Comparison of Time-Domain and Frequency-Domain EMG Features for Gait Phases Classification Using Machine Learning

Muhamad Amirul Sunni Rohim[1], Nurhazimah Nazmi[2]*,
Shin-Ichirou Yamamoto[3], Muhammad Kashfi Shabdin[4], Mohd Asyadi Azam[5]
Malaysia-Japan International Institute of Technology, Universiti Teknologi Malaysia, Kuala Lumpur,
Jalan Sultan Yahya Petra, 54100, Kuala Lumpur, Malaysia[1, 2]
Department of Bio-Science and Engineering, Shibaura Institute of Technology, Fukasaku 307, Saitama-City 337-8570, Japan[2, 3]
Superconductor and Thin Film Laboratory-Faculty of Science-Department of Physics,
Universiti Putra Malaysia, 43400 UPM Serdang, Selangor, Malaysia[4]
Division of Advanced Applied Physics, Institute of Engineering,
Tokyo University of Agriculture and Technology, 2-24-16, Naka-cho, Koganei-shi, Tokyo 184-8588, Japan[4]
Faculty of Industrial and Manufacturing Technology and Engineering,
Universiti Teknikal Malaysia Melaka, Durian Tunggal, Melaka, 76100, Malaysia[5]

***Abstract*—Accurate gait phase detection is essential for biomechanical analysis and the control of wearable assistive devices such as powered prostheses and exoskeletons. Electromyography (EMG) provides a direct representation of neuromuscular activation and offers potential advantages for low-latency, anticipatory gait phase recognition. However, the effectiveness of different EMG feature representations for stance-swing classification has not yet been clearly established. Therefore, this study presents a systematic comparison of time-domain (TD) and frequency-domain (FD) EMG features for gait phase classification. EMG signals were recorded from the tibialis anterior and medial gastrocnemius muscles of ten healthy participants during level walking. After preprocessing and segmentation, TD and FD features were extracted and used as inputs to a support vector machine classifier with a radial basis function kernel. Model performance was evaluated using a leave-one-subject-out cross-validation framework to assess generalization. The results demonstrate that TD features consistently outperform FD features across all evaluation metrics, achieving an accuracy of 0.813 ± 0.112, macro-averaged F1-score (Macro-F1) of 0.812 ± 0.114, and Matthews correlation coefficient (MCC) of 0.672 ± 0.178, compared to FD features with an accuracy of 0.712 ± 0.077, Macro-F1 of 0.708 ± 0.079, and MCC of 0.448 ± 0.159. These findings indicate that TD features more effectively capture the transient amplitude-based neuromuscular patterns associated with gait phase transitions. In addition, TD features offer lower computational complexity, making them well-suited for real-time implementation. Overall, this study highlights the superiority of time-domain EMG representations for reliable and efficient gait phase detection and provides practical guidance for the development of wearable gait monitoring and assistive control systems.**

Keywords—Electromyography (EMG); gait phase detection; stance–swing classification; time-domain features; frequency-domain features; support vector machine (SVM); wearable assistive devices

I. Introduction

Human gait is a cyclic neuromuscular behavior that arises from the coordinated activation of multiple lower-limb muscles to alternately support body weight and enable forward progression. For each leg, the gait cycle is commonly described by two principal phases: stance, during which the foot is in contact with the ground, and swing, during which the foot is lifted and advanced. These phases form the fundamental state representation used in many gait analysis and control frameworks [1,2]. Accurate and low-latency identification of gait phases is essential in biomechanics and rehabilitation engineering, particularly in wearable assistive technologies where mechanical assistance must be synchronized with the user's movement intent. In this context, a comprehensive review of lower-limb exoskeleton systems highlighted that reliable intention recognition, particularly through physiological signals such as electromyography (EMG), is critical for effective operation, while also identifying persistent challenges related to system stability and user comfort [2]. Complementing this system-level perspective, recent advances in machine learning have demonstrated that gait phase recognition can be achieved with high accuracy, exceeding 94% in multi-phase classification tasks, thereby supporting the feasibility of data-driven control strategies [3]. More broadly, artificial intelligence-based approaches have been increasingly adopted for locomotion classification and control, although their clinical validation remains limited [4]. Furthermore, existing exoskeleton systems predominantly rely on rule-based and trajectory-tracking strategies, with relatively low adoption of adaptive control methods despite their potential to improve responsiveness and user-specific adaptation [5]. Collectively, these findings emphasize that robust estimation of gait state and user intent remains central to achieving safe, comfortable, and clinically effective assistance, particularly in real-world applications involving diverse users.

*Corresponding author

Wearable gait phase recognition has traditionally relied on kinematic and kinetic sensors such as inertial measurement units and pressure insoles, which are effective for detecting gait events but primarily reflect motion and contact changes after their occurrence [6,7]. Electromyography (EMG) provides a complementary perspective by directly capturing neuromuscular activation and has therefore been widely investigated as a human–machine interface for gait monitoring and assistive device control. This is attributed to the electromechanical delay, whereby muscle activation precedes observable limb kinematics by approximately 120 ms, enabling earlier insight into movement intent [8]. Similarly, a metric learning-based temporal convolutional network has achieved high classification accuracy while improving robustness to noise disturbances, highlighting the reliability of EMG signals in complex environments [9]. Additional studies have further confirmed the effectiveness of EMG-based recognition using diverse approaches, including multi-scale deep architectures, domain adaptation strategies, and multi-feature fusion methods [10–12]. This anticipatory characteristic, commonly referred to as electromechanical delay, enables EMG-based systems to reduce effective control latency and improve synchronization of assistive actions. It has also motivated the integration of EMG with complementary sensing modalities in multimodal frameworks to enhance performance [13]. Consequently, EMG has attracted increasing attention for online stance–swing recognition in wearable systems. Moreover, the structured activation patterns of lower-limb muscles during gait further support its discriminative capability; for example, tibialis anterior activation occurs before foot contact as a preparatory response to ground interaction, while plantarflexor muscles contribute to propulsion during late stance [14]. In addition, EMG-derived features have been shown to reliably reflect key gait parameters, including stance and swing phases, reinforcing the ability of neuromuscular signals to capture phase-specific characteristics of locomotion [15]. Related studies have also extended EMG-based prediction to more complex locomotor intentions such as directional stepping, further demonstrating the richness of neuromuscular information for wearable control applications [16].

Consequently, a growing body of research has applied machine learning and deep learning techniques to EMG-based gait phase recognition, particularly for real-time exoskeleton control and subject-independent stance–swing estimation. For instance, a BiLSTM-based framework has demonstrated reliable gait phase identification during exoskeleton-assisted walking, with the ability to generalize across different locomotion conditions [8]. Further optimization of deep learning architectures has shown that high classification accuracy can be achieved with significantly reduced computational cost, making such models suitable for real-time deployment [1]. In addition, a metric learning-based temporal convolutional network has been shown to improve both recognition accuracy and robustness to noise disturbances, highlighting the importance of discriminative feature learning for reliable performance in complex environments [9]. Other studies have explored alternative strategies, including domain adaptation to address gait speed variability [11], multi-feature fusion combining time-domain, frequency-domain, and nonlinear descriptors [12], and multimodal sensor fusion frameworks integrating EMG with complementary signals to enhance recognition accuracy [13]. However, despite these methodological advances, an important practical design question remains insufficiently addressed: which EMG feature-domain representation most effectively supports reliable stance–swing classification when evaluated under controlled conditions, including consistent classifier selection and subject-independent validation. Recent surveys of EMG analysis emphasize that feature representation is a major determinant of performance, robustness, and computational efficiency, typically categorizing handcrafted features into time-domain, frequency-domain, and time–frequency domains with different cost–information trade-offs [17]. Within gait phase recognition research, feature selection remains highly heterogeneous, ranging from low-cost time-domain descriptors to hybrid feature sets and end-to-end learned representations, often without systematic comparison to determine which feature domain is most appropriate for stance–swing separability under realistic deployment conditions [8,9,11–13].

The choice of feature domain is particularly important for stance–swing classification because surface EMG during walking is characterized by transient, phase-locked bursts of muscle activation and amplitude modulations that are inherently represented in the time domain [14–17]. Accordingly, time-domain descriptors directly capture the temporal structure and magnitude of neuromuscular activity associated with gait events. In contrast, frequency-domain features summarize spectral properties of the signal that are more closely related to broader physiological conditions, such as muscle fatigue, commonly quantified using measures such as median or mean frequency [18,19]. Although spectral features may provide complementary information in certain EMG analysis tasks, their contribution to gait phase discrimination is less direct, as stance–swing transitions are primarily governed by the timing and amplitude patterns of muscle activation. This is supported by studies demonstrating that anticipatory muscle activity plays a central role in real-time gait phase estimation for exoskeleton control, where temporal alignment of EMG signals enables early detection of gait events [1,8]. In addition, multimodal approaches have shown that EMG contributes most effectively through its temporal continuity and activation patterns rather than through spectral characterization alone [13]. Furthermore, real-time wearable control systems, including powered prostheses and lower-limb exoskeletons, impose strict latency and computational constraints, which necessitate feature representations that balance discriminative capability with efficiency [9,13]. In this context, recent studies on EMG-driven intent estimation highlight a practical accuracy–latency trade-off and consistently report that time-domain features offer a more favorable balance between classification performance and computational efficiency compared with frequency-domain alternatives in time-constrained control applications [20].

Therefore, a systematic and controlled comparison of time-domain versus frequency-domain EMG features for stance–swing classification is needed to clarify which representation more effectively captures the neuromuscular signatures of gait phases under subject-independent validation. In this study, lower-limb EMG recorded from TA and MGAS during walking was used to extract feature sets in the time and frequency domains. A support vector machine (SVM) classifier was used to evaluate stance–swing classification, and performance was assessed under a subject-independent validation framework. By holding the classification model and evaluation protocol constant while varying only the feature-domain representation, the study aims to provide actionable guidance for designing robust, computationally efficient EMG-based gait phase detection pipelines for wearable gait monitoring and assistive device control.

II. Related Work

Wearable gait phase detection has traditionally relied on kinematic and kinetic sensing modalities, particularly inertial measurement units and plantar-pressure or insole systems, as these sensing approaches provide reliable detection of gait events and are well established in prosthetic and exoskeleton applications [6,21,22]. However, these signals primarily reflect the mechanical outcomes of movement rather than the underlying neuromuscular processes. Recent reviews of lower-limb exoskeleton sensing distinguish these mechanical (posterior) signals from earlier physiological signals, emphasizing that motion and foot–ground contact reflect the mechanical outcome of movement, whereas biological signals such as EMG are more directly linked to the neuromuscular processes underlying movement intention and execution [2]. Consequently, electromyography (EMG) has gained increasing attention as a complementary modality for gait-state estimation, with studies explicitly leveraging its anticipatory (ahead-of-motion) characteristics for gait phase recognition and kinematics prediction [8,23].

Building on this physiological advantage, both conventional machine learning and deep learning approaches have been extensively investigated for EMG-based gait recognition. Recent reviews of EMG classification indicate that handcrafted feature pipelines remain widely used, often combined with conventional supervised classifiers such as support vector machines, k-nearest neighbors, and random forests, particularly when dataset sizes are limited [18]. In gait-specific studies, Park et al. [24] systematically compared multiple classifiers, including decision trees, k-nearest neighbors, support vector machines, and neural networks, while Morbidoni et al. [25] applied deep learning for stance–swing classification during level-ground walking. Ling et al. [11] proposed a CNN-based domain-adaptive model to address gait phase recognition under varying walking speeds, Cai et al. [3] introduced an LDA-PSO-LSTM framework for seven-phase recognition, and Guerra et al. [8] employed a bidirectional LSTM model for stance–swing identification during exoskeleton-assisted walking. Collectively, these studies demonstrate a transition toward more adaptive temporal and deep learning architectures, while highlighting that conventional methods remain competitive and widely used in practical applications.

Despite these algorithmic advances, feature representation remains a critical and influential design factor. Recent reviews consistently categorize EMG features into time-domain (TD), frequency-domain (FD), and time–frequency (TF) groups [26,27]. Within this framework, TD features are widely used in gait phase recognition studies due to their ability to capture activation magnitude and timing, whereas FD features primarily characterize spectral properties associated with broader physiological states such as muscle fatigue. TF approaches aim to capture both temporal and spectral information but are less commonly adopted due to increased computational complexity and reduced interpretability [28,29]. This classification is directly relevant to gait phase recognition, where common TD descriptors—such as root mean square, mean absolute value, waveform length, zero-crossings, and slope-sign changes—remain widely used due to their simplicity and effectiveness. For example, Nazmi et al. [30] employed a TD feature set including root mean square, standard deviation, mean absolute value, integrated EMG, and waveform length, showing the continued effectiveness of established TD descriptors in recent gait phase classification studies.

Evidence from gait studies also suggests that task formulation significantly influences the effectiveness of feature-domain representations. In Nazmi et al. [30], binary stance–swing classification achieved an accuracy of approximately 0.98 across all features, suggesting that two-class phase discrimination can be effectively achieved using relatively simple feature representations. In contrast, Cai et al. addressed a more complex seven-phase recognition task using feature-combination selection across TD, FD, and TF domains, while Morbidoni et al. and Guerra et al. utilized deep learning architectures to model stance–swing dynamics directly from EMG signals [3,8,25]. Collectively, these findings indicate that the suitability of feature-domain representations is task-dependent, with simpler TD-based features often sufficient for binary classification, while more complex tasks may require richer or learned representations.

A second major issue is generalization across subjects, which remains a key challenge in EMG-based gait analysis. Di Nardo et al. demonstrated that intra-subject models outperform inter-subject models for EMG-based gait-event prediction, highlighting the significant impact of inter-individual variability [31]. A subsequent study by Di Nardo et al. further showed that both EMG signal processing strategies and sensor configuration complexity influence classification and prediction performance [32]. More recent work has attempted to address this limitation through domain adaptation and rapid recalibration techniques, particularly to accommodate variations such as walking speed and environmental conditions [11]. At a broader methodological level, recent research in wearable robotics emphasizes participant-based data splitting or leave-one-subject-out (LOSO) validation as more realistic evaluation strategies compared to random mixed-subject splits, as these approaches better assess performance on unseen users and reduce optimistic bias caused by cross-subject data leakage [33].

Overall, the literature supports EMG as a robust and promising modality for gait phase recognition. However, existing studies have not systematically compared time-domain and frequency-domain handcrafted features under controlled experimental conditions, particularly when classifier choice and subject-independent evaluation protocols are held constant. This limitation prevents definitive conclusions regarding the most suitable feature-domain representation for reliable stance–swing classification in practical deployment scenarios, thereby defining a focused and well-motivated research gap.

III. Methods

A. Subject

Ten healthy male adults aged between 18 and 40 years participated in this study. This age range was selected to represent individuals with fully developed and stable neuromuscular control [34]. All participants reported no history of neurological disorders, musculoskeletal injuries, or gait impairments, ensuring that the recorded muscle activity reflected normal walking patterns. Prior to participation, all subjects provided written informed consent. The study protocol was reviewed and approved by the Universiti Teknologi Malaysia Research Ethics Committee (Approval No. UTMREC-2025-122).

B. Experimental Protocol

Each participant performed level-ground walking trials in a controlled laboratory environment. During each trial, participants walked along a straight, unobstructed pathway from a marked starting point to an end point and then returned to the starting location, as shown in Fig. 1. To minimize variability caused by footwear differences, all participants wore standardized shoes during the experiment. Participants were instructed to walk at their self-selected comfortable speed in order to preserve natural gait mechanics. This approach allows the recorded muscle activation patterns to represent natural locomotion behavior rather than constrained walking conditions.

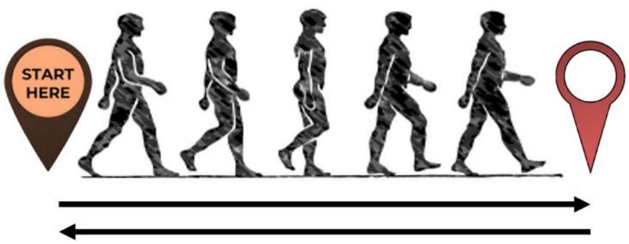

Fig. 1. Experimental protocol for walking trial.

C. Data Collection

EMG signals were recorded from two lower-limb muscles: the tibialis anterior (TA) and the medial gastrocnemius (MGAS). These muscles were selected because they play complementary roles in ankle control during walking. The TA primarily contributes to ankle dorsiflexion and is typically active during the swing phase to ensure foot clearance, whereas the MGAS contributes to plantarflexion and propulsion during the stance phase [35,36]. EMG electrodes were placed on this muscle based on the recommendations of Surface Electromyography for Non-Invasive Assessment (SENIAM) [37]. The electrodes were connected to a Biopac MP150 data acquisition system, and EMG signals were sampled at 1000 Hz, providing sufficient temporal resolution to capture rapid muscle activation changes during gait.

D. Gait Phase Labeling

The gait cycle was divided into two primary phases, which are stance and swing. The stance phase corresponds to the period during which the foot is in contact with the ground, while the swing phase refers to the period when the foot is lifted and moves forward to prepare for the next step [38]. Ground-truth labels for the gait phases were obtained through manual annotation based on visual inspection of the recorded walking sequences. A binary labeling scheme was adopted in which stance was assigned the label 1 and swing was assigned the label 0. In typical human walking, the stance phase occupies approximately 60% of the gait cycle, while the swing phase accounts for the remaining 40% [38]. This temporal distribution reflects the functional roles of the two phases, where stance provides support and propulsion while swing advances the limb for the subsequent step. Fig. 2 shows an example of the EMG signals recorded from the TA and MGAS muscles together with the corresponding stance–swing labels. These labeled gait phases were subsequently used as target classes for supervised classification after preprocessing and feature extraction.

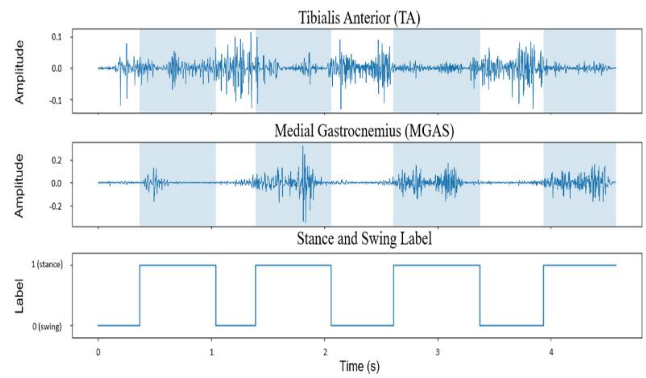

Fig. 2. Stance and swing labeling.

E. EMG Signal Preprocessing

The EMG signals were preprocessed using a sixth-order Butterworth band-pass filter with cutoff frequencies of 20–450 Hz to reduce noise and artifacts [39,40]. This frequency range preserves the dominant EMG spectral components while attenuating motion artifacts and high-frequency noise. After filtering, the EMG signals were segmented using a sliding window approach. Each analysis window had a duration of 200 ms with 50% overlap between consecutive windows, a configuration commonly adopted in EMG analysis to balance temporal resolution with statistical stability [41]. This corresponds to a step size of 100 ms between adjacent windows. For each analysis window, the corresponding gait phase label was determined using a majority voting strategy based on the stance–swing annotations within that window.

F. Feature Extraction

Two categories of features were investigated in this study: time-domain (TD) features and frequency-domain (FD) features. Feature extraction was performed independently for the TA and MGAS signals within each analysis window. The features extracted from both muscles were concatenated to form the final feature vector used for classification.

1) Time-domain features: Time-domain features characterize the amplitude and temporal properties of EMG signals. Six well-established TD features were employed, which are Mean Absolute Value (MAV), Standard Deviation (SD), Root Mean Square (RMS), Integrated EMG (iemg), Waveform Length (WL), and Zero Crossing (ZC), as Shown in Table I. These features have been widely validated in the literature and are known to achieve high classification accuracy [42,43].

TABLE I. TIME DOMAIN (TD) FEATURE EXTRACTION

Time Domain (TD) features	**For each window containing N signal samples *xi*, the following features were extracted.**
Mean Absolute Value (MAV)	$MAV = \frac{1}{N}\sum_{i=1}^{N}\lvert x_i\rvert$
Standard Deviation (SD)	$SD = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(x_i-\mu)^2}$, μ=mean of EMG signal
Root Mean Square (RMS)	$RMS = \sqrt{\frac{1}{N}\sum_{i=1}^{N}x_i^2}$
Integrated EMG (IEMG)	$IEMG = \sum_{i=1}^{N}\lvert x_i\rvert$
Waveform Length (WL)	$WL = \sum_{i=1}^{N-1}\lvert x_{i+1}-x_i\rvert$
Zero Crossing (ZC)	$ZC = \sum_{n=1}^{N-1}\lvert sgn(x_i) - sgn(x_i - 1)\rvert$ where sgn(), $sgn(x) = \begin{cases} 1 & \text{if } x>0 \\ 0 & \text{if } x=0 \\ -1 & \text{if } x<0 \end{cases}$

2) Frequency-Domain features: Frequency-domain (FD) features describe the spectral properties of EMG signals. In this study, five widely used FD features were extracted, namely Mean Frequency (MNF), Median Frequency (MDF), Peak Frequency (PF), Total Spectral Power (TSP), and Spectral Entropy, as summarized in Table II. These features have also been widely validated in the literature on EMG-based classification tasks [44,45].

TABLE II. FREQUENCY DOMAIN (FD) FEATURE EXTRACTION

Frequency Domain (FD) features	**The power spectral density (PSD) of each window was estimated using Welch's method. Let P(j) denote the power spectral density at the j-th frequency bin, where M is the total number of frequency bins.**
Mean Frequency (MNF)	$MNF = \frac{\sum_{j=1}^{MDF} f_j P(j)}{\sum P(j)}$
Median Frequency (MDF)	$\sum_{j=1}^{MDF} P(j) = \sum_{j=MDF}^{M} P(j) = \frac{1}{2}\sum_{j=1}^{M} P(j)$
Peak Frequency (PF)	$PF = \max P(j), \quad j = 1,...,M$
Total Spectral Power (TTP)	$TTP = \sum_{j=1}^{M} P(j)$
Spectral Entropy (SpecEn)	$\text{SpecEn} = -\frac{\sum_{j=1}^{M} P(j)\log_2 P(j)}{\log_2 M}$

G. Support Vector Machine Classification

Support vector machine (SVM) is a supervised learning models that find an optimal hyperplane that separates the data into different classes by transforming it into a higher-dimensional space [46]. This transformation allows SVM to handle nonlinear decision boundaries by implicitly mapping inputs into high-dimensional feature spaces. In addition, SVMs often yield high accuracy and generalize well on high-dimensional data, making them popular for many classification problems [47,48]. In this study, gait phase classification was performed using an SVM classifier with a radial basis function (RBF) kernel. Before classification, the extracted feature vectors were standardized using z-score normalization so that all features contributed equally to the learning process. The classification pipeline, therefore, consisted of a feature scaling stage followed by SVM classification.

The SVM hyperparameters were configured as follows:

- Kernel: radial basis function (RBF)
- Kernel parameter: γ=0.005

The selection of the kernel parameter $\gamma = 0.005$ was based on preliminary empirical evaluation to balance model generalization and decision boundary smoothness. In the context of an RBF kernel, smaller γ values produce smoother decision boundaries and reduce the risk of overfitting, whereas larger values lead to more complex boundaries that may degrade generalization performance, particularly under subject-independent evaluation [49].

H. Model Evaluation

To evaluate the generalization capability of the classifier across participants, a leave-one-subject-out (LOSO) cross-validation strategy was employed. In this approach, the data from one subject was used as the test set while the remaining subjects were used to train the classifier. The process was repeated until each subject had served once as the test subject. Classification performance was assessed using accuracy, precision, recall, F1-score, and Matthews Correlation Coefficient (MCC), as shown in Table III. The F1-score represents the harmonic mean of precision and recall, while the macro-averaged F1-score (Macro-F1) was computed as the average of class-wise F1-scores to ensure balanced evaluation across stance and swing classes. MCC was utilized as a metric to capture agreement beyond chance while providing a balanced evaluation of classification performance across both classes.

TABLE III. PERFORMANCE METRIC

Performance Metric	**Formula**
Accuracy	$\text{Accuracy} = \frac{TP+TN}{TP+TN+FP+FN}$
Precision	$\text{Precision} = \frac{TP}{TP+FP}$
Recall	$\text{Recall} = \frac{TP}{TP+FN}$
F1-score	$\text{F1} = 2\times\frac{Precision\times Recall}{Precision+Recall}$

Macro F1	$\text{Macro F1} = \frac{F1_{swing} + F1_{stance}}{2}$
MCC	$\text{MCC} = \frac{TP \times TN - FP \times FN}{\sqrt{(TP+FP)(TP+FN)(TN+FP)(TN+FN)}}$

IV. Results

A. *Overall Classification Performance (Time vs. Frequency Domain Features)*

The overall classification performance of the time-domain (TD) and frequency-domain (FD) feature sets evaluated using leave-one-subject-out (LOSO) cross-validation is summarized in Table IV and illustrated in Fig. 3. The TD feature set achieved a higher classification accuracy of 0.813 ± 0.112, outperforming the FD feature set with an accuracy of 0.712 ± 0.077. A similar trend was observed for the macro-averaged F1-score (Macro-F1), where TD features achieved 0.812 ± 0.114, compared to 0.708 ± 0.079 for FD features. Furthermore, the Matthews correlation coefficient (MCC) also indicated superior performance for TD features (0.672 ± 0.178) relative to FD features (0.448 ± 0.159).

TABLE IV. Overall Performance Comparison of Time-Domain (TD) and Frequency-Domain (FD) Feature Sets Using LOSO Cross-Validation Based on Fig. 3.

Feature Mode	Accuracy	Macro-F1	MCC
TD	0.813 ± 0.112	0.812 ± 0.114	0.672 ± 0.178
FD	0.712 ± 0.077	0.708 ± 0.079	0.448 ± 0.159

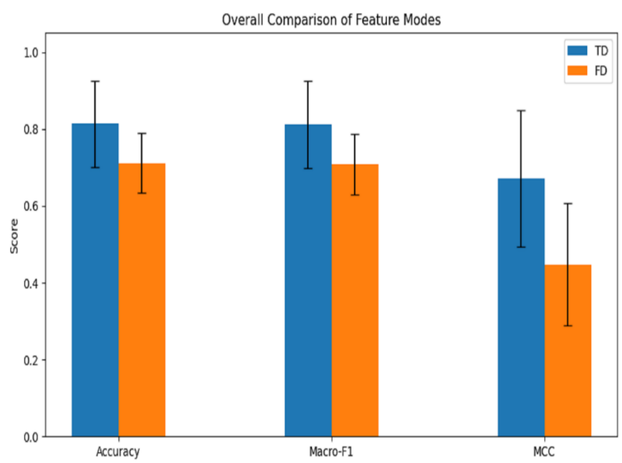

Fig. 3. Comparison of classification performance between time-domain (TD) and frequency-domain (FD) EMG feature sets using SVM under leave-one-subject-out cross-validation. Error bars represent standard deviation across subjects.

B. *Per-Class Performance and Confusion Matrix Analysis*

To further examine class-specific behavior, Fig. 4 presents a heatmap of precision, recall, and F1-score for both swing and stance phases. For the TD feature set, swing phase detection achieved a recall of 0.935 and an F1-score of 0.822, whereas stance detection exhibited a recall of 0.729 and a precision of 0.929. For the FD feature set, both phases exhibited lower performance (swing F1: 0.687, stance F1: 0.729), indicating reduced classification effectiveness.

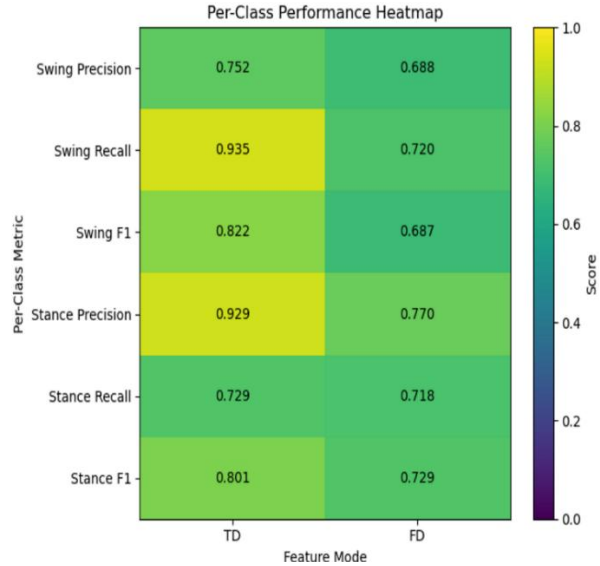

Fig. 4. Per-class performance heatmap.

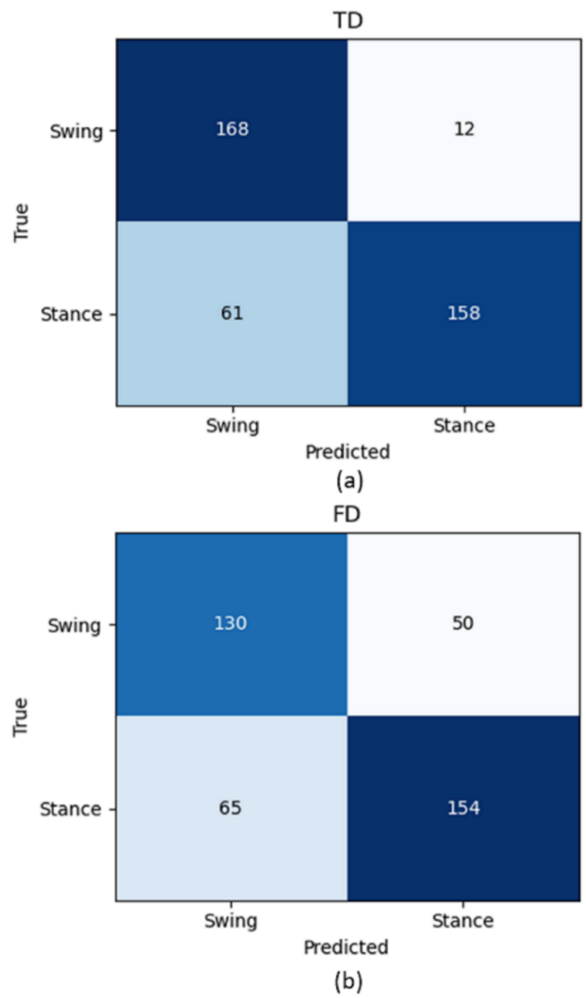

Fig. 5. Confusion matrices across feature modes: (a) Time-domain (TD) (b) Frequency-domain (FD).

Fig. 5 presents the confusion matrices, for the (a) time-domain (TD) and (b) frequency-domain (FD) feature sets, providing further insight into classification errors. Most predictions are concentrated along the diagonal, indicating correct classification of both gait phases. However, a noticeable proportion of stance samples are misclassified as swing, consistent with the lower stance recall observed in Fig. 4. Importantly, Fig. 5(a) exhibits fewer off-diagonal errors compared to Fig. 5(b), confirming its superior classification performance. This reduction in misclassification highlights the effectiveness of TD features in capturing phase-specific amplitude variations during gait transitions.

V. Discussion

A. Interpretation of Feature Performance

In Table IV, the results demonstrate that TD features provide stronger discriminative information for gait phase classification compared to FD features. This is likely because time-domain descriptors capture transient neuromuscular activation patterns, whereas frequency-domain features primarily represent spectral distributions and may be less sensitive to temporal variations. This observation aligns with previous studies reporting that TD features often outperform FD features in dynamic EMG classification tasks [50,51]. The higher MCC values further indicate that TD features maintain more balanced classification performance across classes.

B. Biomechanical Interpretation of Class Differences

The higher recall for the swing phase observed in Fig. 4 can be explained by gait biomechanics. During swing, the tibialis anterior (TA) exhibits a distinct high-amplitude activation burst associated with ankle dorsiflexion and foot clearance [52]. This produces clear and temporally localized EMG patterns, making the phase easier to detect. In contrast, the stance phase involves more distributed and overlapping muscle activations, including contributions from the gastrocnemius and soleus. This results in increased signal variability and reduced separability. Similar findings have been reported in prior work, where swing phases are often classified more accurately than stance phases despite their shorter duration [25].

C. Implications for EMG-Based Gait Phase Detection

The findings highlight the critical role of feature representation in EMG-based gait phase classification. Across all evaluation metrics, TD features consistently outperformed FD features, demonstrating their effectiveness in capturing neuromuscular activation dynamics during locomotion. This result is physiologically justified, as EMG signals predominantly reflect amplitude modulation of muscle activity, which is more directly represented in the time domain [53]. From an application perspective, TD features provide a computationally efficient and reliable basis for real-time gait phase detection. Their strong classification performance indicates that they can effectively support control strategies in assistive technologies, such as powered prostheses and lower-limb exoskeletons, where accurate phase detection is required to synchronize mechanical assistance with user movement [24].

A key limitation of the present framework is that each analysis window is treated as an independent sample, without explicitly modeling temporal dependencies between consecutive gait phases. Consequently, the model does not enforce physiologically consistent phase transitions, which may lead to misclassification near transition boundaries. Recent studies have demonstrated that sequence-aware models, such as recurrent neural networks and temporal convolutional networks, can improve gait phase recognition by capturing temporal continuity and transition constraints [8,54]. Another limitation of the present study is the use of only two lower-limb muscles, namely the tibialis anterior and medial gastrocnemius. Although these muscles provide strong discriminative information due to their roles in dorsiflexion and plantarflexion, gait is inherently governed by coordinated activation across multiple muscle groups. As both muscles are located at the distal segment of the lower limb, the current framework primarily reflects ankle-level control and does not capture the contributions of proximal muscles involved in hip and knee stabilization. Consequently, this limited muscle representation may restrict the generalizability of the findings to more complex locomotion scenarios that require multi-joint coordination and whole-body control. Previous studies have shown that incorporating additional muscles or muscle synergy information can improve classification robustness and better capture the complexity of locomotion [24,55].

Overall, the results demonstrate that TD features provide a physiologically meaningful and computationally efficient representation for gait phase detection. These characteristics make them well-suited for implementation in real-time wearable systems, supporting both assistive device control and rehabilitation-oriented gait monitoring applications.

VI. Conclusion

This study investigated the effectiveness of time-domain (TD) and frequency-domain (FD) electromyography (EMG) features for stance–swing gait phase classification using signals from the tibialis anterior (TA) and medial gastrocnemius (MGAS) muscles during walking. A support vector machine classifier was evaluated under a subject-independent leave-one-subject-out framework to assess generalization performance. The results consistently demonstrate that TD features outperform FD features across all evaluation metrics, including classification accuracy, macro-averaged F1-score, and Matthews correlation coefficient. These findings indicate that amplitude-based temporal characteristics of EMG signals provide more discriminative information for stance–swing classification than spectral representations under the evaluated conditions. Despite these findings, several limitations should be acknowledged. First, the study was conducted on a relatively small dataset consisting of ten healthy participants, which may limit the statistical generalizability of the results. Second, the observed variability in classification performance, particularly in MCC values, reflects inter-subject differences in EMG signal characteristics, including muscle activation patterns and signal quality. Third, the analysis was restricted to two muscles (TA and MGAS) and to predefined TD and FD feature sets, which may not fully capture the complexity of multi-muscle coordination during gait. From an application perspective, TD features offer a computationally efficient and physiologically meaningful representation for real-time gait phase detection. However, the findings should be interpreted within the scope of the experimental setup and feature design choices, particularly under controlled laboratory conditions. These constraints highlight the need for further validation before deployment in real-world assistive systems.

Future work should focus on improving the robustness, scalability, and real-world applicability of EMG-based gait phase classification systems. One important direction is the inclusion of larger and more diverse participant populations, including individuals with gait impairments, to better evaluate generalization across varying neuromuscular conditions. In addition, future studies should investigate multi-muscle configurations to capture a more comprehensive representation of lower-limb coordination. Another key direction involves the integration of temporal modeling approaches. The current study treats EMG windows independently and does not incorporate sequential dependencies between gait phases. Incorporating sequence-aware models, such as recurrent neural networks or temporal convolutional architectures, may improve classification stability and reduce misclassification during phase transitions. Furthermore, while this study compared TD and FD features independently, combining both feature domains may provide complementary information that enhances classification performance. Hybrid feature representations that integrate temporal and spectral characteristics could offer improved robustness, particularly under noisy or variable conditions. Finally, future research should evaluate the proposed approach under real-world conditions, including varying walking speeds, terrains, and sensor noise levels, to assess its suitability for deployment in wearable assistive technologies and continuous gait monitoring applications.

ACKNOWLEDGMENT

The authors would like to acknowledge the contributions of all participants involved in this study. This work was supported by the Universiti Teknologi Malaysia, Geran Penyelidikan Kecerdasan Buatan, 44H54.

REFERENCES

[1] ADDIN Mendeley Bibliography CSL_BIBLIOGRAPHY R. Soldi, B.M.V. Guerra, S. Sozzi, L. Russo, S. Pizzocaro, R. Baptista, A.M. De Nunzio, M. Schmid, S. Ramat, Deep Learning Models Optimization for Gait Phase Identification from EMG Data During Exoskeleton-Assisted Walking, Biomimetics 10 (2025) 1–25, https://doi.org/10.3390/biomimetics10090617.

[2] Y. Sun, Y. Tang, J. Zheng, D. Dong, X. Chen, L. Bai, From sensing to control of lower limb exoskeleton: a systematic review, Annu. Rev. Control 53 (2022) 83–96, https://doi.org/10.1016/j.arcontrol.2022.04.003.

[3] S. Cai, D. Chen, B. Fan, M. Du, G. Bao, G. Li, Gait phases recognition based on lower limb sEMG signals using LDA-PSO-LSTM algorithm, Biomed. Signal Process. Control 80 (2023) 104272, https://doi.org/10.1016/j.bspc.2022.104272.

[4] O. Coser, C. Tamantini, P. Soda, L. Zollo, AI-based methodologies for exoskeleton-assisted rehabilitation of the lower limb: a review, Front. Robot. AI 11 (2024) 1–19, https://doi.org/10.3389/frobt.2024.1341580.

[5] J. de Miguel-Fernández, J. Lobo-Prat, E. Prinsen, J.M. Font-Llagunes, L. Marchal-Crespo, Control strategies used in lower limb exoskeletons for gait rehabilitation after brain injury: a systematic review and analysis of clinical effectiveness, J. Neuroeng. Rehabil. 20 (2023) 1–28, https://doi.org/10.1186/s12984-023-01144-5.

[6] H. Prasanth, M. Caban, U. Keller, G. Courtine, A. Ijspeert, H. Vallery, J. von Zitzewitz, Wearable sensor-based real-time gait detection: A systematic review, Sensors 21 (2021) 2727, https://doi.org/10.3390/s21082727.

[7] Z. Zhang, Z. Wang, H. Lei, W. Gu, Gait phase recognition of lower limb exoskeleton system based on the integrated network model, Biomed. Signal Process. Control 76 (2022) 103693, https://doi.org/10.1016/j.bspc.2022.103693.

[8] B.M.V. Guerra, M. Schmid, S. Sozzi, S. Pizzocaro, A.M. De Nunzio, S. Ramat, A Recurrent Deep Network for Gait Phase Identification from EMG Signals During Exoskeleton-Assisted Walking, Sensors 24 (2024) 6666, https://doi.org/10.3390/s24206666.

[9] J. Liu, C. Wang, B. He, P. Li, X. Wu, Metric Learning for Robust Gait Phase Recognition for a Lower Limb Exoskeleton Robot Based on sEMG, IEEE Trans. Med. Robot. Bionics 4 (2022) 472–479, https://doi.org/10.1109/TMRB.2022.3166543.

[10] X. Shi, X. Zhang, P. Qin, L. Huang, Y. Zhu, Z. Yang, Gait Phase Recognition in Multi-Task Scenarios Based on sEMG Signals, Biosensors 15 (2025) 1–21, https://doi.org/10.3390/bios15050305.

[11] Z.Q. Ling, J.C. Chen, G.Z. Cao, Y.P. Zhang, L.L. Li, W.X. Xu, S. Bin Cao, A Domain Adaptive Convolutional Neural Network for sEMG-Based Gait Phase Recognition Against to Speed Changes, IEEE Sens. J. 23 (2023) 2565–2576, https://doi.org/10.1109/JSEN.2022.3228320.

[12] P. Zhang, J. Zhang, A. Elsabbagh, Lower Limb Motion Intention Recognition Based on sEMG Fusion Features, IEEE Sens. J. 22 (2022) 7005–7014, https://doi.org/10.1109/JSEN.2022.3146446.

[13] Q. Liu, W. Sun, N. Peng, W. Meng, S.Q. Xie, DCNN-SVM-Based Gait Phase Recognition With Inertia, EMG, and Insole Plantar Pressure Sensing, IEEE Sens. J. 24 (2024) 28869–28878, https://doi.org/10.1109/JSEN.2024.3435884.

[14] S.T. Kwak, Y.H. Chang, Fascicle dynamics of the tibialis anterior muscle reflect whole-body walking economy, Sci. Rep. 13 (2023) 1–11, https://doi.org/10.1038/s41598-023-31501-2.

[15] S.H. Liu, A.K. Sharma, B.Y. Wu, X. Zhu, C.J. Chang, J.J. Wang, Estimating gait parameters from sEMG signals using machine learning techniques under different power capacity of muscle, Sci. Rep. 15 (2025) 1–15, https://doi.org/10.1038/s41598-025-95973-0.

[16] E. Anselmino, A. Mazzoni, S. Micera, EMG-based prediction of step direction for a better control of lower limb wearable devices, Comput. Methods Programs Biomed. 254 (2024) 108305, https://doi.org/10.1016/j.cmpb.2024.108305.

[17] J.H. Sul, L. Piyathilaka, D. Moratuwage, S. Dunu Arachchige, A. Jayawardena, G. Kahandawa, D.M.G. Preethichandra, Electromyography Signal Acquisition, Filtering, and Data Analysis for Exoskeleton Development, Sensors 25 (2025) 1–31, https://doi.org/10.3390/s25134004.

[18] S.M. Sid'El Moctar, I. Rida, S. Boudaoud, Comprehensive Review of Feature Extraction Techniques for sEMG Signal Classification: From Handcrafted Features to Deep Learning Approaches, Irbm 45 (2024) 100866, https://doi.org/10.1016/j.irbm.2024.100866.

[19] L. Quesada, D. Verdel, O. Bruneau, B. Berret, M.A. Amorim, N. Vignais, EMG feature extraction and muscle selection for continuous upper limb movement regression, Biomed. Signal Process. Control 103 (2025) 107323, https://doi.org/10.1016/j.bspc.2024.107323.

[20] P.N. Gunaratne, H. Tamura, An EMG-Based GRU Model for Estimating Foot Pressure to Support Active Ankle Orthosis Development, Sensors 25 (2025) 1–22, https://doi.org/10.3390/s25113558.

[21] H.T.T. Vu, D. Dong, H.L. Cao, T. Verstraten, D. Lefeber, B. Vanderborght, J. Geeroms, A review of gait phase detection algorithms for lower limb prostheses, Sensors (Switzerland) 20 (2020) 1–19, https://doi.org/10.3390/s20143972.

[22] V.M. Santos, B.B. Gomes, M.A. Neto, A.M. Amaro, A Systematic Review of Insole Sensor Technology: Recent Studies and Future Directions, Appl. Sci. 14 (2024), https://doi.org/10.3390/app14146085.

[23] B. Wei, Z. Ding, C. Yi, H. Guo, Z. Wang, J. Zhu, F. Jiang, A Novel sEMG-Based Gait Phase-Kinematics-Coupled Predictor and Its Interaction With Exoskeletons, Front. Neurorobot. 15 (2021) 1–14, https://doi.org/10.3389/fnbot.2021.704226.

[24] H. Park, S. Han, J. Sung, S. Hwang, I. Youn, S.J. Kim, Classification of gait phases based on a machine learning approach using muscle synergy, Front. Hum. Neurosci. 17 (2023) 1–12, https://doi.org/10.3389/fnhum.2023.1201935.

[25] C. Morbidoni, A. Cucchiarelli, S. Fioretti, F. Di Nardo, A deep learning approach to EMG-based classification of gait phases during level ground walking, Electron. 8 (2019), https://doi.org/10.3390/electronics8080894.

[26] J. Sun, G. Liu, Y. Sun, K. Lin, Z. Zhou, J. Cai, Application of Surface Electromyography in Exercise Fatigue: A Review, Front. Syst. Neurosci. 16 (2022) 1–8, https://doi.org/10.3389/fnsys.2022.893275.

[27] L. Piyathilaka, J.H. Sul, S. Dunu Arachchige, A. Jayawardena, D. Moratuwage, Advances in EMG Signal Processing and Pattern Recognition: Techniques, Challenges, and Emerging Applications, Electron. 15 (2026) 1–40, https://doi.org/10.3390/electronics15030590.

[28] S. Muceli, R. Merletti, Tutorial. Frequency analysis of the surface EMG signal: Best practices, J. Electromyogr. Kinesiol. 79 (2024) 102937, https://doi.org/10.1016/j.jelekin.2024.102937.

[29] T. Song, Z. Yan, S. Guo, Y. Li, X. Li, F. Xi, Review of sEMG for Robot Control: Techniques and Applications, Appl. Sci. 13 (2023), https://doi.org/10.3390/app13179546.

[30] N. Nazmi, S.I. Yamamoto, M.A.S. Rohim, E.F. Shair, Classification of Gait Phases by Using SVM and ANN Based on EMG Signals, 2024 IEEE Symp. Ind. Electron. Appl. ISIEA 2024 (2024) 1–6, https://doi.org/10.1109/ISIEA61920.2024.10607348.

[31] F. Di Nardo, C. Morbidoni, G. Mascia, F. Verdini, S. Fioretti, Intra-subject approach for gait-event prediction by neural network interpretation of EMG signals, Biomed. Eng. Online 19 (2020) 1–20, https://doi.org/10.1186/s12938-020-00803-1.

[32] F. Di Nardo, C. Morbidoni, A. Cucchiarelli, S. Fioretti, Influence of EMG-signal processing and experimental set-up on prediction of gait events by neural network, Biomed. Signal Process. Control 63 (2021) 102232, https://doi.org/10.1016/j.bspc.2020.102232.

[33] O. Coser, C. Tamantini, M. Tortora, L. Furia, R. Sicilia, L. Zollo, P. Soda, Deep learning for human locomotion analysis in lower-limb exoskeletons: a comparative study, Front. Comput. Sci. 7 (2025), https://doi.org/10.3389/fcomp.2025.1597143.

[34] T. Hortobágyi, T. Vetrovsky, J.S. Brach, M. van Haren, K. Volesky, R. Radaelli, P. Lopez, U. Granacher, Effects of Exercise Training on Muscle Quality in Older Individuals: A Systematic Scoping Review with Meta-Analyses, Sport. Med. - Open 9 (2023), https://doi.org/10.1186/s40798-023-00585-5.

[35] S. Bajelan, W.A. Sparrow, R. Begg, The ankle dorsiflexion kinetics demand to increase swing phase foot-ground clearance: implications for assistive device design and energy demands, J. Neuroeng. Rehabil. 21 (2024) 1–12, https://doi.org/10.1186/s12984-024-01394-x.

[36] B. Kibushi, Muscle coordination patterns in regulation of medial gastrocnemius activation during walking, Hum. Mov. Sci. 90 (2023) 103116, https://doi.org/10.1016/j.humov.2023.103116.

[37] H.J. Hermens, B. Freriks, R. Merletti, D. Stegeman, J. Blok, G. Rau, C. Disselhorst-Klug, G. Hägg, European Recommendations for Surface ElectroMyoGraphy Results of the SENIAM project, Roessingh Research and Development, 1999. https://books.google.com.my/books?id=w7HgOwAACAAJ.

[38] P. Natarajan, R.D. Fonseka, S. Kim, C. Betteridge, M. Maharaj, R.J. Mobbs, Analysing gait patterns in degenerative lumbar spine diseases: a literature review, J. Spine Surg. 8 (2022) 139–148, https://doi.org/10.21037/jss-21-91.

[39] F. Zhang, H. Huang, Source selection for real-time user intent recognition toward volitional control of artificial legs, IEEE J. Biomed. Heal. Informatics 17 (2013) 907–914, https://doi.org/10.1109/JBHI.2012.2236563.

[40] L. Chen, Y. Wang, X. Lu, J. Geng, T. Feng, Modelling and analysis of myoelectric networks based on lower limb motor synergies, IEEE Trans. Neural Syst. Rehabil. Eng. 33 (2025) 3910–3921, https://doi.org/10.1109/TNSRE.2025.3613408.

[41] H. Wang, Q. Tao, X. Zhang, Ensemble Learning Method for the Continuous Decoding of Hand Joint Angles, Sensors 24 (2024) 1–17, https://doi.org/10.3390/s24020660.

[42] C. Angelidou, P. Artemiadis, On Predicting Transitions to Compliant Surfaces in Human Gait via Neural and Kinematic Signals, IEEE Trans. Neural Syst. Rehabil. Eng. 31 (2023) 2214–2223, https://doi.org/10.1109/TNSRE.2023.3272355.

[43] K.H. Lee, J.Y. Min, S. Byun, Electromyogram-based classification of hand and finger gestures using artificial neural networks, Sensors 22 (2022), https://doi.org/10.3390/s22010225.

[44] A. Phinyomark, A. Nuidod, P. Phukpattaranont, C. Limsakul, Feature extraction and reduction of wavelet transform coefficients for EMG pattern classification, Elektron. Ir Elektrotechnika 122 (2012) 27–32, https://doi.org/10.5755/j01.eee.122.6.1816.

[45] P. T., V.K. Elumalai, B. E., Hand gesture classification framework leveraging the entropy features from sEMG signals and VMD augmented multi-class SVM, Expert Syst. Appl. 238 (2024) 121972, https://doi.org/10.1016/j.eswa.2023.121972.

[46] A.C. Tsai, J.J. Luh, T. Te Lin, A novel STFT-ranking feature of multi-channel EMG for motion pattern recognition, Expert Syst. Appl. 42 (2015) 3327–3341, https://doi.org/10.1016/j.eswa.2014.11.044.

[47] R. Guido, S. Ferrisi, D. Lofaro, D. Conforti, An Overview on the Advancements of Support Vector Machine Models in Healthcare Applications: A Review, Inf. 15 (2024), https://doi.org/10.3390/info15040235.

[48] P. Thanh Noi, M. Kappas, Comparison of Random Forest, k-Nearest Neighbor, and Support Vector Machine Classifiers for Land Cover Classification Using Sentinel-2 Imagery, Sensors (Basel). 18 (2017), https://doi.org/10.3390/s18010018.

[49] A. Mastropietro, C. Feldmann, J. Bajorath, Calculation of exact Shapley values for explaining support vector machine models using the radial basis function kernel, Sci. Rep. 13 (2023) 1–11, https://doi.org/10.1038/s41598-023-46930-2.

[50] M. Faisal, I. Khosa, A. Waris, S.O. Gilani, M.J. Khan, F. Hazzazi, M.A. Ijaz, Optimizing the impact of time domain segmentation techniques on upper limb EMG decoding using multimodal features, PLoS One 20 (2025) 1–21, https://doi.org/10.1371/journal.pone.0322580.

[51] S. Mendez-Moreno, L. Espinosa, O. Vital-Ochoa, R. Espinosa-Tanguma, J. Acosta-Elias, Segmentation of Surface Electromyography Signals: A Comparative Analysis of Time and Frequency Domain Methods, Comput. y Sist. 28 (2024) 1783–1797, https://doi.org/10.13053/CyS-28-4-4874.

[52] C.A. Byrne, D.T. O'Keeffe, A.E. Donnelly, G.M. Lyons, Effect of walking speed changes on tibialis anterior EMG during healthy gait for FES envelope design in drop foot correction, J. Electromyogr. Kinesiol. 17 (2007) 605–616, https://doi.org/10.1016/j.jelekin.2006.07.008.

[53] R.K. Mani, B. Nagaraj, Diagnosis and classification of neuromuscular disorders using Bi-LSTM optimized with grey Wolf optimizer for EMG signals, Sci. Rep. 15 (2025) 1–17, https://doi.org/10.1038/s41598-025-03766-2.

[54] T. Li, W. Ma, Y. Zheng, X. Fan, G. Yang, L. Wang, Z. Li, A survey on gait recognition against occlusion: taxonomy, dataset and methodology, PeerJ Comput. Sci. 10 (2024) 1–55, https://doi.org/10.7717/PEERJ-CS.2602.

[55] S. Wang, K. Hase, T. Funato, Computational prediction of muscle synergy using a finite element framework for a musculoskeletal model on lower limb, Front. Bioeng. Biotechnol. 11 (2023) 1–13, https://doi.org/10.3389/fbioe.2023.1130219.