ELECTRICAL ENGINEERING AND ENERGY

OPEN ACCESS

EMG-Based Multi-Class Gesture Recognition with Normalized Muscle Power Evaluation

Enes Halit Aydin[1,2] and Onder Aydemir[2,3]

[1] Dept. of Electrical and Energy, Karadeniz Technical University, Trabzon, Türkiye
[2] Medical Device and Production Application and Research Center, Karadeniz Technical University, Trabzon, Türkiye,
[3] Dept. of Electrical and Electronics Engineering, Karadeniz Technical University, Trabzon, Türkiye

Submitted: 27 October 2025
Accepted: 25 November 2025
Online First: 27 November 2025

Corresponding author
Enes Halit Aydın,
eneshalit@ktu.edu.tr

DOI:10.64470/elene.2025.13

Abstract: The analysis of musculoskeletal system movements using electromyography (EMG) signals is a fundamental requirement in fields such as prosthetic control, human-machine interaction, and neuromuscular rehabilitation. This study presents a comprehensive approach that not only evaluates movement recognition accuracy but also quantitatively assesses the level of muscle force required for each movement. In the study, the muscle loading profile of each hand movement was created using EMG signal energy normalized to the Rest state. Five different classifier models were compared under 5-fold cross-validation (CV) and Leave-One-Subject-Out (LOSO) protocols. The results showed that the Extension movement had the highest normalized power value and that classification accuracy reached its highest level with SVM-RBF (86.95%). Furthermore, Out-of-Bag (OOB) error analysis revealed that the model converged stably around 600–800 trees, while accuracy differences between individuals were attributed to physiological variations. The proposed framework offers a new evaluation perspective for both ergonomic task design and clinical performance monitoring by assessing gesture recognition performance alongside muscle strength requirements.

Keywords Classifier, Muscle, Power Analysis, sEMG.

1. Introduction

The analysis of musculoskeletal movements using surface electromyography (sEMG) provides valuable insight into muscle activation patterns by capturing temporal, amplitude, and spectral characteristics of electrical activity. sEMG has become a fundamental tool in prosthetic control, human–machine interaction, industrial ergonomics, sports biomechanics, and neuromuscular rehabilitation, where quantitative assessment of muscle function is essential for understanding movement performance.

Most studies in the literature have focused on gesture recognition, feature optimization, or improving classification accuracy from sEMG signals. However, the comparative evaluation of physiological parameters such as muscle power, energy expenditure, and activation load has often been overlooked. Importantly, high classification separability of a gesture does not necessarily imply greater muscular effort, since these two metrics depend on distinct physiological mechanisms. Therefore, evaluating both recognition accuracy and energy-related indices can provide a more comprehensive understanding of muscle efficiency, fatigue, and motor control, particularly in rehabilitation and ergonomic task design.

Recent advances in sEMG-based muscle force estimation have applied to a wide variety of modeling strategies. Explainable modeling methods, such as the Element Description Method (EDM) proposed by Sodenaga et al. (2025), achieve deep neural network–level accuracy while maintaining physical interpretability. Deep learning–based architectures, including the two-stage LR–LSTM model by Hua et al. (2022), capture temporal dynamics through recurrent neural networks. Ensemble methods such as Bagged Tree Ensemble (Hajian et al., 2022) and Random Forest (Dick et al., 2024) have been shown to enhance generalization and robustness across sessions. At the same time, event-driven and low-power systems have been explored to improve energy efficiency; for instance, Zanghieri et al. (2025) applied Leaky Integrate-and-Fire neuron encoding for multi-finger force estimation, achieving low latency and high energy efficiency.

Hybrid models combining biomechanical and computational frameworks have also been introduced: Esrafilian et al. (2022) developed an sEMG-driven musculoskeletal–finite element (MS–FE) coupling to predict tissue-level stresses, while Dutra et al. (2021) used state-space and Kalman filtering to achieve continuous grip force estimation. Additionally, Su et al. (2021) proposed a deep convolutional neural network for interaction force estimation in human–robot interaction, and Mathieu et al. (2023) established a unified taxonomy to reconcile differences in biomechanical modeling terminology. Finally, optimization-based frameworks such as the GA-optimized SVM, SVR, and RF models by Mokri et al. (2022) have demonstrated promising real-time potential for robotic rehabilitation control. Despite these advances, most of the aforementioned studies remain accuracy-oriented and do not explicitly address physiological interpretability, computational efficiency, or inter-subject variability. Comparative analyses of the energy cost and muscle activation intensity required for different gestures remain limited. Building on this gap, the present study introduces a power-based classification framework that jointly evaluates gesture recognition performance and normalized muscle activation energy.

Ten hand gestures from forty participants were analyzed using four sEMG channels. A Normalized EMG Power Index (NEPI) was computed relative to resting state to quantify muscle effort objectively.

Subsequently, time-, frequency-, and wavelet-domain features were extracted, and five classical machine learning algorithms—SVM-RBF, Random Forest, Bagged Trees, Linear Discriminant Analysis (LDA), and k-NN—were systematically compared.

The proposed approach provides a dual perspective that integrates classification accuracy with physiological interpretability, offering practical value for prosthetic control, rehabilitation monitoring, and energy-efficient movement analysis. In this respect, the study provides an innovative and comparable basis for both biometric identification and rehabilitation applications in sEMG-based systems. This study provides the following main contributions:

- EMG signals collected from 40 participants performing 10 distinct hand gestures were analyzed to achieve power-based movement recognition.
- Five classic machine learning algorithms were compared, and high accuracy was achieved. The study demonstrated that effective performance can be achieved without the need for deep learning.
- The Extension movement proved to be the most decisive gesture, requiring the highest muscle power expenditure and producing the most distinctive EMG activation pattern compared to other movements.
- The proposed sEMG-based multi-class gesture recognition method reliably grades relative muscle load on a motion-by-motion basis and provides a practical reference for physical therapists, doctors, and self-rehabilitation robots in terms of ergonomic design and motion selection/planning. Therefore, the proposed study is expected to contribute to areas such as selecting appropriate movements in rehabilitation protocols and determining energy-efficient gesture sets in sports and human-computer interaction applications.

Unlike previous studies that primarily focus on classification accuracy, this work integrates normalized

muscle power analysis into the classification framework, providing a dual assessment of recognition performance and physiological energy expenditure.

The subsequent sections of this study are organized as follows: Section 2 and 3 describe the dataset, experimental design, and applied methods. Section 4 presents experimental results. Finally, Section 5 concludes the study with discussions and gives ideas for future research.

2. Detailed Examination of Analyzed Data

The experimental dataset consists of 4-channel sEMG signals collected from 40 participants during the execution of 10 standardized hand gestures. The performed gestures are illustrated in Figure 1.

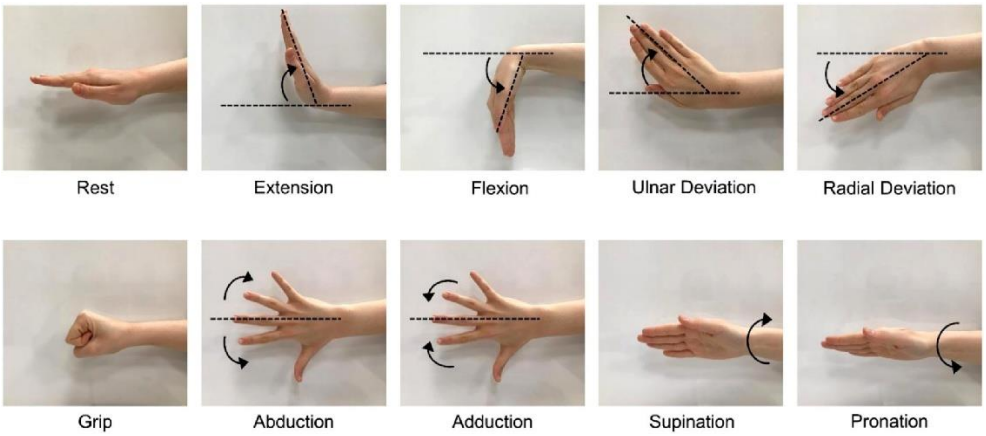

Figure 1 The set of ten hand gestures performed during the experimental recording (Ozdemir et al.,2022).

During data recording, a 4 s rest was taken before and after each gesture, followed by the relevant gesture. These movements were repeated in 5 cycles. One cycle lasted 104 seconds in total. A break of 30 seconds was taken between each cycle. Thus, the whole data collection process took 640 seconds in total The raw sEMG signals were recorded at 2000 Hz. The amplitude of the signals was between -10 and 10 mV and each hand movement was repeated in 4 channels in 5 cycles (Ozdemir et al.,2022). The data set for each participant contains information of size 4 × 1,280,000 (640 s data acquisition time, 2000 Hz sampling). Figure 2 shows the timeline of the data recording (for 1 cycle). The dataset used in this study was obtained from the publicly available hand gesture EMG dataset by Ozdemir et al. (2022)

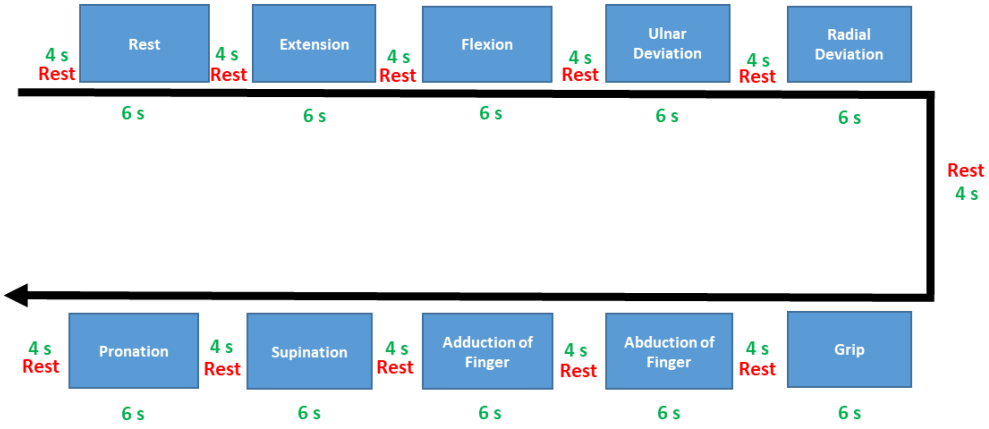

Figure 2 Timeline of a loop used in the data collection phase

When creating sEMG data sets, a series of preprocessing steps are applied to eliminate noise components present in the raw signals. Since sEMG signals inherently contain high noise, filtering processes are generally preferred to reduce the time-dependent non-stationary nature of the signal [23]. Significant amplitude and spectral differences may be observed between sEMG signals obtained from different individuals or from different sessions in the same individual. Therefore, normalization is necessary to reduce individual variability between data and ensure comparability. Normalization was performed by dividing each activation segment by the corresponding rest-phase RMS amplitude for each channel, converting amplitudes to a percentage of resting activity. In this study, Z-score normalization was applied to standardize the distributions of the obtained features. Furthermore, to shorten the model training time and increase prediction accuracy, only the sEMG segments where muscle activation occurred were considered. Signal segments where the muscle was at rest were removed from the data; in this context, 4-second rest segments were filtered out of the dataset, and only 6-second activation periods were included in the analysis.

In this study, the term Rest refers to two different contexts:

(i) Physiological rest periods — 4-second intervals recorded before and after each gesture, which were excluded from model training to focus on active muscle segments.

(ii) Rest gesture — a specific movement class representing a voluntary relaxation posture included in the classification process. This distinction is illustrated in Figures 1 and 2.

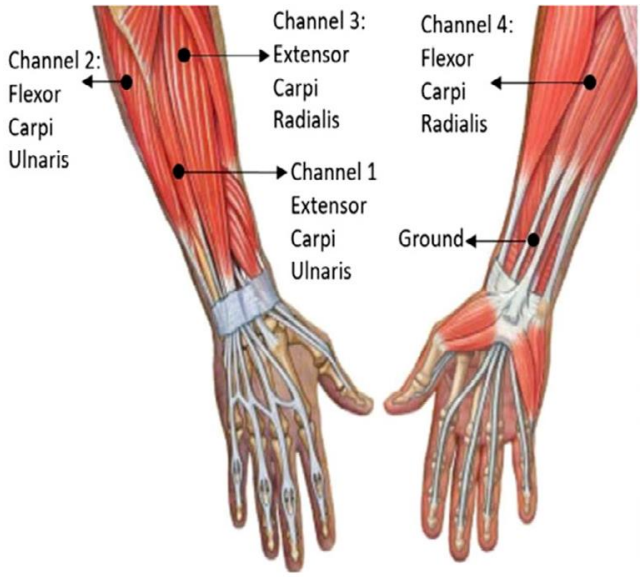

Figure 3 Location of sEMG channels and ground electrode showing posterior and anterior views of the right forearm (Ozdemir et al.,2022).

The electrode placement shown in Fig. 3 is critical for ensuring the reliability and reproducibility of sEMG recordings. Proper channel localization over the flexor and extensor muscle groups allows accurate detection of activation patterns associated with each hand gesture. This configuration minimizes cross-talk between adjacent muscles and provides a balanced representation of forearm muscle activity, which is essential for both power analysis and gesture classification accuracy.

3. Power Analysis and Classification Process

The main aim of this section is to describe the complete analysis workflow, including preprocessing, feature extraction, and classification. Each participant performed 10 distinct hand movements—rest, extension, flexion, ulnar deviation, radial deviation, grasp, abduction, adduction, supination, and pronation—each

repeated five times for consistent signal acquisition. All analyses were performed in MATLAB R2024b on a Windows 11 system (Intel i9, 64 GB RAM).

3.1 Preprocessing

Signals are normalized based on power levels, using the resting phase as a reference in each movement cycle. The resulting Normalized EMG Power Index indicates how much additional muscle energy each gesture requires relative to the resting state. sEMG signals were filtered using a 6th-order Butterworth bandpass filter (20–450 Hz) to remove power-line interference and motion artifacts, preserving the characteristic energy band of EMG activity. Normalization was applied during the training and test phase, and all preprocessing steps (filtering, normalization, segmentation) were independently performed for each of the four sEMG channels.

In this process, the normalized EMG power index was quantitatively calculated by dividing the mean squared amplitude of the EMG signal for each gesture by that of the resting phase. This computation expresses muscle activation intensity as a relative energy ratio, allowing power-based comparison across gestures.

3.2 Feature Extraction

To represent the multidimensional nature of sEMG signals, a rich feature set has been extracted in both time and frequency domains. All these features have been combined, and the characteristics of the four channels for each hand movement have been represented as a single vector. Thus, a feature matrix rich in both amplitude and spectral content has been created. The obtained features were normalized to the [0–1] range using the min–max scaling method to prevent numerical values of different magnitudes from adversely affecting the classification process.

3.2.1 Time Domain Features

The study, we used 5-time attributes commonly used in the literature: waveform length, mean absolute values, root mean square, zero crossings, slope sign change. Time-domain features quantify the amplitude and fluctuation characteristics of the signal. Below are the time-domain feature extraction formulas used in our study. $x = [x_1, x_2, ..., x_n]$ represents a signal sequence consisting of a total of N samples.

Feature 1 Waveform length *(WL)* represents the total amount of absolute deviation in the signal within a given analysis window, providing information about signal complexity and muscle activity intensity. WL indicates the complexity of the signal by calculating the total change in the waveform:

$$\mathrm{WL} = \sum_{i=1}^{N-1} |x_{i+1} - x_i| \quad (1)$$

Here, *x[i]* is the *i*th sample, *N* is the window length. This feature measures the amount of variation in the signal amplitude.

Feature 2 Mean Absolute Value (MAV) calculates the average of the signal's absolute values and provides an estimate of the overall amplitude level of muscle contraction.

$$MAV = \frac{1}{N} * \sum_{i=1}^{N} |x_i| \quad (2)$$

Feature 3 Root Mean Square (RMS) measures the effective magnitude of the signal and reflects the power of the muscle activation during movement. This feature reflects the signal power by emphasizing the effect of high-amplitude samples.

$$RMS = \sqrt{((1/N) * \sum_{i=1}^{N} x_i^2)} \quad (3)$$

Feature 4 Zero Crossing (ZC) counts how many times the signal crosses the zero axis, indicating the frequency of polarity changes related to muscle contraction dynamics. The threshold value (ε) is usually selected as a small number to prevent false crossings caused by noise.

$$ZC = \sum_{i=1}^{N-1}[indicator((x_i * x_{i+1} < 0) \cap |x_i - x_{i+1}| \geq \varepsilon)] \quad (4)$$

Here, the [indicator(·)] function returns 1 if the condition inside is true, and 0 otherwise.

Feature 5 Slope Sign Change (SSC) measures the number of times the slope of the signal changes sign, describing rapid variations in signal shape caused by muscle activation patterns. The threshold value (ε) is again used to ignore small oscillations.

$$SSC = \sum_{i=2}^{N-1}[indicator(\,(x_i - x_{i-1}) * (x_i - x_{i+1}) \geq \varepsilon\,)] \quad (5)$$

Here, the [indicator(·)] function returns 1 if the condition inside is true, and 0 otherwise. In the above formulas; N: The total number of samples (data points) in the signal, x_i: The i-th sample of the signal, ε (epsilon): A small positive threshold value determined to account for noise, Σ (Sigma Sum): Represents the sum symbol, ∩ (Intersection): Represents the "AND" logical operator (both conditions must be satisfied simultaneously)

3.2.2 Frequency Domain Features

In the frequency domain, the average frequency, median frequency, and spectral entropy were calculated using the Welch power spectral density method; this provided information about muscle fatigue and frequency band distribution. Additionally, the 20–450 Hz range was divided into four subbands (20–80, 80–150, 150–250, and 250–450 Hz), and the subband powers were determined for each band.

Feature 6 Welch Power Spectral Density estimates the spectral energy distribution of the signal using the Welch method, where the power spectral density P(f) characterizes the dominant frequency components of the EMG activity.

$$f_{mean} = \frac{\Sigma_f f P(f)}{\Sigma_f P(f)} \quad (6)$$

$$f_{med} = \int_0^{fmed} P(f)df = \frac{1}{2}\int_0^{fmax} P(f)df \quad (7)$$

$$H_{spec} = -\sum_{i=1}^{M} p_i log_{2(}p_{i)},\ p_i = \frac{P(f_i)}{\Sigma_f P(f)} \quad (8)$$

$$BP_{a,b} = \int_a^b P(f)df \quad (9)$$

Here, f_{mean} represents the average frequency, f_{med} represents the median frequency, H_{spec} represents the spectral entropy, and $BP_{a,b}$ represents the power energy in specific sub-bands.

3.2.3 Time-Frequency Domain Features

To increase temporal-frequency resolution, wavelet transform was used; as a result of four-level decomposition with the symlet4 wavelet, the energy content of the signal was calculated in the A4, D4, D3, D2, and D1 components, and the relative energy ratio of each band was determined. This method enabled the precise differentiation of energy distributions for different muscle activities.

Feature 7 Wavelet Energy Ratio is obtained by decomposing the sEMG signal *x(t)* into four levels using the sym4 wavelet and calculating the relative energy in each sub-band to capture frequency-dependent muscle activation. Since all sub-bands have different sampling rates, reconstruction was not performed; instead,

their energy ratios were directly calculated.
The energy of each component:

$$E_{Di} = \sum_{n=1}^{N} D_i(n)^2 \quad (10)$$

$$E_{A4} = \sum_{n=1}^{N} A_4(n)^2 \quad (11)$$

Relative energy coefficients:

$$E_{rel,i} = \frac{E_i}{\Sigma_j(E_j + E_{A4})} \quad (12)$$

These values represent the energy distribution of the signal in the time-frequency domain.

3.2.4 Hjorth Parameter

Feature 8 Hjorth Parameters consists of three statistical descriptors (activity, mobility, and complexity) that represent the variance, frequency content, and structural variability of the signal, respectively.

$$Activity = Var(x) \quad (13)$$

$$Mobility = \sqrt{\frac{Var(\dot{x})}{Var(x)}} \quad (14)$$

$$Comlexty = \frac{\sqrt{\frac{Var(\ddot{x})}{Var(\dot{x})}}}{Mobility} \quad (15)$$

These parameters characterize the distribution and structural complexity of the variance-based frequency components of the signal.

3.2.5 Auto Regressive Coefficients

Fourth-order autoregressive (AR(4)) coefficients were added to the feature set in order to capture the short-term autocorrelation structure of the sEMG signal.

Fourth-order AR model:

$$x_n = -\sum_{k=1}^{4} a_k x_{n-k} + e_n \quad (16)$$

Here, the a_k AR coefficients are the e_n white noise terms. The AR coefficients reflect the short-term autocorrelation structure of the signal.
For each channel, 25 features were computed:

- 5 wavelet energy ratios (A4, D4, D3, D2, D1)
- 5 time-domain features (RMS, MAV, ZC, SSC, WL)
- 3 Hjorth parameters (activity, mobility, complexity)
- 1 total bandpower (20–450 Hz)
- 3 Welch spectral descriptors (mean frequency, median frequency, spectral entropy)
- 4 sub-band powers (20–80, 80–150, 150–250, 250–450 Hz)
- 4 AR(4) coefficients

3.3 Feature Selection

A total of 100 features (25 per channel × 4 channels) were extracted. Among these, the mRMR algorithm selected the top 20 most informative features.

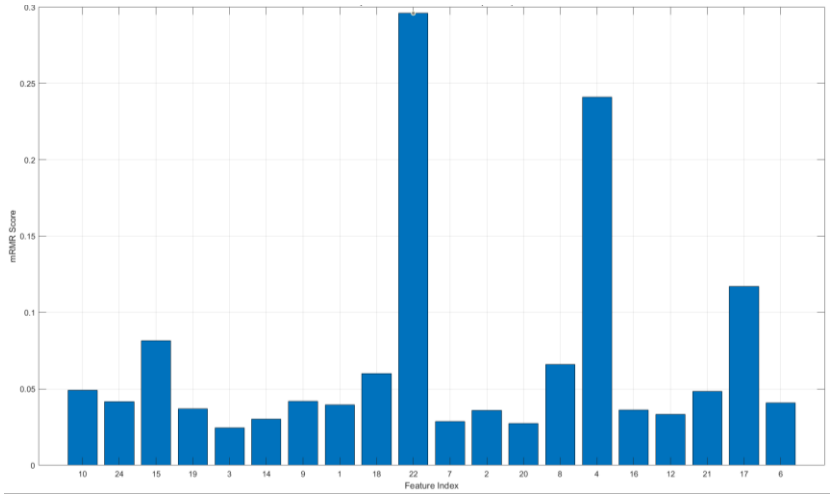

Figure 4 Top-20 most relevant features (mRMR)

Each feature was calculated independently for the four channels, and channel identifiers (e.g., RMS_CH1, RMS_CH2) were preserved to retain spatial information regarding flexor and extensor muscle activation. To reduce redundancy and identify the most discriminative attributes, the minimum redundancy–maximum relevance (mRMR) algorithm was applied to the complete feature set. This algorithm selects features with the highest mutual information with class labels while minimizing correlation among features, ensuring an optimal balance between relevance and uniqueness. Figure 4 illustrates the importance of the scores of the top 20 features selected by mRMR. The X-axis represents the original feature indices rather than sequential order. The analysis showed that time-domain features (RMS, MAV, WL, ZC, SSC) and frequency-domain features (spectral entropy and sub-band powers) provided the strongest discriminative power. Additionally, wavelet energy components and Hjorth mobility and complexity parameters contributed significantly to gesture separability, highlighting that the proposed feature set captures both temporal and spectral characteristics of muscle activation.

Overall, the selected features provide an information-rich representation of sEMG signals, improving both the discriminability and physiological interpretability of the classification model. The mRMR algorithm identified the 20 most informative features out of the 100 extracted features. All selected features originated from Channel 1, indicating that this channel carried the most discriminative information. The selected features included time-domain descriptors (RMS, MAV, WL, ZC, SSC), wavelet energy ratios, spectral measures, sub-band powers, and AR(4) coefficients.

Table 1 lists the Top-20 most informative features selected by mRMR together with their associated channels and feature categories. Interestingly, all selected features originated from Channel 1, indicating that this channel carried the most discriminative information for gesture classification. The selected subset includes time-domain characteristics (RMS, MAV, WL, ZC, SSC), wavelet energy ratios (A4–D3), Hjorth mobility, Welch spectral descriptors (mean/median frequency, entropy), sub-band powers (20–450 Hz), and AR(4) coefficients, demonstrating that both temporal and spectral attributes contributed to discriminative performance.

Table 1. Top-20 Features Selected by the mRMR Algorithm

Rank	Feature Index	Channel	Feature Category	Feature Name
1	10	CH1	Time-domain	Waveform Length (WL)
2	24	CH1	AR(4)	AR coefficient a3
3	15	CH1	Spectral (Welch)	Mean Frequency (fmean)
4	19	CH1	Sub-band Power	Bandpower 80–150 Hz
5	3	CH1	Wavelet Energy Ratio	D3 energy ratio
6	14	CH1	Bandpower (Total)	Total Bandpower 20–450 Hz
7	9	CH1	Time-domain	Slope Sign Change (SSC)
8	1	CH1	Wavelet Energy Ratio	A4 energy ratio
9	18	CH1	Sub-band Power	Bandpower 20–80 Hz
10	22	CH1	AR(4)	AR coefficient a1
11	7	CH1	Time-domain	Mean Absolute Value (MAV)
12	2	CH1	Wavelet Energy Ratio	D4 energy ratio
13	20	CH1	Sub-band Power	Bandpower 150–250 Hz
14	8	CH1	Time-domain	Zero Crossing (ZC)
15	4	CH1	Wavelet Energy Ratio	D2 energy ratio
16	16	CH1	Spectral (Welch)	Median Frequency (fmed)
17	12	CH1	Hjorth	Hjorth Mobility
18	21	CH1	Sub-band Power	Bandpower 250–450 Hz
19	17	CH1	Spectral (Welch)	Spectral Entropy
20	6	CH1	Time-domain	Root Mean Square (RMS)

4. Results and Discussion

The SVM-RBF model achieved the highest accuracy at 86.95%, whereas the Random Forest and Bagged Trees models reached 84.85% and 85.45%, showing comparable performance. The LDA model demonstrated the highest generalization ability in the LOSO scenario. The study calculated normalized average power values for each gesture and compared them relative to the resting state. This determined how much power each movement demanded from the muscles. The results obtained showed that the "Extension" movement had a significantly higher normalized power value compared to other gestures. This finding demonstrates that sEMG-based muscle loading analysis is an effective method for revealing movement-based energy requirements.

4.1 EMG Power Analysis

To comparatively examine the muscle activation level and energy requirements of each hand movement, the energy values obtained from sEMG signals were normalized relative to the resting state. This approach allowed the activation intensity of different muscle groups to be evaluated proportionally based on rest, thus revealing statistically significant physiological strength differences between movements. The normalized EMG power values not only showed how muscle loading changed depending on the type of movement but also supported the biomechanical consistency of the classification results.

Table 2. Relative muscle power requirements based on normalized EMG signal energy for each gesture.

No	GestureName	Normalized Mean Power
1	Rest	1.00
2	**Extension**	**45.76**
3	Flexion	4.79
4	Ulnar Deviation	8.77
5	Radial Deviation	6.39
6	Grip	1.93
7	Abduction	10.43
8	Adduction	3.79
9	Supination	2.14
10	Pronation	2.17

Each value represents the average normalized power computed over 40 participants × 5 repetitions = 200 activation segments per gesture. The normalized power values represent the mean EMG energy of each gesture relative to the resting state (Rest = 1.00). Higher values indicate greater muscle activation and energy expenditure. As summarized in Table 2, the Extension gesture exhibited the highest normalized mean EMG power (45.76), indicating that it required the greatest muscle activation among all movements. In contrast, gestures such as Grip and Supination showed relatively low power values, reflecting minimal muscle energy expenditure.

4.2 Classification Performance

Multiple classification algorithms were evaluated to determine the most effective model for movement detection. The Random Forest (RF) classifier, implemented as a set of 800 decision trees with a minimum leaf size of 1 and Gini impurity as the splitting criterion, achieved high accuracy under 5-fold cross-validation (CV). The number of predictors sampled at each split was determined as $\sqrt{p}$, where p represents the total number of features. Furthermore, RF provided feature importance rankings, revealing the relative contribution of each feature to gesture discrimination.

To evaluate model generalization, RF was compared with four additional algorithms: Support Vector Machine with Radial Basis Function kernel (SVM-RBF), Bagged Trees, Linear Discriminant Analysis (LDA), and k-Nearest Neighbors (k-NN). For the SVM-RBF classifier, the kernel scale (γ) and kernel constraint (C) parameters were optimized using a Bayesian hyperparameter search with 3-fold internal CV over the ranges $\gamma \in [10^{-3}, 10^{3}]$ and $C \in [0{,}1, 100]$. The Bagged Trees model was trained with 400 learning cycles. LDA used a linear discriminant kernel, and k-NN was configured with k = 7 neighbors and standardized features. Both 5-fold CV and LOSO protocols were applied for all classifiers. CV accuracy measured overall recognition performance, while LOSO accuracy evaluated the model's ability to generalize to unseen participants.

The performance of the proposed structure was comprehensively tested using multiple machine learning algorithms under five-fold cross-validation (5-fold CV) and LOSO protocols to evaluate both the accuracy and generalization capability of the model. This evaluation aimed to reveal the model's consistency on data obtained from different participants and the generalizability of the features. Cross-validation measured statistical stability within the dataset, while LOSO analysis enabled the evaluation of subject-independent performance. Thus, the proposed method was analyzed not only in terms of in-dataset accuracy but also in terms of subject-independence. Table 3 compares the performance of five different machine learning algorithms under both 5-fold cross-validation (CV) and leave-one-subject-out (LOSO) validation schemes.

Table 3. Average Performance of Different Classifiers

No	Classifier	CV mean (%)	CV Std. (%)	LOSO mean (%)	LOSO Std. (%)
1	**SVM-RBF**	**86.95**	1.87	62.35	11.63
2	Bagged Trees	85.45	0.99	66.75	13.36
3	Random Forest	84.85	0.63	66.80	13.37
4	LDA	77.25	2.65	**69.65**	**13.55**
5	kNN	71.35	2.67	50.40	11.22

According to the results, the SVM-RBF model achieved the highest accuracy with a CV mean of 86.95%, demonstrating that the RBF kernel effectively captures nonlinear distinctions in EMG power-based features. The Bagged Trees and Random Forest algorithms achieved accuracies of 85.45% and 84.85%, respectively, confirming that ensemble-based approaches are highly effective for power-driven gesture recognition. In the LOSO scenario, although overall accuracy decreased—as expected due to inter-subject variability—the LDA model exhibited the highest generalization ability with 69.65%, suggesting that its low-complexity structure and linear boundaries can better tolerate physiological variability across individuals. Overall, ensemble-based models (Bagged Trees and Random Forest) demonstrated strong short-term learning performance, while SVM-RBF provided the highest overall classification accuracy. In contrast, the generalization capability of LDA under LOSO validation highlights the importance of balancing accuracy and adaptability in real-world EMG-based recognition systems. Future studies should explore hybrid or ensemble model combinations to jointly optimize accuracy and subject-independent generalization performance.

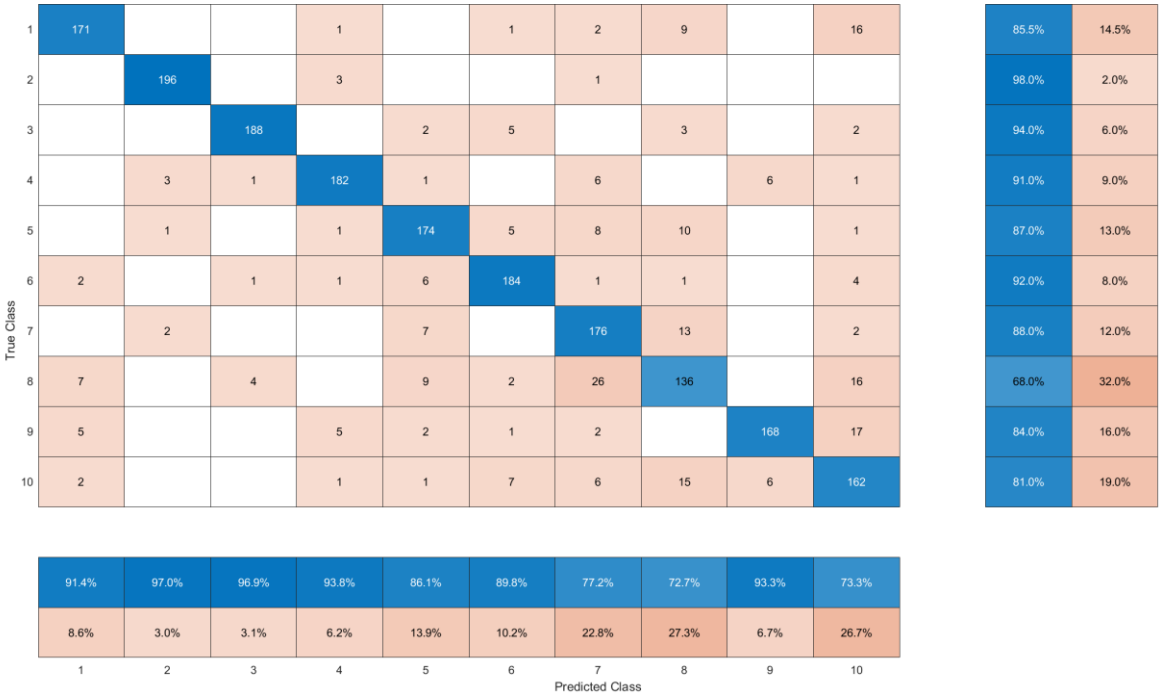

Figure 5 Normalized confusion matrix (SVM_RBF – 10 gestures)

Figure 5 shows the normalized confusion matrix obtained from the SVM_RBF classifier trained using the extracted sEMG features. The rows represent the actual movement classes, while the columns represent the predicted classes. The high values of the diagonal elements for most gestures indicate that the model has achieved high discriminative success. In particular, the accuracy rates for Extension, Flexion, and Abduction movements were observed to be significantly high.

Misclassifications are generally concentrated among movements with similar muscle synergies (e.g., Ulnar Deviation and Radial Deviation); this is due to the physiological overlap of muscle activation patterns. Overall, these results demonstrate that the proposed feature set provides sufficient discriminative power in the time–frequency domain and is reliable for multi-gesture recognition.

Figure 6 shows the distribution of F1-scores obtained for ten different hand movements. The F1-scores in Figure 6 correspond to the SVM-RBF classifier, which achieved the highest CV accuracy. The F1-score is a combined measure of precision and recall metrics, reflecting the balance between false positive and false negative rates for each class. The fact that the F1 scores are close to 1 for movements such as Extension and Abduction indicates that the model makes highly accurate predictions, while the relatively low scores for movements such as Grip and Pronation indicate uncertainties related to class overlaps. These findings confirm that the proposed community-based classifier demonstrates robust generalization capabilities by maintaining the sensitivity-specificity balance across different gestures. The figure shows that the highest F1 value was obtained for Extension, followed by Flexion and Supination. Ulnar Deviation, Radial Deviation, Grip, Abduction, and Pronation were in the middle range, while the lowest F1 values were observed in the Adduction and Rest classes. This distribution indicates that classes with strong and consistent patterns are recognized more reliably, whereas confusion persists in opposing pairs (e.g., Abduction–Adduction, Supination–Pronation) due to similar muscle synergies.

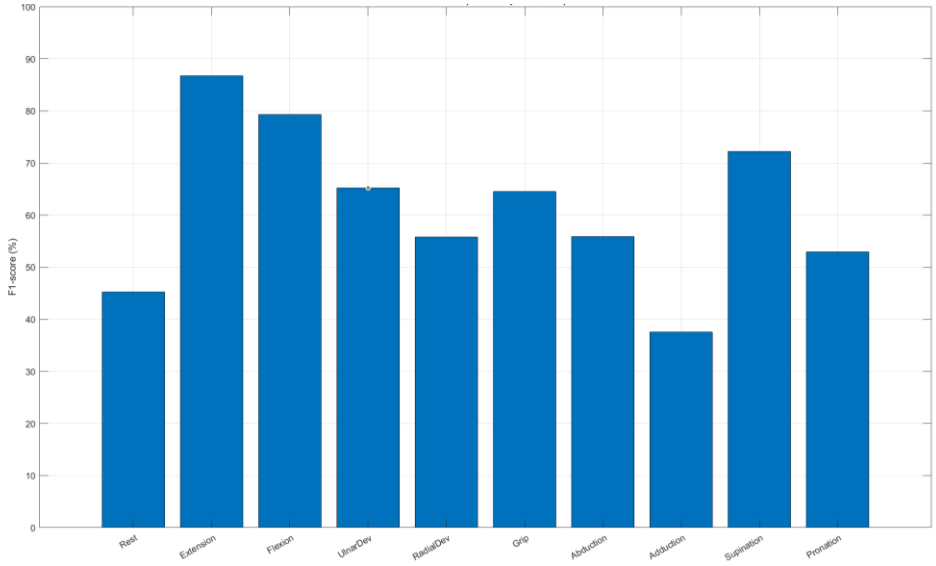

Figure 6 Per-class F1-score

Figure 7 shows the Out-of-Bag (OOB) error rate changing as each decision tree is added during the training of the SVM_RBF model. Although 20 mRMR features were used for classification comparison, Figure 7 illustrates the SVM_RBF model trained on the full feature set (100 features) to present its ensemble convergence behavior. OOB analysis is a built-in evaluation method that directly measures the model's generalization performance and is calculated by taking the average of prediction errors made on samples not used during training. The graph shows that the OOB error rate decreases rapidly as the number of trees increases and that the error curve stabilizes after approximately 600–800 trees. This indicates that the model has reached sufficient learning capacity and that adding more trees does not provide a meaningful improvement in accuracy. The obtained curve form clearly reveals the convergence behavior of the model and confirms that it does not show an overfitting tendency. Thus, it is concluded that the ensemble structure consisting of 800 trees provides an optimal balance in terms of both classification performance and computational cost.

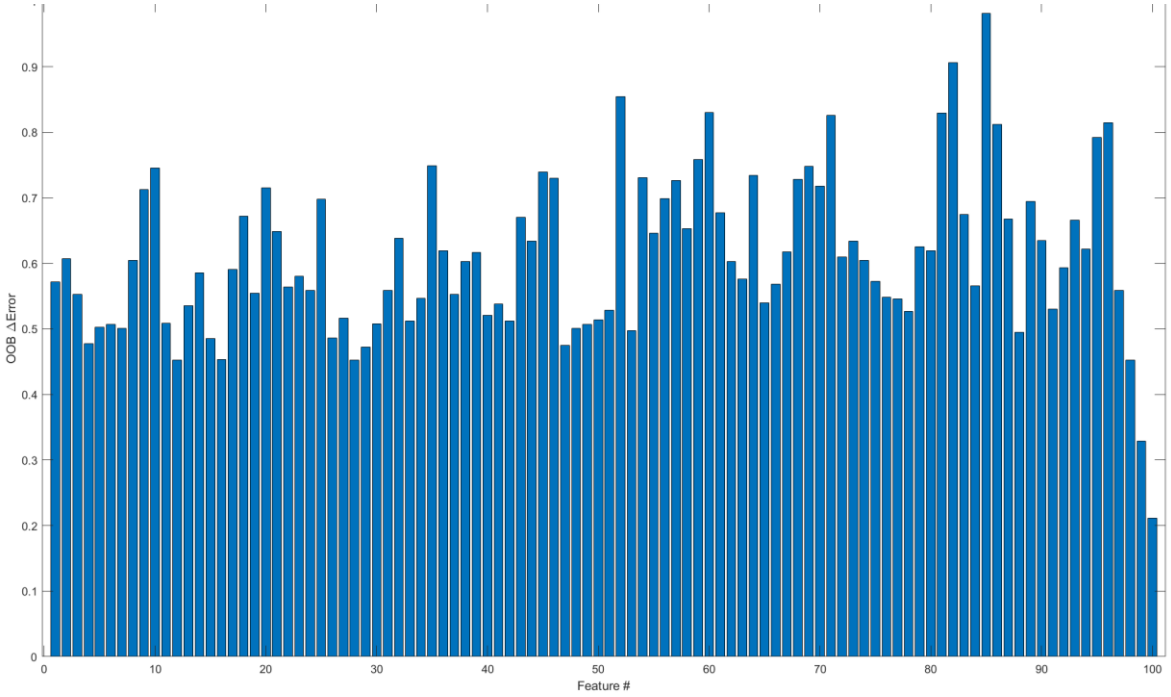

Figure 7 Model learning behavior

In Figure 8, the Extension movement exhibited the highest normalized power values showed the highest normalized power values with a high median and long upper tail. Abduction exhibits a wide intra-class variance along with a medium median. Flexion, UlnarDev, RadialDev, and Grip have lower and more compact power distributions. As expected, the Rest class remained at the baseline level. The power profiles indicate high discriminability potential, particularly for Extension, while the wide variance in Abduction suggests possible class boundary uncertainty.

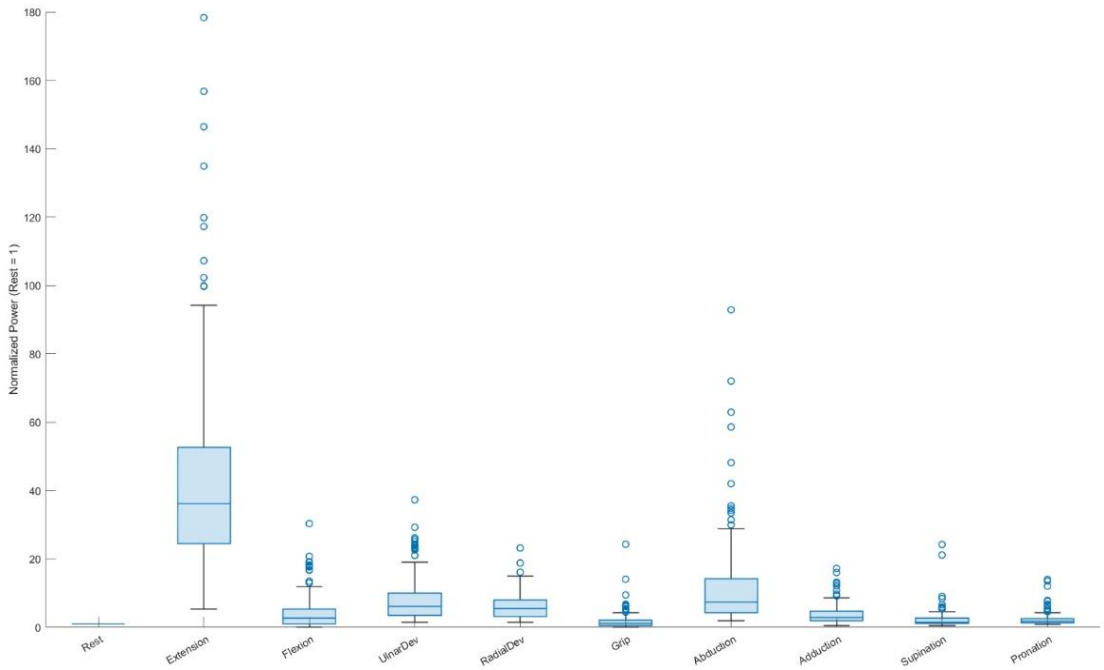

Figure 8 Gesture-wise normalized EMG power (Rest=1, boxplot)

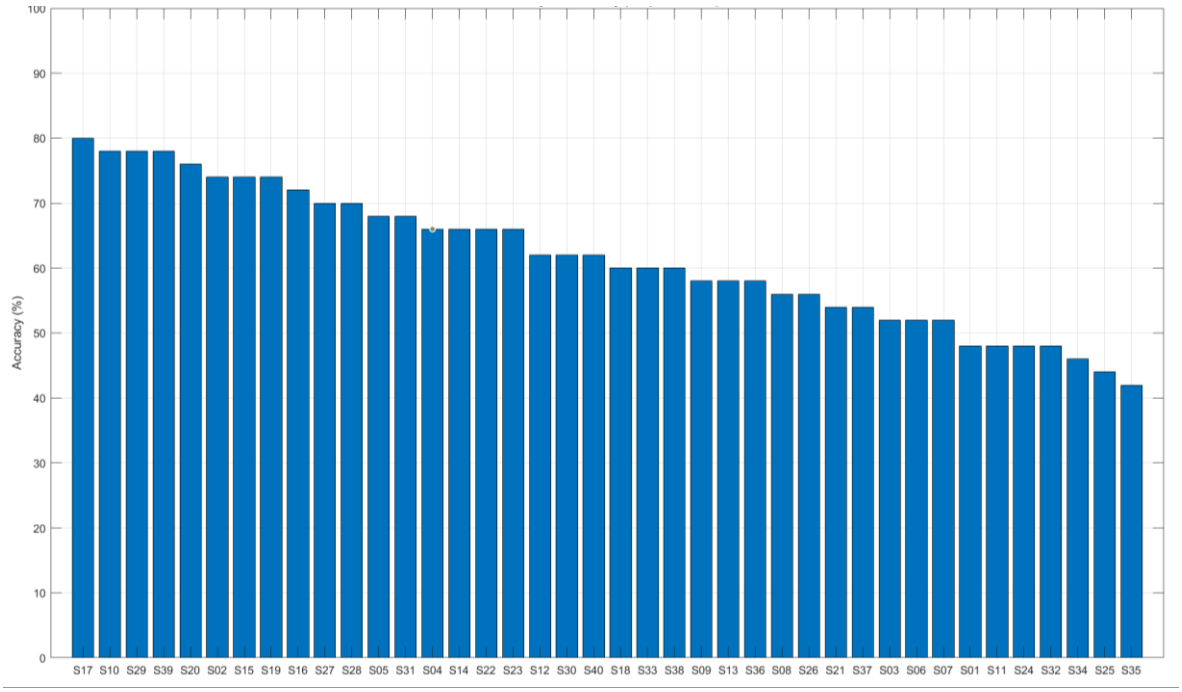

Figure 9 Per-subject Accuracy (CV)

Figure 9 presents the participant-wise accuracy values, revealing a wide dispersion in the range of ~42%–80%. A clustering is observed in a cluster of participants within the 70–80% range and a long tail extending toward 40–60% accuracy. This pattern indicates that electrode placement and skin impedance differences, along with individual muscle activation patterns, significantly affect classification performance, thus demonstrating that subject-specific variance is a dominant error component in subject-independent generalization. The LOSO approach is a protocol where each subject's data is sequentially separated as a test set and the model is trained on the data of all other participants, aiming to evaluate the model's individual-independent generalization capacity. As seen in the graph, accuracy rates show significant differences between participants. While the highest accuracy is achieved in the range of 78–80%, performance has been observed to drop to levels of 45–50% in some participants. This variation stems from physiological factors such as inter-individual muscle activation patterns, electrode placement, skin resistance, and biomechanical differences. In particular, the high accuracy rates of participants S17, S10, and S29 indicate that muscle activation during the movements of these individuals is more consistent and pronounced; in contrast, the low accuracy values of other participants suggest factors such as increased signal noise or low contraction strength.

Overall, these results demonstrate that the proposed model can tolerate moderate inter-subject variability; however, personalized calibration or adaptive normalization strategies could further enhance its robustness and generalization in future applications.

5. Conclusions

This study demonstrates that the proposed power-based framework can effectively capture both the physiological and computational aspects of muscle activity during different hand gestures. The normalized power index provided a quantitative means of comparing the relative energy demand of movements, revealing distinct activation patterns that reflect underlying biomechanical complexity. The findings indicate that gesture classification accuracy is influenced not only by algorithmic efficiency but also by individual differences such as muscle morphology, electrode placement, and skin impedance. These results highlight the importance of integrating physiological interpretability into performance evaluation rather than relying

solely on accuracy metrics. The proposed approach contributes to bridging the gap between signal-based classification and muscle-level energy assessment, offering a more holistic understanding of motor control in sEMG-based systems. Nevertheless, the study was conducted using data from 40 participants and four EMG channels, which may limit its generalization to high-density or dynamic recording conditions. Future work should focus on validating the method in larger, more heterogeneous datasets and exploring adaptive feature selection strategies to improve real-time robustness.

Declaration of Ethical Standards

As the authors of this study, we declare that we comply with all ethical standards.

Credit Authorship Contribution Statement

E.H. Aydın: Conceptualization, Methodology, Data curation, Software, Formal analysis, Investigation, Visualization, Writing – original draft, Writing – review & editing.

Ö. Aydemir: Supervision, Conceptualization, Validation, Methodology, Project administration, Writing – review & editing.

Declaration of Competing Interest

The authors declared that they have no conflict of interest.

Funding / Acknowledgements

This work was supported by The Scientific and Technological Research Council of Türkiye (TÜBİTAK) under Project No. 124E719.

Data Availability

The authors would like to thank the creators of the publicly available sEMG Hand Movement Dataset used in this study for sharing their data openly, which enabled the analysis and evaluation presented in this work. https://www.sciencedirect.com/science/article/pii/S2352340922001330

References

Dick, T. J., Tucker, K., Hug, F., Besomi, M., van Dieën, J. H., Enoka, R. M., & Hodges, P. W. (2024). Consensus for experimental design in electromyography (CEDE) project: Application of EMG to estimate muscle force. *Journal of Electromyography and Kinesiology*, 79, 102910.

Dutra, B., Silveira, A., & Pereira, A. (2021). Grasping force estimation using state-space model and Kalman filter. *Biomedical Signal Processing and Control*, 70, 103036.

Esrafilian, A., Stenroth, L., Mononen, M. E., Vartiainen, P., Tanska, P., Karjalainen, P. A., & Korhonen, R. K. (2022). An EMG-assisted muscle-force driven finite element analysis pipeline to investigate joint- and tissue-level mechanical responses in functional activities: Towards a rapid assessment toolbox. *IEEE Transactions on Biomedical Engineering*, 69(9), 2860–2871.

Hajian, G., Behinaein, B., Etemad, A., & Morin, E. (2022). Bagged tree ensemble modelling with feature selection for isometric EMG-based force estimation. *Biomedical Signal Processing and Control*, 78, 104012.

Hu, R., Chen, X., Zhang, H., Zhang, X., & Chen, X. (2022). A novel myoelectric control scheme supporting synchronous gesture recognition and muscle force estimation. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, 30, 1127–1138.

Hua, S., Wang, C., & Wu, X. (2022). A novel sEMG-based force estimation method using deep-learning algorithm. *Complex & Intelligent Systems*, 8(3), 1949–1961.

Mathieu, E., Crémoux, S., Duvivier, D., Amarantini, D., & Pudlo, P. (2023). Biomechanical modeling for the estimation of muscle forces: Toward a common language in biomechanics, medical engineering, and neurosciences. *Journal of NeuroEngineering and Rehabilitation*, 20(1), 130.

Mokri, C., Bamdad, M., & Abolghasemi, V. (2022). Muscle force estimation from lower-limb EMG signals using novel optimised machine-learning techniques. *Medical & Biological Engineering & Computing*, 60(3), 683–699.

Ozdemir, M. A., Kisa, D. H., Guren, O., & Akan, A. (2022). Dataset for multi-channel surface electromyography (sEMG) signals of hand gestures. *Data in brief*, 41, 107921.

Sitole, S. P., & Sup, F. C. (2023). Continuous prediction of human joint mechanics using EMG signals: A review of model-based and model-free approaches. *IEEE Transactions on Medical Robotics and Bionics*, 5(3), 528–546.

Sodenaga, D., Takeuchi, I., De Silva, D., & Katsura, S. (2025). Force estimation from surface-EMG using element description method. *IEEE Journal of Emerging and Selected Topics in Industrial Electronics*, 6(1), 447–454.

Zanghieri, M., Rapa, P. M., Orlandi, M., Donati, E., Benini, L., & Benatti, S. (2024). Event-based estimation of hand forces from high-density surface EMG on a parallel ultra-low-power microcontroller. *IEEE Sensors Journal*, 24(5), 8124–8135.