluation

The explainability analysis based on Integrated Gradients (IG) offers a comprehensive view of how the G-MASA-TCN model interprets sensor data to classify gait patterns across eight distinct conditions. By incorporating multiple visualization techniques including sample-level IG heatmaps, global feature importance bar plots, and temporal relevance curves, the analysis reveals the spatial and temporal characteristics that guide the model's decision-making. These interpretations provide insight into how specific sensor modalities, representing different anatomical locations and biomechanical movements, contribute to distinguishing healthy gait from pathological gait patterns. The features used in this study include a diverse set of acceleration and gyroscope measurements collected from the head (HE\_Acc, HE\_FreeAcc, HE\_Gyr), lower back (LB\_Acc, LB\_FreeAcc, LB\_Gyr), and bilateral foot sensors (LF\_Acc, LF\_FreeAcc, LF\_Gyr for the left foot; RF\_Acc, RF\_FreeAcc, RF\_Gyr for the right foot). Together, these measurements capture a rich representation of whole-body dynamics during gait.

The first layer of interpretability arises from the IG heatmaps, which illustrate how the model allocates attention to different sensor channels over time for each selected representative sample. Each heatmap contains twelve rows, each corresponding to one of the twelve sensor features, with columns spanning the full temporal sequence

of the gait cycle. The color intensities encode the magnitude of the IG attributions, allowing the viewer to identify which sensor-time combinations most influence the model's prediction. The point of maximum attribution is marked in each heatmap, serving as an indicator of the single most influential sensor reading in the sample. Through these visualizations, it becomes evident that the model's feature utilization is not uniform across classes; rather, each pathology exhibits its own unique pattern of spatial and temporal emphasis.

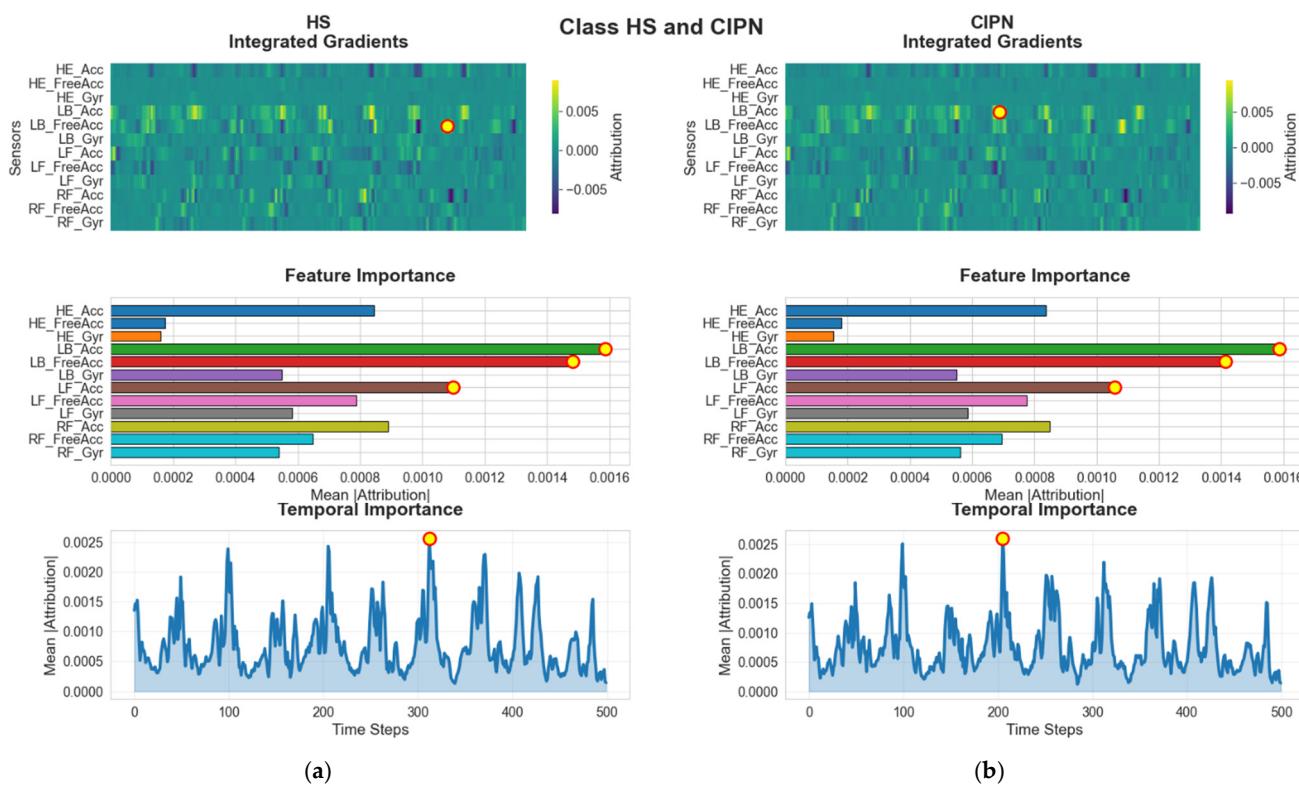
For HSs and participants with CIPN, the heatmaps in Figure 8 consistently highlight the lower-back acceleration sensor (LB\_Acc). This feature captures the linear acceleration of the trunk, which is a central component of the gait process and highly sensitive to deviations in postural control or stability. In both HS and CIPN samples, LB\_Acc displays a pronounced attribution concentration around the mid-stance region of the gait cycle. This similarity may reflect the fundamental importance of trunk dynamics in maintaining balance and gait rhythm, even though CIPN subjects often exhibit subtle alterations due to sensory deficits. Nevertheless, the model successfully captures class-specific signatures, as reflected in the distinct intensity distributions across time. The fact that both groups rely heavily on the LB\_Acc sensor suggests that trunk kinematics contain robust diagnostic information for discriminating against normal gait patterns from those affected by neuropathic conditions.

In contrast to the HS and CIPN groups, the SVA and PD (see Figure 9) cohorts show different attribution patterns aligned with their respective biomechanical abnormalities. For individuals with SVA, the model assigns strongest importance to LF\_Gyr, the angular velocity of the left foot. This sensor primarily reflects rotational foot movements, which may be particularly relevant for identifying stability-related deviations characteristic of SVA. The IG heatmaps show that LF\_Gyr accumulates most of its attribution early in the gait cycle, around frame 10. This finding suggests that the onset of stance or the initial contact phase carries discriminative cues for this population. For subjects with PD, however, the model's focus shifts toward RF\_Acc, the acceleration of the right foot. Parkinsonian gait is known for asymmetries, reduced foot clearance, and irregular acceleration patterns. The concentration of attribution around frame 200 for PD subjects reflects characteristic movement disruptions occurring during the mid-stance or transition phases, when foot stability is challenged. These class-specific patterns demonstrate that the G-MASA-TCN model is attuned to the biomechanical signatures associated with distinct pathologies.

The RIL and ACL groups (See Figure 10) further exhibit unique attribution patterns centered around lower-limb accelerations. For RIL, the model's attention is most strongly associated with LF\_Acc, indicating that the linear acceleration of the left foot serves as a key differentiator for this condition. The temporal peak around frame 100 suggests that mid-stance alterations play a critical role in distinguishing this group. In ACL subjects, however, the dominant feature reverts back to LB\_Acc. The attribution for this cohort is delayed relative to others, reaching its highest intensity around frame 380, near the end of the gait cycle. This timing suggests that late-stance and pre-swing phases when the knee undergoes significant load-bearing and extension may contain the strongest indicators of compensation or deficit resulting from ACL.

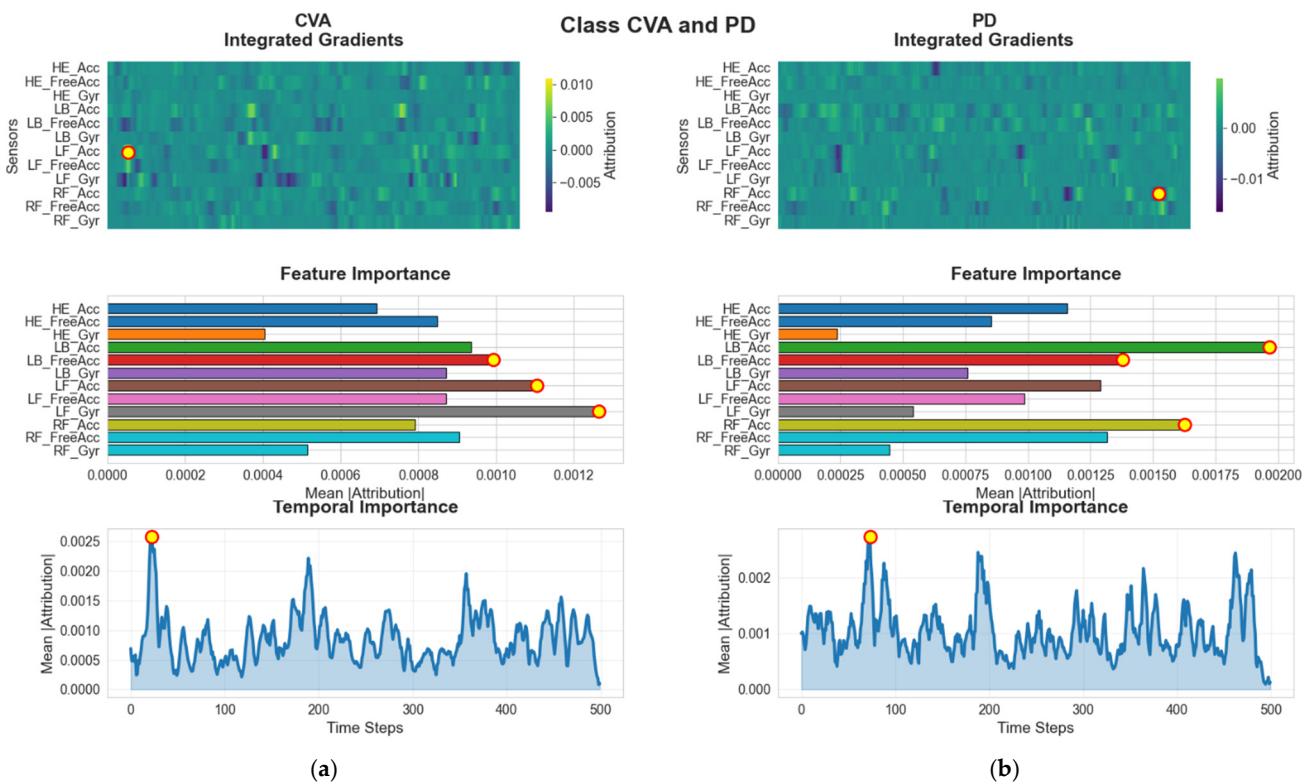
Similarly distinct patterns arise in the HOA and KOA groups (see Figure 11). In HOA samples, the head kinematics specifically HE\_FreeAcc (head acceleration minus gravity) are highlighted as the most influential feature. This sensor captures subtle head-movement adjustments associated with discomfort or instability resulting from HOA. The temporal importance peak around frame 220 suggests that the model identifies instability or compensatory adjustments during late mid-stance or early push-off as the key discriminative

moment. For KOA subjects, the model again prioritizes RF\_Acc, the right-foot acceleration, with its temporal peak around frame 190. KOA commonly influences loading, propulsion, and weight transfer, which explains the reliance on foot acceleration measurements during the mid-gait phases.



**Figure 8.** G-MASA-TCN XAI for Experiment 2 (healthy vs. CIPN) where the red circle with yellow dot is the dominant feature. **(a)** HS: Integrated Gradients (IG) attribution highlights LB\_FreeAcc as the primary feature driving the model’s prediction. The individual-sample IG heatmaps and the global importance panel both show a strong concentration of relevance on LB\_Acc and LB\_FreeAcc. Temporally, the model places the highest emphasis around frame ~200. **(b)** CIPN Cohort: For participants with CIPN, IG attribution similarly identifies LB\_Acc as the dominant discriminatory feature. The global relevance plot confirms its importance, and the temporal IG distribution again peaks around frame ~200, indicating a consistent temporal sensitivity.

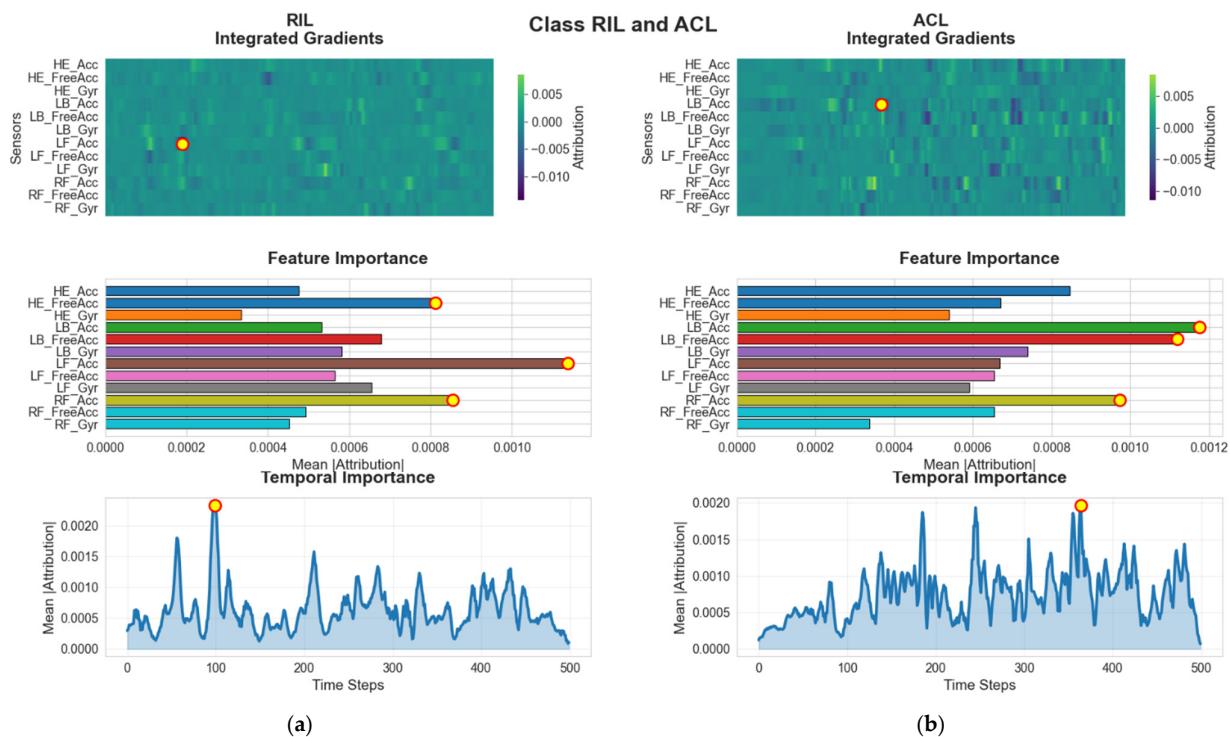
While the IG heatmaps provide sample-level insights, the global feature importance plots aggregate these attributions across time to identify which sensors contribute most consistently across the gait cycle. For each class, the three highest-ranking features reflect the sensors most frequently relied upon by the model. Importantly, the dominant features in the heatmaps always correspond to those with highest mean importance globally. Thus, the importance assigned to LB\_Acc in CIPN and ACL, LF\_Gyr in SVA, RF\_Acc in PD and KOA, LF\_Acc in RIL, and HE\_FreeAcc in HOA is not limited to single time steps but represents a consistent pattern across the entire gait sequence. These global visualizations help reveal the biomechanical drivers characteristic of each pathology: trunk acceleration for neuropathy-related or ligament injuries, foot angular velocity for balance-related abnormalities, head acceleration for hip-related impairments, and foot acceleration for Parkinsonian and osteoarthritic gait deviations.



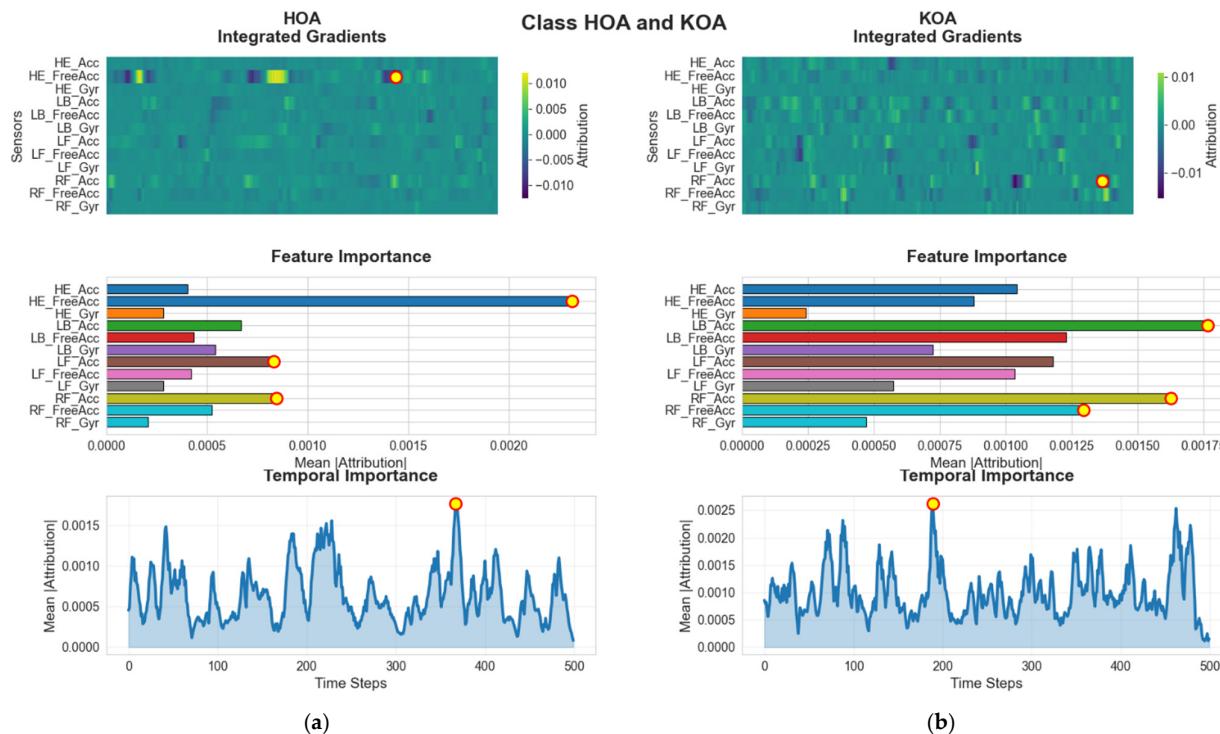
**Figure 9.** G-MASA-TCN XAI for Experiment 2 (SVA vs. PD) where the red circle with yellow dot is the dominant feature. (a) SVA Group: IG explanations reveal that LF\_Gry is the key feature contributing to classification, as shown by both the sample-level heatmaps and the global feature ranking. The model displays a distinct temporal focus near frame ~10, suggesting early-sequence relevance. (b) For PD Group samples, RF\_Acc emerges as the most influential feature. The IG heatmaps and global importance panel consistently highlight this feature, while the temporal relevance pattern reaches its maximum around frame ~200.

The third visualization modality, the temporal importance plot, provides complementary insights by summarizing how the model’s attention evolves across the gait cycle. By averaging attributions across all sensor channels in each time step, the plot identifies critical phases that the model deems most informative. The temporal patterns vary substantially across conditions. HSs and CIPN show their peaks near frame 200, SVA near frame 10, PD again near frame 200, RIL around frame 100, ACL close to frame 380, HOA near frame 220, and KOA around frame 190. These differences emphasize that each pathology manifests discriminative movement signatures at distinct gait phases. By identifying when the model focuses on particular gait events initial contact, mid-stance, push-off, or late-swing, researchers can correlate these moments with clinical gait abnormalities.

Taken together, these three forms of visual explanation offer a rich, multilayered interpretation of the G-MASA-TCN model’s predictions. The IG heatmaps reveal which sensors the model treats as most informative, the global feature importance plots summarize the overall influence of each sensor across time, and the temporal relevance plots pinpoint when in the gait cycle the critical diagnostic information emerges. This integrated analysis enhances understanding of the model’s internal logic, enabling clinicians and researchers to identify not only the anatomical locations most relevant for classification but also the specific phases of movement that contain the strongest biomechanical indicators of pathology. Ultimately, these insights contribute to a more transparent and interpretable gait-classification system, supporting the reliability, clinical credibility, and potential translational application of the model across diverse gait disorders.



**Figure 10.** G-MASA-TCN XAI for Experiment 2 (RIL vs. ACL) where the red circle with yellow dot is the dominant feature. **(a)** RIL Group: IG attributions indicate that LF\_Acc is the dominant feature across both individual and global analyses. The temporal contribution curve identifies a notable relevance peak at frame ~100, marking this interval as critical for classification. **(b)** ACL Group: For individuals with ACL injuries, LB\_Acc stands out as the leading predictive feature, supported by both sample-level IG visualizations and the global feature ranking. Temporal importance increases sharply around frame ~380, highlighting a late-sequence decision region.



**Figure 11.** G-MASA-TCN XAI for Experiment 2 (HOA vs. KOA) where the red circle with yellow dot is the dominant feature. **(a)** HOA Group: The IG results emphasize HE\_FreeAcc as the most informative

feature for distinguishing HOA samples. The global importance analysis confirms its consistent contribution, and the temporal relevance peaks around frame ~220, indicating a mid-sequence focus. **(b) KOA Group:** For the KOA cohort, RF\_Acc is identified as the principal classification feature, clearly visible in both the IG heatmaps and the global feature importance distribution. The model's temporal attention is strongest near frame ~190, marking this region as crucial for prediction.

## 4. Discussion

### 4.1. Result Discussion

The results of this study demonstrate the effectiveness of integrating multi-sensor wearable gait data with advanced temporal deep learning architectures to capture clinically relevant movement signatures across a broad spectrum of neurological and musculoskeletal conditions. By leveraging information from four strategically placed IMUs located on the head, lower back, and both feet, the model was able to exploit both global and segment-specific movement patterns, resulting in highly accurate classification across diverse cohorts. The consistently strong performance across all experimental conditions underscores the richness of wearable inertial data and the suitability of temporal convolutional representations for gait analysis.

A key outcome of this work is the demonstration that gait dynamics, when represented as multi-channel time series, contain sufficiently distinctive patterns to differentiate among multiple pathologies as well as between individuals. The large, clinically annotated dataset used in this study provided a wide range of gait behaviors, including patterns associated with neuropathy, osteoarthritis, ligament injuries, and neurological impairment, alongside healthy control data. The high classification accuracy obtained across repeated cross-validation suggests that these conditions manifest in reproducible sensor-level and temporal signatures. This supports the broader view within the gait analysis community that wearable kinematic data can serve as a reliable proxy for more complex biomechanical measurements traditionally obtained using laboratory motion-capture systems.

Beyond accuracy, the interpretability component of this work provided valuable insights into how different conditions influence movement patterns. The Integrated Gradients analysis revealed that certain sensor channels particularly those associated with lower-back acceleration, foot rotations, and free-acceleration measurements played a dominant role in many classifications. These findings align with established biomechanical understanding: conditions such as neuropathy and osteoarthritis often alter foot-ground interaction forces, trunk stability, and gait symmetry, all of which are reflected in the accelerometer and gyroscope signals captured by wearable IMUs. Temporal attribution further highlighted that specific phases of the gait cycle, such as loading response, mid-stance, or push-off, were more informative for distinguishing pathological gait from normal movement. The convergence of these data-driven insights with clinical expectations strengthens confidence in the relevance and reliability of the extracted gait features. Further, the strong accuracy obtained on real-world gait data highlights the model's ability to generalize effectively to uncontrolled walking conditions, capturing discriminative subject-specific gait signatures while maintaining robustness across different evaluation protocols.

Another notable result is the model's ability to identify individuals with high accuracy in the subject-identification experiment. This indicates that gait, when captured at sufficient temporal resolution, functions as a unique biometric signature. Such findings have important implications for personalized healthcare, long-term patient monitoring, and the development of individualized rehabilitation protocols. They also highlight the robustness of the temporal representation learned by the model, which was able to generalize beyond pathology classification to a fundamentally different task requiring fine-grained discrimination.

Overall, the findings advance the field by showing that high-performing, interpretable gait classification is achievable using lightweight wearable sensors and modern deep learning techniques. The integration of explainable AI ensures that predictions are not treated as black-box outputs but are instead grounded in meaningful sensor-level and temporal evidence. This combination of performance, robustness, and transparency positions the proposed framework as a promising tool for future clinical applications, including automated screening, fall-risk assessment, disease progression monitoring, and personalized rehabilitation. Future work may explore real-time deployment, continuous gait tracking in free-living environments, and the use of domain adaptation to enhance applicability across populations and sensor configurations.

#### 4.2. Comparison with the Original MASA-TCN (EEG Emotion Recognition)

Although our gait classification model retains the core structure of the original MASA-TCN proposed in [26] for EEG-based emotion recognition, several adaptations were introduced to better suit the characteristics of inertial gait data. In the original work, the input consisted of relative power spectral density (rPSD) features computed across multiple frequency bands (typically 5–6), resulting in a tensor of the shape  $(B, C \times f, T)$ , where  $f$  denotes the number of spectral bands. The SAT layer was designed to operate on a per-channel, per-frequency basis, using a context kernel of the size  $(f, k)$  and stride of  $(f, 1)$  to extract band-specific temporal patterns, followed by spatial fusion across electrode channels via a kernel of the size  $(C, 1)$ .

In contrast, our implementation processes raw time-domain IMU signals with 36 channels and no explicit frequency decomposition. The SAT context convolution therefore uses a kernel of the size  $(36, k)$ , treating the full set of sensor streams as a single spatial dimension. This modification preserves the spirit of space-aware modeling but reinterprets “space” as anatomical sensor placement rather than spectral content. While the original model supported both continuous emotion regression (using Concordance Correlation Coefficient loss) and discrete classification, our application focuses exclusively on multi-class discrete classification via cross-entropy, necessitating only the mean-pooling classification head.

Activation functions also differ: while the original model used PReLU throughout, ReLU activation followed by batch normalization is applied in the MAAF block, which has shown superior stability when training on raw acceleration and angular velocity signals prone to high variance. The TCN dilation schedule is fixed at  $\{1, 2, 4\}$  in our version, yielding a consistent receptive field across experiments, whereas the original dynamically adjusted dilations starting from the SAT layer’s  $d_t = 2$ . Despite these differences, the fundamental contribution—adaptive fusion of multi-scale, spatially informed temporal features—is retained and transferred from brain signals to full-body movement dynamics [26].

#### 4.3. Comparison with State-of-the-Art Gait Classifiers

To contextualize the performance and design of our G-MASA-TCN, it is instructive to compare it with leading deep learning architectures previously applied to IMU-based gait classification. One prominent approach is the DeepConvLSTM framework introduced by Ordóñez and Roggen (2016) [38], which combines four convolutional layers for local feature extraction with two bidirectional LSTM layers to model long-range temporal dependencies. While effective, this hybrid architecture relies on sequential processing in the recurrent component, resulting in higher computational cost and memory usage during inference particularly challenging for real-time clinical applications. Moreover, the LSTM component

introduces future context through bidirectionality, violating causality and preventing online deployment.

Another competitive baseline is the attention-augmented TCN-BiGRU model proposed in [40], which stacks dilated temporal convolutional layers with bidirectional GRUs and attention mechanisms. This design achieves a large receptive field of approximately 340 ms via exponential dilation in the TCN:

$$RF_{TCN} = 1 + (k - 1) \sum_{l=1}^L 2^{l-1}, k = 3, L = 8 \quad (32)$$

This yields strong performance on activity recognition tasks; however, the inclusion of bidirectional recurrence again precludes causal inference, and the parameter count remains high ( $\sim 1.0$  M) due to the recurrent module and attention softmax computations. In contrast, our G-MASA-TCN eliminates recurrence entirely, relying instead on dilated convolutions and multi-scale parallelism to achieve an effective receptive field exceeding 1.2 s, sufficient to encompass multiple gait cycles while maintaining full causality and parallel computation.

More recently, Xiong Wei and Zifan Wan (2024) [39] introduced a TCN-Attention model that augments a temporal convolutional backbone with channel-wise and temporal attention mechanisms to dynamically weigh sensor streams and key timestamps. This approach improves interpretability and adaptability to varying signal quality but introduces additional computational overhead from the attention scoring and softmax operations. Our G-MASA-TCN achieves comparable sensor prioritization implicitly through the learned  $1 \times 1$  fusion weights in the MAAF block and the spatially joint 2D convolutions in SAT layers, without requiring explicit attention modules. This results in a simpler, faster model with fewer hyperparameters.

In terms of parameter efficiency, our model (0.92 M parameters) offers a favorable trade-off between capacity and deployability compared to heavier recurrent hybrids ( $> 1.2$  M) and attention-augmented models ( $\sim 0.8\text{--}1.0$  M). Crucially, the explicit multi-scale anchor design and built-in spatial fusion provide structured inductive biases tailored to gait kinematics, potentially leading to better generalization across heterogeneous patient populations compared to generic attention or recurrent mechanisms [38–40].

Although direct comparison with prior methods on the same dataset is not possible due to the dataset's recent release (October 2025), the existing literature highlights several limitations of commonly used baseline models for IMU-based gait analysis. Recurrent models such as GRUs often struggle with long-range temporal dependencies and are sensitive to sequence length and noise [38,51]. Transformer-based models, while powerful, typically require large-scale datasets and incur high computational cost, limiting their effectiveness in moderate-sized clinical datasets [52,53]. Similarly, conventional TCNs rely on fixed dilation patterns, which may restrict their ability to capture heterogeneous temporal dynamics across different gait pathologies [45].

In contrast, the proposed G-MASA-TCN integrates multi-scale temporal modeling and adaptive attention mechanisms, enabling robust learning of both short- and long-term gait characteristics across diverse clinical conditions. This architectural design aligns with recent trends in state-of-the-art IMU-based gait analysis [54] while extending them to a more challenging multi-cohort setting.

#### 4.4. Computational Complexity and Power Consumption Analysis

The proposed G-MASA-TCN integrates multi-scale dilated temporal convolutions, channel-wise attention, and residual temporal modeling. While these design choices enhance representational capacity, they may increase computational demand. Therefore, a detailed complexity analysis in terms of Gigaoperations (GOPs) and an estimation of inference-time power consumption are provided to assess practical deployability.

#### 4.4.1. Operation Count (GOP Analysis)

The dominant computational cost of G-MASA-TCN arises from one-dimensional convolutional layers. For a 1D convolution with input channels,  $C_{\text{in}}$ , output channels,  $C_{\text{out}}$ , the kernel size  $K$ , and the temporal length  $T$ , the number of floating-point operations (FLOPs) is approximated as

$$\text{FLOPs} \approx 2 \times C_{\text{in}} \times C_{\text{out}} \times K \times T \quad (33)$$

where the factor of 2 accounts for multiplication and addition. Using the fixed-length input setting adopted in this work ( $T = 500$ ,  $C = 36$ ), the total complexity per forward pass is summarized as follows:

1. SAT layers (multi-scale temporal extraction)
  - SAT-1:  $36 \rightarrow 64; K = 3; d = 1: \approx 0.007$  GOPs.
  - SAT-2:  $64 \rightarrow 128; K = 5; d = 2: \approx 0.041$  GOPs.
  - SAT-3:  $128 \rightarrow 256; K = 7; d = 4: \approx 0.229$  GOPs.

2. Attention mechanism

The channel-attention module consists of a fully connected projection applied after global average pooling. Its computational cost is negligible relative to convolutional layers (<0.001 GOPs).

3. Temporal convolution block with residual connection
  - TCN-1:  $256 \rightarrow 128; K = 3; d = 2: \approx 0.098$  GOPs.
  - TCN-2:  $128 \rightarrow 64; K = 3; d = 4: \approx 0.025$  GOPs.
  - Residual projection ( $1 \times 1$  convolution):  $\approx 0.008$  GOPs.

4. Fully connected classification head

The dense layers contribute marginally (<0.001 GOPs). Summing all components, the total computational complexity of G-MASA-TCN is approximately 0.41 GOPs per inference sample for a 5 s gait segment.

#### 4.4.2. Power Consumption Estimation

To estimate inference-time energy consumption, we adopt commonly reported hardware efficiency values. Modern GPUs and embedded AI accelerators typically consume approximately 0.5–1.0 nJ per FLOP, while mobile CPUs require slightly higher energy per operation. Assuming a conservative value of 1 nJ/FLOP, the estimated energy per inference is

$$E \approx 0.41 \times 10^9 \times 1 \text{ nJ} = 0.41 \text{ mJ} \quad (34)$$

This low per-sample energy requirement indicates that G-MASA-TCN is suitable for real-time gait analysis and continuous monitoring scenarios. Importantly, inference is causal and convolution-based, avoiding the quadratic complexity associated with self-attention in Transformer models.

#### 4.4.3. Discussion of Efficiency

Compared with baseline architectures evaluated in this study, G-MASA-TCN exhibits a favorable trade-off between accuracy and efficiency. Transformer-based models typically scale as  $O(T^2)$  with sequence length, leading to substantially higher computational and memory costs for long gait sequences. Recurrent architectures such as the GRU require sequential processing, limiting parallelization and increasing latency. In contrast, G-MASA-TCN scales linearly with temporal length ( $O(T)$ ) and supports full parallel inference. Despite integrating multi-scale temporal fusion and attention mechanisms, the model

remains computationally lightweight, enabling deployment on wearable, edge, or clinical monitoring systems without excessive power consumption.

#### 4.5. Real-Time Inference and Latency Analysis

Although G-MASA-TCN requires several hours for offline training due to cross-validation and subject-level learning, real-time deployment depends solely on inference-time complexity. Therefore, latency and throughput are analyzed using GOP-based benchmarking.

As derived in Section 4.4, the proposed G-MASA-TCN requires approximately 0.41 GOPs per forward pass for a 5 s gait segment (500 samples at 100 Hz). Assuming a conservative inference with the input of 100 GOPs on a standard CPU, the expected inference latency is

$$\text{Latency} = \frac{0.41}{100} \approx 4.1 \text{ ms} \quad (35)$$

On embedded AI platforms (300–1000 GOPs), latency reduces to 1.4–0.4 ms, while on modern GPUs (>5000 GOPs), sub-millisecond inference is achieved. These values are significantly lower than the data acquisition interval (5 s), confirming that G-MASA-TCN comfortably satisfies real-time gait analysis requirements. In contrast, Transformer-based architectures scale quadratically with sequence length ( $O(T^2)$ ), leading to higher latency and memory overhead. Recurrent models such as the GRU require sequential processing, further increasing inference delay. G-MASA-TCN, by relying exclusively on parallel dilated convolutions, achieves low latency and deterministic inference time. These results demonstrate that the proposed framework is suitable for real-time clinical gait monitoring and wearable deployment.

#### 4.6. Implications for Human Activity Recognition

The analytical framework developed in this study carries significant implications for the broader human activity recognition (HAR) domain. By integrating interpretable deep learning mechanisms with multi-sensor temporal modeling, the approach demonstrates that robust activity classification can be achieved without sacrificing transparency. This is particularly valuable in healthcare and safety-critical applications, where model accountability and clinical validation are essential. The ability of the framework to quantify sensor-level importance allows practitioners to understand not only what classification outcome was reached but also *why*, enabling deeper insight into the underlying physiological or behavioral patterns present in the data.

A central implication for the HAR field lies in the identification of key sensors and movement phases that consistently contribute to accurate predictions. Such insights can guide system designers in optimizing sensor placement, reducing instrumentation costs, and improving the feasibility of long-term monitoring systems. For example, the observation that trunk acceleration or foot rotational data frequently dominate model decisions suggests that monitoring these anatomical regions may be sufficient for capturing essential movement characteristics in certain applications. Moreover, temporal attribution highlights the specific phases of a movement cycle that carry the greatest informational value, supporting more refined feature engineering and potentially enabling segment-aware detection pipelines.

More broadly, the interpretability results demonstrate that complex human activities, whether related to gait, sports performance, workplace ergonomics, or daily-life monitoring, exhibit structured temporal patterns that can be captured reliably through lightweight wearable sensors. The combination of high accuracy and transparent decision pathways provides a foundation for translating HAR models into practical domains such as physical rehabilitation, athletic performance analysis, behavior tracking, occupational safety, and

smart-home environments. In these contexts, understanding the reasoning behind system outputs is essential for ensuring acceptance by practitioners, promoting user trust, and facilitating regulatory approval when required.

#### 4.7. Limitations and Generalization Considerations

While the framework demonstrates strong performance and interpretability, it is important to acknowledge several factors that influence the reliability and generalization of the findings. First, the Integrated Gradients (IG) method depends on smooth transitions between the baseline and actual inputs. Gait and other natural human activities often contain abrupt changes such as impacts, rapid directional shifts, or irregular compensatory movements which may challenge the smoothness assumption. In such situations, attributions may emphasize or de-emphasize certain regions of the signal, potentially affecting the interpretive clarity. Recognizing these conditions is important when applying IG to activities that exhibit high kinetic variability or irregular timing patterns.

Another consideration relates to the use of temporally averaged attribution measures. While these provide intuitive summaries of the most influential phases of an activity, they may mask more nuanced cross-sensor interactions or interdependencies that are crucial for understanding multidimensional human activities. Many real-world behaviors, whether in clinical populations or non-clinical settings, involve coordinated multi-joint movement patterns in which the interplay between different body segments conveys essential information. Averaging across features, although useful for visualization, may overlook these relationships and should therefore be interpreted with caution.

Generalizability across different populations, activity types, and environmental contexts also warrants consideration. Although the dataset used in this study is large and clinically diverse, real-world HAR applications often involve variations in sensor placement, individual movement styles, footwear, terrain, and daily-life unpredictability. These factors introduce variability that may affect the transferability of the learned representations. As such, extending the evaluation to larger, more heterogeneous datasets including free-living environments outside controlled laboratory settings would strengthen confidence in the model's applicability to broader HAR tasks.

The interpretability insights generated here suggest opportunities for applying similar analytical frameworks in other fields that rely on multivariate physiological or behavioral time series. Domains such as cardiology, sleep analysis, workplace fatigue monitoring, and sports biomechanics all involve structured temporal signals where identifying key sensors and critical time windows can enhance both model performance and domain understanding. The observations from this study underscore the importance of integrating transparent learning mechanisms with rich sensor data to support informed decision-making in technologically assisted human-monitoring systems.

Importantly, this study serves as a pillar for future research, encouraging the community to extend data collection to more diverse and real-world scenarios. Potential directions include incorporating children, rare gait disorders, and everyday walking activities, such as stair climbing, navigating uneven terrain, or carrying objects. By providing this baseline, the current work lays the groundwork for advancing IMU-based gait analysis in complex and clinically relevant environments and inspires further innovations in low-cost, wearable sensor applications.

### 5. Conclusions

The G-MASA-TCN framework captures multi-scale temporal dynamics and sensor-specific contributions for human activity recognition using wearable IMUs. By combining multi-anchor temporal convolutions with spatial attention, the model learns discrimina-

tive movement patterns across diverse populations. Integrated Gradients, global feature importance, and temporal relevance visualizations provide a clear, interpretable view of predictions, with peak contributions highlighted for key sensors and time segments. This approach enhances transparency and trust in deep learning models, demonstrating that high classification accuracy can be achieved without compromising interpretability and supporting its potential application in clinical monitoring, rehabilitation, and real-world human activity analysis.

**Funding:** The author extend their appreciation to the Deanship of Research and Graduate Studies at King Khalid University for funding this study through Large Research Project under grant number RGP2/176/46.

**Data Availability Statement:** The data presented in this study are available through [figshare] at "[https://springernature.figshare.com/articles/dataset/A\\_Dataset\\_of\\_Clinical\\_Gait\\_Signals\\_with\\_Wearable\\_Sensors\\_from\\_Healthy\\_Neurological\\_and\\_Orthopedic\\_Cohorts/28806086?file=53704514](https://springernature.figshare.com/articles/dataset/A_Dataset_of_Clinical_Gait_Signals_with_Wearable_Sensors_from_Healthy_Neurological_and_Orthopedic_Cohorts/28806086?file=53704514)" [accessed on 4 January 2026], reference number [37].

**Conflicts of Interest:** The author declares no conflict of interest.

## References

1. Kim, U.; Lim, J.; Park, Y.; Bae, Y. Predicting fall risk through step width variability at increased gait speed in community dwelling older adults. *Sci. Rep.* **2025**, *15*, 16915. [[CrossRef](#)]
2. Hausdorff, J.M. Gait dynamics, fractals and falls: Finding meaning in the stride-to-stride fluctuations of human walking. *Hum. Mov. Sci.* **2007**, *26*, 555–589. [[CrossRef](#)]
3. Prince, M.J.; Wu, F.; Guo, Y.; Gutierrez Robledo, L.M.; O'Donnell, M.; Sullivan, R.; Yusuf, S. The burden of disease in older people and implications for health policy and practice. *Lancet* **2015**, *385*, 549–562. [[CrossRef](#)]
4. Cruz-Jentoft, A.J.; Sayer, A.A. *Sarcopenia*. *Lancet* **2019**, *393*, 2532–2546. [[CrossRef](#)]
5. Gill, T.M.; Gahbauer, E.A.; Allore, H.G.; Han, L. Transitions Between Frailty States Among Community-Living Older Persons. *JAMA Intern. Med.* **2006**, *166*, 418–423. [[CrossRef](#)]
6. Verghese, J.; Lipton, R.B.; Hall, C.B.; Kuslansky, G.; Katz, M.J.; Buschke, H. Abnormality of Gait as a Predictor of Non-Alzheimer's Dementia. *N. Engl. J. Med.* **2002**, *347*, 1761–1768. [[CrossRef](#)]
7. Chen, S.; Lach, J.; Lo, B.; Yang, G.Z. Toward Pervasive Gait Analysis with Wearable Sensors: A Systematic Review. *IEEE J. Biomed. Health Inform.* **2016**, *20*, 1521–1537. [[CrossRef](#)]
8. Zhang, H.; Guo, Y.; Zanotto, D. Accurate Ambulatory Gait Analysis in Walking and Running Using Machine Learning Models. *IEEE Trans. Neural Syst. Rehabil. Eng.* **2020**, *28*, 187–195. [[CrossRef](#)]
9. Dasgupta, P.; VanSwearingen, J.; Godfrey, A.; Redfern, M.; Montero-Odasso, M.; Sejdic, E. Acceleration Gait Measures as Proxies for Motor Skill of Walking: A Narrative Review. *IEEE Trans. Neural Syst. Rehabil. Eng.* **2021**, *29*, 249–261. [[CrossRef](#)]
10. Sabatini, A.M.; Martelloni, C.; Scapellato, S.; Cavallo, F. Assessment of walking features from foot inertial sensing. *IEEE Trans. Biomed. Eng.* **2005**, *52*, 486–492. [[CrossRef](#)] [[PubMed](#)]
11. Zhao, Q.; Chen, Z.; Landis, C.D.; Lytle, A.; Rao, A.K.; Guo, Y. Gait Analysis with an Integrated Mobile Robot and Wearable Sensor System Reveals Associations Between Cognitive Ability and Dynamic Balance in Older Adults. In Proceedings of the 2022 9th IEEE RAS/EMBS International Conference for Biomedical Robotics and Biomechatronics (BioRob), Seoul, Republic of Korea, 21–24 August 2022; pp. 1–6. [[CrossRef](#)]
12. Figueiredo, J.; Carvalho, S.P.; Vilas-Boas, J.P.; Gonçalves, L.M.; Moreno, J.C.; Santos, C.P. Wearable Inertial Sensor System towards Daily Human Kinematic Gait Analysis: Benchmarking Analysis to MVN BIOMECH. *Sensors* **2020**, *20*, 2185. [[CrossRef](#)]
13. Hori, K.; Mao, Y.; Ono, Y.; Ora, H.; Hirobe, Y.; Sawada, H.; Inaba, A.; Orimo, S.; Miyake, Y. Inertial Measurement Unit-Based Estimation of Foot Trajectory for Clinical Gait Analysis. *Front. Physiol.* **2020**, *10*, 1530. [[CrossRef](#)]
14. Rampp, A.; Barth, J.; Schülein, S.; Gaßmann, K.G.; Klucken, J.; Eskofier, B.M. Inertial Sensor-Based Stride Parameter Calculation From Gait Sequences in Geriatric Patients. *IEEE Trans. Biomed. Eng.* **2015**, *62*, 1089–1097. [[CrossRef](#)]
15. Kobsar, D.; Charlton, J.M.; Tse, C.T.F.; Esculier, J.F.; Graffos, A.; Krowchuk, N.M.; Thatcher, D.; Hunt, M.A. Validity and reliability of wearable inertial sensors in healthy adult walking: A systematic review and meta-analysis. *J. Neuroeng. Rehabil.* **2020**, *17*, 62. [[CrossRef](#)]
16. Gao, J.; Gu, P.; Ren, Q.; Zhang, J.; Song, X. Abnormal Gait Recognition Algorithm Based on LSTM-CNN Fusion Network. *IEEE Access* **2019**, *7*, 163180–163190. [[CrossRef](#)]

17. Sarshar, M.; Polturi, S.; Schega, L. Gait Phase Estimation by Using LSTM in IMU-Based Gait Analysis—Proof of Concept. *Sensors* **2021**, *21*, 5749. [[CrossRef](#)]
18. Lu, Z.; Zhou, H.; Wang, L.; Kong, D.; Lyu, H.; Wu, H. GaitFormer: Two-Stream Transformer Gait Recognition Using Wearable IMU Sensors in the Context of Industry 5.0. *IEEE Sens. J.* **2025**, *25*, 19947–19956. [[CrossRef](#)]
19. Tao, S.; Goh, W.L.; Gao, Y. A Convolved Self-Attention Model for IMU-based Gait Detection and Human Activity Recognition. In Proceedings of the 2023 IEEE 5th International Conference on Artificial Intelligence Circuits and Systems (AICAS), Hangzhou, China, 11–13 June 2023. [[CrossRef](#)]
20. Michael, C.L.; Flynn, D.; Vidal, R.; Reiter, A.; Hager, G.D. Temporal Convolutional Networks for Action Segmentation and Detection. In Proceedings of the 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Honolulu, HI, USA, 21–26 July 2017. [[CrossRef](#)]
21. Addabbo, P.; Bernardi, M.L.; Biondi, F.; Cimitile, M.; Clemente, C.; Orlando, D. Temporal Convolutional Neural Networks for Radar Micro-Doppler Based Gait Recognition. *Sensors* **2021**, *21*, 381. [[CrossRef](#)]
22. Teufl, W.; Lorenz, M.; Mieza, M.; Taetz, B.; Fröhlich, M.; Bleser, G. Towards Inertial Sensor Based Mobile Gait Analysis: Event-Detection and Spatio-Temporal Parameters. *Sensors* **2019**, *19*, 38. [[CrossRef](#)]
23. Collins, S.; Wiggin, M.; Sawicki, G. Reducing the energy cost of human walking using an unpowered exoskeleton. *Nature* **2015**, *522*, 212–215. [[CrossRef](#)]
24. Kang, I.; Hsu, H.; Young, A. The Effect of Hip Assistance Levels on Human Energetic Cost Using Robotic Hip Exoskeletons. *IEEE Robot. Autom. Lett.* **2019**, *4*, 430–437. [[CrossRef](#)]
25. Zhang, T.; Li, Y.; Ning, C.; Zeng, B. Development and Adaptive Assistance Control of the Robotic Hip Exoskeleton to Improve Gait Symmetry and Restore Normal Gait. *IEEE Trans. Autom. Sci. Eng.* **2024**, *21*, 799–809. [[CrossRef](#)]
26. Ding, Y.; Zhang, S.; Tang, C.; Guan, C. MASA-TCN: Multi-Anchor Space-Aware Temporal Convolutional Neural Networks for Continuous and Discrete EEG Emotion Recognition. *IEEE J. Biomed. Health Inform.* **2024**, *28*, 1113–1124. [[CrossRef](#)]
27. Song, T.; Zheng, W.; Song, P.; Cui, Z. EEG Emotion Recognition Using Dynamical Graph Convolutional Neural Networks. *IEEE Trans. Affect. Comput.* **2020**, *11*, 532–541. [[CrossRef](#)]
28. Arjun, A.; Rajpoot, A.S.; Panicker, M.R. Introducing Attention Mechanism for EEG Signals: Emotion Recognition with Vision Transformers. In Proceedings of the 2021 43rd Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC), Guadalajara, Mexico, 1–5 November 2021. [[CrossRef](#)]
29. Wang, S.; Tang, H.; Gao, L.; Tan, Q. Continuous Estimation of Human Joint Angles From sEMGUsing a Multi-Feature Temporal Convolutional Attention-Based Network. *IEEE J. Biomed. Health Inform.* **2022**, *26*, 5461–5472. [[CrossRef](#)]
30. Li, S.; Farha, Y.A.; Liu, Y.; Cheng, M.-M.; Gall, J. MS-TCN++: Multi-Stage Temporal Convolutional Network for Action Segmentation. *IEEE Trans. Pattern Anal. Mach. Intell.* **2023**, *45*, 6647–6658. [[CrossRef](#)]
31. He, Z.; Zhong, Y.; Pan, J. An adversarial discriminative temporal convolutional network for EEG-based cross-domain emotion recognition. *Comput. Biol. Med.* **2022**, *141*, 105048. [[CrossRef](#)]
32. Alarcao, S.M.; Fonseca, M.J. Emotions Recognition Using EEG Signals: A Survey. *IEEE Trans. Affect. Comput.* **2019**, *10*, 374–393. [[CrossRef](#)]
33. Zheng, F.; Hu, B.; Zheng, X.; Zhang, Y. Spatial-temporal features-based EEG emotion recognition using graph convolution network and long short-term memory. *Physiol. Meas.* **2023**, *44*, 065005. [[CrossRef](#)]
34. Li, C.; Zhang, Z.; Zhang, X.; Huang, G.; Liu, Y.; Chen, X. EEG-based Emotion Recognition via Transformer Neural Architecture Search. *IEEE Trans. Ind. Inform.* **2023**, *19*, 6016–6025. [[CrossRef](#)]
35. An, Y.; Xu, N.; Qu, Z. Leveraging spatial-temporal convolutional features for EEG-based emotion recognition. *Biomed. Signal Process. Control* **2021**, *69*, 102743. [[CrossRef](#)]
36. Sheng, W.; Li, X. Multi-task learning for gait-based identity recognition and emotion recognition using attention enhanced temporal graph convolutional network. *Pattern Recognit.* **2021**, *114*, 107864. [[CrossRef](#)]
37. Voisard, C.; Barrois, R.; de L'escalopier, N.; Vayatis, N.; Vidal, P.-P.; Yelnik, A.; Ricard, D.; Oudre, L. A Dataset of Clinical Gait Signals with Wearable Sensors from Healthy, Neurological, and Orthopedic Cohorts. *Sci. Data* **2025**, *12*, 1674. [[CrossRef](#)]
38. Ordóñez, F.J.; Roggen, D. Deep Convolutional and LSTM Recurrent Neural Networks for Multimodal Wearable Activity Recognition. *Sensors* **2016**, *16*, 115. [[CrossRef](#)]
39. Wei, X.; Wang, Z. TCN-attention-HAR: Human activity recognition based on attention mechanism time convolutional network. *Sci. Rep.* **2024**, *14*, 7133. [[CrossRef](#)]
40. He, J.-L.; Wang, J.-H.; Lo, C.-M.; Jiang, Z. Human Activity Recognition via Attention-Augmented TCN-BiGRU Fusion. *Sensors* **2025**, *25*, 5765. [[CrossRef](#)]
41. Molinaro, D.D.; Kang, I.; Camargo, J.; Young, A.J. Biological Hip Torque Estimation using a Robotic Hip Exoskeleton. In Proceedings of the 2020 8th IEEE RAS/EMBS International Conference for Biomedical Robotics and Biomechatronics (BioRob), New York, NY, USA, 29 November–1 December 2020. [[CrossRef](#)]

42. Slade, P.; Kochenderfer, M.J.; Delp, S.L.; Collins, S.H. Personalizing exoskeleton assistance while walking in the real world. *Nature* **2022**, *610*, 277–282. [[CrossRef](#)]
43. Lerner, Z.F.; Damiano, D.L.; Park, H.S.; Gravunder, A.J.; Bulea, T.C. A Robotic Exoskeleton for Treatment of Crouch Gait in Children with Cerebral Palsy: Design and Initial Application. *IEEE Trans. Neural Syst. Rehabil. Eng.* **2017**, *25*, 650–659. [[CrossRef](#)]
44. Luo, Y.; Coppola, S.M.; Dixon, P.C.; Li, S.; Dennerlein, J.T.; Hu, B. A database of human gait performance on irregular and uneven surfaces collected by wearable sensors. *Sci. Data* **2020**, *7*, 219. [[CrossRef](#)]
45. Bai, S.; Kolter, J.Z.; Koltun, V. An Empirical Evaluation of Generic Convolutional and Recurrent Networks for Sequence Modeling. *arXiv* **2018**, arXiv:1803.01271. [[CrossRef](#)]
46. Cho, K.; van Merriënboer, B.; Gulcehre, C.; Bahdanau, D.; Bougares, F.; Schwenk, H.; Bengio, Y. Learning Phrase Representations using RNN Encoder-Decoder for Statistical Machine Translation. *arXiv* **2014**, arXiv:1406.1078. [[CrossRef](#)]
47. Vaswani, A.; Shazeer, N.; Parmar, N.; Uszkoreit, J.; Jones, L.; Gomez, A.N.; Kaiser, L.; Polosukhin, I. Attention Is All You Need. *arXiv* **2017**, arXiv:1706.03762. [[CrossRef](#)]
48. Sundararajan, M.; Taly, A.; Yan, Q. Axiomatic attribution for deep networks. In Proceedings of the 34th International Conference on Machine Learning (ICML), Sydney, Australia, 6–11 August 2017. [[CrossRef](#)]
49. Smilkov, D.; Thorat, N.; Kim Viégas, F.; Wattenberg, M. SmoothGrad: Removing noise by adding noise. *arXiv* **2017**, arXiv:1706.03825. [[CrossRef](#)]
50. Ancona, M.; Ceolini, E.; Öztireli, C.; Gross, M. Towards better understanding of gradient-based attribution methods for deep neural networks. *arXiv* **2018**, arXiv:1711.06104. [[CrossRef](#)]
51. Hammerla, N.Y.; Halloran, S.; Ploetz, T. Deep, convolutional, and recurrent models for human activity recognition using wearables. *arXiv* **2016**, arXiv:1604.08880. [[CrossRef](#)]
52. Zerveas, G.; Jayaraman, S.; Patel, D.; Bhamidipaty, A.; Eickhoff, C. A transformer-based framework for multivariate time series representation learning. *arXiv* **2021**, arXiv:2010.02803. [[CrossRef](#)]
53. Benjamin, S.N.S.; Marlin, M. Multi-time attention networks for irregularly sampled time series. *arXiv* **2021**, arXiv:2101.10318. [[CrossRef](#)]
54. Cristian, J.H.T.K.; Pasluosta, F.; Barth, J.; Schülein, S.; Gaßmann, K.-G.; Klucken, J.; Eskofier, B.M. Mobile Stride Length Estimation With Deep Convolutional Neural Networks. *IEEE J. Biomed. Health Inform.* **2018**, *22*, 354–362. [[CrossRef](#)]

**Disclaimer/Publisher’s Note:** The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.