Evaluating the Effects of Feature set and Hyperparameter Optimization on sEMG-Based Gesture Recognition

Gabriel Molina de Lima[1], **André L. D. Rossi**[2], **José Jair Alves Mendes Junior**[1], **Daniel Prado Campos**[1], **Rafael Gomes Mantovani**[1]

[1] Federal University of Technology – Paraná (UTFPR), Apucarana, Paraná, Brazil

[2] São Paulo State University (Unesp), Bauru, São Paulo, Brazil

gabriell.2019@alunos.utfpr.edu.br, andre.rossi@unesp.br, {jjjunior,danielcampos,rafaelmantovani}@utfpr.edu.br

Abstract. *Gesture recognition using myoelectric signals (sEMG) is a powerful tool for Human-Machine Interfaces (HMIs). This study investigates the classification of sEMG signals with varying sets of gestures and predictive features using Bayesian Optimization for hyperparameter tuning of Machine Learning (ML) algorithms employed for this task. Experiments were conducted on the Ninapro DB2 and DB3 datasets, covering data from both intact and amputated individuals and exploring different ML algorithms. Even though Support Vector Machines (SVMs) showed the most notable improvement through hyperparameter tuning, the best experimental results were achieved by the Random Forest (RF) classifier on both datasets, with an average F-Score of* 0.853 *in DB2 and* 0.720 *in DB3, using only five gestures and nine features. Overall, a larger feature set enhanced signal representation, particularly in reduced gesture sets, while data from amputated individuals posed greater classification challenges.*

1. Introduction

Many human-machine interface applications have extensively employed gesture recognition due to its proven effectiveness [Bi et al. 2019]. Nevertheless, in control of prostheses and orthotics, gesture recognition still struggles to achieve highly accurate models capable of identifying a large number of movements, which is fundamental for the real-world use of such devices. Several datasets have been publicly released to support research on gesture recognition using sEMG signals, enabling the development and benchmarking of machine learning algorithms in a reproducible manner [Atzori et al. 2014, Geng et al. 2016].

In this context, many studies exploring Machine Learning (ML) tasks have been proposed to extract complex features from Electromyography (EMG) signals and identify movement patterns. Two alternatives are commonly explored: classical ML algorithms or Deep Learning (DL) [Qiao et al. 2024, Grattarola et al. 2025]. However, there is no standard for a definitive method to perform gesture classification and feature extraction from EMG signals [Jia et al. 2020].

Recent studies have shown that the performance of EMG-based gesture classification is highly influenced by the selection and dimensionality of the feature set [Sandoval-Espino et al. 2022]. In addition, the effectiveness of many ML algorithms that have been widely applied in this context, such as Support Vector Machine (SVM), Random Forest (RF), and K-Nearest Neighbors (kNN) [Atzori et al. 2014], can be affected by their hyperparameter values [Grattarola et al. 2025, Aviles et al. 2023].

Bayesian optimization techniques have recently demonstrated superior performance in tuning classifiers for biomedical signal tasks, including EMG-based recognition [Grattarola et al. 2025, Qiao et al. 2024].

This study investigates gesture classification using the NinaPro DB2 and DB3 datasets. We compare classical ML algorithms applied to two sets of handcrafted EMG features extracted from time, frequency, and wavelet domains. Additionally, we apply Bayesian optimization to automatically tune hyperparameters for each classifier, aiming to enhance performance and ensure fair comparison across models. The remainder of this paper is organized as follows: Section 2 presents the background on EMG signal processing and machine learning methods. Section 3 reviews related works. Section 4 describes the experimental methodology, including datasets, feature extraction, classification, and optimization. Section 5 presents the results and discussion.

2. Background

This section provides an overview of the main concepts behind sEMG acquisition and ML techniques employed for gesture recognition.

2.1. Surface Electromyography (sEMG)

Surface Electromyography (sEMG) is a non-invasive technique used to capture muscle biopotentials. It provides the raw data required to monitor muscle activity [Ghaffar Nia et al. 2023]. This information is used in diverse applications, notably in pattern recognition for prosthetic control [Huang and Chen 2019]. Understanding EMG acquisition techniques is essential for extracting relevant features for gesture recognition using sEMG signals.

sEMG signals have low peak-to-peak amplitudes, typically ranging from a few tens of microvolts to 1–2 mV during voluntary contractions. Their frequency content is mostly concentrated between 10 Hz and 500 Hz [Merletti and Cerone 2020]. Publicly available databases, such as Ninapro, are pivotal in advancing research in this field. Ninapro offers various datasets employing different signal acquisition methods. Ninapro datasets employ different acquisition methods. DB5 uses Thalmic Myo Armbands at 200 Hz, while DB4 uses Comet electrodes at 2 kHz. Sampling rate is critical, as it impacts signal accuracy and detail [Phinyomark et al. 2018].

Feature extraction is crucial for pattern recognition of sEMG signals. It involves uncovering hidden information about the signal to represent it through features capturing specific patterns [Krishnan and Athavale 2018]. This process often incorporates dimensionality reduction to produce a smaller set of representative data for precise signal description. Feature extraction techniques for biomedical signals fall into four categories: Time Domain (TD), Frequency Domain (FD), Time-Frequency Domain (TFD), and signal decomposition and sparse domain methods [Krishnan and Athavale 2018].

2.2. Machine Learning

Based on the features extracted from sEMG signals, it is possible to apply machine learning algorithms to identify patterns associated with gestures or other types of muscle activity. ML is a field of study dedicated to developing algorithms and models that enable computers to learn, make predictions, and/or make decisions based on data [Marsland 2015].

Depending on the nature of the problem to be solved, different ML algorithms can be adopted. The main categories include reinforcement learning, unsupervised learning, and supervised learning, with the latter being the most widely used in practical applications.

Supervised learning consists of training models using a labeled dataset, in which each instance is represented by a vector of attributes (features) and associated with a label or target. This type of task can be further divided into: i) classification, when the target attribute is categorical, and ii) regression, when the target attribute is continuous [Marsland 2015]. In the context of this work, which uses the Ninapro dataset, the problem is framed as a supervised classification task, in which the goal is to predict the performed gesture based on the features extracted from the sEMG signals.

3. Related Works

Gesture recognition using sEMG has been widely investigated due to its potential in human-machine interfaces, especially in the control of prosthetics and orthoses. Works based on classic ML algorithms, such as SVM, RF, and kNN, remain effective and continue to yield competitive classification performance in this context. In [Khan et al. 2021], the authors used the Cubic-SVM algorithm to classify four gestures: wrist flexion, wrist extension, hand at rest, and closed fist. Features were extracted in the spectral domain of the sEMG signals, achieving an accuracy of 0.989. [Challa et al. 2023] proposed a classifier combining RF and Logistic Regression (LR). They used eight time-domain features extracted from four sEMG channels to classify three gestures: flexing, lifting, and grasping an object. The model achieved average accuracies of 0.966 and 0.94, respectively.

In parallel, the use of DL, especially architectures such as Convolutional Neural Network (CNN) and Long Short Term Memory (LSTM), is being widely explored for its ability to learn complex features directly from sEMG signals automatically. [Huang and Chen 2019] combined spectrograms with CNN and LSTM networks to classify 50 finger and wrist movements. The study used the Ninapro database with 40 individuals. The method increased the average accuracy of basic gesture classification from 0.757 to 0.809, and the overall accuracy from 0.772 to 0.793. These results outperformed a reference method based on SVM with spectrograms, after PCA-based dimensionality reduction. [Ozdemir et al. 2022] used time-frequency images obtained by Hilbert-Huang Transform (HHT), Continuous Wavelet Transform (CWT), and Short-Time Fourier Transform (STFT), with pre-trained CNNs to classify seven specific gestures performed by 30 subjects. The best performance was obtained with HHT and the ResNet-50 architecture, achieving average accuracies of 0.937 (stratified cross-validation) and 0.944 (leave-one-subject-out cross-validation).

Some studies also explored hybrid approaches, combining manual features with those extracted by DL models. [Fajardo et al. 2021] combined features obtained by discrete analysis in the time-spectral domain Mean Absolute Value (MAV), Slope Sign Changes (SSC), peak frequencies, wavelet transform coefficients) with features extracted by CNNs, using a Multilayer Perceptron (MLP) classifier. The experimental results showed average accuracies of 0.815, 0.885, and 0.942 for the classification of 8, 6, and 5 classes of gestures, respectively.

A recent review explored works published between 2018 and 2024, aiming to identify gaps in this research area and evaluate the most commonly used strategies to solve this task [de Lima et al. 2024]. The authors identified some gaps that can be explored in future research, such as using other metrics besides accuracy, such as recall and F-score, especially in unbalanced datasets. In addition, there is a gap in applying parameter optimization techniques and developing models capable of classifying more gestures while maintaining satisfactory performance.

4. Methodology

This section describes the experiments performed to evaluate the impact of feature selection, gesture set size, and hyperparameter tuning on classical ML algorithms applied to gesture classification based on sEMG signals.[1].The main objective was to investigate how these different configurations can influence the performance of the gesture recognition task.

Figure 1 presents an overview of the experimental methodology adopted in this work, organized in five steps: (i) data acquisition from the Ninapro DB2 and DB3 datasets; (ii) signal preprocessing with band-pass filtering, normalization, temporal windowing, and definition of gesture sets (5 and 10 gestures); (iii) feature extraction in the time, frequency, and time-frequency domains, with selection of reduced (3) and full (9) feature sets; (iv) classification using ML algorithms with and without hyperparameter tuning via Bayesian Optimization (BO); and (v) performance assessment for different configurations using classification metrics and statistical tests to compare gesture sets, feature sets, and hyperparameter tuning. The following subsections provide a detailed explanation of each step in the methodology.

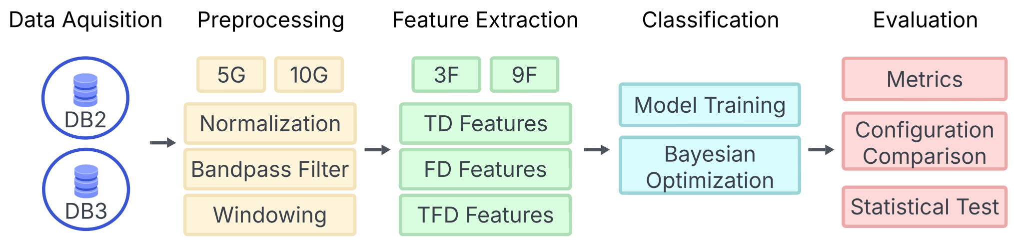

Figure 1. Pipeline adopted for the classification of gestures using sEMG.

4.1. Dataset

Experiments were conducted using the Ninapro[2] dataset collection due to its high-quality recordings for developing and evaluating gesture recognition models. It is also widely used in the literature and facilitates data comparison. The data collection is composed of ten (sub)datasets that sum more than 180 acquisitions from intact and individuals with transradial amputation, covering electromyography, kinematics, inertial signals, clinical, neurocognitive, and eye-hand coordination data [Atzori et al. 2012]. A sample sEMG signal from gestures 2 and 3 of the Ninapro DB2 dataset is shown in Figure 2. This figure illustrates the complexity of sEMG signals, highlighting the challenge of accurately classifying gestures due to signal variability.

[1] https://github.com/GabrielMolinaa/sEMG-Based-Gesture-Recognition-with-BO
[2] https://ninapro.hevs.ch/

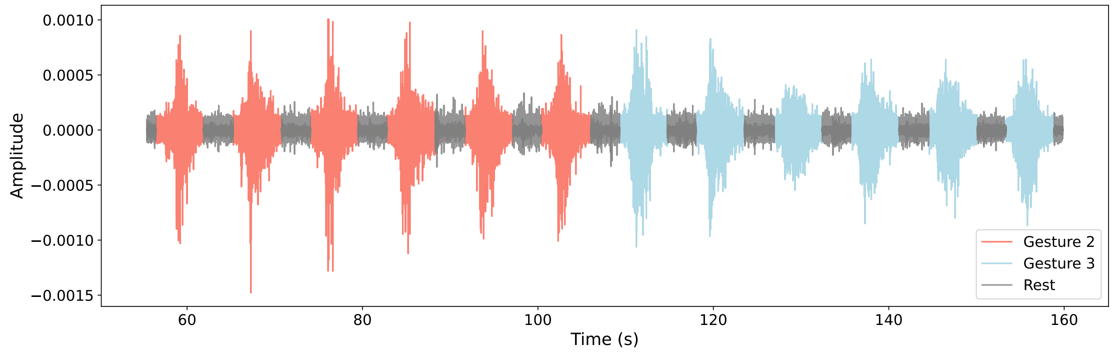

Figure 2. sEMG signal visualization for gestures 2 and 3 from the Ninapro DB2 dataset, including rest intervals.

From all the available datasets, we selected DB2 and DB3. They offer different sensor configurations and gesture types, allowing a comprehensive comparative analysis. DB2 contains signals from 40 intact individuals performing six repetitions of 49 distinct hand and wrist gestures, recorded at 2 kHz with 12 electrodes positioned around the forearm and upper arm [Atzori et al. 2014]. DB3 has data from 11 individuals with transradial amputation who performed six repetitions of the same gestures using the same electrode configuration and sampling rate. The used gestures were the Isometric, isotonic hand configurations and basic wrist movements (EX B DB2) and Grasping and functional movements (EX C DB3), respectively.

Table 1 describes the ten gestures selected in each data set used in the experiments. Two different versions of the datasets were defined, with 10 and 5 gestures, to directly compare the performance of induced models considering different levels of gestural variability. In the case of DB2, the gestures in the set with five classes correspond to positions $\{1, 2, 3, 4, 5\}$ in the table; in DB3, gestures with indexes $\{1, 4, 10, 14, 18\}$ were selected[3]. Different exercises were selected at each base to increase the diversity of movements analyzed, covering simpler and isolated gestures and functional and more complex movements.

Table 1. Gestures used in the experiments with the DB2 and DB3 datasets.

Dataset	Selected Gestures
DB2 (Exercise B)	1. Thumb up; 2. Extension of index and middle, flexion of the others; 3. Flexion of ring and little finger, extension of the others; 4. Thumb opposing base of little finger; 5. Abduction of all fingers; 6. Fingers flexed together in fist; 7. Pointing index; 8. Adduction of extended fingers; 9. Wrist supination (axis: middle finger); 10. Wrist pronation (axis: middle finger)
DB3 (Exercise C)	1. Large diameter grasp; 2. Small diameter grasp (power grip); 3. Fixed hook grasp; 4. Index finger extension grasp; 5. Medium wrap; 6. Ring grasp; 7. Prismatic four fingers grasp; 8. Stick grasp; 9. Writing tripod grasp; 10.Power sphere grasp 14.Prismatic pinch grasp 18.Parallel extension grasp

[3]These indexes follow the same indexes presented in Table 1

4.2. Preprocessing and Feature Extraction

Preprocessing was applied to the raw sEMG signals acquired from the DB2 and DB3 datasets, sampled at 2000 Hz. First, the signals were normalized to reduce inter-subject variability using the repetitions assigned to the training set. Then, a fourth-order Butterworth band-pass filter with a frequency range of 20–500 Hz was applied to remove noise and artifacts outside the relevant sEMG frequency range. After filtering, the data were segmented into overlapping windows of 500 samples (250 ms) with a step of 100 samples (50 ms), a technique known as temporal windowing. The repetitions used for training were $\{1, 3, 4, 6\}$, while the repetitions $\{2, 5\}$ were reserved for testing [Atzori et al. 2012]. The data from all individuals was concatenated, creating a single dataset for training and evaluation purposes. This approach allows the models to generalize across subjects, capturing variability in sEMG signals from different individuals.

Feature extraction was performed on each window of the sEMG signals from the 12 available channels. A set of features was selected based on their relevance for sEMG signal analysis and their use in prior works, particularly in hand gesture recognition tasks. The chosen features include Root Mean Square (RMS); the Hudgins feature set [Hudgins et al. 1993], composed of MAV, Zero Crossing (ZC), SSC, and Waveform Length (WL); spectral entropy, computed using the Welch periodogram [Welch 1967]; mean and median frequencies (MDF and MNF); and the energy of the sub-bands obtained from the marginal Discret Wavelet Transform (mDWT), computed as the sum of the squared wavelet coefficients at each decomposition level. For the DWT, we used the Daubechies 7 (db7) wavelet with a decomposition level of 3, following the methodology adopted in the Ninapro dataset [Atzori et al. 2014].

Two feature sets were then formed to evaluate the impact of feature reduction on classification accuracy. The first and more comprehensive set includes nine features per channel: MAV, RMS, ZC, SSC, WL, MDF, MNF, spectral entropy, and mDWT-based energy. The second, reduced set includes only three features per channel: RMS, spectral entropy, and mDWT-based energy. Feature extraction was performed separately for the training and testing sets.

4.3. Algorithms and Hyperparameter Tuning

Five widely consolidated ML algorithms in the literature were used for classification tasks with sEMG signals: RF, SVM, kNN, Logistic Regression (LR), and Linear Discriminant Analysis (LDA). The choice of these models is justified by the frequency with which they are applied in studies related to the classification of gestures using myoelectric signals [Nia et al. 2023].

Initially, all algorithms were trained using their default hyperparameters provided by the *scikit-learn* library, establishing a baseline for comparing the initial performance of the models. Subsequently, Bayesian Optimization (BO) was applied to tune the hyperparameters of each algorithm individually. This approach aims to evaluate the performance gain provided by optimization and analyze how each model responds to parameter tuning in different gesture and feature configurations. BO is a probabilistic model-based approach to global optimization of hyperparameters, which builds a surrogate model of the objective function (commonly using Gaussian processes) and selects hyperparameter configurations to evaluate based on acquisition functions that balance exploration and exploitation [Snoek et al. 2012]. The specific hyperparameter search spaces used for each model during BO are detailed in Table 2.

Table 2. Search spaces for hyperparameter tuning of each algorithm

Algorithm	Hyperparameters
SVM	C $\in [10^{-2}, 10^{1}]$, $\gamma \in [10^{-3}, 10^{3}]$ (log scale)
RF	n_estimators $\in [100, 600]$, max_depth $\in [3, 20]$
KNN	n_neighbors $\in [1, 30]$
LR	C $\in [10^{-2}, 100]$ (log scale)
LDA	solver $\in \{svd, lsqr, eigen\}$, shrinkage $\in \{auto, None\}$ (if solver is lsqr/eigen)

4.4. Evaluation Methodology

Two main metrics were used to evaluate the models: accuracy and macro F1-score. Choosing the macro average of the F-score allows a more balanced analysis of the models' performance, especially in scenarios with possible imbalances between classes. Statistical tests were applied to compare the performance between the models, verifying whether the differences observed in the metrics are statistically significant. The Friedman test was used for this analysis, followed by the Nemenyi post-hoc test, both recommended for comparisons between multiple classifiers in ML contexts [Demšar 2006].

All experiments were conducted using a global approach, in which training and testing sets were constructed by aggregating data from all individuals within each database. This setup assesses the model's ability to generalize across different subjects, rather than fitting to individual-specific signal patterns.

5. Results and Discussion

Figure 3 depicts the average F-Score values for each ML algorithm on DB2 and DB3 datasets under four experimental configurations: combining gesture sets (5G and 10G) and feature sets (3 and 9 features), with and without hyperparameter tuning. The upper panels depict the results with tuning, while the lower panels present the performances without tuning (baseline).

In DB2, comprising signals from intact individuals, the models generally achieved higher F-Scores across all setups. RF consistently led the rankings, reaching an F-Score of approximately 0.843 in the most favorable configuration (5G, 9 features, with tuning). kNN and SVM followed closely, showing strong performance particularly when the number of gestures was reduced and the full feature set was used. Linear models (LR and LDA) exhibited lower performance across the board, reflecting their limited capacity to capture the complex, multi-class distributions inherent to sEMG data.

In contrast, DB3, comprising signals from individuals with transradial amputation, presented a more challenging scenario. Performance drops were observed across all models and configurations. Even RF, while maintaining a leadership position, showed a maximum F-Score of around 0.72, significantly lower than its performance on DB2. SVM and kNN also displayed reduced effectiveness, whereas the linear models continued to underperform. These trends corroborate prior findings suggesting that sEMG signals from amputees tend to be more variable and less stable, challenging both feature extraction and classification processes.

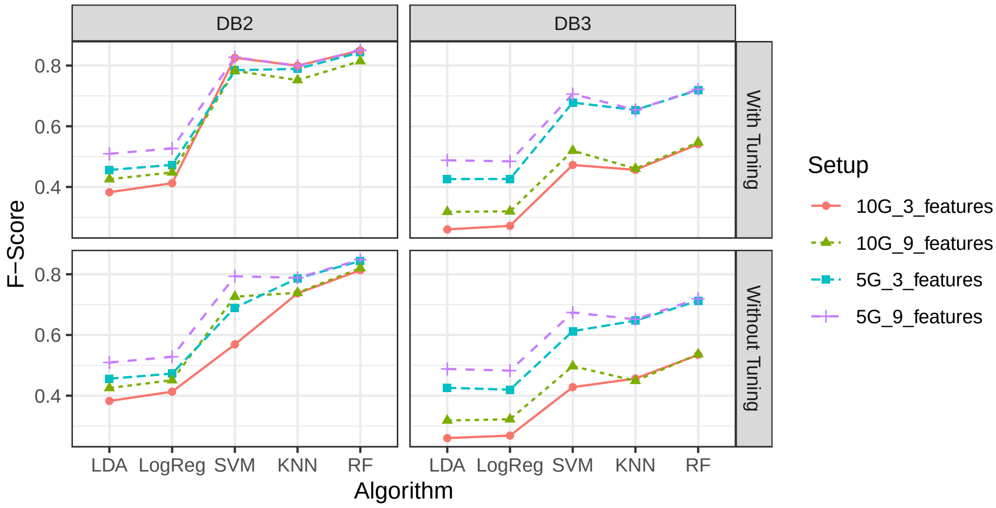

Figure 3. Average results inducing models under different number of gestures (5G, 10G) and features (3, 9).

5.1. Effect of Tuning and Feature Set Size

The results presented in Figure 3 indicate that hyperparameter tuning differently impacted the algorithms. The SVM algorithm was more sensitive to tuning, showing the most significant performance gains, especially on the DB2 dataset. For example, in the configuration with 10 gestures and 9 features, the F-Score of SVM increased from approximately 0.726 to 0.783. In DB3, considering the configuration with 9 features and 5 gestures, the F-Score improved from 0.674 to 0.706, a gain of 3.2 percentage points, while accuracy increased from 67.4% to 70.2%. On the other hand, although the RF algorithm achieved the best overall performance in almost all scenarios, it showed minor sensitivity to tuning. The variation in F-Score before and after optimization was minimal. These results suggest that RF already operates close to its maximum potential with default hyperparameters.

The complete set of features (9) consistently provided better results than the reduced set (3 features). This is due to the expanded set's ability to capture a more comprehensive range of properties of the myoelectric signal, including temporal aspects (RMS, MAV, WL), frequency components (MDF, MNF, spectrum), and time-frequency information (mDWT coefficients). The feature importance plot for RF in Figure 4 supports this observation, highlighting the relevance of temporal and time-frequency features for DB2 and DB3. In DB3, although gains were observed with tuning and feature set expansion, the impact was more limited due to the natural variability and instability of signals from amputees, which presents an additional challenge for modeling. Overall, the combination of fine-tuning and using the complete feature set proved to be the most effective configuration, especially for SVM and RF.

5.2. Challenging Classes and Misclassifications

Figure 5 compares each model's predictions and the ground truth across the samples for DB2 and DB3. The corresponding class color represents correct classifications, while black indicates misclassifications. DB2 heatmap shows that gestures 2 (Extension of index and middle, flexion of the others), 4 (Thumb opposing base of the little finger), 5 (Abduction of all fingers), and 8 (Adduction of extended fingers) are particularly prone to

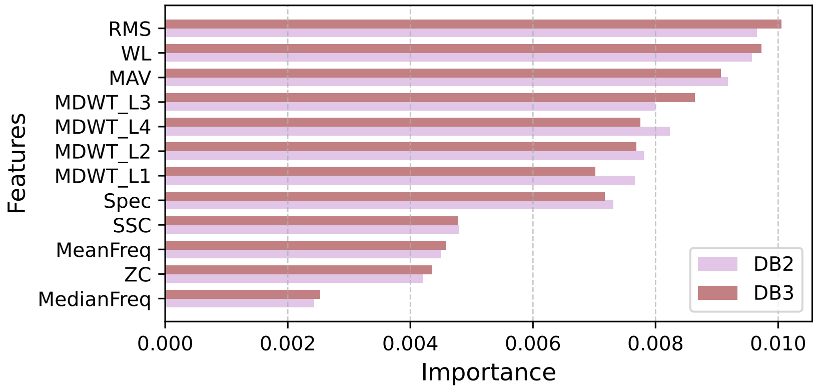

Figure 4. Feature importance for the Random Forest model in the 5-gesture and 9-feature configuration for DB2 and DB3 datasets.

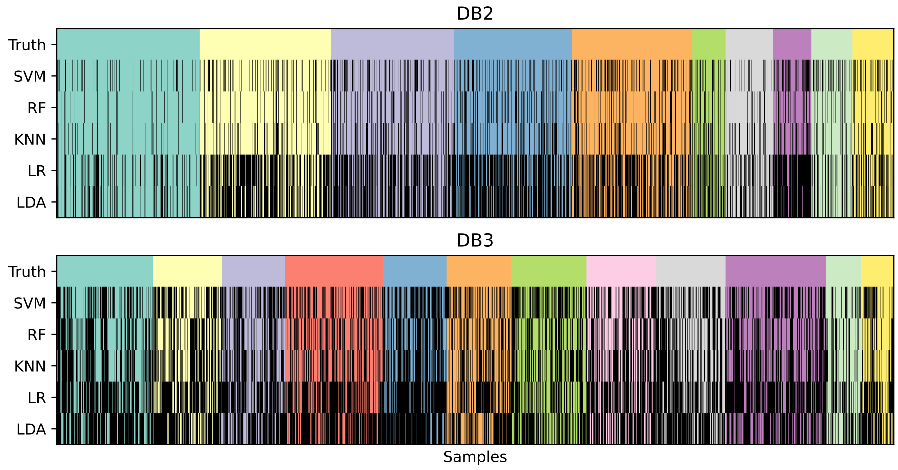

Figure 5. Misclassifications considering predictions obtained by the best induced models for each algorithm.

classification errors. The visual representation highlights scattered regions of misclassification concentrated in these gestures, with interruptions along the sequences, suggesting challenges in accurately identifying their distinct myoelectric patterns. These difficulties are likely related to biomechanical similarities, such as the simultaneous activation of adjacent muscle groups during finger abduction/adduction and thumb opposition, which may produce overlapping signal characteristics. Variations in muscle coordination and recruitment across repetitions can further increase classification complexity, even among non-amputated subjects.

In the DB3 dataset, composed of signals from individuals with transradial amputation, the heatmap reveals more pronounced challenges in classifying certain gestures. The black regions, indicative of errors, are most prominent in gestures 2 (Small diameter grasp), 5 (Medium wrap), 8 (Small diameter grasp), and 9 (Writing tripod grasp), with a consistent pattern of misclassifications for these grasp movements. Notably, three of these four problematic gestures involve different types of grasp, which inherently require complex coordination of multiple muscle groups. The observed difficulties reflect the inherent variability and instability of myoelectric signals from amputees, influenced by individual-specific factors such as residual limb physiology and adaptation strategies. This observation aligns with the literature, which highlights that EMG signals from amputees tend to be less stable and more variable, as different clinical parameters can affect these individuals' ability to reproduce the movements [Atzori et al. 2014].

5.3. Statistical Analysis

A statistical analysis was conducted using the Friedman-Nemenyi test to assess the models and experimental configurations across the DB2 and DB3 datasets. The Critical Difference (CD) diagrams (Figure 6) illustrate the average rankings and significant statistical differences among the models and setups.

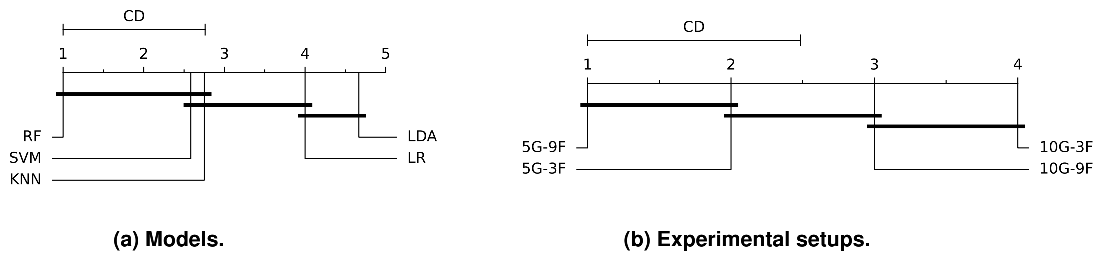

Figure 6. CD diagrams comparing (a) models and (b) experimental setups in the DB2 and DB3 datasets. Configurations connected by horizontal lines do not present statistically significant differences ($\alpha = 0.05$).

Figure 6a indicates that RF, SVM, and kNN did not present statistical differences between them. The linear models (LR and LDA) composed a second group with significantly inferior performance. This demonstrates the consistency and robustness of RF across the different setups, although the differences relative to SVM and kNN were not statistically significant.

Regarding the experimental setups, Figure 6b reveals that the 5G-9F configuration achieved the best average ranking, outperforming the others. However, its performance was not statistically better than the setup with 5G-3F, as a horizontal line connects both. In contrast, the configurations with 10 gestures (10G-3F and 10G-9F) showed lower rankings and significantly inferior performance compared to the 5-gesture setups. This suggests that reducing the number of gestures contributes more to stability and discriminative power than simply increasing the number of features. Overall, the combined statistical analysis highlights the clear advantage of the 5-gesture configurations and the relative similarity among the top-performing models, reinforcing the importance of a balanced experimental design for sEMG gesture classification tasks.

6. Conclusions

This study investigated the classification of myoelectric gestures using sEMG signals, considering different gesture and feature sets and evaluating the impact of hyperparameter tuning. The experiments revealed that Bayesian Optimization significantly enhanced the performance of SVM, while RF proved consistent but showed minor sensitivity to optimization. Using a more comprehensive feature set (9 features) effectively captured diverse information from the signals. Additionally, configurations with a smaller number of gestures (5G) were more stable, especially in the scenario involving amputated individuals (DB3), which presented greater signal variability.

As a continuation, future research could explore the effects of Bayesian Optimization in deep learning models, as this study focused exclusively on classical ML algorithms. Moreover, it is recommended to investigate further the interaction between the number of gestures and the quantity of extracted features, applying the methodology to other sEMG datasets or in contexts with more complex and varied gestures to assess the observed patterns.

Acknowledgments

This work is supported by Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq), grant 406417/2022-9.

References

Atzori, M., Gijsberts, A., Castellini, C., Caputo, B., Hager, A.-G. M., Elsig, S., Giatsidis, G., Bassetto, F., and Müller, H. (2014). Electromyography data for non-invasive naturally-controlled robotic hand prostheses. *Scientific Data*, 1.

Atzori, M., Gijsberts, A., Heynen, S., Hager, A.-G. M., Deriaz, O., van der Smagt, P., Castellini, C., Caputo, B., and Müller, H. (2012). Building the ninapro database: A resource for the biorobotics community. In *2012 4th IEEE RAS & EMBS International Conference on Biomedical Robotics and Biomechatronics (BioRob)*, pages 1258–1265.

Aviles, M., Rodríguez-Reséndiz, J., and Ibrahimi, D. (2023). Optimizing emg classification through metaheuristic algorithms. *Technologies*, 11(4).

Bi, L., Feleke, A. G., and Guan, C. (2019). A review on emg-based motor intention prediction of continuous human upper limb motion for human-robot collaboration. *Biomedical Signal Processing and Control*, 51:113–127.

Challa, K., AlHmoud, I. W., Kamrul Islam, A. K. M., and Gokaraju, B. (2023). Emg-based hand gesture recognition using individual sensors on different muscle groups. In *2023 IEEE Applied Imagery Pattern Recognition Workshop (AIPR)*, pages 1–4.

de Lima, G., Campos, D., and Mantovani, R. (2024). A review on the recent use of machine learning for gesture recognition using myoelectric signals. In *Anais do XXI Encontro Nacional de Inteligência Artificial e Computacional*, pages 180–191, Porto Alegre, RS, Brasil. SBC.

Demšar, J. (2006). Statistical comparisons of classifiers over multiple data sets. *The Journal of Machine Learning Research*, 7:1–30.

Fajardo, J. M., Gomez, O., and Prieto, F. (2021). Emg hand gesture classification using handcrafted and deep features. *Biomedical Signal Processing and Control*, 63:102210.

Geng, W., Du, Y., Jin, W., Wei, W., Hu, Y., and Li, J. (2016). Gesture recognition by instantaneous surface emg images. *Scientific Reports*, 6(1):36571.

Ghaffar Nia, N., Kaplanoglu, E., and Nasab, A. (2023). Emg-based hand gestures classification using machine learning algorithms. In *SoutheastCon 2023*, pages 787–792.

Grattarola, A., Mora, M. C., Cerdá-Boluda, J., and Ortiz, J. V. G. (2025). Grasp pattern recognition using surface electromyography signals and bayesian-optimized support vector machines for low-cost hand prostheses. *Applied Sciences*, 15(3).

Huang, D. and Chen, B. (2019). Surface emg decoding for hand gestures based on spectrogram and cnn-lstm. In *2019 2nd China Symposium on Cognitive Computing and Hybrid Intelligence (CCHI)*, pages 123–126.

Hudgins, B., Parker, P., and Scott, R. (1993). A new strategy for multifunction myoelectric control. *IEEE Transactions on Biomedical Engineering*, 40(1):82–94.

Jia, G., Lam, H.-K., Liao, J., and Wang, R. (2020). Classification of electromyographic hand gesture signals using machine learning techniques. *Neurocomputing*, 401:236–248.

Khan, M. U., Khan, H., Muneeb, M., Abbasi, Z., Abbasi, U. B., and Baloch, N. K. (2021). Supervised machine learning based fast hand gesture recognition and classification using electromyography (emg) signals. In *2021 International Conference on Applied and Engineering Mathematics (ICAEM)*, pages 81–86.

Krishnan, S. and Athavale, Y. (2018). Trends in biomedical signal feature extraction. *Biomedical Signal Processing and Control*, 43:41–63.

Marsland, S. (2015). *Machine learning: an algorithmic perspective*. CRC press.

Merletti, R. and Cerone, G. L. (2020). Tutorial. surface EMG detection, conditioning and pre-processing: Best practices. *J Electromyogr Kinesiol*, 54:102440.

Nia, N. G., Kaplanoglu, E., and Nasab, A. (2023). Emg-based hand gestures classification using machine learning algorithms. In *2023 IEEE SoutheastCon*, pages 787–792.

Ozdemir, M. A., Kisa, D. H., Guren, O., and Akan, A. (2022). Hand gesture classification using time–frequency images and transfer learning based on cnn. *Biomedical Signal Processing and Control*, 77:103787.

Phinyomark, A., N. Khushaba, R., and Scheme, E. (2018). Feature extraction and selection for myoelectric control based on wearable emg sensors. *Sensors*, 18(5).

Qiao, X., Chen, Z., and Liang, S. (2024). semg-based fine-grained gesture recognition via improved lightgbm model.

Sandoval-Espino, J. A., Zamudio-Lara, A., Marbán-Salgado, J. A., Escobedo-Alatorre, J. J., Palillero-Sandoval, O., and Velásquez-Aguilar, J. G. (2022). Selection of the best set of features for semg-based hand gesture recognition applying a cnn architecture. *Sensors*, 22(13).

Snoek, J., Larochelle, H., and Adams, R. P. (2012). Practical bayesian optimization of machine learning algorithms. In Pereira, F., Burges, C., Bottou, L., and Weinberger, K., editors, *Advances in Neural Information Processing Systems*, volume 25. Curran Associates, Inc.

Welch, P. (1967). The use of fast fourier transform for the estimation of power spectra: A method based on time averaging over short, modified periodograms. *IEEE Transactions on Audio and Electroacoustics*, 15(2):70–73.