



**POLITECNICO DI TORINO**  
FACULTY OF ENGINEERING  
Master of Science in Biomedical Engineering

---

**Evaluation of High-Density EMG Feature  
Extraction and Selection to Recognize Lower Limbs  
Movements for a Rehabilitation Exoskeleton**

Supervisors:

**Prof. Marco Gazzoni**

**Prof. Dario Farina**

Candidate:  
**Simone De Giorgi**



**Imperial College  
London**

---

Academic Year 2017 – 2018

---

*...Nino non aver paura di sbagliare un calcio di rigore,  
non è mica da questi particolari che si giudica un giocatore...*

---

*Ai miei Genitori Salvatore e Annunziata,  
a mia Sorella Sara,  
ai miei cari Nonni Gaetano, Cosimo, Rosaria e Concetta  
per tutto quello che mi trasmettete.*

---

## Abstract

In lower limb rehabilitation, the role of the physiotherapist is fundamental, but a lot of studies demonstrate that robotic interventions in rehabilitation and patient's involvement in voluntary movement improve recovery at a faster rate. In this respect, the use of exoskeletons controlled by residual muscle activity can help the patient to reach better results. Many works exist in literature about the myoelectric control of external devices such as prostheses, orthoses and exoskeletons. Myoelectric control is usually performed using standard bipolar EMG. In recent years, it has been demonstrated that sEMG recorded with bipolar EMG is affected by many factors depending on central and peripheral factors; among them the geometrical changes of the muscle during dynamic tasks. In order to collect a more robust signal, representative of the muscle activity, high density sEMG has been proposed. This technique, by detecting EMG from the whole muscle surface, is more robust than bipolar technique. The main aim of this work was to identify the best feature extracted from HD-sEMG signals recorded from lower limb muscles that allows the recognition of daily activities. Nine healthy and voluntary subjects were involved in the experiments. The movements considered in the experimental protocol were: Stand to sit, sit to stand, stair ascending, stair descending, a gait cycle, rest in upright position and rest in a sitting position. Monopolar signals were recorded from 7 muscles using four matrices of 32 electrodes placed on Rectus Femoris, Vastus Medialis, Vastus Lateralis and Tibialis Anterior and two matrices of 64 electrodes placed on Gastrocnemius Medialis and Semitendinosus & Biceps Femoris. Single differential EMG signals were calculated from monopolar signals and used in the analysis. Fourteen features were extracted from epochs 300ms long with overlap 50%, allowing also its calculation for possible real time applications. For the selection of the best feature subset both the computational time and the filter approach, based on information about class separability (J index), were used. The wrapper approach based on Linear Discriminant Analysis (LDA) was then applied in order to compare and to validate the results of the J index with the performance of the LDA classifier. Results show that the two methods are correlated (*Spearman's rho=0.984 and p-value<0.01*). Features calculated in frequency domain, Mean Frequency (MNF) and Median Frequency (MDF), were excluded on the basis of the computational time criterion. Among the other features, Waveform Length(WL) has the best overall performance with the highest J value (*0.061*), the highest accuracy (*98%*) and a low computational time (*9.4ms*). Furthermore, the Integrated EMG (*J=0.0583*), the Mean of Absolute Value (*J=0.0583*) and the Difference Absolute Standard Deviation (*J=0.0582*) are only 5% less performing and can be taken into account for a feature vector creation. In conclusion, the best HD-sEMG feature identified on the basis of a class separability analysis showed also the best performances in the discrimination between daily live tasks. Three other

---

features performed in a similar way. These four features used separately allow to obtain a classification accuracy greater than 96%.

---

# Contents

<b>1</b>	<b>Introduction</b>	<b>12</b>
1.1	Muscle physiology . . . . .	14
1.1.1	Skeletal Muscle Structure . . . . .	14
1.1.2	The Sliding Filament Theory . . . . .	15
1.1.3	Nerve structure . . . . .	16
1.1.4	Motor control mechanisms . . . . .	17
1.1.5	Motor unit (MU) . . . . .	18
1.2	Lower limb muscles . . . . .	19
1.2.1	Vastus Medialis . . . . .	20
1.2.2	Vastus Lateralis . . . . .	20
1.2.3	Rectus Femoris . . . . .	20
1.2.4	Biceps Femoris and Semitendinosus . . . . .	21
1.2.5	Gastrocnemius . . . . .	22
1.2.6	Tibialis Anterior . . . . .	22
1.3	EMG signal . . . . .	23
1.3.1	Motor Unit Action Potential (MUAP) . . . . .	23
1.3.2	Mathematical model of EMG signal . . . . .	25
1.4	Noise sources in EMG signal . . . . .	25
1.4.1	Inherent Noise in the Electrode . . . . .	26
1.4.2	Inherent Instability of the Signal . . . . .	26
1.4.3	Movement Artifact . . . . .	26
1.4.4	Electromagnetic Noise . . . . .	27
1.4.5	Cross Talk . . . . .	27
1.4.6	Internal Noise . . . . .	27
1.5	Emg recording techniques . . . . .	28
1.6	Surface EMG (sEMG) . . . . .	28
1.6.1	Electrodes . . . . .	28
1.6.2	Electrode types . . . . .	28
1.6.3	Electrode shape . . . . .	29
1.6.4	Inter-Electrode Distance (IED) . . . . .	30
1.6.5	Electrodes Placement . . . . .	30

---

<b>2</b>	<b>Materials and Methods</b>	<b>32</b>
2.1	High Density sEMG choice . . . . .	32
2.2	Experimental Set Up . . . . .	33
2.2.1	Muscle Identification following SENIAM recommandations . . . . .	33
2.2.2	Muscles and Matrix correspondence . . . . .	37
2.3	Experimental Protocol . . . . .	38
2.4	Subject reclutation . . . . .	39
2.5	Data acquisition . . . . .	39
2.6	Data Processing . . . . .	41
2.6.1	Filtering Step . . . . .	41
2.6.2	Channels remapping and Single differential calculation . . . . .	42
2.7	Bad channels selection . . . . .	43
2.8	Pattern Recognition . . . . .	44
2.8.1	Feature Extraction . . . . .	45
2.8.2	Computational Time . . . . .	49
2.8.3	Feature Evaluation . . . . .	50
2.9	Linear Discriminant Analysis (LDA) Notions . . . . .	50
2.10	Filter approach - Class separability J Index for feature evaluation . . . . .	54
2.11	Wrapper approach - Classification using LDA Classifier . . . . .	55
2.11.1	Training set and Test set separation . . . . .	56
2.11.2	Linear Classificarion rule . . . . .	56
2.12	Spearman's rank correlation coefficient . . . . .	56
<b>3</b>	<b>Results</b>	<b>58</b>
3.1	Single Differential EMG Signals . . . . .	58
3.2	Computational Time . . . . .	64
3.3	J index performance . . . . .	66
3.4	LDA Classification performance . . . . .	69
3.5	Comparisons: Classification - J index . . . . .	73
<b>4</b>	<b>Conclusions</b>	<b>78</b>
4.1	Future work . . . . .	79
<b>5</b>	<b>Anknowldgment</b>	<b>85</b>

---

**List of Figures**

1	Skeletal Muscle Structure Taken from [1] . . . . .	14
2	Detailed Structure of one Sacromere. Taken from [2] . . . . .	15
3	The Sliding Filament theory cycle. Taken from [3] . . . . .	16
4	Neuron anatomy.Taken from [4] . . . . .	17
5	(A) Focus on the neuromuscular junction; (B) Motor control mechanism scheme. Taken from [5] . . . . .	18
6	(MU) Motor unit organization.Taken from [6] . . . . .	19
7	Muscle activity when action potential comes and the differnt phases of the muscle activity. Taken from [7] and modified. . . . .	19
8	Muscle of anterior thigh: Vastus medialis, Vastus Lateralis, Rectus Femoris.Taken from [8] . . . . .	21
9	Posterior muscles of thigh included Biceps Femoris and Semitendinosus.Taken from [9] . . . . .	21
10	Gastrocnemius muscle.Taken from [10] . . . . .	22
11	Muscles of the anterior leg included Tibialis Anterior.Taken from [11] . . . . .	22
12	Representation of a motor unit (MU) and of a motor unit action potential (MUAP). Zoomed view of the source. Taken from [12] . . . . .	24
13	Schematic representation of the generation of the motor unit action potential. Taken from [13] . . . . .	24

14	Models of the electrode–skin impedance: (a) simplified electrical model of the electrode–gel interface: $E_{hc}$ is the half-cell potential at the metal–electrolyte junction, the parallel $R_pC_p$ takes into account the polarizability and the capacitive behavior of the junction, $R_s$ describes the resistive behavior of the electrolyte gel, $V_{noise}$ is the associated noise component; (b) generalized model of the electrode–skin interface. The electrode–electrolyte junction described in (a) is expanded in order to take into account the effect of the conductive gel ( $R_p$ , $C_p$ , and $R_s$ ) and the electrolyte–skin interface. $E_{epi}$ is the half-cell potential due to differences in the ionic concentrations between the gel and the superficial layer of the skin, the parallel $R_{epi}C_{epi}$ characterizes the skin impedance, and $R_{sub}$ is the resistive component associated to the subcutaneous tissue layer. Equivalent noise generators are not indicated for simplicity but are present at all interfaces. Taken from [14] . . . . .	31
15	Quattrocento amplifier for the acquisition of surface/intramuscular monopolar or bipolar EMG signals and technical specifications. Taken from [15] . . . . .	33
16	Vatus Medialis muscle identification. Taken from [16] . . . . .	34
17	Vastus Lateralis muscle identification. Taken from [16] . . . . .	34
18	Tibialis Anterior muscle identification. Taken from [16] . . . . .	35
19	Rectus Femoris muscle identification. Taken from [16] . . . . .	35
20	Gastrocnemius medialis muscle identification. Taken from [16] . . . . .	36
21	a)Semitendinosus and b)Biceps Femiris muscle identification. Taken from [16]	36
22	a) Matrices of 32 electrodes positioned on Vastus Medialis, Vastus Lateralis, Tibialis Anterior, Rectus Femoris; b)Matrices of 64 electrodes positioned on Gastrocnemius Medialis and Biceps Femoris&Semitendinosus . . . . .	37
23	a) Matrix of 64 electrodes. b) Matrix of 32 electrodes . . . . .	38
24	GUI realized in MATLAB for visualization and EMG recording . . . . .	39
25	GUI realized in MATLAB for visualization and EMG recording . . . . .	40
26	Mask of band-pass filter . . . . .	42
27	Mask of band-stop filter . . . . .	42
28	a) channels sequence during acquisition. b) directions of single differential calculation on tibialis anterior. c) directions of single differential calculation on vastus lateralis,vastus medialis and rectus femoris. . . . .	43

29	a) channels sequence during acquisition using matrices of 64 electrodes. b) directions of single differential calculation on Gastrocnemius medialis and Biceps Femoris&Semitendinosus . . . . .	43
30	Bad channel selection GUI . . . . .	44
31	Segmentation with 50% overlap. The difference between segment and increment correspondes to overlap. Increment is the time to wait going from one segment to the next one. . . . .	46
32	Schematic steps to calculate a lower dimensional subspace in LDA technique. Taken from [17] . . . . .	51
33	Projection of the original samples on the lower dimensional space of LDA. Taken from [17] . . . . .	52
34	On the left positive correlation, in the middle negative correlation, on the right no significative correlation . . . . .	57
35	A) SD HD-EMG signals recorded from Semitendinosus and Biceps Femoris muscles. A matrix of 64 electrodes was used. A.1) expanded view of 200ms of signal in which the different muscle activities between the two muscles are evident. The tasks duration is 1.8s. . . . .	59
36	A) SD HD-EMG signals recorded from Gastrocnemius medialis muscle. A matrix of 64 electrodes was used. A.1) expanded view of 150ms of signal in which the different activation areas can be observed. In red the most activated one. The tasks duration is 1.8s. . . . .	59
37	A) SD HD-EMG signals recorded from Vastus medialis muscle. A matrix of 32 electrodes was used. A.1) expanded view of 180ms of signal in which the different activation areas can be observed. In red the most activated one. The tasks duration is 1.8s . . . . .	60
38	SD HD-EMG signals recorded from Gastrocnemius medialis muscle during two tasks: A) stair descending and B) stand to sit. A matrix of 32 electrodes was used. A.1) expanded view of 400ms of signal during the stair descending task. B.1) expanded view of 200ms of signal during the stand to sit task. In A.1) and B.1) the different activation areas can be observed. In red the most activated ones. The tasks duration is 1.8s. . . . .	61

39	SD HD-EMG signals recorded from Rectus Femoris muscle during two tasks: A) gait cycle and B) stand to sit. A matrix of 32 electrodes was used. A.1) expanded view of 200ms of signal during the gait cycle task. B.1) expanded view of 200ms of signal during the stand to sit task. In A.1) and B.1) the different activation areas can be observed. In red the most activated ones. The tasks duration is 1.8s. [0.3cm] . . . . .	62
40	SD HD-EMG signals recorded from Tibialis Anterior muscle during six tasks: A) stand to sit, B) stair ascending, C) gait cycle, D) star descending, E) sit to stand, F) rest sitting. A matrix of 32 electrodes is used. The tasks duration is 1.8s . . . . .	63
41	Bar plot about the computational time of all 14 features. In green both features in frequency domain with an high computational cost. . . . .	65
42	Bar plot about the computational time of all features except those in frequency domain. . . . .	65
43	Boxplot about J index of 12 features from one subject. . . . .	66
44	Average value of J index of 12 features from one subject (Error bar: standard deviation on fifteen repetitions). . . . .	67
45	Boxplot about J index of 12 features from one subject. . . . .	67
46	Average value of J index of 12 features from one subject (Error bar: standard deviation on fifteen repetitions). . . . .	68
47	Average J index behavior about 12 features in 9 subjects . . . . .	68
48	Boxplot about classification of 12 features from one subject. . . . .	70
49	Average accuracy about the 12 features from one subject (Error bar: Standard deviation on fifteen repetitions tested). . . . .	71
50	Boxplot about classification of 12 features from one subject. . . . .	71
51	Average accuracy about the 12 features from one subject (Error bar: Standard deviation on fifteen repetitions tested . . . . .	72
52	Average Accuracy behavior about 12 features in 9 subjects . . . . .	72
53	Trend between average classification accuracy (X axis) and average J index(Y axis) among repetitions of Subject 2 in Table 3 . . . . .	74
54	Trend between average classification accuracy (X axis) and average J index(Y axis) among repetitions of Subject 8 in Table 3 . . . . .	74

---

55	On the left there is the average J index and on the right the average classification accuracy among subjects. (Error bar: standard deviation) . . . . .	75
56	Trend between average classification accuracy (X axis) among subjects and average J index(Y axis) among subject. . . . .	76
57	Feature average J index ranking among subjects . . . . .	76
58	Feature average accuracy ranking among subjects . . . . .	77

## List of Tables

1	Computational time . . . . .	64
2	Average Accuracy on fifteen repetitions. . . . .	70
3	Spearman's rank correlation coefficient and p-value between average J index and average accuracy per each subject. Null Hypotesis of no correlation. . . .	73

## 1 Introduction

Nowadays Stroke is one of the main causes of serious long-term disability as well as morbidity and mortality [20]. Unfortunately, only in Europe, people affected by Stroke are about one million [21] and in U.S.A. alone, approximately 800.000 people have a Stroke episode every year. The majority of victims suffers from neurological and sensory-motor deficits, so they need a period of rehabilitation in order to achieve functional independence. Gait recovery, with the final aim of walking to perform daily activities, is the major objective in the rehabilitation program for stroke patients. In this respect, hemiparesis is one of the most common manifestation of Stroke. In order to cope with that problem task-oriented repetitive movements can improve movement coordination and muscular strength in patients. Manual assistance of therapists during the daily life activities plays a very important role in this rehabilitation field. Many reviews in the last years considered traditional gait training, in combination with high-technology rehabilitation approaches [22], evaluating in this way the benefits of this. Infact, the clinician can obtain highly controlled repetitive and intensive training by means of robotic rehabilitation overcoming the burden of clinical staff. The usage of robotic interventions in training tasks is expected to improve the recovery at a faster rate, and to resume daily activities sooner; this is the reason why different kind of technologies have been designed. Body weight support with treadmill training machines, such as Lokomat [23], LOPES [24] and ALEX [25] have been used during gait training but it is known that patient's involvement in voluntary movement is important to get better results in rehabilitation [26]. Furthermore, under physiological conditions, such as overground walking, cognitive engagement can be increased as well as sensory inputs and central neuronal circuits become to be activated reaching a massive neural regeneration effects. For this reason, ambulatory exoskeletons with control strategies based on Emg signals such as H2 [27] can be more effective to reinstate neuroplasticity and improve motor functions. In this project, in order to recognize which daily living movement (among 7) the patient wanted to do, HD-EMG (High-Density-Electromyography) is used in combination with an evaluation of feature extraction and selection to find the best feature subset. It is more robust than the bipolar technique usually used in myoelectric control applications, and it allows to take into account the whole muscle surface. In general feature selection can be performed following two different ways: filter approach based on a statistical criterion and the wrapper approach based on the classification accuracy. The former, compared to the latter has a simpler implementation and

the computational cost is neatly lower. In this study it is demonstrated, using the Spearman's rank correlation coefficient, that both methods are equivalent because they get the same feature ranking, which means that when the J index (for the filter approach) increases, also the LDA (for the wrapper approach) classification accuracy increases. For this reason, in this work, Fisher linear discriminant index (J index) is used to find the best subset of features among the fourteen ones studied, even if two of them are immediately eliminated because of the high computational time that doesn't allow eventual real time applications. All the features evaluated were found in literature among different papers and they are in both frequency and time domain. The relevance of this work is that the majority knowledge of feature selection is about upper limbs and few works are referred to lower limbs. Furthermore, the most relevant is the fact that there is no evident researches that applies these techniques on High-Density EMG.

Background knowledge about muscle physiology, lower limb muscles used in this project and EMG fundamentals and artifacts are condensed in the next subchapters.

## 1.1 Muscle physiology

In the human body there are three different muscle types which differ in their forms and functions: skeletal, cardiac and smooth. Cardiac muscle fibers, placed in the heart and the bottom of large veins, as well as smooth muscle cells are managed through the autonomic (sympathetic and parasympathetic) nervous system. In contrast, skeletal muscle fibers are controlled by the somatic nervous system.

### 1.1.1 Skeletal Muscle Structure

Skeletal muscle forms the vast majority of our muscle tissue. It constitutes the 40% of the body weight and contains 50% – 75% of the body proteins. Skeletal muscle has multiple functions.

It is a key component of the locomotor system. It has the ability to convert the chemical energy into mechanical energy therefore generating movement. In this way, a subject can be independent and able to carry out daily activities. Skeletal muscle is also a reserve of carbohydrates and amino acids for different tissues such as skin, heart and brain.

Skeletal muscle contributes to the maintenance of blood glucose levels during periods of starvation. The secondary disability to impaired skeletal muscle is particularly debilitating. The good maintenance of skeletal muscle health is crucial for the prevention of various diseases and the psychological stress due to the disability. The skeletal muscles' architecture could be described as a bundle of muscle fibers held all together through associated connective tissue called fascia (collagen fibers).

Muscle fibers are multinucleated, post-mitotic cylinders which means that they consist of many myofibrils composed by 'Actin' and 'Myosin' proteins and can't change in number, but just in size, and are fully developed by birth, as Frontera's and Ochala's skeletal muscle review [2] explains (Figure 1):

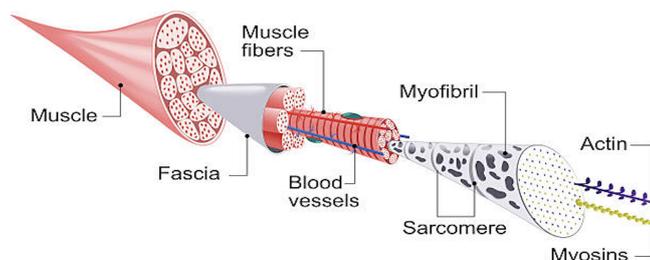


Figure 1: Skeletal Muscle Structure Taken from [1]

Myofibrils consist of very fine contractile fibres, which extend in group in parallel columns along the length of striated muscle fibres. In each muscle fibre there are hundreds of myofibrils. Myofibrils consist of sectional sarcomeres ordered in series and separated each other by Z lines. The myofibrils are made up of two types of protein filaments which give the muscle its striped appearance. The thick filaments are composed by myosin, and the thin filaments are composed by actin, along with two other muscle proteins, tropomyosin and troponin. Every sarcomere is characterized by different bands and zones as showed in (Figure 2);

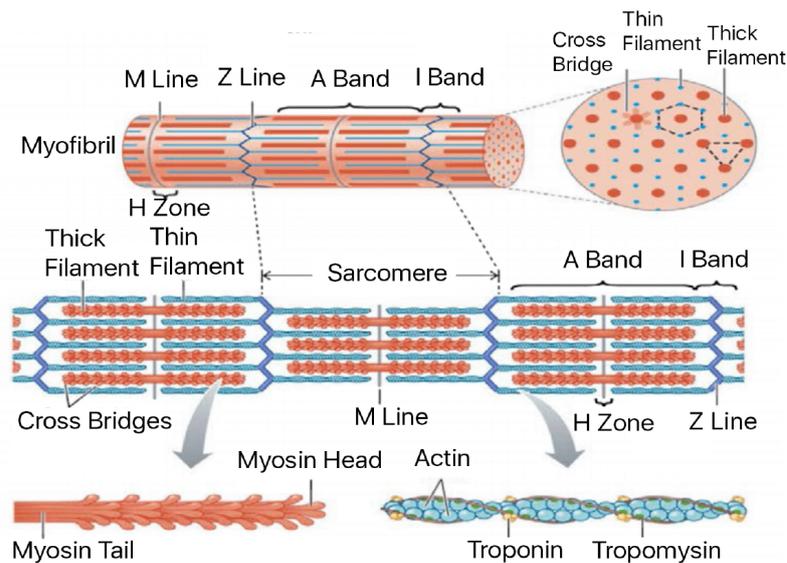


Figure 2: Detailed Structure of one Sarcomere. Taken from [2]

The 'I Band' indicates the region where 'Thin Filaments' are not superimposed by 'Thick Filaments'. The 'A Band' indicates the interacting region between 'Actin' and 'Myosin', the 'M Line' indicates the centre of this region. The 'H Zone' only contains thick filaments which are not covered by thin filaments. 'Cross Bridges' indicate zones of attachments of myosin with actin. The inner structure of myosin, with its 'Myosin Head' and 'Myosin Tail' as well as actin with its two component proteins 'Troponin' and 'Tropomyosin' are shown. Muscular contraction is caused by the interaction between actin and myosin as they temporarily bind to each other and are released.

### 1.1.2 The Sliding Filament Theory

The sliding filament theory explains contraction and relaxation behaviour of muscles by means of pulling and relaxation of actin and myosin filaments. Before muscle contraction happens, the Sarcoplasmic Reticulum (SR) has to release calcium ions into the cytosol, the

calcium ions will bind to troponin which causes the troponin-tropomyosin complexes to move away from the myosin binding sites on actin. Once this happens contraction can begin. The process is similar to rowing, each myosin head rows forward using the actin as a base to pull from. (Figure 3);

- Myosin heads contain a site for ATP binding and ATPase enzymes. ATP is hydrolyzed into ADP freeing up a Phosphate molecule, this energizes the myosin head.
- Myosin attaches to the myosin binding site on actin and releases the free phosphate group. This forms the cross-bridge between the actin and myosin threads.
- Phosphate release by myosin triggers the actual power contraction stroke. ADP pocket opens up on the myosin head. Myosin head rotates and releases the ADP. As the head rotates it moves in a ratchet like motion and pulls the thin filaments closer to the center of the M line of the sarcomere.
- When the ratchet motion is over the myosin awaits another ATP so it can be released from the actin binding site. If an ATP is available the myosin detaches and begins another stroke of the cycle. If not the Myosin remains locked in place and the muscle remains locked up as well this is called “rigor-mortis” .

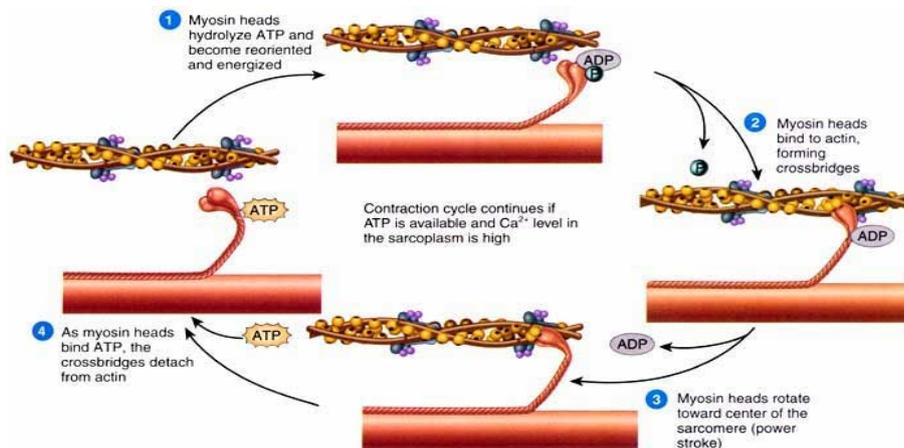


Figure 3: The Sliding Filament theory cycle. Taken from [3]

### 1.1.3 Nerve structure

Each neuron is made by a cell body and by an axon. The former contains a nucleus which is the centre of operation for the neuron and dendrites. The latter carries the electrical impulses away from the cell body and towards the muscle. The final part of an axon is known

as axon terminals and end at synaptic knobs which have contact with the muscle. The axon is surrounded by a Myelin sheath which acts to insulate the nerve. That sheath is not continuous but has breaks called nodes of Ranvier. The impulse jumps from one node to another, allowing a faster conduction speed. (Figure 4)

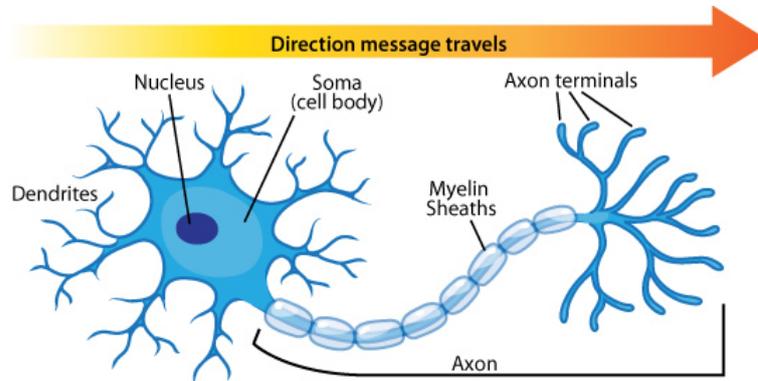


Figure 4: Neuron anatomy. Taken from [4]

#### 1.1.4 Motor control mechanisms

Motor programming takes place in the premotor cortex, the supplementary motor area, and other associated areas of the cortex. Primary motor cortex is organized as a contralateral motor map: this means that the right side of the body is mapped on left hemisphere, while the left one is mapped on the right hemisphere. Inside motor cortex, the pyramidal tract spreads impulses (in the form of nerve impulses) that control the execution of determined voluntary movements. These impulses move from motor cortex to the lower motorneurons, located into the ventral horn of the spinal cord (which is the lower part of the central nervous system). Spinal cord gives rise to the peripheral nerves that reach out into all areas of our bodies, also muscles, where a Neuromuscular Junction is formed (Figure 5). Inputs from these areas, from the cerebellum and, to some extent, from the basal ganglia converge to the primary motor cortex and excite or inhibit the various neurons of the primary motor cortex. The outputs from the primary motor cortex have a powerful influence on interneurons and motoneurons of the brain stem and of the spinal cord. There exists a link between the corticospinal tract and alpha ( $\alpha$ )-motoneurons, providing direct cortical control of muscle activity

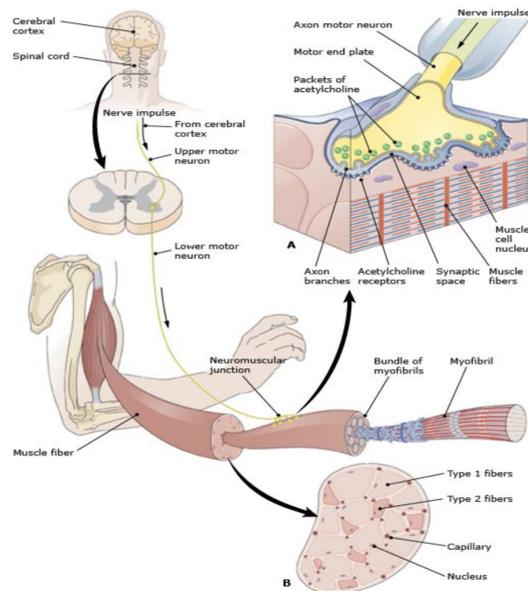


Figure 5: (A) Focus on the neuromuscular junction; (B) Motor control mechanism scheme. Taken from [5]

### 1.1.5 Motor unit (MU)

A motor unit (MU) is composed by a single motoneuron (alpha motoneuron in the spinal cord) and all of the muscle fibres it innervates. A motor unit can contain from tens to thousands of muscle fibres. When muscles produce large powerful movements, they contain motor units with a large numbers of fibres, and for small intricate movements they contain only a few fibres per motor unit. The synaptic knobs of the neuron meet the muscle fibres in correspondence of the neuromuscular junction. When an impulse reaches the neuromuscular junction, the Acetylcholine neurotransmitter is released and it filters across the synaptic cleft (Figure 5A). After that the muscular contraction happens by means of the sliding filament theory (Figure 6).

A single muscle stimuli which comes from the brain to the skeletal muscles is called 'twitch'. It consists of three different phases: 'Latent Period', 'Contraction Phase' and 'Relaxation Phase'. The Latent Period is the shortest time interval between an Action Potential and the muscle excitation due to impulse travel time through nerves or chemical events in muscle fibers. In the Contraction Phase the muscle shortens and produces tension. During the Relaxation Phase the muscle goes back to its original length.

The mechanical response to repeated stimulation depends on the rate of the stimulation. A muscle, like other excitable tissues, has a period following its action potential during which

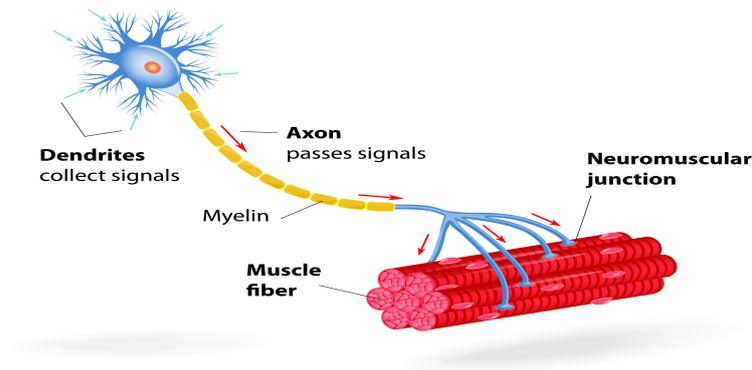


Figure 6: (MU) Motor unit organization. Taken from [6]

the membrane will not respond to stimulation regardless of the strength. This period is called refractory period. Therefore, a second pulse within that interval will not cause any response. Each stimulus causes a twitch. If stimuli are delivered slowly enough, the tension in the muscle will relax between successive twitches. If stimuli are delivered at high frequency, the twitches will overlap, resulting in tetanic contraction. (Figure 7)

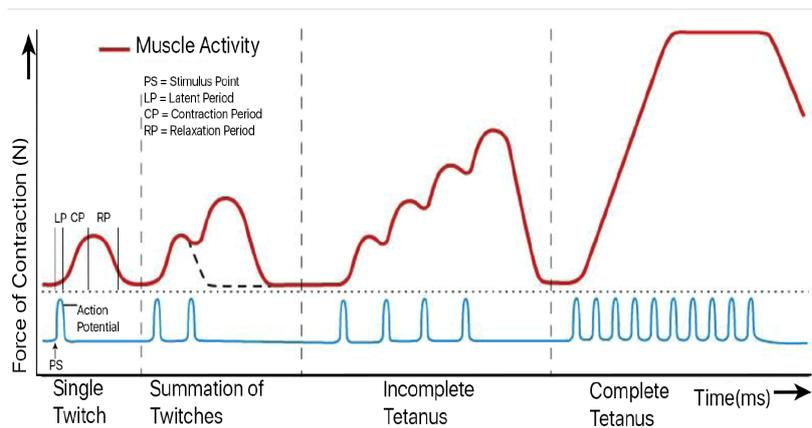


Figure 7: Muscle activity when action potential comes and the different phases of the muscle activity. Taken from [7] and modified.

## 1.2 Lower limb muscles

The muscles of lower limb include those of the thigh, leg and foot; in this work we take into account seven different muscles: five of them from the thigh and two from the leg. Respectively *Vastus Medialis*, *Vastus Lateralis*, *Rectus Femoris*, *Biceps Femoris*, *Semitendinosus*, *Gastrocnemius medialis* and *Tibialis Anterior*.

### **1.2.1 Vastus Medialis**

Vastus medialis is part of Quadriceps positioned on the medial side of the anterior thigh. This muscle originates from the intertrochanteric line and medial lip of linea aspera. It is involved in knee extension and patella stabilization. (Figure 8)

### **1.2.2 Vastus Lateralis**

Vastus lateralis is an anterior muscle of the thigh and also the largest of quadriceps femoris group. It is on the side of the thigh and originates from different parts of the femur as the upper intertrochanteric line, lower border greater trochanter, lateral side gluteal tuberosity, upper half lateral lip of linea aspera and lateral intermuscular septum. The main functionality of this muscle is the knee extension and the patella stabilization. (Figure 8)

### **1.2.3 Rectus Femoris**

Rectus Femoris is an anterior muscle of the thigh and also part of quadriceps femoris group. It originates from the ilium superiorly to the acetabulum and it flexes the thigh at the hip point and extends at the knee joint. (Figure 8)

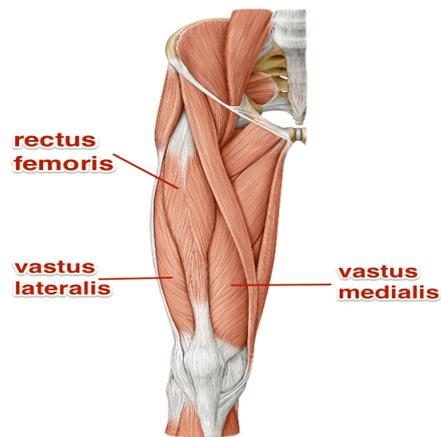


Figure 8: Muscle of anterior thigh: Vastus medialis, Vastus Lateralis, Rectus Femoris. Taken from [8]

#### 1.2.4 Biceps Femoris and Semitendinosus

Biceps Femoris is a muscle located in the the posterior side of the thigh. As its name says, it has two heads of origin: the long heads originates from the inner and lower impression on the back part of the tuberosity of the ischium, and the short head arises from the lateral lip of the linea aspera. Both heads of the biceps femoris are involved in knee flexion. (Figure 9)

Also Semitendinosus is a muscle located in the the posterior side of the thigh between Semimembranosus and biceps Femoris. It is involved in extension of the thigh at the hip, flexion of the leg at the knee and internal rotation of the knee when it is flexed. (Figure 9)



Figure 9: Posterior muscles of thigh included Biceps Femoris and Semitendinosus. Taken from [9]

### 1.2.5 Gastrocnemius

Gastrocnemius is the most superficial muscle of all the muscles in the back portion of the leg. It is made of two muscular regions, the medial head and lateral head, which attach to the medial and lateral sides of the femur. Both the heads plantarflex the foot at the ankle and flex the leg at the knee. (Figure 10)

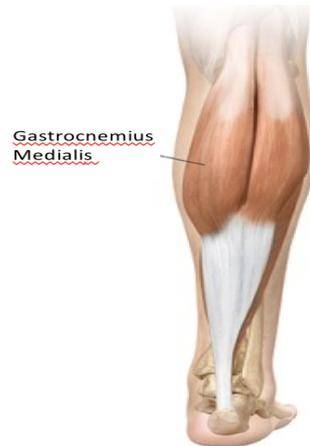


Figure 10: Gastrocnemius muscle. Taken from [10]

### 1.2.6 Tibialis Anterior

The tibialis anterior muscle is a long, narrow muscle in the anterior compartment of the lower leg. It originates from the lateral condyle and proximal body of the tibia and runs down the shin laterally to the tibia. It performs dorsiflexion and inversion of the ankle. (Figure 11)

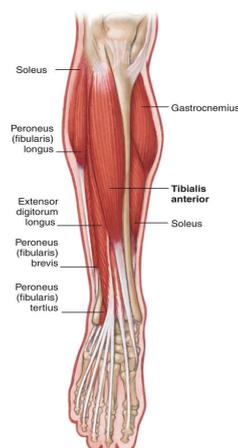


Figure 11: Muscles of the anterior leg included Tibialis Anterior. Taken from [11]

### 1.3 EMG signal

EMG signal is a signal contains information of the collective electric signal from muscles, which is controlled by the nervous system and produced during muscle contraction. The signal represents the anatomical and physiological properties of muscles.

#### 1.3.1 Motor Unit Action Potential (MUAP)

An action potential propagating down a motoneuron activates all the branches of the motoneuron; these in turn activate all the muscle fibers of a motor unit. When the postsynaptic membrane of a muscle fiber is depolarized, the depolarization propagates in both directions (Figure 12) along the fiber. The membrane depolarization, by means of a movement of ions, generates an electromagnetic field around the muscle fibers. The potential or voltage can be detect using an electrode located around that field, whose time excursion is known as *Muscle Fiber Action Potential (MFAP)*.

A schematic representation of this situation is shown in (Figure 13). In that representation the  $n$  represents the total number of muscle fibers of one motor unit whose action potential is detected by the electrode. For simplicity, only the muscle fibers from one motor unit are depicted. On the right side of (Figure 13). there are the action potentials of the respective muscle fibers. The individual muscle fiber action potentials represent the contribution that each active muscle fiber makes to the signal detected at the electrode site.

The depolarizations of the muscle fibers of one motor unit overlap in time and the resultant signal present at the detection site will constitute a spatial-temporal superposition of the contributions of the individual action potentials. The resultant signal is known as *Motor Unit Action Potential (MUAP)* and will be designated as  $h(t)$  in (Figure 13) and described by Equation 1.

In normal muscle, the peak-to-peak amplitude of a MUAP detected with indwelling electrodes (needle or wire) may range from a few microvolts to 5 mV, with a typical value of 500  $\mu\text{V}$  and the time duration of MUAPs may also vary greatly, ranging from less than 1 to 13 ms [13]. In addition, to sustain a muscle contraction, the motor units must be repeatedly activated. The resulting sequence of MUAPs is called *Motor Unit Action Potential Train (MUAPT)* and it is explained by Equation 2. [13].

EMG signal (3) is simply the algebraic summation of the generated MUAPTs and background interference.

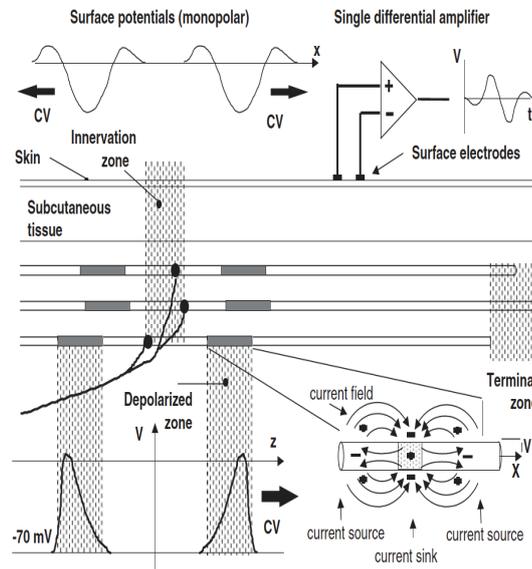


Figure 12: Representation of a motor unit (MU) and of a motor unit action potential (MUAP). Zoomed view of the source. Taken from [12]

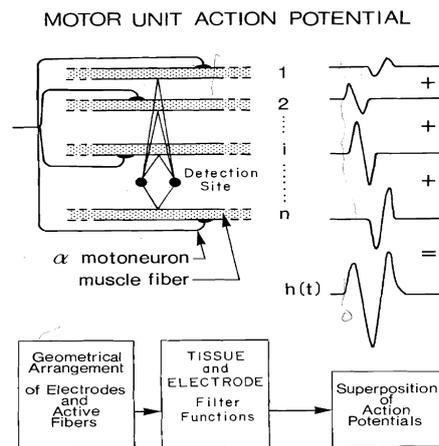


Figure 13: Schematic representation of the generation of the motor unit action potential. Taken from [13]

### 1.3.2 Mathematical model of EMG signal

To demonstrate in a mathematical way what explained in the previous section, it is possible to start from the definition of a Muscle Fiber Action Potential (*MFAP*), going through the definitions of a Motor Unit Action Potential (*MUAP*) and Motor Unit Action Potential Train (*MUAPT*) to arrive at the final mathematical formulation of the (*EMG*) signal.

In particular way :

$$MUAP_j(t) = \sum_{i=1}^n MFAP_i(t - \tau_i) s_i \quad (1)$$

where  $n$  is the number of muscle fibers in the  $j^{th}$  MU of a contracting muscle,  $\tau_i$  is the temporal delay of  $MFAP_i(t)$  and  $s_i$  is a binary variable which is "1" if fiber  $i$  fires and is "0" if it does not fire.

$$MUAPT_j(t) = \sum_{i=1}^M MUAP_{ji}(t - \delta_{ji}) \quad (2)$$

where  $M$  is the number of times that the  $j^{th}$  MU fires,  $\delta_{ji}$  is the  $i^{th}$  firing time of  $MU_j$  and  $MUAP_{ji}(t)$  is the  $i^{th}$  MUAP generated by  $MU_j$  during its  $i^{th}$  firing.

$$EMG(t) = \sum_{J=1}^N MUAPT_j(t) + n(t) \quad (3)$$

where  $N$  is the number of active MUs,  $MUAPT_j(t)$  is the MUAP generated by the  $j^{th}$  MU, and  $n(t)$  is background noise.

### 1.4 Noise sources in EMG signal

EMG signal recorded from every muscle is mixed of various noise signals or artifacts. The attributes of the EMG signal depend on the internal structure of the subject, including the individual skin formation, blood flow velocity, measured skin temperatures, the tissue structure (muscle, fat, etc.), the measuring site, and more. Because of those, different types of noise signals can be found within the EMG signals. This could have a bad effect on the results of feature extraction and on the diagnosis of the EMG signals. For this reason various denoising methods have been developed [28].

#### 1.4.1 Inherent Noise in the Electrode

Inherent Noise is a kind of noise common to all the electronic equipments that generate electrical noise. The frequency range in which this noise is, goes from 0 Hz to several thousand Hz. This noise can be eliminated by using intelligent circuit design and high quality instruments.

#### 1.4.2 Inherent Instability of the Signal

The frequency components between 0 and 20 Hz are very unstable because they are affected by the firing rate of the motor units. The firing rate of the motor units is quasi-random so that these components of the signal create a noise considered as unwanted. The numbers of active motor units, motor firing rate and mechanical interaction between muscle fibers can change the behavior of the information in the EMG signal [29].

#### 1.4.3 Movement Artifact

Movement artifact is a noise that appears every time the cables move during the connection to the amplifier. Another cause that creates movement artifacts is the interface between the surface of the electrode and the skin. Muscle fibers generate electric activity whenever muscles are active [30]. When the muscle is activated, the length of the muscle decreases and the muscle, skin and electrodes move with respect to one another. So, the electrodes that record the signal close to the muscle groups will show movement artifacts. The frequency range of the motion noise is usually 1–10 Hz with a voltage amplitude comparable to that of the EMG. The movement artifact can be removed efficiently by means of recessed electrodes, in which a conductive gel layer is interposed between the skin surface and the electrode-electrolyte interface. Another kind of movement artifact is due to the potential difference between skin layers. In this specific case recessed electrodes can not remove this artifact. However, this type of artifact is attenuated reducing the skin impedance [31]. Tam and Webster [32] found that a good treatment of the skin before recording, reduces these artifacts. The virtual movement between skin surface electrodes and the innervations zones of the underlying motor units can cause another type of motion artifact. Mesin et al. [33] explain the effect of the innervation zone on amplitude, frequency and conduction velocity estimates from the signal and the effect of electrodes placed close to IZ so that is better avoiding to record near that zone(s).

#### 1.4.4 Electromagnetic Noise

The human body behaves like an antenna—the surface of the body is continuously in contact with electric and magnetic radiation, which is the source of electromagnetic noise [28]. Electromagnetic sources from the environment superimpose the unwanted signal, or cancel the signal recorded from a muscle. The amplitude of the electromagnetic noise is often one to three times greater than the EMG signal. The human body's surface continuously emits electromagnetic radiation. The dominant ambient noise has a frequency of 50 Hz in Europe (60 Hz in USA) and comes from power sources and it is also called Power-Line Interference (PLI).

#### 1.4.5 Cross Talk

Cross talk signal is a type of signal recorded from groups of muscle that are not directly observed, but placed near the muscle the user is interested in. In this way Cross talk contaminates the EMG signal desired and induces the user to the error and to a bad interpretation of the EMG information. Crosstalk depends on the many physiological parameters [34], and can be minimized by choosing electrode size and inter-electrode distances carefully. Crosstalk increases with increasing subcutaneous fat thickness and has a different shape with respect to the signals detected directly over an active muscle and has a broader bandwidth than these signals. Neither the cross-correlation coefficient analysis nor a high pass filtering method have effect on reducing crosstalk. In addition Electrocardiogram (ECG) spikes can also interfere measurements that involve small muscle groups which are close to the heart even if in this case the study interests involve the lower limbs muscles. Anyway ECG spikes are at the frequency of about 1 Hz, so they can be attenuated by means of a band-pass filter.

#### 1.4.6 Internal Noise

Internal Noise is a type of noise that affect the quality of the EMG due to different factors, such as biochemical and physiological that take place because of the number of muscle fibers per unit, the anatomy and in addition the depth and location of active fibers, as well as amount of tissue between contracting muscles and electrodes. Hemingway et al. [35] showed that the electromyographic activity is inversely proportional to the thickness of the subcutaneous tissue between the surface electrode and active muscles. When the former increases the latter decreases and viceversa.

## 1.5 Emg recording techniques

EMG signal is the electrical activity of a muscle's motor units, and can be recorded in different ways such as surface EMG (sEMG) and intramuscular EMG,

Specifically:

- **Intramuscular EMG:** Invasive technique for measuring muscle electrical activity resulting from the muscle activity. In this case needles are used to detect the EMG signal.
- **Surface EMG:** Non-invasive technique for recording muscle electrical activity resulting from the muscle activity and by means of different number and kind of electrodes.

In this specific case the main interest is on the surface EMG (sEMG) and for this reason it is going to be studied deeply.

## 1.6 Surface EMG (sEMG)

Here is some characteristics about surface EMG.

### 1.6.1 Electrodes

The most common electrode derivations used in sEMG include bipolar, monopolar, and Laplacian configuration. When bipolar derivation is applied, sEMG signal is the voltage difference between a pair of recording surface electrodes aligned along the length of the skin surface of the muscle [14]. Monopolar electrode configuration measures a difference between the electrode on active site (the muscle) and a common reference electrode on non-active site that typically is on bony area [36].

Laplacian configuration uses typically one central surface electrode and number of surrounding electrodes. Research interest is going now towards high-density sEMG (HD-sEMG) which use electrodes matrices to record the muscle activity and this technique is used in this project.

### 1.6.2 Electrode types

A surface electrode is characterized by factors may strongly influence the recorded sEMG signal such as physical dimension, shape, technology, and constituent materials.

Surface EMG electrodes can be classified considering the materials and the technologies adopted for their manufacturing [14]. We can distinguish between dry and non-dry or wet

electrodes.

Wet electrodes require conductive gel, hydrogel or sponge saturated with an electrolyte between the electrode and the skin, but can provide high quality sEMG signals as well as they require skin preparation (e.g. shaving and skin abrasion), in order to reduce skin–electrode impedance and motion artifacts [37]. In addition, the wet electrodes are not good for long-term use due to the drying of the conductive gel. It could cause irritation, discomfort, skin allergy and inflammation. These electrodes are often self-adhesive, so they can be easily applied for a dynamic use. Instead nowadays dry electrodes do not require conductive gel and skin preparation, and they can get signal quality comparable to wet electrodes [38]. For this reason, the dry electrodes are the most used today.

The material of the electrode affects its electrochemical behavior. It is possible to distinguish between polarizable and non polarizable electrode. On the one hand the Polarizable ones (e.g. gold, platinum and iridium electrodes) are characterized by capacitive behavior because only displacement current passes between the skin and electrode. On the other hand non-polarizable electrodes (e.g. galvanized and sintered Ag/AgCl electrodes) behave like resistors since they allow a free flow of charge across the electrode–skin interface. It is important to specify that it is difficult to have a perfectly non-polarizable or polarizable electrode. Anyway polarizable electrodes are not recommended for sEMG measurements because of their high sensitivity to motion artifacts [14]. Commonly used non-polarizable silver–silverchloride (Ag/AgCl) electrodes, made of a silver metal surface plated with a thin layer of silver chloride, are highly stable.

### 1.6.3 Electrode shape

There are no information in literature about the recommendations for the sEMG-electrode shape and in a lot of studies the shape of electrodes is not indicated. By the way also SENIAM does not define a clear recommendation or criteria to choose a specific shape instead of another in order to obtain a better signal recorded. In general circular, rectangular or bar electrodes are the most diffused. Electrodes matrices (rectangular shape) are used in HD-sEMG. The sEMG signal is always dominated by superficial sources closest to typically within 10–20 mm. IED is also critical to HD-sEMG systems. The sampling rate in space is related to IED, and too long IED may result the sampling of surface potentials at the rate below Nyquist frequency, and thus generating spatial aliasing. A recent study [66] demonstrated that in order to avoid spatial aliasing maximum IED of 10 mm should be used, and recommended the

IED has to be below 10 mm for HD-sEMG systems

#### **1.6.4 Inter-Electrode Distance (IED)**

The Inter-Electrode Distance (IED) is very important to considerate for evaluating the recorded EMG signal, in fact the muscle volume considered by the recording sistem is a sphere with radius equal to the IED. For this reason Bipolar EMG channels are preferred in sEMG interfaces because they are more tolerant to noise than the monopolar ones.

Bipolar EMG channels are preferred in sEMG interfaces as they are more tolerant to noise than for example monopolar ones [39] even if Merletti et al. say that it contains the entire information which can be recorded from the detection volume [14]. The larger the inter-electrode distance (IED) is, the wider the pick-up volume sampled is. It implies higher amplitude of the signal but less spatial specificity.

#### **1.6.5 Electrodes Placement**

The placement of electrodes is important to improve the signal-to-noise ratio of sEMG signals placing the electrodes close to the EMG signal source. The electrodes are separated from the muscle of interest by a layered volume conductor composed by subcutaneous tissue (adiposetissue and other soft tissues), and the skin (Figure 14), acting as a spatial lowpass filter smoothing the detected MUAPs and thus decreasing their amplitude and frequency content [14]. The impedance quantification is an issue without a unique solution because its value depends on time and on many parameters such as the materials of the electrode, the gel chemical composition, the electrode size, the surface structure, and the skin treatment. Furthermore during long term recordings, the gel dryies changing skin properties.

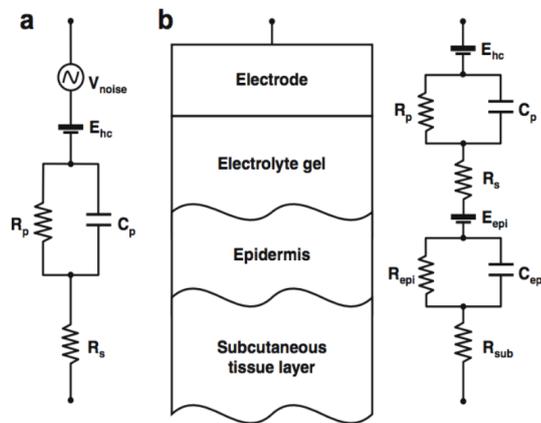


Figure 14: Models of the electrode–skin impedance: (a) simplified electrical model of the electrode–gel interface:  $E_{hc}$  is the half-cell potential at the metal–electrolyte junction, the parallel  $R_p C_p$  takes into account the polarizability and the capacitive behavior of the junction,  $R_s$  describes the resistive behavior of the electrolyte gel,  $V_{noise}$  is the associated noise component; (b) generalized model of the electrode–skin interface. The electrode–electrolyte junction described in (a) is expanded in order to take into account the effect of the conductive gel ( $R_p$ ,  $C_p$ , and  $R_s$ ) and the electrolyte–skin interface.  $E_{epi}$  is the half-cell potential due to differences in the ionic concentrations between the gel and the superficial layer of the skin, the parallel  $R_{epi} C_{epi}$  characterizes the skin impedance, and  $R_{sub}$  is the resistive component associated to the subcutaneous tissue layer. Equivalent noise generators are not indicated for simplicity but are present at all interfaces. Taken from [14]

## 2 Materials and Methods

In this section the main steps useful to find the best feature subset are presented. In particular way there are information about the experimental setup, the muscles and subjects involved in the experiments, and the movements performed. Furthermore, is explained the definition of the fourteen features, in time and frequency domain, taken into account in this study and the approach (filter approach - J index) used to evaluate their quality in terms of class separability, choosing then the best feature subset. To validate the chosen approach, it was compared to another approach (wrapper approach) from the literature, based on a LDA classifier. The former, compared to the latter is simpler in terms of implementation and computation. The correlation between both approaches is obtained by means of the Spearman's rank correlation coefficient. Before the application of the J index, a previous discrimination in terms of calculation time is performed, in order to exclude from the study those features that need too much time to be calculated.

### 2.1 High Density sEMG choice

Many works exist in literature about the myoelectric control of external devices such as prostheses, orthoses and exoskeletons. Myoelectric control is usually performed using standard bipolar EMG. In recent years, it has been demonstrated that sEMG recorded with bipolar EMG is affected by many factors depending on central and peripheral factors; among them the geometrical changes of the muscle during dynamic tasks. In order to collect a more robust signal, representative of the muscle activity, high density sEMG has been proposed. This technique, by detecting EMG from the whole muscle surface, is more robust than bipolar technique. In section 3.1 it is demonstrated that the muscle activations are not spatially equal on the muscle surface, in fact during a specific task there is a different activity in amplitude and shape from one zone to another. Furthermore the most muscle area activated during different tasks is not always the same. For this reason we can't take into account only a specific area from every muscle.

## 2.2 Experimental Set Up

In the following the experimental setup is described. It has been identified after some pilot experiments focused to identify the best electrode positioning using the SENIAM recommendations and taking into account the technological limits. Six flexible electrode matrices of different number of channels, produced by OT-Bioelettronica, were used to record the EMG signals from seven lower limb muscles. In particular were used four matrices of 32 electrodes and two of 64 electrodes and all of them were connected to OT Quattrocento using all of the eight input of 16 channels (from IN1 to IN8) and two MULTIPLE IN of 64 channels (Figure 15). The electrodes placement has to be on the belly of the muscles and not in correspondence of the tendons. Before electrodes placement, skin preparation is mandatory in order to avoid the presence of body oil, body salt, hair and dead cells. The skin has to be shaved and then cleaned using an abrasive gel applied on the entire electrodes site and rubbing with a gauze pad. In this way a reduction of skin impedance would enhance performance of the monitoring electrodes.



Figure 15: Quattrocento amplifier for the acquisition of surface/intramuscular monopolar or bipolar EMG signals and technical specifications. Taken from [15]

### 2.2.1 Muscle Identification following SENIAM recommendations

Before electrodes placement, the muscle identification is very important to record the sEMG from the right area of interest. In order to find muscles, the SENIAM recommendations are followed with the help of the user manual palpation asking to the subject the execution of specific movements.

Here are the SENIAM recommendations muscle per each muscle studied:

- **Vastus Medialis:** Respect to the distal portion of the line between the medial side of the patella and the anterior superior iliac spine, the right landmark is on a line oriented  $50^\circ$  (Figure 16). During the identification of Vastus Medialis, the subject had to be at the edge of a table. He had to perform an isometric contraction during a knee extension. The shank had to be at  $45^\circ$  [16].

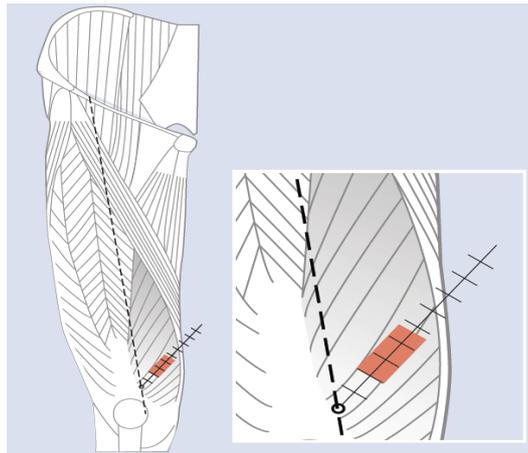


Figure 16: Vastus Medialis muscle identification. Taken from [16]

- **Vastus Lateralis:** Respect to the distal portion of the line between the lateral side of the patella and the anterior superior iliac spine, the right landmark is on a line  $20^\circ$  oriented. During the identification of Vastus Lateralis, the subject had to be at the edge of a table. He had to perform an isometric contraction during a knee extension. The shank had to be at  $80^\circ$  [16].

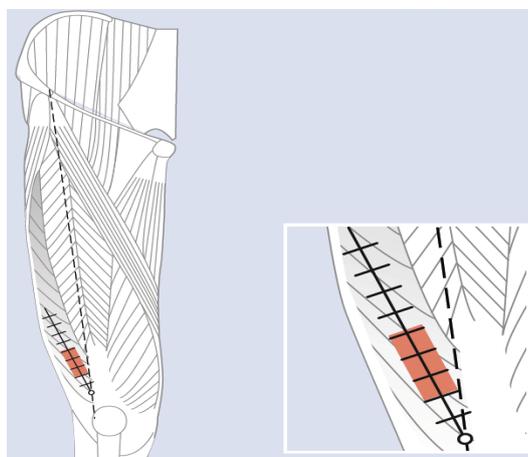


Figure 17: Vastus Lateralis muscle identification. Taken from [16]

- **Tibialis Anterior:** The right landmark was chosen at around the 19 % of the line between the tibial tuberosity and the intermalleolar line. During the identification of Tibialis Anterior the subject, seated on a chair, had to perform an isometric contraction during an ankle dorsal extension [16].

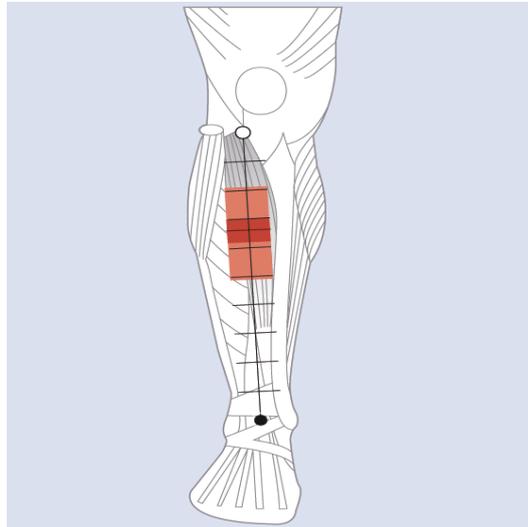


Figure 18: Tibialis Anterior muscle identification. Taken from [16]

- **Rectus Femoris:** The right landmark is between 0% and 50% of the line between the superior side of the patella and the anterior superior iliac spine. During the identification of Rectus Femoris the subject, seated at the edge of the table had to perform an isometric contraction during an 80° extension of the knee [16].

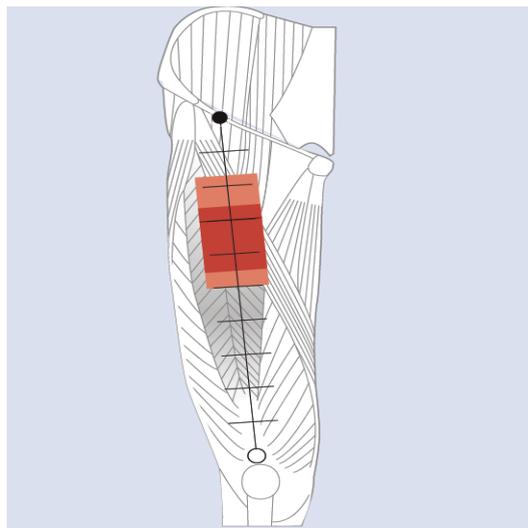


Figure 19: Rectus Femoris muscle identification. Taken from [16]

- **Gastrocnemius Medialis:** The right landmark is between the 87% and 100% of the line between the medial side of the Achilles tendon insertion and the medial side of the popliteal cavity. The subject had to be in upright position [16].

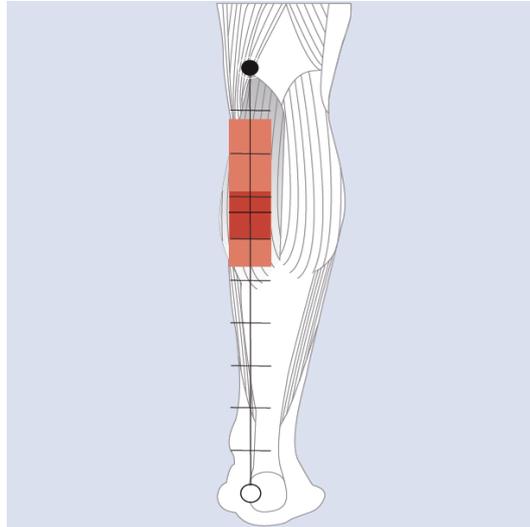


Figure 20: Gastrocnemius medialis muscle identification. Taken from [16]

- **Semitendinosus and Biceps femoris:** In this case the matrix was fixed midway between the landmarks of both muscles. The landmark for the Semitendinosus is between 0% and 26% of the line between the ischial tuberosity and the medial side of the popliteal cavity. The landmark for the Biceps Femoris is between 0% and 22% of the line between the ischial tuberosity and the lateral side of the popliteal cavity. The subject had to be prone with the knee flexed at  $45^\circ$  performing an isometric contraction during flexion of the knee placed in internal rotation for the Semitendinosus and external for the Biceps Femoris [16].

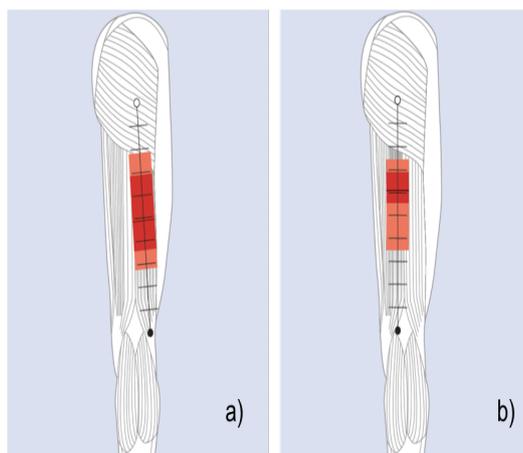


Figure 21: a) Semitendinosus and b) Biceps Femoris muscle identification. Taken from [16]

### 2.2.2 Muscles and Matrix correspondence

In Figure 22 you can see a picture about the matrices fixed on the subject's muscles before movement executions. In the table there is the correspondance between muscle and matrix used on the surface of that muscle and in Figure 23 we can see the detection side of matrices.

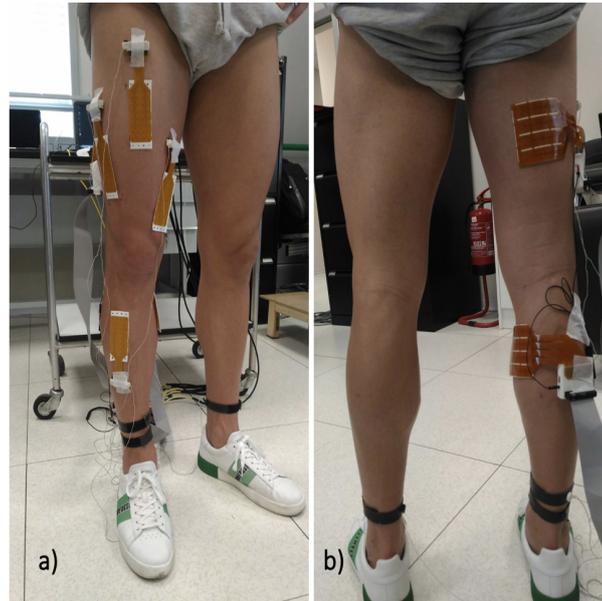


Figure 22: a) Matrices of 32 electrodes positioned on Vastus Medialis, Vastus Lateralis, Tibialis Anterior, Rectus Femoris; b) Matrices of 64 electrodes positioned on Gastrocnemius Medialis and Biceps Femoris&Semitendinosus

Muscle	Matrix
Vastus Medialis	32 electrodes
Vastus Lateralis	32 electrodes
Tibialis Anterior	32 electrodes
Rectus Femoris	32 electrodes
Gastrocnemius medialis	64 electrodes
Semitendinosus and Biceps Femoris	64 electrodes

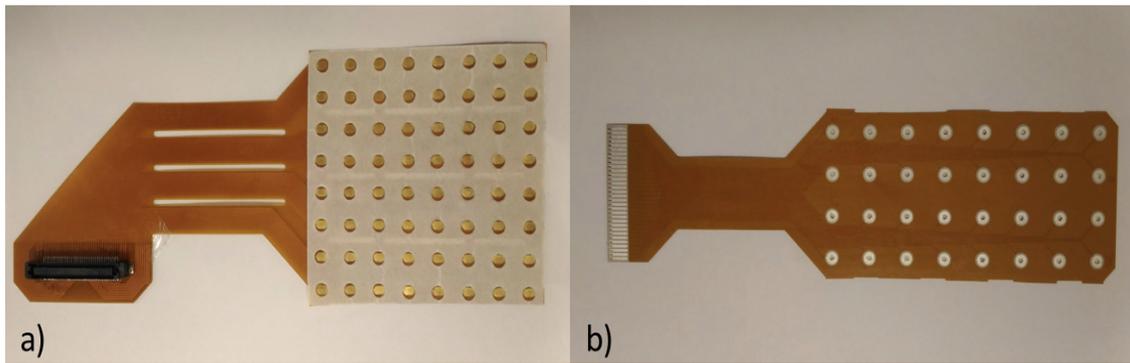


Figure 23: a) Matrix of 64 electrodes. b) Matrix of 32 electrodes

### 2.3 Experimental Protocol

The experimental protocol consists of 7 different tasks and all of these, except for both rest tasks, were repeated 15 times in order to record more repetitions for the same task. All the tasks are movements that every person usually does during a daily life routine. In particular:

- **Sitting:** At the beginning, the subject is in upright position with an ankle angle of  $180^\circ$  and then he sits down on a chair thus obtaining an ankle angle of  $90^\circ$ . Feet must not move.
- **Standing:** At the beginning the subject is seated on a chair with an ankle angle of  $90^\circ$  and then he starts to stand up until reaching upright position with an ankle angle of  $180^\circ$ . Feet must not move.
- **Stair ascending:** At the beginning, the subject is in upright position. When the movement starts, the subject climbs a step 18 cm high. This task ends when both feet are on the step and the person is in upright position again. The first limb to move is the right one that is the sensorized one as well.
- **Stair descending:** At the beginning, the subject is in upright position with both feet on the step. When the movement begins, the subject goes down from the step until reaching the floor with both feet. The first limb to go down has to be the right one that is the sensorized one as well.
- **Gait cycle:** This task consists of a single gait cycle going from one heel off of the sensorized leg to the next one.
- **Rest sitting:** The subject is resting, seated on a chair.

- **Rest in upright position:** The subject stands in an upright position.

## 2.4 Subject reclutation

Nine healthy subjects participated to the study (age between 24 and 27 years). All the subjects were right-handed and the sensorized lower limb was the dominant one.

## 2.5 Data acquisition

In order to record the EMG signals an interactive GUI (Figure 24) was realized in MATLAB. The only software realized by OT-Bioettronica, called OT-Biolab, is just to record and to visualize signals.

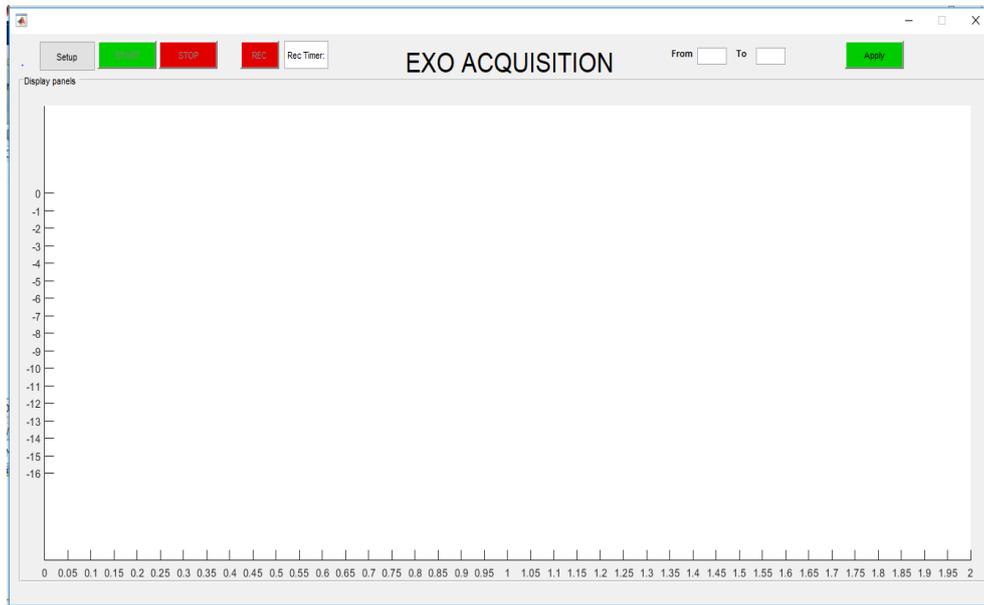


Figure 24: GUI realized in MATLAB for visualization and EMG recording

In particular, five push buttons are available on the GUI and they are activated in different conditions. First of all, the *SET UP*. botton is clickable in order to choose the right configuration that the user is gonna use during his experiment. Once *SET UP* is selected, a window will open as shown in (Figure 25):

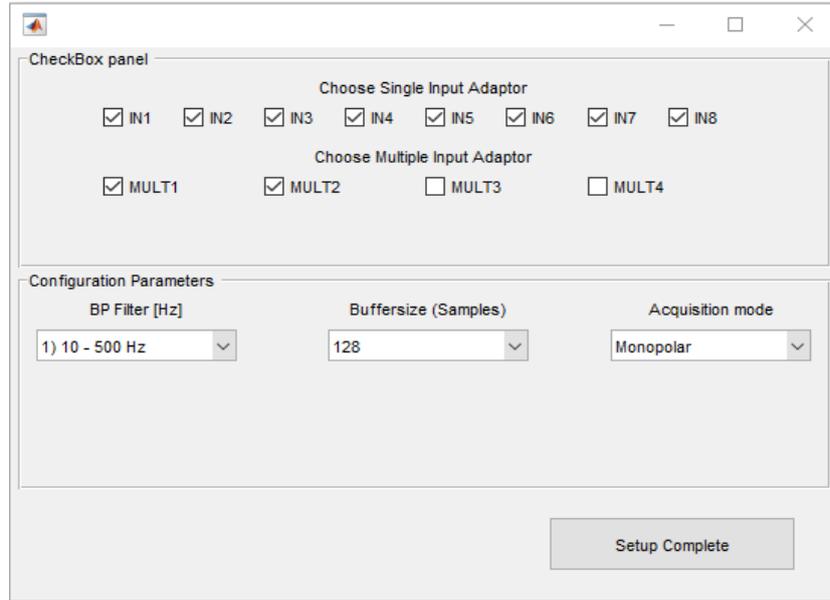


Figure 25: GUI realized in MATLAB for visualization and EMG recording

In the upper part, there is a checkbox panel, where it is possible to select which single and multiple adaptors are physically connected to the Quattrocento. In the lower part instead, configuration parameters such as *bandwidth* of the filter, *buffersize* and *acquisition mode* are selectable by means of three popup menu. About the filter, the user can choose between two bandpass filters *10-500* and *20-400*. The possible buffer sizes are *256*, *512*, *1024* and *2048* and the acquisition mode *monopolar* or *single differential*. During the experiments, in this project, all the single inputs and two of the multiple ones have been selected as well as 10-500 as bandwidth, 256 for the buffersize and the acquisition mode is monopolar. Once finished the setup, you can go ahead using the *Setup Complete* button. After that, the *Start* button becomes active so that you can start the acquisition. During the acquisition only 16 channels are plotted (IN1 by default) because of computational limits, and for this reason you can move from the channels of one matrix to the channels of another one or within the same muscle using the *From...To...* text boxes and pushing *Apply*. At this moment, in case of the user want to start recording, you can push on *REC* which becomes *STOP* to block the recording and to save that acquisition in a *.mat* file. Next to that a stopwatch shows the recording duration. Every recorded task is triggered by the subject who uses a mechanic botton pushed at the beginning and at the end of each repetition.

## 2.6 Data Processing

This section is focused on data processing which is divided in filtering, to know what filters were used before visualization and algorithm execution, and in channel remapping and single differential calculation in order to see the right spatial information during visualization and to compute the right single differential calculation which is the kind of signal used in whole the algorithm

### 2.6.1 Filtering Step

Once recorded all the 15 repetitions of the 7 tasks, the data processing is gonna start. First of all an additional filtering step is added (in addition to the hardware filter 10-500 Hz) in this part of algorithm. In particular, a band pass filter with bandwidth [20-500], and a recursive band-reject (notch) filter for removing 50 Hz are used. More specifically:

- **Band-Pass filter:** This filter of order 4 and bandwidth [20 500] was realized using the matlab function *Butter* in order to find the transfer function coefficients(a,b). In this case the transfer function is expressed in terms of b and a as:

$$H(z) = \frac{B(z)}{A(z)} = \frac{b(1) + b(2)z^{-1} + \dots + b(n+1)z^{-n}}{a(1) + a(2)z^{-1} + \dots + a(n+1)z^{-n}} \quad (4)$$

The filtering is executed using the MATLAB function *filtfilt* to avoid phase distortion and the consequent filter delay. The choice of this bandwidth is in accordance with the literature. In particular Carlo J. De Luca *et al.* [40] explain that spectral analysis reveals that the movement artifact component continues to influence the EMG shape until approximately 20Hz. At the same time 10Hz filter does not totally remove the artifact component as well as 30Hz filter attenuates the artifacts but also remove the lower frequency component of sEMG signal. About the upper band limit it is known that [39] the dominant energy (about 95%) of an sEMG signal is limited to harmonics up to 400–500 Hz, so components higher than 500 are considered as noise.

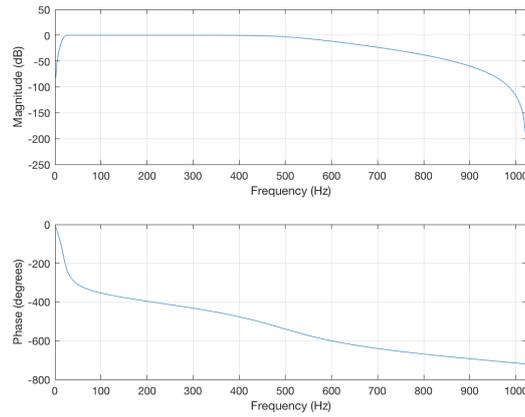


Figure 26: Mask of band-pass filter

- **Band-reject filter:** This filter of order 2 and rejected band centred in 50Hz, which is the frequency corresponding to the power-line interference in UK and in Italy where the algorithm was tested. The MATLAB function used was *Butter* with *Stop* instruction. The transfer function formula (4) used in this case is the same to that used for the band pass filter. The filtering is executed using the MATLAB function *filtfilt* to avoid phase distortion.

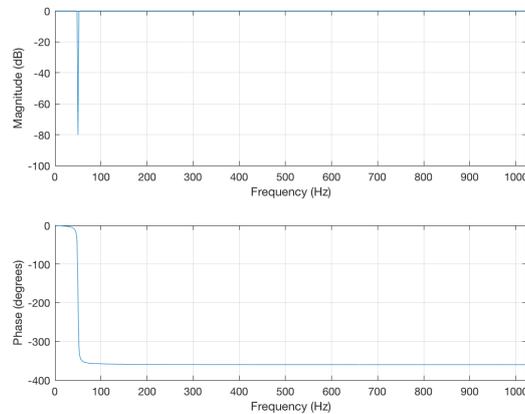


Figure 27: Mask of band-stop filter

### 2.6.2 Channels remapping and Single differential calculation

After filtering, since the EMG signals are saved as monopolar and we need to calculate the single differential (Figure 12) among those channels, a channel remapping is necessary in order to obtain differential signals along each column of the matrices following the direction of the muscle fibers. It happens because each electrode has a number that indicates the number

of the input pin to which it belongs, and a different orientation of the matrix on the muscle requires a remapping, since the columns, on which the single differential has to be calculated, change in base on the different matrix orientation.

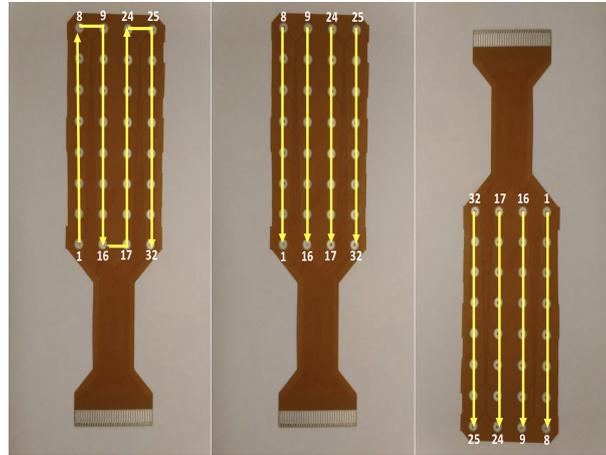


Figure 28: a) channels sequence during acquisition. b) directions of single differential calculation on tibialis anterior. c) directions of single differential calculation on vastus lateralis, vastus medialis and rectus femoris.

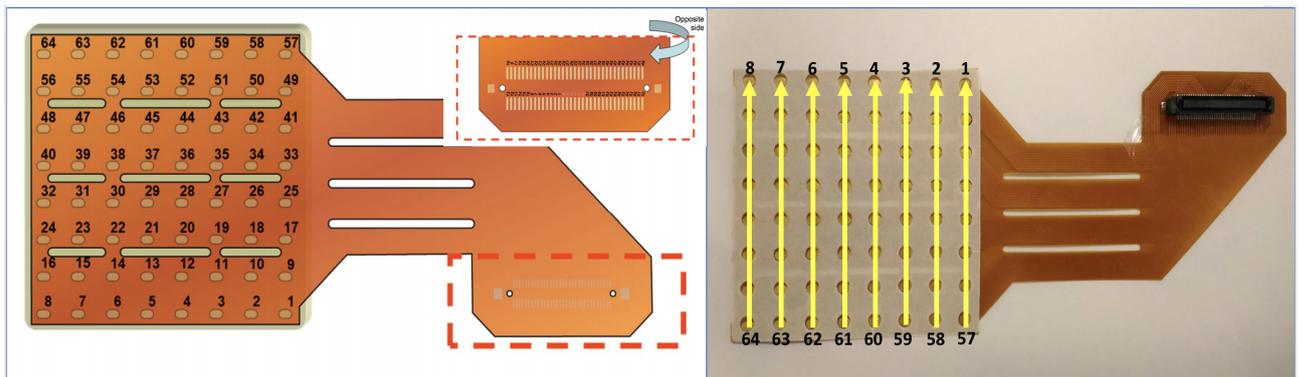


Figure 29: a) channels sequence during acquisition using matrices of 64 electrodes. b) directions of single differential calculation on Gastrocnemius medialis and Biceps Femoris&Semitendinosus

## 2.7 Bad channels selection

After single differential signals calculation, an interactive GUI was designed in MATLAB (Figure 30) in order to look at the signals divided in matrices in order to see the contribution of each muscle in separate way.

This GUI allows the user to move from one repetition to another, using *Next Rep*, and within a repetition from one muscle to another using *Next Matrix* and *Prev.Matrix*. It often happens that there are bad channels, mainly because of a low quality electrode-skin contact. In these cases it is mandatory to delete the bad channels and, to do that the user can write

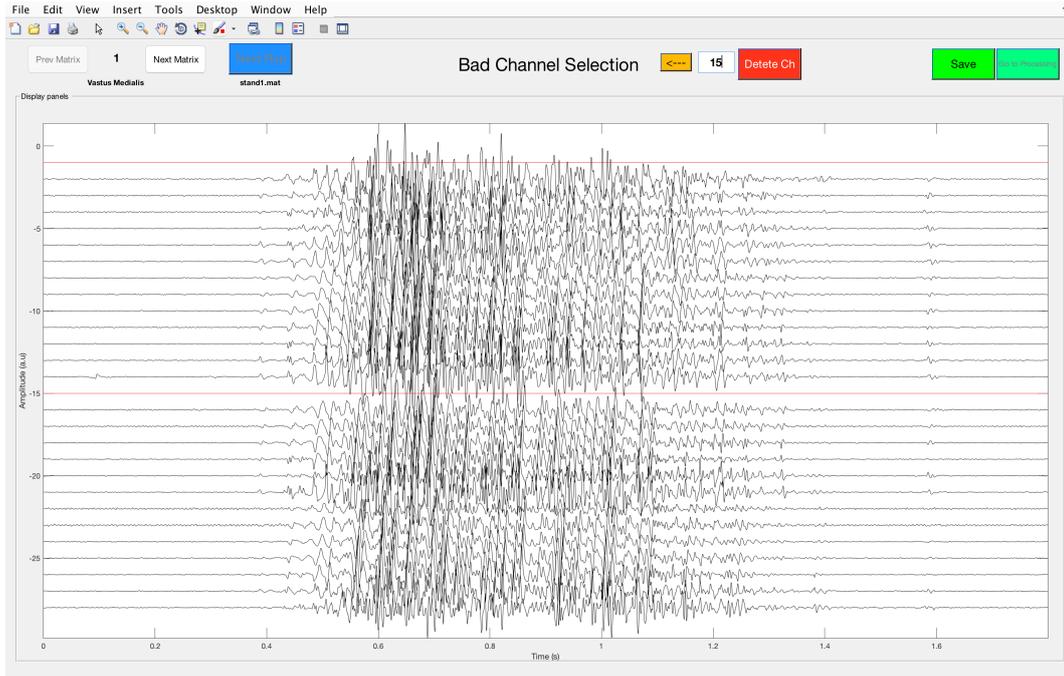


Figure 30: Bad channel selection GUI

the number of the channel to delete in the text box, pushing then the button *Delete Ch*. So, the bad channel will become red. If the user deletes the wrong channels he can come back pushing on the arrow "*<-*". Obviously if the first matrix is selected, you can't go to the previous matrix as well as if you select the last one matrix you can't go to the next muscle. Another check was added when you write the channel number to delete, infact if you digit a number that is out of range for the total channels of that matrix, nothing is gonna happens. When the last repetition of the last task is analyzed, the *Save* button is enabled, and pushing on that one all the bad channels will be saved in a vector, usefull when the user selects *Go to Processing* to proceed to the processing, letting the main algorithm start.

When a channel of a matrix is deleted for one repetition of a task, that channel will be deleted for all the tasks so that the same subset of electrodes is used for all tasks for *Within class scatter matrix* ( $S_W$ ) and *Between class scatter matrix* ( $S_W$ ) calculation in the following sections.

## 2.8 Pattern Recognition

The recognition of limb motions from the electromyogram (EMG) plays an important role in rehabilitation engineering. For instance, it is pretty evident that different technologies have beneficial effect on the rehabilitation progress of stroke patients, and such kind of

therapies require detection of the movement the patient would like doing. Furthermore, by extracting the intention, the patient can control a big amount of devices. In fact EMG pattern recognition systems have been widely used in many human-machine interface applications. Pattern recognition-based approach is based on the assumption that the classifier is able to recognize the input values and to assign each input value to one of a given set of classes. It is possible thanks to a training session that allows the classifier to identify the class of belonging during the test session. Input values are the values of features extracted from the sEMG signal, and classes correspond to different tasks the patient is doing. After feature extraction, feature selection is very important in order to select those features more significant than others.

### 2.8.1 Feature Extraction

Fourteen features in time and in frequency domain were computed on epochs of 300ms with 50% overlap (Figure 31). All the recordings have different duration among them (1.8 seconds give or take) but, since it is necessary to have signals of the same length, *resample* MATLAB function was used in order to obtain the same number of windows among all the repetitions. In this way all the repetitions of all the tasks were resampled to 1.8s. The choice of epoch 300ms long with 150ms of overlap was driven by the fact that overlapped window approach is more appropriate in sEMG control systems because it produces better classification accuracy and it avoids long latency in real time application in which 300ms are the maximum time allowed between the movement intention and the movement executed by the device. Actually, the minimum length for a segment to individuate a movement is 200ms which is the minimum interval between two different contractions [41]. In particular, looking at (Figure 31) we can see that the effective time (Incr.) to wait before getting the next segment ready is 150ms.

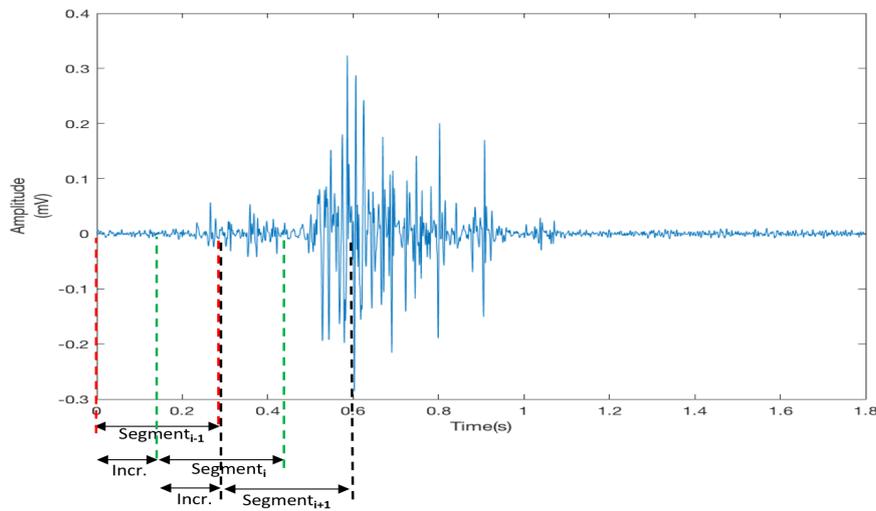


Figure 31: Segmentation with 50% overlap. The difference between segment and increment corresponds to overlap. Increment is the time to wait going from one segment to the next one.

Here are the fourteen features taken into account in this study and both their definition and mathematical formulations:

- **Variance of EMG (VAR)**

Variance of EMG (VAR) is another power index. Generally, variance is defined as an average of square values of the deviation of that variable; since the mean value of EMG signal is close to zero [42], variance of the EMG signal can also be defined as

$$VAR = \frac{1}{N-1} \sum_{i=1}^N x_i^2 \quad (5)$$

- **Integrated EMG (IEMG)**

Integrated EMG is commonly used in EMG non-pattern recognition and clinical application as onset detection index [43]. It is defined as the summation of absolute values of the EMG signal amplitude:

$$IEMG = \sum_{i=1}^N |x_i| \quad (6)$$

- **Mean Absolute Value (MAV)**

Mean of Absolute value is one of the most used feature in EMG analysis and in literature

it can be found under different names such as ARV, IAV or AAV. [44] It is defined as the average of EMG signal in a segment:

$$MAV = \frac{1}{N} \sum_{i=1}^N |x_i| \quad (7)$$

- **Modified Mean Absolute Value 1(MAV1)**

Modified Mean Absolute Value 1 is an extension of MAV using weighting window function [45] . It is shown as

$$MAV1 = \frac{1}{N} \sum_{i=1}^N w_i |x_i| \quad (8)$$

$$w(n) = \begin{cases} 1, & \text{if } 0.25N \leq i \leq 0.75N \\ 0.5, & \text{otherwise} \end{cases}$$

- **Modified Mean Absolute Value 2(MAV2)**

Modified Mean Absolute Value 2 is a variation of MAV1 and the smooth window is improved in this method using continuous weighting window function [45]. In particular

$$MAV2 = \frac{1}{N} \sum_{i=1}^N w_i |x_i| \quad (9)$$

$$w(n) = \begin{cases} 1, & \text{if } 0.25N \leq i \leq 0.75N \\ \frac{4i}{N}, & \text{if } 0.25N > i \\ \frac{4i-N}{N}, & \text{if } 0.75N < i \end{cases}$$

- **Mean Frequency(MNF)**

This feature estimates the mean frequency of the signal in a time segment [46] and [47]:

$$MNF = \frac{\sum_{j=1}^M f_j P_j}{\sum_{i=1}^M P_i} \quad (10)$$

$f_i$  is the frequency of the spectrum at frequency bin  $j$  and  $f_i$  is the intensity of frequency spectrum.  $P_j$  is the EMG power spectrum at frequency bin  $j$ , and  $M$  is length of the frequency bin.

- **Median Frequency(MDF)**

MDF is a frequency at which the EMG power spectrum is divided into two regions with equal amplitude [47] or it can also defined as a half of the total power:

$$MDF = \sum_{j=1}^{MDF} P_j = \frac{1}{2} \sum_{j=1}^M P_j \quad (11)$$

$h_i$  is the frequency and  $f_i$  is the intensity of frequency spectrum.

- **Root Mean Square(RMS)**

Root Mean Square (RMS) is modeled as amplitude modulated Gaussian random process whose RMS is related to the constant force and non-fatiguing contraction. It relates to standard deviation, which can be expressed as

$$RMS = \sqrt{\frac{1}{N} \sum_{i=1}^N x_i^2} \quad (12)$$

- **Waveform Length(WL)**

Waveform length (WL) is the cumulative length of the waveform over the time segment. WL is related to the waveform amplitude, frequency and time.

$$WL = \sum_{i=1}^{N-1} |x_{i+1} - x_i| \quad (13)$$

- **Absolute value of the 3rd, 4th, and 5th temporal moment (TM3,TM4,TM5)**

Absolute value of the 3rd, 4th, and 5th temporal moment Temporal moment is a statistical analysis that was proposed in study of Saridis and Gootee [48] to be used in control of a prosthetic arm. Normally, the absolute value was taken to greatly reduce the within class separation for the odd moment case. Since the first moment and the second moment are similar to the MAV and VAR features, respectively, the third, fourth, and fifth moments (TM3, TM4, and TM5) were used and evaluated as well. They can be respectively expressed as:

$$TM3 = \left| \frac{1}{N} \sum_{i=1}^{N-1} x_i^3 \right| \quad (14)$$

$$TM4 = \frac{1}{N} \sum_{i=1}^{N-1} x_i^4 \quad (15)$$

$$TM5 = \left| \frac{1}{N} \sum_{i=1}^{N-1} x_i^5 \right| \quad (16)$$

- **Log detector (LOG)**

This feature also provides an estimate of the muscle contraction force [49]. However, definition of the non-linear detector, which implicitly estimates muscle force, is changed to be based on logarithm and log detector (LOG) feature. It is defined as:

$$LOG = e^{\frac{1}{N} \sum_{i=1}^{N-1} \log(|x_i|)} \quad (17)$$

- **Difference absolute standard deviation value (DASDV)**

This feature looks like RMS feature, in other words, it is a standard deviation value of the wavelength (Kim et al., 2011), as can be defined by

$$DASDV = \sqrt{\frac{1}{N-1} \sum_{i=1}^{N-1} (x_{i+1} - x_i)^2} \quad (18)$$

### 2.8.2 Computational Time

Since in many applications the computational time is crucial, also the time that the algorithm needs to obtain the value of each feature about one class was calculated in this study. In our case in fact, after recording signals from a movement, we would like to calculate a feature, to classify that movement and to send a control signal within 300ms which is the maximum time to considerate an application as real time. In addition, we want to make a decision as soon as possible in order to avoid significant delay from the subject intention and the movement of the end effector. Even though the time is important we can't evaluate the feature quality only in base on the computational time but it is essential to exclude from the study those features that need too much time to be calculated. In the current study, it was calculated on a MacBook Pro (2.3 GHz Intel Core i5), using MATLAB R2017b.

Since both MNF and MDF features, calculated in the frequency domain, as it can be seen in the next sections, need a lot of time to be extracted, from now on only the other twelve features will be taken into account.

### 2.8.3 Feature Evaluation

In literature there are two possible ways to evaluate feature quality for the final aim of selecting an optimal feature subset: the **wrapper** and the **filter** approach. On the one hand the wrapper approach uses the classification rate of a classifier and on the other hand the filter approach evaluates feature quality in class separability using statistical criterion methods [50]. In this project, for the reason why the filter approach need a lower computational cost and a simple implementation, it was used in order to find the best features that allow to recognize different movements. In particular J index (based on LDA theory), known also as Fisher Linear Discriminant Index (FLDI) [51] is used for this approach. Furthermore, the obtained results were compared to the classification accuracy of an LDA (Linear Discriminant Analysis) classifier demonstrating the validity of this method by means of Spearman's correlation. In general a good quality in class separation from the filter approach means that the classification rate should be as high as possible [52].

## 2.9 Linear Discriminant Analysis (LDA) Notions

This section provides the basic knowledge useful to understand the next sections about class separability index (J index) and classification.

Linear Discriminant Analysis also known as Fisher's linear discriminant analysis is widely used as dimensionality reduction technique during pre-processing in pattern recognition and machine learning applications to find a linear combination of features to separate two or more classes reducing in addition computational costs. To achieve that, Fisher's propose was essentially the maximization of the distance among the means of all classes and the minimization of the spreading within each class. To obtain this goal three steps are necessary: 1) the calculation of the separability among different classes (*between-class scatter matrix*), 2) the calculation of the distance between the mean and the samples of each class (*within-class scatter matrix*) and, 3) the projection onto a the lower dimensional space which maximizes the between-class variance and minimizes the within class variance. Figure 32 and Figure 33 explain in a schematic way LDA steps going from feature data matrix to data after projection

in case of 3 classes.

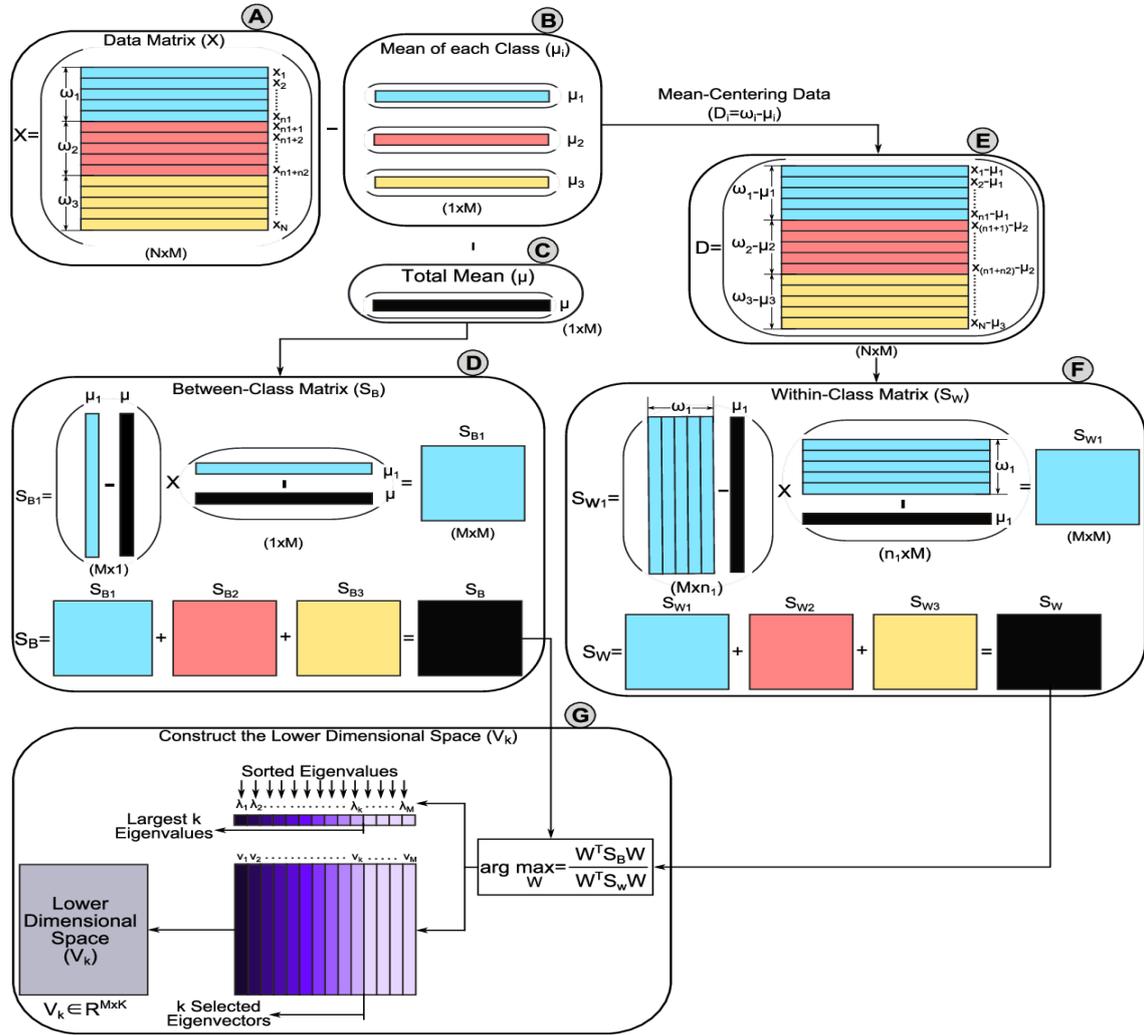


Figure 32: Schematic steps to calculate a lower dimensional subspace in LDA technique. Taken from [17]

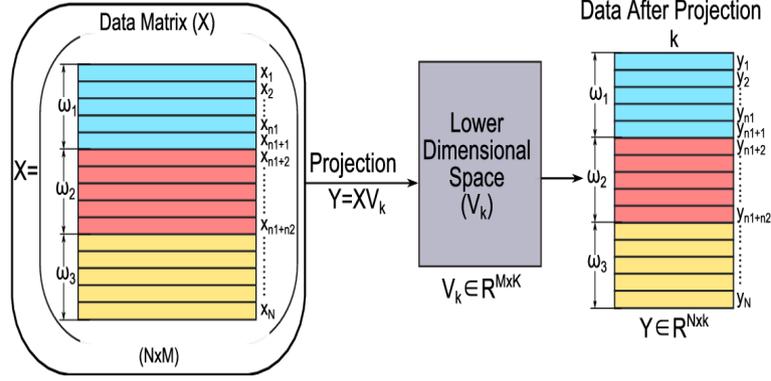


Figure 33: Projection of the original samples on the lower dimensional space of LDA. Taken from [17]

Generalizing what Figure 32 and Figure 33 show, let's assume to have a Data Matrix  $X$  ( $N \times M$ ) made by the samples of  $c$  classes (Figure 32.A);

**Within-Class scatter matrix** ( $M \times M$ ) (Figure 32.F) is calculated as:

$$S_W = S_{W1} + S_{W2} + \dots + S_{Wc} = \sum_{j=1}^c \sum_{i=1}^{n_j} (x_{ij} - \mu_j)(x_{ij} - \mu_j)^T \quad (19)$$

where

- $c$  represents the total number of classes.
- $n$  represents the total number of observations for each class.
- $\mu_j$  ( $1 \times M$ ) (Figure 32.B) is the mean vector of  $j$ -th class:

$$\mu_j = \frac{1}{n_j} \sum_{i=1}^{n_j} x_{ij} \quad (20)$$

- $x_{ij}$  represents the  $i$ -th sample in the  $j$ -th class.

**Between-Class scatter matrix** ( $M \times M$ ) (Figure 32.D) is calculated as:

$$S_B = S_{B1} + S_{B2} + \dots + S_{Bc} = \sum_{j=1}^c n_j (\mu_j - \mu)(\mu_j - \mu)^T \quad (21)$$

where

-  $\mu$  ( $1 \times M$ ) is the mean vector of class mean vectors (Figure 32.C):

$$\mu = \frac{1}{C} \sum_{j=1}^c \mu_j \quad (22)$$

After the calculation of Within-Class scatter matrix and Between class scatter matrix, it is essential to use them in the Fisher's criterion (23) in order to find the transformation matrix  $W$  (Figure 32.G) that allow the projection of original data on the lower dimensional space.

$$\operatorname{argmax}_W \frac{W^T S_B W}{W^T S_W W} \quad (23)$$

The transformation matrix  $W$  can be obtained by solving the generalized eigenvalue problem:

$$S_W W = \lambda S_B W \quad (24)$$

Where  $\lambda$  are the eigenvalues of  $W$ . The solution to this problem can be reached calculating eigenvalues ( $\lambda = (\lambda_1, \lambda_2, \dots, \lambda_M)$ ) and eigenvectors ( $V = (v_1, v_2, \dots, v_M)$ ) of  $W$  have to be found. If  $S_W$  is not singular  $W$  can be calculated as:

$$W = S_W^{-1} S_B \quad (25)$$

In particular an eigenvector is a non zero vector that represents one axis of the LDA space, and provides the direction of new space. Each eigenvector has his own eigenvalue which is a scalar quantity about the magnitude of the corresponding eigenvalue telling us how much it is robust. Thus if the dimensionality of our observations is  $M$  like in (Figure 32.A), we will have  $M$  eigenvalues and  $M$  eigenvectors, but we have to take into account the eigenvectors with the  $k$  highest eigenvalues because it means they are better to differentiate different classes increasing between class distance and decreasing within class variance of each class (Figure 32.G). After choosing the best  $k$  eigenvectors the projection follows the formula:

$$Y = X V_k \quad (26)$$

Thus the initial  $M$  dimensionality is reduced to  $K$ , going through a lower dimensional space.

## 2.10 Filter approach - Class separability J Index for feature evaluation

Filter approach is the main criterion followed in this study. Once twelve features are extracted from EMG signals, it is essential to find which feature is better than others to separate the 7 tasks executed by the subject, in terms of between classes distance and within class spread. To do that, Fisher's discriminant is used for the calculation of the *Between class scatter matrix* ( $S_B$ ) and the *Within class scatter matrix* ( $S_W$ ). Using these two matrices a specific index (*J index*), that gives us information about the quality of that feature, can be calculated [54]. Thus, in this way, every feature will have its correspondent *J value*. In particular, the starting point is a feature matrix  $X$  (described in Figure 32), one per feature, ( $N_{windows} \times N_{channels}$ ) where:

- $N_{windows}$  are the observations for the 7 classes (tasks) and it is calculated using the number of windows of 300ms with 50% of overlap contained in a repetition. Each class has got 15 repetitions. In one repetition 1.8s long there are 11 windows. Every repetition of each class has 11 windows each, because each one of them (reps.) was resampled in time domain to 1.8s. Every  $J$  value is calculated taking into account one repetition for all the classes and it is repeated as many times as the total number of repetitions (i.e. The first  $J$  value is calculated on the first repetitions of all the classes, the second  $J$  value on the second repetitions of all the classes, ecc). In this way  $N_{windows}$  is the number of windows (11) per repetition multiply the number of classes (7), thus the rows of  $X$  are 77 and they represent the observations referred to one repetition of seven classes. In total there are 15 combinations which means 15 value of  $J$  per features.

- $N_{channels}$  is the total number of the recorded channels, excluded bad channels found as indicated in Section 2.7. This number is the same for all repetitions of all the classes.

Each element of Matrix  $X$  is the values of the feature during the  $n$ -th observation in the  $n$ -th channel.

In the following the steps for the calculation of *J index* for each feature are reported..

### 1. Within-class scatter matrix $S_W$ ( $N_{channels} \times N_{channels}$ )

$$S_W = S_{W1} + S_{W2} + S_{W3} + S_{W4} + S_{W5} + S_{W6} + S_{W7} = \sum_{j=1}^7 \sum_{i=1}^{11} (x_{ij} - \mu_j)(x_{ij} - \mu_j)^T \quad (27)$$

where

-  $\mu_j$  ( $1 \times M$ ) is the mean vector of j-th class:

$$\mu_j = \frac{1}{11} \sum_{i=1}^{11} x_{ij} \quad (28)$$

-  $x_{ij}$  represents the i-th observation in the j-th class. In our case the number of classes is 7

## 2. Between-class scatter matrix $S_B$ ( $Nchannels \times Nchannels$ )

$$S_B = S_{B1} + S_{B2} + S_{B3} + S_{B4} + S_{B5} + S_{B6} + S_{B7} = \sum_{j=1}^7 11(\mu_j - \mu)(\mu_j - \mu)^T \quad (29)$$

where

-  $\mu$  ( $1 \times M$ ) is the mean vector of class mean vectors:

$$\mu = \frac{1}{7} \sum_{j=1}^7 \mu_j \quad (30)$$

Once that  $SW$  and  $SB$  are calculated for each feature, we obtain the twelve (one per feature) values of the  $J$  index for each combination of the leave one out. It is expressed as the ratio between the trace of  $SB$  and the trace of  $SW$ :

$$J = \frac{Trace(S_B)}{Trace(S_W)} \quad (31)$$

This  $J$  index is an unbounded index, thus the best situation is to have a smaller scatter of information within one class, and a larger scatter of information among the classes, the best feature is that one with the highest  $J$  value. The most important thing is that doesn't need a high computational cost and it has a simple implementation.

### 2.11 Wrapper approach - Classification using LDA Classifier

Wrapper approach is used in this study to compare the results of  $J$  index and to validate its results demonstrating that the proposed method ( $J$  Index) is in accord with the classification accuracy for each feature. Classification of movements was performed using a multi-class

Linear Discriminant Analysis (LDA).

### 2.11.1 Training set and Test set separation

To classify the 7 different tasks, all the data from 15 repetitions for each class were divided in 14 repetitions per class as *training set* and 1 repetition per class as *test set*, using in this way a *leave one out* method. In particular, the training set matrix is the matrix  $X$  (Nwindows x Nchannels) described in Figure 32 . The number of rows of  $X$  Matrix is 11 windows per repetition, multiply 14 repetitions (in each class) taken in to account by the leave one out, multiply 7 classes. The total is 1078 observations. The number of columns corresponds to the number of channels minus the number of bad channels. The same reasoning is applied to the test set, but taking into account only one repetition. Each element of Matrix  $X$  is the values of the feature during the n-th observation in the n-th channel. The classification of all the twelve features is repeated for the 15 combinations of the leave one out. In this way we get 15 classification rates for each feature; one per each combination of the leave one out.

### 2.11.2 Linear Classification rule

The method used to classify is general, but for a better explanation let's consider two classes only. We call as  $\bar{y}_1$  the projection of group 1 average on  $a$  and  $\bar{y}_2$  the projection of group 2 average on  $a$  which is the discriminant direction. Let's also assume that  $\bar{y}_1 > \bar{y}_2$ . Being  $x_0$  the observation we want to classify and  $y_0$  his projection on  $a$ ,  $x_0$  will be assigned to the class whose average it is closest to along  $a$ . For example, if  $|y_0 - \bar{y}_1| < |y_0 - \bar{y}_2|$   $x_0$  will be assigned to the 1st class and to class 2 if the opposite inequality hold. It is known as *linear classification rule* [53].

## 2.12 Spearman's rank correlation coefficient

Spearman's rank correlation coefficient, indicated with the Greek letter  $\rho$  (rho) is a nonparametric measure of rank correlation. In statistics it is widely used to assess how well the relationship between two variables is described as a monotonic function. It can be defined as 32:

$$\rho = \frac{6 \sum d_j^2}{n(n^2 - 1)} \quad (32)$$

Spearman's correlation can assume different values from -1 to 1, in particular:

- **0-0.19** the correlation is **very weak**
- **0.20-0.39** the correlation is **weak**
- **0.40-0.59** the correlation is **moderate**
- **0.60-0.79** the correlation is **strong**
- **0.80-1.00** the correlation is **very strong**

If the value is positive (+1) it means that there is a perfect positive correlation, if the value is negative (-1) it means that there is a perfect negative correlation and the 0 value means no correlation between the two variables (Figure 34). This coefficient was calculated in matlab by means of the MATLAB function "*corr*" which gives a further information about the *p-value* between the two variables with the null hipotesis that there is no correlation between them.

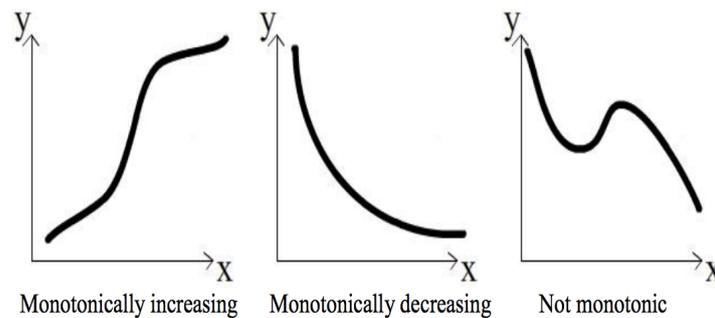


Figure 34: On the left positive correlation, in the middle negative correlation, on the right no significant correlation

### 3 Results

In this section, the signals recorded are showed as well as all the results about the methodes explained before are presented

#### 3.1 Single Differential EMG Signals

Signals recorded from muscles during different tasks are showed here. In particular way we can appreciate how the muscle activations are not spatially equal on the muscle surface. Furthermore, the muscles activations within the same muscle are different not just in amplitude but also the shape of the action potentials changes spatially in the majority of cases. Signals are plotted for all channels of the matrices so that it can be easier to see and to understand these differences. Since the information recorded spatially is not always the same, it demostratrates that using High Density-EMG the redundancy can be considered to be minimal and by means of a big number of electrodes we can get the entire information. In Figure 35 we can see the single differential EMG signals recorded by each channel of a matrix of 64 electrodes from Semitendinosus and Biceps Femoris muscles during a gait cycle 1.8s long (Figure 35 A). In the expanded view (Figure 35 A.1) it is evident the different behaviour of the two muscles. Infact the signal shape on the left side of the matrix is different from that one in the right side. The four matrix columns on the left are fixed on semitendinosus muscles, instead the other four on the Biceps Femoris muscle.

Let's see now how within a muscle during the same task the muscle activation is different among the channels. In Figure 36 we have the single differntial EMG signals recorded by a matrix of 64 electrodes during a gait cycle 1.8s long from the medial portion of Gastrocnemius muscle (Figure 36 A). We can see (Figure 36 A.1) how columns 1,2,3,4,5 have a bigger amplitude than columns 6, 7, and 8, which means that the area taken in to account on the left side of the muscle portion (red) is more activated during the time interval that we expanded. As well as the amplitude also the shape of the signals seems to be different between the two zones.

Something similar can be observed also in Figure 37 which shows the single differential EMG signals recorded by a matrix of 32 electrodes from Rectus Femoris muscle during a sit to stand task 1.8s long (Figure 37 A). For example, in the expanded view of the signal portions selected (Figure 37 A.1) we can observe that the central area of the matrix corresponds to the most activated muscle zone. As well as the amplitude also the shape of the signals seems

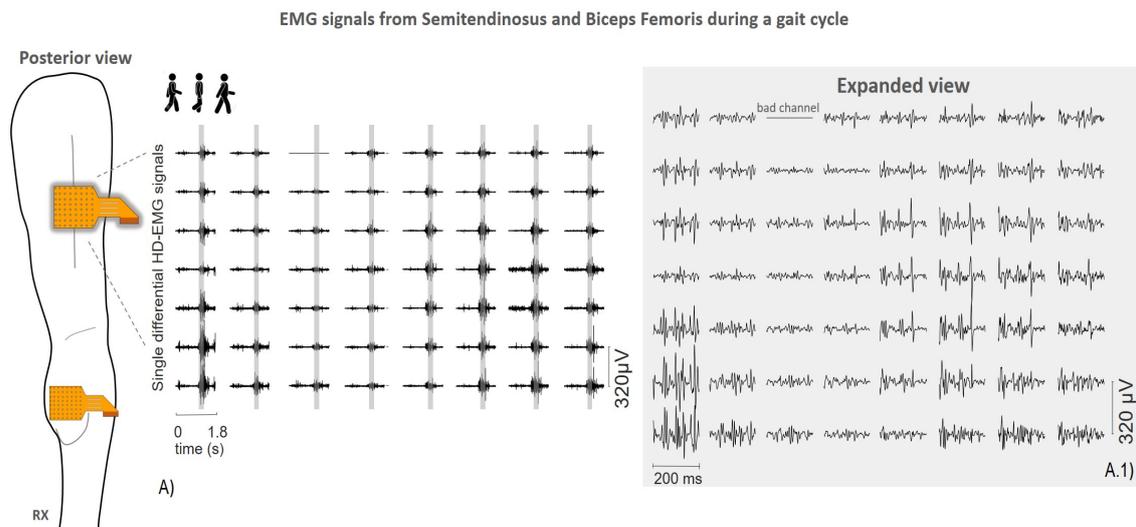


Figure 35: A) SD HD-EMG signals recorded from Semitendinosus and Biceps Femoris muscles. A matrix of 64 electrodes was used. A.1) expanded view of 200ms of signal in which the different muscle activities between the two muscles are evident. The tasks duration is 1.8s.

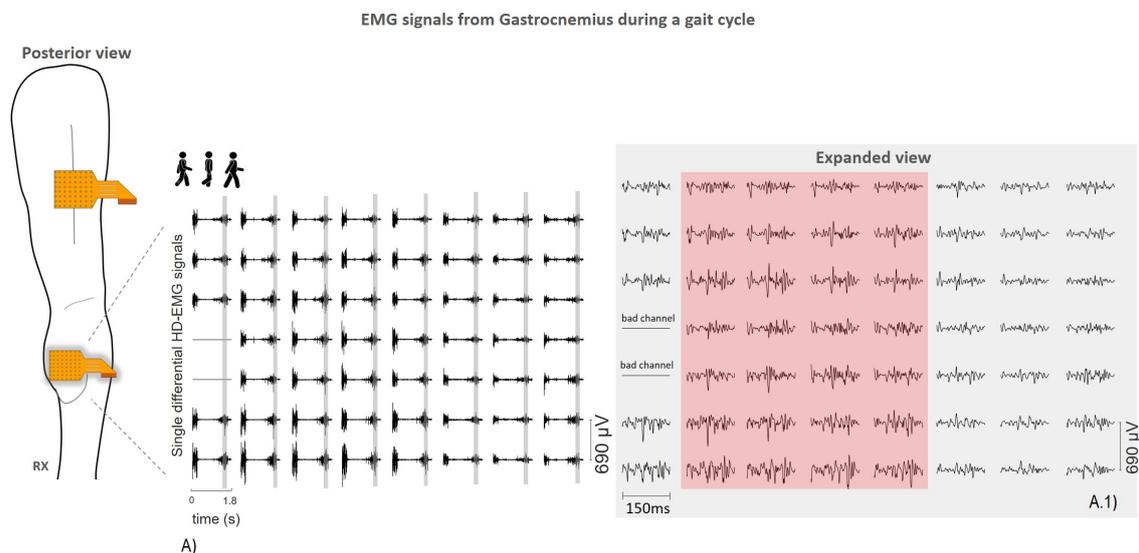


Figure 36: A) SD HD-EMG signals recorded from Gastrocnemius medialis muscle. A matrix of 64 electrodes was used. A.1) expanded view of 150ms of signal in which the different activation areas can be observed. In red the most activated one. The tasks duration is 1.8s.

to be different comparing for instance column 3 with column 4.

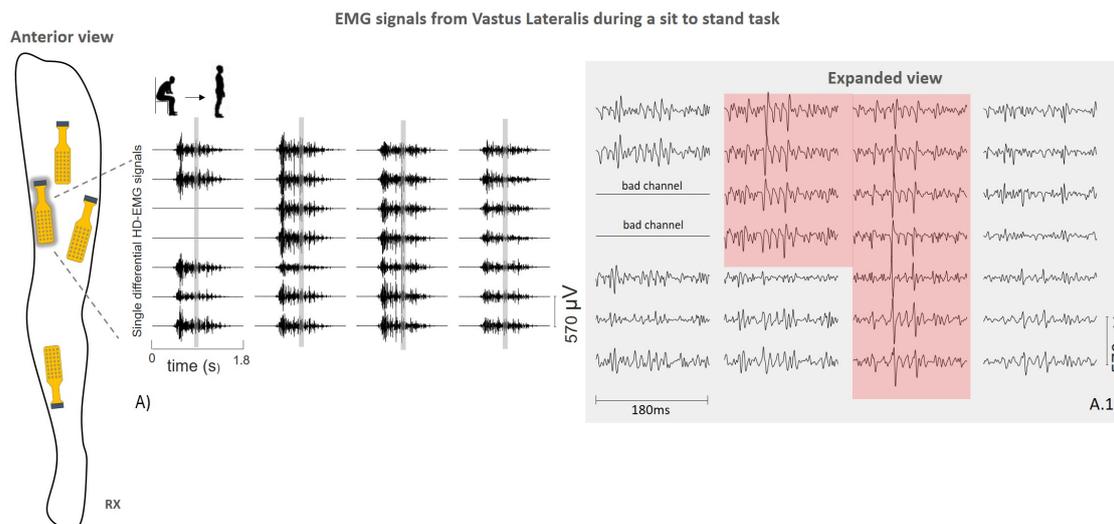


Figure 37: A) SD HD-EMG signals recorded from Vastus medialis muscle. A matrix of 32 electrodes was used. A.1) expanded view of 180ms of signal in which the different activation areas can be observed. In red the most activated one. The tasks duration is 1.8s

Analyzing the activity of a muscle by comparing the signals recorded during two different tasks, we can observe that, as we expected, the muscle has a different activation periods among tasks. Very important is that there is not only difference of amplitude and shape among channels within a muscle during the same task. Infact the same muscle has a different spatial behaviour from one task to another so that the most activated area is not always the same. In Figure 38 the single differential EMG signals recorded from Vastus Lateralis muscle during stair descending and stand to sit task 1.8s long, are showed. A matrix of 32 electrodes was used. We can see that the activation periods depend on the task that the subject does (Figure 38 A, B). In the stair descending task (Figure 38 A), within the portion time selected to be expanded (Figure 38 A.1), the most activated area is the one on the left side (red) of the matrix, but also the shape of the action potentials are different going from the first two columns to the 3rd and 4th columns. Instead, during the stand to sit task (Figure 38 B), in the expanded view (Figure 38 B.1) we can observe that the most activated area (red) is in correspondence of the first column, the inferior half of the second one and almost the entire fourth column. A big difference in shape can be seen going from one column to another.

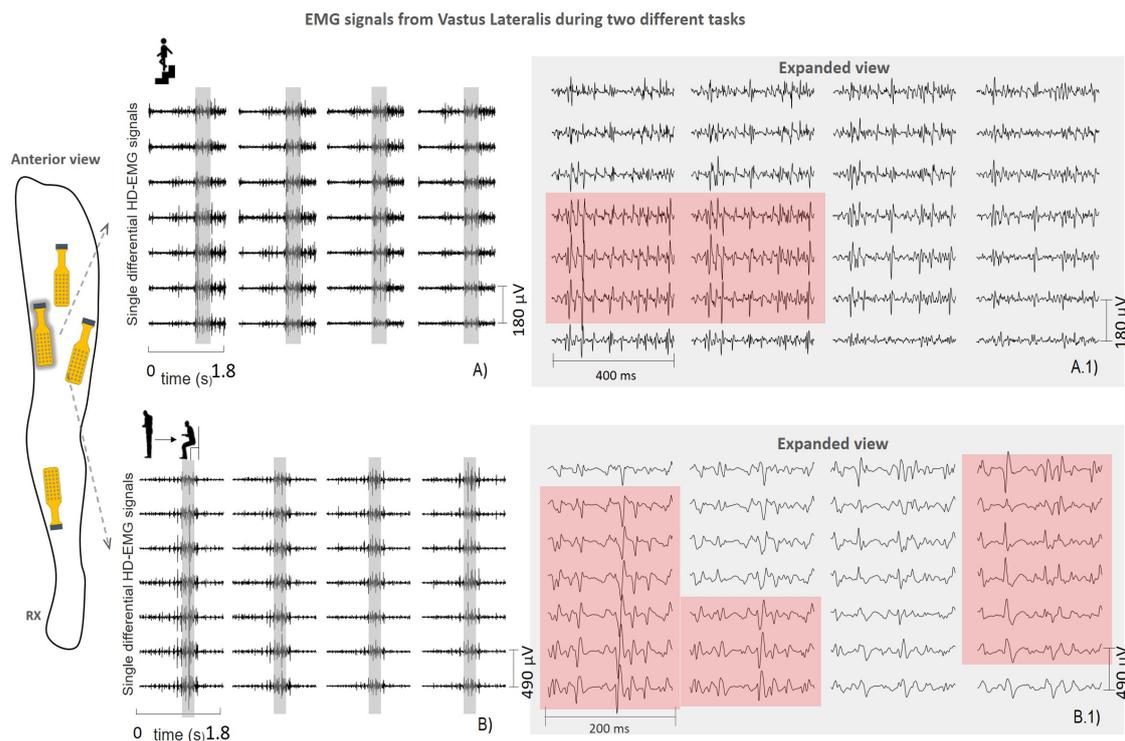


Figure 38: SD HD-EMG signals recorded from Gastrocnemius medialis muscle during two tasks: A) stair descending and B) stand to sit. A matrix of 32 electrodes was used. A.1) expanded view of 400ms of signal during the stair descending task. B.1) expanded view of 200ms of signal during the stand to sit task. In A.1) and B.1) the different activation areas can be observed. In red the most activated ones. The tasks duration is 1.8s.

Analyzing Figure 39, the single differential EMG signals recorded from Rectus Femoris during both gait cycle and stand to sit tasks 1.8s long are plotted. A matrix of 32 electrodes was used. We can see, how right it is, that the activation bursts are different between both tasks (Figure 39 A, B) but analyzing carefully the portion time expanded we can see that with the same muscle but during different tasks the most activated area (red) is not always the same. In fact during the gait cycle the most activated zone is the lower half of the matrix. There is no evident difference in shape among the columns but it is not completely true if you compare the superior part with the inferior part of the matrix. Talking about the stand to sit expanded view, we can see how the signal with the higher amplitude (red) is in the left half of the matrix.

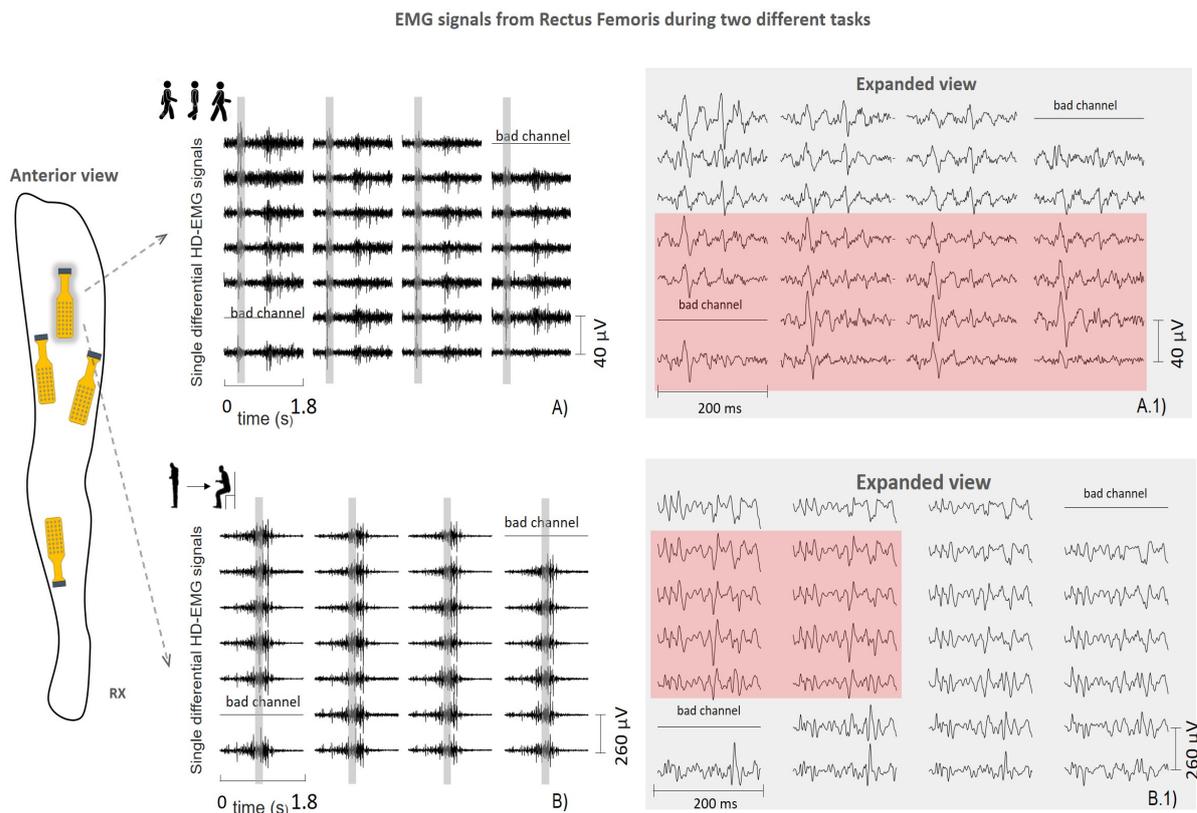


Figure 39: SD HD-EMG signals recorded from Rectus Femoris muscle during two tasks: A) gait cycle and B) stand to sit. A matrix of 32 electrodes was used. A.1) expanded view of 200ms of signal during the gait cycle task. B.1) expanded view of 200ms of signal during the stand to sit task. In A.1) and B.1) the different activation areas can be observed. In red the most activated ones. The tasks duration is 1.8s.

Let's see now how the muscles have a different activation instants in base on the task. In figure 40 we can see the single differential signals recorded from the Rectus Femoris muscle during 6 different task. The matrix used here is made by 32 electrodes. During the experiment 7 tasks were performed, but it was decided to show only one between the two rest periods; in particular, the rest period in upright position is not plotted. The tasks that we can find in this figure are: stand to sit (Figure 40A), stair ascending Figure (40B), a gait cycle (Figure 40C), star descending (Figure 40D), sit to stand (Figure 40E) and rest sitting (Figure 40F).

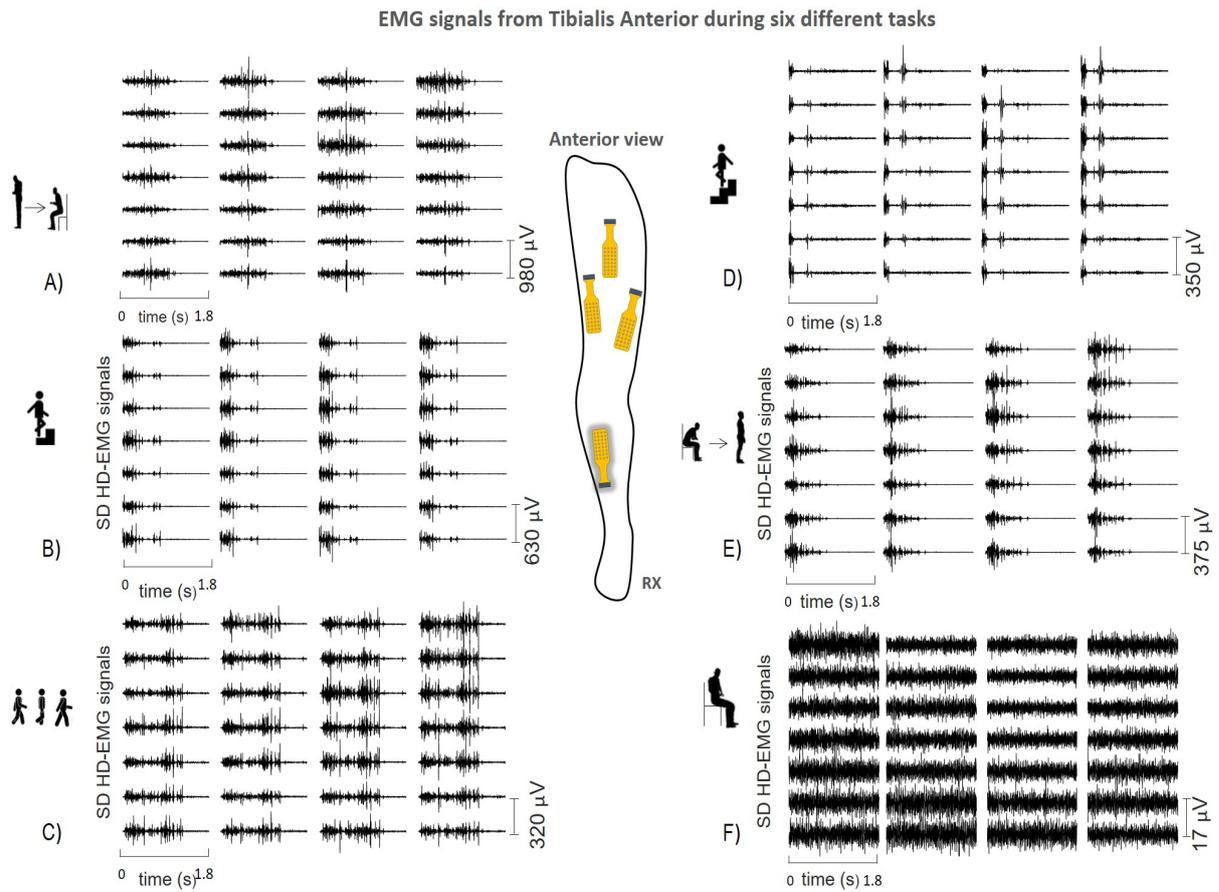


Figure 40: SD HD-EMG signals recorded from Tibialis Anterior muscle during six tasks: A) stand to sit, B) stair ascending, C) gait cycle, D) stair descending, E) sit to stand, F) rest sitting. A matrix of 32 electrodes is used. The tasks duration is 1.8s

### 3.2 Computational Time

First of all, all the features were evaluated in terms of computational time in order to exclude from the study those features that doesn't allow to detect a movement in real time, which means 300ms maximum from the movement intention to the device motion. In Table 1 the average (among all windows of all repetitions of all classes) computational time to calculate each feature on a signal 300ms long for all channels, is indicated. In Figure 41 and Figure 42 the bar graph about the computational time with and without features in frequency doimain are plottet for a better visualization. Furthermore, computational time was calculated on a MacBook Pro (2.3 GHz Intel Core i5), using MATLAB R2017b.

Table 1: Computational time

<i>Feature</i>	<i>Computational Time (s)</i>
VAR	8.2
IEMG	8.0
WL	9.4
MNF	1614
RMS	14
MAV	12.7
MDF	1593
MAV1	16.1
MAV2	16.9
TM3	69.9
TM4	70
TM5	69.9
LogDetect	24
DASDV	33.8

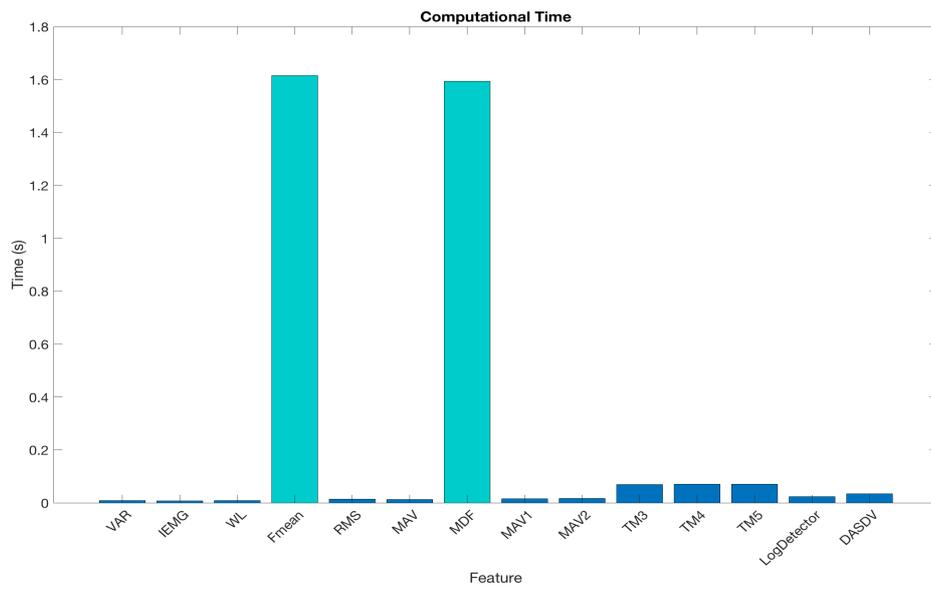


Figure 41: Bar plot about the computational time of all 14 features. In green both features in frequency domain with an high computational cost.

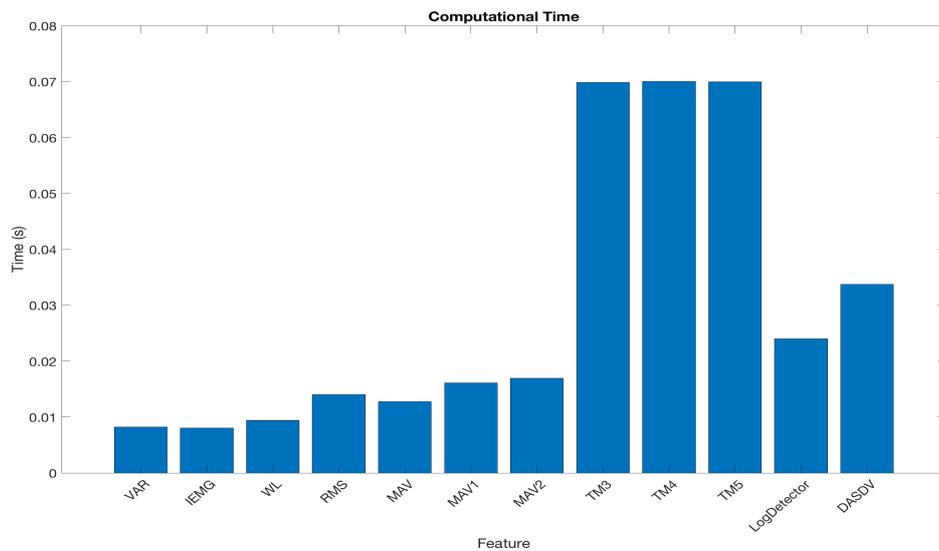


Figure 42: Bar plot about the computational time of all features except those in frequency domain.

It is easy to see how both MNF and MDF features need too much time, about 1.61s and 1.59s respectively, to be calculated for all channels on a windows 300ms long. So much time doesn't allow the use of those feature in real time applications as well as it would be dangerous to wait for all that time before helping the subject in movements.

### 3.3 J index performance

Here are the Boxplot about the twelve features taken into account in this work after a previous selection about computational time. The boxplot of each feature includes information about the J values considering together one repetition per class at a time (for example the first J value is about all the first repetitions of all the classes, the second J value is about all the second repetitions of all the classes, ecc. ). It is repeated for the fifteen repetitions. In particular, since the features evaluation using the J index is pretty similar among subject, only the most significative ones are reported in Figure 43 and Figure 45. In each boxplot the central red line indicates the median, the bottom and top edges indicate the 25th and 75th percentiles, respectively. The whiskers are extended to the most extreme data points not considered outliers. The outliers are indicated by means of a red cross. In Figure 44 and Figure 46 you can see the respective average J value calculated on the J values of 15 repetitions. An additional plot in Figure 47 shows the behavior of the average J value about the 15 features for all the subjects.

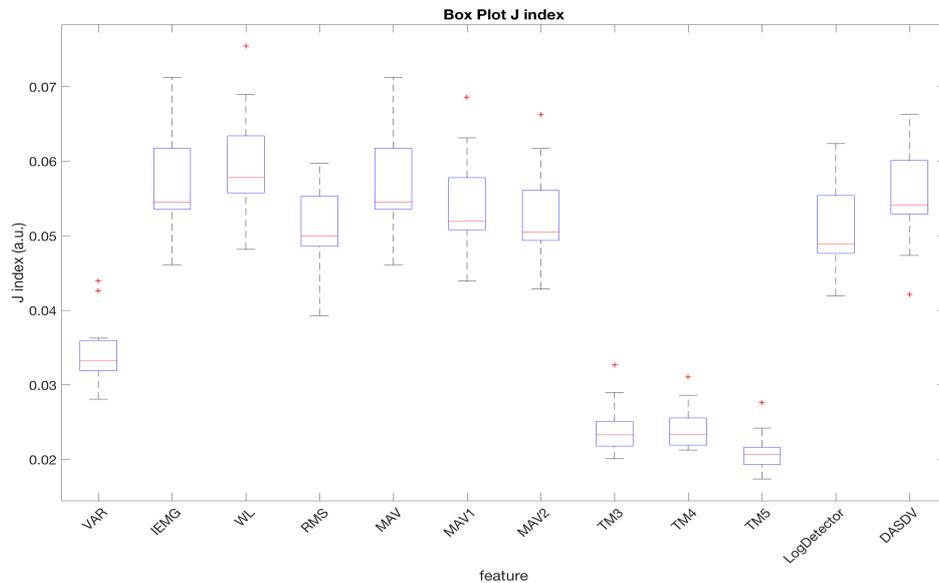


Figure 43: Boxplot about J index of 12 features from one subject.

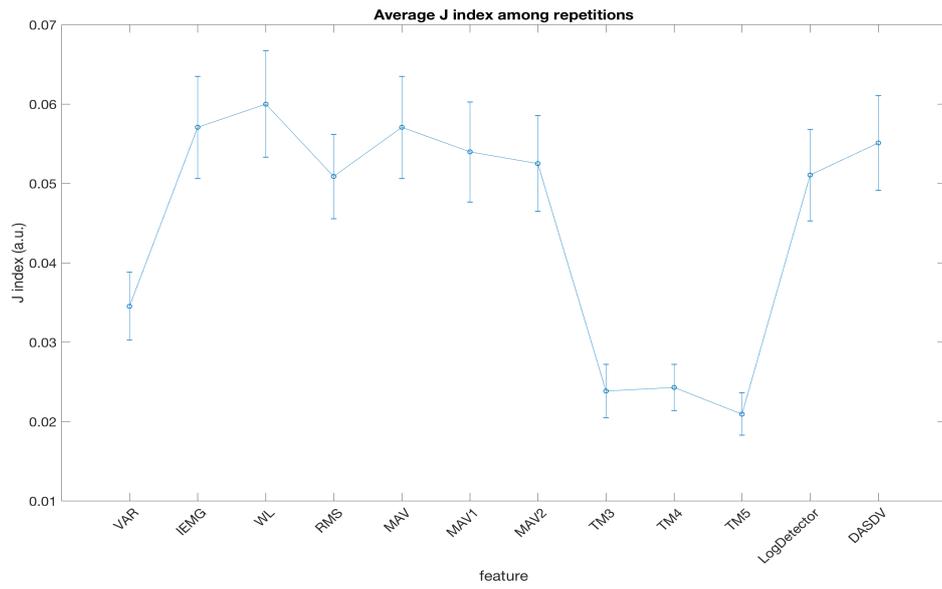


Figure 44: Average value of J index of 12 features from one subject (Error bar: standard deviation on fifteen repetitions).

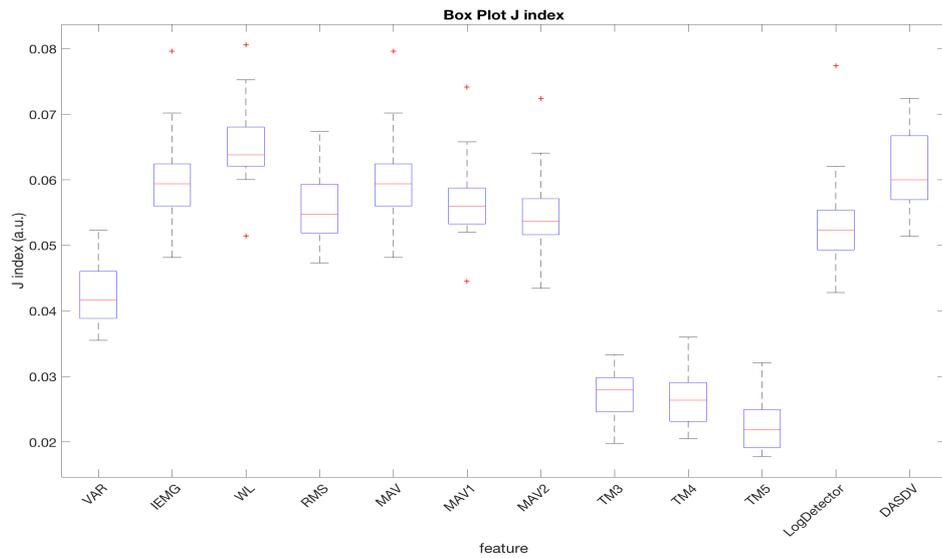


Figure 45: Boxplot about J index of 12 features from one subject.

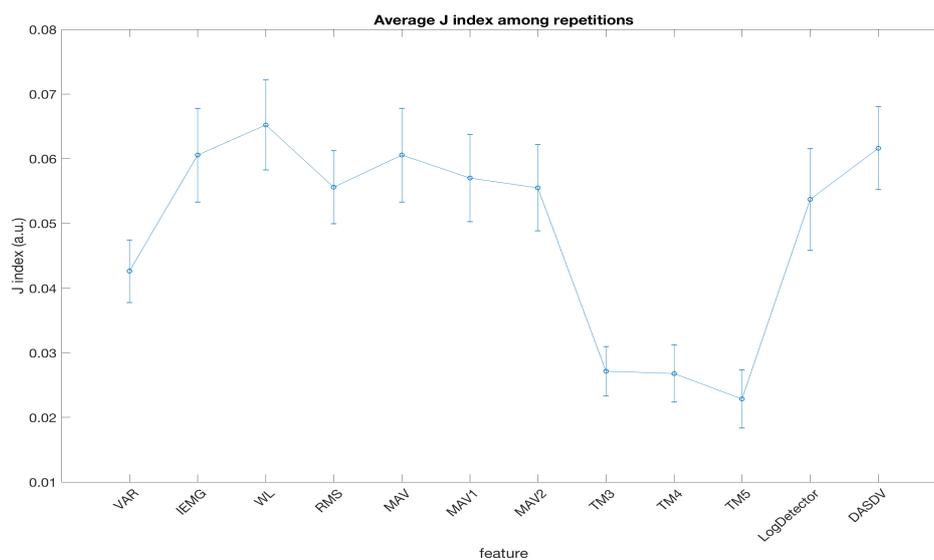


Figure 46: Average value of J index of 12 features from one subject (Error bar: standard deviation on fifteen repetitions).

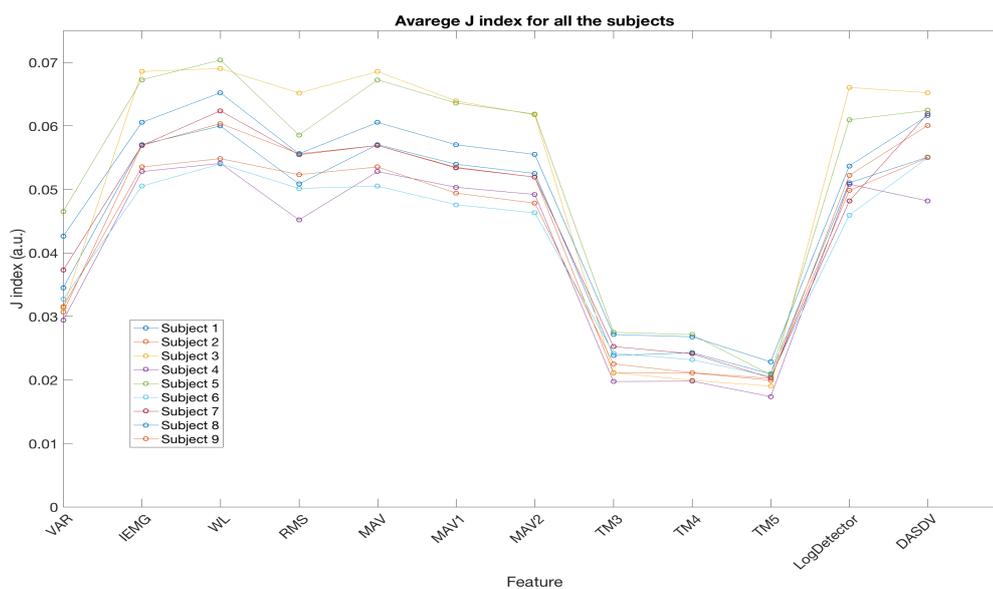


Figure 47: Average J index behavior about 12 features in 9 subjects

As we can see in Figure 43, Figure 45, Figure 44 and Figure 46 in both subjects, it is immediate to appreciate that the feature with the highest value of J is Waveform Length (WL) and the lowest J index belongs to Variance (VAR), Absolute value of 3rd (TM3), 4th (TM4) and 5th (TM5) temporal moment. Instead, in Figure 47 you can observe the behaviour of J index in each of 9 subjects. What it very clear is that the behaviour is pretty similar among subjects which means that there is a repeatability of the results as well as the best feature is always WL and the worst VAR, TM3, TM4, TM5.

### 3.4 LDA Classification performance

Here are the Boxplot about the 15 features taken into account in this work. The boxplot of each feature is built using the 15 Classification rates found from the 15 combinations of the leave one out applied to the 15 repetitions of each task. In particular, for each combination, 14 repetitions per class were used as training set and one repetition per task as test set. Since the features evaluation using LDA classifier is pretty similar among subject, only the most representative ones (2 subjects) are reported in Figure 48 and Figure 50. In each boxplot the central red line indicates the median, the bottom and top edges indicate the 25th and 75th percentiles, respectively. The whiskers are extended to the most extreme data points not considered outliers. The outliers are indicated by means of a red cross. In Figure 49 and 51 there is the average accuracy value for each feature calculated on the fifteen repetitions tested and an error bar about the respective standard deviations. An additional plot in Figure 52 shows the behavior of the average accuracy about the 15 features for all the nine subjects. In Table 2 there are the average accuracy values of each feature on all the subjects; also the standard deviation is indicated.

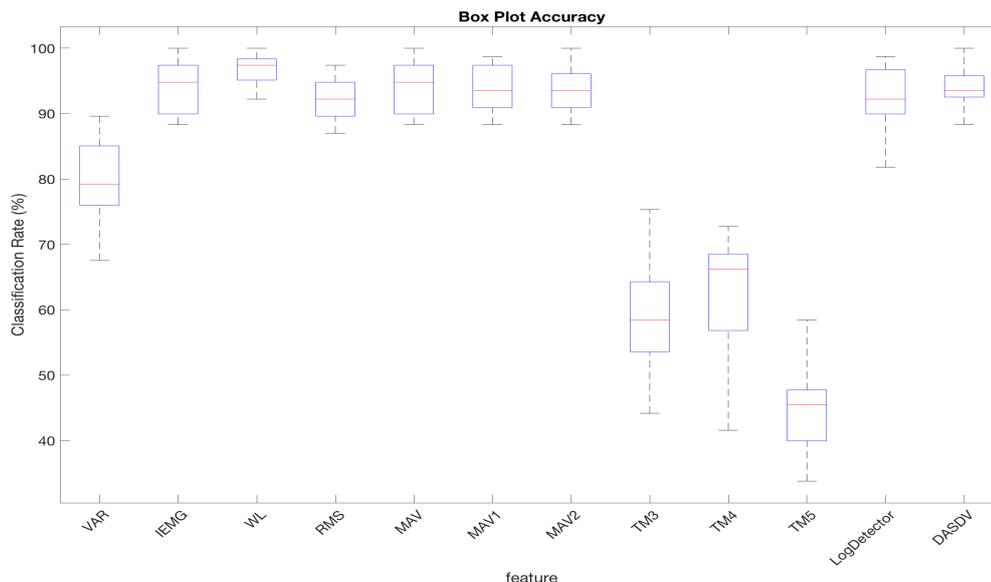


Figure 48: Boxplot about classification of 12 features from one subject.

As we can see in Figure 48, Figure 49, Figure 50 and Figure 51, in both subjects, it is immediate to appreciate that the feature with the highest value of classification accuracy is Waveform Length (WL) and the lowest one belongs to Variance (VAR), Absolute value of 3rd (TM3), 4th (TM4) and 5th (TM5) temporal moment. Instead, in Figure 52 you can observe the behaviour of classification accuracy of all the nine subjects. What it very clear is that the behaviour is pretty similar among subjects which means that there is a repeatability of the results as well as the best feature is always WL and the worst VAR, TM3, TM4, TM5.

Table 2: Average Accuracy on fifteen repetitions.

	Subject 1	Subject 2	Subject 3	Subject 4	Subject 5	Subject 6	Subject 7	Subject 8	Subject 9
VAR± STD	79,82±7,0	93,76±3,7	94,89±3,8	77,22±5,8	87,96±8,5	96,19±3,6	77,74±3,7	76,27±3,6	79,82±7,5
IEMG±STD	94,02±4,0	99,13±1,0	98,87±1,2	94,54±3,3	98,44±4,0	99,48±1,3	96,62±1,9	94,63±2,2	94,97±3,9
WL±STD	96,45±2,5	99,22±1,2	99,91±0,3	96,53±2,9	98,70±4,3	99,82±0,6	97,14±2,1	96,19±2,2	98,09±1,8
RMS±STD	92,29±3,4	99,22±1,1	98,44±1,1	93,24±2,2	98,78±3,3	99,56±0,9	95,84±1,7	92,12±2,7	94,19±3,9
MA±STD	94,02±4,1	99,13±1,0	98,87±1,2	94,54±3,3	98,44±4,0	99,48±1,3	96,62±1,9	94,63±2,2	94,97±3,9
MAV1±STD	94,02±3,3	99,04±1,2	98,78±1,3	93,85±3,4	98,70±2,9	99,30±1,7	96,53±1,6	93,16±2,5	95,06±3,6
MAV2±STD	93,77±3,4	99,22±1,18	98,78±1,25	93,59±2,88	98,18±3,73	99,30±1,7	95,67±2,9	92,90±2,3	95,63±3,5
TM3±STD	58,26±8,1	73,85±3,9	72,55±4,0	49,78±7,3	56,45±7,7	62,42±6,7	57,31±5,3	53,67±5,8	45,62±6,2
TM4±STD	62,07±9,0	76,62±6,9	83,29±2,3	52,98±4,7	64,15±7,4	73,93±8,1	59,30±6,8	50,12±5,7	46,06±6,7
TM5±STD	45,28±7,2	64,24±6,7	57,14±6,1	36,70±7,7	42,33±5,7	51,08±6,2	38,00±6,1	40,77±3,8	37,74±5,8
LogDet±STD	92,20±4,9	97,74±2,0	98,87±1,4	90,56±4,4	98,18±4,5	99,13±1,8	93,41±3,4	91,42±3,0	92,81±4,8
DASDV±STD	94,02±2,8	99,04±1,7	99,65±0,6	96,01±3,6	98,96±3,6	99,74±1,0	95,58±2,0	94,63±2,4	95,15±2,7

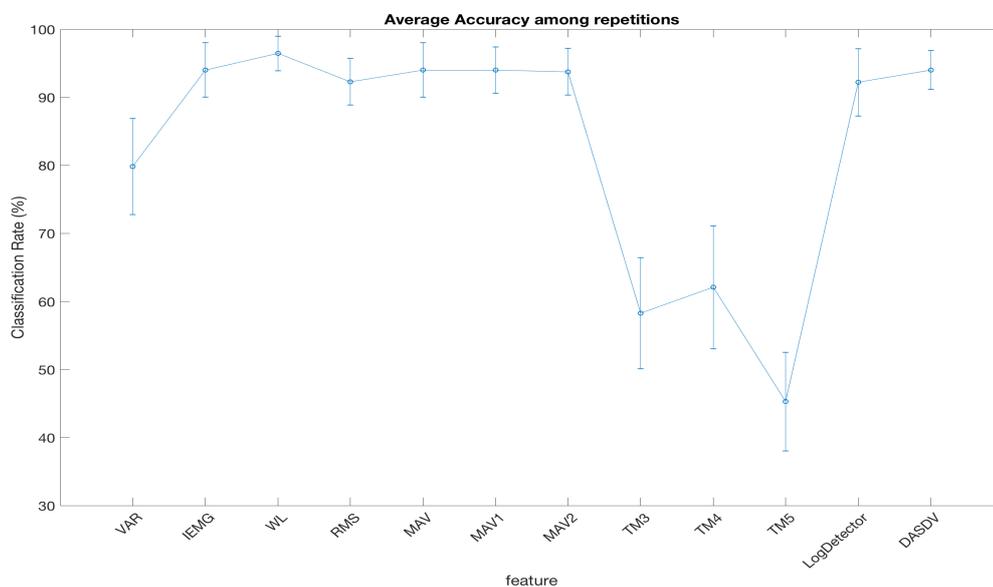


Figure 49: Average accuracy about the 12 features from one subject (Error bar: Standard deviation on fifteen repetitions tested).

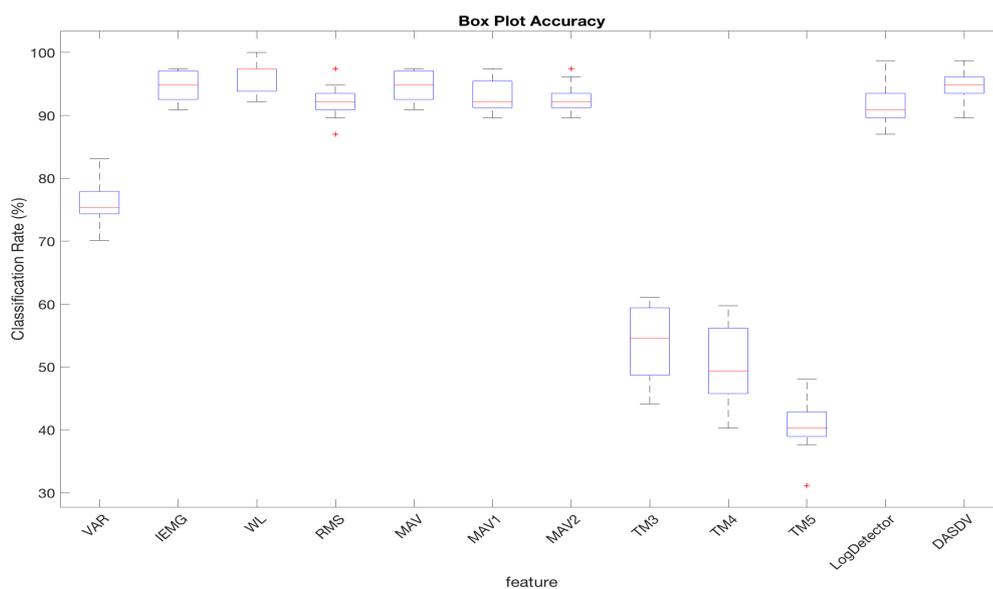


Figure 50: Boxplot about classification of 12 features from one subject.

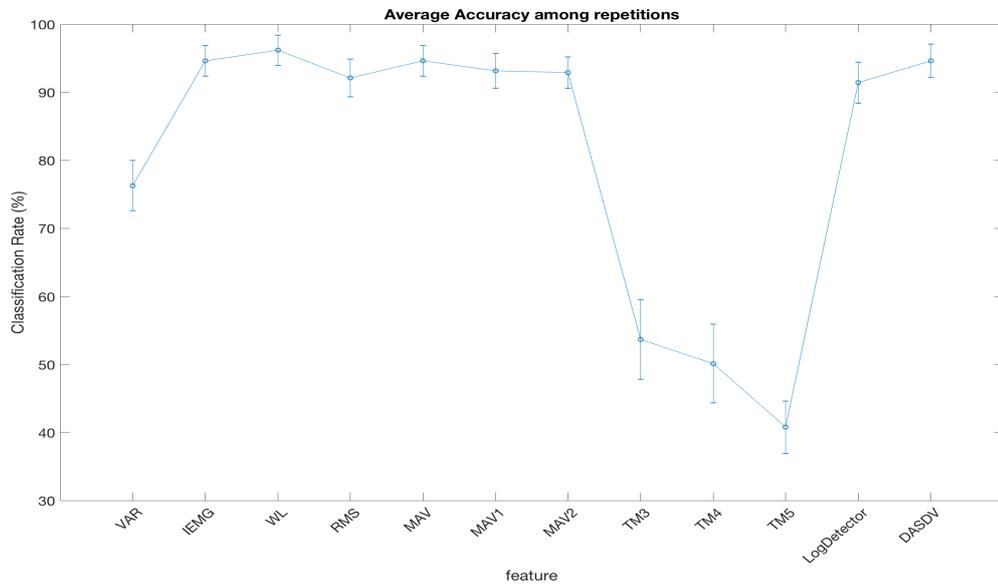


Figure 51: Average accuracy about the 12 features from one subject (Error bar: Standard deviation on fifteen repetitions tested)

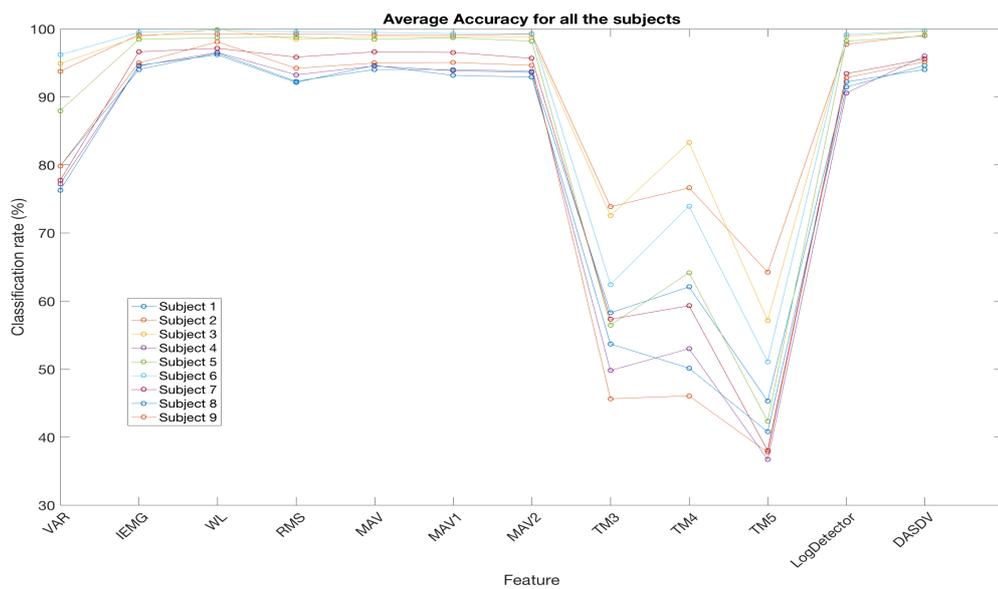


Figure 52: Average Accuracy behavior about 12 features in 9 subjects

### 3.5 Comparisons: Classification - J index

In this section we compare the results obtained from the J index with the results of classification accuracy. To quantify the correlation between both models, Spearman's rho was used and in Table 3 we can see the value of rho coefficient and the p-value for each subject calculated between the average J values and the average LDA accuracy, among the fifteen repetitions, for all the features .

Table 3: Spearman's rank correlation coefficient and p-value between average J index and average accuracy per each subject. Null Hypotesis of no correlation.

	<i>Spearman's rho</i>	<i>p-value</i>
<b>Subject 1</b>	0.933	<0.01
<b>Subject 2</b>	0.773	<0.01
<b>Subject 3</b>	0.926	<0.01
<b>Subject 4</b>	0.844	<0.01
<b>Subject 5</b>	0.714	<0.01
<b>Subject 6</b>	0.963	<0.01
<b>Subject 7</b>	0.879	<0.01
<b>Subject 8</b>	0.986	<0.01
<b>Subject 9</b>	0.879	<0.01

In particulare you can see that for seven subjects, (Subject1, Subject3, Subject4, Subject6, Subject7, Subject8, Subject9), the rho'coefficient is higher than 0.8 which means that the correlation between both methods is "very strong" and for two subjects (Subject2, Subject5) the correlation is considered as "strong" because the rho value is between 0.6 and 0.79. For every subject we find a  $p\text{-value} < 0.01$  rejecting the null hypotesis that both approaches are scorrelated. Hence there is a very strong correlation between the use of the Jindex and the LDA classifier. In Figure 53 ( $\rho=0.773$ ) and Figure 54 ( $\rho=0.986$ ) there are the classification accuracy-J Index trend for all the features about two subjects: The best one and the second last in terms of rho value . In both subjects you can see that when the classification accuracy of one feature increases, also the J value of that feature increases. The relation between both approaches is not complitely linear but it is strongly monotonically increasing.

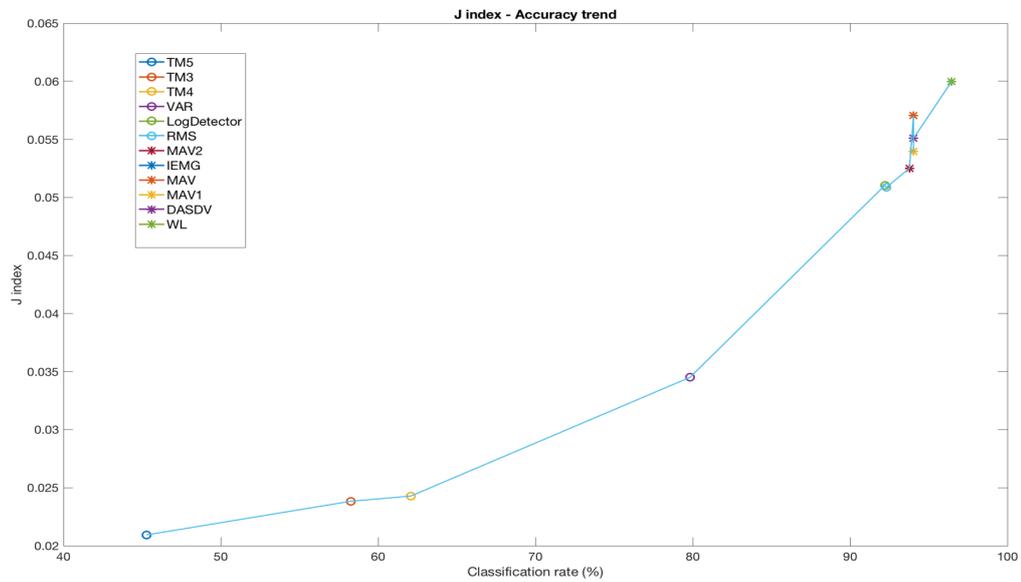


Figure 53: Trend between average classification accuracy (X axis) and average J index(Y axis) among repetitions of Subject 2 in Table 3

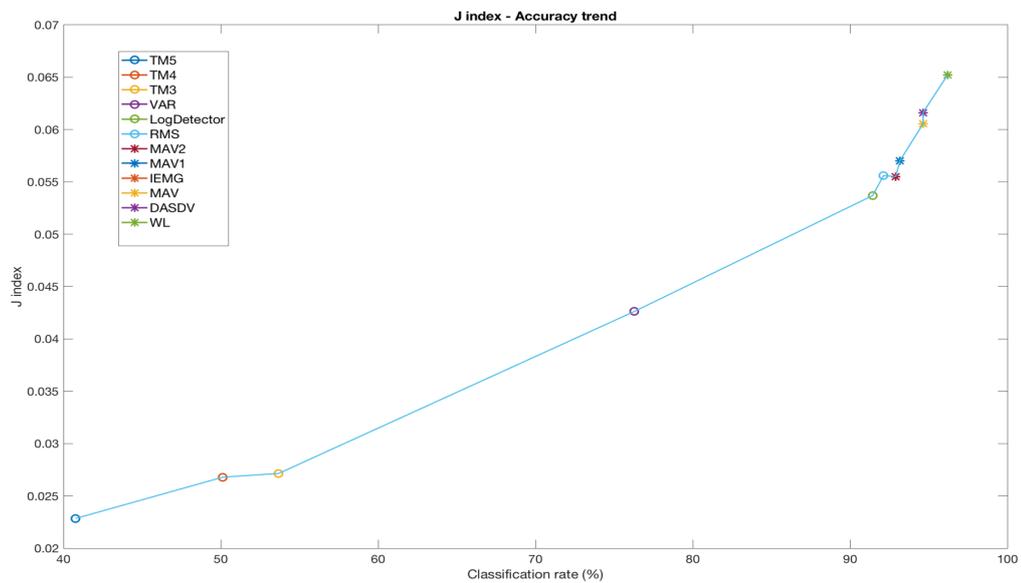


Figure 54: Trend between average classification accuracy (X axis) and average J index(Y axis) among repetitions of Subject 8 in Table 3

In Figure 55 there is the average behaviour among subject of all the features about the J index values and the LDA accuracy. You can see how the trend of the average J index is pretty similar to the trend of the average accuracy. In fact in both approaches Waveform Length (WL) is the most performing feature in terms of J value and accuracy. At the same time the features less performing are the Variance (VAR), the Absolute value of 3st (TM3), 4th(TM4) and 5th(TM5) order.

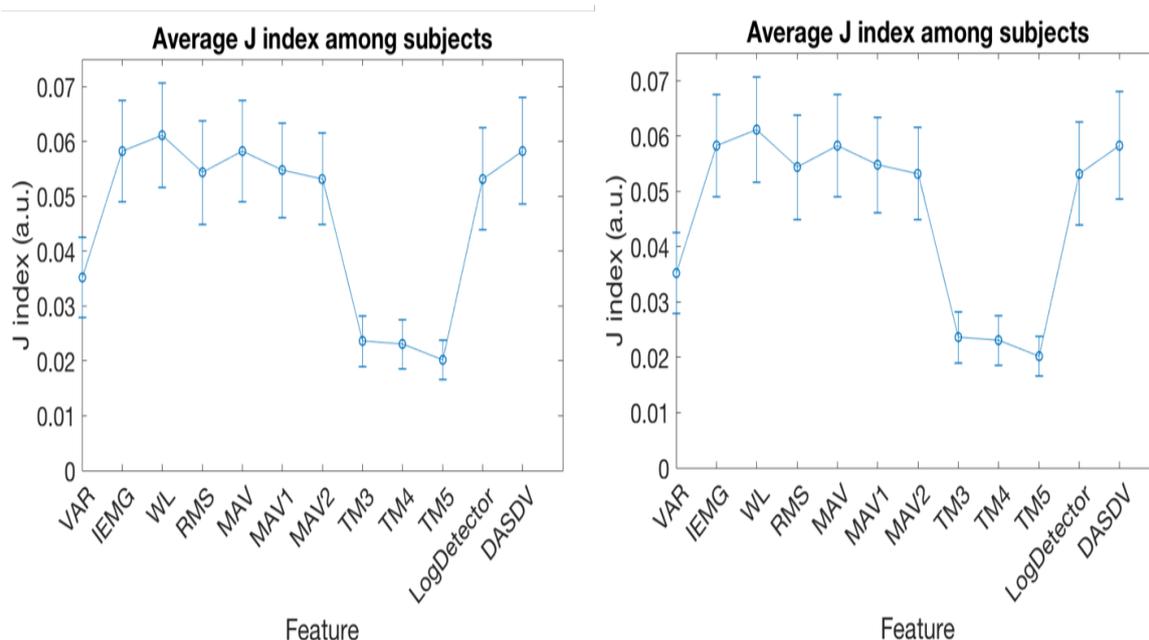


Figure 55: On the left there is the average J index and on the right the average classification accuracy among subjects. (Error bar: standard deviation)

In Figure 56 the average J index and the average classification accuracy among subject (Figure 55) are used to plot their trend. The rho coefficient between them is 0.984 which means the trend is strongly monotonically increasing and the null hypothesis that says that both approaches are uncorrelated is rejected ( $p\text{-value} < 0.01$ ). In particular, in Figure 57 and Figure 58 there is the feature ranking about J index and LDA classifier. As we can see the two rankings are pretty similar, in fact TM5, TM4 and TM3 are the last performing features and WL ( $J\text{index}=0.0611$ ) is the best one. IEMG ( $J\text{index}=0.0583$ ), MAV ( $J\text{index}=0.0583$ ) and DASDV ( $J\text{index}=0.0582$ ) are 5% less performing than WL. Being both of them equivalent approaches, the results of the proposed methods are validated and for this reason we can choose that best subset of features relying on J index ranking.

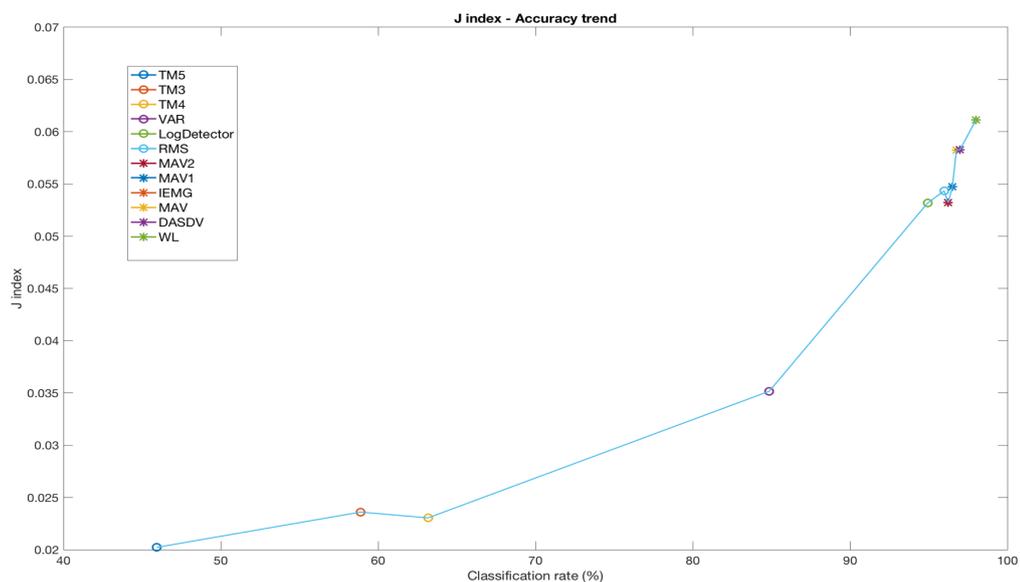


Figure 56: Trend between average classification accuracy (X axis) among subjects and average J index(Y axis) among subject.

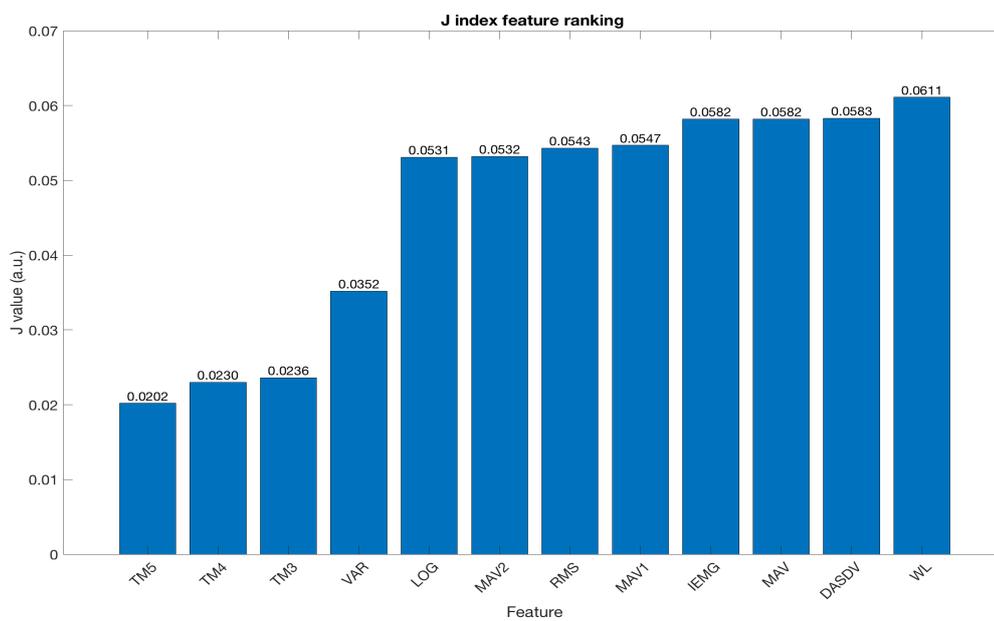


Figure 57: Feature average J index ranking among subjects

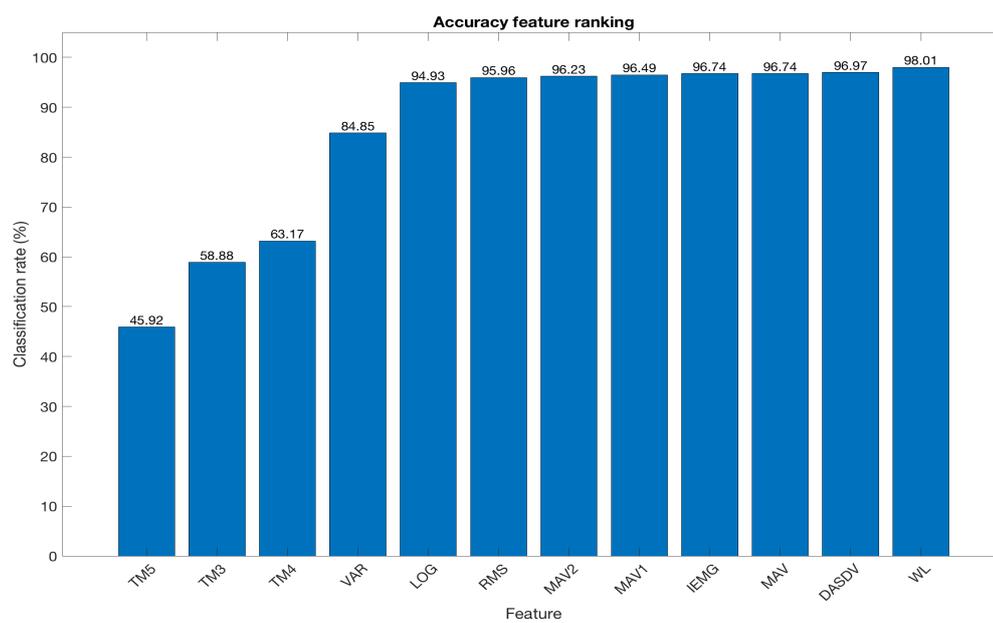


Figure 58: Feature average accuracy ranking among subjects

## 4 Conclusions

Currently, Stroke is one of main causes of serious long term disability, morbidity and mortality. All around the world 2 million people are affected by this pathology and they mainly need rehabilitation in order to achieve functional independence. Respect to technologies with body weight support and treadmill that are widely used, the use of a system with the direct patient involvement allows a faster and better results. For this reason an exoskeleton with a control strategy based on the patient intention is the ideal solution. One of the first step to realize that is the recognition of the patient intentions. To do that a pattern recognition is essential and with this study we try to find the best feature subset that allows the movements recognition. In this project High-Density sEMG, which is more robust than the bipolar technique usually used in myoelectric control applications, is used in order to take into account the whole muscle surface. It recorded as monopolar signal from seven muscles by means of four matrices of 32 channels fixed on Rectus Femoris, Vastus Lateralis, Vastus Medialis Tibialis Anterior, and two matrices of 64 electrodes fixed on Gastrocnemius Medialis and Semitendinosus&Biceps Femoris. After that a remapping algorithm was necessary within each matrix, in base on its orientation on the muscle belly, in order to calculate the single differential signal for each electrode pair. Before processing all the data, a bad channels selection was mandatory so that all the channels without information, because of a low quality electrode-skin contact, could be eliminated. The number of bad channels found in every subject is not significant respect to the total number available. So many channels are used because we want to get all the spatial information from the muscle surface since the muscle has a different spatially behaviour during each single task, as well as a specific area is activated in different ways from one task to another (Section 3.1). Nine healthy and voluntary subjects participated to the study and they performed fifteen repetitions for seven different tasks: sit to stand, stand to sit, stair ascending, stair descending, a gait cycle, rest in upright position and rest in sitting position. In particular fourteen features were studied and calculated on epochs 300ms long with overlap 50%. First of all, computational time was calculated per each feature and both features in frequency domain Mean Frequency (MNF) and Median Frequency (MDF) were excluded from the study because they needed around 1.5s for one windows on all channels. Computational time is very important to take a decision in real time: if it is longer than 300ms it is impossible to actuate that. After that, to evaluate feature quality, the filter approach based on J index which gives information about class separability in terms of distance between

classes and spread within each class, was used. This method was compared and validated with the wrapper approach which use the classification accuracy of an LDA classifier. In fact, by means of Spearman's rho coefficient was evaluated a strong correlation between both approaches for two subjects ( $0.6 < \rho < 0.79$ ) and a very strong correlation in seven subjects ( $\rho > 0.8$ ) as you can see in Table 3. Looking at the average trend among all the subjects for the J index and for the classification accuracy, the correlation between the filter and the wrapper approach is very strong with a  $\rho = 0.984$  and a  $p\text{-value} < 0.01$  rejecting the null hypothesis of no correlation. In Figure 56 you can see that the relation is monotonically increasing, which means that when the classification accuracy of one feature increases also the J value increases. Looking at the feature ranking obtained from both approaches considering the average J index (Figure 57) and the average classification accuracy (Figure 58) among all the subjects you can see that they are pretty equal and it is demonstrated that the proposed method is a good method of feature selection and its strength is the simplicity of implementation and computation. To conclude, the best feature subset is found by means of J index and as we can see in Figure 57 and in particular TM5 ( $J\text{index} = 0.0202$ ), TM4 ( $J\text{index} = 0.0230$ ) and TM3 ( $J\text{index} = 0.0236$ ) are the worst performing features and WL ( $J\text{index} = 0.0611$ ) is the best one. Moreover, IEMG ( $J\text{index} = 0.0583$ ), MAV ( $J\text{index} = 0.0583$ ) and DASDV ( $J\text{index} = 0.0582$ ), that are only 5% less performing than WL, can be considered for the creation of a feature vector in combination with WL.

#### 4.1 Future work

In future studies it is interesting to explore a bigger set of features or new features made by the combination of those just studied, for example combining IEMG, MAV and DASDV with WL, in order to be able to see if new features could have better performance. It is interesting also to see how the Jindex works compared to other classifiers and not just with the Linear Discriminant Analysis (LDA) like in this work even if it depends on which classifier will be used to classify movements during real-time applications. A big challenge would be to apply this work on after stroke patients in order to see if the correlation between J index and LDA classifier continues to be true and if the best feature is still Waveform Length (WL) like in the healthy subjects studied here. A further work to investigate is to check if WL continues to be the best feature also combining this project to a previous channel selection algorithm, in order to take into account only the most relevant channels among those recorded with the matrices.

---

## References

- [1] Skeletal muscle structure. [Online]. Available: <https://www.dreamstime.com/stock-illustration-structure-skeletal-muscle-each-fiber-has-many-bundles-myofilaments-each-bundle-called>
- [2] W. R. Frontera and J. Ochala, "Skeletal muscle: a brief review of structure and function," *Calcified tissue international*, vol. 96(3), pp. 183–195, March 2015.
- [3] The sliding filament theory of muscle contraction. [Online]. Available: <http://www.thealevelbiologist.co.uk/the-sliding-filament-theory-of-muscle-contraction>
- [4] A nervous journey. [Online]. Available: <https://askabiologist.asu.edu/neuron-anatomy>
- [5] Neuromuscular anatomy and physiology. [Online]. Available: <https://aneskey.com/neuromuscular-anatomy-and-physiology/>
- [6] Endoneurium. [Online]. Available: <http://craftbrewswag.info/endoneurium/>
- [7] Tucci kinesiology. [Online]. Available: <http://tuccikinesiology.weebly.com/muscle-twitch-graph.html>
- [8] Yoganatomy. [Online]. Available: <https://www.yoganatomy.com/quadriceps-muscles/>
- [9] Duke anatomy. [Online]. Available: <https://it.pinterest.com/pin/657525614316163008/>
- [10] Fairview. [Online]. Available: <https://www.custompilatesandyoga.com/health/gastrocnemius-learn-your-muscles/>
- [11] The free dictionary. [Online]. Available: [http://medical-dictionary.thefreedictionary.com/\\_/viewer.aspx?path=MosbyMD&name=tibialis-anterior.jpg&url=http%3A%2F%2Fmedical-dictionary.thefreedictionary.com%2Ftibialis%2Banterior](http://medical-dictionary.thefreedictionary.com/_/viewer.aspx?path=MosbyMD&name=tibialis-anterior.jpg&url=http%3A%2F%2Fmedical-dictionary.thefreedictionary.com%2Ftibialis%2Banterior)
- [12] E. Clancy, E. Morin, and R. Merletti, "Sampling, noise-reduction and amplitude estimation issues in surface electromyography," *J Electromyogr Kinesiol.*, vol. 12(1), pp. 1–16, February 2002.
- [13] C. J. De Luca and J. V. Basmajian, *Muscles Alive: Their Functions Revealed by Electromyography*, 1985, ch. Description and Analysis of the EMG Signal, pp. 65–70.
- [14] R. Merletti, A. Botter, A. Troiano, E. Merlo, and M. A. Minetto, "Technology and instrumentation for detection and conditioning of the surface electromyographic signal: State of the art," *Clinical Biomechanics*, vol. 24(2), pp. 122–134, February 2009.

- 
- [15] Ot-bioelettronica. [Online]. Available: [http://www.otbioelettronica.it/index.php?option=com\\_content&view=article&id=67&lang=en](http://www.otbioelettronica.it/index.php?option=com_content&view=article&id=67&lang=en)
- [16] M. Barbero, R. Merletti, and A. Rainoldi, *Atlas of Muscle Innervation Zones*, 2015, ch. Part II, pp. 85–136.
- [17] semanticscholar. [Online]. Available: <https://www.semanticscholar.org/paper/Linear-discriminant-analysis%3A-A-detailed-tutorial-Tharwat-Gaber/2235fb17d9ee62c4fd489a1e29e36e3f97d0778f>
- [18] E. Ostertagova and O. Ostertag, “Methodology and application of one-way anova,” *American Journal of Mechanical Engineering*, vol. 1(7), pp. 256–261, November 2013.
- [19] Enguneering statistics handbook. [Online]. Available: <https://www.itl.nist.gov/div898/handbook/prc/section4/prc438.htm>
- [20] V. Feigin, C. Lawes, D. Bennett, S. Barker-Collo, and V. Parag, “Worldwide stroke incidence and early case fatality reported in 56 population-based studies: a systematic review,” *Lancet Neurol*, vol. 8(4), pp. 355–69, April 2009.
- [21] M. Bortole, A. Venkatakrishnan, F. Zhu, J. Moreno, G. Francisco, J. Pons, and J. Contreras-Vidal, “The h2 robotic exoskeleton for gait rehabilitation after stroke: early findings from a clinical study,” *Journal of NeuroEngineering and Rehabilitation*, vol. 17, pp. 12–54, June 2015.
- [22] A. Moseley, A. Stark, I. Cameron, and A. Pollock, “Treadmill training and body weight support for walking after stroke,” *Cochrane Database Syst Rev.*, vol. 19(4), October 2005.
- [23] M. Gery Colombo, M. Matthias Joerg, B. Reinhard Schreier, and M. Volker Dietz, “Treadmill training of paraplegic patients using a robotic orthosis,” *Journal of Rehabilitation Research and Development*, vol. 37(6), pp. 693–700, November/December 2000.
- [24] J. Veneman, R. Kruidhof, E. Hekman, R. Ekkelenkamp, E. Van Asseldonk, and H. van der Kooij, “Design and evaluation of the lopes exoskeleton robot for interactive gait rehabilitation.” *IEEE Trans Neural Syst Rehabil Eng.*, vol. 15(3), pp. 379–86, September 2007.
- [25] S. Banala, S. Kim, S. Agrawal, and J. Scholz, “Robot assisted gait training with active leg exoskeleton (alex).” *IEEE Trans Neural Syst Rehabil Eng.*, vol. 17(1), pp. 2–8, February 2009.

- 
- [26] A. Kaelin-Lang, L. Sawaki, and L. Cohen, “Role of voluntary drive in encoding an elementary motor memory.” *J Neurophysiol.*, vol. 93(2), pp. 1099–103, February 2005.
- [27] B. Magdo, V. Anusha, Z. Fangshi, M. Juan C, F. Gerard E, P. Jose L, and J. L. Contreras-Vidal, “The h2 robotic exoskeleton for gait rehabilitation after stroke: early findings from a clinical study,” *Journal of NeuroEngineering and Rehabilitation*, vol. 12:54, June 2015.
- [28] H. C. Rubana, B. I. R. Mamun, A. Mohd Alauddin Bin Mohd, A. A. B. Ashrif, C. Kalaivani, and T. G. Chang, “Surface electromyography signal processing and classification techniques,” *Sensor(Basel)*, vol. 13(9), p. 12431–12466, September 2013.
- [29] M. Raez, M. Hussain, and F. Mohd-Yasin, “Techniques of emg signal analysis: detection, processing, classification and applications,” *Biol Proced Online*, vol. 8, pp. 11–35, 2006.
- [30] C. J. De Luca, *Webster J.G., editor. Encyclopaedia of Medical Devices and Instrumentation*. New York, NY, USA: Wiley, 1988, ch. Electromyography, pp. 1111–1120.
- [31] R. Merletti and P. Parker, *Electromyography, Physiology, Engineering, and Noninvasive Applications*. New York, NY, USA: John Wiley and Sons, 2014, ch. 4, p. 84.
- [32] H. Tam and J. Webster, “Minimizing electrode motion artifact by skin abrasion,” *IEEE Trans Biomed Eng.*, vol. 24(2), pp. 134–138, March 1977.
- [33] L. Mesin, R. Merletti, and R. A., “Surface emg: the issue of electrode location.” *J Electromyogr Kinesiol.*, vol. 19(5), pp. 719–727, October 2009.
- [34] S. Viljoen, T. Hanekom, and D. Farina, “Effect of characteristics of dynamic muscle contraction on crosstalk in surface electromyography recordings,” *SOUTH AFRICAN INSTITUTE OF ELECTRICAL ENGINEERS*, vol. 98(1), pp. 18–28, March 2007.
- [35] M. Hemingway, H. Biedermann, and J. Inglis, “Electromyographic recordings of paraspinal muscles: variations related to subcutaneous tissue thickness,” *Biofeedback Self Regul.*, vol. 20(1), pp. 39–49, March 1995.
- [36] H. J. Hermens, B. Freriks, C. Disselhorst-Klug, and G. Rau, “Development of recommendations for semg sensors and sensor placement procedures,” *Journal of Electromyography and Kinesiology*, vol. 10(5), pp. 361–374, October 2000.
- [37] H. Tam and J. G. Webster, “Minimizing electrode motion artifact by skin abrasion,” *IEEE Transactions on Biomedical Engineering*, vol. 24(2), pp. 134–139, March 1977.

- 
- [38] P. Laferriere, A. D. C. Chan, and E. D. Lemaire, "Surface electromyographic signals using a dry electrode," *Medical Measurements and Applications Proceedings (MeMeA)*, pp. 77–80, 30 April-1 May 2010.
- [39] M. Hakonena, P. Harri, and A. Visala, "Current state of digital signal processing in myoelectric interfaces and related applications," *Biomedical Signal Processing and Control*, vol. 18, pp. 1334–359, March 2015.
- [40] C. DeLuca, L. Gilmore, M. Kuznetsov, and S. H. Roy, "Filtering the surface emg signal: Movement artifact and baseline noise contamination," *Journal of Biomechanics*, vol. 43, pp. 1573–1579, January 2010.
- [41] M. A. Oskoei and H. Hu, "Support vector machine-based classification scheme for myoelectric control applied to upper limb," *Browse Journals & Magazines*, vol. 55(8), pp. 1956–65, March 2008.
- [42] A. Phinyomark, P. Phukpattaranont, and C. Limsakul, "Feature reduction and selection for emg signal classification," *Expert Systems with Applications*, vol. 39(8), pp. 7420–21, June 2012.
- [43] R. Merletti, "Standards for reporting emg data." *Journal of Electromyography and Kinesiology*, vol. 6(1),III-IV, 1996.
- [44] B. Hudgins, P. Parker, and R. Scott, "A new strategy for multifunction myoelectric control," *IEEE Transactions on Biomedical Engineering*, vol. 40(1), pp. 82–94, 1993.
- [45] P. Angkoon, L. Chusak, and P. Pornchai, "A novel feature extraction for robust emg pattern recognition," *Journal of computing*, vol. 1(1), Decembere 2009.
- [46] P. Euljoon and G. M. Sanford, "Fatigue compensation of the electromyographic signal for prosthetic control and force estimation," *Browse Journals & Magazines*, vol. 40(10), October 1993.
- [47] H. Oskoei, M. A. and Hu and J. Q. Gan, "Manifestation of fatigue in myoelectric signals of dynamic contractions produced during playing pc games," *Proceedings of EMBS 2008 30th Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, pp. 315–318, 20-25 August 2008.

- 
- [48] S. G. N. and G. T. P., “Emg pattern analysis and classification for a prosthetic arm,” *IEEE Transactions on Biomedical Engineering*, vol. BME-29(6), p. 403?412, June 1982.
- [49] M. Zardoshti-Kermani, B. C. Wheeler, and K. Badie, “Emg feature evaluation for movement control of upper extremity prostheses,” *IEEE TRANSACTIONS ON REHABILITATION ENGINEERING*, vol. 3(4), pp. 324–333, December 1994.
- [50] A. A. Adewuyi, L. J. Hargrove, and T. A. Kuiken, “Evaluating emg features and classifier selection for application to partial hand prosthesis control,” *Frontiers in Neurobotics*, October 2016.
- [51] Ga-based feature subset selection for myoelectric classification. [Online]. Available: [https://www.researchgate.net/publication/221258125\\_GA-Based\\_Feature\\_Subset\\_Selection\\_for\\_Myoelectric\\_Classification](https://www.researchgate.net/publication/221258125_GA-Based_Feature_Subset_Selection_for_Myoelectric_Classification)
- [52] C. Gongrit, P. Angkoon, H. Huosheng, P. Pornchai, and L. Chusak, “Evaluation of emg feature extraction for classification of exercises in preventing falls in the elderly,” *The 10th International PSU Engineering Conference*, vol. 8(4), 14-15 May 2012.
- [53] Discriminant analysis and supervised classification. [Online]. Available: [http://www2.stat.unibo.it/montanari/Didattica/Multivariate/Discriminant\\_analysis.pdf](http://www2.stat.unibo.it/montanari/Didattica/Multivariate/Discriminant_analysis.pdf)
- [54] X. Xugang, T. Minyan, M. M. Seyed, and L. Zhizeng, “Evaluation of feature extraction and recognition for activity monitoring and fall detection based on wearable semg sensors.” *Sensors (Basel)*, vol. 17(6), May 2017.
- [55] P value. [Online]. Available: <http://felix.unife.it/Didattica/Master-1415/25125-Wikipedia-p-value.pdf>

## 5 Acknowledgment

Al termine di questo progetto, vorrei ringraziare con grande stima il Professor Marco Gazzoni per avermi dato la possibilità di svolgere il mio lavoro di tesi in un ambito della ricerca che era esattamente quello che avrei desiderato. Lo ringrazio enormemente per avermi dato la possibilità di trascorrere cinque mesi della mia vita in una delle più belle città al mondo; chi lo avrebbe mai detto. Ringrazio anche il Lisin, per quello che mi ha trasmesso durante questi mesi. Un grazie va poi al Professor Dario Farina che mi ha ospitato al dipartimento di bioingegneria dell'Imperial College, permettendomi di confrontarmi con una realtà tra le migliori al mondo.

Giunto a questo momento però non posso non fermarmi un momento per riavvolgere quel nastro che frame dopo frame mi fa ritornare in mente persone, situazioni ed esperienze che in tutti questi anni mi hanno portato ad essere così come sono, a superare ostacoli apparentemente insormontabili e ad ottenere i traguardi raggiunti. Un immenso grazie va prima di tutto ai miei genitori, la mia fonte di vita e di educazione. Grazie mamma e grazie papà per l'educazione che mi avete dato, per avermi concesso tutto quello di cui ho sempre avuto bisogno, ma facendomelo guadagnare col sacrificio. Grazie a voi per avermi permesso di arrivare fin qui, di essermi stati vicini durante ogni sessione e durante i miei momenti di debolezza. Grazie per essere stati la mia guida, ma soprattutto per aver lasciato sempre me al volante, libero di scegliere quello che più mi avrebbe reso felice. Un Grazie a mia sorella Sara, da sempre il cervellone di casa; non posso che augurarti un immenso in bocca al lupo per i prossimi anni di studio con la consapevolezza che anche in futuro, fra i due, continuerò io ad essere ricordato come "fratello di".

Come non ringraziare il mio gruppo di Amici: Alfredo, Antonio, Enrico, Igor, Fabrizio, Francesco, Pierangelo. Rigorosamente in ordine alfabetico perché diversamente non saprei chi mettere come primo nome. Un gruppo formatosi negli anni spensierati delle elementari e medie e che ha saputo coltivare il meraviglioso sentimento dell'amicizia negli anni delle superiori, per poi rafforzarlo con la lontananza dell'università. Dico grazie a tutti voi perché è con tutti voi che ho condiviso le mie gioie, le mie delusioni e come non includere le mie ipocondrie. Devo dirvi grazie perché siete stati un esempio puro da seguire. Vi confesso che non riesco ad immaginare un futuro senza voi anche perché in fondo se siamo un muro che regge intatto dopo tutto questo tempo, un motivo ci sarà.

Un immenso grazie mi sento di rivolgerlo a tutte le persone che hanno condiviso con me

gli anni universitari e in particolar modo ad Alberto per avermi sopportato e supportato per 5 anni, per essere stato l'ideale compagno di camera e il cuoco che tutti vorrebbero in casa;

Poi arriva il mitico gruppo della Magistrale: Antonio, Matteo, Roberto e Marco. Molto più del vincente gruppo di laboratorio che siamo stati. Una vera e propria famiglia. Mi vengono in mente i km fatti a piedi per studiare in corso Peschiera o i laboratori che degeneravano rigorosamente in ricerche su come fregare just eat. Le parole purtroppo non sono abbastanza per ricordare quello che abbiamo condiviso. Spero che tutti questi ricordi continueranno a vivere a lungo fra le nostre menti. Siamo riusciti a creare un gruppo che sarà la cosa che più mi mancherà a partire da oggi.

Un grazie speciale mi sento di rivolgerlo al mio liceo "G. Stampacchia", perché a distanza di 6 anni dire di essere stato un suo studente mi riempie ancora di immenso orgoglio. Ovviamente devo includere tutti i miei professori, quelli che hanno creduto in me e quelli che a tratti lo hanno fatto un po' meno, perché in fondo se si vuole una medaglia bisogna accettare entrambe le facce.

Un ruolo non meno importante dell'istruzione è stato quello degli anni passati a vestire una maglia di colore rosso-blu che non finirò mai di amare, il sogno di tanti bambini e che ho avuto l'onore di vivere da vicino e di indossare per tanti anni. In fondo tanti lati della mia personalità si sono formati su un campo di calcio, tra una panchina e un posto da titolare. Un grazie speciale al Tricase Calcio, ai miei compagni e a tutti gli Uomini che mi hanno allenato e insegnato a crescere.

Un grazie immenso vorrei dedicarlo a me stesso, per essere stato caparbio, forte e deciso. Per aver affrontato e superato gli ostacoli incontrati, magari buttandone a volte giù qualcuno, ma riuscendo sempre ad uscirne da vincitore.

Per ultimo e non perché meno importanti vorrei dire grazie alle persone da cui tutto questo è potuto nascere: i miei fantastici Nonni. Se sono qui a scrivere questo, o a gioire dei miei successi lo devo principalmente a voi, un tesoro inestimabile.

Grazie a Torino, la città che mi ha saputo accogliere quando ancora fanciullo.

Grazie a Tricase e alla mia terra che come una vera madre è sempre pronta a riaccogliere i suoi figli.