Politecnico di Torino

Politecnico di Torino

Corso di Laurea Magistrale in Ingegneria Biomedica

A.a. 2025/2026

Graduation Session Marzo 2026

EMG Signal Deconvolution as Preprocessing for Enhanced Hand Gesture Recognition

Supervisor:
Luca Mesin

Candidate:
Marcelo C. Venturini

Abstract

This thesis investigates a computationally efficient strategy to improve Surface Electromyography (sEMG)-based hand gesture recognition under constraints typical of wearable myoelectric interfaces. In particular, it evaluates whether single-channel Electromyography (EMG) deconvolution can yield robust and discriminative feature representations for wrist and hand gesture classification using few electrodes.

This work proposes an online-compatible single-channel deconvolution pipeline that operates independently on each channel and produces a Cumulative Weighted Firing (CWF) pattern, avoiding the computational burden of multi-channel decomposition. Offline experiments are conducted on a publicly available dataset comprising recordings from 10 subjects acquired with a 64-channel array and annotated into six wrist/hand gestures plus rest; to emulate a practical low-density configuration, the analysis is restricted to four Single Differential (SD) channels obtained from eight monopolar electrodes. Building on the deconvolved representation, the thesis introduces features that summarize the temporal organization and sparsity of the estimated CWF (e.g., synchronization, spike rate, Gini index, and entropy-based descriptors), and compares them against standard amplitude/-complexity features extracted from the raw signal. A hybrid filter–wrapper feature selection approach is adopted to obtain compact models while accounting for reliability with confidence-based abstention and inference time. Two classifiers are considered, Linear Discriminant Analysis (LDA) and Non-Linear Logistic Regression (NLR).

In the offline evaluation performed on held-out data from the public dataset, deconvolution based descriptors improve reliability compared to raw features. With LDA, the macro F1-score increases from $97.21 \pm 0.91\%$ (raw-only) to $98.54 \pm 0.55\%$ (deconvolved-only) and $98.71 \pm 0.46\%$ (combined). With NLR, the deconvolved-only set achieves the highest accuracy ($99.33 \pm 0.74\%$) and macro F1-score ($99.07 \pm 0.78\%$). The complete pipeline is then deployed in a task-based online protocol, implemented in a custom real-time GUI, on five volunteers performing an extended gesture set; the best online performance is obtained with LDA and features extracted solely from the deconvolved signal (macro F1-score $96.73 \pm 0.44\%$, accuracy $96.63 \pm 0.58\%$, abstention rate $15.64 \pm 1.62\%$).

Table of Contents

List of Figures VI

Acronyms XII

1 Introduction 1
- 1.1 Context and motivation . . . 1
- 1.2 Goal . . . 2
- 1.3 Thesis structure . . . 2

2 Physiological and Anatomical Background 4
- 2.1 Skeletal Muscle and Motor Units . . . 4
  - 2.1.1 Skeletal Muscle Structure . . . 4
  - 2.1.2 Excitability and Action Potential Propagation . . . 7
  - 2.1.3 Motor Units and Recruitment . . . 7
  - 2.1.4 Muscles Involved in the Studied Gestures . . . 9
- 2.2 EMG Signal Generation . . . 12
  - 2.2.1 Volume Conduction and Spatial Filtering . . . 12
  - 2.2.2 Motor Unit Action Potential Propagation . . . 16
  - 2.2.3 Interference Pattern Formation . . . 18

3 State of the Art 25
- 3.1 Introduction . . . 25
- 3.2 Surface EMG for Hand Gesture Recognition . . . 25
  - 3.2.1 Applications and Challenges . . . 25
  - 3.2.2 High Abandonment Rates of Prostheses . . . 26
- 3.3 Pattern Recognition Pipeline . . . 26
  - 3.3.1 Standard Classification Pipeline . . . 26
- 3.4 Classifiers . . . 27
  - 3.4.1 Linear Discriminant Analysis (LDA) . . . 27
  - 3.4.2 Support Vector Machine (SVM) . . . 28
  - 3.4.3 Non-Linear Logistic Regression (NLR) . . . 28

3.4.4 Deep Learning Techniques . . . 29
3.5 Abstention Rate and Delay Issues . . . 30
3.5.1 High Abstention Rate: Advantages and Challenges . . . 30
3.5.2 Offline vs Online Classification Performance . . . 31
3.5.3 Impact of Muscle Fatigue on Classification Performance . . . 31
3.6 Electrode Configurations . . . 33
3.6.1 High-Density vs. Low-Density EMG . . . 33
3.7 High-Density EMG Decomposition . . . 37
3.7.1 Principles and Methodology . . . 37
3.7.2 Challenges and Limitations . . . 38
3.8 Conclusion . . . 40

4 Single-Channel EMG Deconvolution Method 41
4.1 Signal Model and Kernel Definition . . . 41
4.2 Inverse Problem Formulation and Regularization . . . 43
4.2.1 L1 Formulation and IRLS Optimization . . . 44
4.3 Online Implementation of the Deconvolution Algorithm . . . 46
4.4 Assumptions and Limitations . . . 48

5 EMG Data Acquisition and Feature Extraction 50
5.1 Participants and Experimental Protocol . . . 50
5.2 EMG Acquisition Setup . . . 51
5.2.1 Signal Quality and Preprocessing . . . 52
5.3 Muscles and Electrode Placement . . . 52
5.4 Feature Extraction . . . 54
5.4.1 Features Common to Both Raw and Deconvolved Signals . . . 55
5.4.2 Features Exclusive to Raw EMG Signals . . . 59
5.4.3 Features Exclusive to Deconvolved EMG Signals . . . 60
5.4.4 Summary . . . 62

6 Feature Selection for Gesture Classification 64
6.1 Feature Selection Overview . . . 64
6.1.1 Motivation and Challenges . . . 65
6.1.2 Feature Selection Methodologies . . . 66
6.2 Filter-Based Feature Selection . . . 67
6.2.1 Data Preprocessing for Feature Selection . . . 67
6.2.2 Pearson Correlation Analysis . . . 68
6.2.3 One-Way Analysis of Variance (ANOVA) . . . 69
6.2.4 Construction of Feature Sets . . . 70
6.3 Wrapper-Based Feature Selection Using Genetic Algorithm . . . 71
6.3.1 Genetic Algorithm Fundamentals . . . 71

6.3.2 Chromosome Encoding 72
6.3.3 Fitness Function 73
6.3.4 Validation Strategy Within GA 76
6.3.5 Genetic Operators 77
6.3.6 GA Hyperparameters 79

7 Results: Feature Selection and Classification Performance 81
7.1 Filter Stage Feature Selection Results 81
7.1.1 Raw Signal Feature Set (FS-Raw) 81
7.1.2 Deconvolved Signal Feature Set (FS-Dec) 84
7.1.3 Combined Feature Set (FS-Combined) 87
7.1.4 Filtered Feature Set Composition 90
7.2 Genetic Algorithm Wrapper Optimization Results 91
7.2.1 Linear Discriminant Analysis Results 91
7.2.2 Non-linear Logistic Regression Results 95
7.3 Classification Performance on Test Set 98
7.3.1 Linear Discriminant Analysis 98
7.3.2 Non-linear Logistic Regression 106
7.4 Online Gesture Classification 112
7.4.1 Online Experimental Setup 112
7.4.2 Real-time Processing Pipeline 113
7.4.3 Online Results 113

8 Conclusions and Future Work 118
8.1 Conclusions 118
8.2 Future Work 120

Bibliography 122

List of Figures

2.1 Three Connective Tissue Layers [12]. 5

2.2 Microscopic anatomy of a skeletal muscle fiber [13]. 6

2.3 Recording of an Action Potential (AP) in a muscle fiber [13]. 7

2.4 Individual muscle fiber potentials are recorded in sum as the Motor Unit Action Potential (MUAP) [19]. 8

2.5 Variation in Motor Unit (MU) size across the motor unit pool. [10]. 8

2.6 Motor Unit (MU) recruitment strategies and the Cinderella motor units [26]. 10

2.7 Muscles of the anterior fascial compartment of the forearm. (a) Superficial view. (b) Several muscles have been removed to reveal the flexor digitorum superficialis. (c) Deep muscles of the anterior compartment. [13]. 11

2.8 Muscles of the posterior fascial compartment of the right forearm. (a) Superficial muscles, posterior view. (b) Deep posterior muscles, superficial muscles removed. [13]. 13

2.9 Spatial potential distributions at the skin surface for motor units at different tissue depths. Deeper sources (yellow dot) produce broader, lower-amplitude distributions compared to superficial sources (white dot) [26]. 14

2.10 Schematic representation of the volume conductor model showing the detection volume for surface EMG. The detection volume is defined here as the region of space in the muscle in which a fiber generates a surface potential with energy higher or equal to 1/100 of that of the potential generated by the reference fiber. The model consists of anisotropic muscle tissue (a) and isotropic fat (b) and skin (c) layers [32]. 15

2.11 Schematic of Motor Unit Action Potential (MUAP) propagation detected by a linear electrode array. The left panel shows the muscle with colored lines marking the Innervation Zone (IZ). The right panel displays the resulting Single Differential (SD) signals, where three distinct Motor Unit Action Potentials (MUAPs) show bidirectional propagation originating from these zones. [34]. 16

2.12 Propagating vs. non-propagating components across 10 Single Differential (SD) channels. The red dashed line follows the propagating component (depolarization wave) traveling along the muscle fibers, while the blue dashed line marks the End Of Fiber (EOF) effect [34]. 17

2.13 Schematic drawing of the propagation of a Motor Unit Action Potential (MUAP) in time and space. The electrodes are arranged symmetrically with respect to the Innervation Zone (IZ), with P_1 and P_2 on the left, and P_3 and P_4 on the right. The graphs on the right show the detected potentials versus time. Due to the symmetrical positioning of P_2 and P_3 relative to the IZ, the differential potential detected between them would be near zero. [34] 18

2.14 Schematic representation of the fibers of two motor units (MU1 and MU2) that are innervated by two motoneurons (MN1 and MN2) in two innervation zones (IZ1 and IZ2): (a) the summation of the skin contributions of the propagating action potentials, as detected by a differential amplifier, are depicted as MUAP1 and MUAP2; (b) the interference signal [20]. 19

2.15 Surface electromyography recordings from the Tibialis Anterior, Peroneus Longus (PL), and Tibia bone during muscle activation. The presence of electrical activity on the Tibia (bone) demonstrates crosstalk i.e. the detection of signals originating from nearby active muscles rather than from the recording site itself [26]. 21

2.16 The curves illustrate the spectral and amplitude manifestations of neuromuscular fatigue, with characteristic decreases in Conduction Velocity (CV), Median Frequency (MNF), Mean Frequency (MDF) and increases in amplitude parameters (Average Rectified Value (ARV), Root Mean Square (RMS)) [26]. 23

3.1 Low-density EMG armband for gesture recognition. The Myo armband (Thalmic Labs) features a sparse circumferential array of 8 dry bipolar sEMG sensors. This configuration is representative of commercial low-density interfaces, providing discrete muscle activity signals for classifying hand gestures with low computational overhead and setup time [62]. 34

3.2 High-density EMG sleeve interface. The NeuroLife® sleeve features a dense grid of up to 150 embedded electrodes covering the entire forearm circumference. This configuration enables the high-resolution capture of spatiotemporal muscle activation patterns from both superficial and deep muscle groups, providing a rich dataset for decoding complex hand gestures and continuous joint kinematics compared to traditional sparse electrode setups. [65]. 35

3.3 Blind source separation (BSS) pipeline. (A)sEMG signals are acquired via a high-density electrode grid during isometric contraction (e.g., 35% MVC trapezoidal profile). The raw interference pattern is a summation of multiple MUAPs. (B) Blind Source Separation (BSS) algorithms separate the mixed signal into constituent Motor Unit (MU) spike trains (b and bi), identifying individual firing instances. (C) Spike-triggered averaging uses the identified firing times to extract the unique MUAP waveforms across the electrode array. (D) Raster plot showing the firing patterns of reliable motor units, color-coded by recruitment threshold [64]. 38

5.1 Experimental setup. (A) 1) Real Time target machine Speedgoat running the Simulink model. 2) Vicon motion capture system. 3) Cometa system bipolar EMG. 4) Sessantaquattro high-density EMG. 5) Sensorized glove from Bielefeld University. (B) Placement of the acquisition sensors: 23 reflective markers for the Vicon system, 10 bipolar EMG electrodes, 2 grids for HD-sEMG acquisition, and sensorized glove[7]. 53

6.1 Classifier performance as a function of feature dimensionality, illustrating the curse of dimensionality and the existence of an optimal number of features [87] 65

6.2 Two-point crossover: chromosome segments between two randomly selected cut points (dashed lines) are exchanged between parents to create offspring, resulting in new candidate feature subsets. 78

7.1 Pearson correlation matrix for the raw EMG feature set (`FS-Raw`, $N_f = 19$). The pronounced red blocks indicate high redundancy among magnitude-related descriptors. 82

7.2 Hierarchical clustering dendrogram for Raw Signal Features (`FS-Raw`) at distance threshold 0.05 (horizontal line). Branches and features are color-coded to indicate redundant feature groups. 83

7.3 Pearson correlation matrix for the deconvolved signal feature set (FS-Dec, $N = 26$ features). The heatmap highlights the low correlation (blue regions) of deconvolution features (lower right) compared to the redundant magnitude-related descriptors. 85
7.4 Hierarchical clustering dendrogram for FS-Dec at distance threshold 0.05 (horizontal line). Branches and features are color-coded to indicate redundant feature groups. 86
7.5 Feature redundancy heatmap for the combined feature set (FS-C, $N = 45$). 88
7.6 Hierarchical clustering dendrogram for FS-Combined at distance threshold 0.05 (horizontal line). Branches and features are color-coded to indicate redundant feature groups. 89
7.7 Selection frequency within the top 100 elite solutions for the raw feature set (FS-Raw, mean fitness: 0.0360). The optimization process achieved absolute convergence: a specific subset was selected in 100% of the best solutions. 92
7.8 Selection frequency within the top 100 solutions of the Deconvolved Signal Features (FS-Dec) population (mean fitness: 0.0221). While metrics like Maximum Fractal Length $(\mathrm{MFL})_d$, Slope Sign Change $(\mathrm{SSC})_d$, Sync, CoA, and GI exhibit near-saturation ($> 90\%$), features such as Waveform Length $(\mathrm{WL})_d$ and Moment of 3 order $(\mathrm{V3})_d$ show moderate variability across the solution pool. 93
7.9 Selection frequency within the top 10 solutions (mean fitness: 0.0213). As the algorithm converges, the subset crystallizes: WL_d and $\mathrm{V3}_d$ achieve 100% selection rates, while previously present features like Variance $(\mathrm{Var})_d$ and Willison Amplitude $(\mathrm{WAMP})_d$ are fully eliminated. 93
7.10 Feature selection frequency within the top 100 solutions of the combined feature set (Combined Features (FS-C), mean fitness: 0.0210). 94
7.11 Selection frequency refined to the top 10 elite solutions for the combined feature set (FS-C, mean fitness: 0.0192). 95
7.12 Selection frequency within the top 10 elite solutions for the FS-Raw feature set using the NLR classifier (mean fitness: 0.0242). 96
7.13 Selection frequency within the top 10 solutions of the FS-Dec population (mean fitness: 0.0117). Features like MFL_d, SSC_d, Sync, CoA, and GI exhibit near-saturation ($> 90\%$) 97
7.14 Selection frequency within the top 10 elite solutions for the FS-C feature set using the NLR classifier (mean fitness: 0.0138). In contrast to LDA, the algorithm fails to converge on secondary features. 98
7.15 Comparison of Macro F1-Score distributions for the FS-Dec, FS-Raw, and FS-C feature sets. Brackets indicate statistically significant differences between the methods ($p < 0.05$). 99

7.16 Confusion matrix for the FS-Raw feature set. The matrix is row-normalized, where diagonal cells represent the correct classification rate (Recall) for each gesture class and off-diagonal cells indicate misclassification percentages. 102
7.17 Confusion matrix for the FS-Dec feature set. The matrix is row-normalized, where diagonal cells represent the correct classification rate (Recall) for each gesture class and off-diagonal cells indicate misclassification percentages. 103
7.18 Confusion matrix for the FS-C feature set. The matrix is row-normalized, where diagonal cells represent the correct classification rate (Recall) for each gesture class and off-diagonal cells indicate misclassification percentages. 103
7.19 Safety profile per gesture class using LDA. For each gesture, the grouped bars represent the performance of the FS-Raw (left), FS-Dec (middle), and FS-C (right) feature sets. The stacked segments visualize the distribution of classification outcomes. 104
7.20 Accuracy vs. Abstention Rate trade-off curves for the FS-Raw, FS-Dec, and FS-C feature sets. The solid lines represent the mean accuracy calculated on accepted samples, while the shaded regions indicate the standard deviation. The graph illustrates the progression of classification reliability as the system is permitted to abstain from uncertain predictions. 105
7.21 Comparison of Macro F1-Score distributions for the FS-Dec, FS-Raw, and FS-C feature sets. Brackets indicate statistically significant differences between the methods ($p < 0.05$). 106
7.22 Confusion matrix for the FS-Raw feature set. The matrix is row-normalized, where diagonal cells represent the correct classification rate (Recall) for each gesture class and off-diagonal cells indicate misclassification percentages. 108
7.23 Confusion matrix for the FS-Dec feature set. The matrix is row-normalized, where diagonal cells represent the correct classification rate (Recall) for each gesture class and off-diagonal cells indicate misclassification percentages. 109
7.24 Confusion matrix for the FS-C feature set. The matrix is row-normalized, where diagonal cells represent the correct classification rate (Recall) for each gesture class and off-diagonal cells indicate misclassification percentages. 109
7.25 Safety profile per gesture class using the NLR classifier. For each gesture, the grouped bars represent the performance of the FS-Raw (left), FS-Dec (middle), and FS-C (right) feature sets. The stacked segments visualize the distribution of classification outcomes. 110

7.26 Accuracy vs. Abstention Rate trade-off curves for the FS-Raw, FS-Dec, and FS-C feature sets. The solid lines represent the mean accuracy calculated on accepted samples, while the shaded regions indicate the standard deviation. 111
7.27 Confusion matrix showing the online performance of the chosen classifier and feature set. The matrix is row-normalized: diagonal cells represent the correct classification rate (Recall) for each gesture class, while off-diagonal cells indicate misclassification percentages. . 115
7.28 Safety profile per gesture class obtained from online testing. For each gesture, the grouped bars represent the performance of the FS-Raw (left), FS-Dec (middle), and FS-C (right) feature sets. The stacked segments visualize the distribution of outcomes (correct, abstention, and misclassification). 116

Acronyms

`FS-C`	Combined Features
`FS-Dec`	Deconvolved Signal Features
`FS-Raw`	Raw Signal Features
ANN	Artificial Neural Network
ANOVA	Analysis of Variance
AP	Action Potential
AR	Abstention Rate
ARV	Average Rectified Value
Brst	Burstiness
BSS	Blind Source Separation
CKC	Convolution Kernel Compensation
CNN	Convolutional Neural Network
CoA	Co-Activation
CoV	Coefficient of Variation
CV	Conduction Velocity
CWF	Cumulative Weighted Firing
EMG	Electromyography
EOF	End Of Fiber
FRI	Fourth Root Integral
GA	Genetic Algorithm
GI	Gini Index
GUI	Graphic User Interface
HC	Hand Closing

HD-EMG High-Density Electromyography
HO Hand Opening

IASD Integrated Absolute of Second Derivative
IED Inter-Electrode Distance
IRLS Iterative Reweighted Least Squares
ISI Inter-Spike Interval
ISIE Interspike Interval Entropy
IZ Innervation Zone

Krt Kurtosis

Lat Latency
LD Logarithm Detector
LD-EMG Low-Density Electromyography
LDA Linear Discriminant Analysis
LSTM Long Short-Term Memory

M2 second-order moment
M4 fourth-order moment
MAVS Mean Absolute Value Slope
MDF Mean Frequency
MFL Maximum Fractal Length
MLE Maximum Likelihood Estimation
MLP Multi-Layer Perceptron
MNF Median Frequency
MU Motor Unit
MUAP Motor Unit Action Potential
MVC Maximum Voluntary Contraction

NLR Non-Linear Logistic Regression

P Pinch
PFP Progressive FastICA Peel-off
PSD Power Spectral Density

R Rest
RD Radial Deviation
RI Root Integral
RMS Root Mean Square

SAI	Simple Absolute Integral
SD	Single Differential
sEMG	Surface Electromyography
Skw	Skewness
SR	Spike Rate
SRI	Square Root Integral
SSC	Slope Sign Change
SSI	Simple Square Integral
SVM	Support Vector Machine
Sync	Synchronization
UD	Ulnar Deviation
V3	Moment of 3 order
Var	Variance
WAMP	Willison Amplitude
WE	Wrist Extension
WF	Wrist Flexion
WL	Waveform Length
WP	Wrist Pronation
WS	Wrist Supination
ZC	Zero Crossing

Chapter 1

Introduction

Hand function relies on the coordinated activation of multiple muscles and joints to produce dexterous movements ranging from simple grasps to fine wrist rotations. In assistive and rehabilitation technologies, decoding these movements is a central requirement for intuitive control of prosthetic hands, exoskeletons, and human–computer interfaces [1]. Among the available biosignals, Surface Electromyography (sEMG) provides a direct, non-invasive measurement of muscle activation and remains one of the most practical sensing modalities for wearable myoelectric interfaces [2].

Despite decades of progress, the performance gap between laboratory demonstrations and daily-life usability remains substantial. In particular, upper-limb prostheses still exhibit high abandonment rates, indicating that reliability, comfort, responsiveness, and ease of use are at least as important as raw classification accuracy [3]. From a signal processing perspective, sEMG-based gesture recognition is challenged by inter- and intra-subject variability, electrode placement sensitivity, crosstalk, and fatigue-related signal changes, all of which may alter feature distributions and reduce the robustness of pattern-recognition systems in practical use [4].

1.1 Context and motivation

Conventional sEMG-based gesture recognition systems operate on short signal windows and extract features directly from the interference Electromyography (EMG). Although this approach has been widely adopted, the resulting descriptors are derived from a signal that reflects neural drive only indirectly and is affected by amplitude cancellation, electrode placement, crosstalk, and other sources of variability [5, 4]. These factors can reduce robustness in practical myoelectric control, especially when low latency and low channel count are required.

A possible way to enrich the representation is to move toward information more closely related to Motor Unit (MU) discharge behavior. High-Density Electromyography (HD-EMG) and multi-channel decomposition methods can provide this type of information, but they generally require many electrodes and computationally demanding processing pipelines that are less compatible with wearable real-time systems [6]. This motivates the approach investigated in this thesis: using single-channel sEMG deconvolution as a lightweight preprocessing step and evaluating whether descriptors extracted from the resulting Cumulative Weighted Firing (CWF) can improve gesture recognition while preserving an online-compatible pipeline. In line with this objective, although the original Reach&Grasp dataset provides 64 monopolar electrodes, the present study restricts the analysis to 8 monopolar signals forming 4 Single Differential (SD) channels [7].

1.2 Goal

The goal of this thesis is to assess whether single-channel sEMG deconvolution can improve feature-based hand gesture recognition under low-complexity, online-compatible conditions. The study focuses on a subset of wrist and hand movements from the public Reach&Grasp dataset [7] and evaluates the effect of deconvolution-derived representations on both offline classification performance and online operation under latency constraints.

More specifically, the objectives are:

- implement a single-channel deconvolution pipeline (kernel estimation and Iterative Reweighted Least Squares (IRLS)-based inverse solution) to estimate a cumulative firing activity signal from each SD channel [8];
- define and extract compact feature representations from (i) raw sEMG, (ii) deconvolved signals, and (iii) their combination, to isolate the contribution of deconvolution-based descriptors;
- apply a sequential filter–wrapper feature selection strategy to reduce redundancy and identify small, high-performing feature subsets suitable for real-time inference;
- quantify the reliability–responsiveness trade-off introduced by confidence-based abstention, and validate the proposed pipeline both offline (held-out testing) and online (streaming inference).

1.3 Thesis structure

The remainder of this thesis is organized as follows:

- **Chapter 2** introduces the physiological and anatomical background required to interpret sEMG, including MU recruitment principles and the main mechanisms underlying signal generation.
- **Chapter 3** reviews the state of the art in sEMG-based hand gesture recognition, with emphasis on practical constraints for prosthetic control, classifier choices, abstention mechanisms, and electrode configurations.
- **Chapter 4** presents the single-channel sEMG deconvolution method adopted in this work.
- **Chapter 5** describes the dataset and acquisition setup, and details the preprocessing and feature extraction procedures for raw and deconvolved signals.
- **Chapter 6** introduces the feature selection methodology, combining correlation based filtering and wrapper optimization to obtain compact feature subsets.
- **Chapter 7** reports the experimental results, comparing feature sets and classifiers in offline evaluation and presenting the online performance of the real-time pipeline.
- **Chapter 8** summarizes the main findings of this thesis and discusses possible directions for future work.

Chapter 2

Physiological and Anatomical Background

Hand gestures involve the coordinated activation of intrinsic muscles within the hand and extrinsic muscles in the forearm, together with their associated tendons and joints, which produce finely controlled multi-joint movements [9]. Understanding these anatomical structures and physiological processes underlying both hand movement and Surface Electromyography (sEMG) generation is essential for extracting information about how the nervous system controls movement from Electromyography (EMG) signals [10]. From the perspective of surface electromyography, these structures define which muscles contribute to a gesture, how their fibers are oriented, and how far they lie from the recording electrodes, all of which strongly influence the recorded signals [11].

2.1 Skeletal Muscle and Motor Units

2.1.1 Skeletal Muscle Structure

Skeletal muscles are organs that have the ability to contract and generate movement. They consist of skeletal muscle fibers, blood vessels, nerve fibers and connective tissue. The role of the connective tissue is to provide structural support to the muscle, separate the different structures within each muscle and to define a pathway for the blood vessels and axons.

Muscle fibers are grouped into bundles called fascicles, and multiple fascicles together form the whole muscle. [12] These structures are surrounded by three connective tissue layers (mysia) as seen in Fig. 2.1:

- **Epimysium**: dense connective tissue that surrounds the entire muscle and contributes to force transmission to the tendon.

- **Perimysium**: connective tissue that groups muscle fibers into fascicles, providing pathways for nerves and blood vessels.
- **Endomysium**: thin layer of connective tissue that surrounds each individual muscle fiber and helps maintain the extracellular environment.

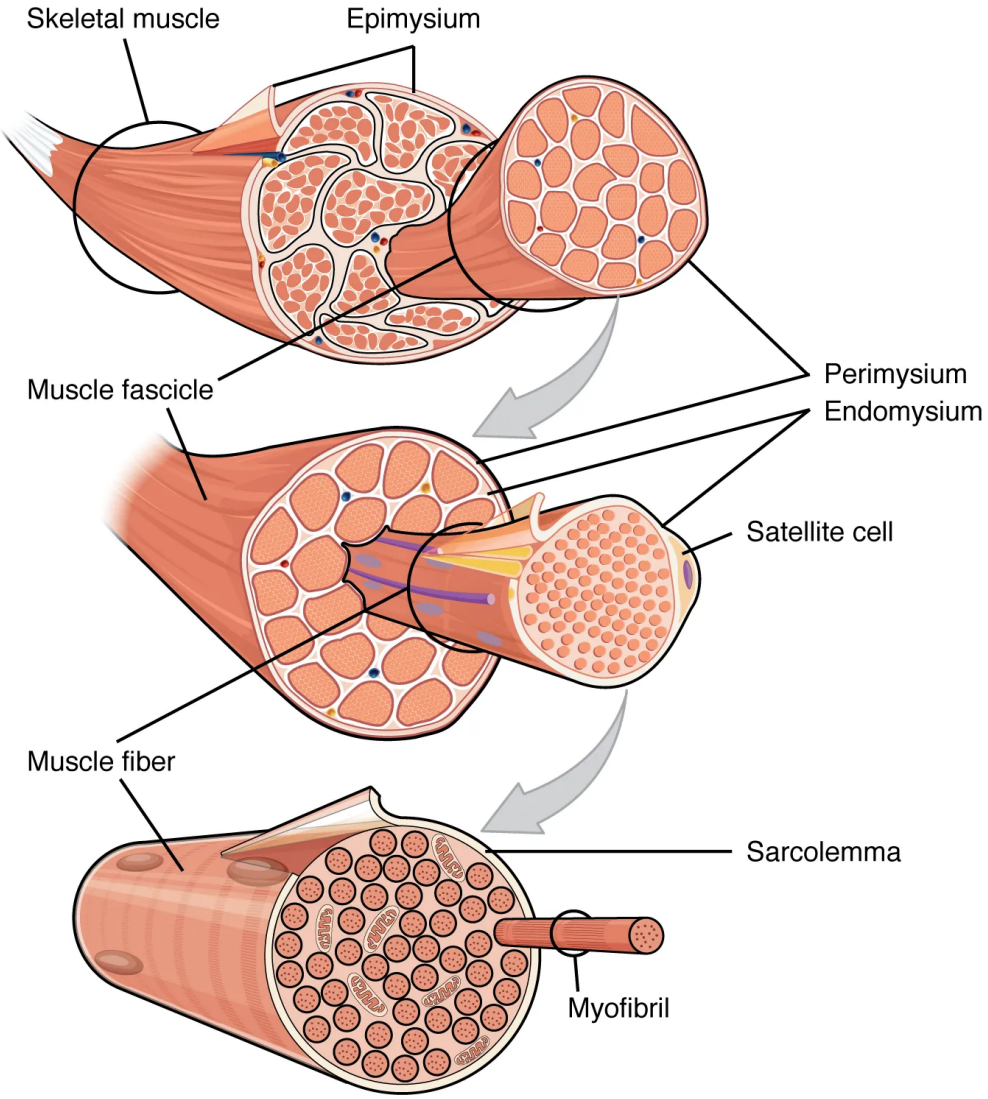

Figure 2.1: Three Connective Tissue Layers [12].

A single fascicle contains many muscle fibers, typically ranging from hundreds to several thousands depending on the muscle and its functional role. [12]

It is important to note that this anatomical organization into fascicles does not define distinct activation regions of the muscle, i.e. it is not a functional

organization; functional activation is determined by motor units, which will be discussed in Section 2.1.3.

At a lower level of organization, each muscle fiber is surrounded by its plasma membrane, the sarcolemma, with diameters typically between 10 μm and 100 μm and lengths that can reach several tens of centimeters in large muscles [13]. The sarcolemma hosts the neuromuscular junction, where the terminal axons of a motor neuron synapses onto the muscle fiber and transmits the excitation signals that trigger contraction [12].

Muscle fibers contain hundreds to thousands of myofibrils that occupy most of the cell volume. [12] These myofibrils are composed of repeating sarcomeres, the fundamental contractile units of skeletal muscle, in which actin and myosin filaments slide past each other during contraction; sarcomere shortening leads to fiber shortening and force transmission through the mysia to the tendon [13].

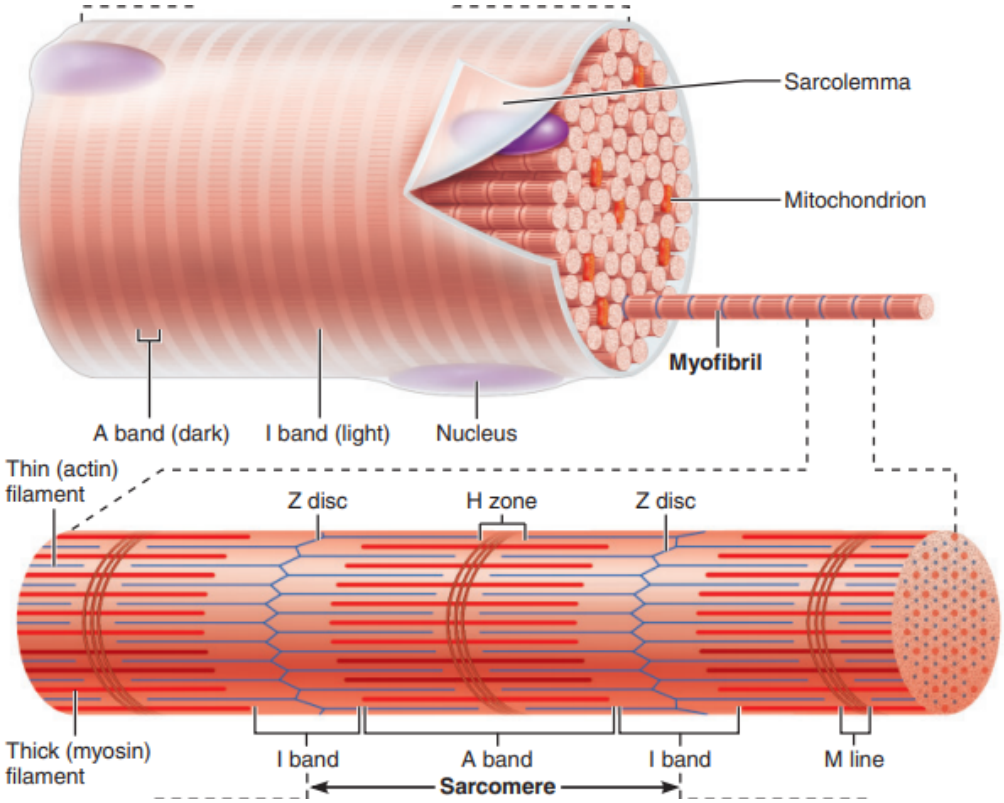

Figure 2.2: Microscopic anatomy of a skeletal muscle fiber [13].

This hierarchical organization, together with fiber orientation and depth, will influence how Motor Unit (MU) activity is represented in surface EMG recordings, as discussed in the following subsections. [14]

2.1.2 Excitability and Action Potential Propagation

Like all living cells, muscle fibers maintain a transmembrane electrical gradient, with the intracellular potential typically between -60 and -90 mV relative to the extracellular space caused by the asymmetrical distributions of Na^+, K^+ and Cl^+ [15]. These fibers are electrically excitable cells; by regulating ion fluxes through voltage and ligand-gated membrane channels, the sarcolemma can rapidly transition from its resting potential to a propagating Action Potential (AP) along the fiber surface. [16] The upstroke of this action potential is produced by a brief, large increase in sodium permeability via voltage-gated Na^+ channels, whereas subsequent opening of voltage-gated K^+ channels drives repolarization of the membrane [17].

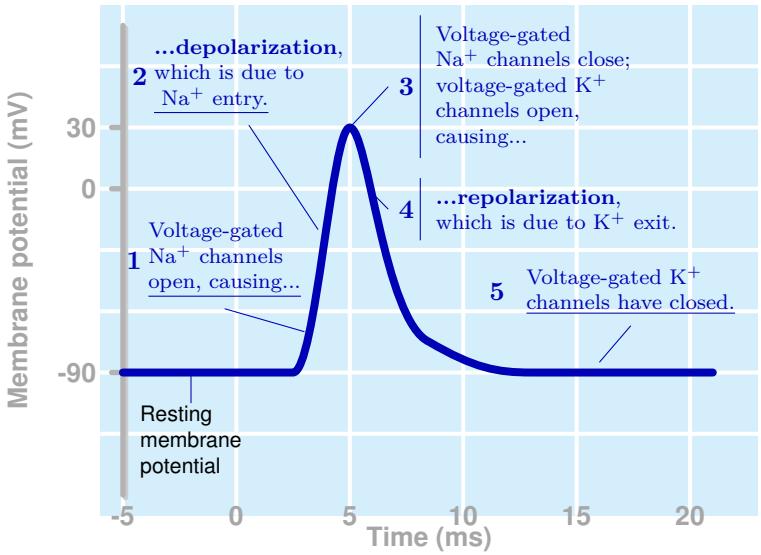

Figure 2.3: Recording of an Action Potential (AP) in a muscle fiber [13].

The AP generates the contraction of the muscle by travelling along the cell membrane and carrying the signal into the interior of the muscle fiber where conformational changes in intracellular receptors cause the influx of Ca^{2+} [16]. The increase in Ca^{2+} concentration in the sarcoplasm initiates the contraction of the muscle fiber by its contractile units, or sarcomeres [12].

2.1.3 Motor Units and Recruitment

The MU is the basic functional unit of the neuromuscular system and consists of the entire motoneuron, including its dendrites and axon, and the muscle fibers innervated by the axon [18]. Each motoneuron can innervate anywhere between 4

muscle fibers to several hundreds, the average size of motor units within a muscle depends on whether the muscle's purpose is to exert fine control or if it is a weight bearing muscle [12].

Functionally, the activation of a motoneuron triggers action potentials in all its innervated muscle fibers. The spatial and temporal summation of these single fiber action potentials constitutes the Motor Unit Action Potential (MUAP) [19]. These potentials propagate bidirectionally from the Innervation Zone (IZ) toward the tendons and last between 10 and 15 ms [20, 21].

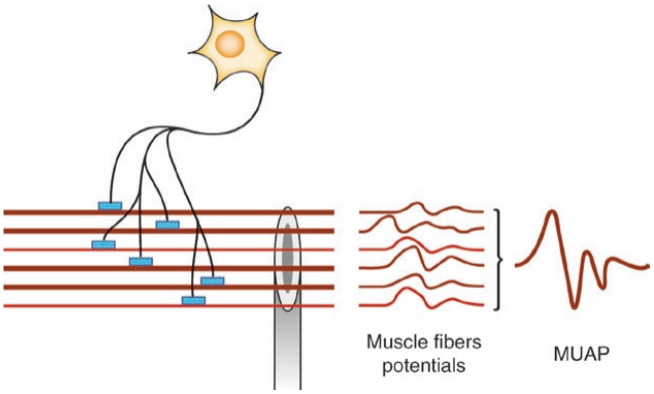

Figure 2.4: Individual muscle fiber potentials are recorded in sum as the Motor Unit Action Potential (MUAP) [19].

Within a muscle, MU sizes are not uniformly distributed: most MU innervate relatively few fibers, whereas a smaller number innervate many fibers, a pattern often captured by exponential-like distributions in modeling studies [10].

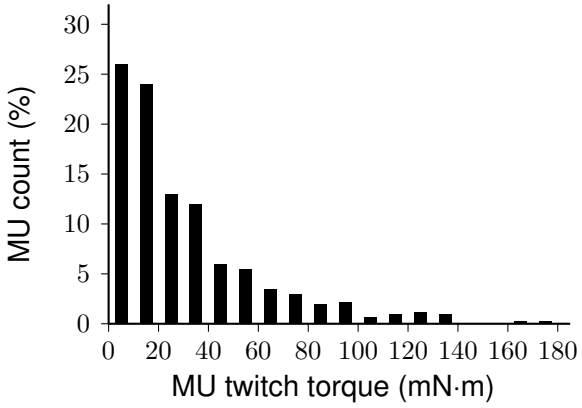

Figure 2.5: Variation in Motor Unit (MU) size across the motor unit pool. [10].

Muscle fibers corresponding to the same motor unit can occupy from 10% to

70% of the cross-sectional area of a muscle, meaning that the density of muscle fibers is around 10 times larger than the density of motor units [10]. MU size tends to covary with twitch force, EMG twitch amplitude, and recruitment threshold, so that small MUs are weak, produce small EMG potentials, and are recruited at low force levels, whereas large MUs are strong, have larger EMG potentials, and are only recruited as force demands increase (Figure 2.6) [22, 23].

This organization is a manifestation of Henneman's size principle: as synaptic drive to the motoneuron pool rises, motoneurons are recruited in an orderly fashion from the smallest, lowest-threshold cells to progressively larger, higher-threshold cells [22]. A key mechanism is that smaller MUs have higher input resistance, so a given synaptic current generates a larger change in membrane potential and brings them to firing threshold earlier than larger MUs [24, 22].

In addition to recruitment, the nervous system modulates muscle force by changing the mean firing rate of active MUs (rate coding). At low contraction levels, force increments are achieved mainly by recruiting additional low-threshold units, whereas at higher levels, further increases in force rely more on raising the discharge rate of already active, larger MUs [24]. From the EMG perspective, this means that both the number of active MUs and their firing rates shape the interference pattern, with higher forces reflecting not only the recruitment of larger units but also their increased discharge frequency (Figure 2.6) [25].

These recruitment and rate-coding strategies determine the temporal structure of motor unit firing patterns. Extracting these patterns from surface EMG could provide features that are more directly related to motor control strategies than traditional amplitude or spectral features, which are confounded by volume conductor effects and amplitude cancellation (see Section 2.2.3).

2.1.4 Muscles Involved in the Studied Gestures

In this work, surface EMG is acquired exclusively from extrinsic hand muscles located in the forearm, which generate most of the force for hand opening/closing and wrist movements, and therefore dominate the myoelectric patterns associated with the studied gestures [12, 27].

The main anterior (flexor/pronator) compartment muscles considered are [28]:

- **Flexor carpi radialis (FCR)**: flexes and abducts the wrist, contributing to wrist flexion and radial deviation during closing or combined flexion–grip tasks.
- **Palmaris longus (PL)**: assists in wrist flexion and tenses the palmar aponeurosis, adding to EMG activity during general wrist flexion and gripping.
- **Flexor carpi ulnaris (FCU)**: flexes and adducts the wrist, active in wrist

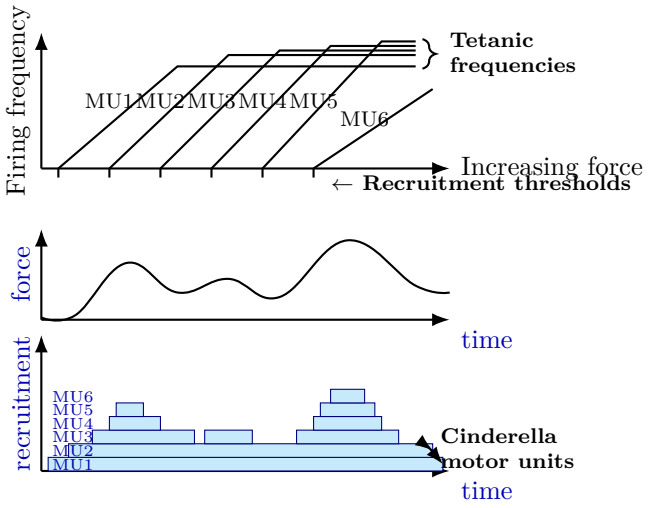

Figure 2.6: Motor Unit (MU) recruitment strategies and the Cinderella motor units [26].

flexion with ulnar deviation and in stabilizing the wrist during strong hand closing.

- **Flexor digitorum superficialis (FDS)**: flexes the proximal interphalangeal joints of fingers 2–5 and assists in wrist flexion, providing a major contribution to hand closing and fist formation.
- **Flexor digitorum profundus (FDP)**: flexes the distal interphalangeal joints of fingers 2–5 and contributes to overall finger flexion during strong grips.
- **Pronator teres (PT)**: pronates the forearm and assists with elbow flexion, contributing to wrist/forearm pronation gestures.
- **Pronator quadratus (PQ)**: deep distal pronator that produces forearm pronation across the radioulnar joints, particularly in sustained pronation tasks.

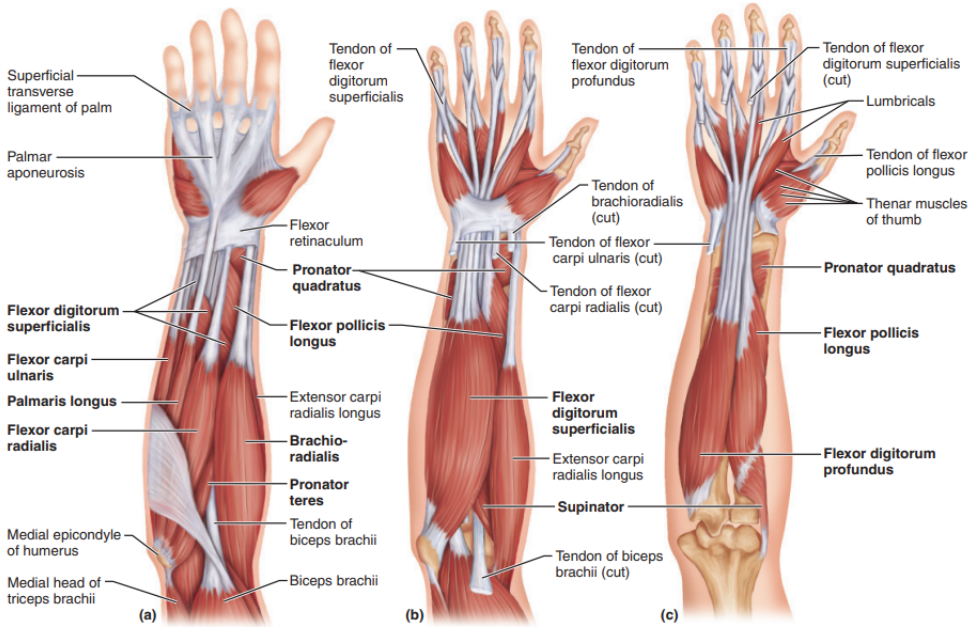

Figure 2.7: Muscles of the anterior fascial compartment of the forearm. (a) Superficial view. (b) Several muscles have been removed to reveal the flexor digitorum superficialis. (c) Deep muscles of the anterior compartment. [13].

The main posterior (extensor/supinator) compartment muscles considered are [28]:

- **Brachioradialis (BR)**: superficial lateral forearm muscle that flexes the elbow and helps bring the forearm toward a neutral position, assisting both pronation and supination transitions.
- **Extensor carpi radialis longus (ECRL)**: extends and abducts the wrist, contributing to wrist extension and radial deviation.
- **Extensor carpi radialis brevis (ECRB)**: assists ECRL in extending and abducting the wrist and stabilizes the wrist during finger flexion.
- **Extensor carpi ulnaris (ECU)**: extends and adducts the wrist, active in wrist extension with ulnar deviation.
- **Extensor digitorum (ED)**: extends fingers 2–5 and assists in wrist extension, providing the main extrinsic contribution to hand opening.
- **Extensor indicis (EI)**: provides independent extension of the index finger, supporting selective index opening gestures.
- **Extensor digiti minimi (EDM)**: extends the little finger, contributing to selective movements of digit 5 and lateral hand opening.
- **Supinator (SUP)**: wraps around the proximal radius and supinates the forearm, particularly during low-load supination of the wrist/hand.

Collectively, these muscles generate the EMG patterns associated with hand opening and closing, wrist flexion and extension, and wrist pronation and supination investigated in this study. [13]. Because these extrinsic muscles are relatively large and superficial, their activation can be detected reliably with multi-channel surface EMG, and distinct spatial patterns over the forearm have been shown to correspond to different wrist and finger motions [29].

2.2 EMG Signal Generation

2.2.1 Volume Conduction and Spatial Filtering

The surface EMG signal reflects the summed electrical activity of muscle fibers that are active during a contraction, and originates from the depolarization and repolarization zones that travel along their membranes. During surface recordings, these sources are separated from the electrodes by several biological tissues which together form a volume conductor that acts as a spatial and temporal low-pass filter

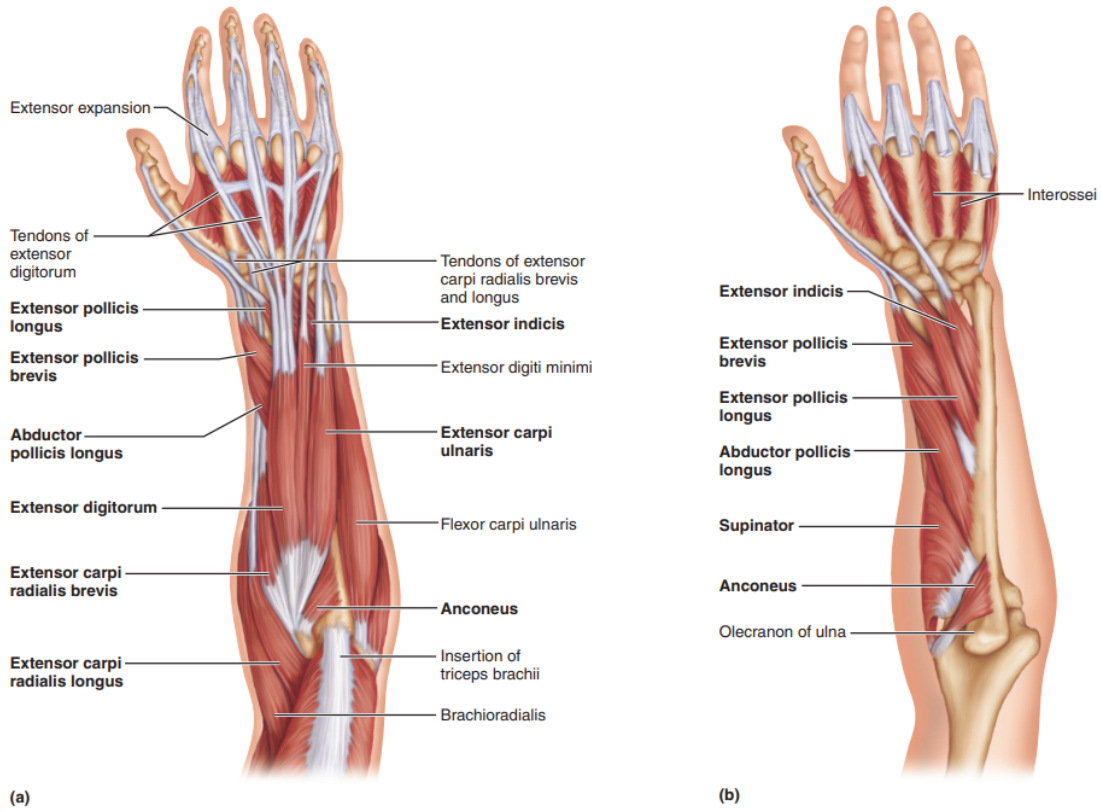

Figure 2.8: Muscles of the posterior fascial compartment of the right forearm. (a) Superficial muscles, posterior view. (b) Deep posterior muscles, superficial muscles removed. [13].

on the potential distribution at the skin surface [30, 20]. For intramuscular EMG, the deforming effect of the intervening tissues is relatively limited because the electrodes lie very close to the active fibers, whereas for sEMG the volume conductor substantially alters the detected signal [10]. As action potentials propagate through successive tissue layers (muscle, fat, skin), their high-frequency components are progressively attenuated, resulting in a smoothed and lower-amplitude representation of the motor unit action potential on the skin surface (Figure 2.9). The filtering intensity is directly proportional to the thickness of the volume conductor, with subcutaneous fat thickness being the most variable factor across individuals; variations from 3 mm to 7 mm can reduce MUAP amplitude by up to 50–70% [23].

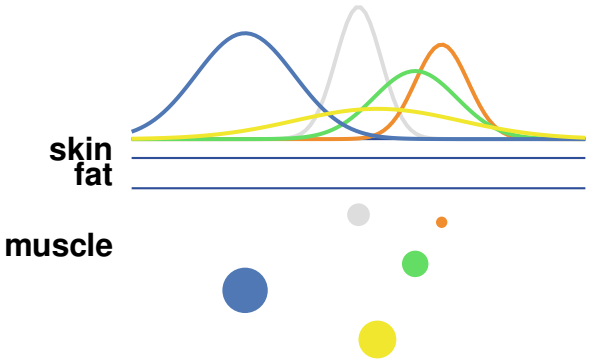

Figure 2.9: Spatial potential distributions at the skin surface for motor units at different tissue depths. Deeper sources (yellow dot) produce broader, lower-amplitude distributions compared to superficial sources (white dot) [26].

The detection volume defines the spatial region from which surface electrodes can reliably detect motor unit activity, and its extent depends on the electrode configuration, Inter-Electrode Distance (IED), and tissue properties [31]. For a Single Differential (SD) configuration with 10 mm IED, commonly used in forearm EMG recordings, the detection volume typically extends 20–40 mm in depth and laterally from the electrode pair [23, 32]. The detection field is further shaped by the anisotropy of muscle tissue, which conducts current approximately 5–10 times better along the fibers than transversally [33]. Consequently, MUs located deeper or farther laterally contribute progressively less to the recorded signal due to spatial filtering effects. This selectivity implies that surface EMG predominantly reflects the activity of superficial motor units within a limited portion of the muscle cross-section, which may not be fully representative of the entire motor unit pool [31, 25].

These spatial filtering properties have several important implications for interpreting sEMG signals in pattern recognition applications. First, the recorded signal

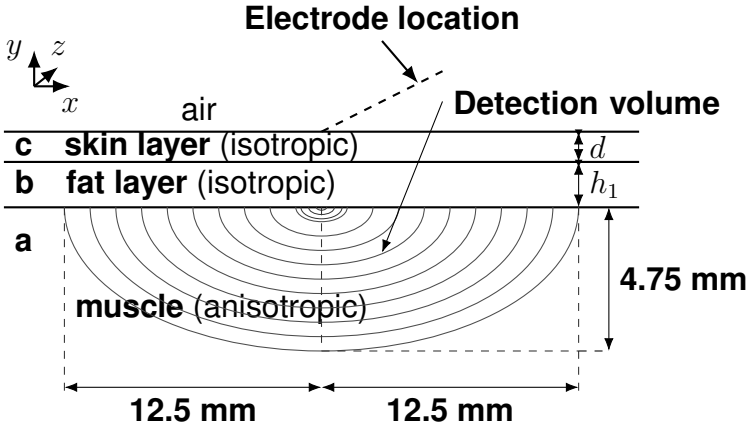

Figure 2.10: Schematic representation of the volume conductor model showing the detection volume for surface EMG. The detection volume is defined here as the region of space in the muscle in which a fiber generates a surface potential with energy higher or equal to 1/100 of that of the potential generated by the reference fiber. The model consists of anisotropic muscle tissue (a) and isotropic fat (b) and skin (c) layers [32].

amplitude is weighted toward larger, more superficial MUs, creating a detection bias that must be considered when estimating muscle activation levels [25]. Second, the tissues attenuate the high-frequency components of the source potentials, effectively smoothing the signal; this means the bandwidth of sEMG is significantly narrower than that of the source currents or intramuscular recordings [10]. Third, electrode placement relative to the IZ and tendon regions critically affects the balance between propagating and non-propagating signal components, as will be discussed in Section 2.2.2 [31].

2.2.2 Motor Unit Action Potential Propagation

When a MU fires, the AP is initiated at the neuromuscular junction and propagates bidirectionally from the innervation zone towards both muscle-tendon junctions [20, 23]. All fibers within a MU share similar diameters and histological properties, causing their AP to travel at approximately the same Conduction Velocity (CV), typically ranging from 3 to 6 m/s [23]. This synchronous propagation causes the individual fiber contributions to sum coherently at the skin surface, creating a distinctive MUAP waveform that serves as a temporal and spatial "fingerprint" for each MU [20].

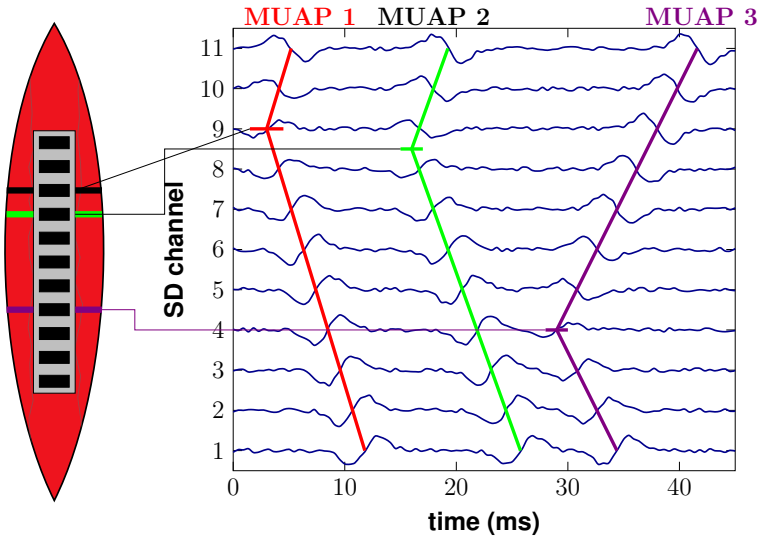

Figure 2.11: Schematic of Motor Unit Action Potential (MUAP) propagation detected by a linear electrode array. The left panel shows the muscle with colored lines marking the Innervation Zone (IZ). The right panel displays the resulting Single Differential (SD) signals, where three distinct MUAPs show bidirectional propagation originating from these zones. [34].

The surface detected MUAP is not a simple traveling wave but comprises two distinct signal components with different origins and propagation characteristics [23]. The propagating component represents the traveling depolarization wave moving along the fibers toward the tendons, whereas the extinction component occurs at the two muscle-tendon junctions, where APs terminate as they reach the high-impedance boundary of the tendon [20, 23]. These stationary components can significantly affect the detected signal morphology and contribute to crosstalk between muscles.

The extinction phenomenon, commonly referred to as the End Of Fiber (EOF) effect, arises from the asymmetric termination of the action potential at the tendon. The propagating action potential can be modeled as a tripole consisting of a central depolarization zone (current source) flanked by two repolarization zones (current sinks) [20]. As this tripole approaches the tendon, the leading dipole progressively shrinks and vanishes, leaving the trailing dipole momentarily unbalanced. This creates a sharp, non-propagating transient in the monopolar signal that can be observed as a stationary "tail" following the propagating wave (Figure 2.12) [20]. Although SD detection partially attenuates these non-propagating components through spatial filtering, they remain significant contributors to the overall signal, particularly for deeper motor units where the EOF amplitude can exceed that of the propagating component [20]. These stationary signals also constitute the primary source of crosstalk between adjacent muscles, as they spread spatially more than propagating components and are less effectively canceled by differential recordings [31].

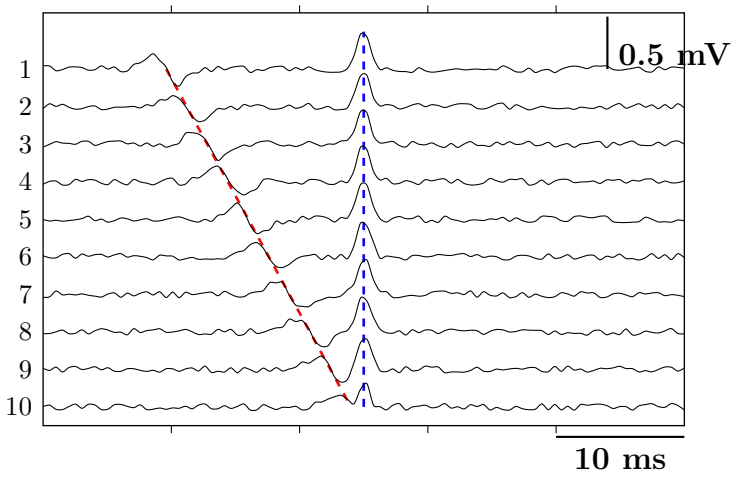

Figure 2.12: Propagating vs. non-propagating components across 10 Single Differential (SD) channels. The red dashed line follows the propagating component (depolarization wave) traveling along the muscle fibers, while the blue dashed line marks the End Of Fiber (EOF) effect [34].

The morphology of the detected MUAP depends critically on electrode location relative to the innervation zone and tendon regions. When detection electrodes are placed correctly between the innervation zone and a tendon, the MUAP appears as a clear biphasic waveform, reflecting the passage of the depolarization zone under the electrode pair in one direction only [31, 20]. A key feature of MUAP propagation is the polarity reversal across the innervation zone: electrodes on the distal side of the IZ detect a signal with inverted polarity relative to those on the proximal side, a property frequently exploited to locate the innervation zone experimentally using linear electrode arrays(Figure 2.13) [20]. For reliable EMG interpretation and minimal contamination from non-propagating components, electrodes should be placed in the region between the innervation zone and the nearest tendon, avoiding the IZ itself where signal amplitude approaches zero due to phase cancellation(Figure 2.13) [31, 20].

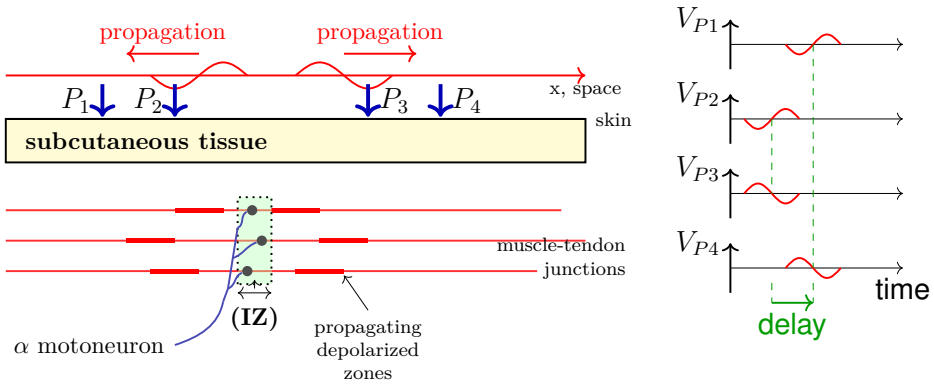

Figure 2.13: Schematic drawing of the propagation of a Motor Unit Action Potential (MUAP) in time and space. The electrodes are arranged symmetrically with respect to the Innervation Zone (IZ), with P_1 and P_2 on the left, and P_3 and P_4 on the right. The graphs on the right show the detected potentials versus time. Due to the symmetrical positioning of P_2 and P_3 relative to the IZ, the differential potential detected between them would be near zero. [34]

2.2.3 Interference Pattern Formation

Superposition and Amplitude Cancellation

The sEMG represents a complex interference pattern arising from the asynchronous summation of MUAPs from multiple active MUs [25]. The mathematical representation of this interference signal can be expressed as:

$$s(t) = \sum_{i=1}^{M} \sum_{j=-\infty}^{\infty} \varphi_i(t)\delta(t - \tau_{ij}) + n(t) \tag{2.1}$$

where M is the number of active MUs, $s(t)$ is the recorded EMG signal, $\varphi_i(t)$is the shape of the MUAP for the i-th MU, τ_{ij} is the j-th time at which the i-th MU discharges APs and $\delta(t - \tau_{ij})$ is the delta function representing a MU discharge at time τ_{ij} [25]; $n(t)$ represents additive noise. This model indicates that the EMG signal contains information about both the neural drive (discharge times τ_{ij}) and the muscle fiber electrical properties (AP shapes $\varphi_i(t)$) [25].

Equivalently, this can be expressed in convolution operator as:

$$s(t) = \sum_{m=1}^{N} M_m(t) * F_m(t) + n(t) \tag{2.2}$$

where $M_m(t)$ is the m-th MUAP, $*$ denotes convolution, $F_m(t)$ is the firing pattern of the m-th MU, and N is the number of active MUs.

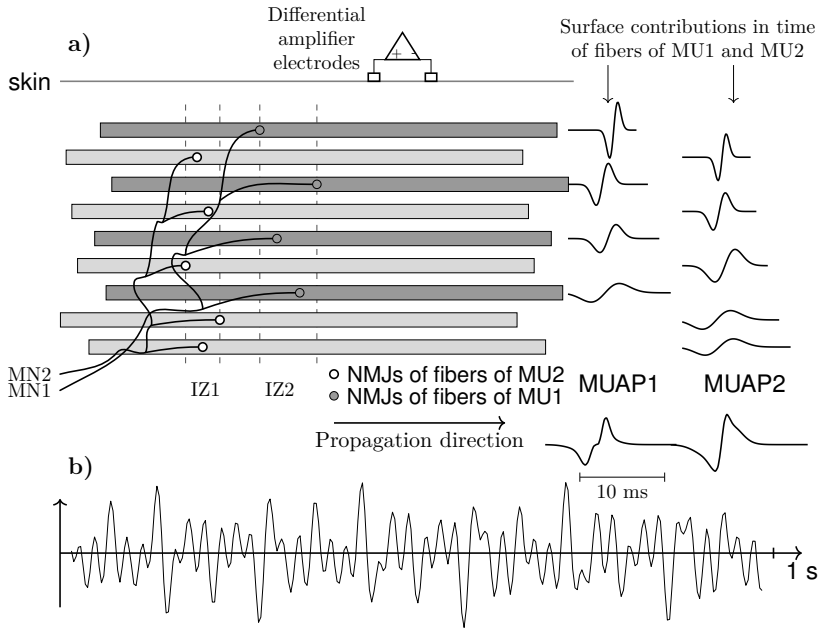

Figure 2.14: Schematic representation of the fibers of two motor units (MU1 and MU2) that are innervated by two motoneurons (MN1 and MN2) in two innervation zones (IZ1 and IZ2): (a) the summation of the skin contributions of the propagating action potentials, as detected by a differential amplifier, are depicted as MUAP1 and MUAP2; (b) the interference signal [20].

A fundamental phenomenon affecting interference pattern analysis is amplitude

cancellation, which occurs due to the algebraic summation of positive and negative phases of overlapping MUAPs [25, 35].

Amplitude cancellation is formally expressed by the inequality:

$$\sigma_{EMG} = \sum_{i=1}^{M} \sqrt{P_i} < \sum_{i=1}^{M} \sigma_i \tag{2.3}$$

where σ_i and P_i are the amplitude and power for the train of action potentials discharged by the i-th motor unit [25]. This phenomenon indicates that the estimation of EMG amplitude from an interference signal is less than that obtained by summing the amplitudes of the individual motor unit action potentials [25]. The amplitude cancellation term corresponds to the difference between the sum of the amplitudes of the action potential trains and the amplitude of the EMG. Amplitude cancellation increases monotonically with the number of activated motor units, and is greatest when the action potential shapes are the same for all motor units [25]. Critically, rectification of an EMG signal does not eliminate the influence of amplitude cancellation i.e. a rectified EMG signal has the same amount of amplitude cancellation as an interference EMG signal [25, 10].

Power Spectral Characteristics

In contrast to its effect on EMG amplitude estimates, amplitude cancellation does not influence the power spectrum of an interference EMG signal [25]. The distribution of signal power in an interference EMG is the sum of the powers of the trains of MUAPs, provided the trains are uncorrelated [25]. Each train's power spectral density comprises two components: the spectrum of the surface-detected AP and their repetition rate [35]. The latter produces peaks at frequencies that are multiples of the average discharge rate. Assuming independence among MUAP trains, their individual power spectra add linearly to form the overall spectrum of the interference EMG signal [35].

The useful energy of the sEMG signal is primarily contained within the frequency range of approximately 20-500 Hz, with the dominant spectral power concentrated in the 50-150 Hz range [14, 31]. This bandwidth is determined by the physiological properties of muscle fiber action potentials and the spatial filtering effects of subcutaneous tissues [14].

Crosstalk

Crosstalk refers to the detection of EMG signals originating from muscles adjacent to the target muscle, and represents a significant source of signal contamination in sEMG recordings [35, 31]. This phenomenon is particularly relevant in forearm

recordings where multiple muscles are anatomically close and share similar activation patterns during hand gestures.

The mechanisms underlying crosstalk are primarily related to the volume conductor properties and the non-propagating components of motor unit action potentials [31]. Contrary to intuitive expectations, crosstalk signals do not necessarily exhibit lower frequency content despite originating from greater distances [35]. This counter-intuitive behavior occurs because signals recorded at distant locations are dominated by extinction components (end of fiber effects) rather than propagating components [35]. The extinction components at muscle-tendon junctions contain higher frequency content than propagating signals, and because they spread spatially more than propagating waves, they dominate the crosstalk contribution [31].

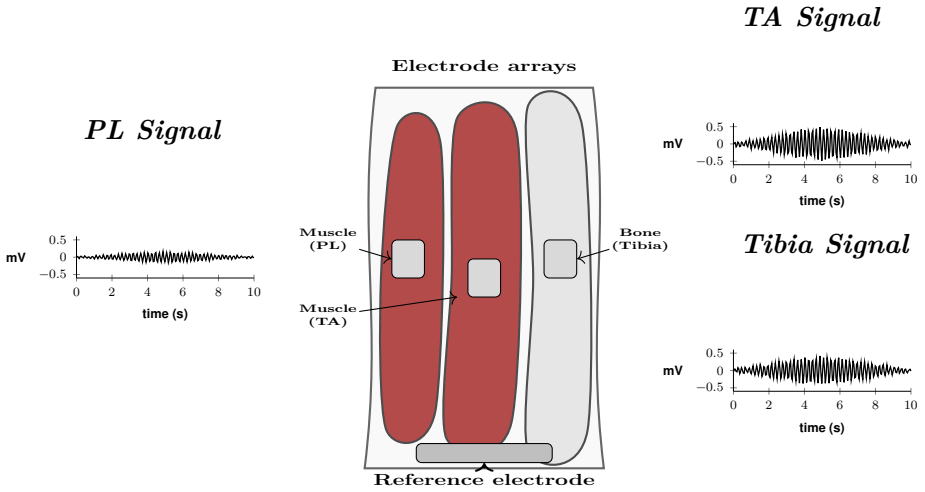

Figure 2.15: Surface electromyography recordings from the Tibialis Anterior, Peroneus Longus (PL), and Tibia bone during muscle activation. The presence of electrical activity on the Tibia (bone) demonstrates crosstalk i.e. the detection of signals originating from nearby active muscles rather than from the recording site itself [26].

Experimental studies have demonstrated that in the forearm, crosstalk contamination can be substantial: between 32% to 50% in wrist-dedicated flexors and 11% to 25% in digit-dedicated flexors during gripping tasks [36], with approximately 40% common signal detected between adjacent electrode pairs (3 cm spacing) around the proximal forearm [37]. These values are considerably higher than those observed in larger, more spatially separated muscles such as the leg [31], reflecting the anatomical proximity and overlapping functions of extrinsic hand muscles. The amount of crosstalk depends on several factors:

- **Inter-Electrode Distance (IED)**: Larger IED increases the detection volume

and susceptibility to distant sources, but very small IED may compromise signal quality due to electrode shunting [14].

- **Electrode location**: Placement near muscle boundaries increases crosstalk from adjacent muscles. Optimal placement in the midline of the muscle belly, between the innervation zone and tendon, minimizes this effect [31].

- **Subcutaneous tissue thickness**: Thicker fat layers increase spatial filtering, which paradoxically may not reduce crosstalk proportionally due to the prevalence of extinction components in distant signals [35].

- **Muscle architecture**: Pennate muscles with oblique fiber orientations may generate stronger lateral signal spread compared to parallel fibered muscles [14].

In EMG pattern recognition applications for hand and wrist gestures using forearm recordings, crosstalk represents a particularly challenging issue because:

1. Multiple extrinsic muscles with overlapping functions are densely packed within a limited anatomical space.

2. Many hand gestures involve coordinated activation of both flexor and extensor muscle groups, increasing the likelihood of crosstalk contamination.

3. The relationship between gesture kinematics and muscle activation is many-to-many: different muscles can contribute to similar movements, and the same muscle may participate in different gestures.

While crosstalk cannot be completely eliminated with conventional sEMG electrode configurations, understanding its sources and characteristics is essential for interpreting the recorded signals and developing robust pattern recognition algorithms that can accommodate this inherent limitation [25].

Muscle Fatigue and Spectral Changes

Muscle fatigue is characterized by a progressive decline in the force-generating capacity during sustained or repeated contractions, and is accompanied by distinctive changes in the sEMG signal that occur before overt force decrements become apparent [31, 14]. Understanding how fatigue shapes the EMG signal is important for interpreting recordings during prolonged or repetitive hand gesture tasks.

Physiological Mechanisms The primary physiological changes during muscle fatigue that affect the EMG signal include [38]:

- **Reduced muscle fiber Conduction Velocity (CV)**: Accumulation of extracellular K^+ and intracellular H^+ alters membrane excitability, progressively reducing the CV of action potentials along muscle fibers [14].
- **Increased MU synchronization**: Fatigue may be associated with increased common drive and synchronization among active MUs [35].

Spectral Manifestations The most prominent and well documented effect of muscle fatigue on the surface EMG is a progressive shift of the power spectrum toward lower frequencies during sustained contractions [31, 14]. This spectral compression is primarily attributable to the decrease in muscle fiber CV, which scales the EMG power spectrum [35].

The spectral changes during fatigue are typically quantified using spectral descriptors such as:

- **Median Frequency (MNF)**: The frequency that divides the power spectrum into two regions of equal power. Decreases monotonically during sustained contractions above ∼30% Maximum Voluntary Contraction (MVC) [31].
- **Mean Frequency (MDF)**: The average frequency of the power spectrum, weighted by the spectral power at each frequency. Shows similar declining trend but with higher variability than median frequency [35].

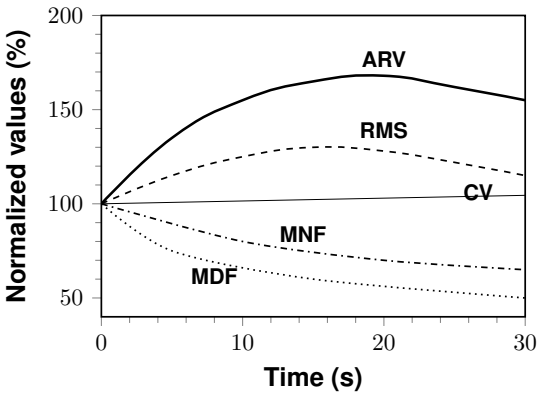

Figure 2.16: The curves illustrate the spectral and amplitude manifestations of neuromuscular fatigue, with characteristic decreases in Conduction Velocity (CV), Median Frequency (MNF), Mean Frequency (MDF) and increases in amplitude parameters (Average Rectified Value (ARV), Root Mean Square (RMS)) [26].

For pattern recognition applications, fatigue induced changes in EMG features (both spectral and amplitude based) represent a source of within session non stationarity that can degrade classifier performance over time. Robust algorithms

must either be invariant to these changes or adapt to them through online learning strategies. Understanding the physiological origins of fatigue related signal modifications is essential for developing such adaptive approaches.

Implications for Signal Processing

The interference phenomena associated with the superposition and partial cancellation of MUAPs have important implications for the interpretation and processing of sEMG signals. In particular, the asynchronous summation of multiple MU contributions limits the extent to which conventional signal features can be directly interpreted in physiological terms.

Time-domain amplitude based features, which are widely employed in myoelectric pattern recognition, provide an effective representation of overall muscle activation but remain an indirect measure of the underlying neural drive. Their interpretation is influenced by amplitude cancellation, electrode configuration, and volume conductor properties, whose combined effects cannot be reliably predicted or compensated [35, 25].

These considerations do not diminish the practical utility of conventional EMG features, but they motivate the exploration of complementary signal representations that aim to more explicitly capture neural activation mechanisms. Advanced signal processing techniques, including EMG decomposition and deconvolution, have therefore been proposed to extract information related to MU firing behavior from the interference signal. In particular, single-channel EMG deconvolution seeks to estimate MU firing times, yielding features that may be less sensitive to volume conductor variability and more directly linked to neural control strategies [8, 39]. These approaches are reviewed in Chapter 3.

Chapter 3

State of the Art

3.1 Introduction

Hand gesture recognition systems based on Surface Electromyography (sEMG) have attracted considerable attention in the research community and in prosthetics applications. These systems seek to translate muscle electrical activity into control signals for prosthetic devices, exoskeletons, or human-computer interfaces. Despite significant technological advances, challenges persist, particularly regarding the high abandonment rates of prosthetic devices [3, 40].

This chapter presents the current state of research in Electromyography (EMG) based hand gesture recognition, with particular emphasis on applications in prosthetic control. The chapter examines the principal challenges faced by such systems, including signal acquisition, pattern recognition methodologies, classification algorithms, and factors affecting real time performance.

3.2 Surface EMG for Hand Gesture Recognition

3.2.1 Applications and Challenges

sEMG represents a non invasive method for detecting the electrical activity of muscles through electrodes placed on the skin surface. This technique has found widespread application in the control of myoelectric prostheses, exoskeletons, and human-computer interfaces [40, 41].

The primary application of sEMG based gesture recognition lies in prosthetic hand control. Despite considerable research efforts and technological advances, the overall abandonment rate for upper limb prostheses remains approximately 44%, with myoelectric devices accounting for 92% of abandoned prostheses [3, 6]. This high abandonment rate is primarily attributed to inadequate control mechanisms

that fail to manage the prostheses' available degrees of freedom in a natural and intuitive manner [40].

Several challenges affect the performance of sEMG based hand gesture recognition systems:

- **Signal variability**: sEMG signals exhibit substantial variability due to factors such as muscle fatigue, electrode displacement, skin impedance changes, and variations in muscle contraction intensity [42, 43].
- **Motion artifacts**: Movement artifacts introduce noise into the signal, complicating the extraction of useful information [44].
- **Inter-subject variability**: Differences in muscle anatomy, skin properties, and electrode placement across individuals lead to significant variations in signal characteristics [45, 46].
- **Crosstalk**: As described in Section 2.2.3 Electrical activity from nearby muscles can contaminate the signal from the target muscle making it harder to detect activation patterns for each gesture..

3.2.2 High Abandonment Rates of Prostheses

Recent studies have demonstrated that innovations in prosthetic technology over the past decade have not significantly reduced abandonment rates [3]. A comprehensive study examining prosthetic usage among upper-limb amputees revealed that 44% of users rejected their devices, with comfort (60.87%), weight (52.17%), and function (43.48%) identified as the primary complaints [3].

The rejection rate varies according to the level of amputation, with below-elbow amputees showing an acceptance rate of 71.4% compared to 50% for above elbow amputees [3]. Prosthetic training quality emerged as a significant factor, with perceived training usefulness showing a strong correlation with device acceptance ($p = 0.003$) [3].

These findings underscore the necessity for improved control strategies that are intuitive, responsive, and capable of managing multiple degrees of freedom simultaneously, thereby reducing the cognitive burden on the user [40, 6].

3.3 Pattern Recognition Pipeline

3.3.1 Standard Classification Pipeline

The conventional pattern recognition pipeline for EMG based hand gesture recognition comprises several fundamental stages [42, 44]:

1. **Signal acquisition**: sEMG signals are recorded using surface electrodes positioned on the forearm or residual limb.

2. **Preprocessing**: Raw signals undergo filtering to remove noise, artifacts, and baseline drift. Typical preprocessing includes bandpass filtering (typically 20-500 Hz) and notch filtering at 50/60 Hz to eliminate power line interference [10].

3. **Segmentation**: The continuous signal is divided into analysis windows (epochs). Window lengths typically range from 150 ms to 300 ms, representing a trade off between classification accuracy and system responsiveness [42, 43].

4. **Feature extraction**: Relevant features are extracted from each window. These may include time-domain features (e.g., mean absolute value, root mean square, waveform length), frequency-domain features (e.g., mean frequency, median frequency), or time-frequency features (e.g., wavelet coefficients) [45].

5. **Dimensionality reduction**: Feature selection or extraction techniques are applied to reduce computational complexity and remove redundant information [45].

6. **Classification**: A trained classifier assigns the feature vector to one of the predefined gesture classes [42].

7. **Post-processing**: Decision smoothing or majority voting may be applied to reduce spurious classifications [41].

3.4 Classifiers

3.4.1 Linear Discriminant Analysis (LDA)

Linear Discriminant Analysis (LDA) is among the most widely used classifiers in EMG based pattern recognition for hand gesture and prosthesis control, serving as both a practical baseline and the gold standard for benchmarking [47, 48]. Its popularity stems from computational simplicity and real-time feasibility on embedded hardware [47].

Theoretically, LDA is a linear generative classifier assuming class-conditional Gaussian distributions with a shared covariance matrix across classes [47]. In practice, LDA is implemented as a multi-class regressor (one-vs-all) paired with compact time-domain features extracted over short sliding windows (typically 250 ms with 200 ms overlap), which minimizes latency and computational burden [48]. A key advantage is that LDA can provide class scores or posterior probabilities, enabling confidence-based rejection (abstention) to filter uncertain predictions [49].

Performance data consistently show high accuracy: one optimization study on upper-limb prosthetic control reported LDA achieving 98.5% classification (std 1.4) and 98.5% F1-score (std 1.5) with 5 electrodes, with 39.5% mean abstention on healthy subjects [42]. On three trans-radial amputees, LDA yielded classification accuracy of 94.7–96.5% with similar abstention rates [42]. In the Hannes prosthesis study, LDA showed no significant performance difference from non-linear alternatives (e.g., Non-Linear Logistic Regression (NLR)) when hyperparameters were optimized, confirming its role as a reliable reference algorithm [48].

Post-processing (e.g., majority vote) further stabilizes real-time control by reducing isolated misclassifications [47]. However, performance can degrade outside laboratory conditions due to electrode shift, muscle fatigue, and limb position variations, underscoring the importance of robustness oriented design choices [47].

3.4.2 Support Vector Machine (SVM)

Support Vector Machine (SVM) construct separating hyperplanes that maximize the margin between classes, which generally yields good generalization in high-dimensional EMG feature spaces [42, 47]. In hand gesture recognition, this is beneficial given the noisy, correlated, multi-channel nature of the input.

A key strength of SVM is the ability to model non linear decision boundaries via kernel functions, most commonly radial basis function and low degree polynomial kernels, which provide flexible yet computationally feasible decision surfaces for real time control [43]. Reported accuracies for EMG based gesture recognition range from about 83.5% up to 98% in able bodied subjects [42], and SVM often outperforms k-Nearest Neighbor methods, especially for complex movement patterns and generalization to unseen data [43]. In trans-radial amputees, mean accuracies between 73% and 84% have been obtained in the Hannes prosthesis framework [48].

At the same time, SVM typically attains its best performance only when a relatively larger number of EMG channels is available: accuracy tends to saturate with 5–6 electrodes, whereas Non-Linear Logistic Regression reaches comparable performance with 3–4 sensors [48]. This increases interface complexity and sensitivity to electrode placement. Furthermore, because the method produces discrete decisions rather than calibrated probabilities, abstention or confidence based rejection must be realized through additional post-processing stages (e.g., thresholds on the decision function or temporal voting) rather than being inherently embedded in the model [47].

3.4.3 Non-Linear Logistic Regression (NLR)

NLR extends standard logistic regression by augmenting the input with polynomial combinations of theEMGchannels, so that non linear class boundaries can be

represented while still producing class probabilities as outputs [42]. In practice, low polynomial degrees (e.g.D=2) are generally sufficient for EMG-based hand gesture recognition and avoid an unnecessary growth of the feature space [42].

In comparative evaluations on upper-limb prosthetic control, NLR has been shown to reach high performance with relatively few electrodes. In one study, a configuration with three EMG channels and D=2 achieved mean accuracies close to 99% in able bodied subjects, with an average Abstention Rate (AR) of about 67% when a conservative likelihood threshold was applied [42]. Alternative methods such as LDA, SVM, Artificial Neural Network (ANN) and recursive least squares required four to six electrodes to reach their performance plateau and generally operated with lower abstention levels for comparable settings [42].

NLR presents similar robustness limitations to the other feature based methods considered, since performance depends on the chosen EMG feature set and can degrade under distribution shifts induced by electrode displacement, posture changes or fatigue [42, 47]. In addition, increasing the polynomial degree beyond D=2 rapidly inflates the number of derived features without statistically significant improvements in F1-score, which raises computational cost and the risk of overfitting [42].

3.4.4 Deep Learning Techniques

Deep learning architectures, encompassing both Artificial Neural Networks (ANNs) and Convolutional Neural Networks (CNNs), have demonstrated promising results in EMG based gesture recognition [44, 50]. These methods leverage multiple layers of non linear transformations to automatically extract hierarchical feature representations from EMG signals.

Multi-Layer Perceptrons (MLPs), a fundamental form of ANN, with 2-3 hidden layers containing 40-64 nodes have achieved accuracies of 87.25% [50]. CNNs have emerged as powerful tools for automatic feature extraction from raw EMG signals, eliminating the need for manual feature engineering [44, 50]. A study employing CNN architecture achieved 96.48% accuracy using R-CNN network with wavelet transform feature extraction [50].

Hybrid architectures combining CNN with Long Short-Term Memory (LSTM) networks have demonstrated exceptional performance. A CNN-LSTM model achieved approximately 99% accuracy on the DualMyo dataset and 97% on multi-subject datasets [50]. The CNN component extracts spatial features while LSTM captures temporal dynamics, creating a comprehensive representation of gesture patterns [44].

However, deep learning approaches present significant limitations for practical implementation. These methods require substantial computational resources and large training datasets, typically comprising thousands of labeled samples to achieve optimal performance [43]. The need for extensive data collection and

computational power potentially limits their applicability in embedded systems or real-time applications, particularly for portable prosthetic devices where energy efficiency and rapid response times are critical requirements.

3.5 Abstention Rate and Delay Issues

3.5.1 High Abstention Rate: Advantages and Challenges

Abstention, also referred to as rejection or confidence-based classification, represents a mechanism whereby the classifier refrains from making a decision when confidence falls below a predetermined threshold [42]. This approach aims to reduce classification errors by rejecting uncertain predictions. Importantly, abstention mechanisms are naturally compatible with regression-based methods such as LDA and NLR, which output continuous probability distributions. In contrast, pure classifiers like SVM require additional post-processing to implement abstention, as they produce discrete class predictions without inherent confidence measures.

The implementation of abstention mechanisms demonstrates several advantages:

- **Improved classification accuracy**: By rejecting low-confidence predictions, overall accuracy increases. Studies show that NLR with abstention achieved 99.0% classification accuracy compared to 80.3% for ANN without abstention [42].

- **Reduced false activations**: Abstention prevents unintended movements caused by misclassifications, particularly important for prosthetic control where safety is paramount [41].

- **Robustness to signal variability**: Abstention mechanisms help maintain system stability during periods of poor signal quality or atypical muscle activation patterns [42].

However, high ARs introduce significant challenges:

- **Increased response delay**: When a high percentage of windows are rejected, the effective sampling rate decreases, introducing latency between user intention and device response [42].

- **User frustration**: Frequent rejection of intended gestures can lead to user dissatisfaction and contribute to device abandonment [3].

3.5.2 Offline vs Online Classification Performance

A critical distinction exists between offline and online (real-time) classification performance that significantly impacts practical system deployment. Offline classification involves evaluating algorithms on pre-recorded datasets with fixed temporal characteristics, while online systems must operate in real-time with dynamic signal conditions and immediate feedback requirements.

Real-time systems face several challenges that degrade performance relative to offline benchmarks [51]. Studies comparing offline and online performance reveal significant accuracy disparities. Research by Abbaspour et al. demonstrated that while LDA achieved 97.9% offline accuracy, real-time accuracy dropped to 69.3% [52]. Similarly, Maximum Likelihood Estimation (MLE) showed 97.0% offline accuracy but only 68.4% in real-time conditions [52]. These performance gaps typically ranging from 10-20 percentage points stem from multiple factors including reduced temporal information availability, computational constraints, and the inability to leverage complete signal segments in real time decision-making [53].

Window length represents a critical parameter influencing this performance trade-off. Smith et al. investigated analysis windows ranging from 50 to 550 ms and found that classification error decreased significantly with longer windows ($p < 0.01$) [54]. However, longer windows introduce controller delays that can impair real-time usability. Their research demonstrated that window length had a more pronounced effect on two-channel systems compared to four-channel configurations, suggesting that increased spatial sampling can partially compensate for reduced temporal information [54]. Studies indicate that windows of 100-300 ms contain the highest information content for EMG signals, representing an optimal balance between accuracy and responsiveness for real-time prosthetic applications [55]. Window lengths over 300 mS introduce a noticeable delay for the user and it has been proven that there was no significant enhancement for window lengths larger than 300 ms [56]

3.5.3 Impact of Muscle Fatigue on Classification Performance

Muscle fatigue represents a significant challenge for real-time myoelectric control systems, as it induces systematic changes in the characteristics of the EMG signal that degrade the accuracy of the classification over time. The performance degradation manifests with substantial quantitative impact: classifiers trained exclusively on non-fatigued data exhibit dramatic accuracy drops when tested on fatigued signals. Ao et al. demonstrated that K-NN, SVM, Decision Tree, and CNN classifiers achieved 96.7%, 89.0%, 87.3%, and 97.5% accuracy respectively on non-fatigued test data, but performance plummeted to 53.3%, 55.4%, 45.8%, and

64.8% when evaluated on fatigued conditions, representing accuracy reductions of 30-40% [57]. Similarly, Díaz-Amador and Mendoza-Reyes reported that while LDA achieved approximately 93% accuracy under non-fatigued conditions, classification performance deteriorated progressively with fatigue: moderate fatigue conditions yielded an accuracy $\approx 79\%$, while high fatigue states resulted in values of accuracy $\approx 61\%$ [58].

These substantial performance decrements come from fundamental physiological changes in the EMG signal during sustained contractions [10, 38]. Muscle fatigue induces concurrent alterations on the feature distributions on which classifiers are trained [57, 58, 59]. Separability Index and Repeatability Index—metrics quantifying interclass distances and pattern consistency respectively—both decrease systematically as fatigue level increases [58]. This explains the classifier performance drop, as decision boundaries optimized for non-fatigued feature spaces become increasingly misaligned with fatigued signal distributions. Song et al. emphasized that pattern recognition systems utilizing signals from multiple EMG channels are especially susceptible to fatigue effects, as fatigue levels differ across muscles even within the same limb [60].

Several mitigation strategies have demonstrated efficacy in reducing fatigue-related degradation, though each approach presents inherent limitations. Multi-condition training, incorporating data from non-fatigued, moderately fatigued, and highly fatigued states during classifier training, increased accuracy by 25% compared to conventional training [58]. However, this approach incurred a critical trade-off: accuracy on non-fatigued data decreased from 93% to 91.3%, indicating that improved robustness to fatigue necessarily degrades performance in optimal conditions [58]. Selective classification schemes employing separate classifiers for each fatigue condition maintained stable performance across all states, but presented two major shortcomings: first, the requirement to train and maintain three independent classifiers increases computational load substantially; second, and more critically, this approach requires an additional classifier to assess fatigue level in real-time, introducing a dual stage classification scheme with associated latency and potential error propagation.

Adaptive classification strategies that continuously retrain classifiers with correctly classified samples demonstrated robust performance, maintaining accuracy and sensitivity exceeding 90% while limiting false positive rates below 10% throughout sustained contractions [58]. However, determining when adaptation should occur remains a challenge. Premature adaptation causes excessive retraining risks and incorporates misclassified samples that corrupt the model, while delayed adaptation fails to prevent performance degradation [58]. Data augmentation approaches using Wasserstein Generative Adversarial Networks with Gradient Penalty to transform fatigued EMG signals into non-fatigued distributions achieved substantial recovery, improving classification accuracy from 64.8% to 92.1% for

CNN classifiers [57]. Despite these promising results, this method requires extensive computational resources for training, necessitating high-performance GPU hardware that precludes embedded implementation in wearable prosthetic systems [57]. Song et al. proposed adapting hyperbox boundaries in Fuzzy Min-Max Neural Networks based on observed consistent temporal trends in feature variations during sustained contractions [60], though this approach requires user specific calibration and assumes consistent fatigue progression patterns that may not generalize across individuals or varying contraction intensities.

3.6 Electrode Configurations

3.6.1 High-Density vs. Low-Density EMG

Electrode configuration significantly influences system performance, cost, and usability. Two principal approaches exist: Low-Density Electromyography (LD-EMG) and High-Density Electromyography (HD-EMG).

Low-Density EMG

LD-EMG typically employs 2-8 bipolar electrode pairs positioned over target muscles [6]. This configuration offers several advantages:

- **Simplicity**: Fewer electrodes reduce setup time and complexity
- **Cost effectiveness**: Lower hardware costs make the system more accessible.
- **Portability**: Compact systems suitable for wearable applications [61].
- **Real time capability**: Reduced data processing requirements lower the computational burden, reducing classification delay and improving user experience.

Research demonstrates that carefully optimized LD-EMG configurations can achieve high classification accuracy. Prakash and Kunju systematically investigated electrode optimization for wrist-worn gesture recognition systems, demonstrating that four electrodes strategically placed at the wrist achieved 85.9% accuracy, 85.93% precision, 86.27% recall, and 86.09% F1-score when classifying 17 hand gestures using an SVM classifier [61]. Their analysis revealed that reducing electrodes from six to four maintained comparable performance while simplifying device complexity and enhancing user comfort. A separate study using only 3 EMG channels with NLR classifier achieved 99.0% classification accuracy for 10 gestures [42].In summary, these findings confirm that LD-EMG systems can deliver high classification performance comparable to complex arrays while offering superior

computational efficiency and hardware affordability. However, the reduced spatial resolution inherently limits the information available for source separation and increases susceptibility to signal non-stationarities such as electrode shifts or muscle fatigue. Consequently, significant potential for improvement lies in developing robust signal processing strategies and adaptive algorithms that can compensate for these physiological variations without increasing hardware complexity.

Figure 3.1: Low-density EMG armband for gesture recognition. The Myo armband (Thalmic Labs) features a sparse circumferential array of 8 dry bipolar sEMG sensors. This configuration is representative of commercial low-density interfaces, providing discrete muscle activity signals for classifying hand gestures with low computational overhead and setup time [62].

High-Density EMG

HD-EMG employs electrode grids or arrays (typically 64-128 electrodes) with inter-electrode distances of 8-10 mm [6]. This configuration provides several distinct advantages over conventional sparse electrode systems:

- **Spatial resolution**: Detailed mapping of muscle activity distribution enables capture of fine-grained spatial patterns correlated with movement information [63, 4].
- **Motor Unit (MU) identification**: Capability to decompose signals into individual MU contributions, facilitating physiologically interpretable control strategies [64].
- **Robustness to electrode shift**: Redundant spatial information compensates for electrode displacement, with advanced deep learning architectures maintaining above 70% accuracy even with 75% channel dropout [6].

- **Advanced feature extraction**: Spatial patterns enable sophisticated analysis techniques, including convolutional processing of sEMG images that exploit both instantaneous and temporal spatial distributions [6].

Figure 3.2: High-density EMG sleeve interface. The NeuroLife® sleeve features a dense grid of up to 150 embedded electrodes covering the entire forearm circumference. This configuration enables the high-resolution capture of spatiotemporal muscle activation patterns from both superficial and deep muscle groups, providing a rich dataset for decoding complex hand gestures and continuous joint kinematics compared to traditional sparse electrode setups. [65].

Studies employing HD-EMG with machine learning and deep learning algorithms report performance metrics ranging substantially depending on gesture complexity, validation protocols, and system configurations. Qiu et al. systematically investigated electrode placement effects using 256-channel HD-sEMG across three forearm regions (distal wrist, mid-forearm, proximal elbow), revealing that single-degree-of-freedom gestures achieved 98.63% (intra-subject), 79.73% (inter-day), and 75.47% (inter-subject) accuracy when sampled from the distal wrist, whereas multi-finger synergistic gestures performed optimally with signals from mid-forearm and proximal elbow regions [4]. Deep learning approaches have demonstrated particular efficacy in exploiting HD-EMG spatial characteristics: Chen et al. showed that 3D CNNs processing sequential sEMG images achieved 98.9% accuracy on static gestures (CapgMyo DB-+a dataset) and 90.7% on dynamic finger movements (CSL-HDEMG dataset containing 27 gestures), outperforming 2D CNN architectures by 15.3-18.6% on dynamic tasks due to superior spatiotemporal feature extraction capabilities [66]. Massa et al. employed Graph Neural Networks to process raw 128-channel HD-EMG data for recognizing 65 fine-grained gestures from 20 volunteers, achieving 95% accuracy while reducing computational cost by 53% through explainable AI based graph topology refinement [67]. Yang et al. demonstrated that HD-EMG enables highly responsive gesture-based control, with fully-connected neural networks achieving 99.6% initial accuracy and 93.8% after

one hour of continuous use across 13 participants controlling mobile manipulator robots [68]. Performance metrics across the literature indicate that HD-EMG systems achieve accuracy ranges of 75-99% depending on gesture set complexity (8-65 gestures), subject population (able-bodied vs. clinical), and validation rigor (intra-subject vs. inter-subject) [69, 50].

However, HD-EMG presents significant challenges that limit its practical deployment in wearable prosthetic systems. Massa et al. conducted systematic analysis on 128-channel systems, revealing that feature extraction from raw HD-EMG signals requires 62.88 ms on average

[67]. Their explainable AI analysis demonstrated that spatial redundancy inherent to high electrode density degrades classification efficiency, with only a small subset (approximately 50%) of electrodes contributing meaningfully to classification decisions while the majority provide redundant or noisy information. This redundancy necessitates computationally expensive dimensionality reduction algorithms (principal component analysis, feature selection) that introduce processing delays incompatible with the <300 ms controller delay threshold required for acceptable prosthetic user experience [67, 6]. Additional shortcomings include:

- **Complexity**: Requires sophisticated acquisition hardware (64-128 channel amplifiers, high sampling rates 1-2 kHz) and complex signal processing pipelines involving spatial filtering, artifact rejection, and multi-dimensional feature extraction [10, 67].
- **Cost**: High-density acquisition systems are substantially more expensive than LD-EMG configurations, with commercial 128-channel systems costing 10-20× more than 4-8 channel alternatives [6].
- **Setup time**: Proper electrode grid placement, skin preparation for 64-128 sites, and conductive gel application requires substantially more time respect to a LD-EMG setup.
- **Computational burden**: Processing multiple channels requires various optimization approaches, including computationally efficient feature design, rapid decomposition algorithms, and employing parallel computing technologies such as FPGA and GPU to accelerate the processing of HD-EMG signals [70].
- **Signal interpretability**: The high-dimensional feature space (>500 features) obscures physiological interpretation, complicating clinical assessment of neuromuscular function and system debugging during deployment [67].
- **Practicality**: Not suitable for daily use in prosthetic applications due to gel drying (requiring reapplication every 2-4 hours), electrode grid maintenance, cable management for 64-128 channels, and user discomfort during prolonged wear [40, 6].

3.7 High-Density EMG Decomposition

3.7.1 Principles and Methodology

HD-EMG decomposition aims to separate the interference EMG signal into constituent Motor Unit Action Potential (MUAP) trains [71]. This process provides direct access to the neural drive to muscles, enabling detailed investigation of motor control strategies.

Interestingly, the requirements for successful decomposition do not always match with those necessary for extracting global features from the EMG [64, 72]. While global feature extraction typically benefits from bipolar electrode configurations and larger inter-electrode distances to reduce volume conductor effects, decomposition requires maximizing the spatial variation in action potential shapes across recording channels. For this reason, electrode arrays should be centered approximately above primary innervation zones for fusiform muscles to ensure maximal variation in action potential shapes across channels [64].

The Blind Source Separation (BSS) approach represents the widely adopted method for HD-EMG decomposition. BSS methods, particularly those based on Convolution Kernel Compensation (CKC) and Progressive FastICA Peel-off (PFP), attempt to identify MU spike trains by iteratively optimizing separation filters that maximize the sparseness of individual MU discharge patterns [73, 74].

Validation using concurrent intramuscular and sEMGrecordings demonstrates that HD-EMG decomposition achieves substantial accuracy in MU discharge identification. The two-source validation method, first introduced for intramuscular EMG [71] and later adapted for surface recordings [75], compares decomposition results from simultaneously recorded signals. Studies report sensitivity ranging from 84% to 92% across different muscles during low-force isometric contractions (2.5-20% maximum voluntary contraction) [75]. Alternative validation approaches using reconstruct-and-test procedures with synthetic signals report average accuracies of 92.5% (range 77-97%) [76].

The yield of decomposition, measured by the number of successfully identified motor units, varies substantially across studies and conditions. High-yield decomposition systems report 16-40 MUAP trains per contraction, with typical yields of 20-30 MUAP trains depending on muscle characteristics, contraction force levels, and electrode configurations [76]. Independent validation comparing CKC and PFP decomposition methods on the same recordings found high agreement (97.85±1.85% matching rate) for an average of 10.6±4.3 commonly identified motor units per trial [73], providing supportive evidence for the accuracy of both approaches.

An important limitation of current decomposition methods is that they identify only a small proportion of the active MU population. Simulation studies indicate that decomposition typically identifies approximately 23% of active MUs [5], with

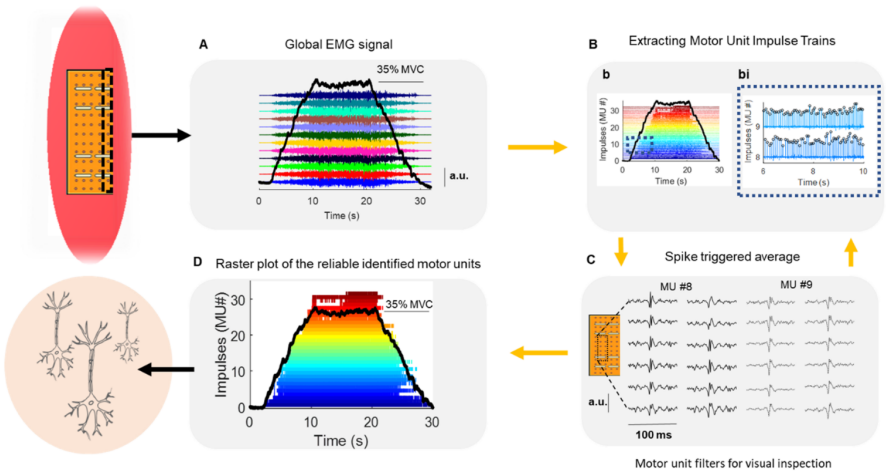

Figure 3.3: Blind source separation (BSS) pipeline. (A)sEMG signals are acquired via a high-density electrode grid during isometric contraction (e.g., 35% MVC trapezoidal profile). The raw interference pattern is a summation of multiple MUAPs. (B) BSS algorithms separate the mixed signal into constituent MU spike trains (b and bi), identifying individual firing instances. (C) Spike-triggered averaging uses the identified firing times to extract the unique MUAP waveforms across the electrode array. (D) Raster plot showing the firing patterns of reliable motor units, color-coded by recruitment threshold [64].

the detected units explaining most of the interference signal amplitude despite representing a minority of the active population. This bias toward larger, more superficial MUs means that decomposition provides limited but highly accurate information about a specific subset of the motor neuron pool.

3.7.2 Challenges and Limitations

Despite significant advances, several challenges limit the practical application of HD-EMG decomposition:

- **Signal quality requirements**: Decomposition accuracy degrades significantly with poor signal-to-noise ratio. Empirical evidence suggests that the ratio of signal power to noise power should exceed 2:1 at low contraction forces for reliable decomposition [64].
- **Contraction level limitations**: Decomposition yield typically decreases by approximately 30% at forces exceeding 35-70% maximum voluntary contraction due to the temporal and spatial superimposition of Action Potentials (APs) [77, 64]. This decline is exacerbated by the recruitment of larger, high-threshold

MUs, whose high-energy potentials obscure the smaller waveforms of low-threshold units through amplitude cancellation effects [78], thereby limiting the applicability of decomposition methods for high-force motor control studies.

- **Muscle geometry dependence**: Performance varies significantly across different muscles. Fusiform muscles with parallel fibers yield better decomposition results than pennate muscles, where fibers orient at various angles to the line of action [64]. The spatial distribution of MU territories within the muscle also influences decomposition yield, with more dispersed territories facilitating better discrimination between MUs [77].

- **Anatomical and subject-specific variations**: Subcutaneous tissue thickness substantially affects decomposition quality, with studies demonstrating a negative correlation between muscle-electrode distance and the number of identified MUs [77, 79].

- **Computational complexity**: Multi-channel algorithms have been reported to require on the order of 3–4 s of processing per 1 s of decomposed activity on a standard PC [76], and BSS/CKC-based methods for HD-EMG still involve large iterative optimizations that make fully real-time operation challenging despite recent speed-ups from fast gradient CKC variants [80].

- **Electrode array positioning**: Decomposition quality depends critically on proper electrode placement. As mentioned previously, arrays should be centered approximately above primary innervation zones for fusiform muscles to ensure maximal variation in action potential shapes across channels [64]. However, BSS methods are less sensitive to array positioning in pennate muscles due to the inherent spatial complexity of action potential propagation in these architectures.

- **Manual intervention requirement**: Current decomposition methods require expert visual inspection and manual refinement of spike trains to achieve reported accuracy levels [71]. While automated versions of decomposition algorithms exist, they typically achieve lower accuracy and reduced MU yield compared to expert-guided semi-automatic approaches [81].

- **Non-stationarity and signal variability**: MUAP waveforms may change gradually during sustained contractions due to shifts in muscle fiber position, changes in intracellular action potential characteristics during fatigue, and variations in muscle fiber conduction velocity [64]. These non-stationary characteristics challenge decomposition algorithms that assume relatively stable action potential templates. While modern BSS methods incorporate template updating mechanisms to track these changes [71], excessive variability

can degrade decomposition accuracy and necessitate more frequent manual intervention.

3.8 Conclusion

The state of the art in EMG based hand gesture recognition reveals significant progress alongside persistent challenges. While classification accuracies exceeding 95% are achievable with optimized systems, practical implementation in prosthetic devices remains constrained by factors including signal variability, computational complexity, user comfort, and real-time performance requirements.

The high abandonment rate of myoelectric prostheses (44%) underscores the gap between laboratory achievements and clinical reality [3]. Future research must address not only classification accuracy but also system robustness, user adaptability, and practical usability considerations that determine long term device acceptance.

High-density EMG decomposition offers detailed insights into motor control strategies through identification of individual MU contributions [64]. However, the computational complexity, manual intervention requirements, and practical limitations of HD-EMG systems restrict their applicability primarily to research environments. The development of alternative preprocessing approaches that can extract relevant neural information while maintaining compatibility with low-density, practical EMG systems represents a promising direction for advancing myoelectric control beyond current limitations.

Chapter 4

Single-Channel EMG Deconvolution Method

This thesis adopts a single-channel Surface Electromyography (sEMG) deconvolution framework for estimating the Cumulative Weighted Firing (CWF) activity from a Single Differential (SD) channel, and its subsequent adaptation as a preprocessing step for myoelectric control [8, 82]. The proposed method models the interference Electromyography (EMG) as the convolution between a stereotyped Motor Unit Action Potential (MUAP) kernel and an unknown firing pattern. Finally, it uses an Iterative Reweighted Least Squares (IRLS) approach to recover this pattern.

The proposed deconvolution approach trades individual Motor Unit (MU) resolution for a robust estimate of cumulative MU activity in the detection volume of a single bipolar channel. Instead of separating individual MUs, the method estimates a cumulative firing pattern with low-frequency content reflecting the distribution of MU discharge rates and high-frequency, sparse structure approximating the timing of the underlying MUAPs weighted by their amplitudes. This cumulative pattern can be computed in real time and is hypothesised to provide an alternative representation of the raw EMG signal that may be advantageous for subsequent feature extraction and pattern classification in myoelectric control applications [82].

4.1 Signal Model and Kernel Definition

Equation 2.1 can be rewritten as:

$$s(t) = \sum_{m=1}^{N} \Big(M_m(t) * F_m(t) \Big) + n(t), \tag{4.1}$$

where:

- N is the number of active MUs,
- $M_m(t)$ is the MUAP waveform of the m-th MU in SD configuration,
- $F_m(t) = \sum_j \delta(t - t_{m,j})$ is its spike train (Dirac impulses at discharge times $t_{m,j}$),
- $n(t)$ collects model errors and measurement noise,
- $*$ denotes convolution.

The proposed single-channel deconvolution method introduces two key modelling simplifications to the classical EMG model:

1. **Shape approximation**: all SD MUAPs are approximated by a single prototype waveform (kernel) $K(t)$, up to a scaling factor:

$$M_m(t) \approx A_m K(t), \tag{4.2}$$

   where A_m is an amplitude factor reflecting MU size and position relative to the detection volume.

2. **Cumulative firing pattern**: the contributions of all MUs are aggregated into a single, amplitude-weighted firing pattern

$$f(t) = \sum_{m=1}^{N} A_m F_m(t) = \sum_{m=1}^{N} \sum_j A_m \, \delta(t - t_{m,j}), \tag{4.3}$$

   so that the EMG can be approximated as

$$s(t) \approx (K * f)(t) + n(t). \tag{4.4}$$

Under this model, deconvolution reduces to estimating $f(t)$ from the measured $s(t)$, given a suitably chosen kernel $K(t)$.

Kernel choice and adaptation

For SD sEMG signals from mucles with fibres approximately parallel to the skin and electrodes placed distal to the Innervation Zone (IZ), the MUAP is well approximated by the first derivative of a Gaussian. Accordingly, the kernel is parameterized as:

$$K(t) = \frac{d}{dt}\left[\exp\left(-\frac{t^2}{2\sigma^2}\right)\right], \tag{4.5}$$

where σ controls the temporal width of the kernel and thus its Power Spectral Density (PSD). The medium-high frequency portion of the EMG PSD (typically above $\sim$40 Hz) primarily reflects MUAP shape, whereas the low-frequency part ($<\sim$40 Hz) reflects MU firing statistics. The kernel variance σ^2 is therefore tuned so that the PSD of $K(t)$ matches the medium-high-frequency PSD of the recorded EMG.

Practically, the procedure is:

1. Compute the PSD of the band-pass filtered EMG.
2. Identify a frequency band that contains most of the MUAP-related power, typically around the EMG median frequency and its spread.
3. Vary σ and select the value that minimizes the discrepancy (e.g., in least-squares sense) between the kernel PSD and the EMG PSD within this band.

This adaptive kernel design compensates, to first order, for variations in MUAP duration due to conduction velocity or tissue filtering and improves the robustness of the deconvolution.

4.2 Inverse Problem Formulation and Regularization

Given the model

$$s(t) \approx (K * f)(t) + n(t), \tag{4.6}$$

estimating $f(t)$ is an ill-posed deconvolution problem. In discrete time, for an analysis window of L samples, the convolution can be written as a matrix–vector product:

$$\mathbf{s} \approx \mathbf{A}\mathbf{x} + \mathbf{n}, \tag{4.7}$$

where:

- $\mathbf{s} \in \mathbb{R}^L$ is the EMG sample vector,
- $\mathbf{x} \in \mathbb{R}^L$ is the discretized cumulative firing pattern,
- $\mathbf{A} \in \mathbb{R}^{L \times L}$ is a lower-triangular Toeplitz matrix that contains time-shifted copies of the kernel K,
- $\mathbf{n}$ is noise and model mismatch.

A straightforward least-squares solution

$$\min_{\mathbf{x}} \|\mathbf{s} - \mathbf{A}\mathbf{x}\|_2^2 \tag{4.8}$$

is numerically unstable due to the ill-conditioning of $\mathbf{A}$. Tikhonov (ridge) regularization is introduced to stabilize the inversion by penalizing the energy of $\mathbf{x}$:

$$\min_{\mathbf{x}} \ \|\mathbf{s} - \mathbf{A}\mathbf{x}\|_2^2 + \alpha\|\mathbf{x}\|_2^2 \tag{4.9}$$

where $\alpha > 0$ controls the trade-off between data fit and solution smoothness. In practice, α is set to a small fraction (e.g. 1%) of the largest eigenvalue of $\mathbf{A}^\top\mathbf{A}$, yielding a condition number in the range 10^2–10^3 and providing a stable baseline solution. The corresponding normal equations are:

$$(\mathbf{A}^\top\mathbf{A} + \alpha\mathbf{I})\mathbf{x} = \mathbf{A}^\top\mathbf{s} \tag{4.10}$$

With solution:

$$\mathbf{x} = (\mathbf{A}^\top\mathbf{A} + \alpha\mathbf{I})^{-1}\mathbf{A}^\top\mathbf{s} \tag{4.11}$$

This L_2-regularized solution, however, tends to produce smooth, non-sparse $\mathbf{x}$ and is sensitive to outliers, which is undesirable when the target is a sequence of sharp firing bursts.

4.2.1 L1 Formulation and IRLS Optimization

To better approximate the sparse, spike-like structure of the cumulative firing pattern, the objective is reformulated in terms of the L_1 norm of an augmented residual. Defining

$$\tilde{\mathbf{A}} = \begin{bmatrix} \mathbf{A} \\ \sqrt{\lambda}\,\mathbf{I} \end{bmatrix}, \qquad \tilde{\mathbf{s}} = \begin{bmatrix} \mathbf{s} \\ \mathbf{0} \end{bmatrix}, \tag{4.12}$$

the optimization problem can be written as

$$\arg\min_{\mathbf{x}} \ \|\tilde{\mathbf{s}} - \tilde{\mathbf{A}}\mathbf{x}\|_1. \tag{4.13}$$

The L_1 norm penalizes large residuals less severely than the L_2 norm while strongly suppressing small, unnecessary fluctuations. This leads to a solution in which many entries of $\mathbf{x}$ are exactly zero, and the non-zero entries concentrate around effective firing events. However, there is no analytical solution for this L_1-regularized problem.

The proposed method therefore uses IRLS to approximate the L_1 solution via a sequence of weighted L_2 problems [8]:

1. **Initialization**:
   Compute an initial estimate $\mathbf{x}_{(0)}$ from the L2–Tikhonov solution:

$$\mathbf{x}_{(0)} = (\mathbf{A}^\top \mathbf{A} + \lambda \mathbf{I})^{-1} \mathbf{A}^\top \mathbf{s}. \tag{4.14}$$

2. **Iterative reweighting**:
   For iteration $k = 1, \ldots, K$ (typically $K \approx 10$):

   - Compute the current residual:

$$\mathbf{r}_{(k-1)} = \tilde{\mathbf{A}} \mathbf{x}_{(k-1)} - \tilde{\mathbf{s}}. \tag{4.15}$$

   - Define a diagonal weight matrix $\mathbf{W}_{(k)}$ with entries:

$$w_i = \frac{1}{\sqrt{|r_i| + \epsilon}}, \tag{4.16}$$

     where ϵ is a small constant to avoid division by zero. Large residuals get smaller weights, approximating the L_1 behaviour. In the numerical implementation, the weights are additionally normalized by a common scalar factor, which does not alter the minimizer.

   - Solve the weighted least-squares problem:

$$\mathbf{x}_{(k)} = \arg\min_{\mathbf{x}} \left\| \mathbf{W}_{(k)} (\tilde{\mathbf{s}} - \tilde{\mathbf{A}} \mathbf{x}) \right\|_2^2 \tag{4.17}$$

     This has the solution:

$$\mathbf{x}_{(k)} = \left(\tilde{\mathbf{A}}^\top (\mathbf{W}_{(k)})^2 \tilde{\mathbf{A}} \right)^{-1} \tilde{\mathbf{A}}^\top (\mathbf{W}_{(k)})^2 \tilde{\mathbf{s}}. \tag{4.18}$$

3. **Non-negativity constraint**:

   In parallel with the IRLS solution, a secondary refinement is applied using the Landweber method to enforce strict non-negativity and further stabilize the estimate [8].

   Starting from the initial Tikhonov estimate $\mathbf{x}_{(0)}$, the Landweber iteration refines the estimate through gradient descent with explicit positivity projection:

$$\mathbf{x}_{\text{pos},j} = \max\left(0, \mathbf{x}_{\text{pos},j-1} - \mu \mathbf{A}^\top (\mathbf{A} \mathbf{x}_{\text{pos},j-1} - \mathbf{s})\right), \qquad \mathbf{x}_{\text{pos},0} = \mathbf{x}_{(0)}. \tag{4.19}$$

   where $\mathbf{A}$ is the unregularized convolution matrix, $\mu = 0.9/(100\lambda)$ is the step size ensuring convergence, and the $\max(0, \cdot)$ operator enforces non-negativity element-wise [8].

The algorithm runs for a fixed number of iterations (typically 10), after which:

$$\mathbf{x}_{\text{dec}} = \mathbf{x}_{\text{pos},J} \tag{4.20}$$

is taken as the final non-negative estimate. This projection is essential because MU firing patterns cannot contain negative spike amplitudes; any negative components are artifacts that would distort downstream feature extraction [8].

After a fixed small number of iterations (in this case 10), the algorithm converges to a sparse estimate $\mathbf{x}_{\text{dec}}$ that serves as the deconvolved signal, i.e., the estimated CWF.

4.3 Online Implementation of the Deconvolution Algorithm

The deconvolution procedure described in the previous section was originally implemented for offline analysis, where the full signal is available in advance and can be processed without strict timing constraints. In that formulation, the kernel is estimated directly from the signal being analyzed, a Toeplitz convolution matrix is then constructed from the estimated kernel, and the solution is refined through multiple iterations of both Landweber projection and IRLS. While this approach is suitable for offline processing, it is not directly compatible with real-time operation, where the signal must be handled causally and each update must be completed within the available decision interval.

To make the method compatible with online use, the implementation was reorganized into two stages: a calibration stage, in which the channel-specific operators are estimated and precomputed, and a streaming stage, in which those operators are reused to process the incoming signal in real time. The purpose of this reformulation was to preserve the deconvolution-based representation while shifting the most expensive computations outside the real-time loop.

During calibration, the kernel was estimated independently for each differential channel pair using the same bandwidth-based procedure adopted in the offline algorithm. Once the kernel was obtained, the matrices required by the deconvolution stage were computed once and stored for later reuse. In particular, the convolution matrix $\mathbf{A}$ associated with the estimated kernel, the regularized system matrix, and the Gram matrix $\mathbf{G} = \mathbf{B}^T\mathbf{B}$ were precomputed. In addition, the pseudo-inverse matrix $\mathbf{P}_{\text{inv}} = (\mathbf{A}^T\mathbf{A})^{-1}\mathbf{A}^T$ was calculated in advance to evaluate the initial ℓ_2 solution. This allowed the baseline signal estimation to be performed online through a fixed matrix-vector multiplication ($\mathbf{x} = \mathbf{P}_{\text{inv}}\mathbf{b}$), rather than recomputing the corresponding linear solve at each update step. The Tukey window used to taper the input segment was also generated in advance. As a result, the online stage did

not need to rebuild or invert system matrices at each update step, allowing the deconvolution of each new window to be executed as a sequence of fixed operator applications.

A second modification concerned the temporal organization of the processing windows. The offline implementation operated on short overlapping epochs and stitched their valid portions to reconstruct the full deconvolved signal. In the online version, this idea was reformulated in a causal manner using a rolling buffer. The real-time pipeline processed windows of 250 ms (500 samples at 2 kHz) with 150 ms overlap (300 samples), resulting in a hop size of 100 ms (200 samples). At each update, the newest chunk of data was appended to the buffer and the oldest samples were discarded, so that the deconvolution stage always operated on the most recent 250 ms interval while producing an updated output every 100 ms.

Only the central valid portion of each processed window was retained in the streaming output. After deconvolving the current 250 ms buffer, the samples corresponding to the overlap margins were discarded and only the central segment of duration equal to the hop size was written into the stitched deconvolution streams. With an overlap of 150 ms, this corresponds to trimming 75 ms from each boundary and keeping the central 100 ms portion. This strategy follows the same approach as the offline application of the method, namely reducing boundary artifacts and ensuring temporal consistency across consecutive windows. The resulting stitched streams were then used for feature extraction and classification.

In the present implementation, the same sliding-window structure was used for both deconvolution and the subsequent feature extraction stage. This choice was mainly motivated by the need to preserve a simple and synchronized pipeline, while also ensuring a direct comparison between raw and deconvolved signal representations under identical temporal conditions. Although deconvolution and classification may in principle benefit from different temporal supports, decoupling the two stages would have introduced additional buffering, synchronization, and hyperparameter selection requirements, making it more difficult to isolate the specific contribution of the deconvolution stage within the scope of this work.

The iterative part of the solver was also simplified to meet the latency requirements of the online classifier. In the offline implementation, the sparse solution was obtained using 10 iterations of the IRLS method, while the positivity-constrained estimate was refined through 10 Landweber iterations. In the online implementation, the number of IRLS iterations was reduced to 2 and the Landweber refinement was reduced to 5 iterations. This choice decreased the computational cost of each update while preserving the main properties of the recovered representation.

To further reduce computational cost, matrix operations were reformulated. In the offline implementation, the weighted least-squares step in the IRLS loop is written explicitly through a diagonal weighting matrix. In the online version, the same weighting is applied through element-wise row scaling, avoiding the repeated

construction of a diagonal matrix at each iteration. Similarly, the gradient term in the Landweber projection was expressed using the precomputed Gram matrix $\mathbf{G}$, ensuring that the online computation only required matrix-vector and vector operations. The regularization parameter was also intentionally adjusted in the online implementation to maintain sparsity while preventing an ill-conditioned system matrix.

An additional adaptation of the online implementation was the update of the kernel-related matrices during operation. After the initial calibration, the precomputed matrices were reused online for each channel pair, but they were refreshed periodically every 30 s using newly acquired data. This mechanism was introduced to account for possible changes in signal characteristics during prolonged use caused by muscle fatigue. Therefore, the online deconvolution operated with fixed precomputed operators over short time intervals, while still allowing occasional recalibration of the kernel and associated matrices.

To assess the fidelity of the online approximation, the deconvolved output produced by the online implementation was compared with the corresponding offline reference on the same signal segments across the four selected channels. The two implementations showed a correlation of 0.85 ± 0.04 and a normalized root mean square error below 3%, indicating that the proposed online simplifications preserved the main structure of the deconvolved signal while reducing computational cost.

Overall, the online implementation adapts the offline deconvolution framework to real-time operation by reorganizing the main computational steps. Kernel estimation and the most computationally demanding matrix operations are moved outside the real-time loop, while the per-window solver is simplified and executed on a rolling buffer. In this way, the deconvolution stage remains compatible with the timing constraints of the online gesture-classification pipeline while producing a CWF that remains close to the offline reference.

4.4 Assumptions and Limitations

The validity of the proposed method relies on several assumptions:

- muscle fibres are approximately parallel to the skin, and the SD electrode pair is aligned along the fibre direction and placed outside the innervation zone, so that MUAPs have similar biphasic shapes and propagate unidirectionally.
- A single kernel can reasonably approximate the MUAPs of different MUs in the detection volume, up to amplitude scaling.
- Baseline artefacts and non-myoelectric noise are sufficiently suppressed by preprocessing filters.

Violations of these assumptions (e.g. strong pinnation, multiple innervation zones under the electrodes, severe misalignment, or dominant movement artefacts) can degrade the interpretability of the deconvolution output. Nonetheless, both simulation and experimental validations indicate that the proposed method remains surprisingly robust to moderate deviations from the ideal model, and can extract useful MU population information from single-channel sEMG in realistic conditions [39].

Chapter 5

EMG Data Acquisition and Feature Extraction

This study utilizes data from the publicly available Reach&Grasp dataset [7], a multimodal repository containing electromyographic, kinematic, and tactile recordings from healthy subjects performing various upper-limb movements. The dataset was originally collected at the Rehab Technologies Lab of the Istituto Italiano di Tecnologia (Genova, Italy) and is structured according to the Brain Imaging Data Structure (BIDS) standard. While the complete dataset encompasses multiple recording modalities, this work focuses exclusively on the Surface Electromyography (sEMG) data from a subset of simple wrist and hand movements.

5.1 Participants and Experimental Protocol

The Reach&Grasp dataset includes recordings from 10 healthy, right-handed participants (4 females and 6 males) aged 24-36 years (mean $\pm$ SD: 30.3 ± 4.0 years), with no known neuromuscular impairments. All participants provided written informed consent, and all experimental procedures were approved by the local ethical committee (CER Liguria Ref. 11554, October 18, 2021) in accordance with the Declaration of Helsinki.

For the present study, a subset of six simple wrist and hand movements was selected from the original dataset, which contains 16 different tasks ranging from simple single-degree-of-freedom movements to complex reaching and manipulation actions. The six gestures analyzed were:

- **Hand opening (HO)**: extension of the fingers
- **Hand closing (HC)**: flexion of the fingers

- **Wrist extension (WE)**: lifting the hand backwards so that the back of the hand faces the forearm
- **Wrist flexion (WF)**: bending the hand so that the palm faces the forearm
- **Wrist pronation (WP)**: rotating the forearm so the palm faces downward
- **Wrist supination (WS)**: rotating the forearm so the palm faces upward
- **Rest (R)**: relaxed state with the elbow flexed at 90° and wrist aligned with the forearm

Each participant performed 10 consecutive repetitions of each gesture within the same recording session. Subjects were instructed to start from a standardized rest position with the elbow flexed at 90° and the wrist aligned with the forearm, and to execute each movement at a natural, self-selected pace without strict timing constraints. A 1-minute rest period was provided between tasks to prevent muscular fatigue. The average execution time for simple movements was 1.53 ± 0.66 seconds per repetition [7].

While the original dataset includes additional modalities (motion capture kinematics via Vicon system and CyberGlove, as well as tactile data from a sensorized glove), these were used exclusively during the labeling and segmentation process to identify precise onset and offset times for each gesture. The analyses presented in this work rely solely on the sEMG signals.

5.2 EMG Acquisition Setup

High-density surface Electromyography (EMG) was recorded using the Sessantaquattro system (OT Bioelettronica, Torino, Italy), a 64-channel portable acquisition device specifically designed for multi-channel EMG recordings. The system employs two pre-gelled adhesive monopolar electrode grids (REF GR10MM0804, LOT GR0319), each consisting of 32 electrodes arranged in a configuration optimized for forearm coverage.

Although the acquisition system provides 64 monopolar channels, in this work only 8 channels forming 4 Single Differentials (SDs) were retained for feature extraction and classification, in order to keep the pipeline lightweight and closer to a low-density wearable setting.

The key technical specifications of the acquisition system are:

- **Number of channels**: 64 (two grids of 32 electrodes each)
- **Sampling frequency**: 2000 Hz (2 kHz)

- **Analog-to-digital resolution**: 16 bits
- **Input range**: ±18.75 mV peak-to-peak
- **Hardware filtering**: First-order high-pass filter with cutoff frequency at 10.5 Hz (applied internally to remove DC component)

The reference electrode was placed on the ulnar styloid process to establish a stable reference potential. Signal acquisition was synchronized across all channels through a hardware voltage trigger, ensuring temporal alignment with other recording systems used during data collection.

5.2.1 Signal Quality and Preprocessing

To ensure data quality and reliability, the dataset authors applied digital filtering to the raw EMG signals prior to public release [7]. Specifically, a 4th-order Butterworth bandpass filter with cutoff frequencies at 10 Hz and 500 Hz was applied to isolate the physiologically relevant EMG frequency content. Subsequently, a 6th-order notch filter centered at 50 Hz was used to attenuate power-line interference originating from the European electrical grid. These preprocessing steps ensure that the downloaded data contains minimal low-frequency drift, high-frequency noise, and electrical artifacts, providing a clean signal baseline for subsequent feature extraction and classification analyses.

5.3 Muscles and Electrode Placement

The two electrode grids were positioned on the right forearm to capture activity from the primary extrinsic muscles responsible for wrist and finger movements. The electrodes were placed proximally, starting approximately 5 cm below the olecranon, and arranged to cover the entire circumference of the forearm.

Grid 1 (Channels 1–32): Flexor Compartment

The first electrode grid was positioned over the anterior (volar) surface of the forearm to record activity from the flexor muscles. This grid captures signals from the following anatomical structures (see Chapter 2 for detailed muscle anatomy):

- **Flexor carpi radialis (FCR)**: flexes and abducts the wrist
- **Palmaris longus (PL)**: assists in wrist flexion
- **Flexor carpi ulnaris (FCU)**: flexes and adducts the wrist
- **Flexor digitorum superficialis (FDS)**: flexes the proximal interphalangeal joints of fingers 2–5

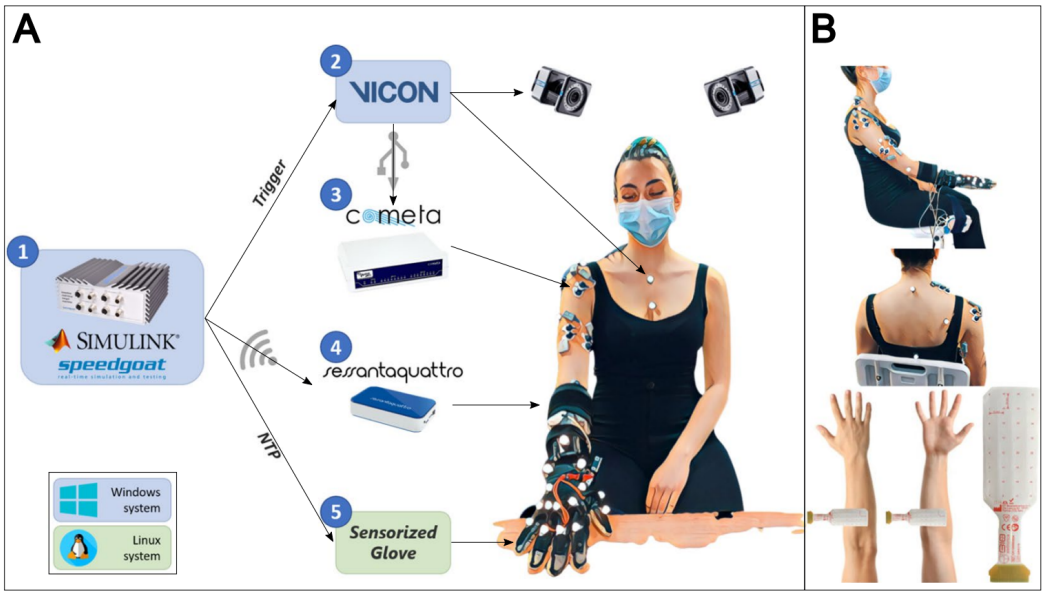

Figure 5.1: Experimental setup. (A) 1) Real Time target machine Speedgoat running the Simulink model. 2) Vicon motion capture system. 3) Cometa system bipolar EMG. 4) Sessantaquattro high-density EMG. 5) Sensorized glove from Bielefeld University. (B) Placement of the acquisition sensors: 23 reflective markers for the Vicon system, 10 bipolar EMG electrodes, 2 grids for HD-sEMG acquisition, and sensorized glove[7].

- **Flexor digitorum profundus (FDP)**: flexes the distal interphalangeal joints of fingers 2–5
- **Pronator teres (PT)**: pronates the forearm
- **Pronator quadratus (PQ)**: deep pronator of the forearm

These muscles are primarily recruited during hand closing, wrist flexion, and forearm pronation gestures.

Grid 2 (Channels 33–64): Extensor Compartment

The second electrode grid was positioned over the posterior (dorsal) surface of the forearm to record activity from the extensor muscles. This grid captures signals from:

- **Brachioradialis (BR)**: flexes the elbow and assists in pronation/supination
- **Extensor carpi radialis longus (ECRL)**: extends and abducts the wrist
- **Extensor carpi radialis brevis (ECRB)**: extends and abducts the wrist

- **Extensor carpi ulnaris (ECU)**: extends and adducts the wrist
- **Extensor digitorum (ED)**: extends fingers 2–5 and assists wrist extension
- **Extensor indicis (EI)**: extends the index finger
- **Extensor digiti minimi (EDM)**: extends the little finger
- **Supinator (SUP)**: supinates the forearm

These muscles are primarily recruited during hand opening, wrist extension, and forearm supination gestures.

5.4 Feature Extraction

The extraction of meaningful features from EMG signals is critical for accurate classification and myoelectric control. Raw EMG(unprocessed) signals contain high-dimensional, noisy information that requires transformation into a compact, discriminative feature space. This section describes the mathematical formulation and physiological interpretation of the features extracted from both raw and deconvolved EMG signals in this study.

Features were extracted using a sliding window approach with 250 ms epochs and 150 ms overlap (60% overlap, corresponding to 100 ms increment), consistent with real-time prosthetic control requirements. This configuration provides smoother predictions and improved classification accuracy compared to disjoint segmentation, while maintaining response times below the 300 ms threshold required for real-time prosthetic control [56]. Based on the deconvolution framework applied to sEMG, features were categorized into three groups: (1) features common to both raw and deconvolved signals, (2) features exclusive to raw EMG signals, and (3) features exclusive to deconvolved EMG signals.

Mathematical Notation In the formulations presented below, $x[n]$ denotes a generic discrete-time signal sequence of length N and applies to both the raw interference signal and the deconvolved spike train.

To distinguish the signal source in the subsequent analysis (Chapter 7), features extracted from the deconvolved signal will be denoted by the subscript $_d$ (e.g., RMS_d), while features extracted from the raw EMG will retain the standard acronym (e.g., RMS).

5.4.1 Features Common to Both Raw and Deconvolved Signals

These features capture fundamental signal properties related to amplitude, energy and complexity in the time domain. Each feature was computed on both the raw signal and the deconvolved signal.

Root Mean Square (RMS)

The Root Mean Square (RMS) represents the signal power and is modeled as an amplitude modulated Gaussian random process, relating to constant force and non-fatiguing muscle contractions. RMS is widely used in prosthetic applications due to its computational efficiency and robust representation of muscle contraction intensity.

$$\text{RMS} = \sqrt{\frac{1}{N}\sum_{n=1}^{N} x[n]^2} \tag{5.1}$$

Average Rectified Value (ARV)

Average Rectified Value (ARV) represents the mean absolute amplitude of the signal and is commonly used as an onset detection index and provides a robust estimate of overall muscle activity:

$$\text{ARV} = \frac{1}{N}\sum_{n=1}^{N} |x[n]| \tag{5.2}$$

Simple Square Integral (SSI)

Simple Square Integral (SSI) provides an energy index that is proportional to the total power within the analysis window and is sensitive to changes in muscle activation levels. SSI expresses the energy of the signal as:

$$\text{SSI} = \sum_{n=1}^{N} x[n]^2 \tag{5.3}$$

Simple Absolute Integral (SAI)

The Simple Absolute Integral (SAI) is the most basic amplitude-based feature, computed as the sum of absolute values of the signal within a window. This feature provides a direct measure of the cumulative magnitude of muscle activity, proportional to the total electrical output regardless of the signal's temporal structure.

SAI is defined as:

$$\text{SAI} = \sum_{n=1}^{N} \Big| x[n] \Big| \tag{5.4}$$

Root Integral Features (SRI, FRI)

The Root Integral (RI) features are designed to increase sensitivity to low-amplitude activity by applying sub-linear power transformations to the signal magnitude before summation. Exponents smaller than 1 compress large values and relatively expand small ones, enhancing the contribution of low-level activations within the window.

The Square Root Integral (SRI) is defined as:

$$\text{RI} = \sum_{n=1}^{N} \Big| x[n] \Big|^{1/2} \tag{5.5}$$

A Fourth Root Integral (FRI) variant further emphasizes very small amplitudes:

$$\text{FRI} = \sum_{n=1}^{N} \Big| x[n] \Big|^{1/4} \tag{5.6}$$

Both features provide smoothed measures of signal intensity that are more sensitive to small-amplitude EMG activity than standard energy-based measures (e.g., SSI), making them suitable for discriminating between low levels of muscle activation.

Willison Amplitude (WAMP)

Willison Amplitude (WAMP) measures the number of times the difference between adjacent samples exceeds a predefined threshold and is related to Motor Unit (MU) firing rates and muscle contraction force.

$$\text{WAMP} = \sum_{n=1}^{N-1} f(|x[n+1] - x[n]|) \tag{5.7}$$

where $f(x) = 1$ if $x \geq T_{wamp}$, and $f(x) = 0$ otherwise.

V-Order Feature (V3)

The V3 feature is a non-linear detector that estimates muscle contraction force through a cube-root transformation:

$$\text{V3} = \left(\frac{1}{N} \sum_{n=1}^{N} |x[n]|^3 \right)^{1/3} \tag{5.8}$$

This feature emphasizes large amplitude events while maintaining sensitivity to signal magnitude variations.

Logarithm Detector (LD)

The Logarithm Detector (LD) provides an estimate of muscle contraction force through a logarithmic transformation:

$$\mathrm{LD} = \exp\left(\frac{1}{N}\sum_{n=1}^{N}\log(|x[n]|)\right) \tag{5.9}$$

LD is a non-linear detector similar to V3, but uses logarithmic scaling which can be less sensitive to extreme outliers while still capturing muscle activation levels.

Waveform Length (WL)

Waveform Length (WL) measures the cumulative length of the waveform, indicating signal complexity:

$$\mathrm{WL} = \sum_{n=1}^{N-1}|x[n+1] - x[n]| \tag{5.10}$$

WL provides information about waveform amplitude, frequency, and duration, making it a comprehensive feature for movement classification.

Higher-Order Moments (M2, M4)

Higher-order moments quantify the power of the first derivative of the signal, capturing rapid changes in amplitude at different scales:

$$\mathrm{M2} = \sum_{n=1}^{N-1}(x[n+1] - x[n])^2 \tag{5.11}$$

$$\mathrm{M4} = \sum_{n=1}^{N-1}(x[n+1] - x[n])^4 \tag{5.12}$$

M2 is sensitive to overall signal transitions and sudden changes. M4, by raising the derivative to the fourth power, emphasizes large amplitude changes while suppressing smaller fluctuations, making it more selective for detecting sharp spikes or transient events. Together, these features provide complementary information about signal dynamics at different amplitude scales.

Maximum Fractal Length (MFL)

Maximum Fractal Length (MFL) estimates signal complexity through a logarithmic transformation of the cumulative squared derivative:

$$\text{MFL} = \log\left(\sqrt{\sum_{n=1}^{N-1}(x[n+1]-x[n])^2}\right) \tag{5.13}$$

This feature estimates signal complexity through a logarithmic transformation of the cumulative squared derivative. It provides a measure of waveform irregularity, with higher values indicating more complex, jagged signal patterns.

Integrated Absolute of Second Derivative (IASD)

Integrated Absolute of Second Derivative (IASD) captures the cumulative absolute changes in the signal's first derivative:

$$\text{IASD} = \sum_{n=1}^{N-2}|x'[n+1]-x'[n]| \tag{5.14}$$

where $x'[n] = x[n+1] - x[n]$. IASD acts as a high-frequency filter, reducing low-frequency noise while emphasizing rapid signal transitions.

Slope Sign Change (SSC)

Slope Sign Change (SSC) counts the number of times the slope of the signal changes sign, detecting peaks and valleys:

$$\text{SSC} = \sum_{n=2}^{N-1} f\left[(x[n]-x[n-1])\times(x[n]-x[n+1])\right] \tag{5.15}$$

where $f(x) = 1$ if $x \geq T_{SSC}$, and $f(x) = 0$ otherwise. SSC provides frequency information while remaining computationally efficient. In this study, the threshold was set to 0 to capture all directional changes.

Variance (VAR)

Variance (Var) represents the power of the signal around its mean and is calculated as:

$$\text{VAR} = \frac{1}{N-1}\sum_{n=1}^{N}(x[n]-\mu)^2 \tag{5.16}$$

where μ is the mean of the signal. Var measures the dispersion of the signal around its mean.

Kurtosis (KRT)

Kurtosis (Krt) measures the "tailedness" of the signal distribution, indicating the concentration of samples around the mean:

$$\mathrm{KRT} = \frac{\frac{1}{N}\sum_{n=1}^{N}(x[n]-\mu)^4}{\left(\frac{1}{N}\sum_{n=1}^{N}(x[n]-\mu)^2\right)^2} \tag{5.17}$$

where μ is the mean of the signal. Higher kurtosis values indicate samples concentrated near the mean with few extreme values, suggesting lower muscle activation. Lower kurtosis values indicate a flatter distribution with more extreme values, suggesting higher muscle activation or burst activity.

5.4.2 Features Exclusive to Raw EMG Signals

These features leverage characteristics specific to the raw sEMG interference pattern, extracting frequency information via zero-crossings and quantifying temporal amplitude trends across the analysis window.

Zero Crossing (ZC)

Zero Crossing (ZC) counts the number of times the signal crosses the zero amplitude level, providing frequency information in the time domain:

$$\mathrm{ZC} = \sum_{n=1}^{N-1} \mathbf{1}(x[n]\,x[n+1] < 0 \ \wedge \ |x[n]-x[n+1]| \geq T_{zc}) \tag{5.18}$$

A threshold $T_{zc} = 0.05$ is applied to both samples surrounding each potential zero crossing; this guards against spurious counts caused by baseline noise and quantization artifacts when the signal oscillates near zero with very small amplitude.

Mean Absolute Value Slope (MAVS)

Mean Absolute Value Slope (MAVS) captures temporal dynamics by measuring the difference in average amplitude between consecutive segments:

$$\mathrm{MAVS} = \frac{1}{\lfloor N/2 \rfloor}\left(\sum_{n=1}^{\lfloor N/2 \rfloor}|x[n]| - \sum_{n=\lfloor N/2 \rfloor+1}^{N}|x[n]|\right) \tag{5.19}$$

This feature quantifies changes in muscle activation over time within the analysis window.

5.4.3 Features Exclusive to Deconvolved EMG Signals

The following features are specific to the deconvolved representation and capture neural drive characteristics, synchronization patterns, and temporal firing statistics that are not accessible from raw EMG.

Spike Rate (SR)

Spike Rate (SR) estimates the average firing rate of active MUs:

$$\mathrm{SR} = \frac{\mathrm{spike\_count}}{\mathrm{duration}} \tag{5.20}$$

where spikes are detected using `findpeaks` with a minimum peak height threshold ($\mu+3\sigma$), where μ and σ are the mean and standard deviation of the deconvolved signal.

Coefficient of Variation (CoV)

Coefficient of Variation (CoV) quantifies the regularity of MU firing:

$$\mathrm{CoV} = \frac{\sigma_{\mathrm{ISI}}}{\mu_{\mathrm{ISI}}} \tag{5.21}$$

where μ_{ISI} and σ_{ISI} are the mean and standard deviation of Inter-Spike Interval (ISI). Higher CoV values indicate greater firing irregularity, which may reflect changes in neural control strategies.

Synchronization (Sync)

Synchronization (Sync) measures the average correlation between the deconvolved signal of a given channel and all other channels:

$$\mathrm{Sync} = \frac{1}{M-1}\sum_{i\neq j}\rho(x_i, x_j) \tag{5.22}$$

where ρ is the Pearson correlation coefficient, M is the number of channels, and i, j index the channels. This feature captures common drive and synchronization across motor unit populations.

Burstiness (BRST)

Burstiness (Brst) quantifies the proportion of short inter-spike intervals, indicating burst firing patterns:

$$\text{BRST} = \frac{\sum \mathbf{1}(\text{ISI} < T_B}{\text{total number of ISIs}} \tag{5.23}$$

where the burst threshold T_B was set to 50 ms.

Co-Activation (CoA)

Co-Activation (CoA) measures the temporal coincidence of spikes across channels within a narrow time window:

$$\text{CoA} = \frac{1}{N_s \times (M-1)} \sum_{i,j} \mathbf{1}(|\Delta t_{ij}| \leq w) \tag{5.24}$$

where N_s is the total spike count in the reference channel, M is the number of channels, Δt_{ij} is the time difference between spike i in the reference channel and spike j in a different channel, $w = 20$ ms is the coincidence window, and the indicator function $\mathbf{1}(\cdot)$ counts spike pairs within the window. CoA ranges from 0 (no synchrony) to 1 (perfect synchrony).

Latency (Lat)

Latency (Lat) estimates the temporal offset between the first spike in a reference channel (channel 1) and the first spike in the current channel:

$$\text{Lat} = \frac{\text{spike\_time}_{\text{current}} - \text{spike\_time}_{\text{reference}}}{\text{sampling frequency}} \tag{5.25}$$

If no spikes are detected, the latency is set to the epoch duration, indicating absence of synchronized activation.

Skewness(SKW)

Skewness (Skw) measures the asymmetry of the signal distribution:

$$\text{SKW} = \frac{\frac{1}{N}\sum_{n=1}^{N}(x[n]-\mu)^3}{\left(\frac{1}{N}\sum_{n=1}^{N}(x[n]-\mu)^2\right)^{3/2}} \tag{5.26}$$

Positive skewness indicates a distribution with a longer tail on the positive side, while negative skewness indicates the opposite. In deconvolved signals, skewness can reflect the asymmetry of MU firing patterns.

Gini Index (GI)

The Gini Index (GI) is a measure of statistical dispersion that quantifies the inequality in the distribution of deconvolved MU firing amplitudes. It ranges from 0 (perfect equality, all amplitudes identical) to 1 (perfect inequality, one spike dominates). This feature captures the concentration of MU activity across detected spikes.

The Gini Index is computed as:

$$\mathrm{GI} = \frac{2\sum_{i=1}^{n} i \cdot x_{\mathrm{s}}[i]}{n\sum_{i=1}^{n} x_{\mathrm{s}}[i]} - \frac{n+1}{n} \tag{5.27}$$

where x_{s} is the deconvolved spike amplitudes sorted in ascending order and n is the total number of detected spikes. Lower GI values indicate more uniform spike amplitudes (homogeneous recruitment), while higher values indicate sparse, dominant spikes (heterogeneous recruitment).

Interspike Interval Entropy (ISIE)

The Interspike Interval Entropy (ISIE)) quantifies the regularity and predictability of motor unit firing patterns by measuring the Shannon entropy of the ISI distribution. This feature captures the temporal organization of MU recruitment: high entropy indicates irregular, variable firing intervals, while low entropy indicates highly regular, stereotyped patterns.

ISIE is defined as:

$$\mathrm{ISIE} = -\sum_{j=1}^{k} p_j \log_2(p_j) \tag{5.28}$$

where p_j is the probability of the j-th ISI bin computed from a 10-bin histogram of ISIs and k is the number of non-empty bins. ISI values are normalized by the sampling frequency f_c. When fewer than three spikes are detected, ISIE is assigned a value of 0.

5.4.4 Summary

A total of 28 distinct features were extracted: 17 features common to both signal types (computed on both raw and deconvolved EMG, yielding 34 feature values), 2 features exclusive to raw EMG, and 9 features exclusive to deconvolved EMG. Computing the same feature on both raw and deconvolved signals enables a direct assessment of how the deconvolution algorithm influences feature values and, consequently, classification performance.

Not all features are meaningful for both signal representations due to their different statistical and morphological properties. For instance, the deconvolved

signal is strictly positive, rendering the ZC feature uninformative. Conversely, several features were defined only on the deconvolved signal because its sparse representation of weighted cumulative firing patterns allows detailed characterization of spike timing, amplitude, and temporal organization, which cannot be reliably accessed from the raw EMG.

Chapter 6

Feature Selection for Gesture Classification

This chapter presents the methodology adopted for feature selection and classifier training. Given the high dimensionality of the feature space extracted from both raw and deconvolved signals (see Chapter 5), a systematic approach to feature selection is essential to reduce computational cost, minimize overfitting, and improve classifier generalization performance.

The chapter is organized as follows: Section 6.1 introduces the main feature selection methodologies and motivates the hybrid filter-wrapper strategy adopted in this work. Section 6.2 describes the filter-based stage, including data preprocessing, Pearson correlation analysis, one-way Analysis of Variance (ANOVA) ranking, and the construction of the three feature sets. Section 6.3 then presents the wrapper-based optimization using a genetic algorithm, including chromosome encoding, fitness function design, validation within the optimization process, genetic operators, and the selected hyperparameters.

6.1 Feature Selection Overview

The extraction of features from both raw and deconvolved EMG signals, as described in Chapter 5, results in a high-dimensional feature space. Without appropriate dimensionality reduction, this abundance of features poses significant challenges for classifier training and real-time implementation of myoelectric control systems[83]. This section presents the theoretical framework and methodological rationale for the hybrid filter-wrapper feature selection approach adopted in this work.

6.1.1 Motivation and Challenges

The Curse of Dimensionality

High dimensional feature spaces constitute a problematic in EMG pattern recognition. As the number of features increases, redundant and noisy information can erase previously well defined classification boundaries and cause overfitting of the training data [84]. This phenomenon, known as the curse of dimensionality, manifests in two primary ways: first, classifiers require exponentially more training data to maintain generalization performance as dimensionality increases due to the sparsity of the data in this higher dimension feature space[85]; second, the risk of overfitting escalates as models may learn spurious correlations specific to the training set rather than capturing the underlying discriminative structure [86].

In the context of EMG based hand gesture recognition, the number of extracted features grows rapidly due to the combination of multiple recording channels, diverse feature types, and the inclusion of both raw and deconvolved signal representations. This can easily result in feature vectors with hundreds of dimensions when all combinations are considered.

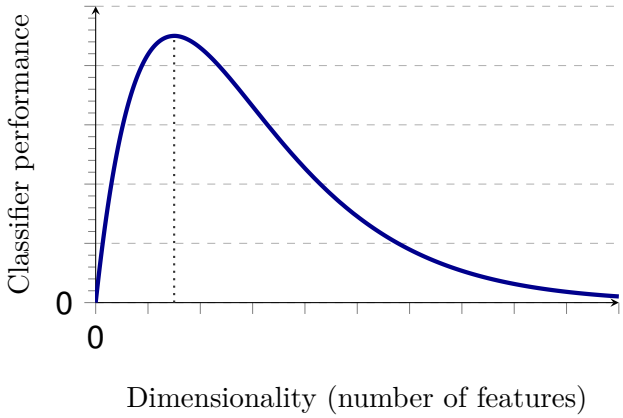

Figure 6.1: Classifier performance as a function of feature dimensionality, illustrating the curse of dimensionality and the existence of an optimal number of features [87]

Computational Constraints for Real-Time Control

The computational cost of feature extraction varies considerably across different feature types. Simple time-domain features such as RMS or Var require minimal computation, whereas spectral analysis and deconvolution based features demand significantly greater processing power. When deployed on embedded or wearable systems with limited battery capacity, reducing the number of computed features

directly impacts system autonomy and user experience [6].

Furthermore, classifier complexity scales with the dimensionality of the input feature vector. High-dimensional inputs increase both training time and inference latency, potentially violating real-time constraints. Feature selection thus serves a dual purpose: improving classification generalization while simultaneously reducing computational burden [88].

Redundancy and Correlation Among EMG Features

EMG feature selection is characterized by the presence of substantial redundancy among commonly used features. Many time-domain features are mathematically interrelated or capture similar signal properties [89]. For example, RMS, ARV, SAI, and V3 all reflect signal energy or amplitude, differing primarily in normalization or weighting schemes [90]. Similarly, ZC, SSC, and WAMP all encode frequency information through threshold based counts of signal transitions [91].

When features extracted from raw and deconvolved signals are combined, additional redundancy emerges. If deconvolution successfully isolates motor unit firing patterns while preserving discriminative information present in the raw signal, then certain features computed from both representations may be highly correlated.

6.1.2 Feature Selection Methodologies

Feature selection methods can be taxonomically organized into three broad categories—filter methods, wrapper methods, and embedded methods [92]. Although embedded methods are efficient, they are model-specific and incompatible with the standard formulation of Linear Discriminant Analysis (LDA) utilized in this work. Consequently, this section focuses on filter and wrapper methods. These approaches provide a versatile framework for reducing dimensionality that is independent of the specific constraints of the classification algorithm.

Filter Methods

Filter methods assess the relevance of features independently of any specific classifier, relying instead on intrinsic statistical properties of the data [92]. Common filter criteria include correlation coefficients (e.g., Pearson or Spearman), mutual information, Fisher's discriminant ratio, and ANOVA F-statistics [45, 83]. These methods rank features according to their individual discriminative power or their redundancy with respect to other features [45].

The primary advantage of filter methods is computational efficiency: feature evaluation does not require iterative training of classification models, making them suitable for high-dimensional datasets and exploratory analysis [92, 93].

Additionally, filter methods are model-agnostic, producing rankings that can be reused across different classifiers without recomputation [92].

However, filter methods suffer from a fundamental limitation: they ignore feature interactions and classifier specific characteristics. A feature that appears individually uninformative may become highly discriminative when combined with other features, a synergy that univariate filter criteria cannot detect [92]. Moreover, different classifiers may benefit from different feature subsets depending on their decision boundaries (e.g., linear vs. nonlinear) [93]. Consequently, filter-selected features may not optimize classification performance for a specific model.

Wrapper Methods

Wrapper methods address the limitations of filter approaches by directly incorporating classifier performance into the feature selection process. A wrapper method evaluates candidate feature subsets by training the target classifier on each subset and measuring its predictive accuracy (or another performance metric) through cross-validation or a held-out validation set. Feature subsets are iteratively added, removed, or recombined using search strategies such as sequential forward selection (SFS), sequential backward elimination, or metaheuristic algorithms like genetic algorithms [45].

The key advantage of wrapper methods is that they are inherently classifier-specific and capable of capturing feature interactions and complementarities [86]. By optimizing directly for classification performance, wrappers can identify feature subsets that yield superior accuracy compared to filter-selected features [93].

Nonetheless, wrapper methods are computationally expensive. Each candidate subset requires full model training and validation, resulting in high computational cost when the feature space is large or when cross-validation is employed. Additionally, wrapper methods are susceptible to overfitting, particularly when the number of candidate features is large relative to the number of training samples [92].

Given the complementary strengths and weaknesses of filter and wrapper methods, a hybrid two-stage approach was adopted in this work to balance computational efficiency, feature interpretability, and classifier-specific optimization.

6.2 Filter-Based Feature Selection

6.2.1 Data Preprocessing for Feature Selection

Feature selection operates on preprocessed dataset designed to facilitate the identification of discriminative features while maintaining ecological validity across the subject population. This preprocessing stage encompasses three steps: within-subject normalization, cross-subject concatenation, and movement-only epoch selection.

Each step addresses specific challenges inherent to multi-subject, multi-channel sEMG recordings.

Within-subject z-score normalization was applied independently to each subject's feature distribution to remove inter-subject variability in EMG amplitude. The normalization is expressed as:

$$z_{ij} = \frac{x_{ij} - \mu_i}{\sigma_i}$$

where x_{ij} denotes the feature value for the j-th epoch from subject i, μ_i is the mean feature value across all epochs from subject i, and σ_i is the corresponding standard deviation. This standardization ensures that the feature selection procedure identifies features with consistent discriminative capacity across subjects, rather than features that correlate with subject-specific amplitude variations.

The biological justification for this approach arises from the substantial inter subject variability in sEMG signals due to differences in subcutaneous tissue thickness (3–7 mm), muscle fiber alignment, electrode-skin impedance, and anatomical variations in forearm composition.

Rest epochs (labeled as gesture class 0) were systematically excluded from the feature selection analysis. Although discriminating rest from active movement is clinically important, the rest class was substantially overrepresented in the dataset relative to individual gesture classes. Including rest epochs during feature selection would bias the process toward amplitude-dependent features that excel at exploiting this class imbalance, rather than features that robustly discriminate among active gestures where the classification challenge actually resides. By excluding rest, feature selection prioritizes the more demanding problem of distinguishing among hand gestures based on their distinct muscle activation patterns.

6.2.2 Pearson Correlation Analysis

Following data pre-processing, Pearson correlation analysis was employed to identify and group redundant features. For each pair of features (f_a, f_b) in the pre-processed dataset, the Pearson correlation coefficient was computed across all epochs as:

$$r(f_a, f_b) = \frac{\sum_{i=1}^{N}(f_{a,i} - \overline{f_a})(f_{b,i} - \overline{f_b})}{\sqrt{\sum_{i=1}^{N}(f_{a,i} - \overline{f_a})^2}\sqrt{\sum_{i=1}^{N}(f_{b,i} - \overline{f_b})^2}}$$

where N is the total number of epochs, $f_{a,i}$ and $f_{b,i}$ are the values of features a and b at epoch i, and $\overline{f_a}$ and $\overline{f_b}$ are the respective means. Features were grouped into correlation clusters using hierarchical agglomerative clustering. A dissimilarity metric was defined as $d = 1 - |r|$, and an average-linkage method was employed to construct the dendrogram which was cut at a distance threshold of 0.05. This meant that within each group approximately 90% of the variance in one feature

can be predicted from the other ($r^2 = 0.90$). At this threshold, retaining both features provides minimal additional discriminative information while substantially increasing model complexity and the risk of overfitting.

6.2.3 One-Way Analysis of Variance (ANOVA)

One-way ANOVA was applied to rank features within each correlation group according to their ability to discriminate among the six gesture classes. This approach enabled the selection of the single most discriminative feature from each redundant group.

ANOVA Formulation

For each feature f in a correlation group, a one-way ANOVA was conducted to compare how the mean feature values differ across the six gesture classes. The test statistic is the F-value:

$$F = \frac{\text{MS}_{\text{between}}}{\text{MS}_{\text{within}}} = \frac{\sum_{k=1}^{K} n_k(\overline{f}_k - \overline{f})^2/(K-1)}{\sum_{k=1}^{K}\sum_{i=1}^{n_k}(f_{ki} - \overline{f}_k)^2/(N-K)}$$

where:

- $K = 6$ is the number of gesture classes,
- n_k is the number of epochs in gesture class k,
- $N = \sum_{k=1}^{K} n_k$ is the total number of epochs,
- $\overline{f}_k$ is the mean feature value for class k,
- $\overline{f}$ is the grand mean across all epochs,
- f_{ki} is the feature value for the i-th epoch in class k.

The F-statistic quantifies the ratio of between-group (gesture-related) variance to within-group (gesture-independent or noise-related) variance. Higher F-values indicate stronger discrimination among gesture classes, as the variance attributable to differences in gesture identity becomes large relative to the residual within-gesture variance.

Within each correlation group, the feature with the highest F-statistic was selected as the most discriminative representative of that group. This ranking approach has several practical advantages: the F-statistic directly reflects the ability of a feature to separate gesture classes, is computationally efficient, and provides a single, interpretable ranking criterion that generalizes across features with different units and distributions.

Assumptions and Limitations

While ANOVA provides an effective framework for feature ranking, several critical assumptions require discussion, particularly in the context of EMG feature selection.

The standard ANOVA formulation assumes that observations (epochs) are independent. However, in this dataset, temporal windows are extracted from continuous EMG signals with substantial overlap (as specified in Section 5). This overlapping window structure introduces temporal autocorrelation: successive epochs contain highly correlated feature values due to the slow temporal dynamics of hand gesture execution and EMG signal continuity. This violation of independence can lead to deflated estimates of the within-group variance, thereby inflating F-statistics and potentially overestimating the discriminative power of features. ANOVA also assumes that feature distributions within each gesture class are approximately normally distributed and that the variance of each feature is approximately equal across gesture classes.

However, in the present context, ANOVA is used not for hypothesis testing (where false positives are problematic) but for feature ranking to select the best representative from each correlation group. The relative ranking of F-statistics is the primary quantity of interest, and strict statistical significance testing of individual features is not performed.

6.2.4 Construction of Feature Sets

Three distinct feature sets were constructed to isolate and evaluate the contribution of raw EMG features, deconvolved EMG features, and their combination. Each set underwent the filter-based selection procedure (Pearson correlation grouping and ANOVA F-score ranking) independently, followed by wrapper-based optimization using a genetic algorithm (detailed in Section 6.3). This sequential filter-wrapper approach enabled systematic evaluation of whether deconvolution-based features enhance classification performance.

Feature Set 1: Raw Signal Features (`FS-Raw`) This set comprises 19 features: the 17 common features computed from raw EMG plus the 2 raw-exclusive features. This set provides a baseline for assessing classification performance achievable with conventional EMG feature extraction, without the preprocessing advantages of single-channel deconvolution.

Feature Set 2: Deconvolved Signal Features (`FS-Dec`) This set comprises 26 features: the 17 common features computed from deconvolved EMG plus the 9 deconvolution-exclusive features. Comparison of `FS-Dec` classification performance

to `FS-Raw` provides direct evidence of whether deconvolution preprocessing enhances feature quality and discriminative power for gesture recognition.

Feature Set 3: Combined Features (`FS-C`) This set comprises all 45 features (17 common + 2 raw-exclusive + 9 deconv-exclusive). This set tests the hypothesis that raw and deconvolved features capture non-redundant information about muscle activation patterns, and that their combination enables superior classification performance.

6.3 Wrapper-Based Feature Selection Using Genetic Algorithm

6.3.1 Genetic Algorithm Fundamentals

Genetic Algorithms (GAs) represent a class of population-based metaheuristic optimization techniques inspired by the principles of natural selection and biological evolution.

In the context of feature selection for EMG based gesture recognition, GAs offer several advantages over exhaustive search methods. The number of possible feature subsets grows exponentially with the number of candidate features: for N_f features, there exist 2^{N_f} possible subsets. Exhaustive evaluation becomes computationally prohibitive even for moderately sized feature sets. GAs circumvent this limitation by efficiently exploring the solution space through stochastic search guided by fitness-based selection [94].

The fundamental principle underlying GAs is the iterative refinement of a population of candidate solutions through the application of biologically-inspired operators: selection, crossover, and mutation. Solutions with fitness values that better satisfy the optimization objective are more likely to survive and contribute their genetic material to subsequent generations. This process mimics natural selection, progressively improving solution quality across generations [95].

A key strength of GAs in wrapper-based feature selection lies in their ability to handle multi-objective optimization problems. Unlike filter methods that evaluate features independently based on univariate statistical criteria, wrapper methods using GAs can simultaneously optimize multiple conflicting objectives, such as classification accuracy, model complexity, and computational efficiency [96]. This capability is particularly relevant for real-time myoelectric control applications, where both high classification performance and low computational latency are critical requirements.

6.3.2 Chromosome Encoding

In binary genetic algorithms applied to feature selection, each candidate solution is represented as a chromosome—a binary vector of length N_f, where N_f corresponds to the total number of features available after filter-based dimensionality reduction (Section 6.2). This encoding scheme directly maps the feature selection problem to a binary optimization framework, where each gene (bit) in the chromosome indicates the inclusion or exclusion of a specific feature.

Formally, a chromosome $\mathbf{c}$ is defined as:

$$\mathbf{c} = [c_1, c_2, \ldots, c_{N_f}] \quad \text{where} \quad c_i \in \{0, 1\}, \quad i = 1, 2, \ldots, N_f \tag{6.1}$$

where:

$$\begin{aligned} c_i = 1 &\quad \text{indicates that feature } i \text{ is selected for classification} \\ c_i = 0 &\quad \text{indicates that feature } i \text{ is excluded from classification} \end{aligned}$$

The initial population of chromosomes is generated through uniform random sampling from the binary solution space. Each gene c_i in each chromosome is independently initialized with equal probability, ensuring that the initial population exhibits sufficient diversity to explore the feature subset space effectively [94].

To guarantee practical feasibility and prevent the selection of excessively large or minimal feature subsets, an admission criterion was imposed on candidate solutions. This constraint requires that the number of selected features $N_s = \sum_{i=1}^{N_f} c_i$ must fall within a valid range defined by classifier-specific and feature-set-specific bounds

These bounds were determined empirically during preliminary experiments to balance classification performance with model complexity. Chromosomes that violated the admission criterion were not evaluated and were excluded from the GA search process.

Table 6.1 summarizes the admission criterion bounds for each classifier and feature set combination.

Table 6.1: Admission criterion bounds for feature subset cardinality.

Feature Set	**Bounds**
FS-Raw	5–11
FS-Dec	5–11
FS-Combined	5–15

6.3.3 Fitness Function

Multi-Objective Optimization

The fitness function serves as the objective measure that guides the evolutionary search process toward optimal feature subsets. For real-time myoelectric prosthetic control, optimizing classification accuracy alone is insufficient; practical deployment requires simultaneous consideration of system usability (Abstention Rate (AR)) and computational efficiency (inference latency). Consequently, the fitness function is formulated as a weighted combination of three components:

1. **Classification performance**: quantified by macro-averaged F1-score across all gesture classes and subjects
2. **AR penalty**: penalizes excessive classifier abstention that reduces system responsiveness.
3. **Computational cost penalty**: penalizes high feature extraction and inference time that violates real-time constraints

Macro F1-Score

Classification performance is quantified using the macro-averaged F1-score, computed independently for each subject and subsequently averaged across the entire subject population. The F1-score represents the harmonic mean of precision and recall, providing a balanced measure that accounts for both false positives and false negatives.

For each gesture class $c \in \{1, 2, \ldots, C\}$, precision and recall are defined as:

$$\text{Precision}_c = \frac{\text{TP}_c}{\text{TP}_c + \text{FP}_c} \tag{6.2}$$

$$\text{Recall}_c = \frac{\text{TP}_c}{\text{TP}_c + \text{FN}_c} \tag{6.3}$$

where TP_c, FP_c, and FN_c denote true positives, false positives, and false negatives for class c, respectively.

The class-specific F1-score is then computed as:

$$F1_c = \frac{2 \cdot \text{Precision}_c \cdot \text{Recall}_c}{\text{Precision}_c + \text{Recall}_c} \tag{6.4}$$

The macro-averaged F1-score treats all gesture classes equally, regardless of their representation in the dataset:

$$F1_{\text{macro}} = \frac{1}{C} \sum_{c=1}^{C} F1_c \tag{6.5}$$

This averaging strategy is particularly appropriate for hand gesture recognition, where all gestures are functionally important for prosthetic control and should be recognized with comparable accuracy. In contrast, micro-averaging or weighted F1-scores would disproportionately emphasize majority classes, potentially degrading recognition performance for less frequent but equally critical gestures.

To aggregate performance across the subject population, the macro F1-score for each subject k is computed independently, and the final fitness component $F1_{\text{total}}(\mathbf{c})$ represents the average across all N_{sub} subjects:

$$F1_{\text{total}}(\mathbf{c}) = \frac{1}{N_{\text{sub}}} \sum_{k=1}^{N_{\text{sub}}} F1_{\text{macro}}^{(k)}(\mathbf{c}) \tag{6.6}$$

where $F1_{\text{macro}}^{(k)}(\mathbf{c})$ is the macro F1-score for subject k evaluated using feature subset $\mathbf{c}$.

Abstention Rate Penalty

For a given classification decision on epoch t, abstention occurs when the maximum posterior probability falls below a predefined confidence threshold τ_{conf}:

$$\mathrm{A}(t) = \begin{cases} 1 & \text{if } \max_c P(c \mid \mathbf{x}_t) < \tau_{\text{conf}} \\ 0 & \text{otherwise} \end{cases} \tag{6.7}$$

where $P(c \mid \mathbf{x}_t)$ represents the posterior probability of class c given feature vector $\mathbf{x}_t$ for epoch t.

The abstention rate $\mathrm{AR}(\mathbf{c})$ for a feature subset $\mathbf{c}$ is computed as the proportion of epochs for which the classifier abstains, averaged across all subjects:

$$\mathrm{AR}(\mathbf{c}) = \frac{1}{N_{\text{sub}}} \sum_{k=1}^{N_{\text{sub}}} \frac{1}{N_{\text{e}}^{(k)}} \sum_{t=1}^{N_{\text{e}}^{(k)}} \mathrm{A}^{(k)}(t) \tag{6.8}$$

where $N_{\text{e}}^{(k)}$ is the total number of classification epochs for subject k.

To balance accuracy and responsiveness, a quadratic penalty is applied to abstention rates exceeding a tolerable threshold $\mathrm{AR}_{\text{free}}$:

$$\text{Penalty}_{\text{AR}}(\mathbf{c}) = \lambda_{\text{AR}} \cdot \max\left(0, \mathrm{AR}(\mathbf{c}) - \mathrm{AR}_{\text{free}}\right)^2 \tag{6.9}$$

where λ_{AR} is the penalty weight coefficient and $\mathrm{AR}_{\text{free}}$ represents the acceptable abstention rate below which no penalty is incurred.

The quadratic formulation ensures that moderate abstention rates near $\mathrm{AR}_{\text{free}}$ incur minimal penalty, while excessive abstention is heavily penalized. This penalty structure reflects the nonlinear relationship between abstention rate and user experience: small increases in abstention near zero have minimal impact, whereas high abstention rates fundamentally compromise system utility.

Computational Cost Penalty

The computational cost penalty quantifies the total time required for feature extraction and classifier inference, penalizing feature subsets that violate real-time latency requirements.

For feature subset $\mathbf{c}$, the prediction time $t_{\text{pred}}(\mathbf{c})$ comprises two components:

$$t_{\text{pred}}(\mathbf{c}) = t_{\text{e}}(\mathbf{c}) + t_{\text{i}}(\mathbf{c}) \tag{6.10}$$

where:

- $t_{\text{e}}(\mathbf{c})$ is the time required to extract the selected features from preprocessed EMG signals
- $t_{\text{i}}(\mathbf{c})$ is the time required for classifier inference given the extracted feature vector

Computational cost was measured empirically using the target deployment platform (specified hardware and software environment). For each feature subset $\mathbf{c}$, prediction time was measured across all classification epochs for all subjects, and the average value was computed:

$$t_{\text{pred}}(\mathbf{c}) = \frac{1}{N_{\text{sub}}} \sum_{k=1}^{N_{\text{sub}}} \frac{1}{N_{\text{e}}^{(k)}} \sum_{t=1}^{N_{\text{e}}^{(k)}} \left[t_{\text{e}}^{(k)}(t, \mathbf{c}) + t_{\text{i}}^{(k)}(t, \mathbf{c}) \right] \tag{6.11}$$

Measurements were conducted on a standard desktop computing platform representative of embedded prosthetic control systems: Intel Core i7-9700K CPU (3.6 GHz base clock), 16 GB RAM, running MATLAB r2024b. All timing measurements included only feature extraction and inference computations. Each measurement was repeated 100 times and averaged to reduce variability from operating system scheduling and background processes.

Fitness Function Formulation

$$\text{Fitness}(\mathbf{c}) = 1 - F1_{\text{total}}(\mathbf{c}) + \lambda_{\text{AR}} \cdot \max\left(0, \text{AR}(\mathbf{c}) - \text{AR}_{\text{free}}\right)^2 + \lambda_t \cdot t_{\text{pred}}(\mathbf{c}) \tag{6.12}$$

Parameter values adopted for the multi-objective optimization are summarized in Table 6.2. The value of the confidence threshold τ_{conf} was optimized for each feature subset $\mathbf{c}$ such that it minimized the fitness value.

The fitness function is evaluated for each chromosome in the population by training the classifier (LDA or NLR) on the selected feature subset and computing classification performance, AR, and prediction time through cross-validation on the training data (detailed in Section 6.3.4).

Table 6.2: Fitness function parameter values.

Parameter	Symbol	Value
Abstention tolerance	$\mathrm{AR}_{\mathrm{free}}$	0.10
Abstention penalty weight	λ_{AR}	0.5
Computational penalty weight	λ_t	5×10^{-4} ms^{-1}

6.3.4 Validation Strategy Within GA

The validation strategy employed during GA based feature selection must balance two conflicting requirements: computational efficiency (to enable evaluation of hundreds of candidate feature subsets across multiple generations) and unbiased performance estimation (to avoid selecting feature subsets that exploit idiosyncrasies of the training data) [97].

In this work, fitness evaluation for each chromosome is conducted using a subject specific 5-fold cross-validation protocol. For each subject $k \in \{1, 2, \ldots, N_{\mathrm{sub}}\}$, the available training data is partitioned into 5 disjoint folds. The classifier is trained on 4 folds and validated on the remaining fold, with this process repeated 5 times such that each fold serves once as the validation set. The macro F1-score, abstention rate, and prediction time are computed for each fold and averaged across the 5 folds to obtain subject-specific performance metrics $F1^{(k)}_{\mathrm{macro}}(\mathbf{c})$, $\mathrm{AR}^{(k)}(\mathbf{c})$, and $t^{(k)}_{\mathrm{pred}}(\mathbf{c})$.

The subject-specific metrics are subsequently aggregated across all subjects by averaging:

$$F1_{\mathrm{total}}(\mathbf{c}) = \frac{1}{N_{\mathrm{sub}}} \sum_{k=1}^{N_{\mathrm{sub}}} F1^{(k)}_{\mathrm{macro}}(\mathbf{c}) \tag{6.13}$$

$$\mathrm{AR}(\mathbf{c}) = \frac{1}{N_{\mathrm{sub}}} \sum_{k=1}^{N_{\mathrm{sub}}} \mathrm{AR}^{(k)}(\mathbf{c}) \tag{6.14}$$

$$t_{\mathrm{pred}}(\mathbf{c}) = \frac{1}{N_{\mathrm{sub}}} \sum_{k=1}^{N_{\mathrm{sub}}} t^{(k)}_{\mathrm{pred}}(\mathbf{c}) \tag{6.15}$$

This aggregation strategy ensures that the fitness function evaluates the *average* generalization performance across the subject population, preventing the GA from optimizing feature subsets that perform well for only a subset of subjects.

The choice of 5-fold cross-validation represents a trade-off between bias, variance, and computational cost. Smaller values of k (e.g., 3-fold) reduce computational expense but increase the variance of performance estimates and risk bias due to smaller training set sizes. Larger values (e.g., 10-fold or leave-one-out cross-validation) provide lower-variance estimates but substantially increase computational cost. Studies on cross-validation for model selection indicate that 5-fold cross-validation

achieves a favorable balance for datasets of moderate size (hundreds to thousands of samples per subject) [98].

6.3.5 Genetic Operators

Selection

Roulette wheel selection was employed in this work. In roulette wheel selection, each individual in the population is assigned a selection probability proportional to its fitness value. Parents are selected stochastically according to these probabilities, such that higher-fitness individuals are more likely to be selected, but lower-fitness individuals still have a non-zero probability of reproduction [99].

Formally, the selection probability for individual i is defined as:

$$P_i = \frac{\text{Fitness}(\mathbf{c}_i)}{\sum_{j=1}^{N_{\text{pop}}} \text{Fitness}(\mathbf{c}_j)} \tag{6.16}$$

where N_{pop} is the population size. Parents are selected by sampling from the population according to these probabilities.

The stochastic nature of roulette wheel selection provides a balance between exploitation (preferring better fitness solutions) and exploration (maintaining worse fitness solutions that may carry useful genetic material for future generations) [95].

This property makes roulette wheel selection well-suited for multi-objective fitness functions that combine classification accuracy, AR penalty, and computational cost, as in the present study.

Crossover

Crossover is the primary operator for generating new candidate solutions by combining genetic material from two parent chromosomes. Crossover exploits the building block hypothesis: high-quality solutions often contain subsequences (building blocks) that contribute positively to fitness, and combining building blocks from different parents can produce offspring with superior fitness [99].

Two-point crossover was chosen for this study because it provides a balanced approach between genetic recombination and preservation of parent structure. Unlike uniform crossover, which treats each gene independently, two-point crossover preserves larger building blocks of parent solutions, which can be beneficial when features exhibit dependencies or when subsequences encode meaningful patterns [100]. In feature selection for EMG classification, two-point crossover may help preserve feature combinations that interact synergistically in the classifier.

The crossover probability p_c determines the proportion of parents in each generation that undergo crossover; parents not selected for crossover are copied

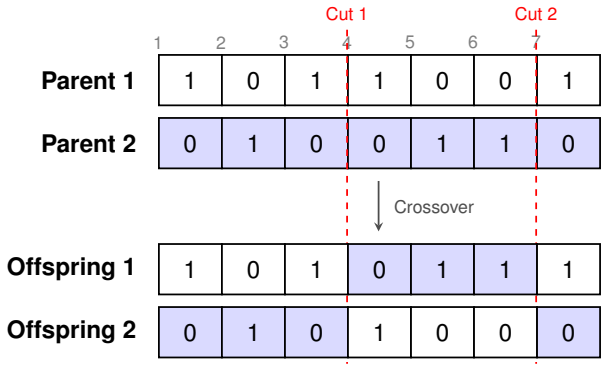

Figure 6.2: Two-point crossover: chromosome segments between two randomly selected cut points (dashed lines) are exchanged between parents to create offspring, resulting in new candidate feature subsets.

directly to the offspring population without recombination. The specific crossover probabilities (p_c) adopted in this work vary based on the computational constraints of the chosen classifier [99]. These values are detailed in Section 6.3.6

Mutation

Mutation introduces random variation into offspring chromosomes, preventing premature convergence and enabling the GA to escape local optima by exploring regions of the solution space not reachable through crossover alone. Mutation has been implemented as bit-flip mutation: each gene c_i in an offspring chromosome is independently flipped (inverted) with a small probability p_m.

Formally, the bit-flip mutation operator is defined as:

$$c_i^{(\text{mut})} = \begin{cases} 1 - c_i & \text{with probability } p_m \\ c_i & \text{with probability } 1 - p_m \end{cases} \quad \text{for } i = 1, 2, \ldots, N_f \tag{6.17}$$

where $c_i^{(\text{mut})}$ is the mutated value of gene i.

The mutation probability p_m is a critical parameter that controls the balance between exploration and exploitation. High mutation rates ($p_m \gg 1/N_f$) induce excessive randomness, disrupting useful building blocks and slowing convergence. Low mutation rates ($p_m \ll 1/N_f$) provide insufficient exploration, increasing the risk of premature convergence. The specific mutation probabilities (p_m) adopted in this study are tailored to the population size and are detailed in Section 6.3.6.

6.3.6 GA Hyperparameters

The performance and convergence behavior of GAs are strongly influenced by hyperparameter settings. These must be carefully selected to balance exploration, exploitation, and computational efficiency:

Population size ($N_{\mathbf{pop}}$): Determines the number of candidate solutions per generation, balancing search space exploration with computational cost.

Parent fraction (f_p): Defines the proportion of elite individuals selected to survive and reproduce, maintaining selective pressure toward high-performing regions.

Number of generations ($T_{\mathbf{max}}$): Sets the absolute upper limit on evolutionary iterations.

Stopping criterion (S_c): Terminates the GA early to save computation if the best fitness value fails to improve for S_c consecutive generations.

Number of independent runs: Addresses the stochastic nature of the GA. To assess stability, the optimization was executed 10 independent times for the LDA classifier. In contrast, due to its significantly higher computational cost, the GA was executed only once for the Non-Linear Logistic Regression (NLR) model.

Table 6.3 summarizes the GA hyperparameter values adopted in this study. Because the NLR classifier carries a significantly higher computational cost than LDA, the hyperparameters were adapted accordingly. For NLR, the population size (N_{pop}) and early stopping patience (S_c) were reduced to maintain feasible execution times. To compensate for the reduced exploration capability of a smaller population, the mutation probability (p_m) was doubled to preserve genetic diversity. Furthermore, a stricter parent fraction ($f_p = 0.4$) was applied alongside a higher crossover probability ($p_c = 0.9$) to ensure rigorous recombination among only the most elite solutions.

Table 6.3: Genetic algorithm hyperparameters for each feature set and classifier.

Symbol	**LDA**			**NLR**		
	Raw	**Dec**	**Comb**	**Raw**	**Dec**	**Comb**
N_{pop}	150	175	250	130	140	200
f_p	0.7	0.7	0.7	0.4	0.4	0.4
$T_{\max}$	100	100	100	100	100	100
p_c	0.8	0.8	0.8	0.9	0.9	0.9
p_m	0.05	0.05	0.05	0.1	0.1	0.1
S_c	50	50	50	30	30	30

Chapter 7

Results: Feature Selection and Classification Performance

This chapter presents the results of the feature selection process and the evaluation of the proposed gesture-classification pipeline. The analysis is organized into four main parts: (1) results of the filter-based feature selection stage, including correlation grouping and Analysis of Variance (ANOVA)-based ranking; (2) outcomes of the Genetic Algorithm (GA)-based wrapper optimization; (3) offline classification performance on held-out test data; and (4) online evaluation of the real-time pipeline, including the experimental setup, streaming implementation, and performance obtained during task-based testing.

7.1 Filter Stage Feature Selection Results

This section presents the results of the filter stage of feature selection for the three feature sets. For each set, Pearson correlation analysis was first used to identify redundant feature groups, and one-way ANOVA was then applied within each group to select the most discriminative representative feature.

7.1.1 Raw Signal Feature Set (FS-Raw)

The correlation analysis for the raw Electromyography (EMG) feature set (Figure 7.1, $N_f = 19$) reveals a feature space characterized by extensive collinearity. The heatmap is dominated by a large block of high positive correlation centered on magnitude-related descriptors and extending to event-count and complexity

descriptors such as Zero Crossing (ZC), Willison Amplitude (WAMP), and Maximum Fractal Length (MFL). This widespread redundancy indicates that, for the raw interference pattern, a broad range of mathematical formulations converge on encoding closely related aspects of signal magnitude and waveform activity.

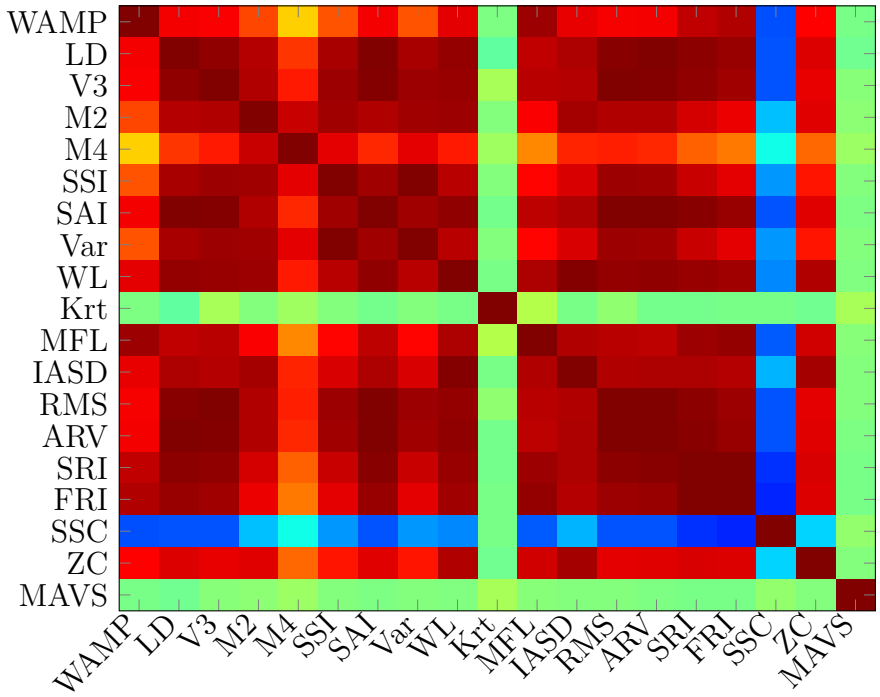

Figure 7.1: Pearson correlation matrix for the raw EMG feature set (`FS-Raw`, $N_f = 19$). The pronounced red blocks indicate high redundancy among magnitude-related descriptors.

Despite this high redundancy, specific features exhibit distinct independence. The Slope Sign Change (SSC) feature appears as a pronounced blue band across the matrix, indicating a negative yet low correlation with the dominant magnitude-related block. Furthermore, Kurtosis (Krt) and Mean Absolute Value Slope (MAVS) manifest as distinct green bands, evidencing their orthogonality and their capacity to capture statistical properties of the raw signal distribution that are not correlated to magnitude.

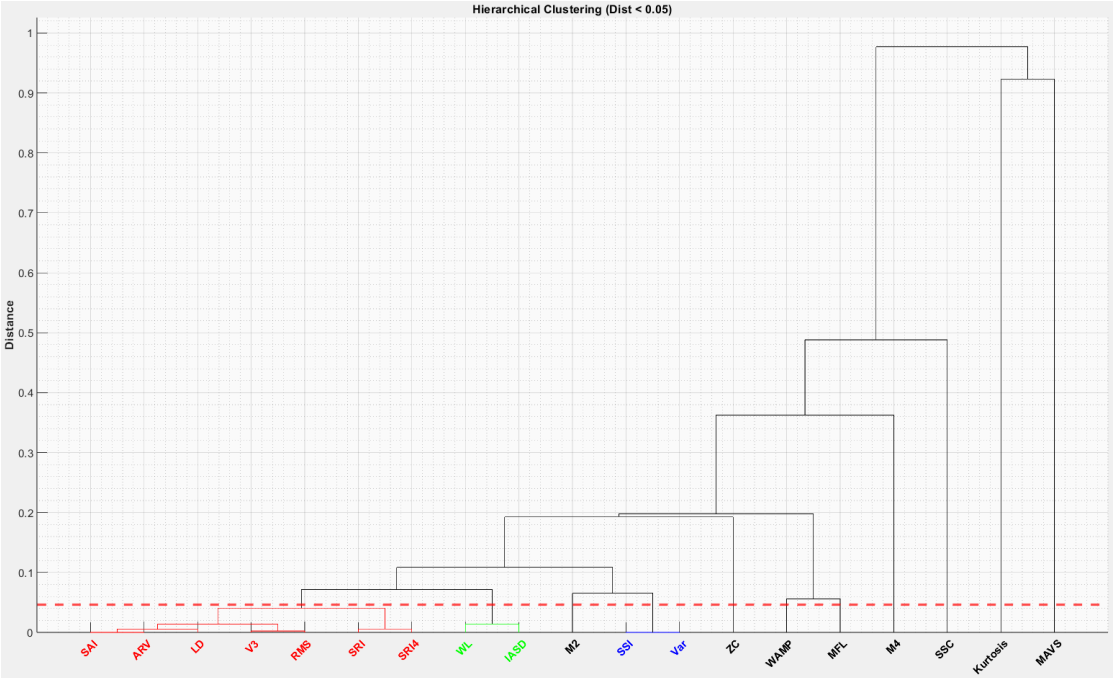

Figure 7.2: Hierarchical clustering dendrogram for Raw Signal Features (`FS-Raw`) at distance threshold 0.05 (horizontal line). Branches and features are color-coded to indicate redundant feature groups.

The hierarchical clustering dendrogram (Figure 7.2) partitions the feature space at the 0.05 distance threshold, identifying three primary structural behaviors:

- **Primary Magnitude-Related Cluster:** The most significant structure is the tight merging of seven descriptors related to signal magnitude (Simple Absolute Integral (SAI), Average Rectified Value (ARV), Logarithm Detector (LD), Moment of 3 order (V3), Root Mean Square (RMS), Square Root Integral (SRI), Fourth Root Integral (FRI)) at extremely low linkage distances. This confirms that linear, root, and logarithmic scalings of the raw signal magnitude yield nearly identical information content.
- **Intermediate Groupings:** Two additional compact groups are observed. Waveform Length (WL) and Integrated Absolute of Second Derivative (IASD) merge early, reflecting their shared dependence on combined amplitude and frequency changes. Likewise, Simple Square Integral (SSI) and Variance (Var) merge at a small distance due to their very similar mathematical formulation.
- **Distinct Singletons:** A subset of features remains isolated at the 0.05 threshold. These features join the hierarchy at significantly higher linkage

distances (approaching 1.0 for MAVS and Krt), highlighting their role as independent descriptors that provide unique variance complementary to the redundant magnitude-related measures.

Consequently, the filter-based selection process heavily consolidates the dominant magnitude-related cluster into a single representative, while retaining the independent higher-order statistics and specific event-based descriptors that resist redundancy.

As summarized in Table 7.1, within the magnitude-related group, FRI exhibited the highest F-statistic ($F > 3200$), surpassing widely used features like RMS and ARV. This result implies that, for raw Surface Electromyography (sEMG), sub-linear scaling, which emphasizes low-level muscle activation while compressing high-amplitude peaks, provides a more robust separation of hand gestures than linear descriptors. Additionally, between WL and IASD, WL was selected, confirming its efficacy in capturing the combined amplitude and frequency characteristics of the raw interference pattern.

Table 7.1: ANOVA-based representative feature selection for FS-Raw.

Candidate features	**Selected feature**
FRI, SRI, SAI, ARV, RMS, LD, V3	FRI
WL, IASD	WL
SSI, Var	SSI

7.1.2 Deconvolved Signal Feature Set (FS-Dec)

The correlation heatmap for the deconvolved feature set (Figure 7.3, $N_f = 26$) reveals a pronounced structural dichotomy in the feature space. The matrix is dominated by a large block of high correlation (indicated by deep red regions) comprising magnitude-related descriptors (e.g., RMS_d, SAI_d, SSI_d, Var_d). This redundancy confirms that, despite their differing mathematical formulations, these metrics largely encode the same underlying information regarding the instantaneous magnitude of the Cumulative Weighted Firing (CWF) pattern.

In sharp contrast, features encoding temporal firing statistics and sparsity-related properties, specifically Synchronization (Sync), Latency (Lat), Interspike Interval Entropy (ISIE), and Gini Index (GI), exhibit a predominantly low correlation profile (cyan/blue regions) against the main magnitude-related block. This lack of linear dependence ($|r| < 0.3$) suggests that these features extract a subspace

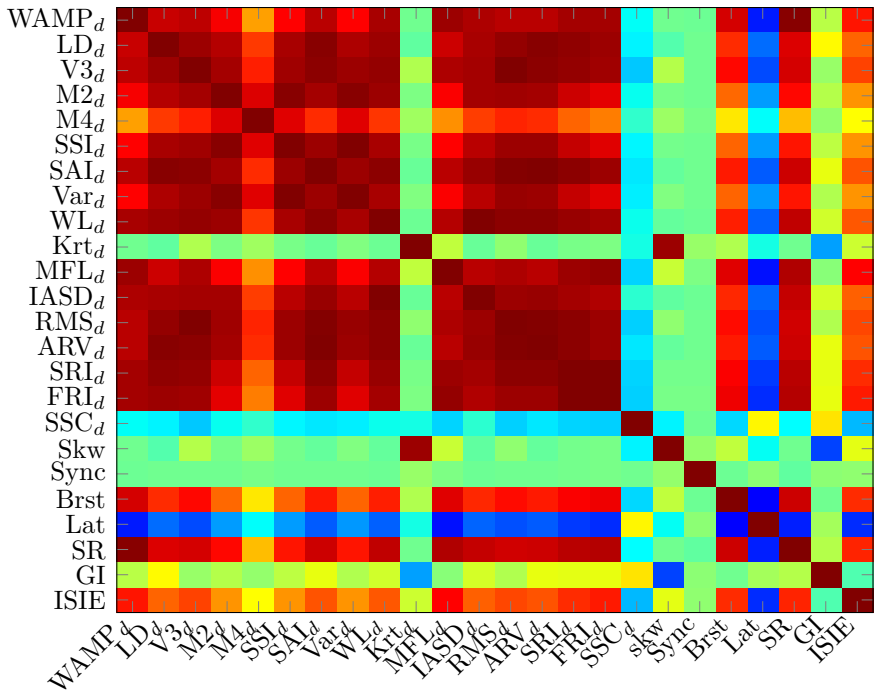

Figure 7.3: Pearson correlation matrix for the deconvolved signal feature set (`FS-Dec`, $N = 26$ features). The heatmap highlights the low correlation (blue regions) of deconvolution features (lower right) compared to the redundant magnitude-related descriptors.

of information orthogonal to the signal magnitude, capturing the temporal organization and sparsity of Motor Unit (MU) firings rather than simple activation intensity. The hierarchical clustering analysis (Figure 7.4) quantitatively confirms

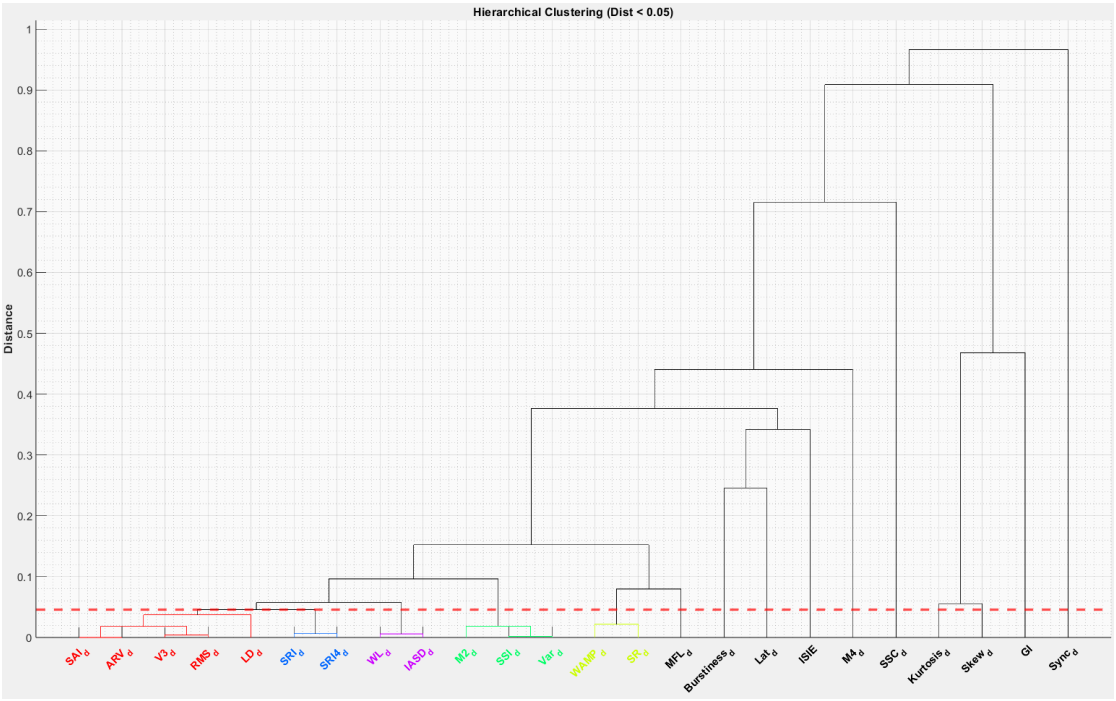

Figure 7.4: Hierarchical clustering dendrogram for FS-Dec at distance threshold 0.05 (horizontal line). Branches and features are color-coded to indicate redundant feature groups.

these observations, partitioning the feature space into 17 distinct correlation groups at the 0.05 distance threshold ($|r| \geq 0.95$). The topology of the dendrogram highlights two main structural patterns:

- **Highly Redundant Compact Groups:** The magnitude, power, and complexity descriptors collapse into tight clusters at near-zero distance, corroborating the strong collinearity observed in the heatmap.
- **Independence of Deconvolution-Derived Features:** The remaining 12 groups consist entirely of isolated singletons that resist merging until significantly higher linkage distances.

This structure indicates that while deconvolution still produces several highly redundant amplitude-related descriptors, it effectively expands the useful feature space by providing numerous independent variables that characterize the temporal structure of the MU firing pattern.

The ANOVA rankings within these correlated groups revealed a preference for higher-order amplitude descriptors over more conventional amplitude-related features. As summarized in Table 7.2, among $V3_d$, RMS_d, ARV_d, SAI_d, and LD_d, $V3_d$ achieved the highest F-score. This suggests that emphasizing high-amplitude components of the estimated CWF through cubic scaling provides better discrimination of gesture intent than conventional amplitude descriptors.

A similar trend emerged for the pair FRI_d and SRI_d. Although FRI_d achieved a marginally higher F-statistic, SRI_d was selected as the representative feature because the difference in discriminative power was negligible and did not justify the higher computational cost associated with the fourth-root operation.

Table 7.2: ANOVA-based representative feature selection for FS-Dec.

Candidate features	**Selected feature**
$V3_d$, RMS_d, ARV_d, SAI_d, LD_d	$V3_d$
FRI_d, SRI_d	SRI_d
WL_d, $IASD_d$	WL_d
Var_d, SSI_d, $M2_d$	Var_d
$WAMP_d$, Spike Rate (SR)	$WAMP_d$

7.1.3 Combined Feature Set (FS-Combined)

The correlation heatmap for the combined feature set (Figure 7.5, $N_f = 45$) presents a pronounced checkerboard pattern that visually quantifies the interaction between the raw and deconvolved signal representations, highlighting two distinct forms of redundancy:

1. **Internal Redundancy:** As previously seen, large diagonal blocks within each domain show that magnitude-related descriptors remain highly collinear.

2. **Inter-Domain Redundancy:** Significant off-diagonal correlations appear between homologous features. For instance, the raw amplitude-related features (RMS, ARV, SAI) show near-perfect linear dependence with their deconvolved counterparts (RMS_d, ARV_d, SAI_d), indicating that the signal's magnitude information is largely conserved across the deconvolution process.

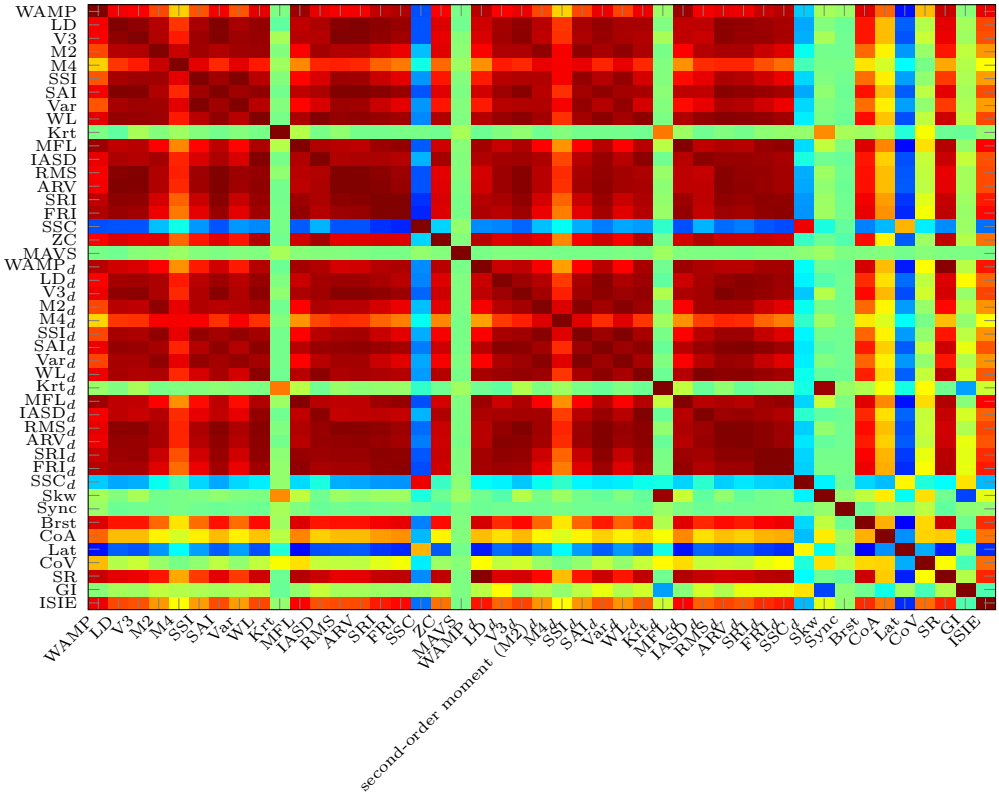

Figure 7.5: Feature redundancy heatmap for the combined feature set (`FS-C`, $N = 45$).

However, the matrix also reveals a vast region of low correlation corresponding to features computed from the CWF. Features such as SSC_d, Krt_d, and the specialized deconvolved metrics (Lat, Coefficient of Variation (CoV), Sync, ISIE) exhibit minimal correlation with the dominant magnitude-related blocks.

The hierarchical clustering dendrogram (Figure 7.6) partitions the 45-dimensional feature space into 25 distinct groups revealing two main structural patterns:

- **Inter-Domain Fusion:** The left side of the dendrogram displays mixed clusters where homologous features from both domains merge. The **Yellow Cluster** groups amplitude-related descriptors (SAI, ARV, V3, RMS) from both raw and deconvolved signals. Similarly, the **Cyan Cluster** and **Green Cluster** fuse the complexity (WL, IASD) and root-integral (SRI, FRI) features with their deconvolved counterparts.

- **Isolation of Deconvolution-derived Features:** The right side of the dendrogram consists of isolated singletons. Features specific to the deconvolved

domain—Burstiness (Brst), Lat, ISIE, Co-Activation (CoA), CoV, Sync—do not merge with any raw-signal features.

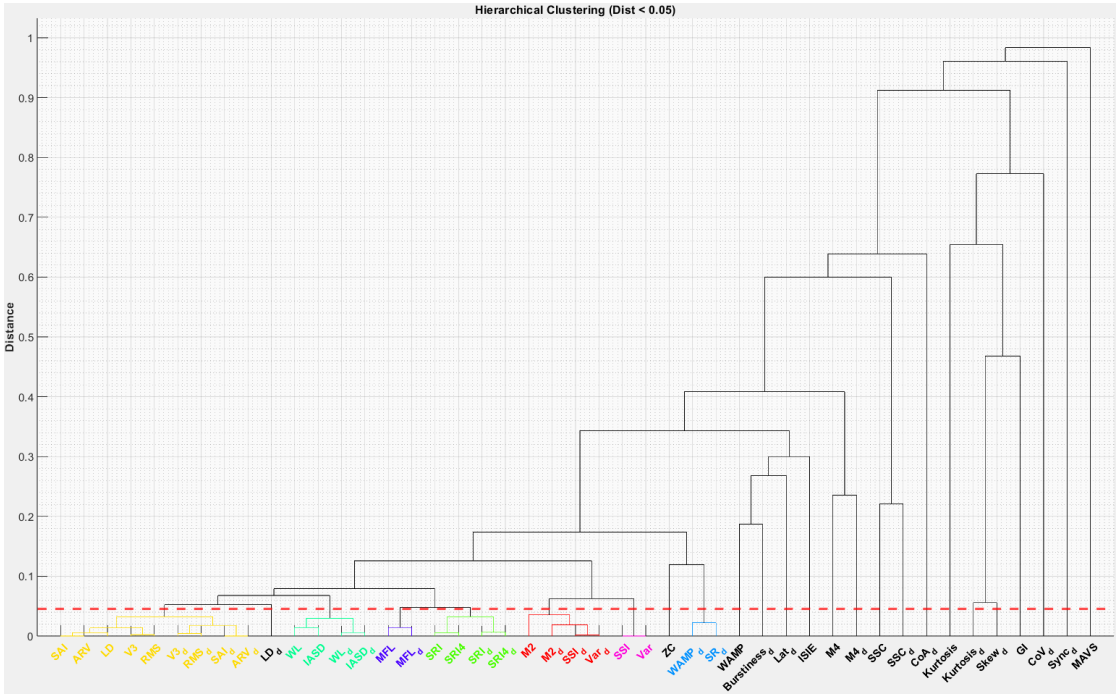

Figure 7.6: Hierarchical clustering dendrogram for FS-Combined at distance threshold 0.05 (horizontal line). Branches and features are color-coded to indicate redundant feature groups.

As summarized in Table 7.3, among the magnitude related features, the selected representative was extracted from the raw signal, suggesting that the deconvolution process, by sparsifying the signal, reduces the global magnitude information that might be more effectively captured by the raw EMG. Conversely, between MFL and MFL_d, MFL_d was selected, suggesting that the temporal structure of the estimated spike train provides a cleaner substrate for fractal analysis than the stochastic interference pattern of the raw EMG. In contrast, the pair WL and IASD with their deconvolved counterparts was represented by the raw WL, while the set Var_d, SSI_d, M2_d, and M2 retained Var_d, indicating that the relative advantage of raw and deconvolved features depends on the specific signal property being described.

Table 7.3: ANOVA-based representative feature selection for FS-Combined.

Candidate features	**Selected feature**
SAI, ARV, RMS, LD, V3, $V3_d$, RMS_d, ARV_d, SAI_d	ARV
WL, WL_d, IASD, $IASD_d$	WL
MFL, MFL_d	MFL_d
FRI, SRI, FRI_d, SRI_d	FRI
Var_d, SSI_d, $M2_d$, M2	Var_d

7.1.4 Filtered Feature Set Composition

Table 7.4: Filtered Feature Set Composition After Filter-Based Selection

Feature Set	**Reduction**	**Features Retained**
FS-Raw	42%	WAMP, M2, M4, SSI, MAVS WL, Krt, MFL, FRI, SSC, ZC
FS-Dec	34%	$WAMP_d$, $V3_d$, $M4_d$, Var_d, WL_d Krt_d, MFL_d, SRI_d, SSC_d, Skw Sync, Brst, CoA, Lat CoV, GI, ISIE
FS-Combined	44%	WAMP, M4, SSI, WL, Krt ARV, FRI, SSC, ZC, MAVS $WAMP_d$, LD_d, $M4_d$, Var_d, Krt_d MFL_d, SSC_d, Skw, Sync Brst, CoA, Lat, CoV, GI, ISIE

7.2 Genetic Algorithm Wrapper Optimization Results

7.2.1 Linear Discriminant Analysis Results

Raw Signal Features (`FS-Raw`)

The selection frequency within the elite solutions for `FS-Raw` is presented in Figure 7.7. The search for the optimal raw feature vector achieved absolute convergence. The GA identified a 7-dimensional subset—comprising WAMP, SSI, WL, Krt, MFL, FRI, and SSC—which appeared in 100% of the top-performing individuals. All other candidates, including standard metrics like MAVS and ZC, were systematically rejected.

This specific composition reflects a sophisticated balance between avoiding redundancy and maximizing information capture across different signal characteristics. The inclusion of FRI validates the filter-stage findings where sub-linear scaling showed high discriminative power. By selecting FRI alongside the power-based SSI, the algorithm effectively combines sensitivity to low-amplitude muscle activation with the detection of high-energy events, creating a dynamic range coverage that linear features like ARV cannot achieve in isolation.

Furthermore, the simultaneous selection of WL, MFL, and SSC highlights the necessity of complexity and frequency information for interpreting the raw interference pattern. Crucially, the algorithm explicitly favored WAMP over both ZC and fourth-order moment (M4). This indicates that for raw signals, estimating contraction intensity through threshold-based amplitude changes provides more robust separability than simple zero-crossing counts. The retention of Krt further enhances this vector by adding statistical information regarding the sparsity and "tailedness" of the signal distribution, a characteristic orthogonal to simple energy or frequency measures.

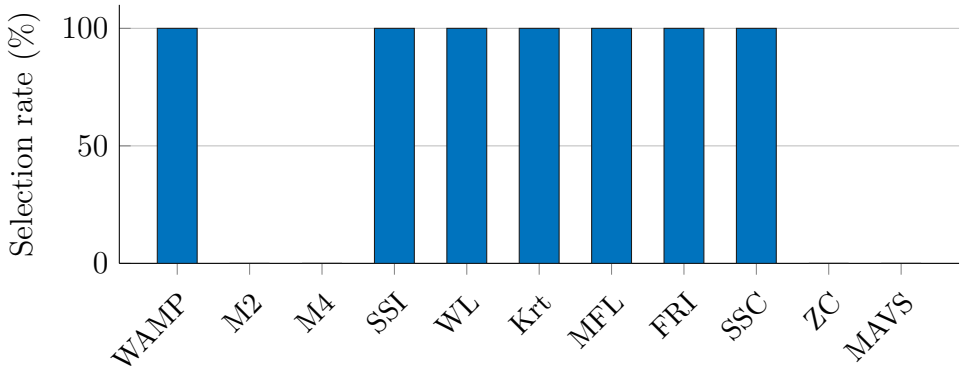

Figure 7.7: Selection frequency within the top 100 elite solutions for the raw feature set (`FS-Raw`, mean fitness: 0.0360). The optimization process achieved absolute convergence: a specific subset was selected in 100% of the best solutions.

Deconvolved Signal Features (`FS-Dec`)

The evolutionary trajectory of the feature selection process for the deconvolved dataset (Deconvolved Signal Features (`FS-Dec`)) is illustrated in Figure 7.8 and Figure 7.9. In the early optimization stage (Top 100 individuals, Figure 7.8), the GA explores a broad search space where features exclusive to the deconvolved signal are prioritized. Features such as MFL_d, SSC_d, and the temporal synchronization metrics (CoA, Sync) exhibit near-saturation ($> 90\%$ selection rate), indicating that temporal firing structure and pattern complexity provide the most robust initial discriminative power. Conversely, standard amplitude metrics like Var_d and WAMP_d show moderate variability (20–60%), suggesting they are interchangeable or less critical in the broader solution pool.

As the algorithm converges toward the best solutions (Figure 7.9), achieving a mean fitness of 0.0213 (an improvement over the `FS-Raw` baseline of 0.0360), the feature selection stabilizes into a consistent subset. In contrast to the absolute convergence observed in the raw feature set, the optimization process for the deconvolved set exhibited slight variability in the top tier. A decisive filtering effect occurs within the amplitude domain: while standard energy metrics like Var_d were present in the earlier generations, they are systematically eliminated in the final solutions. In their place, the algorithm converges on V3_d, selecting it in 100% of the elite individuals alongside WL_d.

This selection pattern suggests that for deconvolved signals, standard variance-based energy measures are suboptimal. The preference for V3_d—which utilizes a cubic transformation—indicates that emphasizing higher-amplitude events (such as the dominant spikes in the deconvolved train) provides better class separability than the mean energy estimation of RMS_d or Var_d. The final elite vector balances

different information domains: it combines non-linear amplitude features ($V3_d$) and signal complexity (WL_d) with discharge pattern statistics (CoA, GI), demonstrating that a diverse feature set effectively exploits the information recovered by the deconvolution process.

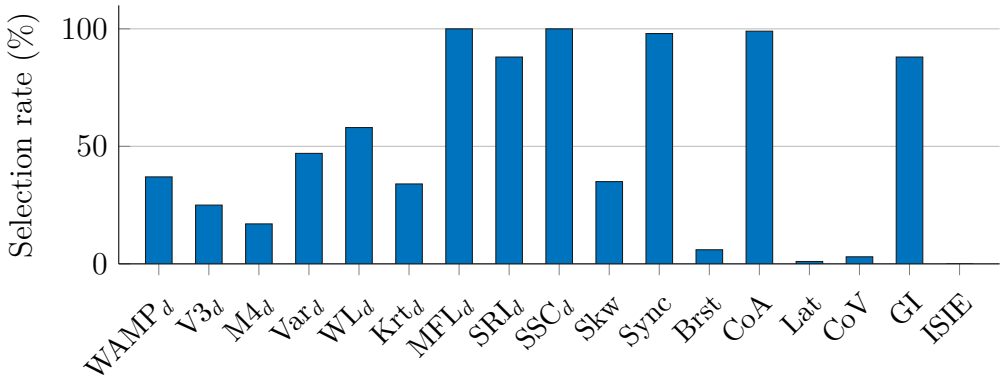

Figure 7.8: Selection frequency within the top 100 solutions of the `FS-Dec` population (mean fitness: 0.0221). While metrics like MFL_d, SSC_d, Sync, CoA, and GI exhibit near-saturation ($> 90\%$), features such as WL_d and $V3_d$ show moderate variability across the solution pool.

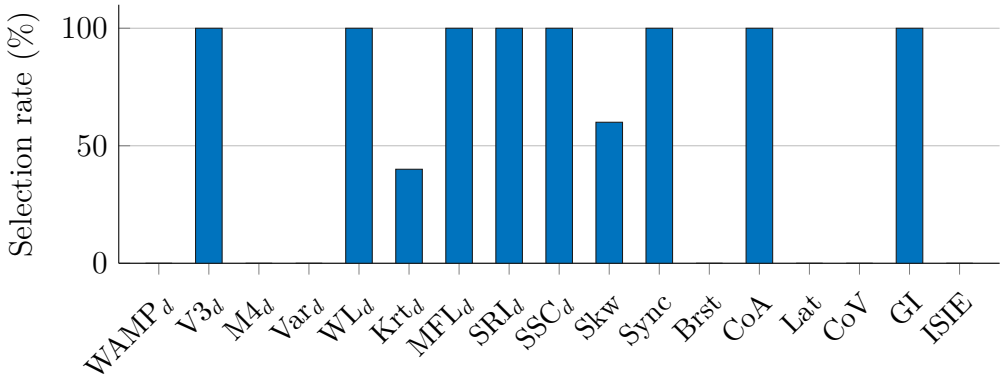

Figure 7.9: Selection frequency within the top 10 solutions (mean fitness: 0.0213). As the algorithm converges, the subset crystallizes: WL_d and $V3_d$ achieve 100% selection rates, while previously present features like Var_d and $WAMP_d$ are fully eliminated.

Combined Features (`FS-C`)

The feature selection results for the combined dataset (Combined Features (`FS-C`)) are illustrated in Figure 7.10 and Figure 7.11. This scenario allowed the GA to freely select from both raw and deconvolved domains, effectively testing the complementarity of the two signal representations. In the broad search phase (Top 100 solutions, Figure 7.10), the algorithm identified a hybrid core: WAMP and SSC_d achieved saturation, while temporal descriptors like GI, Sync, and CoA maintained high prevalence ($> 60\%$) across the population.

As the population converged to the elite solutions (Figure 7.11), achieving the lowest overall mean fitness of 0.0192, the optimal subset crystallized into multi-domain vector. A clear domain specialization emerged: the algorithm preferred raw features for energy and intensity estimation (universally selecting WAMP and heavily favoring WL) while selecting deconvolved features for complexity and synchronization (SSC_d, CoA).

The algorithm's preference for raw amplitude features can be attributed to the fundamental difference in signal structure: the inherent sparsity of the deconvolved signal does not favor the use of traditional energy estimators. Within the short analysis windows employed, the dense interference pattern of the raw signal might provide a more stable and consistent measure of overall muscle activation. Conversely, structural characteristics—such as slope changes and co-activation patterns—are more discriminative when extracted from the sparse CWF pattern.

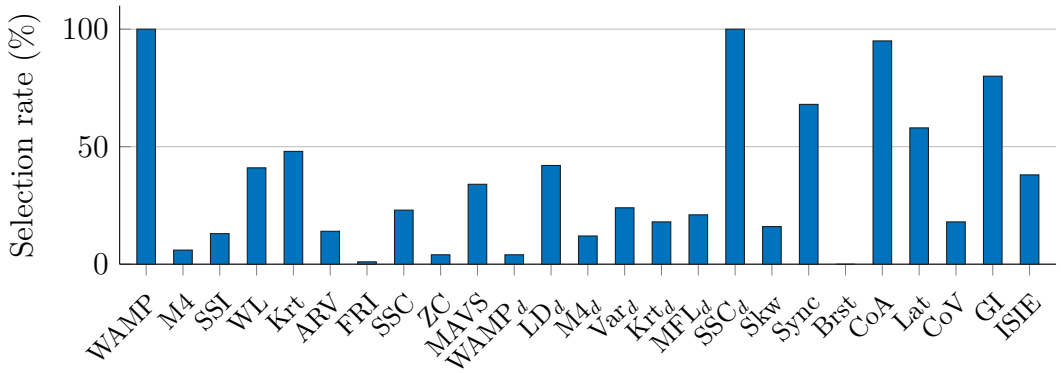

Figure 7.10: Feature selection frequency within the top 100 solutions of the combined feature set (`FS-C`, mean fitness: 0.0210).

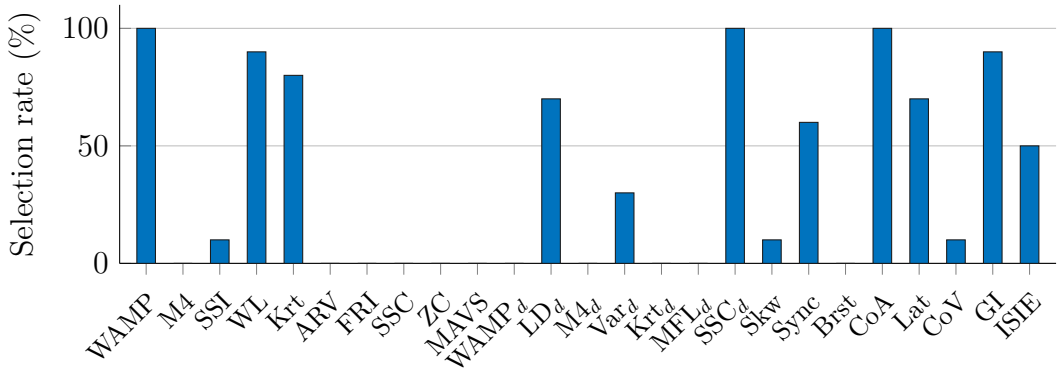

Figure 7.11: Selection frequency refined to the top 10 elite solutions for the combined feature set (`FS-C`, mean fitness: 0.0192).

Table 7.5: Summary of the best solution found by the Genetic Algorithm for each feature set using LDA. The table lists the minimum cost function value (Best Fitness), the dimensionality of the selected subset (N_{feat}), and the specific features comprising the optimal vector.

Feature Set	Best Fitness	N_{feat}	Selected Feature Vector
`FS-Raw`	0.0360	7	WAMP, SSI, WL, Krt, MFL, FRI, SSC
`FS-Dec`	0.0213	9	$V3_d$, WL_d, MFL_d, SRI_d, SSC_d, Skw, Sync, CoA, GI
`FS-C`	**0.0192**	10	WAMP, WL, Krt, LD_d, Var_d, SSC_d, Sync, CoA, Lat, GI

7.2.2 Non-linear Logistic Regression Results

Raw Signal Features (`FS-Raw`)

The feature selection results for the FS-Raw set highlight how linear and non-linear models prioritize information differently. While the Linear Discriminant Analysis (LDA) classifier (Figure 7.7) relied on WL and FRI, the Non-Linear Logistic Regression (NLR) classifier (Figure 7.12) completely discarded WL and reduced FRI to a 20% selection rate. Instead, NLR shifted its preference to ZC to capture frequency information and MAVS to encode temporal amplitude dynamics, with both achieving a 100% selection rate. This substitution suggests that while the need for spectral and temporal information remains constant, the optimal mathematical representation of these properties depends on the classifier.

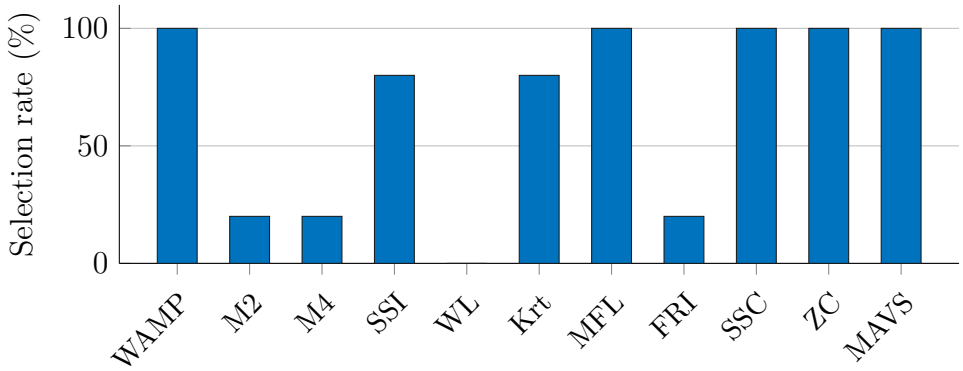

Figure 7.12: Selection frequency within the top 10 elite solutions for the FS-Raw feature set using the NLR classifier (mean fitness: 0.0242).

Deconvolved Signal Features (`FS-Dec`)

The feature selection profile of NLR is presented in Figure 7.13. Similar to the results obtained with LDA, there is a core of structural descriptors that comprise MFL_d, SSC_d and Sync which were selected in all of the elite solutions of the GA.

A notable difference is observed in the treatment of amplitude features. With the LDA classifier, WL and V3_d, were retained in 100% of the folds. Conversely, the NLR model largely suppressed these variables, reducing their selection rates to approximately 40% and 20%, respectively. Instead, higher order statistics like Krt_d and GI with selection rates exceeding 90% were favoured. This suggests that the nonlinear decision boundary prioritizes signal complexity and inequality distribution over amplitude characteristics.

Additionally, the rejection of temporal features was consistent across both classifiers. Brst, Lat, and ISIE resulted in a 0% selection rate in both NLR and LDA. This confirms that for `FS-Dec`, discriminative information is encoded primarily in the morphological and frequency domain rather than in temporal onset characteristics.

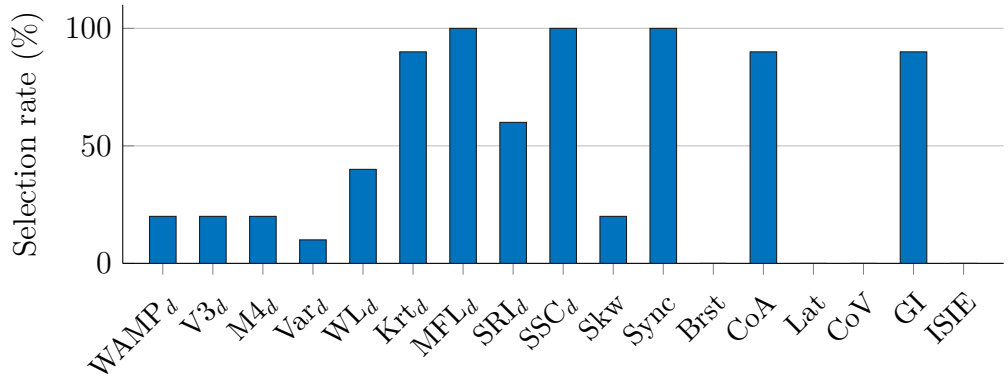

Figure 7.13: Selection frequency within the top 10 solutions of the `FS-Dec` population (mean fitness: 0.0117). Features like MFL$_d$, SSC$_d$, Sync, CoA, and GI exhibit near-saturation ($> 90\%$)

Combined Features (`FS-C`)

The feature selection analysis for the Combined Feature Set (FS-C) using the NLR classifier (Figure 7.14) identifies a diverse set of top performing features. Notably, WAMP and SSC$_d$ were retained in 100% of the solutions, identifying them as the most robust features.

Apart from these two core features, the NLR model displays a pattern of optimization instability, which is likely a consequence of the reduced number of iterations necessitated by its high computational cost. Unlike the LDA classifier, which achieved high convergence on a sparse set of secondary features (Figure 7.11), NLR failed to crystallize on a sparse set of features. Instead, it distributed selection pressure across a range of features, with SSC (70%), Brst (70%), and ZC (50%) appearing in the majority of solutions but lacking consistency.

This divergence shows a shift in information preference. While the LDA prioritized spatial coordination (CoA) and waveform characteristics (WL), the NLR classifier did not select these metrics with the same frequency. In their place, the model prioritized frequency and temporal intermittency features (SSC, Brst). This suggests that NLR discriminates based on spectral characteristics and firing patterns, rather than on the waveform complexity or muscle co-activation patterns favored by LDA.

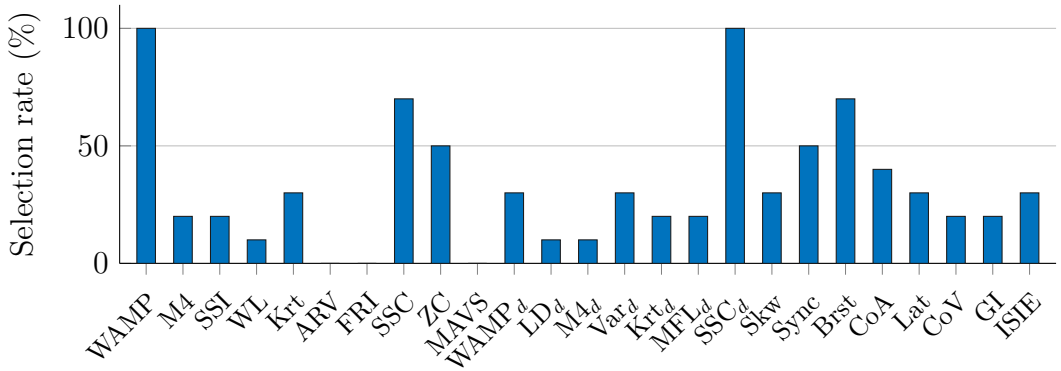

Figure 7.14: Selection frequency within the top 10 elite solutions for the FS-C feature set using the NLR classifier (mean fitness: 0.0138). In contrast to LDA, the algorithm fails to converge on secondary features.

Table 7.6: Summary of the best solution found by the Genetic Algorithm for each feature set using NLR. The table lists the minimum cost function value (Best Fitness), the dimensionality of the selected subset (N_{feat}), and the specific features comprising the optimal vector.

Feature Set	Best Fitness	N_{feat}	Selected Feature Vector
`FS-Raw`	0.0242	7	WAMP, SSI, Krt, MFL, SSC, ZC, MAVS
`FS-Dec`	**0.0111**	7	WL_d, Krt_d, MFL_d, SSC_d, Sync, CoA, GI
`FS-C`	0.0128	6	WAMP, SSC, $WAMP_d$, SSC_d, Skw, CoA

7.3 Classification Performance on Test Set

7.3.1 Linear Discriminant Analysis

Macro F1-Score

Figure 7.15 reports the distribution of per-subject Macro F1-scores on the test set for LDA across `FS-Raw`, `FS-Dec`, and `FS-C`. Statistical significance between feature sets was evaluated using a Friedman test on paired subject results. Post-hoc pairwise comparisons were then performed on the mean ranks using the Tukey–Kramer critical value to account for multiple comparisons. The analysis indicates a significant overall effect ($p < 0.05$), with both `FS-Dec` and `FS-C` significantly outperforming `FS-Raw`. Conversely, the difference between `FS-Dec` and `FS-C` does

not reach statistical significance, suggesting that most of the gain is captured by the deconvolution derived features, while the added raw descriptors provide a smaller and less systematic benefit.

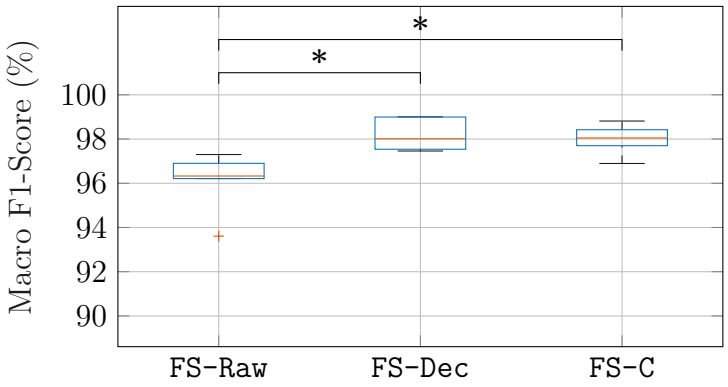

Figure 7.15: Comparison of Macro F1-Score distributions for the FS-Dec, FS-Raw, and FS-C feature sets. Brackets indicate statistically significant differences between the methods ($p < 0.05$).

This trend is reflected in the global metrics reported in Table 7.7. The combined representation `FS-C` achieves the best mean performance (Accuracy $98.80 \pm 0.52\%$, Macro F1-score $98.71 \pm 0.46\%$) and the lowest variance, indicating more stable generalization across subjects. `FS-Dec` is closely competitive (Macro F1-score $98.54 \pm 0.55\%$) and yields the lowest mean Abstention Rate (AR) ($26.36 \pm 6.69\%$), with `FS-C` nearly identical ($26.44 \pm 9.61\%$) and both clearly improving over `FS-Raw` ($29.36 \pm 12.54\%$). Importantly, accuracy and F1-score are computed only on accepted epochs, whereas AR explicitly quantifies the proportion of rejected epochs, reflecting the safety mechanism introduced by the confidence threshold (Section 6.3.3). The confidence threshold τ_{conf} is subject-dependent and selected during training. Therefore, differences in AR between feature sets reflect how each feature set affects the distribution of posterior probabilities and the frequency of borderline decisions. In contrast to other works presented in Chapter 3, the observed AR values (approximately 26–29% on average) constitute a favorable trade-off, remaining well below the high rejection rates reported by several abstention based approaches.

The per-class results in Table 7.8 clarify which gestures benefit the most and how abstention interacts with reliability. For dynamic wrist gestures, deconvolution-based representations deliver substantial gains relative to `FS-Raw`, particularly for Wrist Extension (WE) and Wrist Flexion (WF), where `FS-Dec` yields large improvements alongside markedly reduced variability. The combined set `FS-C` is especially advantageous for Wrist Pronation (WP), achieving the highest F1-score ($99.4 \pm 1.2\%$) while notably reducing abstention ($31.9 \pm 18.0\%$) compared to `FS-Raw` and `FS-Dec` (both around 44%). This pattern suggests that raw descriptors

Table 7.7: Global Performance Summary. Comparison of global reliability metrics (Mean ± SD) across feature sets. *FS-Combined* yields the most stable performance, achieving the highest F1-score and accuracy with the lowest variance.

Feature Set	**Accuracy** [%]	**F1-Score** [%]	**Abstention Rate** [%]
FS-Raw	97.49 ± 1.02	96.21 ± 0.91	29.36 ± 12.54
FS-Dec	98.58 ± 0.70	98.54 ± 0.55	**26.36 ± 6.69**
FS-Combined	**98.80 ± 0.52**	**98.71 ± 0.46**	26.44 ± 9.61

complement the features obtained from the CWF pattern by discriminating a subset of previously ambiguous WP epochs. Conversely, Hand Closing (HC) illustrates the same complementarity from the opposite direction: `FS-Dec` underperforms for this class ($96.8 \pm 2.7\%$), while `FS-C` restores performance ($98.5 \pm 1.6\%$), implying that amplitude information contained in raw features remains informative for certain gestures. Overall, the LDA results support the thesis claim that single-channel deconvolution derived features provide discriminative information that generalizes across subjects, and that combining them with standard raw descriptors can further improve stability while preserving a moderate AR compatible with responsive and safe control.

Table 7.8: Per-Class F1-score and Abstention. Detailed breakdown of the F1-Score and Abstention Rate per class (Mean ± SD). *FS-Combined* shows notable consistency, particularly in minimizing variance for complex wrist gestures (WP, WS).

Class	FS-Raw		FS-Dec		FS-C	
	F1 [%]	Abs. Rate [%]	F1 [%]	Abs. Rate [%]	F1 [%]	Abs. Rate [%]
R	97.8 ± 1.0	29.8 ± 15.0	98.4 ± 1.5	30.0 ± 10.7	**98.8 ± 0.7**	29.0 ± 15.7
HO	98.0 ± 1.8	23.8 ± 11.1	**99.1 ± 0.8**	13.6 ± 8.6	98.9 ± 0.5	19.0 ± 6.2
HC	98.4 ± 1.7	14.2 ± 6.3	96.8 ± 2.7	11.6 ± 6.8	**98.5 ± 1.6**	18.1 ± 7.4
WP	97.9 ± 3.2	44.0 ± 22.0	99.0 ± 1.0	44.5 ± 24.7	**99.4 ± 1.2**	31.9 ± 18.0
WS	96.6 ± 3.0	40.0 ± 30.2	**99.0 ± 1.0**	38.1 ± 28.6	98.8 ± 1.7	44.0 ± 29.7
WF	97.1 ± 3.9	26.4 ± 18.4	**99.4 ± 0.6**	14.2 ± 5.1	99.2 ± 1.2	15.5 ± 8.1
WE	94.6 ± 3.5	30.4 ± 16.7	**98.2 ± 0.8**	23.4 ± 6.6	97.5 ± 1.4	25.1 ± 7.1

Confusion Matrices and Misclassification Patterns

Figures 7.16–7.18 show the test-set confusion matrices for LDA, the off-diagonal cells represent the residual misclassifications that persist even after low-confidence epochs are rejected, providing insight into the common misclassifications of each feature representation.

With `FS-Raw` (Fig. 7.16), the strongest and most systematic confusions are directed toward the Rest (R) class, especially for the more dynamic wrist gestures. For instance, WE, WP, and Wrist Supination (WS) exhibit non-negligible leakage into **R** (e.g., WE→**R** ≈ 8.4%, WP→**R** ≈ 5.2%, WS→**R** ≈ 4.8%), while cross-gesture confusions remain comparatively limited. This pattern suggests that raw descriptors primarily struggle to classify weak/transition activations, rather than confusing one active gesture for another.

The deconvolution based feature set `FS-Dec` (Fig. 7.17) reduces the percentage of false negatives and sharpens the diagonal for most active classes. In particular, wrist gestures such as WP, WS, and WF reach near perfect recall (around 99%), and WE improves substantially relative to `FS-Raw` (recall $\approx$ 96.5% with WE$\rightarrow$**R** reduced to $\approx$ 3.5%). The remaining misclassifications correspond primarily of leakage into R, indicating that when mistakes occur they are primarily associated with epochs whose activation patterns remain ambiguous even after transforming the signal into its CWF.

Finally, the combined feature set `FS-C` (Fig. 7.18) preserves the overall "clean" structure achieved by `FS-Dec` while improving selected classes that benefit from conventional amplitude features. A representative example is HC, where `FS-C` reduces the leakage into R compared to `FS-Dec` (recall $\approx$ 98.7% vs. $\approx$ 96.9%), while maintaining very low cross-gesture confusions for WP, WS, and WF. Overall, the three matrices consistently show that the dominant residual error mode is *movement*$\rightarrow$R rather than movement$\rightarrow$movement, which aligns with the broader interpretation of this system: the primary difficulty is robustly detecting subtle activations, while discrimination among clearly active patterns is already strong once the confidence gate is passed.

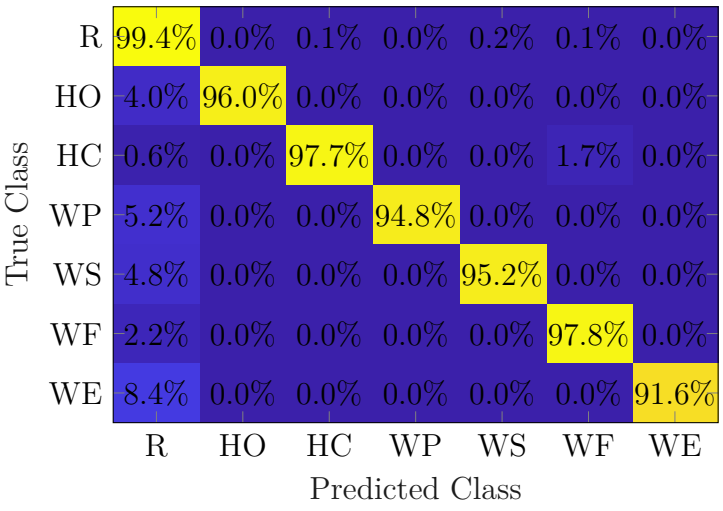

Figure 7.16: Confusion matrix for the `FS-Raw` feature set. The matrix is row-normalized, where diagonal cells represent the correct classification rate (Recall) for each gesture class and off-diagonal cells indicate misclassification percentages.

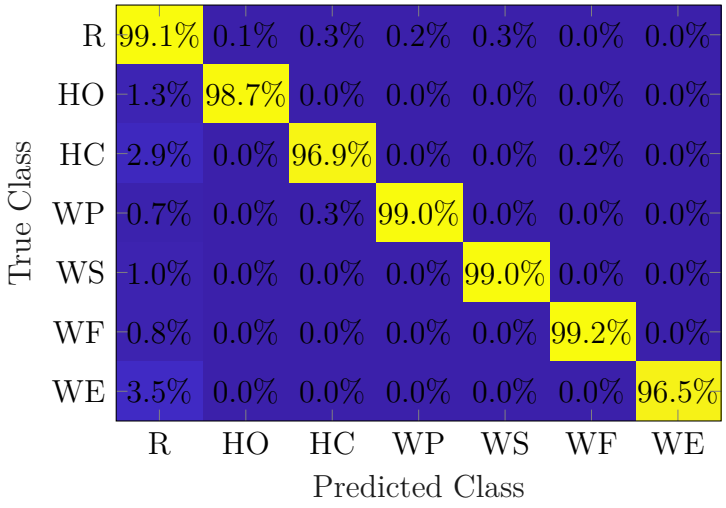

Figure 7.17: Confusion matrix for the `FS-Dec` feature set. The matrix is row-normalized, where diagonal cells represent the correct classification rate (Recall) for each gesture class and off-diagonal cells indicate misclassification percentages.

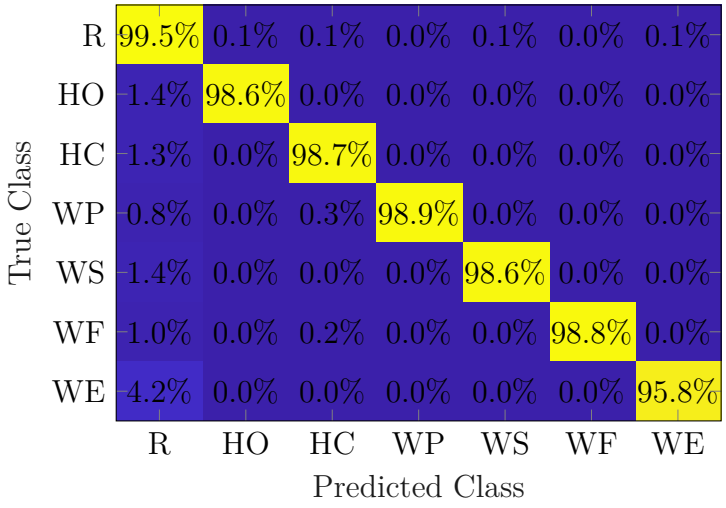

Figure 7.18: Confusion matrix for the `FS-C` feature set. The matrix is row-normalized, where diagonal cells represent the correct classification rate (Recall) for each gesture class and off-diagonal cells indicate misclassification percentages.

Confidence Threshold Analysis

Figure 7.19 reports the safety profile obtained with the subject-specific confidence threshold τ_{conf} selected during training. For each gesture class, the grouped stacked

bars partition the test epochs into correct predictions, abstentions, and misclassifications, providing a direct visual comparison of how outcomes vary across the three feature sets. A consistent trend is that the most complex wrist gestures (WP, WS) exhibit the highest AR across feature sets, indicating that these classes concentrate the greatest uncertainty and therefore drive most rejections.

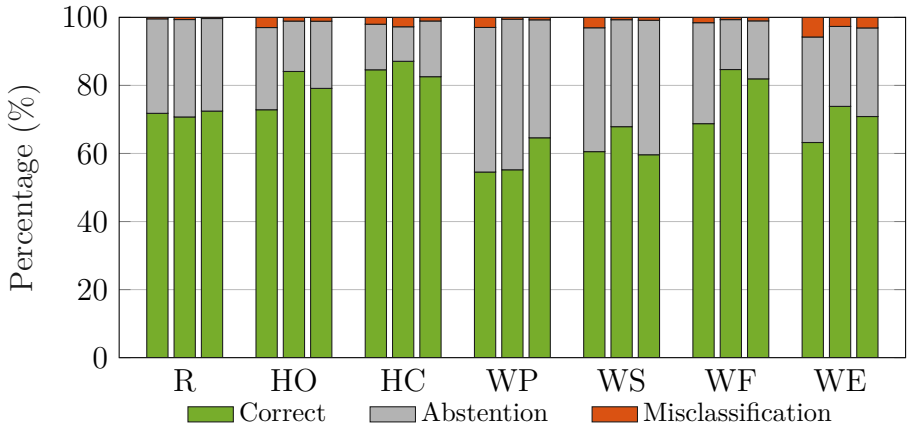

Figure 7.19: Safety profile per gesture class using LDA. For each gesture, the grouped bars represent the performance of the FS-Raw (left), FS-Dec (middle), and FS-C (right) feature sets. The stacked segments visualize the distribution of classification outcomes.

The feature sets mainly differ in how many uncertain epochs they reject versus classify correctly. As observed in Figure 7.19, with `FS-Raw` (left bars), the system relies more heavily on rejection for WP and WS, reflecting a less separable feature space for these gestures under a linear classifier. In contrast, `FS-Dec` (middle bars) generally increases the portion of correctly classified epochs with only minimal orange segments, suggesting that the deconvolution-based features increase decision confidence and reduce misclassifications. The combined set `FS-C` (right bars) maintains consistently low misclassification rates across all classes and, particularly for difficult gestures like WP, further converts part of the ambiguous decisions into correct ones compared to the raw baseline, yielding a highly balanced per-class outcome distribution.

The effect of varying τ_{conf} is shown in Fig. 7.20, which plots accuracy against AR. As abstention is allowed to increase, the accuracy on accepted samples rises and reaches a plateau, consistent with progressively rejecting the most ambiguous epochs. Across the full sweep, `FS-Dec` and `FS-C` consistently dominate `FS-Raw`, reaching higher accuracy for the same abstention level and showing tighter variability bands. This indicates that the deconvolution based representations improve the reliability of decisions not only at the chosen operating point, but throughout the threshold

range, delivering a more favorable safety responsiveness trade-off on the test set.

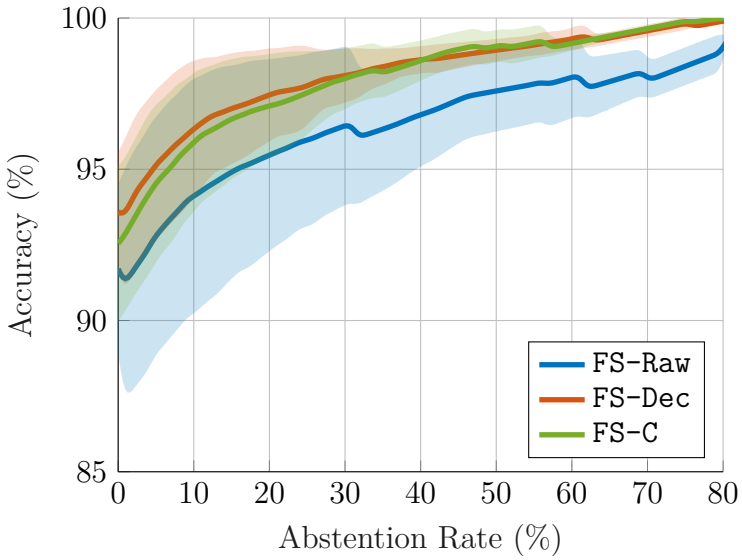

Figure 7.20: Accuracy vs. Abstention Rate trade-off curves for the FS-Raw, FS-Dec, and FS-C feature sets. The solid lines represent the mean accuracy calculated on accepted samples, while the shaded regions indicate the standard deviation. The graph illustrates the progression of classification reliability as the system is permitted to abstain from uncertain predictions.

7.3.2 Non-linear Logistic Regression

Macro F1-Score

Using a second-degree NLR, the Macro F1-score distributions in Fig. 7.21 show that both `FS-Dec` and `FS-C` improve over `FS-Raw` in a consistent way across subjects. The Friedman test on paired subject scores indicates statistically significant differences ($p < 0.05$) between `FS-Raw` and both `FS-Dec` and `FS-C`. In contrast to the results obtained for LDA, `FS-Dec` and `FS-C` remain closely matched, suggesting that once the feature space contains deconvolution derived features, the additional contribution from raw descriptors is comparatively small for this classifier.

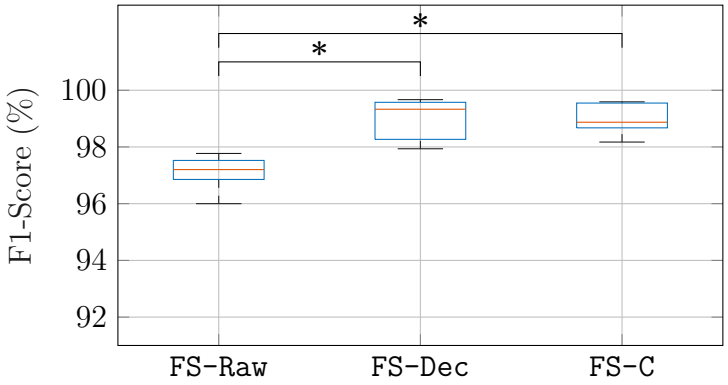

Figure 7.21: Comparison of Macro F1-Score distributions for the FS-Dec, FS-Raw, and FS-C feature sets. Brackets indicate statistically significant differences between the methods ($p < 0.05$).

As shown in Table 7.9, `FS-Dec` achieves the best mean performance and the lowest mean AR, while `FS-C` is very close in mean accuracy/F1 but shows the smallest dispersion across subjects. Compared to `FS-Raw`, the main difference is not only higher performance on the accepted samples but also a substantially lower AR.

Table 7.9: Global Performance Summary. Comparison of global reliability metrics (Mean ± SD) across feature sets. *FS-Combined* yields the most stable performance, achieving the highest F1-score and accuracy with the lowest variance.

Feature Set	Accuracy [%]	F1-Score [%]	Abstention Rate [%]
FS-Raw	98.35 ± 0.75	97.87 ± 0.95	32.29 ± 10.22
FS-Dec	**99.33 ± 0.74**	**99.07 ± 0.78**	**25.08 ± 3.70**
FS-Combined	99.21 ± 0.37	98.95 ± 0.54	25.94 ± 5.20

The per-class breakdown in Table 7.10 shows that the largest performance differences are concentrated in the complex wrist movements, where `FS-Raw` becomes unstable across subjects. The most evident case is WS, where `FS-Raw` exhibits very low and highly variable performance ($80.4 \pm 39.5\%$) together with a high AR ($53.6 \pm 30.4\%$). Both `FS-Dec` and `FS-C` largely remove this failure mode, lifting WS close to 98% while reducing abstention. A similar pattern appears for WP: `FS-Raw` combines lower F1-score ($95.2 \pm 5.5\%$) with extremely high abstention ($62.5 \pm 19.7\%$), whereas `FS-Dec`/`FS-C` improve reliability and cut abstention by more than 20%. In contrast, simpler gestures remain near optimum for all representations (e.g., Hand Opening (HO), HC), with `FS-Raw` marginally leading on HO, indicating that raw descriptors are already sufficient for those classes while deconvolution derived features are most impactful where coordinated neuromuscular control strategies drive separability.

Table 7.10: Per-Class F1-score and Abstention. Detailed breakdown of the F1-Score and Abstention Rate per class (Mean ± SD). The results show that while FS-Raw performs well on simple grasps (HO, HC), FS-Dec and FS-C achieve superior reliability on complex wrist movements (WP, WS, WF).

Class	FS-Raw		FS-Dec		FS-C	
	F1 [%]	Abs. Rate [%]	F1 [%]	Abs. Rate [%]	F1 [%]	Abs. Rate [%]
R	98.3 ± 1.1	31.5 ± 16.8	**99.2 ± 1.2**	25.1 ± 8.9	99.2 ± 0.7	26.2 ± 10.3
HO	**99.8 ± 0.6**	30.3 ± 17.3	99.7 ± 0.4	20.5 ± 4.9	99.5 ± 0.8	28.1 ± 13.1
HC	99.0 ± 1.5	16.2 ± 5.8	**99.1 ± 1.4**	16.0 ± 8.8	99.0 ± 0.9	16.4 ± 11.2
WP	95.2 ± 5.5	62.5 ± 19.7	97.7 ± 2.5	39.2 ± 14.9	**98.2 ± 2.0**	40.2 ± 12.3
WS	80.4 ± 39.5	53.6 ± 30.4	**98.3 ± 2.3**	40.7 ± 22.8	98.0 ± 1.9	35.6 ± 21.5
WF	98.9 ± 2.3	31.2 ± 20.8	**99.9 ± 0.4**	18.6 ± 4.4	99.7 ± 0.5	18.8 ± 9.3
WE	97.2 ± 2.2	24.4 ± 7.7	**99.7 ± 0.5**	27.7 ± 6.8	99.2 ± 0.5	26.2 ± 5.5

Confusion Matrices and Misclassification Patterns

Figures 7.22–7.24 report the test-set confusion matrices for the NLR classifier, computed with the same convention adopted in the LDA section. Across all three feature sets, the matrices are strongly diagonal and movement-to-movement confusions remain marginal, indicating that when the classifier commits to a decision, the predicted gesture is typically correct and rarely mistakes an active gesture for another one.

The main discrepancy between feature sets is concentrated in the complex wrist gestures, where `FS-Raw` exhibits a clear degradation. In Fig. 7.22, (WP) shows the

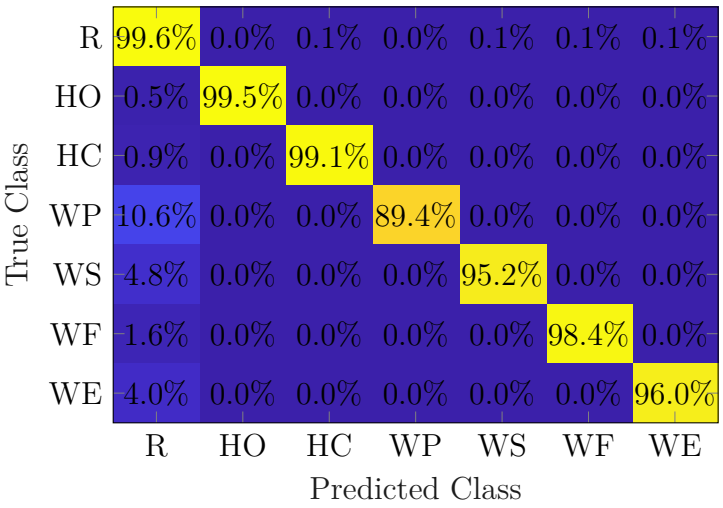

Figure 7.22: Confusion matrix for the `FS-Raw` feature set. The matrix is row-normalized, where diagonal cells represent the correct classification rate (Recall) for each gesture class and off-diagonal cells indicate misclassification percentages.

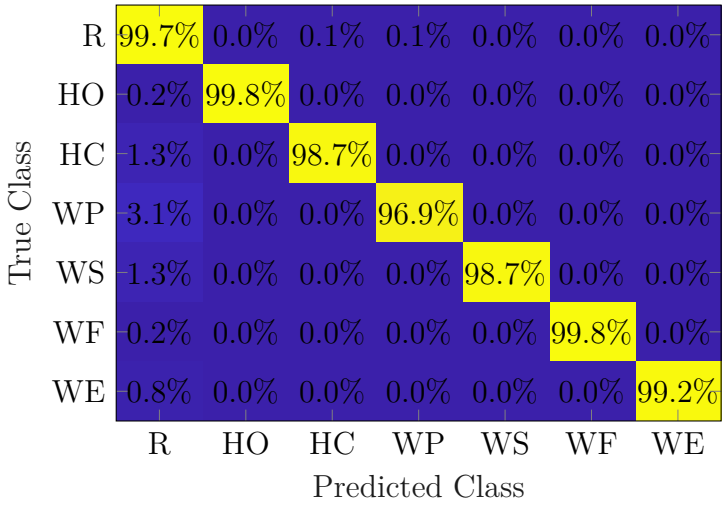

Figure 7.23: Confusion matrix for the `FS-Dec` feature set. The matrix is row-normalized, where diagonal cells represent the correct classification rate (Recall) for each gesture class and off-diagonal cells indicate misclassification percentages.

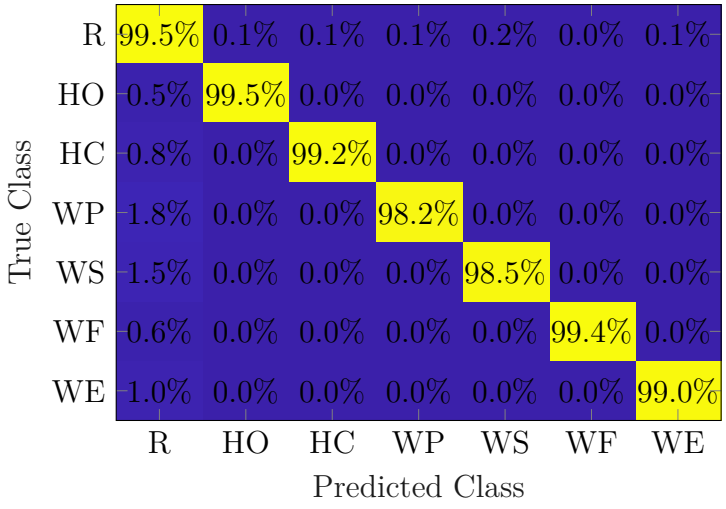

Figure 7.24: Confusion matrix for the `FS-C` feature set. The matrix is row-normalized, where diagonal cells represent the correct classification rate (Recall) for each gesture class and off-diagonal cells indicate misclassification percentages.

lowest recall (89.4%), dominated by WP→R errors (10.6%), with a similar but smaller tendency for WS and WE to collapse into R (4.8% and 4.0%, respectively). The introduction of deconvolution features (Fig. 7.23) largely decreases this misclassification pattern, WP recall increases to 96.9% (WP→R reduced to 3.1%), and the remaining wrist classes approach ceiling performance. The combined set

(Fig. 7.24) further tightens the diagonal for WP, reaching 98.2% recall with WP→R reduced to 1.8%. Overall, for NLR the benefit of `FS-Dec` and `FS-C` is primarily the stabilization of the wrist rotation classes.

Confidence Threshold Analysis

Figure 7.25 summarizes the per-class outcome distribution of the NLR classifier. The most evident pattern is the very small misclassification component across classes: when the model produces an accepted decision, it is rarely wrong, and most ambiguous epochs are handled through abstention. As expected, the most complex gestures (WP, WS) concentrate the largest gray segments, confirming that these gestures are the primary source of uncertain epochs.

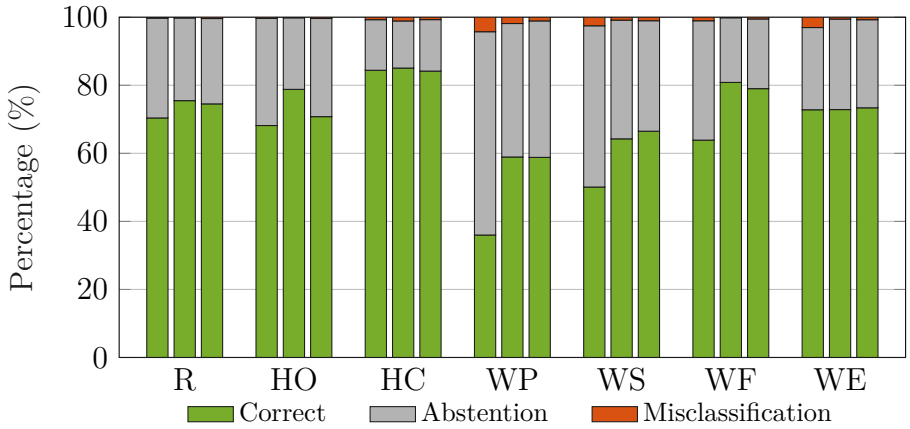

Figure 7.25: Safety profile per gesture class using the NLR classifier. For each gesture, the grouped bars represent the performance of the FS-Raw (left), FS-Dec (middle), and FS-C (right) feature sets. The stacked segments visualize the distribution of classification outcomes.

The feature representation mainly affects how much of this uncertainty can be converted into correct classifications without inflating errors. As observed in Figure 7.25, with `FS-Raw` (left bars), WP and WS show the highest reliance on abstention, indicating that even with a non-linear boundary the raw space leaves these classes comparatively ambiguous. Both deconvolution-based sets convert abstentions to correct classifications for the wrist gestures while keeping misclassifications close to zero (`FS-Dec` and `FS-C`, middle and right bars, respectively). In practice, this means that `FS-Dec` and `FS-C` improve the latency of WP/WS decisions without paying for it in additional misclassifications, which is the intended behavior of the confidence mechanism.

Figure 7.26 extends the same observation by sweeping the confidence threshold

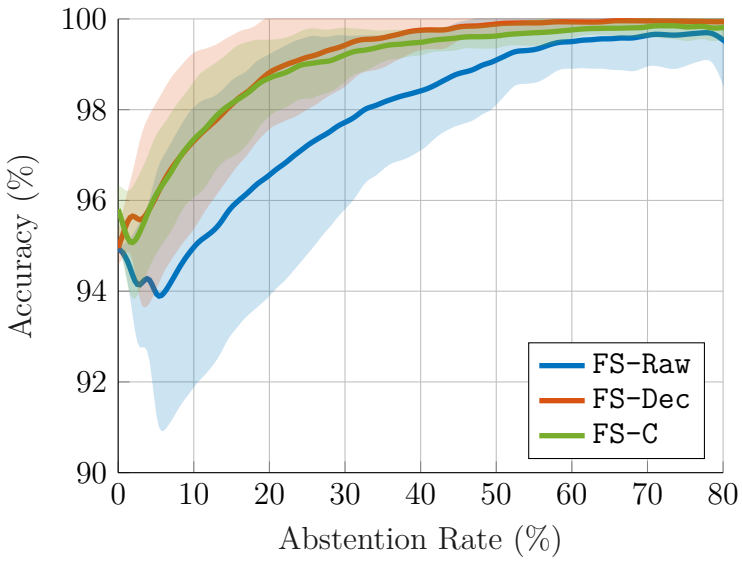

Figure 7.26: Accuracy vs. Abstention Rate trade-off curves for the FS-Raw, FS-Dec, and FS-C feature sets. The solid lines represent the mean accuracy calculated on accepted samples, while the shaded regions indicate the standard deviation.

and averaging the resulting curves. At low abstention levels, `FS-Dec` and `FS-C` start from a higher accuracy baseline and show tighter variability bands than `FS-Raw`, indicating a more consistent reliability. As abstention increases, all feature sets approach a near-ceiling regime; however, `FS-Raw` requires a higher AR to close the gap, whereas the deconvolution-based sets reach high accuracy earlier in the sweep. Overall, the curves confirm that the benefit of `FS-Dec`/`FS-C` is not limited to a single threshold choice, but persists across the full confidence range.

7.4 Online Gesture Classification

This section introduces the online (real-time) gesture classification results. In contrast to the offline evaluation on held-out test data (Section 7.3), the online setting operates on streaming EMG under latency constraints and preserves the abstention mechanism to reduce unsafe misclassifications.

7.4.1 Online Experimental Setup

Training and test separation

The training data was shared across all classifier–feature-set configurations, and model fitting was completed before online testing started. The only parameter that was updated during inference was the kernel used in deconvolution to account for frequency changes characteristic of fatigue(Section 3.5.3). Both classifiers (LDA and NLR) were evaluated online using all feature representations considered offline (FS-Raw, FS-Dec, FS-C), ensuring that comparisons reflect the effect of feature representations and classifier performance rather than intra-subject variability across sessions.

Task-based protocol and GUI feedback

Online evaluation used a task-based protocol with fixed duration target cues separated by rest intervals. The gestures used included the offline set and were extended with three additional gestures-Ulnar Deviation (UD), Radial Deviation (RD), and Pinch (P)—to increase task complexity given the near saturated offline performance.

A dedicated Graphic User Interface (GUI) provided real-time feedback by displaying the current target gesture, the remaining time within the trial, and the next required gesture.

Confidence-threshold calibration

A first cycle of the online protocol was used to determine an optimal confidence threshold τ_{conf} using the cost function in Eq. 6.12. Once selected, τ_{conf} was held constant for the remainder of the online run.

7.4.2 Real-time Processing Pipeline

The real-time pipeline matches the offline preprocessing and feature definitions described in Chapter 5, but is executed on a streaming input. Signal conditioning includes stateful notch and low-pass filtering to preserve continuity between successive segments.

Buffered deconvolution and trimming

The implementation uses a trimming strategy to reduce boundary artifacts. Deconvolution is computed on an extended context buffer that includes additional samples before and after the portion of interest; the unreliable edges are then removed by trimming a fixed margin. Only the remaining central portion aligned with the current update is retained and appended to a continuously stitched deconvolved stream. Features are extracted from the most recent processed samples so that raw, deconvolved, and combined feature sets remain time-aligned, while introducing a fixed and controlled algorithmic delay determined by the trimming margin.

7.4.3 Online Results

This subsection reports online performance for each classifier (LDA, NLR) and each feature set (`FS-Raw`, `FS-Dec`, `FS-C`).

Online testing was conducted on five able-bodied volunteers performing the extended set of gestures. To mitigate the impact of the user's reaction time and the delay introduced by cue switching, the first three predictions after each task change (approximately 300 ms) were excluded from the analysis. This choice prevents penalizing the classifier for the initial delay in motor initiation.

Table 7.11 shows that LDA with `FS-Dec` achieved the best overall online performance, with a Macro-F1 of **96.73%** and accuracy of **96.63%**. Although its AR is slightly higher than the other LDA feature sets, this behavior is consistent with a more conservative decision strategy: `FS-Dec` appears to better capture class-specific structure during training and, when uncertainty arises online, it tends to reject ambiguous epochs rather than committing to an incorrect command. From a usability perspective, this trade-off is often desirable in assistive interfaces, where limiting unsafe activations can be more important than maximizing the number of issued commands, and the observed AR remains acceptable for real-time control.

Table 7.11: Online performance summary. Mean ± SD across subjects.

Classifier	Feature set	Macro-F1 (%)	Acc (%)	AR (%)
LDA	FS-Raw	95.19 ± 0.59	95.37 ± 0.75	7.87 ± 2.27
LDA	**FS-Dec**	**96.73±0.44**	**96.63±0.58**	15.64 ± 1.62
LDA	FS-C	95.68 ± 0.94	95.95 ± 0.71	10.38 ± 1.91
NLR	FS-Raw	84.18 ± 3.74	88.01 ± 2.23	5.66 ± 4.05
NLR	FS-Dec	93.37 ± 1.84	93.82 ± 1.84	**2.12±1.59**
NLR	FS-C	92.22 ± 4.04	93.01 ± 2.48	8.26 ± 5.64

In contrast, NLR performances were lower online than in offline testing, which is plausibly explained by the reduced amount of training data used for online sessions compared to the offline dataset. In this setting, `FS-Dec` still provided the strongest NLR results, supporting the idea that deconvolution-derived descriptors remain informative in real-time conditions.

Table 7.12: Per-Class F1-score and Abstention (Online, LDA FS-Dec). Clean F1-score (accepted predictions only) and abstention rate per class (Mean ± SD across sessions).

Class	**F1[%]**	**Abs. Rate [%]**
R	96.4 ± 1.3	10.1 ± 1.3
HO	98.2 ± 1.9	1.3 ± 0.7
HC	98.7 ± 0.7	2.9 ± 2.4
WP	99.0 ± 1.1	3.6 ± 3.2
WS	97.6 ± 1.4	35.7 ± 15.5
WF	96.5 ± 3.0	32.9 ± 9.8
WE	96.7 ± 2.7	41.2 ± 8.5
UD	98.3 ± 1.6	28.3 ± 8.7
RD	96.9 ± 1.4	9.2 ± 6.2
P	92.9 ± 4.5	13.4 ± 16.0

The per-class breakdown in Table 7.12 confirms that most gestures achieve high clean F1-scores, with particularly strong performance on HO/HC and WP. Higher abstention is concentrated in the more demanding wrist movements (WS, WF, WE, and UD), indicating larger overlap in the underlying EMG patterns and a greater tendency of the system to reject uncertain epochs for these classes. This behavior aligns with a safety-oriented design: the system preserves high reliability

for accepted decisions while selectively reducing decisions for movements that are harder to separate online.

The confusion matrix in Fig. 7.27 shows a strong diagonal structure, indicating that the deconvolution-based feature set preserves high class separability also in real-time operation. In particular, HO/HC and WP/WS present near-ceiling recall, while most off-diagonal terms remain limited to a few percent. A recurrent misclassification is observed where WF and WE are predicted as UD (3.7% and 2.2%, respectively). Although not ideal, this confusion is biomechanically plausible because ulnar deviation shares forearm muscle involvement and contraction patterns with wrist flexion/extension; sEMG can therefore exhibit partially overlapping signatures depending on posture and effort. Importantly, these errors are considerably less frequent than the correct classifications and are less critical than confusions between antagonistic hand actions (e.g., HO vs. HC). Overall, the confusion matrix is consistent with the offline findings and suggests that deconvolution helps maintain discriminability across gesture classes when using Low-Density Electromyography (LD-EMG) configurations.

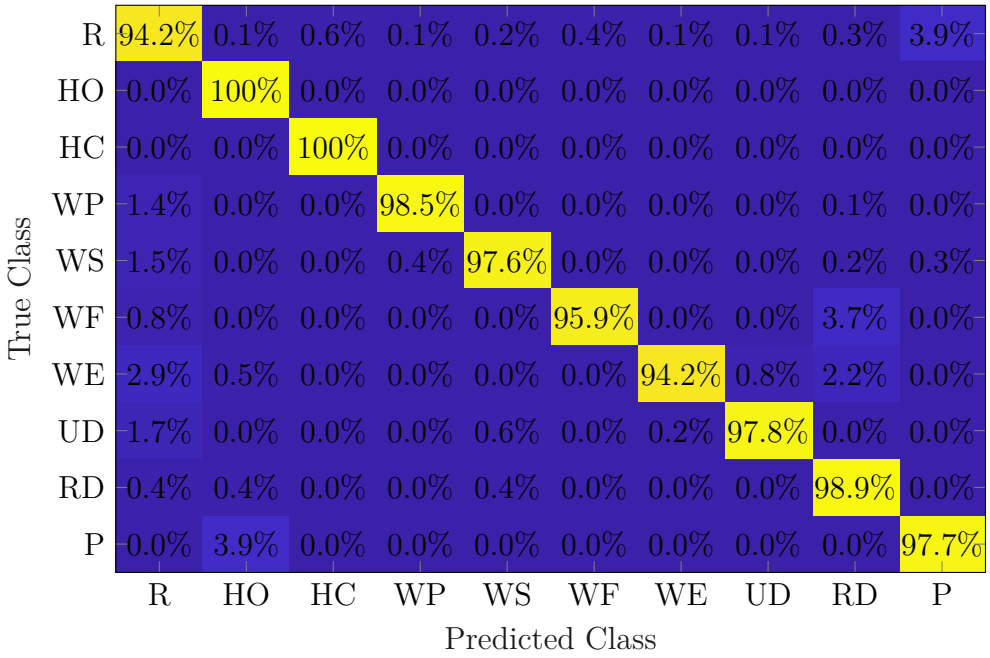

Figure 7.27: Confusion matrix showing the online performance of the chosen classifier and feature set. The matrix is row-normalized: diagonal cells represent the correct classification rate (Recall) for each gesture class, while off-diagonal cells indicate misclassification percentages.

Misclassifications involving R are also informative. Some of the misclassifications

may be explained by residual user-delay effects when transitioning into R, where incomplete relaxation or transient tonic activity can resemble low-level activation patterns. This interpretation is consistent with the lower F1-score for P, which is mainly driven by a reduction in precision: R was sometimes predicted as P (3.9%). Because R was the most numerous class, even a small fraction of R→P errors can substantially increase false positives for P and thus reduce its precision and F1-score.

From a control perspective, the grouped safety profile in Fig. 7.28 provides a clear visualization of how the different feature sets manage uncertainty. Most notably, `FS-Dec` (middle bars) consistently yields the lowest misclassification rates (orange segments) across the evaluated gestures. While `FS-Dec` occasionally exhibits a higher abstention rate than `FS-Raw`—particularly for complex wrist movements like WS, WF, and WE—this is a direct consequence of its superior class characterization. By defining tighter and more accurate decision boundaries within a restricted feature space, the model is better equipped to recognize ambiguous epochs. Consequently, instead of guessing and making an error, it safely discards these epochs as abstentions, effectively minimizing critical mistakes.

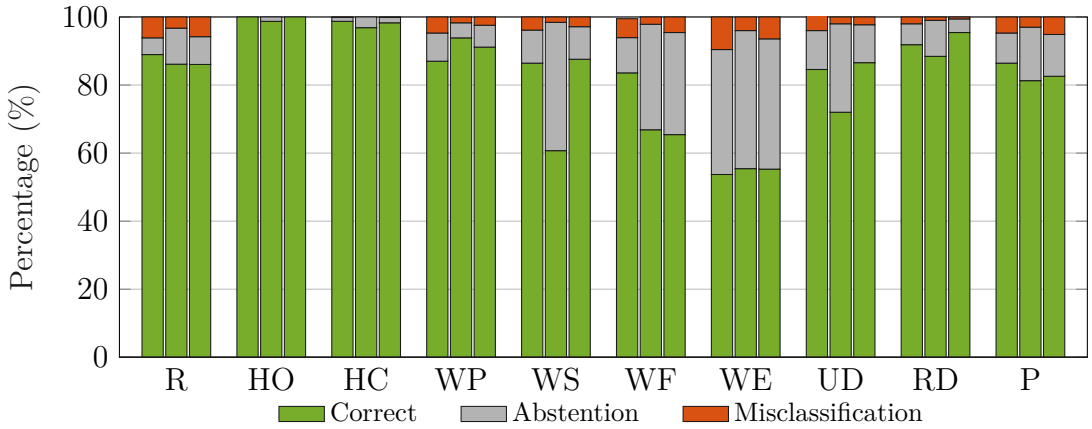

Figure 7.28: Safety profile per gesture class obtained from online testing. For each gesture, the grouped bars represent the performance of the FS-Raw (left), FS-Dec (middle), and FS-C (right) feature sets. The stacked segments visualize the distribution of outcomes (correct, abstention, and misclassification).

Overall, the online evaluation supports that features extracted from the deconvolved signal improve robustness in real-time myoelectric control. In particular, **LDA + FS-Dec** yields the most favorable reliability–coverage trade-off under the tested conditions: the slightly higher AR is accompanied by the best global accuracy and Macro-F1, indicating that the classifier remains conservative when uncertainty is high while maintaining strong performance on accepted predictions.

Crucially, these results confirm the performance trends observed offline and demonstrate that deconvolution is a useful preprocessing technique for EMG signals acquired with low-density configurations, enabling accurate and safety-oriented online classification without additional temporal smoothing.

Latency and Real-Time Responsiveness

In addition to classification reliability, online usability depends on the delay between the onset of the gesture and the corresponding output by the classifier. In this work, latency is defined as the time between the produced decision and the center of the analysis window used to compute that decision. With a window length of $W = 250\,\text{ms}$ and hop size $H = 100\,\text{ms}$, the window center precedes the window end by $W/2 = 125\,\text{ms}$. Denoting by t_{end} the timestamp of the last sample in the window and by t_{out} the time at which the classifier output becomes available, the effective latency is:

$$L_{\text{eff}} = t_{\text{out}} - \left(t_{\text{end}} - \frac{W}{2}\right) = (t_{\text{out}} - t_{\text{end}}) + \frac{W}{2}. \tag{7.1}$$

This definition captures both the intrinsic delay introduced by the size of the analysis window and the additional delay due to processing, buffering, and software execution $(t_{\text{out}} - t_{\text{end}})$.

To complement latency, time between accepted decisions is also reported. This measure reflects the effective command update rate perceived by the user. Let $t_{\text{acc}}^{(k)}$ be the time at which the k-th accepted decision is produced; then the inter-decision interval is defined as:

$$\Delta t_{\text{acc}}(k) = t_{\text{acc}}^{(k)} - t_{\text{acc}}^{(k-1)}. \tag{7.2}$$

This metric is particularly relevant for selective classification: higher ARs can increase Δt_{acc}, potentially reducing responsiveness even when reliability improves. For reporting, L_{eff} and Δt_{acc} are summarized using robust statistics (median and 95th percentile) across the online stream, providing both typical and worst-case responsiveness.

Table 7.13: Online latency and responsiveness (LDA). Effective decision latency (Eq. 7.1) and time between accepted decisions (Eq. 7.2). Values reported as median [95th percentile] across the online stream.

Feature set	L_{eff} [ms]	Δt_{acc} [ms]	AR [%]
FS-Raw	129[133]	100[105]	8.75
FS-Dec	219[238]	100[202]	16.87
FS-C	225[242]	100[204]	10.39

Chapter 8

Conclusions and Future Work

8.1 Conclusions

The results support the main hypothesis of the thesis. First, the proposed deconvolution pipeline proved compatible with real-time operation and low computational burden. Since the processing is performed independently on each channel, the method avoids the complexity typically associated with multi-channel decomposition algorithms and can be integrated into a streaming pipeline with controlled latency. This demonstrates that discharge-inspired signal representations are not restricted to offline analysis or high-density laboratory acquisitions, and that highly discriminative control information can be successfully extracted from a reduced, practical number of electrodes.

Second, deconvolution was shown to be an effective preprocessing strategy for gesture recognition. The estimated CWF patterns enabled the extraction of features describing temporal organization, sparsity, and coordination properties of the signal, yielding a representation that consistently improved over conventional raw-signal descriptors. While the offline results suggest that raw and deconvolved representations may still retain some complementary information, the online evaluation provided the clearest practical indication: under real-time conditions, the deconvolution-based feature set achieved the strongest performance. This supports the view that deconvolution is not only useful, but particularly advantageous in the setting most relevant for wearable myoelectric control.

The online findings are especially relevant because they move beyond the optimistic conditions of held-out offline testing. In the real-time protocol, the deconvolution-based representation preserved strong class separability and achieved the best classification performance on the extended gesture set, showing that its

advantage is not limited to offline evaluation. This suggests that deconvolution allows for the extraction of a feature representation that transfers well to online control, where causality, latency, and temporal variability make the problem more challenging.

A related observation concerns the behavior of the classifiers and their interaction with the selected features. The feature selection process demonstrated that the optimal feature subset is highly dependent on the chosen classifier. Because LDA and NLR partition the feature space differently—linearly versus non-linearly—they rely on different combinations of descriptors to maximize class separability. Furthermore, while NLR achieved excellent results in the offline experiments, it did not perform equally well in the online scenario, where only a limited amount of training data was available before real-time testing. This suggests that the data-constrained online condition particularly penalized NLR's performance, likely due to overfitting on the small calibration set. In contrast, LDA proved much more reliable in this setting. When paired with its specifically optimized deconvolution-based feature set, LDA achieved the best overall online performance. This result highlights that strong offline performance does not automatically translate into superior online usability, and reinforces the importance of co-optimizing the feature set and classifier under realistic operating conditions.

Another important outcome of this work is the role of abstention. The introduction of a confidence-based rejection mechanism reduced false positive decisions and improved the reliability of the system, which is particularly important in assistive applications where an incorrect activation may be more problematic than a missed one. The results showed that, when properly tuned, abstention improves the operating point of the classifier by rejecting uncertain outputs and increasing effective classification performance, especially in the more demanding online scenario.

More broadly, the findings confirm the difficulty of extracting large performance gains in a mature problem space such as sEMG-based gesture recognition. However, even moderate improvements become highly valuable when achieved under the strict constraints of real-time processing and wearable applicability. Deconvolution provides precisely this kind of meaningful, practical advancement, improving classification accuracy without sacrificing computational efficiency or hardware simplicity.

Overall, this thesis shows that low-channel, online-compatible deconvolution is feasible, computationally practical, and beneficial for myoelectric gesture recognition. Most importantly, it shows that deconvolution can provide a stronger signal representation than conventional raw sEMG features in the online setting that is most relevant for real-world assistive applications, while offline results further suggest that it can also interact constructively with standard representations.

8.2 Future Work

A first natural continuation of this work is the design of a dedicated acquisition device tailored to the LD-EMG configuration adopted in this thesis. Since the proposed pipeline only relies on four channels, a custom hardware solution could be developed to acquire only the required channels. This would represent an important step toward a fully wearable implementation and would allow the proposed methodology to be tested on hardware designed according to the requirements of the system.

A closely related development is the implementation of the complete processing and classification chain on embedded hardware. Although the present results indicate that the method is compatible with real-time constraints, an embedded realization would make it possible to measure latency, memory occupation, processor load, and energy consumption directly on the target platform. This would also enable a more rigorous comparison between raw and deconvolution-based pipelines under the computational constraints of portable systems, and would help identify the most critical stages for further optimization.

Beyond computational optimization, evaluating the system's robustness against muscular fatigue is a critical next step. In particular, dedicated fatigue-inducing protocols should be designed to test whether the kernel update strategy is sufficient to compensate for the signal changes caused by muscular fatigue during prolonged operation. This would provide a more realistic assessment of the stability of the deconvolution process and of the stability of the extracted features over time.

Similarly, future studies must address the forearm's position effect. Variations in arm and forearm posture are known to alter sEMG characteristics and may reduce classification performance in practical use. For this reason, the proposed pipeline should be tested under controlled protocols involving multiple limb positions, in order to quantify performance degradation and determine whether the current feature set and calibration strategy remain reliable across postural changes.

From a clinical perspective, it is imperative to assess the proposed methodology on target users, including amputee subjects when feasible. Testing on end-users is necessary to verify whether the advantages observed in the present work remain valid under the physiological and operational conditions of real prosthetic control. This would also make it possible to identify practical limitations that cannot be fully captured in experiments on intact subjects.

Finally, building upon the successful online validation achieved in this work, the complete system should be evaluated in longitudinal sessions during activities of daily living. While the current real-time protocol rigorously demonstrated algorithmic reliability and responsiveness, future developments should expand to quantify long-term usability metrics. These include daily setup time, calibration repeatability across multiple days, and user interaction with the interface over

extended periods. Assessing these human factors represents the final stage in translating the proposed low-density, deconvolution-based pipeline into a clinically deployable wearable device.

Bibliography

[1] Néstor J. Jarque-Bou, Joaquín L. Sancho-Bru, and Margarita Vergara. «A Systematic Review of EMG Applications for the Characterization of Forearm and Hand Muscle Activity during Activities of Daily Living: Results, Challenges, and Open Issues». In: *Sensors* 21.9 (2021), p. 3035. DOI: 10.3390/s21093035 (cit. on p. 1).

[2] Dario Farina, Ning Jiang, Hubertus Rehbaum, Aleš Holobar, Bernhard Graimann, Hans Dietl, and Oskar C. Aszmann. «The extraction of neural information from the surface EMG for the control of upper-limb prostheses: emerging avenues and challenges». In: *IEEE Transactions on Neural Systems and Rehabilitation Engineering* 22.4 (2014), pp. 797–809. DOI: 10.1109/TNSRE.2014.2305111 (cit. on p. 1).

[3] Stefan Salminger, Heiko Stino, Lukas H. Pichler, Clemens Gstoettner, Agnes Sturma, Johannes A. Mayer, Michael Szivak, and Oskar C. Aszmann. «Current rates of prosthetic usage in upper-limb amputees – have innovations had an impact on device acceptance?» In: *Disability and Rehabilitation* 44.14 (2022), pp. 3708–3713. DOI: 10.1080/09638288.2020.1866684 (cit. on pp. 1, 25, 26, 30, 40).

[4] Feng Qiu, Xinyu Liu, and Xudong Ye. «Influence of electrode placement on the recognition of different gesture categories using high-density sEMG». In: *Frontiers in Neuroscience* 19 (2025), p. 1750792. DOI: 10.3389/fnins.2025.1750792 (cit. on pp. 1, 34, 35).

[5] Dario Farina, Aleš Holobar, Roberto Merletti, and Roger M. Enoka. «Decoding the neural drive to muscles from the surface electromyogram». In: *Clinical Neurophysiology* 121.10 (2010), pp. 1616–1623. DOI: 10.1016/j.clinph.2009.10.040 (cit. on pp. 1, 37).

[6] Debora Quadrelli, Michele Canepa, Dario Di Domenico, Nicolò Boccardo, Michela Chiappalone, and Matteo Laffranchi. «Advances in HD-EMG interfaces and spatial algorithms for upper limb prosthetic control». In: *Frontiers in Neuroscience* 19 (2025), p. 1655257. DOI: 10.3389/fnins.2025.1655257 (cit. on pp. 2, 25, 26, 33–36, 66).

[7] Dario Di Domenico et al. «Reach&Grasp: a multimodal dataset of the whole upper-limb during simple and complex movements». In: *Scientific Data* 12.1 (2025), p. 233. DOI: 10.1038/s41597-025-04552-5 (cit. on pp. 2, 50–53).

[8] Luca Mesin. «Single channel surface electromyogram deconvolution to explore motor unit discharges». In: *Medical & Biological Engineering & Computing* 57.9 (2019), pp. 2045–2054. DOI: 10.1007/s11517-019-02010-0 (cit. on pp. 2, 24, 41, 44–46).

[9] Johnston JA, Bobich LR, and Santello M. «Coordination of intrinsic and extrinsic hand muscle activity as a function of wrist joint angle during two-digit grasping.» In: *Neurosci Lett.* 474(2) (Apr. 2010), pp. 104–108 (cit. on p. 4).

[10] R. Merletti and D. Farina. *Surface Electromyography: Physiology, Engineering, and Applications*. 445 Hoes Lane Piscataway, NJ 0885: IEEE Press, 2016 (cit. on pp. 4, 8, 9, 14, 16, 20, 27, 32, 36).

[11] Pereira Botelho D, Curran K, and Lowery MM. «Anatomically accurate model of EMG during index finger flexion and abduction derived from diffusion tensor imaging.» In: *PLOS Computational Biology* 15 (Aug. 2019), pp. 1–24 (cit. on p. 4).

[12] J. Gordon Betts and et al. *Anatomy and Physiology*. OpenStax, Rice University, 2022 (cit. on pp. 4–9).

[13] Katja Hoehn, Lawrence W. Haynes, and Matthew A. Abbott. *Marieb Human Anatomy & Physiology*. 12th ed. Hoboken, NJ: Pearson, 2024 (cit. on pp. 6, 7, 11–13).

[14] Roberto Merletti and Glenda L. Cerone. «Tutorial. Surface EMG detection, conditioning and pre-processing: Best practices.» In: *Journal of Electromyography and Kinesiology* 54 (June 2020), p. 102440 (cit. on pp. 6, 20, 22, 23).

[15] Robert L. Barchi. «Abnormal Excitability of the Sarcolemma Can Affect Muscle Function». In: *Basic Neurochemistry: Molecular, Cellular and Medical Aspects*. Ed. by George J. Siegel, Bernard W. Agranoff, R. Wayne Albers, Stephen K. Fisher, and Michael D. Uhler. 6th ed. Philadelphia: Lippincott-Raven, 1999 (cit. on p. 7).

[16] C. McCuller, R. Jessu, and A. L. Callahan. «Physiology, Skeletal Muscle». In: *StatPearls*. Treasure Island (FL): StatPearls Publishing, July 2023 (cit. on p. 7).

[17] Faizan Khan, John Heckmatt, and Adnan Qureshi. «Physiology, Neuromuscular Junction». In: *StatPearls*. Treasure Island, FL: StatPearls Publishing, 2025 (cit. on p. 7).

[18] Jacques Duchateau and Roger M. Enoka. «Human motor unit recordings: origins and insight into the integrated motor system». In: *Brain Research* 1409 (2011), pp. 42–61. DOI: 10.1016/j.brainres.2011.06.011 (cit. on p. 7).

[19] Scott Francis Davis and Alan David Kaye, eds. *Principles of Neurophysiological Assessment, Mapping, and Monitoring*. 2nd ed. Cham: Springer, 2020 (cit. on p. 8).

[20] Isabella Campanini, Andrea Merlo, Catherine Disselhorst-Klug, Luca Mesin, Silvia Muceli, and Roberto Merletti. «Fundamental Concepts of Bipolar and High-Density Surface EMG Understanding and Teaching for Clinical, Occupational, and Sport Applications: Origin, Detection, and Main Errors». In: *Sensors* 22.11 (2022), p. 4150. DOI: 10.3390/s22114150 (cit. on pp. 8, 14, 16–19).

[21] Daniel Dumitru, John C. King, and William E. Rogers. «Motor unit action potential components and physiologic duration». In: *Muscle & Nerve* 22.6 (1999), pp. 733–741. DOI: 10.1002/(SICI)1097-4598(199906)22:6<733::AID-MUS10>3.0.CO;2-6 (cit. on p. 8).

[22] Lorne M. Mendell. «The size principle: a rule describing the recruitment of motoneurons». In: *Journal of Neurophysiology* 93.6 (2005), pp. 3024–3026 (cit. on p. 9).

[23] Eike Petersen and Philipp Rostalski. «A comprehensive mathematical model of motor unit pool organization, surface electromyography, and force generation». In: *Frontiers in Physiology* 10 (2019), p. 176 (cit. on pp. 9, 14, 16, 17).

[24] C. J. Heckman and Roger M. Enoka. «Motor Unit». In: *Comprehensive Physiology* 2 (2012), pp. 2629–2682. DOI: 10.1002/cphy.c100087 (cit. on p. 9).

[25] Dario Farina, Roberto Merletti, and Roger M. Enoka. «The extraction of neural strategies from the surface EMG: an update». In: *Journal of Applied Physiology* 117.11 (2014), pp. 1215–1230. DOI: 10.1152/japplphysiol.00162.2014 (cit. on pp. 9, 14, 16, 18–20, 22, 24).

[26] Laboratorio di Ingegneria del Sistema Neuromuscolare (LISiN). *Official Website*. Politecnico di Torino. 2024. URL: https://www.lisin.polito.it/ (cit. on pp. 10, 14, 21, 23).

[27] M. Latarjet and A. Testut. *Compendio de anatomía descriptiva*. Barcelona: Salvat, 1958 (cit. on p. 9).

[28] E. Okwumabua, M. A. Sinkler, and B. Bordoni. «Anatomy, Shoulder and Upper Limb, Hand Muscles». In: *StatPearls [Internet]*. Updated 2023 Jul 24. Treasure Island (FL): StatPearls Publishing, July 2023. URL: https://www.ncbi.nlm.nih.gov/books/NBK537229/ (cit. on pp. 9, 12).

[29] Marco Gazzoni, Nicolò Celadon, Daniela Mastrapasqua, Matteo Paleari, Valentina Margaria, and Paolo Ariano. «Quantifying Forearm Muscle Activity during Wrist and Finger Movements by Means of Multi-Channel Electromyography». In: *PLOS ONE* 9.10 (2014), e109943. DOI: 10.1371/journal.pone.0109943 (cit. on p. 12).

[30] John V. Basmajian and Carlo J. De Luca. *Muscles Alive: Their Functions Revealed by Electromyography*. 5th. Baltimore: Williams & Wilkins, 1985. ISBN: 978-0683004144 (cit. on p. 14).

[31] Carlo J. De Luca. «The Use of Surface Electromyography in Biomechanics». In: *Journal of Applied Biomechanics* 13.2 (1997), pp. 135–163. DOI: 10.1123/jab.13.2.135 (cit. on pp. 14, 16–18, 20–23).

[32] Dario Farina, Mauro Fosci, and Roberto Merletti. «Motor unit recruitment strategies investigated by surface EMG variables». In: *Journal of Applied Physiology* 92.1 (2002), pp. 235–247. DOI: 10.1152/japplphysiol.00650.2001 (cit. on pp. 14, 15).

[33] Dario Farina and Alberto Rainoldi. «Compensation of the effect of subcutaneous tissue layers on surface EMG: a simulation study». In: *Journal of Electromyography and Kinesiology* 9.6 (1999), pp. 381–391. DOI: 10.1016/S1050-6411(99)00010-4 (cit. on p. 14).

[34] Roberto Merletti. *Teaching Module 5: Basic Electrophysiology and sEMG generation*. Project CoMES, LISiN, Politecnico di Torino. Accessed: 2026-01-04. 2018. URL: https://www.robertomerletti.it/en/emg/material/teaching/module5/ (cit. on pp. 16–18).

[35] Dario Farina, Roberto Merletti, and Roger M. Enoka. «The extraction of neural strategies from the surface EMG». In: *Journal of Applied Physiology* 96.4 (2004), pp. 1486–1495. DOI: 10.1152/japplphysiol.01070.2003 (cit. on pp. 20–24).

[36] Yong-Ku Kong and Brian D. Lowe. «Crosstalk effect on surface electromyogram of the forearm flexors during a power grip task». In: *Journal of Electromyography and Kinesiology* 20.6 (2010), pp. 1223–1229. DOI: 10.1016/j.jelekin.2010.08.001 (cit. on p. 21).

[37] Jeremy P. M. Mogk and Peter J. Keir. «Crosstalk in surface electromyography of the proximal forearm during gripping tasks». In: *Journal of Electromyography and Kinesiology* 13.1 (2003), pp. 63–71. DOI: 10.1016/S1050-6411(02)00071-8 (cit. on p. 21).

[38] D. G. Allen, G. D. Lamb, and H. Westerblad. «Skeletal Muscle Fatigue: Cellular Mechanisms». In: *Physiological Reviews* 88.1 (2008), pp. 287–332 (cit. on pp. 22, 32).

[39] Luca Mesin, Emiliano Robert, Gennaro Boccia, and Taian Martins Vieira. «Investigation of Motor Units Activity: Comparison of Single Channel Surface EMG Deconvolution and Blind Source Separation of Multichannel Data». In: *IEEE Access* 12 (2024), pp. 45456–45468. DOI: 10.1109/ACCESS.2024.3380005 (cit. on pp. 24, 49).

[40] Andrea Marinelli, Nicolò Boccardo, Jacopo Tessadori, Dario Di Domenico, Marianna Semprini, and Lorenzo De Michieli. «Active upper limb prostheses: a review on current state and upcoming breakthroughs». In: *Progress in Biomedical Engineering* 5.1 (2023), p. 012001. DOI: 10.1088/2516-1091/acb5d6 (cit. on pp. 25, 26, 36).

[41] Nana Wang, Jianwei Niu, Xuefeng Liu, Dongqin Yu, Guogang Zhu, Xinghao Wu, Mingliang Xu, and Hao Su. «BeyondVision: An EMG-driven Micro Hand Gesture Recognition Based on Dynamic Segmentation». In: *Proceedings of the Thirty-Third International Joint Conference on Artificial Intelligence (IJCAI-24)*. IJCAI, 2024, pp. 6044–6052. DOI: 10.24963/ijcai.2024/668 (cit. on pp. 25, 27, 30).

[42] Andrea Marinelli et al. «Performance Evaluation of Pattern Recognition Algorithms for Upper Limb Prosthetic Applications». In: *IEEE Transactions on Neural Systems and Rehabilitation Engineering* 31 (2023), pp. 2626–2637. DOI: 10.1109/TNSRE.2023.3283307 (cit. on pp. 26–30, 33).

[43] Camila Montecinos, Jessica Espinoza, Mónica Zamora-Zapata, Viviana Meruane, and Ruben Fernandez. «Improving Fast EMG Classification for Hand Gesture Recognition: A Comprehensive Analysis of Temporal, Spatial, and Algorithm Configurations for Healthy and Post-Stroke Subjects». In: *Sensors* 25.22 (2025), p. 6980. DOI: 10.3390/s25226980 (cit. on pp. 26–29).

[44] T. Prabhavathy, Vinodh Kumar Elumalai, Balaji E, and Dhanasekaran Sandhiya. «A surface electromyography based hand gesture recognition framework leveraging variational mode decomposition technique and deep learning classifier». In: *Biomedical Signal Processing and Control* 88 (2024), p. 105652. DOI: 10.1016/j.bspc.2023.105652 (cit. on pp. 26, 29).

[45] Hiba Hellara, Rim Barioul, Salwa Sahnoun, Ahmed Fakhfakh, and Olfa Kanoun. «Comparative Study of sEMG Feature Evaluation Methods Based on the Hand Gesture Classification Performance». In: *Sensors* 22.22 (2022), p. 8840. DOI: 10.3390/s22228840 (cit. on pp. 26, 27, 66, 67).

[46] Kunkun Zhao, Zhisheng Zhang, Haiying Wen, and Alessandro Scano. «Intra-Subject and Inter-Subject Movement Variability Quantified with Muscle Synergies in Upper-Limb Reaching Movements». In: *Biomimetics* 6.4 (2021), p. 63. DOI: 10.3390/biomimetics6040063 (cit. on p. 26).

[47] Nawadita Parajuli et al. «Real-Time EMG Based Pattern Recognition Control for Hand Prostheses: A Review on Existing Methods, Challenges and Future Implementation». In: *Sensors* 19.20 (2019), p. 4596. DOI: 10.3390/s19204596 (cit. on pp. 27–29).

[48] Andrea Marinelli, Nicolò Boccardo, Lorenzo De Michieli, and Marianna Semprini. «A comparative optimization procedure to evaluate pattern recognition algorithms on Hannes prosthesis». In: *2023 IEEE International Workshop on Metrology for Industry 4.0 & IoT (MetroInd4.0&IoT)*. IEEE, 2023, pp. 364–369. DOI: 10.1109/MetroInd4.0IoT57462.2023.10180145 (cit. on pp. 27, 28).

[49] Erik J. Scheme, Bernard S. Hudgins, and Kevin B. Englehart. «Confidence-Based Rejection for Improved Pattern Recognition Myoelectric Control». In: *IEEE Transactions on Biomedical Engineering* 60.2 (2013), pp. 521–531. DOI: 10.1109/TBME.2012.2226189 (cit. on p. 27).

[50] Mai H. Abdelaziz, Wael A. Mohamed, and Ayman S. Selmy. «Hand Gesture Recognition Based on Electromyography Signals and Deep Learning Techniques». In: *Journal of Advances in Information Technology* 15.2 (2024), pp. 255–263. DOI: 10.12720/jait.15.2.255-263 (cit. on pp. 29, 36).

[51] Reza Bagherian Azhiri, Mohammad Esmaeili, and Mehrdad Nourani. «Real-Time EMG Signal Classification via Recurrent Neural Networks». In: *2021 IEEE International Conference on Bioinformatics and Biomedicine (BIBM)*. IEEE, 2021, pp. 2628–2635. DOI: 10.1109/BIBM52615.2021.9669872 (cit. on p. 31).

[52] Sara Abbaspour, Autumn Naber, Max Ortiz-Catalan, Hamid GholamHosseini, and Maria Lindén. «Real-Time and Offline Evaluation of Myoelectric Pattern Recognition for the Decoding of Hand Movements». In: *Sensors* 21.16 (2021), p. 5677. DOI: 10.3390/s21165677 (cit. on p. 31).

[53] W. Li, P. Shi, and H. Yu. «Gesture Recognition Using Surface Electromyography and Deep Learning for Prostheses Hand: State-of-the-Art, Challenges, and Future». In: *Frontiers in Neuroscience* 15 (2021), p. 621885. DOI: 10.3389/fnins.2021.621885 (cit. on p. 31).

[54] Lauren H. Smith, Levi J. Hargrove, Blair A. Lock, and Todd A. Kuiken. «Determining the optimal window length for pattern recognition-based myoelectric control: balancing the competing effects of classification error and controller delay». In: *IEEE Transactions on Neural Systems and Rehabilitation Engineering* 19.2 (2011), pp. 186–194. DOI: 10.1109/TNSRE.2010.2100828 (cit. on p. 31).

[55] Ali H. Al-Timemy, Guido Bugmann, and Javier Escudero. «Adaptive Windowing Framework for Surface Electromyogram-Based Pattern Recognition System for Transradial Amputees». In: *Sensors* 18.8 (2018), p. 2402. DOI: 10.3390/s18082402 (cit. on p. 31).

[56] John Jairo Villarejo-Mayor, Regina Mamede Costa, Anselmo Frizera-Neto, and Teodiano Freire Bastos. «Dexterous hand gestures recognition based on low-density sEMG signals for upper-limb forearm amputees». In: *Biomedical Signal Processing and Control* 62 (2020), p. 102122. DOI: 10.1016/j.bspc.2020.102122 (cit. on pp. 31, 54).

[57] Jinxin Ao, Shili Liang, Tao Yan, Rui Hou, Zong Zheng, and JongSong Ryu. «Overcoming the effect of muscle fatigue on gesture recognition based on sEMG via generative adversarial networks». In: *Biomedical Signal Processing and Control* 99 (2025), p. 106886. DOI: 10.1016/j.bspc.2024.106886 (cit. on pp. 32, 33).

[58] Roberto Díaz-Amador and Miguel A. Mendoza-Reyes. «Towards the reduction of the effects of muscle fatigue on myoelectric control of upper limb prostheses». In: *Mobile Service Robotics: Proceedings of the 20th International Conference on Climbing and Walking Robots and the Support Technologies for Mobile Machines (CLAWAR 2017)*. Ed. by Karsten Berns, Manuel Felipe Silva, and Mohammad Osman Tokhi. World Scientific, 2017, pp. 675–682. DOI: 10.1142/9789813231047_0085 (cit. on p. 32).

[59] Ran Wang, David H. Fukuda, Jeffrey R. Stout, Edward H. Robinson, Amelia A. Miramonti, Maren S. Fragala, and Jay R. Hoffman. «Evaluation of Electromyographic Frequency Domain Changes during a Three-Minute Maximal Effort Cycling Test». In: *Journal of Strength and Conditioning Research* 29.10 (2015), pp. 2716–2724. DOI: 10.1519/JSC.0000000000000958 (cit. on p. 32).

[60] Jae-Hoon Song, Jin-Woo Jung, and Zeungnam Bien. «Robust EMG Pattern Recognition to Muscular Fatigue Effect for Human-Machine Interaction». In: *IEEE Transactions on Industrial Electronics* 54.6 (2007), pp. 3103–3113. DOI: 10.1109/TIE.2007.907047 (cit. on pp. 32, 33).

[61] Kartik S. Prakash and Nissan Kunju. «An optimized electrode configuration for wrist wearable EMG-based hand gesture recognition using machine learning». In: *Biomedical Signal Processing and Control* 92 (2024), p. 106093. DOI: 10.1016/j.bspc.2024.106093 (cit. on p. 33).

[62] Zhen Zhang, Changxin He, and Kuo Yang. «A Novel Surface Electromyographic Signal-Based Hand Gesture Prediction Using a Recurrent Neural Network». In: *Sensors* 22.6 (2022), p. 2170. DOI: 10.3390/s22062170 (cit. on p. 34).

[63] Jonathan Lundsberg, Anders Björkman, Nebojsa Malesevic, and Christian Antfolk. «Inferring position of motor units from high-density surface EMG». In: *Scientific Reports* 14.1 (2024), p. 3858. DOI: 10.1038/s41598-024-54405-1 (cit. on p. 34).

[64] Alessandro Del Vecchio, Aleš Holobar, Deborah Falla, Francesco Felici, Roger M. Enoka, and Dario Farina. «Tutorial: Analysis of motor unit discharge characteristics from high-density surface EMG signals». In: *Journal of Electromyography and Kinesiology* 53 (2020), p. 102426. DOI: 10.1016/j.jelekin.2020.102426 (cit. on pp. 34, 37–40).

[65] Nicholas Tacca, Collin Dunlap, Sean P. Donegan, James O. Hardin, Eric Meyers, Michael J. Darrow, Samuel Colachis, Andrew Gillman, and David A. Friedenberg. «Wearable high-density EMG sleeve for complex hand gesture classification and continuous joint angle estimation». In: *Scientific Reports* 14.1 (2024), p. 23337. DOI: 10.1038/s41598-024-73898-w (cit. on p. 35).

[66] Jiangcheng Chen, Sheng Bi, George Zhang, and Guangzhong Cao. «High-Density Surface EMG-Based Gesture Recognition Using a 3D Convolutional Neural Network». In: *Sensors* 20.4 (2020), p. 1201. DOI: 10.3390/s20041201 (cit. on p. 35).

[67] Silvia Maria Massa, Daniele Riboni, and Kianoush Nazarpour. «Explainable AI-Powered Graph Neural Networks for HD EMG-Based Gesture Intention Recognition». In: *IEEE Transactions on Consumer Electronics* 69.4 (2023), pp. 1118–1129. DOI: 10.1109/TCE.2023.3283307 (cit. on pp. 35, 36).

[68] Jehan Yang, Kent Shibata, Douglas Weber, and Zackory Erickson. «High-density electromyography for effective gesture-based control of physically assistive mobile manipulators». In: *npj Robotics* 3.1 (2025), p. 2. DOI: 10.1038/s44182-025-00018-3 (cit. on p. 36).

[69] Jonathan Lundsberg. «Computational Methods for High-Density Surface EMG: Decomposition, Localization, and Modelling». Paper IV: Generalizable gesture classification of HDsEMG using volume representations. PhD thesis. Lund, Sweden: Lund University, 2025 (cit. on p. 36).

[70] Donald James Reynolds, Aashin Shazar, and Xiaorong Zhang. «Design and Validation of a Sensor Fault-Tolerant Module for Real-Time High-Density EMG Pattern Recognition». In: *IEEE Sensors Journal* 19.21 (2019), pp. 9896–9904. DOI: 10.1109/JSEN.2019.2928236 (cit. on p. 36).

[71] Ronald S. LeFever and Carlo J. De Luca. «A Procedure for Decomposing the Myoelectric Signal Into Its Constituent Action Potentials—Part I: Technique, Theory, and Implementation». In: *IEEE Transactions on Biomedical Engineering* BME-29.3 (1982), pp. 149–157. DOI: 10.1109/TBME.1982.324881 (cit. on pp. 37, 39).

[72] Marco Barbero, Roberto Merletti, and Alberto Rainoldi. *Atlas of Muscle Innervation Zones: Understanding Surface Electromyography*. Milan, Italy: Springer, 2012. ISBN: 978-88-470-2463-2. DOI: 10.1007/978-88-470-2463-2 (cit. on p. 37).

[73] Maoqi Chen, Aleš Holobar, Xu Zhang, and Ping Zhou. «Progressive FastICA Peel-Off and Convolution Kernel Compensation Demonstrate High Agreement for High Density Surface EMG Decomposition». In: *Neural Plasticity* 2016 (2016), p. 3489540. DOI: 10.1155/2016/3489540 (cit. on p. 37).

[74] Aleš Holobar and Damjan Zazula. «Multichannel Blind Source Separation Using Convolution Kernel Compensation». In: *IEEE Transactions on Signal Processing* 55.9 (2007), pp. 4487–4496. DOI: 10.1109/TSP.2007.896108 (cit. on p. 37).

[75] Aleš Holobar, Marco Alessandro Minetto, Alberto Botter, Francesco Negro, and Dario Farina. «Experimental Analysis of Accuracy in the Identification of Motor Unit Spike Trains From High-Density Surface EMG». In: *IEEE Transactions on Neural Systems and Rehabilitation Engineering* 18.3 (2010), pp. 221–229. DOI: 10.1109/TNSRE.2010.2041593 (cit. on p. 37).

[76] S. Hamid Nawab, Shey-Sheen Chang, and Carlo J. De Luca. «High-yield decomposition of surface EMG signals». In: *Clinical Neurophysiology* 121.10 (2010), pp. 1602–1615. DOI: 10.1016/j.clinph.2009.11.092 (cit. on pp. 37, 39).

[77] Daniela S. de Oliveira et al. «Neural decoding from surface high-density EMG signals: influence of anatomy and synchronization on the number of identified motor units». In: *Journal of Neural Engineering* 19.4 (2022), p. 046029. DOI: 10.1088/1741-2552/ac823d (cit. on pp. 38, 39).

[78] Dario Farina, Corrado Cescon, Francesco Negro, and Roger M. Enoka. «Amplitude Cancellation of Motor-Unit Action Potentials in the Surface Electromyogram Can Be Estimated With Spike-Triggered Averaging». In: *Journal of Neurophysiology* 100.1 (2008), pp. 431–440. DOI: 10.1152/jn.90365.2008 (cit. on p. 39).

[79] Alessandro Sampieri, Gioi Spinello, Martino V. Franchi, Francesco Campa, Antonio Paoli, Tatiana Moro, and Andrea Casolo. «Greater muscle electrode distance and fat mass affect motor units identification from high-density surface EMG in the vastus lateralis muscle». In: *Scientific Reports* 15.1 (2025), p. 41122. DOI: 10.1038/s41598-025-24966-w (cit. on p. 39).

[80] Chuang Lin, Ziwei Cui, Chen Chen, Yanhong Liu, Chen Chen, and Ning Jiang. «A fast gradient convolution kernel compensation method for surface electromyogram decomposition». In: *Biomedical Signal Processing and Control* 79 (2023), p. 104261. DOI: 10.1016/j.bspc.2022.104261 (cit. on p. 39).

[81] Maoqi Chen and Ping Zhou. «A Novel Framework Based on FastICA for High Density Surface EMG Decomposition». In: *IEEE Transactions on Neural Systems and Rehabilitation Engineering* 24.11 (2016), pp. 1177–1187. DOI: 10.1109/TNSRE.2015.2412038 (cit. on p. 39).

[82] Maxime Bourges, Ganesh R. Naik, and Luca Mesin. «Single Channel Surface Electromyogram Deconvolution Is a Useful Pre-Processing for Myoelectric Control». In: *IEEE Transactions on Biomedical Engineering* 69.5 (May 2022), pp. 1767–1774. DOI: 10.1109/TBME.2021.3131650 (cit. on p. 41).

[83] Jie Liu, Xiaoyan Li, Guanglin Li, and Ping Zhou. «EMG Feature Assessment for Myoelectric Pattern Recognition and Channel Selection: A Study with Incomplete Spinal Cord Injury». In: *Medical Engineering & Physics* 36.9 (2014), pp. 1234–1240. DOI: 10.1016/j.medengphy.2014.07.004 (cit. on pp. 64, 66).

[84] Lei Yu and Huan Liu. «Efficient Feature Selection via Analysis of Relevance and Redundancy». In: *Journal of Machine Learning Research* 5 (2004), pp. 1205–1224 (cit. on p. 65).

[85] Evan Campbell, Angkoon Phinyomark, and Erik Scheme. «Current Trends and Confounding Factors in Myoelectric Control: Limb Position and Contraction Intensity». In: *Sensors* 20.6 (2020), p. 1613. DOI: 10.3390/s20061613 (cit. on p. 65).

[86] Yongtao Shi, Yuefeng Zheng, and Xiaotong Bai. «A multiple filter-wrapper feature selection algorithm based on process optimization mechanism for high-dimensional omics data analysis». In: *PLOS ONE* 16.5 (2021), e0251783. DOI: 10.1371/journal.pone.0251783 (cit. on pp. 65, 67).

[87] Sridhar Krishnan and Yashodhan Athavale. «Trends in biomedical signal feature extraction». In: *Biomedical Signal Processing and Control* 43 (2018), pp. 41–63. DOI: 10.1016/j.bspc.2018.02.023 (cit. on p. 65).

[88] Wafaa N. Abdelrazik, Ahmed El-Bialy, Hamed A. Ibrahim, and Bassam A. Hemade. «Enhancing EMG-based gesture recognition for prosthetic control: Impact of dimensionality reduction on machine learning classifier». In: *Sensors and Actuators A: Physical* 381 (2025), p. 117120. DOI: 10.1016/j.sna.2024.117120 (cit. on p. 66).

[89] Angkoon Phinyomark, Franck Quaine, Sylvie Charbonnier, Christine Serviere, Franck Tarpin-Bernard, and Yann Laurillau. «Feature extraction of the first difference of EMG time series for EMG pattern recognition». In: *Computer Methods and Programs in Biomedicine* 117.2 (2014), pp. 247–256. DOI: 10.1016/j.cmpb.2014.06.013 (cit. on p. 66).

[90] Parvatam Ramya Chandrika, Omkar S. Powar, and Krishnan Chemmangat. «Feature Selection and Ranking in EMG Analysis for Hand Movement Classification». In: *2018 IEEE Distributed Computing, VLSI, Electrical Circuits and Robotics (DISCOVER)*. IEEE, 2018, pp. 206–211. DOI: 10.1109/DISCOVER.2018.8674092 (cit. on p. 66).

[91] Angkoon Phinyomark, Pornchai Phukpattaranont, and Chusak Limsakul. «Feature reduction and selection for EMG signal classification». In: *Expert Systems with Applications* 39.8 (2012), pp. 7420–7431. DOI: 10.1016/j.eswa.2012.01.102 (cit. on p. 66).

[92] Shilan S. Hameed, Olutomilayo Olayemi Petinrin, Abdirahman Osman Hashi, and Faisal Saeed. «Filter-Wrapper Combination and Embedded Feature Selection for Gene Expression Data». In: *International Journal of Advances in Soft Computing and its Applications* 10.1 (2018), pp. 90–105 (cit. on pp. 66, 67).

[93] Q. Li, A. Zhang, Z. Li, and Y. Wu. «Improvement of EMG Pattern Recognition Model Performance in Repeated Uses by Combining Feature Selection and Incremental Transfer Learning». In: *Frontiers in Neurorobotics* 15 (2021), p. 699174. DOI: 10.3389/fnbot.2021.699174 (cit. on pp. 66, 67).

[94] R. V. Schulte, E. C. Prinsen, H. J. Hermens, and J. H. Buurke. «Genetic Algorithm for Feature Selection in Lower Limb Pattern Recognition». In: *Frontiers in Robotics and AI* 8 (2021), p. 710806. DOI: 10.3389/frobt.2021.710806 (cit. on pp. 71, 72).

[95] Padmini Sahu, Bikesh Kumar Singh, and Neelamshobha Nirala. «An improved feature selection approach using global best guided Gaussian artificial bee colony for EMG classification». In: *Applied Soft Computing* 146 (2023), p. 110688. DOI: 10.1016/j.asoc.2023.110688 (cit. on pp. 71, 77).

[96] Guopeng Liu, Jianbin Ma, Tongle Hu, and Xiaoying Gao. «A feature selection method with feature ranking using genetic programming». In: *Connection Science* 34.1 (2022), pp. 1146–1168. DOI: 10.1080/09540091.2022.2049702 (cit. on p. 71).

[97] Gavin C. Cawley and Nicola L. C. Talbot. «On Over-fitting in Model Selection and Subsequent Selection Bias in Performance Evaluation». In: *Journal of Machine Learning Research* 11.70 (2010), pp. 2079–2107. URL: https://jmlr.org/papers/v11/cawley10a.html (cit. on p. 76).

[98] Ron Kohavi. «A Study of Cross-Validation and Bootstrap for Accuracy Estimation and Model Selection». In: *Proceedings of the 14th International Joint Conference on Artificial Intelligence (IJCAI)*. Vol. 2. Montreal, Canada: Morgan Kaufmann, 1995, pp. 1137–1143 (cit. on p. 77).

[99] David E. Goldberg. *Genetic Algorithms in Search, Optimization, and Machine Learning*. Reading, MA: Addison-Wesley, 1989. ISBN: 0-201-15767-5 (cit. on pp. 77, 78).

[100] Kenneth A. De Jong. «Are genetic algorithms function optimizers?» In: *Parallel Problem Solving from Nature 2*. Ed. by R. Männer and B. Manderick. Amsterdam: North-Holland, 1992, pp. 3–13 (cit. on p. 77).