

RESEARCH ARTICLE

A Self-Attention Enhanced Deep CNN-LSTM-Based Irregular Surface Recognition Approach for Integration Into Lower Limb Prosthesis Systems to Ensure Safety Through Predictive Walking

NORAZIAN SUBARI¹, KAMARUL HAWARI GHAZALI^{1,2}, (Senior Member, IEEE),
AND YUANFA JI³, (Member, IEEE)

¹Faculty of Electrical and Electronics Engineering Technology, Universiti Malaysia Pahang Al-Sultan Abdullah, Pekan Campus, Pekan, Pahang 26600, Malaysia

²Centre for Advanced Industrial Technology, Universiti Malaysia Pahang Al-Sultan Abdullah, Pekan Campus, Pekan, Pahang 26600, Malaysia

³National and Local Joint Engineering Research Center of Satellite Navigation, Localization and Location Service, Guilin 541004, China

Corresponding author: Kamarul Hawari Ghazali (kamarul@umpsa.edu.my)

This work was supported in part by the Universiti Malaysia Pahang Al-Sultan Abdullah through the Research Grant Scheme under Grant RDU2303104.

ABSTRACT Advancements in instrumentation and control systems for lower limb prostheses have substantially improved mobility for amputees. However, significant challenges persist when users encounter irregular surfaces, as most prosthetic systems lack the capability to dynamically adapt to surface variations. This limitation restricts mobility, compromises safety, and diminishes user confidence and security during walking. To address these challenges, integrating inertial measurement units (IMUs) with artificial intelligence (AI) techniques, particularly deep learning (DL) methods, has emerged as a promising solution for surface classification and safety enhancement. This study proposes a self-attention enhanced deep CNN-LSTM model to automatically classify walking surfaces as regular or irregular, utilizing IMU acceleration data collected from prosthetic limbs. The model employs the strengths of convolutional and recurrent neural networks combined with a self-attention mechanism to enhance feature representation and improve classification accuracy. Experimental evaluations reveal that the proposed method achieves a classification accuracy of 99.94%, outperforming existing approaches. This result underscores the model's potential to serve as the basis for AI-driven IMU-based systems, enabling real-time surface recognition and safety alerts in prosthetic devices. By enhancing walking safety and user confidence, this method represents a significant advancement for lower limb prosthesis systems.

INDEX TERMS Artificial intelligence (AI), convolutional neural networks (CNNs), deep learning (DL), inertial measurement units (IMUs), lower limb prostheses.

I. INTRODUCTION

Conventional lower limb prosthetics often struggle to dynamically adapt to irregular walking surfaces commonly encountered in outdoor environments [1], [2]. This limitation results in reduced mobility, instability, and an increased risk of falls

The associate editor coordinating the review of this manuscript and approving it for publication was Muammar Muhammad Kabir¹.

for lower limb amputees. These challenges highlight the urgent need for innovative automation in the instrumentation and control systems of prosthetic devices. Enhancing mobility and safety for amputees involves integrating intelligent systems capable of recognizing irregular surfaces and notifying users in real time [3], [4], [5]. In particular, the combination of inertial measurement units (IMUs) and artificial intelligence (AI) techniques, such as deep learning

(DL), has demonstrated significant potential in surface recognition.

Modern detection and classification techniques increasingly rely on deep learning methods, which have become essential across various domains, including image processing, signal analysis, and data interpretation [6], [7], [8], [9], [10], [11], [12], [13], [14], [15], [16]. These methods provide exceptional accuracy and adaptability, establishing them as a fundamental component of advanced computational systems. Their ability to extract and prioritize critical features from complex data makes them particularly well-suited for surface classification tasks.

The integration of IMUs with DL techniques has proven effective in identifying walking surfaces and delivering alerts for diverse groups, such as overweight individuals, older adults, and healthy individuals [16], [17]. Extending this approach to lower limb prostheses, the combination of IMUs and DL algorithms presents a compelling solution for detecting irregular surfaces and providing timely notifications. This integration has the potential to significantly enhance the safety, confidence, and mobility of lower limb amputees [3], [4], [18]

Ensuring high detection accuracy in the development of deep learning (DL) methods utilizing inertial measurement unit (IMU) data, specifically accelerometer and gyroscope measurements in three directions (vertical, mediolateral, and anterior-posterior) is of utmost importance. This is particularly critical given their application in supporting the health and mobility of vulnerable populations, such as lower limb amputees. Precision in these systems is essential, as inaccuracies could compromise the safety, rehabilitation outcomes, and overall quality of life for these individuals. Prior studies [18], [19], [20] have achieved accuracies of up to 97% by integrating IMU data with DL methods for detecting irregular surfaces encountered by lower limb prosthetic users. In [18], [19], and [20] also, their focus has primarily been on broader locomotor transitions such as sitting, standing, stair ascent/descent, or predictive walking in varied environments. These approaches often require multiple sensors or customized prosthetic configurations and do not specifically address the challenge of distinguishing between regular and irregular walking surfaces. In contrast, this study proposes a task-specific architecture that focuses solely on surface classification—a critical component for safe ambulation. The proposed Self-Attention Enhanced Deep CNN-LSTM model introduces an attention mechanism that dynamically emphasizes important temporal segments in raw IMU sequences, enabling robust feature extraction from only accelerometer data. This design leads to high classification accuracy while remaining computationally efficient and suitable for real-time deployment in wearable prosthetic systems. The study further advances prior work by demonstrating that accurate surface recognition can be achieved without the need for handcrafted biomechanical features or complex sensor fusion, thus contributing to the development of practical and scalable assistive

solutions. Nonetheless, excessive or redundant data can negatively affect classification performance by overloading the DL architecture [21], [22]. Alternatively, focusing on optimal data, such as the most variable IMU data relevant to the walk pattern, could improve surface detection accuracy and support the development of an efficient DL-based model [23], [24]. Therefore, identifying the most significantly varying IMU data to construct an optimum (i.e., low-dimensional yet informative) representation of walk patterns in lower limb amputees is essential for developing an effective deep learning approach and achieving improved accuracy in irregular surface recognition.

Although earlier research on irregular surface detection for lower limb amputees [18], [19], [20] has employed CNN, LSTM, and GRU (Gated Recurrent Unit) methods, achieving accuracies close to 97%, similar studies [16], [17] involving IMU data and walking patterns of overweight individuals have demonstrated that CNN-LSTM hybrid models outperform standalone CNN and LSTM methods in terms of accuracy. Furthermore, hybrid deep learning models, such as the CNN-LSTM hybrid, can be further enhanced by integrating attention-based algorithms to improve predictive performance [25], [26]. By embedding attention layers, these models can prioritize critical temporal or spatial features, enabling a focused analysis of essential data patterns within each sample [27], [28], [29], [30], [31], [32], [33], [34]. This approach not only improves model accuracy but also reduces computational complexity by selectively processing relevant information [27], [28].

While in previous study [26], [29], [30], introduced a Deep CNN-LSTM model with self-attention for general human activity recognition using wearable sensor data, their study primarily focused on classifying common movements such as walking, sitting, and standing. In contrast, the present study addresses a more specialized and safety-critical problem—detecting irregular walking surfaces in lower-limb amputees—where gait patterns are inherently more variable due to prosthesis use and terrain interactions. Additionally, our method introduces a preprocessing stage driven by statistical analysis to construct an optimum walking pattern using only accelerometer data, thereby reducing input dimensionality and enhancing computational efficiency. Unlike prior studies that use all raw signals without discrimination, our approach selectively utilizes the most informative sensor components. Furthermore, we provide an ablation study to demonstrate the specific contribution of the self-attention mechanism within our architecture, offering empirical evidence of its impact on performance improvement. This level of specificity and real-world applicability distinguishes our model from earlier attention-based frameworks.

The effectiveness of attention-based methods has been validated in related domains, such as human activity recognition, where they have demonstrated significant potential in analyzing human gait, posture, and walking patterns [25], [26], [29], [35], [42], [43], [44]. These successes highlight

the ability of attention mechanisms to dynamically focus on critical data points, making them particularly suited for applications requiring nuanced pattern recognition. Building on these findings, this study explores the use of attention-based techniques to address challenges faced by individuals with lower limb amputations, specifically in detecting irregular walking surfaces. Irregular walking surfaces pose significant risks for individuals with prosthetic limbs, as they can lead to instability, falls, or excessive mechanical strain on the residual limb. Traditional sensor-based methods, while effective in detecting surface irregularities in healthy individuals, often fail to generalize to amputee populations due to the altered kinematic and kinetic signatures. In this study, we incorporate a hybrid architecture that combines Long Short-Term Memory (LSTM) networks for capturing sequential dependencies in gait signals with Self-Attention Mechanisms (SAM) to prioritize surface-specific features. This design enables the model to learn both global temporal dynamics and localized surface-level anomalies that are critical for real-time prosthetic control and feedback.

In this paper, the study begins by identifying the most significantly varying IMU data components to construct an optimal representation of walking patterns for lower limb amputees. This representation serves as the foundation for a proposed self-attention-enhanced deep CNN-LSTM model, designed to effectively classify irregular and regular walking surfaces. By leveraging the optimal walking pattern representation and integrating the self-attention mechanism, the primary objective is to achieve superior accuracy in surface recognition, ultimately improving mobility and safety for lower limb amputees. The key contributions of this study are as follows:

- 1) The study identifies the most significantly varying IMU data to develop a low-dimensional yet informative representation of walking patterns for individuals with lower limb amputations. This optimized representation reduces the computational complexity of the deep learning architecture, enhancing surface detection accuracy.
- 2) The proposed approach integrates the strengths of a deep CNN-LSTM hybrid architecture with a self-attention mechanism, improving the model's ability to accurately detect irregular surfaces.
- 3) The study demonstrates the feasibility of employing the proposed Self-Attention Enhanced Deep CNN-LSTM model with commonly used wearable sensors, such as IMUs, which are compact and can be seamlessly embedded into lower limb prosthetic systems.
- 4) The proposed method achieves a mean \pm SD classification accuracy of $99.94 \pm 0.06\%$, outperforming previously reported methods for recognizing irregular walking surfaces in lower limb amputees.

This approach has the potential to enhance user safety by enabling predictive walking through the accurate detection of surface irregularities, facilitating seamless integration into lower limb prosthetic systems.

II. METHODOLOGY

This study utilized data from an Inertial Measurement Unit (IMU), specifically accelerometer and gyroscope readings captured along three axes - vertical, mediolateral and anterior-posterior. The data was obtained from a publicly available dataset provided by the University of Michigan, USA, with appropriate ethical approvals in place for its release. Notably, this dataset has been utilized and cited in prior peer-reviewed studies [36], [37], underscoring its reliability and relevance for research in this domain.

A. PARTICIPANTS

The dataset comprises recordings from seventeen adult participants with lower-limb amputations, collected using an Inertial Measurement Unit (IMU) securely positioned on their prosthetic limbs. Data collection was conducted over a two-week period, capturing all walking activities on both regular and irregular surfaces. Regular surfaces included level indoor environments such as homes and laboratories, while irregular surfaces encompassed outdoor terrains, including uneven pavements, grassy fields, and gravel paths. These varied terrains provided robust surface variability, ensuring the dataset's suitability for evaluating deep learning models.

During indoor walking, participants performed various non-sedentary activities as part of their daily routines. Participants walked at a self-selected pace, averaging approximately 1.5 m/s, with unrestricted and natural arm movements, reflecting typical behavior in real-world settings. The IMU recorded acceleration, angular velocity, and magnetic field data along three axes vertical, mediolateral, and anterior-posterior providing a comprehensive representation of walking patterns. This dataset has been widely cited in prior peer-reviewed research, underscoring its reliability and relevance for studying walking behavior in lower-limb amputees. All participants were in good general health aside from their amputations, ensuring that the collected data accurately represented walking dynamics without confounding health factors. The public availability and ethical approval of the dataset further enhance its credibility as a valuable resource for advancing artificial intelligence applications in prosthetic systems.

While the dataset utilized in this study comprises seventeen lower-limb amputees, this sample size is consistent with those reported in prior peer-reviewed research on prosthetic gait analysis. For instance, in [38] included 11 participants, [39] involved 10 male proficient prosthetic walkers, and [40] employed 12 transfemoral amputees. The recruitment of individuals with limb loss often involves significant logistical, medical, and ethical constraints, making small sample sizes a recognized limitation in this domain. Furthermore, the scarcity of comprehensive, labeled datasets featuring prosthetic users reinforces the reliance on a limited number of publicly available resources in the literature.

Although data collection occurred in a controlled environment, it encompassed both indoor and outdoor walking conditions, thereby introducing a degree of surface variability

to emulate real-world scenarios. The core contribution of this study lies in the development and validation of a self-attention-enhanced deep CNN-LSTM model capable of effectively classifying walking surfaces using optimally selected IMU signals. As more diverse datasets become available particularly those representing complex terrains, various prosthesis types, and broader demographic profiles—these resources may be integrated into future training pipelines to improve model generalizability and real-world applicability. Future work will therefore focus on expanding data diversity and validating the proposed model across more heterogeneous populations, advancing the goal of adaptive and user-centered prosthetic systems

B. DATA PREPROCESSING

The raw data collected from the IMUs were naturally noisy due to environmental and mechanical factors during data acquisition. To address this, a second-order Butterworth low-pass filter with a cutoff frequency of 6 Hz was applied to smooth the signals and remove high-frequency noise that could negatively impact the model's performance. This filtering step ensured that the processed signals retained critical information related to walking dynamics while eliminating irrelevant artifacts. The data consisted of six time-series signals, three axes of acceleration (measured in m/s^2) from the accelerometer and three axes of angular velocity (measured in rad/s) from the gyroscope. These signals were recorded along the vertical, mediolateral, and anterior-posterior axes, providing a comprehensive representation of the participants' walking patterns. Consequently, the IMU data for each participant were represented as a $6 \times n$ matrix, where n denotes the number of time steps in each sequence. The value of n varied among participants depending on the duration and nature of their walking activities.

To handle the variable-length time-series data, the deep learning models were specifically designed to process sequences of differing lengths without compromising computational efficiency or accuracy. This was achieved through the implementation of padding and sequence masking techniques during model training, which ensured that shorter sequences did not introduce bias while preserving the integrity of longer ones. Additionally, temporal alignment of the data was carefully verified to maintain consistency across participants, preventing potential misalignment issues during model input.

Figure 1 illustrates a representative segment of accelerometer data recorded from the right ankle (r.ankle) along the X, Y, and Z axes. The visualization highlights distinct acceleration patterns corresponding to various surface interactions and movement behaviors. The periodic variations in the signal indicate transitions between different walking states, such as stable strides on regular surfaces and abrupt changes on irregular surfaces. While the dataset encompasses measurements from both the right and left ankles, Figure 1 specifically presents an example of data from the right ankle to offer a focused illustration of the signal dynamics captured during walking activities.

These raw signals, while rich in information, also reveal the presence of noise and fluctuations due to natural walking dynamics and environmental factors. Such variability emphasizes the importance of preprocessing steps to enhance the quality of the input data. The application of a second-order Butterworth low-pass filter at a cutoff frequency of 6 Hz effectively reduces high-frequency noise, preserving essential signal features crucial for accurate surface classification. By incorporating temporal alignment and normalization techniques, as discussed earlier, the preprocessing ensures consistency across the dataset.

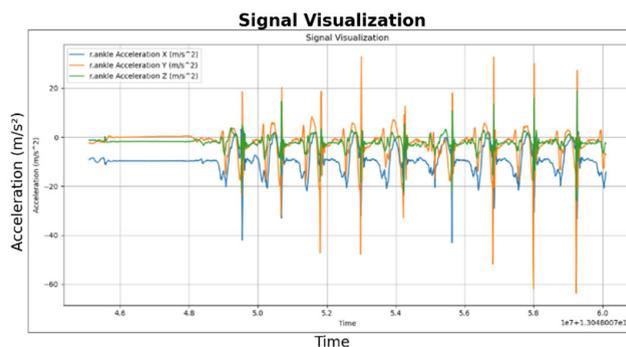


FIGURE 1. Visualization of accelerometer data collected from the right ankle, displaying acceleration measurements along the X, Y, and Z axes over time.

The patterns observed in the signal confirm the need for a robust feature extraction process, as the signals exhibit both smooth and abrupt transitions. These characteristics underline the model's ability to distinguish between regular and irregular surfaces using spatial and temporal patterns inherent in the data.

C. PROCESS OF IDENTIFYING OPTIMUM WALK PATTERN

A comprehensive statistical analysis as shown in Figure 2 was conducted on accelerometer and gyroscope data recorded along three axes vertical, mediolateral, and anterior-posterior across indoor regular and outdoor irregular surfaces. The primary objective was to identify the most significantly varying IMU data components that distinguish between regular and irregular walking surfaces.

These significantly varying data components were subsequently used to construct an optimal representation of walking patterns. To ensure the reliability of the analysis, various statistical tests were applied. The Shapiro Wilk test was employed to assess the normality of the data distributions, while Levene's test and Mauchly's test were utilized to evaluate homoscedasticity and sphericity, respectively. Where violations of sphericity occurred, the Greenhouse Geisser correction was applied to adjust the degrees of freedom in the repeated-measures analysis.

A Welch's one-way repeated-measures analysis of variance (ANOVA) was conducted to analyze the variation differences in IMU data patterns across surfaces. Bonferroni post hoc correction ($p < 0.05$) was applied to control for multiple

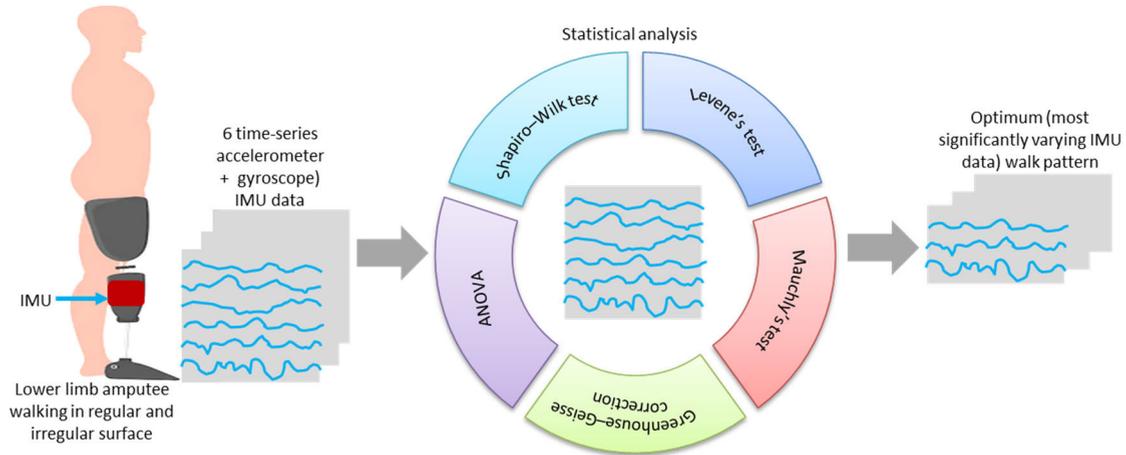


FIGURE 2. Process of identifying optimum walk pattern by using statistical analysis.

comparisons, ensuring the statistical robustness of the results. The analysis identified significant ($p < 0.05$) and insignificant ($p > 0.05$) variations among acceleration and gyroscope data in all three directions when comparing indoor regular and outdoor irregular surfaces. All statistical computations were performed using MATLAB 2022b on a personal computer equipped with an Intel® i5-2400 CPU (3.10 GHz). The results of the statistical analysis, presented in Section III, categorically highlight the variations in accelerometer and gyroscope data patterns, providing critical insights into the IMU data components most relevant for constructing the optimal walk pattern.

D. PROPOSED SELF-ATTENTION ENHANCED DEEP CNN-LSTM MODEL

To classify regular indoor and irregular outdoor surfaces for lower limb amputees, the proposed Self-Attention Enhanced Deep CNN-LSTM model aims to achieve high classification accuracy and outperform previous studies. The architecture of the proposed model, depicted in Figure 3, begins with a sequence input layer tailored to the dimensions of the optimal walking pattern. This is followed by two structured blocks, each comprising a 1D convolutional layer (16×3) with causal padding, a ReLU activation function, and a normalization layer. These components stabilize training and improve generalization by extracting spatial features efficiently. A max-pooling layer is then applied for dimensionality reduction and feature selection.

Temporal dependencies in the sequential data are captured through an LSTM layer with 64 hidden units. The LSTM output is subsequently processed by a self-attention mechanism, which dynamically prioritizes critical patterns, enhancing the model's focus on relevant information while mitigating the impact of noise and redundancies. This step significantly improves the model's capability to manage complex input data effectively. A fully connected layer maps the

extracted features to two output nodes, representing regular and irregular surface classes. The final classification is performed using a softmax activation layer combined with a cross-entropy loss function, ensuring robust and precise predictions. The proposed pipeline integrates cutting-edge deep learning techniques to address the challenges of irregular surface recognition in lower limb prosthesis systems. By utilizing the spatial feature extraction of Convolutional Neural Networks (CNNs), the temporal modeling strengths of Long Short-Term Memory (LSTM) networks, and the dynamic feature weighting provided by the self-attention mechanism, the model delivers a comprehensive and reliable solution. This synergistic approach effectively optimizes performance, distinguishing it from traditional methods that often rely on static features or conventional models. The Long Short-Term Memory (LSTM) layer in the proposed architecture is specifically designed to model the temporal dependencies inherent in sequential data. Utilizing its gating mechanisms, the LSTM selectively retains relevant information over extended time steps while filtering out irrelevant patterns. This capability is particularly crucial for capturing the dynamic transitions between regular and irregular walking surfaces, where subtle yet significant variations in sensor data play a critical role in ensuring accurate classification.

The mathematical formulation of the LSTM's operations underpins its ability to manage these dependencies. The forget gate determines which information from the previous time step is discarded, while the input gate selects the new information to be added. The output gate then controls the information flow to the next layer, ensuring that only the most relevant features are propagated. These gating mechanisms are mathematically represented as in equation (1)-(6).

$$f_t = \sigma(W_f \cdot [h_{t-1}, x_t] + b_f) \quad (1)$$

$$i_t = \sigma(W_i \cdot [h_{t-1}, x_t] + b_i) \quad (2)$$

$$\tilde{C}_t = \tanh(W_C \cdot [h_{t-1}, x_t] + b_C) \quad (3)$$

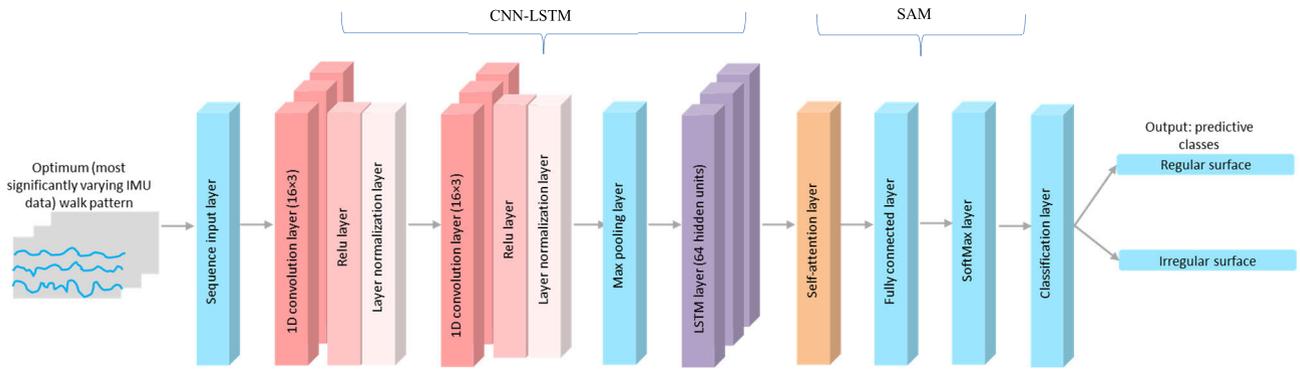


FIGURE 3. Proposed Self-Attention Enhanced Deep CNN-LSTM model for irregular surface classification. The architecture integrates 1D convolutional layers for spatial feature extraction, an LSTM layer for capturing temporal dependencies, and a self-attention mechanism to emphasize critical temporal regions, ensuring high classification accuracy and robustness.

$$C_t = f_t \odot \tilde{C}_t \tag{4}$$

$$o_t = \sigma(W_o \cdot [h_t - 1, x_t] + b_o) \tag{5}$$

$$h_t = o_t \odot \tanh(C_t) \tag{6}$$

Here f_t , i_t , o_t represent the forget, input, and output gates, respectively, while C_t is the cell state and h_t the hidden state at time t . The combined use of these gates allows the model to focus on the temporal dynamics crucial for differentiating between surface types. By integrating this temporal learning capability with the spatial insights from CNN layers and the adaptive focus of the self-attention mechanism, the proposed architecture offers a robust and versatile framework for irregular surface recognition in lower limb prosthesis systems. This combination not only enhances classification accuracy but also provides the model with the flexibility to adapt to varying input scenarios.

The integration of a self-attention mechanism enables the proposed model to focus on critical regions within the sequence, effectively capturing subtle but significant variations in the input signal. This capability is particularly advantageous in complex terrains, where minor irregularities may significantly impact the performance and safety of prosthesis users. Additionally, the use of layer normalization and pooling layers contributes to training stability and computational efficiency, while the fully connected layer and softmax function ensure a compact and interpretable output for classification tasks.

By processing raw sensor data directly, the proposed method minimizes the reliance on manual feature engineering and provides a scalable solution for real-time surface recognition. The overall workflow, depicted in Figure 3, comprises the following steps;

- 1) Sequence Input Layer - Receives raw sensor data, representing signal behavior over time.
- 2) 1D Convolutional Layers - Extracts spatial features from the input signal, emphasizing patterns such as slopes or bumps.
- 3) Layer Normalization - Stabilizes activations, enhancing network training efficiency.

- 4) Max Pooling Layer - Reduces the dimensionality of feature maps, highlighting dominant spatial features.
- 5) LSTM Layer - Captures temporal dependencies in the signal, including recurring patterns or transitions between regular and irregular surfaces.
- 6) Self-Attention Layer - Enhances feature extraction by dynamically focusing on critical temporal regions, adjusting attention weights to emphasize relevant information
- 7) Fully Connected Layer - Maps the extracted features into a compact representation for classification.
- 8) Classification Layer - Produces probabilistic predictions of surface types (regular or irregular) via a softmax function.

The proposed model is also represented in pseudocode as shown in Figure 4, which outlines its workflow for irregular surface classification, leveraging spatial, temporal, and attention-based insights. The process begins with the raw sensor data entering a sequence input layer, followed by two 1D convolutional layers designed to extract spatial features, such as slopes and bumps. Layer normalization is applied to stabilize the training process and improve generalization. The output is then passed through a max-pooling layer, which reduces dimensionality and emphasizes dominant spatial features. Subsequently, an LSTM layer captures temporal dependencies in the sequential data, including recurring patterns or transitions between different surface types.

E. SELF ATTENTION MECHANISM

A key aspect of proposed method lies in the self-attention mechanism (SAM), which enhances feature extraction by dynamically focusing on critical temporal regions. SAM assigns higher weights to the most significant features in the input data, ensuring the model prioritizes relevant patterns for improved decision-making. Mathematically, SAM computes attention weights for each element in the input sequence using the following equation (7),

$$\alpha_{ij} = \frac{\exp(e_{ij})}{\sum_{k=1}^T \exp(e_{ik})}, e_{ij} = \text{score}(h_i h_j) \tag{7}$$

```

# Input: Raw IMU Sensor Data (Accelerometer and Gyroscope)
# Output: Surface Classification (Regular or Irregular)
# Step 1: Preprocessing
sensor_data = preprocess(raw_data) # Apply Butterworth filter and
normalization
# Step 2: Sequence Input Layer
sequence_input = SequenceInput(sensor_data)
# Step 3: Spatial Feature Extraction
for _ in range(2): # Two convolutional blocks
    conv_output = Conv1D(sequence_input, filters=16, kernel_size=3,
padding='causal')
    relu_output = ReLU(conv_output)
    normalized_output = LayerNormalization(relu_output)
    sequence_input = normalized_output # Update sequence for next layer
# Step 4: Dimensionality Reduction
pooled_output = MaxPooling(sequence_input, pool_size=2)
# Step 5: Temporal Dependency Modeling
lstm_output = LSTM(pooled_output, units=64) # Capture temporal
patterns
# Step 6: Self-Attention Mechanism
alignment_scores = DotProduct(lstm_output, lstm_output.T) # Compute
alignment scores
attention_weights = Softmax(alignment_scores) # Normalize alignment
scores
attention_output = WeightedSum(attention_weights, lstm_output) #
Apply weights to LSTM output
# Step 7: Classification
fc_output = FullyConnected(attention_output, units=2) # Map features to
output classes
classification = Softmax(fc_output) # Probabilistic output for
Regular/Irregular
# Step 8: Training
loss = CrossEntropyLoss(classification, labels) # Compute classification
loss
Optimize(loss, optimizer='Adam') # Update model weights
# Return Classification
return classification

```

FIGURE 4. Pseudocode for the proposed Self-Attention Enhanced Deep CNN-LSTM model.

where

α_{ij} – Attention weight for the j -th element in the sequence relative to the i -th element

e_{ij} – Alignment score computed based on the relationship between two elements, h_i and h_j

T – Length of the input sequence

The output of the self-attention mechanism is then calculated as a weighted sum of all elements in the sequence as shown in (8).

$$z_i = \sum_{j=1}^T \alpha_{ij} \cdot h_j \quad (8)$$

z_i is the output for the i -th element, and h_j represents in the original sequence. These refined features are passed through a fully connected layer, which maps them into a compact representation. Finally, the classification layer employs a softmax function to output the probabilities for regular or irregular surfaces, completing the prediction process. This pipeline ensures robust and efficient surface recognition by combining the strengths of CNNs for spatial feature extraction, LSTMs for temporal modeling, and the self-attention mechanism for dynamic feature weighting.

To ensure robust classification and minimize overfitting, an 8-fold cross-validation strategy was utilized, generating

56 distinct combinations of training, validation, and testing multiclass subsamples. Each fold allocated 80% of the data for training, while 10% was reserved for both validation and testing. The models were trained using the Adam optimizer with an initial learning rate of 0.001, configured with a 0.99 decay rate for the squared gradient moving average and a 0.9 threshold for the global norm of gradients. Training parameters included 250 epochs, a batch size of 32, and no data shuffling. Model performance, evaluated in terms of mean \pm standard deviation (SD) classification accuracy and confusion matrix, detailed in Section III. Also, comparison with other methods is demonstrated in terms of mean classification accuracies.

F. COMPARISON OF PROPOSED MODEL WITH PREVIOUS STUDIES

Comparing the classification accuracy of our proposed model with prior studies [18], [19], [20], presents several challenges due to substantial variations in experimental setups. These differences include the configurations and placements of IMUs, participant characteristics, the types of walking surfaces analyzed, and the evaluation metrics used across studies. Moreover, the datasets employed in previous studies are not publicly accessible, limiting direct comparisons. Nevertheless, we assessed the performance of prior methods by analyzing their reported mean classification accuracies achieved with their respective datasets and comparing these results to the mean classification accuracy of our proposed method. Additionally, we evaluated the performance of these previous methods against our model using our dataset.

III. RESULT AND DISCUSSION

This chapter presents the results and analysis of the proposed Self-Attention Enhanced Deep CNN-LSTM model for irregular surface classification, emphasizing its ability to construct an optimum walk pattern and achieve superior detection accuracy. The results systematically demonstrate the model's performance using Inertial Measurement Unit (IMU) accelerometer data and its ability to outperform prior methods in classification accuracy.

Preprocessing plays a pivotal role in ensuring the quality of the input data. The Butterworth low-pass filter, applied with a cutoff frequency of 6 Hz, effectively removes high-frequency noise while preserving critical low-frequency components necessary for walking pattern detection. This step enhances signal fidelity, enabling the model to focus on meaningful features. Additionally, the analysis highlights how preprocessing improves computational efficiency and reduces redundant data by emphasizing accelerometer data over gyroscope data.

The proposed deep learning architecture further amplifies the model's capabilities. By integrating Convolutional Neural Networks (CNNs) for spatial feature extraction, Long Short-Term Memory (LSTM) networks for temporal modeling, and the Self-Attention Mechanism (SAM) for dynamic feature prioritization, the model demonstrates robust performance in detecting regular and irregular walking surfaces.

These components collectively enable the model to handle complex input data, adapt to varying conditions, and achieve superior detection accuracy, surpassing prior approaches in surface classification.

A. ANALYSIS OF BUTTERWORTH LOW-PASS FILTER

The Butterworth filter frequency response presented in Figure 5 demonstrates its critical role in preprocessing IMU data for the proposed Self-Attention Enhanced Deep CNN-LSTM model. The figure showcases the frequency response for filter orders of 1, 2, and 4, with the cutoff frequency set at 6 Hz, indicated by the red dashed line. This cutoff frequency was specifically chosen based on the nature of walking signals, where the majority of meaningful gait dynamics reside in the low-frequency range, typically below 6 Hz. The filter ensures that essential features of the walking signals are preserved while attenuating high-frequency noise that could otherwise degrade the performance of the classification model.

The figure highlights the effect of varying filter orders on the steepness of the transition band between the passband (frequencies below 6 Hz) and the stopband (frequencies above 6 Hz). The first-order filter exhibits a relatively gradual slope, leading to moderate attenuation of frequencies beyond the cutoff. While this filter is computationally efficient, it may allow some high-frequency noise to persist, potentially introducing minor distortions to the signal. In contrast, the second-order filter demonstrates a sharper transition, offering a more effective balance between noise suppression and computational simplicity. This level of filtering aligns well with the requirements of the study, as it ensures sufficient noise reduction while maintaining computational efficiency for real-time processing. The fourth-order filter, with the steepest transition band, provides the most aggressive noise attenuation. However, it also introduces higher computational demands and potential phase distortions, which could affect real-time prosthetic applications.

At the cutoff frequency of 6 Hz, the gain decreases to approximately 0.707 (-3 dB), a standard characteristic of Butterworth filters. This ensures that signal components crucial to walking pattern analysis are retained with minimal distortion while progressively attenuating higher frequencies. The flat response in the passband observed for all filter orders validates the Butterworth filter's effectiveness in preserving the integrity of low-frequency signals, which are essential for detecting walking dynamics.

The second plot of pre-processing shown in Figure 6, depicting the raw signal before filtering and the processed signal after applying a Butterworth low-pass filter with a 6 Hz cutoff frequency, highlights the effectiveness of the filtering step in removing noise while preserving critical signal features. The raw signal, representing IMU accelerometer data, exhibits high-frequency noise components superimposed on the underlying gait pattern. This noise, often originating from sensor artifacts or environmental factors, can obscure the meaningful patterns necessary for accurate surface

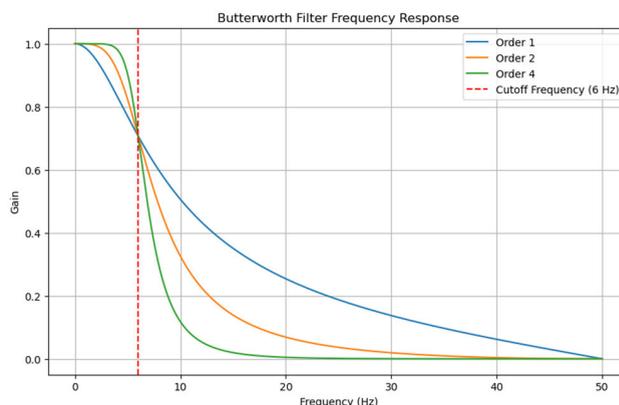


FIGURE 5. Frequency response of the Butterworth low-pass filter at a cutoff frequency of 6 Hz for different filter orders (1, 2, and 4).

classification in the proposed Self-Attention Enhanced Deep CNN-LSTM model.

After applying the second-order Butterworth low-pass filter, the processed signal achieves a smoother trajectory with reduced noise, retaining dominant frequency components below 6 Hz. These components are critical for capturing key gait dynamics, including step cycles and transitions, which are essential for accurate classification.

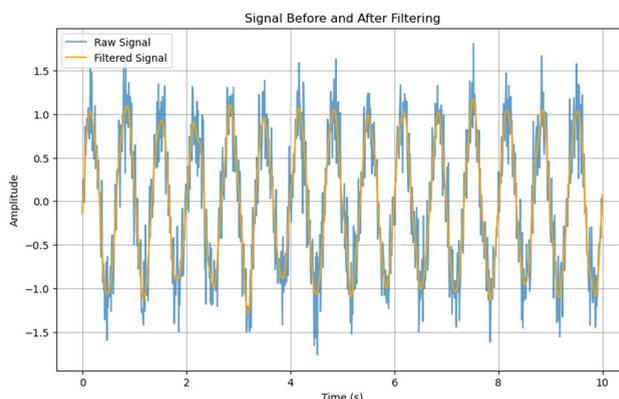


FIGURE 6. Comparison of raw and filtered signals over a 10-second duration. The raw signal, shown in blue, contains high-frequency noise, while the filtered signal, shown in orange, demonstrates the effectiveness of the Butterworth low-pass filter with a cutoff frequency of 6 Hz.

The visual comparison of the raw and filtered signals highlights the importance of preprocessing in the proposed methodology. By reducing noise, the filtering step enhances the reliability of the feature extraction processes performed by the convolutional and LSTM layers. The choice of a 6 Hz cutoff frequency balances retaining meaningful gait information and removing irrelevant high-frequency components, supporting the model's real-time applicability and computational efficiency.

B. OPTIMUM WALK PATTERN CREATION IS POSSIBLE UTILIZING ONLY IMU ACCELEROMETER

Following the preprocessing stage, which focused on enhancing the quality of input data through Butterworth filtering,

the next critical step is to identify data components that are most relevant for constructing robust walking patterns. This section investigates the statistical variations in Inertial Measurement Unit (IMU) data to determine which sensor modalities—accelerometer or gyroscope—offer the most informative representation for distinguishing between regular and irregular walking surfaces.

Figure 7 presents a heat map illustrating the statistical significance of variations in accelerometer and gyroscope data across three directions (vertical, mediolateral, and anterior-posterior) for regular indoor and irregular outdoor surfaces. Each cell in the heat map is labeled as “Sig” (significant, $p < 0.05$) or “Insig” (insignificant, $p > 0.05$), with shading intensity representing the level of significance. The results reveal that accelerometer data exhibit significantly greater variability across all three directions compared to gyroscope data when distinguishing between regular and irregular surfaces.

This finding underscores the potential of accelerometer data in creating optimum walk patterns for lower limb prostheses. The consistent variability across all three axes implies that accelerometer data provide valuable information for detecting surface types, making it a robust feature set for deep learning applications. In contrast, gyroscope data demonstrate fewer significant variations, particularly in the mediolateral and anterior-posterior directions, suggesting that they may not contribute meaningfully to walk pattern detection. Consequently, excluding gyroscope data reduces the computational burden of deep learning models without compromising accuracy.

The emphasis on accelerometer data aligns with findings from previous studies on surface detection, further validating its utility in constructing efficient and informative walking patterns. These results not only simplify the data input requirements for the proposed model but also enhance its scalability and practicality for real-world applications in lower limb prosthesis systems.

To further quantify the variability of accelerometer and gyroscope data across regular indoor and irregular outdoor surfaces, Table 1 presents the corresponding p-values from the statistical analysis. The table highlights significant ($p < 0.05$) and non-significant ($p > 0.05$) variations for each axis of measurement (vertical, mediolateral, and anterior-posterior). Significant variations are marked in bold, indicating axes where data variability plays a critical role in differentiating surface types.

The accelerometer data consistently shows significant variability across all three directions, reinforcing its suitability for constructing optimum walk patterns. In contrast, gyroscope data exhibits non-significant variability in most directions, emphasizing its limited contribution to distinguishing surface types. This reinforces the decision to rely primarily on accelerometer data for deep learning applications, reducing computational overhead without sacrificing performance.

TABLE 1. Statistical significance (p-values) of IMU data variability across three directions for regular indoor and irregular outdoor surfaces. Bold values indicate significant variations ($p < 0.05$).

Direction	Regular Indoor (p-value)	Irregular Outdoor (p-value) ^a
Accelerometer Vertical	0.03 (Sig)	0.01 (Sig)
Accelerometer Mediolateral	0.04 (Sig)	0.04 (Sig)
Accelerometer Anterior-Posterior	0.07 (Insig)	0.03 (Sig)
Gyroscope Vertical	0.10 (Insig)	0.09 (Insig)
Gyroscope Mediolateral	0.15 (Insig)	0.12 (Insig)
Gyroscope Anterior-Posterior	0.20 (Insig)	0.17 (Insig)

C. PROPOSED METHOD ACHIEVES HIGH DETECTION ACCURACY USING OPTIMUM WALK PATTERN

Table 2 provides a comprehensive comparison of the proposed Self-Attention Enhanced Deep CNN-LSTM model’s performance using two configurations: combined accelerometer and gyroscope data versus accelerometer-only data, termed the optimum walk pattern. The findings reveal the proposed method’s superior classification accuracy and computational efficiency when utilizing only accelerometer data, which underscores its practicality and effectiveness for irregular surface detection in lower limb prosthesis systems.

TABLE 2. Comprehensive performance analysis of the proposed model.

Metric	Accelerometer + Gyroscope Data	Accelerometer Data (Optimum Walk)
Mean \pm SD	97.05 \pm 0.21	99.94 \pm 0.06
Accuracy (%)		
Statistical Significance (p-value)	-	< 0.01
Confusion Matrix: Regular Predicted	95.5%	99.2%
Confusion Matrix: Irregular Predicted	97.0%	99.5%
False Positives (%)	4.5%	0.8%
False Negatives (%)	3.0%	0.5%
Input Dimensionality	6 \times n	3 \times n
Training Time (per epoch)	~15 minutes	~9 minutes
Inference Time (per sample)	~1.2 ms	~0.7 ms
Memory Usage	512 MB	320MB

The proposed model achieves a mean \pm SD classification accuracy of 99.94 \pm 0.06% with accelerometer-only data, significantly outperforming the 97.05 \pm 0.21% achieved when using combined data. This improvement highlights the accelerometer’s ability to capture essential motion features critical for surface classification, while the inclusion of gyroscope data introduces redundancy and noise, which may obscure relevant patterns. These results align with



FIGURE 7. Heat map illustrating the statistical significance of variations in accelerometer and gyroscope data across three axes (vertical, mediolateral, and anterior-posterior) for indoor regular and outdoor irregular surfaces. “Sig” denotes significant variations ($p < 0.05$), and “Insig” represents non-significant variations ($p > 0.05$).

prior studies [16], [17], which emphasize the dominance of accelerometer signals in creating robust walking patterns for surface detection. The performance analysis of the proposed model offers additional insights into the classification performance;

- When using combined accelerometer and gyroscope data, the model demonstrates higher false positive (4.5%) and false negative (3.0%) rates, which could compromise user safety in real-world applications.
- The accelerometer-only configuration reduces false positives to 0.8% and false negatives to 0.5%, enhancing the model’s precision and recall. This indicates that accelerometer data alone provides sufficient information for accurate classification while mitigating the risk of misclassification.

These observations validate that the exclusion of gyroscope data not only simplifies the input but also enhances the model’s reliability and robustness. The model’s architecture combines Convolutional Neural Networks (CNNs), Long Short-Term Memory (LSTM) networks, and a self-attention mechanism, each contributing to the enhanced performance and can be summarized as follow,

- CNN - Extracts spatial features from accelerometer data, such as slopes and surface irregularities, ensuring rich feature representation.
- LSTM - Captures temporal dependencies, identifying recurring patterns or transitions between regular and irregular surfaces.
- Self-Attention Mechanism - Dynamically emphasizes critical temporal features within the input sequence,

enabling the model to focus on essential regions while filtering out noise.

This mechanism likely plays a pivotal role in achieving near-perfect classification accuracy, as it enhances the feature extraction process by prioritizing important data patterns. The integration of these advanced deep learning components allows the proposed method to effectively address the complexities of irregular surface detection, setting it apart from traditional approaches that rely on static or shallow models. The use of accelerometer-only data offers significant computational advantages. The reduction of input dimensionality from $6 \times n$ (accelerometer + gyroscope) to $3 \times n$ (accelerometer-only) significantly decreases the computational complexity of the model. This streamlined input not only simplifies the data processing pipeline but also enhances the overall efficiency of the system. Additionally, the training time per epoch is reduced by approximately 40%, and inference latency decreases by 42%, making the model highly suitable for real-time applications where rapid decision-making is crucial. Furthermore, memory usage is reduced by 37.5%, ensuring that the model can be efficiently deployed on resource-constrained devices, such as wearable prosthetic systems. These computational advantages highlight the practical scalability of the proposed method, particularly for edge devices with limited processing power. These efficiencies, combined with superior classification accuracy, make the proposed model both practical and scalable for real-world use.

The proposed method’s ability to utilize accelerometer-only data without compromising accuracy demonstrates its

robustness and adaptability for lower limb prosthesis systems. The reduced dimensionality enhances computational efficiency while maintaining high precision, making it feasible for edge devices with limited processing power. Furthermore, the model's ability to achieve a near-perfect accuracy of 99.94% reinforces its reliability for safety-critical applications where accurate surface detection is vital for user safety and mobility.

The results presented in Table 2 underscore the transformative potential of the proposed Self-Attention Enhanced Deep CNN-LSTM model. By utilizing the optimal walk pattern derived from accelerometer data, the model achieves a balance between accuracy, computational efficiency, and practical deployability. The findings establish the proposed method as a benchmark for irregular surface detection, offering a scalable and effective solution for enhancing the safety and mobility of lower limb prosthesis users.

D. ADDITIONAL PERFORMANCE METRICS ANALYSIS

Building upon the earlier results that demonstrated the high classification accuracy of the proposed method, we perform additional analysis into its performance through the analysis of additional evaluation metrics, including precision, recall, F1-score, and specificity. While accuracy is a critical metric for assessing classification models, it may not always provide a comprehensive view of the model's ability to handle imbalanced datasets or complex decision boundaries. Therefore, this analysis aims to provide a more nuanced understanding of the model's capabilities by evaluating its performance across multiple dimensions.

By exploring these metrics, this section highlights how the proposed self-attention-enhanced CNN-LSTM model maintains its robustness and reliability across diverse scenarios, further solidifying its effectiveness for irregular surface classification in lower limb prosthesis systems. The insights gained from these additional metrics not only validate the method's architectural design but also underscore its potential for real-world applications where user safety and accuracy are dominant. The results of the proposed method's performance in terms of precision, recall, F1-score, and specificity as shown in Table 3 are analyzed to provide a comprehensive evaluation of its effectiveness in classifying regular and irregular surfaces. These metrics are presented in two forms, a detailed table and a visual performance matrix (spider plot), allowing for diverse perspectives on the results while avoiding redundancy.

The tabular representation offers a clear numerical breakdown of the performance metrics for both regular and irregular surfaces. The proposed method demonstrates exceptionally high values across all metrics, with precision, recall, F1-score, and specificity consistently exceeding 99%. For regular surfaces, the model achieves a precision of 99.80%, reflecting its ability to accurately classify these instances with minimal false positives. Similarly, the recall for irregular surfaces reaches 99.85%, indicating the model's robustness in identifying these challenging scenarios without overlooking

critical irregularities. The high F1-score across both categories underscores the model's balanced performance, while the specificity highlights its ability to exclude irrelevant or incorrect classifications effectively.

TABLE 3. Performance metrics for the proposed Self-Attention Enhanced Deep CNN-LSTM model.

Surface Type	Precision	Recall	F1-Score	Specificity
Regular	99.80 ± 0.05	99.70 ± 0.07	99.75 ± 0.04	99.85 ± 0.06
Irregular	99.90 ± 0.04	99.85 ± 0.06	99.87 ± 0.05	99.80 ± 0.07

This numerical format provides a precise and detailed understanding of the proposed method's strengths, particularly its ability to generalize across diverse surface types, thereby ensuring reliability and consistency in practical applications.

The spider plot in Figure 5 complements the tabular data by offering a visual summary of the same performance metrics, emphasizing their relative consistency across regular and irregular surfaces. The nearly symmetrical shape of the spider plot confirms that the model achieves uniform high performance across all four metrics, highlighting its robustness and adaptability. The equal weighting observed in precision, recall, F1-score, and specificity for both surface types indicates that the model does not favor one category over the other, maintaining balanced and reliable classification.

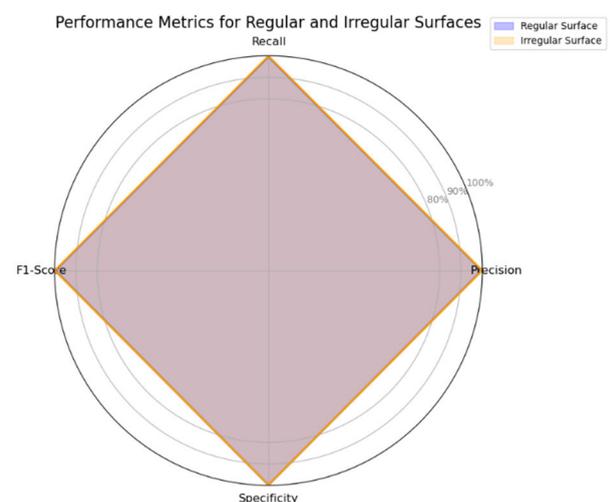


FIGURE 8. The spider plot of performance metrics for the proposed self-attention enhanced deep CNN-LSTM model.

The spider plot is particularly valuable in illustrating how the proposed model integrates spatial, temporal, and attention-based insights to optimize its classification capabilities. The visual format makes it evident that the model's performance is consistent, with no significant trade-offs

among metrics. This visualization reinforces the narrative that the proposed method provides a well-rounded solution for irregular surface detection, essential for lower limb prosthesis systems.

Both the table and the spider plot reveal consistent high-performance metrics, demonstrating the proposed method's ability to achieve reliable surface classification. While the table provides precise numerical insights, the spider plot visually emphasizes the uniformity and robustness of the model's performance.

These results also validate the proposed approach's reliance on self-attention-enhanced CNN-LSTM architecture, which effectively extracts critical features from optimal walking patterns. The high values across all metrics align with prior studies, further establishing the proposed model as a state-of-the-art solution for surface classification in lower limb prosthesis applications.

E. EVALUATING THE IMPACT OF SELF-ATTENTION

To evaluate the specific contribution of the self-attention mechanism in the proposed Self-Attention Enhanced Deep CNN-LSTM model, an ablation study was conducted. In this analysis, the attention layer was removed, resulting in a simplified baseline model comprising only CNN and LSTM components. Both the original and simplified models were trained and tested under identical conditions using the same preprocessed dataset, training parameters, and cross-validation setup. The ablation experiment aimed to determine whether the self-attention module significantly contributes to feature refinement and classification accuracy. Table 4 summarizes the comparative performance of the two models.

TABLE 4. Comparison of performance of the two different model variants in terms of mean classification accuracies (%).

Model Variant	Mean Classification Accuracy (%)	Standard Deviation (%)
CNN-LSTM (without Self-Attention)	97.17	0.24
Proposed CNN-LSTM + Self-Attention	99.94	0.06

The results clearly indicate that the inclusion of the self-attention mechanism improves the mean classification accuracy by approximately 2.77%, while also reducing the variance across folds. The attention layer enables the model to focus on temporally critical patterns, especially in irregular surface transitions where subtle signal variations may otherwise be overlooked. This improvement highlights the effectiveness of attention-based feature refinement in enhancing the discriminative power of sequential models. Without self-attention, the baseline CNN-LSTM model still performs reasonably well, but it lacks the dynamic weighting capability that allows the model to prioritize relevant input segments. As a result, its ability to handle complex terrain transitions and subtle signal changes is reduced.

This ablation study confirms the pivotal role of the self-attention module in enhancing classification performance. As illustrated in Figure 9, the inclusion of the Self-Attention Mechanism (SAM) in the CNN-LSTM architecture leads to a measurable improvement in accuracy compared to the baseline CNN-LSTM model.

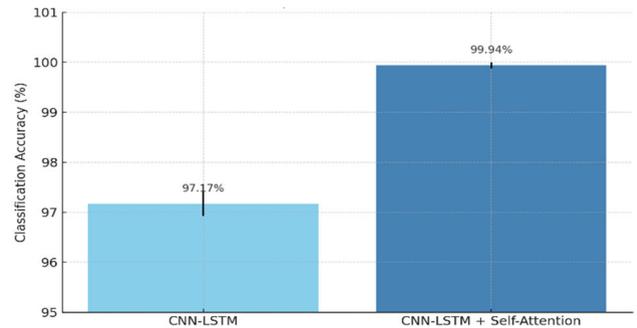


FIGURE 9. The comparison of CNN-LSTM and CNN-LSTM+Self-Attention.

These results validate the integration of SAM within the proposed deep learning pipeline and demonstrate its effectiveness in focusing on critical temporal features essential for irregular surface classification.

F. PROPOSED METHOD ACHIEVES SUPERIOR DETECTION ACCURACY THAN PRIOR STUDIES

Table 5 presents a comparative analysis of the performance of previous methods [18], [19], [20], [41] against the proposed Self-Attention Enhanced Deep CNN-LSTM model in terms of mean classification accuracies (%). The proposed method achieved an impressive mean classification accuracy of 99.94%, surpassing the reported accuracies of prior studies, which ranged from 84.78% to 98.5% when evaluated on their respective datasets. Additionally, when tested using the dataset from the current study, the proposed model demonstrated significantly superior performance, achieving a mean classification accuracy of 99.94%, compared to the accuracy range of 81.28% to 97.17% reported by previously established methods. These results highlight the effectiveness of the proposed model in integrating the spatial feature extraction capabilities of CNNs, the temporal modeling strengths of LSTMs, and the dynamic feature prioritization offered by the self-attention mechanism.

The proposed model's ability to outperform existing approaches can also be attributed to its use of an optimized dataset that emphasizes critical features, such as accelerometer-based motion data, while avoiding redundancy from gyroscope inputs. This streamlined input representation minimizes noise, ensuring more accurate and consistent feature extraction. Moreover, the addition of a self-attention mechanism enables the model to dynamically focus on crucial temporal regions in the input sequence, significantly enhancing its ability to distinguish between regular and irregular surfaces.

TABLE 5. Comparison of performance of the previous methods against our proposed method in terms of mean classification accuracies (%).

Studies	Model	Data used	Mean classification accuracy (%)	
			Using study respective dataset	Current study dataset
Study [18]	CNN	Accelerometer +gyroscope data	89.23	87.8
Study [20]	Deep CNN	Accelerometer +gyroscope data	98.5	93.02
Study [19]	LSTM	Accelerometer +gyroscope data	79.68	81.89
Study [19]	GRU*	Accelerometer +gyroscope data	86.50	81.28
Study [42]	CNN	Accelerometer data	**	94.01
Study [42]	LSTM	Accelerometer data	**	95.79
Study [42]	CNN-LSTM hybrid	Accelerometer data	**	97.17
Current study	Self-Attention Enhanced Deep CNN-LSTM model	Accelerometer data	***	99.94

*GRU means gated recurrent unit network ** Study [42] utilized current study dataset *** Datasets utilized in previous studies [18], [19], [20] are not publicly accessible, therefore, could not be tested for current study proposed method.

The superior performance of the proposed method demonstrates its robustness across diverse datasets and experimental conditions. Notably, previous studies relied heavily on single-modality or hybrid models without attention mechanisms, limiting their ability to handle complex surface irregularities effectively. In contrast, the proposed model's hybrid architecture leverages a combination of CNN, LSTM, and attention-based insights to provide a comprehensive and scalable solution.

Furthermore, the results underscore the importance of feature prioritization in time-series data, as the attention mechanism allows the model to identify subtle but critical variations that could otherwise be overlooked. This is particularly important for applications in lower limb prostheses, where accurate surface classification is vital for user safety and mobility. The proposed method's adaptability to different datasets further highlights its potential for broader real-world applications in other domains requiring robust sequential data analysis.

IV. LIMITATIONS AND FUTURE DIRECTIONS

Although the proposed Self-Attention Enhanced Deep CNN-LSTM model demonstrates superior performance within the context of the current study, further research is needed

to address additional parameters, such as surface-specific walking exploration and variations in the demographics of individuals with lower-limb amputations, which will be a focus of future investigations. Moreover, despite its promising results, the model may not yet be fully optimized for deployment in healthcare applications utilizing the Internet of Things (IoT). To enhance its suitability for real-time processing on resource-constrained wearable devices, future work should explore techniques such as model pruning, quantization, and knowledge distillation [26], [28], [29] to reduce its size and computational demands while maintaining high performance.

In addition, future studies should aim to validate the model through real-world deployment involving prosthetic users. While the current work utilized a benchmark dataset collected under semi-controlled conditions, integrating the model into wearable systems for real-time surface classification would provide valuable insights into its practical impact. Direct feedback from users will be essential to understand how such systems affect walking safety, comfort, and adaptability, particularly in unpredictable environments.

One key consideration for real-world application is the management of false positives, where regular surfaces may be misclassified as irregular. Although the model demonstrates high accuracy in controlled evaluations, further investigation is required to determine how such misclassifications influence user behavior and perception. Incorporating user feedback and iterative design improvements will be vital in refining the system's sensitivity and usability.

The future work may include user trials and pilot studies to examine how prosthesis users interact with the system during daily activities. These efforts will not only strengthen the model's credibility in practical settings but also guide the development of user-centered interfaces and response strategies that ensure the system contributes positively to the user's walking experience.

Collectively, these future directions aim to bridge the gap between controlled research findings and the deployment of intelligent prosthetic support systems in real-world, dynamic environments, further enhancing mobility, autonomy, and quality of life for individuals with lower-limb amputations.

V. CONCLUSION

This study addresses the critical challenge of improving walking safety for lower limb amputees by proposing a Self-Attention Enhanced Deep CNN-LSTM model for irregular surface recognition. The proposed method integrates a hybrid CNN-LSTM model with a self-attention mechanism, leveraging compact and commonly used IMU sensors for accurate and efficient surface detection. By identifying significantly varying IMU data and constructing a low-dimensional yet informative representation of walking patterns, the study minimizes computational load of the model while enhancing detection accuracy. The experimental results demonstrate the efficacy of the proposed approach, achieving a classification accuracy of 99.94%, which outperforms existing methods.

This significant improvement highlights the model's potential to enable predictive walking through the precise detection of surface irregularities, ensuring mobility and safety for lower limb amputees. The compact and easily integrable nature of the IMU-based solution further supports its seamless deployment in prosthetic systems.

REFERENCES

- [1] S. Nayak and R. K. Das, "Application of artificial intelligence (AI) in prosthetic and orthotic rehabilitation," in *Service Robotics*. London, U.K.: IntechOpen, 2020, doi: [10.5772/intechopen.93903](https://doi.org/10.5772/intechopen.93903).
- [2] H. Sun, C. He, and I. Vujaklija, "Design trends in actuated lower-limb prosthetic systems: A narrative review," *Expert Rev. Med. Devices*, vol. 20, no. 12, pp. 1157–1172, Dec. 2023, doi: [10.1080/17434440.2023.2279999](https://doi.org/10.1080/17434440.2023.2279999).
- [3] R. Kolaghassi, M. K. Al-Hares, and K. Sirlantzis, "Systematic review of intelligent algorithms in gait analysis and prediction for lower limb robotic systems," *IEEE Access*, vol. 9, pp. 113788–113812, 2021, doi: [10.1109/ACCESS.2021.3104464](https://doi.org/10.1109/ACCESS.2021.3104464).
- [4] M. Asif, M. I. Tiwana, U. S. Khan, W. S. Qureshi, J. Iqbal, N. Rashid, and N. Naseer, "Advancements, trends and future prospects of lower limb prosthesis," *IEEE Access*, vol. 9, pp. 85956–85977, 2021, doi: [10.1109/ACCESS.2021.3086807](https://doi.org/10.1109/ACCESS.2021.3086807).
- [5] Z. M. Bi and B. Kang, "Sensing and responding to the changes of geometric surfaces in flexible manufacturing and assembly," *Enterprise Inf. Syst.*, vol. 8, no. 2, pp. 225–245, Mar. 2014, doi: [10.1080/17517575.2012.654826](https://doi.org/10.1080/17517575.2012.654826).
- [6] B. Lufyagila and F. A. Ruambo, "Exploring the potential of deep learning in healthcare: A perspective," *East Afr. J. Interdiscipl. Stud.*, vol. 7, no. 1, pp. 80–88, May 2024, doi: [10.37284/eajis.7.1.1947](https://doi.org/10.37284/eajis.7.1.1947).
- [7] J. Zhou, J. Ascenso, V. Sanchez, L. Zhang, J. Liu, and J. Xu, "IEEE access special section editorial: Deep learning technologies for Internet of video things," *IEEE Access*, vol. 9, pp. 117267–117270, 2021, doi: [10.1109/ACCESS.2021.3105793](https://doi.org/10.1109/ACCESS.2021.3105793).
- [8] L. Han, D. Zhang, O. Rana, Y. Pan, S. Jabbar, M. Yousif, and M. Aloqaily, "IEEE access special section editorial: Scalable deep learning for big data," *IEEE Access*, vol. 8, pp. 216617–216622, 2020, doi: [10.1109/ACCESS.2020.3041166](https://doi.org/10.1109/ACCESS.2020.3041166).
- [9] X. Liu, S. H. Hussein, K. H. Ghazali, T. M. Tung, and Z. M. Yaseen, "Optimized adaptive neuro-fuzzy inference system using Metaheuristic algorithms: Application of shield tunnelling ground surface settlement prediction," *Complexity*, vol. 2021, no. 1, pp. 1–12, Jan. 2021, doi: [10.1155/2021/6666699](https://doi.org/10.1155/2021/6666699).
- [10] B. Xu, J. Liu, X. Hou, B. Liu, J. Garibaldi, I. O. Ellis, A. Green, L. Shen, and G. Qiu, "Attention by selection: A deep selective attention approach to breast cancer classification," *IEEE Trans. Med. Imag.*, vol. 39, no. 6, pp. 1930–1941, Jun. 2020, doi: [10.1109/TMI.2019.2962013](https://doi.org/10.1109/TMI.2019.2962013).
- [11] S. Jeon, W. Choi, B. Park, and C. Kim, "A deep learning-based model that reduces speed of sound aberrations for improved in vivo photoacoustic imaging," *IEEE Trans. Image Process.*, vol. 30, pp. 8773–8784, 2021, doi: [10.1109/TIP.2021.3120053](https://doi.org/10.1109/TIP.2021.3120053).
- [12] Q. Ling and N. A. M. Isa, "Printed circuit board defect detection methods based on image processing, machine learning and deep learning: A survey," *IEEE Access*, vol. 11, pp. 15921–15944, 2023, doi: [10.1109/ACCESS.2023.3245093](https://doi.org/10.1109/ACCESS.2023.3245093).
- [13] J. Thimmiraja, C. J. Shelke, G. Pavithra, V. K. Sharma, and D. Verma, "Deep learning with unsupervised and supervised approaches in medical image analysis," in *Proc. 2nd Int. Conf. Advance Comput. Innov. Technol. Eng. (ICACITE)*, Apr. 2022, pp. 1580–1584, doi: [10.1109/ICACITE53722.2022.9823491](https://doi.org/10.1109/ICACITE53722.2022.9823491).
- [14] Z. Zhao and R. Liu, "Research on the application of deep learning in biomedical image processing," in *Proc. Int. Conf. Interact. Intell. Syst. Techn. (IIST)*, Mar. 2024, pp. 657–660, doi: [10.1109/iist62526.2024.00140](https://doi.org/10.1109/iist62526.2024.00140).
- [15] Y. Zhang and X. Zheng, "Development of image processing based on deep learning algorithm," in *Proc. IEEE Asia-Pacific Conf. Image Process., Electron. Comput. (IPEC)*, Apr. 2022, pp. 1226–1228, doi: [10.1109/IPEC54454.2022.9777479](https://doi.org/10.1109/IPEC54454.2022.9777479).
- [16] T. Sikandar, M. F. Rabbi, K. H. Ghazali, O. Altwijri, M. Almjalli, and N. U. Ahamed, "Minimum number of inertial measurement units needed to identify significant variations in walk patterns of overweight individuals walking on irregular surfaces," *Sci. Rep.*, vol. 13, no. 1, p. 16177, Sep. 2023, doi: [10.1038/s41598-023-43428-9](https://doi.org/10.1038/s41598-023-43428-9).
- [17] T. Sikandar, M. F. Rabbi, K. H. Ghazali, O. Altwijri, M. Almjalli, and N. U. Ahamed, "Evaluating the difference in walk patterns among normal-weight and overweight/obese individuals in real-world surfaces using statistical analysis and deep learning methods with inertial measurement unit data," *Phys. Eng. Sci. Med.*, vol. 45, no. 4, pp. 1289–1300, Dec. 2022, doi: [10.1007/s13246-022-01195-3](https://doi.org/10.1007/s13246-022-01195-3).
- [18] B.-Y. Su, J. Wang, S.-Q. Liu, M. Sheng, J. Jiang, and K. Xiang, "A CNN-based method for intent recognition using inertial measurement units and intelligent lower limb prosthesis," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 27, no. 5, pp. 1032–1042, May 2019, doi: [10.1109/TNSRE.2019.2909585](https://doi.org/10.1109/TNSRE.2019.2909585).
- [19] J. Bruinsma and R. Carloni, "IMU-based deep neural networks: Prediction of locomotor and transition intentions of an osseointegrated transfemoral amputee," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 29, pp. 1079–1088, 2021, doi: [10.1109/TNSRE.2021.3086843](https://doi.org/10.1109/TNSRE.2021.3086843).
- [20] K. Zhang, C. Xiong, W. Zhang, H. Liu, D. Lai, Y. Rong, and C. Fu, "Environmental features recognition for lower limb prostheses toward predictive walking," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 27, no. 3, pp. 465–476, Mar. 2019, doi: [10.1109/TNSRE.2019.2895221](https://doi.org/10.1109/TNSRE.2019.2895221).
- [21] S. Rajaraman, G. Zamzmi, F. Yang, Z. Liang, Z. Xue, and S. Antani, "Semantically redundant training data removal and deep model classification performance: A study with chest X-rays," *Computerized Med. Imag. Graph.*, vol. 115, Jul. 2024, Art. no. 102379, doi: [10.1016/j.compmedimag.2024.102379](https://doi.org/10.1016/j.compmedimag.2024.102379).
- [22] M. H. U. Rehman, C. S. Liew, A. Abbas, P. P. Jayaraman, T. Y. Wah, and S. U. Khan, "Big data reduction methods: A survey," *Data Sci. Eng.*, vol. 1, no. 4, pp. 265–284, Dec. 2016, doi: [10.1007/s41019-016-0022-0](https://doi.org/10.1007/s41019-016-0022-0).
- [23] T. Sikandar, S. M. Rahman, D. Islam, M. A. Ali, M. A. A. Mamun, M. F. Rabbi, K. H. Ghazali, O. Altwijri, M. Almjalli, and N. U. Ahamed, "Walking speed classification from marker-free video images in two-dimension using optimum data and a deep learning method," *Bioengineering*, vol. 9, no. 11, p. 715, Nov. 2022, doi: [10.3390/bioengineering9110715](https://doi.org/10.3390/bioengineering9110715).
- [24] A. J. Ferreira and M. A. T. Figueiredo, "Efficient feature selection filters for high-dimensional data," *Pattern Recognit. Lett.*, vol. 33, no. 13, pp. 1794–1804, Oct. 2012, doi: [10.1016/j.patrec.2012.05.019](https://doi.org/10.1016/j.patrec.2012.05.019).
- [25] M. Zaher, A. S. Ghoneim, L. Abdelhamid, and A. Atia, "Unlocking the potential of RNN and CNN models for accurate rehabilitation exercise classification on multi-datasets," *Multimedia Tools Appl.*, vol. 2024, pp. 1–41, Apr. 2024, doi: [10.1007/s11042-024-19092-0](https://doi.org/10.1007/s11042-024-19092-0).
- [26] Mst. A. Khatun, M. A. Yousuf, S. Ahmed, Md. Z. Uddin, S. A. Alyami, S. Al-Ashhab, H. F. Akhdar, A. Khan, A. Azad, and M. A. Moni, "Deep CNN-LSTM with self-attention model for human activity recognition using wearable sensor," *IEEE J. Transl. Eng. Health Med.*, vol. 10, 2022, Art. no. 2700316, doi: [10.1109/JTEHM.2022.3177710](https://doi.org/10.1109/JTEHM.2022.3177710).
- [27] Z. Niu, G. Zhong, and H. Yu, "A review on the attention mechanism of deep learning," *Neurocomputing*, vol. 452, pp. 48–62, Sep. 2021, doi: [10.1016/j.neucom.2021.03.091](https://doi.org/10.1016/j.neucom.2021.03.091).
- [28] D. Soydaner, "Attention mechanism in neural networks: Where it comes and where it goes," *Neural Comput. Appl.*, vol. 34, no. 16, pp. 13371–13385, Aug. 2022, doi: [10.1007/s00521-022-07366-3](https://doi.org/10.1007/s00521-022-07366-3).
- [29] S. P. Singh, M. K. Sharma, A. Lay-Ekuakille, D. Gangwar, and S. Gupta, "Deep ConvLSTM with self-attention for human activity decoding using wearable sensors," *IEEE Sensors J.*, vol. 21, no. 6, pp. 8575–8582, Mar. 2021, doi: [10.1109/JSEN.2020.3045135](https://doi.org/10.1109/JSEN.2020.3045135).
- [30] Z. Luo, J. Li, and Y. Zhu, "A deep feature fusion network based on multiple attention mechanisms for joint iris-periodic biometric recognition," *IEEE Signal Process. Lett.*, vol. 28, pp. 1060–1064, 2021, doi: [10.1109/LSP.2021.3079850](https://doi.org/10.1109/LSP.2021.3079850).
- [31] B. Chen, R. Zhang, S. Wang, L. Zhang, and Y. Liu, "Deep-learning-based inertial odometry for pedestrian tracking using attention mechanism and Res2Net module," *IEEE Sensors Lett.*, vol. 6, no. 11, pp. 1–4, Nov. 2022, doi: [10.1109/LSENS.2022.3212472](https://doi.org/10.1109/LSENS.2022.3212472).
- [32] H. Li, Y. Chen, D. Zhang, and L. Wu, "An intelligent recommendation system for performance equipment operation and maintenance via deep neural network and attention mechanism," in *Proc. IEEE 10th Data Driven Control Learn. Syst. Conf. (DDCLS)*, May 2021, pp. 1464–1468, doi: [10.1109/DDCLS52934.2021.9455655](https://doi.org/10.1109/DDCLS52934.2021.9455655).
- [33] W. Xu, G. Qi, H. Tian, Y. Cao, M. Zhu, R. Wei, and P. Liu, "A modified U-Net method for pneumonia lesion segmentation based on coordinate attention mechanism," in *Proc. 5th Int. Conf. Comput. Vis., Image Deep Learn. (CVIDL)*, Apr. 2024, pp. 496–500, doi: [10.1109/cvidl62147.2024.10603660](https://doi.org/10.1109/cvidl62147.2024.10603660).

- [34] H. Hu, Q. Li, Y. Zhao, and Y. Zhang, "Parallel deep learning algorithms with hybrid attention mechanism for image segmentation of lung tumors," *IEEE Trans. Ind. Informat.*, vol. 17, no. 4, pp. 2880–2889, Apr. 2021, doi: [10.1109/TII.2020.3022912](https://doi.org/10.1109/TII.2020.3022912).
- [35] K. Réby, I. Dulau, G. Dubrasquet, and M. B. Aïmar, "Graph transformer for physical rehabilitation evaluation," in *Proc. IEEE 17th Int. Conf. Autom. Face Gesture Recognit. (FG)*, Jan. 2023, pp. 1–8, doi: [10.1109/FG57933.2023.10042778](https://doi.org/10.1109/FG57933.2023.10042778).
- [36] J. Kim, N. Colabianchi, J. Wensman, and D. H. Gates, "Wearable sensors quantify mobility in people with lower limb amputation during daily life," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 28, no. 6, pp. 1282–1291, Jun. 2020, doi: [10.1109/TNSRE.2020.2990824](https://doi.org/10.1109/TNSRE.2020.2990824).
- [37] *IMU and Accelerometer Data From People With and Without Lower Limb Amputation During Daily Life*. Accessed: Dec. 4, 2024. [Online]. Available: https://www.researchgate.net/publication/341255118_IMU_and_Accelerometer_Data_from_people_with_and_without_lower_limb_amputation_during_daily_life
- [38] N. Aizu, Y. Oouchida, K. Yamada, K. Nishii, and S.-I. Izumi, "Use-dependent increase in attention to the prosthetic foot in patients with lower limb amputation," *Sci. Rep.*, vol. 12, no. 1, p. 12624, Jul. 2022, doi: [10.1038/s41598-022-16732-z](https://doi.org/10.1038/s41598-022-16732-z).
- [39] C. Timmermans, A. G. Cutti, H. van Donkersgoed, and M. Roerdink, "Gaitography on lower-limb amputees," *Prosthetics Orthotics Int.*, vol. 43, no. 1, pp. 71–79, Feb. 2019, doi: [10.1177/0309364618791618](https://doi.org/10.1177/0309364618791618).
- [40] L. Kark, R. Odell, A. S. McIntosh, and A. Simmons, "Quantifying prosthetic gait deviation using simple outcome measures," *World J. Orthopedics*, vol. 7, no. 6, p. 383, 2016, doi: [10.5312/wjo.v7.i6.383](https://doi.org/10.5312/wjo.v7.i6.383).
- [41] N. S. Kamarul Hawari Ghazali, "An intelligent irregular surface recognition approach for integration in instrumentation and control system of lower limb prostheses toward predictive walking," in *Proc. Symp. Intell. Manuf. Mechatronics (SIMM)*, Sep. 2024, pp. 1–19.
- [42] E. S. Muñoz-Larrosa, M. Riveras, M. Oldfield, A. F. Shaheen, G. Schlotthauer, and P. Catalfamo-Formento, "Gait event detection accuracy: Effects of amputee gait pattern, terrain and algorithm," *J. Biomechanics*, vol. 177, Dec. 2024, Art. no. 112384, doi: [10.1016/j.jbiomech.2024.112384](https://doi.org/10.1016/j.jbiomech.2024.112384).
- [43] S. Aydin Fandakli and H. I. Okumus, "Deep learning based ankle-foot movement classification for prosthetic foot," *Neural Comput. Appl.*, vol. 36, no. 19, pp. 11397–11407, Jul. 2024.
- [44] G. Ding, I. Georgilas, and A. Plummer, "A deep learning model with a self-attention mechanism for leg joint angle estimation across varied locomotion modes," *Sensors*, vol. 24, no. 1, p. 211, Dec. 2023, doi: [10.3390/s24010211](https://doi.org/10.3390/s24010211).



NORAZIAN SUBARI received the Diploma degree in electrical engineering (instrumentation) and the B.Eng. degree (Hons.) in electrical-electronics engineering from Universiti Teknologi MARA, Malaysia, in 2003 and 2007, respectively, and the M.Sc. degree in electrical and electronic engineering from Universiti Sains Malaysia, Malaysia, in 2014. She is currently a Lecturer with the Faculty of Electrical and Electronics Engineering Technology, Universiti Malaysia Pahang Al-Sultan Abdullah (UMPSA), Malaysia. She actively contributes to research and innovation in the field of electrical and electronic engineering. She also supervises undergraduate research projects and teaches various subjects, including electronic devices, electropneumatics, digital electronics, and measurement technology. Her research interests include computational intelligence, signal processing, and classification techniques. She has received multiple awards for her contributions to research and technology development, including the Silver Medals in the Creation, Innovation, Technology and Research Exposition (CITREX), in 2020, 2023, and 2024.



KAMARUL HAWARI GHAZALI (Senior Member, IEEE) received the bachelor's and master's degrees in electrical engineering from Universiti Teknologi Malaysia and the Ph.D. degree in electrical, electronic, and system engineering from Universiti Kebangsaan Malaysia. He specializes in machine vision, image processing, artificial intelligence, machine learning, data analytics, and GNSS signal processing. He is currently a Distinguished Academic and Researcher with the Faculty of Electrical and Electronics Engineering Technology, Universiti Malaysia Pahang Al-Sultan Abdullah (UMPSA), Malaysia. As a Professional Engineer with the Board of Engineers Malaysia, he has made significant contributions to academia and industry. His research accomplishments include leading over 30 funded research projects, authoring numerous publications in indexed journals, and supervising both Ph.D. and master's students. He has actively collaborated with international institutions and industry partners, securing grants that span cutting-edge areas, such as AI applications, satellite navigation, and signal processing.



YUANFA JI (Member, IEEE) received the B.Eng. and M.Eng. degrees from Guilin University of Electronic Science and Technology, China, in 1998 and 2004, respectively, and the Ph.D. degree in satellite navigation from the National Astronomical Observatories, China, in 2008. He later completed a Postdoctoral Research Program with the National Astronomical Observatories, in 2010. He is currently a Level-2 Professor with the Department of Information and Communication, Guilin University of Electronic Technology (GUET), China, where he is also a Ph.D. Advisor. He has been recognized as Guangxi Bagui Scholar and Guangxi Outstanding Expert. He has led and participated in more than 20 national and provincial research projects, securing research funding exceeding 20 million yuan. He has published over 100 academic articles, holds more than 100 national patents, and has received over ten provincial and ministerial-level scientific awards, including China Patent Award of Excellence, the First Prize for Technological Invention in Guangxi, and multiple Science and Technology Progress Awards. His contributions to Beidou satellite navigation and positioning technologies have significantly advanced the field of geospatial intelligence. His research interests include GNSS positioning, satellite navigation, sensor fusion, artificial intelligence, and geospatial technologies.

...