UNIVERSITÀ POLITECNICA DELLE MARCHE
FACOLTÀ DI INGEGNERIA

Corso di Laurea Magistrale in Biomedical Engineering

Handwriting Characterization: a Myoelectric-based Pattern Recognition Approach

Relatore:
Prof. Sandro Fioretti

Tesi di Laurea di:
Federico Barbarossa

Correlatori:
Ing. Alessandro Mengarelli
Ing. Andrea Tigrini

Acknowledgements

For the development of this thesis, i would like to thank my Supervisor Sandro Fioretti, co-Supervisors Alessandro Mengarelli and Andrea Tigrini, and Professor Federica Verdini for helping me constantly and efficiently during my research internship and for reviewing the present manuscript. Moreover, i would like to strongly thank my fellow students Alessio Di Lello, Giacomo Verde and Giacomo Covella. Special thanks to my family, my friends and my girlfriend for supporting me over these years.

Contents

Abstract

Introduction. This thesis aims to propose a robust and reliable handwriting recognition method based on myoelectric analysis. We aim to provide a stepwise approach to create a classification model capable of distinguish correctly among 10 different handwritten characters classes, exploiting only the informative content hidden inside sequential arm muscle activation when a specific handwriting task is performed. Along with this main goal, we investigated some aspects that are essential when a solid handwritten character classification model is to be created. From this work, four main objectives are distinguished. First of all, an evaluation of the optimal recording electrodes setup is performed. In addition, it is conducted an analysis on single time-domain, frequency-domain and time-frequency domain features about their suitability to be used for classification purposes. Moreover, an evaluation on which set of features can achieve optimal classification accuracy is investigated, and finally, the effect of Principal Component Analysis (PCA) as feature sets dimensionality reduction technique is explored.

Materials and methods. 11 healthy subjects are asked to write 10 digits following a standard template. Signals are acquired by 8 electrodes, 4 on wrist and 4 on forearm. Once the signals are processed and segmented, 27 features are extracted and further aggregated into 6 feature sets (see Section 2.8): Hudgins, Du, Phinyomark 1, Phinyomark 2, TDAR and DWTC. These latter are used to train and validate models outcoming from Linear Discriminant Analysis (LDA), Support Vector Machine (SVM) and Random Forest (RF) classification architectures. Performance metrics are calculated to access the degree of classification accuracy of each feature sets for all classifiers.

Results and discussion. Our research carried out that, for handwriting pattern recognition problems, it is essential to combine the myoelectric information coming from wrist and forearm contemporaneously, thus a 8-channels setup is required. It has been shown that time domain features could better predict over 10 classes when used to train and test the three classifiers. Moreover, TDAR feature set is the best one in terms of accuracy and precision, while DWTC is not suitable when SVM, LDA and RF are used for this application. Hudgins feature set is a good alternative when real-time problems are to be faced. Elsewhere, though PCA can drastically reduce dimensionality of feature sets, it strongly affects the overall performance, severely decreasing classification precision. Hence, a linear method for dimensionality reduction is not recommended for the considered feature sets and for the complexity of the motor task taken into account.

1 Introduction

Handwriting is, among all human gestures, one of the most complex tasks to be accomplished. The ability of writing numbers, words and sentences fills a fundamental role in a socio-cultural context since it represents, in spite of the extensive implications of computers, an important everyday means of communication and a significant mode of annotating and apprehending textual data. Even though it might seem like a simple and obvious action, handwriting requires a complex hierarchical control of muscles of forearm and hand, whose activations must be precisely coordinated by the brain.

Central nervous system must integrate visuo-spatial information and regulate all fine contractions of muscles in order to adjust the movement to finally map a graphical sign into a 2D orthographic representation, following a comprehensible writing pattern [1]. Therefore, the inner complexity of handwriting arises mainly from 2 aspects: the involvement of a network of brain structures whose interconnections are writing specific, and the actuator muscles at the level of the forearm and wrist that must carry out the graphical sign according to the motor command. For a right-handed writer, the strongly left-lateralized activations are related to the control of the right hand by left hemispheric sensorimotor regions [2]. Complexity of handwriting task reflects on the level of brain areas activations. Since it is a task that does not usually stand alone, generally, a complex synergic activation of many brain zones is required to coordinate muscles for writing and performing other classical actions that go along with writing, like reading, elaborating sentences and understanding the meaning of phrases. For this reason, it is difficult to establish which specific brain area is responsible for the pure writing. Typically, left superior frontal sulcus/middle frontal gyrus area, left intraparietal sulcus/superior parietal area and right cerebellum are considered as primarily writing-specific brain areas while others are related to non-specific motor (primary motor and sensorimotor cortex, supplementary motor area, thalamus, and putamen) or linguistic processes (ventral premotor cortex, posterior/inferior temporal cortex) [1]. Figure 1 shows brain areas activations during handwriting task performed by a right-handed subject. Note the strong lateralization effect: major electrical activity of brain occurs in left hemisphere that is responsible for right arm muscle control.

It has been shown that handwriting task shows high variability between subjects [3, 4, 5]. The large variation in handwriting styles over individuals is mainly dependent on the writing method picked up as a kid, but also on personal preference, the copying of style variants from peers in adolescence, and later in age, by the amount of writing experience [3].

Variability of handwritten track can concern static and dynamic properties. A static variation may occur in the size or in the shape of written symbols, thus same letters may present different height or width, while a dynamic variation can occur in sequence of strokes or in writing pattern. Indeed, the degree of variation depends on the style and speed of writing, with fast writing usually showing less legibility and greater variation [6]. Despite the highest variability is mostly observed between different subjects, the handwritten track can significantly vary even in a intra-subject scenario [3]. Personal writing style is strongly affected by external environment conditions (like light conditions

and location) and internal condition (like tiredness or hurry), hence influencing, from time to time, the manner of writing. Variability of handwriting does not concern only the shape of the tracks but also the way in which actuator muscles of each individual perform writing task. Biological signals related to forearm muscle contractions, namely electromypgraphic signals (EMG), show important differences when acquired on subjects other. Considering the complex and non-stationary nature of the myoelectric handwriting signals, it is almost impossible to find any correlation between signals of different subjects: the task of writing a character involves simultaneous muscular activations in one writer that are completely different from another one. Moreover, the sEMG signal complexity can also be explained by the fact that there are various ways in which a motor task can be performed, that is the motor redundancy [7].

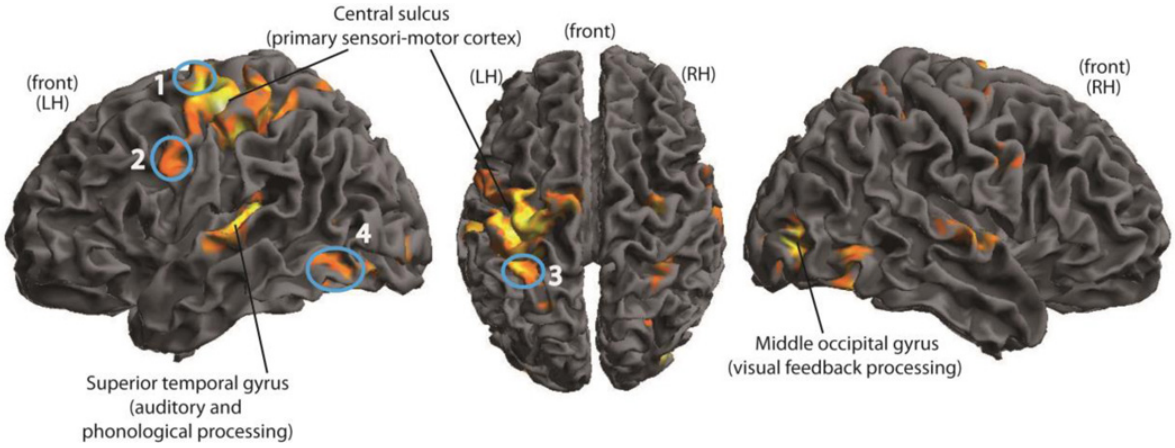


Figure 1: Brain activations during handwriting. In sequence, left view, top view and right view of brain. A fMRI brain imaging study conducted by Longcamp et al [2] put in evidence the 4 brain nodes (blue circles) which activation is consistent during handwriting task. 1-dorsal premotor cortex, 2-ventral premotor cortex, 3-superior parietal cortex, 4-fusiform gyrus. Evident lateralization effect can: this is an example of right-handed writer.

1.1 Actuators muscles of handwriting

Movements of the pen in the plane of the paper during handwriting are executed by three muscle groups that are physiologically capable of independent operation. Anatomical reference of forearm and hand muscle is represented in Figure 2.

- Muscles in the superior part of the forearm, that are extensor carpi radialis brevis, extensor carpi ulnaris, extensor/flexor digitorum and flexor carpi radialis. Radial abduction and ulnar abduction of the wrist joint are commonly used in giving letters width. It should be noted that, in the case of left-handed writers, those writing with the hand in a hooked position above the line of writing exchange the roles of thumb/finger and wrist movements, so that the latter is responsible for letter height [8].

- Muscle of the hand, that are opponens pollicis, first dorsal interosseus, medial slip/lateral slip, abductor pollicis brevis. Flexion and extension of thumb, index, and second fingers is usually used to give letters their height.

- Muscles of the upper arm near the shoulder, usually not considered in analysis since they are largely responsible for gross movement of the pen across the page.

All these muscles play a major role in the handwriting process. However, it is difficult to analyse the hand movements based on all these muscles because of their unique complexity and intricacy. In particular, muscles of the hand and wrist are so closely linked that it is impossible to voluntarily omit a muscle from the synergy of which it belongs [9]. Hence, for simplification purposes, the hand motion is usually formulated at the wrist and forearm in literature [10, 11, 12, 13, 14].

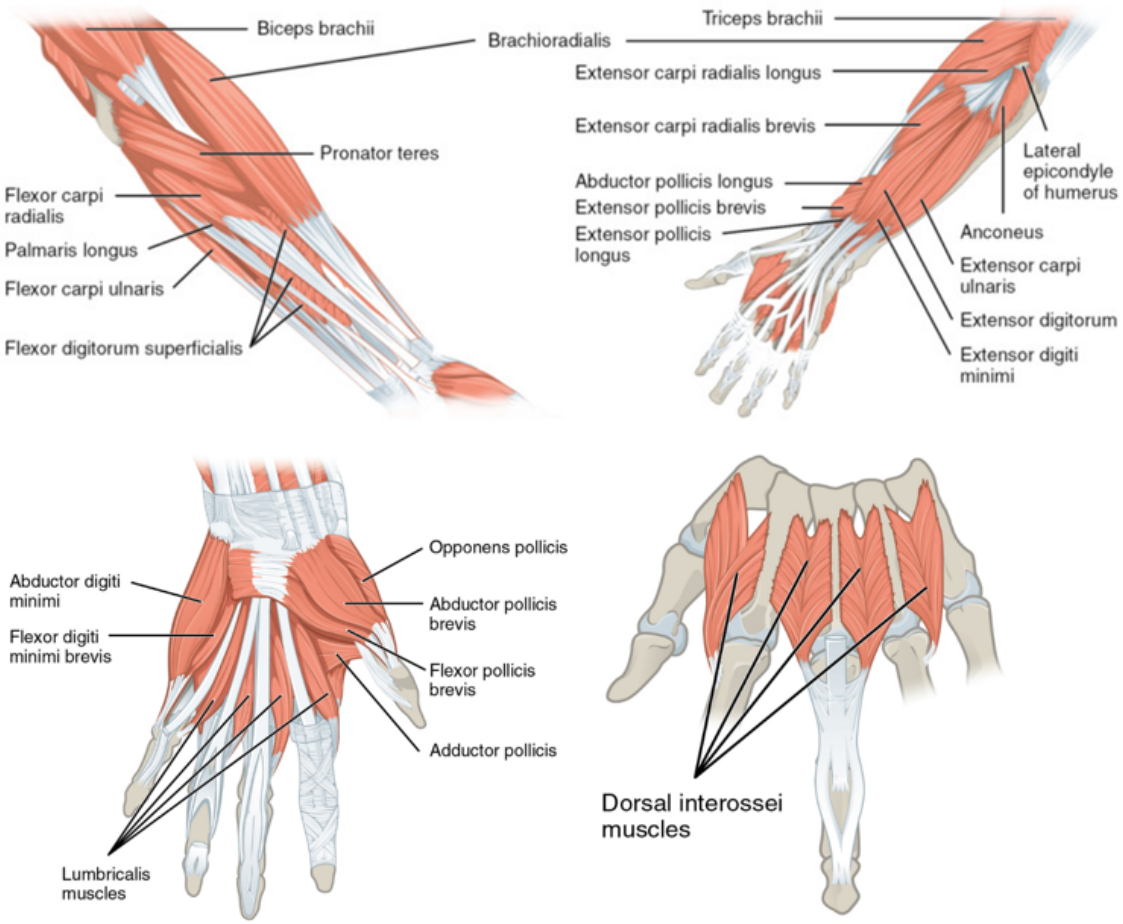


Figure 2: Muscles of forearm. Top left panel: palmar view of left forearm superficial muscles. Top right panel: dorsal view of left forearm superficial muscles. Bottom left panel: dorsal view of right hand superficial muscles. Bottom right panel: dorsal view of left hand interossei muscles. Figure taken from [15].

1.2 Handwriting recognition

Handwriting recognition is the process through which a machine can recognize handwritten letters, symbols, characters and numbers of every alphabet in an automatic way. Handwriting recognition can be performed both in an online and offline form. In case of online character recognition system, signs are processed while they are under creation and the machine recognizes the tracks in real time, while the user writes. The improvement of this technique, accompanied with recent CPU performance, makes online handwriting recognition suitable for practical use especially in pen-input devices like tablet or keypads [16]. Applications of online handwriting recognition are numerous: reading postal addresses, bank check amounts, and forms. Furthermore, online character recognition plays an important role for digital libraries, allowing the entry of image textual information into computers by digitization, image restoration, and recognition methods [17]. Offline handwriting recognition, by contrast, is performed after the writing is completed. Despite the handwritten signal is processed a posteriori (once the writer has completed his task), it finds a lot of application in situation in which it is not necessary a simultaneous response. An advantage of offline recognition is that it can be done at any time after the document has been written [17]. Offline handwriting recognition systems are less accurate than online systems because only spatial information is available for offline systems, while both spatial and temporal information is available for online systems [17]. There are mainly three types of approach to handwriting recognition: handwriting recognition based on handwritten traces image analysis, handwriting recognition based on kinematic analysis and handwriting recognition based on myoelectric analysis. The latter approach represents a novelty item with respect to classical literature, where instead the first and second methods have received much larger attentions. The difference among them is that, for each of the methods, a different type of input data is used: the first one exploits 2D character image (usually acquired by touch screen tablets), the second one uses kinematic data like acceleration and velocity of segments of hand, wrist, and forearm, and the third one uses the information contained in EMG signals to recognize handwritten characters.

- **Handwriting recognition by traces image analysis.** The first method relies on the analysis of captured images of handwritten pattern traces that are automatically recognized by algorithms. More precisely, such way to perform character recognition is the process of detecting and recognizing characters from input image and converts it into ASCII or other equivalent machine editable form [18]. It is the most widespread method, especially with growing of powerful hardware like low delay electronic pen and fine touch sensitive surfaces and displays. The development of many software applications, capable of interpreting movement of pen and fast translating the strokes into digitalized text, leaded this handwriting recognition method to be the most adopted. Both offline and online handwriting characters recognition systems exploit image segmentation techniques using algorithms that divide up the text into writing units and recognize the geometrical shape of each one.

- **Handwriting recognition by kinematic analysis.** In this approach, kinematic variables are used to recognize handwritten traces. Usually, horizontal and vertical velocity, maximum displacements of hand and forearm segments, angular and linear accelerations, trajectory curvatures are the main features that are extracted from signals acquired with instrumentation like electronic gyroscope, accelerometer sensors and Inertial Measurement Units (IMU). Sensors are directly connected to the hand of the subject performing the handwriting task.
  In 2017, Dash et al. [19] designed an approach called AirScript, for visualizing and recognizing characters written in the air. In this case, kinematic data from the Myo armband was exploited to obtain a 2D coordinates sequence and visualize a handwritten digit as an image. One year later, Roy et al. [20] implemented a mixed method, namely an analysis of kinematics data combined with image trace analysis. Single stroke digits were written in the air and were detected by a marker placed on the hand. Thus, a kinematic camera recorded the handwriting trajectory. Once each frame underwent segmentation, the marker tip position was used to approximate the handwriting trajectory and to obtain a final image of the reconstructed handwritten digit. Handwriting recognition by kinematic analysis is particularly addicted to assess the effects of neuro-degenerative diseases, since they dramatically affect musculature neural control leading the patient to a progressive degeneration of fineness of movement, especially handwriting [21]. In 2018, Gadherian et al [22] proposed an algorithm for handwriting recognition based on the analysis of some kinematic features extracted from 43 subjects. Their study brought to access a reliable handwriting-based tool for early detection of Alzheimer's disease.

- **Handwriting recognition by myoelectric analysis.** EMG signals have been widely used and applied as a control signal in numerous man–machine interface applications and have also been deployed in many clinical and industrial applications including diagnoses of neuromuscular diseases, controlling assistive devices like prosthetic and orthotic devices [13, 23]. Handwriting recognition by EMG signals relies on the analysis of information hidden inside myoelectric data. Myoelectric signal (MES) contains rich information that can be exploited to access the subject's movement intention in the form of a muscular contraction, using surface EMG electrodes. EMG signals represent the input of this approach. Generally, after processing of raw signals, features are extracted from signal windows and are used to train and validate a proper classifier aimed to perform pattern recognition [24]. In the present study, we followed this type of approach for handwriting recognition, focusing on the aspect related to myoelectric signals (MES) generated by the movement (contraction and relaxation) of forearm and hand muscles (at wrist level) during handwriting tasks.

1.3 State of the art in the handwriting recognition by myoelectric analysis

Previous studies have relied largely on the analyses of handwritten traces or kinematic analysis of handwriting. Nevertheless, myoelectric analysis associated with handwriting has received little attention and it is still a subject that has to be deeply explored. Some interesting research was published over last decade. Simao et al. [25] made a recent review of the sEMG-based systems using pattern recognition techniques and describing sensitive processes as signal acquisition and filtering. Their study focused on new human-machine interaction modalities [25]. Recently, advances in biomedicine have made possible the project and the development of myoelectric prostheses capable of performing real arm or leg movements [26]. Neural prostheses assure the recovery of limbs mobility and assist the communication of subjects with disabilities. Further to this point, Reza Boostani et al [27] proposed their criteria for extracting the best features to have a high rate of motion classification for controlling an artificial hand to replicate gestures. In 2009, Linderman et al. [28] studied a robust method in which sEMG signals generated by hand and forearm muscles during handwriting activity were reliably translated into both algorithm-generated handwriting traces and font characters using decoding algorithms. Linear discriminant was used as classification algorithm. The outcome of their research, as a result, validated the feasibility of recognizing handwriting solely from EMG signals. Their results are supported by solid value of discrimination accuracy (90.4%) between classes. In literature, mainly two solutions are proposed when sEMG signals are to be used to solve problems of handwriting recognition. Among the most widespread methods are Template Matching [28] and Dynamic Time Warping (DTW) [29]. The first technique proposed by [28] consists in creating a template (averaging the EMGs over all detected epochs yielding a generic template) that is entered in the template matching analysis where a sliding window is moved along the EMG recordings. In this way the EMGs and the template are continuously correlated in order to refine the EMG sequencies, that are further recognized using proper classification algorithms. On the other side, in 2010, Huang et al. [29] proposed a novel sEMG based handwriting recognition method. Dynamic time warping (DTW) algorithm was introduced to align two sEMG time series in the time axis. Two time series, which have approximately the same overall wave-forms, are not close to each other in Euclidean distance. To overcome constraints imposed by geometric distance, they used DTW algorithm to warp one time series non-linearity and calculate the distance with the other time series in a more accurate way. These distance values represent discriminating features extracted from the signal. DTW algorithm is particularly addicted when dealing with sequences in which individual components have characteristics that largely vary over time, and for which the simple linear expansion or compression of the two sequences does not bring satisfactory results. An improvement of DTW algorithm was carried out by Chengzhang et al. three years later [30]. This work aimed to even improve classification process and enhance recognition accuracy. By applying several modifications to the older model, the results of offline analysis showed an important increasing in the average recognition accuracy (9.20 % greater with respect to [29], boosting classification accuracy to 93%). In 2020, Beltran-Hernandez et al [7] proposed a new approach where powerful

Deep Learning (DL) architectures for feature extraction and sequence recognition are used. However, despite the application of such architectures carried out promising results (94.8% overall classification accuracy), deep learning approach resulted more computationally expensive than machine learning, thus it might be practically more difficult to be applied in an online scenario. For this reason, the choice of a proper classifier is an important aspect that has been also discussed in literature. Neural Networks (NN), like convolutional NN, are addicted to the signal processing and feature extraction, when there are huge amounts of data to deal with. Even Long Short-Term Memory (LSTM) network can be used to recognize sEMG signals, supported by the fact that LSTMs have also found application successfully to gesture recognition [25]. Anyway, machine learning algorithms have been demonstrated to be a valid alternative in terms of computational burden and recognition accuracy [28].

1.4 Aims of this work

The main goal of this research is to propose a robust and reliable handwriting recognition approach based on myoelectric analysis. A solid model that can recognize handwritten character solely from EMG signal, represents, in addition to an innovation in myoelectric pattern recognition, also a tool with interesting practical implications, such as rehabilitation for people with pathologies that involve degeneration of neuromuscular control. The information that is extracted from the EMG signal, can be exploited to help the patient with a real-time correction of handwriting or with the possible rebound from it.

We aim to provide a stepwise approach to create a classification model capable of distinguish correctly among 10 different handwritten characters classes, exploiting only the informative content hidden inside sequential arm muscle activation when a specific handwriting task is performed. In the interests of achieving this challenging goal, four parallel essential aspects should be carefully investigated. They are so far explored in this thesis.

First of all, when an EMG pattern recognition problem is to be faced, the choice of a proper electrodes setup is essential. Therefore, it is our objective to investigate the optimal setup of electrodes placement over the arm, and evaluate the best configuration to maximize the retrieval of myoelectric information and increase the multi-class predictive performance of the system when handwritten characters must be recognized.

Features extracted by EMG signal are a way to access myoelectric informative content. For this reason, a proper choice and selection of them is essential to achieve satisfactory results in pattern recognition problems. This thesis aims to investigate what features, singularly, can provide good classification performances in terms of recognition accuracy. Thus, an evaluation of discriminating behaviour of single time domain, frequency domain and time frequency domain features extracted by EMG signal, is carried out in this study. A total number of 27 features will be thoroughly explored and analysed.

In pattern recognition problems, it is common practice to group single features in sets to exploit synergically the power of single features. Further to this point, we aggregated features according six classical feature sets found in literature [31]. We aim to investigate which are the most suitable feature

sets for this type of application, comparing their multi-class discriminating behaviour and analysing their robustness in terms of classification accuracy.

Moreover, this work aims to investigate the effect of a dimensionality reduction technique, namely Prinicpal Component Analysis, when it is used to decrease the feature set dimension. A comparison between original and reduced dimension feature set is useful to investigate if a linear dimensionality reduction method is appropriate for handwriting recognition purposes, considering the feature sets taken into account and the complexity of motor task.

The following 4 macro-sections of this thesis are *Materials and methods* (chapter 2) that describes in detail the setup, the writing task, the signal processing phase, the extraction and aggregation of features and the classification phase. In addition, a brief overview of electromyography is reported. *Results* (chapter 3) is the section dedicated to all the results that we have carried out from experiments while in *Discussion* and *Conclusion* (chapter 4), the results are thoroughly explained and discussed. Moreover, further possible enhancements in myoelectric based handwriting recognition are proposed.

2 Materials and methods

2.1 An overview of electromyography

Electromyography is an experimental technique concerned with the development, recording and analysis of myoelectric signals [32]. The myoelectric signal is the electrical manifestation of the neuromuscular activation associated with a contracting muscle [33]. Rich information contained inside MES can be extracted to access anatomical and physiological properties of muscles as well as the control scheme of the peripheral nervous system.

2.1.1 EMG signal generation

Motor unit (MU) (Figure 3) is the smallest functional unit, formed by motor neuron (with its dendrites and axons) and the muscle fibres innervated by the axon, that describes the neural control of the muscular contraction [34]. When a contraction of specific muscle fiber is required, motor units related to that muscle are activated. Typically, groups of motor units often work together to coordinate the contractions of a single muscle. When a motor unit is activated by CNS, all its muscle fibers contract. In particular, the activation of an alpha-motor anterior horn cell, induced by the central nervous system or reflex, results in the conduction of the excitation along the motor nerve [34, 35]. After the release of electric stimulus at the motor end-plates, an end-plate potential is formed at the muscle fiber innervated by this motor unit. The diffusion characteristics of the muscle fiber membrane are briefly modified and Na+ ions flow in. This causes a destabilization in resting membrane potential (around -70/-80 mV [36]). Hence, this ignites a membrane depolarization (due to the increasing of Na+ influx) which is immediately restored by backward exchange of ions within the active ion pump mechanism, i.e. the repolarization. Now, if the Na+ ions influx exceeds, a certain threshold value of membrane potential (around -55 mV [36]), an action potential (AP) is generated. AP is characterized by a fast succession of strong depolarization and an immediate repolarization phase, followed by a late hyperpolarization period of the membrane (Figure 4) [35]. Starting from the motor end plates, the action potential spreads along the muscle fiber in both directions and inside the muscle fiber through a tubular system. Linked chemical processes (electro-mechanical coupling) finally produce a shortening of the contractile elements of the muscle cell. The EMG signal is based upon action potentials at the muscle fiber membrane resulting from depolarization and repolarization processes as described above. This depolarization zone, after initial excitation, travels along the muscle fiber at a velocity of 2-6m/s and passes through electrode site [35]. Indeed, if a probe is placed in that field, it will detect potential differences causally related with the action potential of the fiber (Figure 5). The depolarization-repolarization cycle forms a depolarization wave or electrical dipole which travels along the surface of a muscle fiber. Depending on the spatial distance between the two probes, the dipole forms a potential difference between the electrodes. Because a motor unit consists of many muscle fibers, the electrode pair detects the magnitude of all innervated fibers within this motor unit, depending on the probe spatial distance and resolution [35]. Typically, action potentials of the fibers

are summed up to obtain the motor unit action potential (MUAP), which has a triphasic pattern, as shown in Figure 6. MUAPs differ in form and size depending on the geometrical fiber orientation in ratio to the electrode site. Superposition of all MUAPs crafts the final EMG signal.

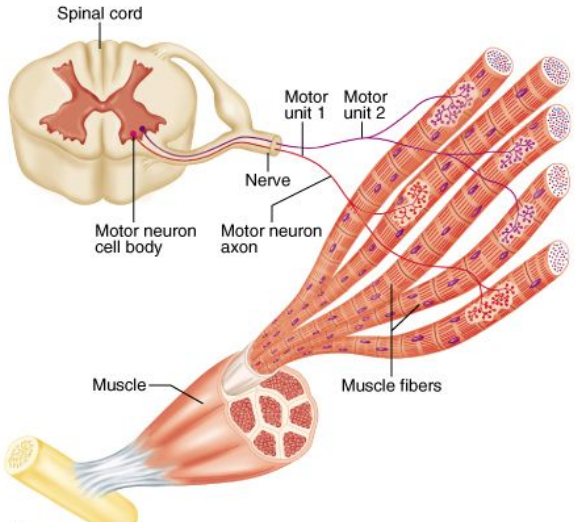


Figure 3: Motor unit. Axons of motor neuron extend from spinal cord to muscles. There each axons divide into a number of terminals forming neuromuscular junction with muscle fibers. Image taken from [37].

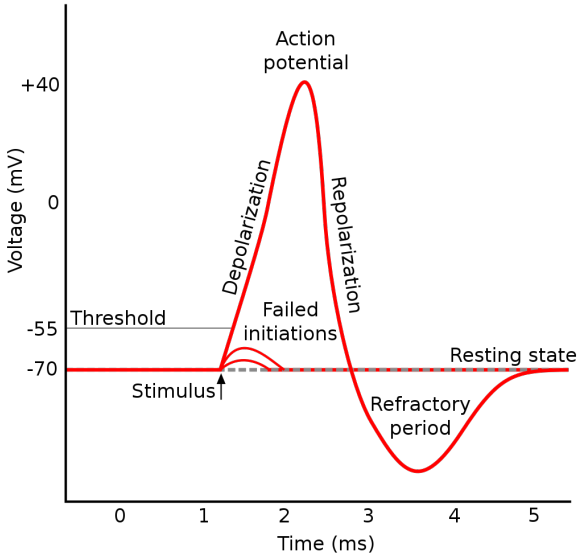


Figure 4: Action potential. A stimulus is applied at time 1 ms, which raises the membrane potential above -55 mV (threshold potential). After the stimulus is applied, the membrane potential rapidly rises to a peak potential of $+40$ mV at time 2 ms. Potential then drops and overshoots to -90 mV at time 3 ms, and finally the resting potential of -70 mV is reestablished at time 5 ms. Image adopted from [38].

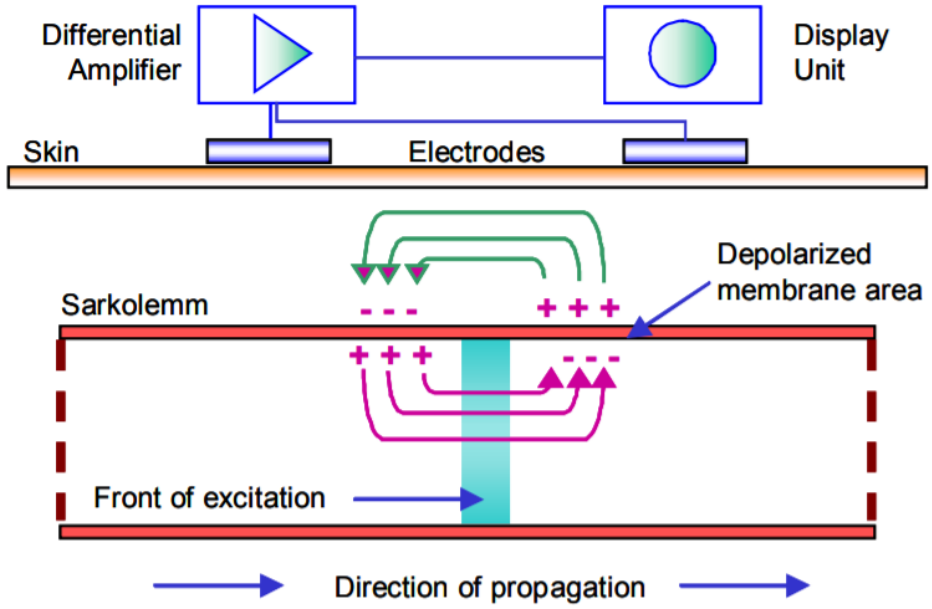


Figure 5: Depolarization zone propagating along muscle fiber. Electrical dipole generated by recursive depolarization and repolarization succession can be detected at skin level by superficial electrodes. Figure taken from [35].

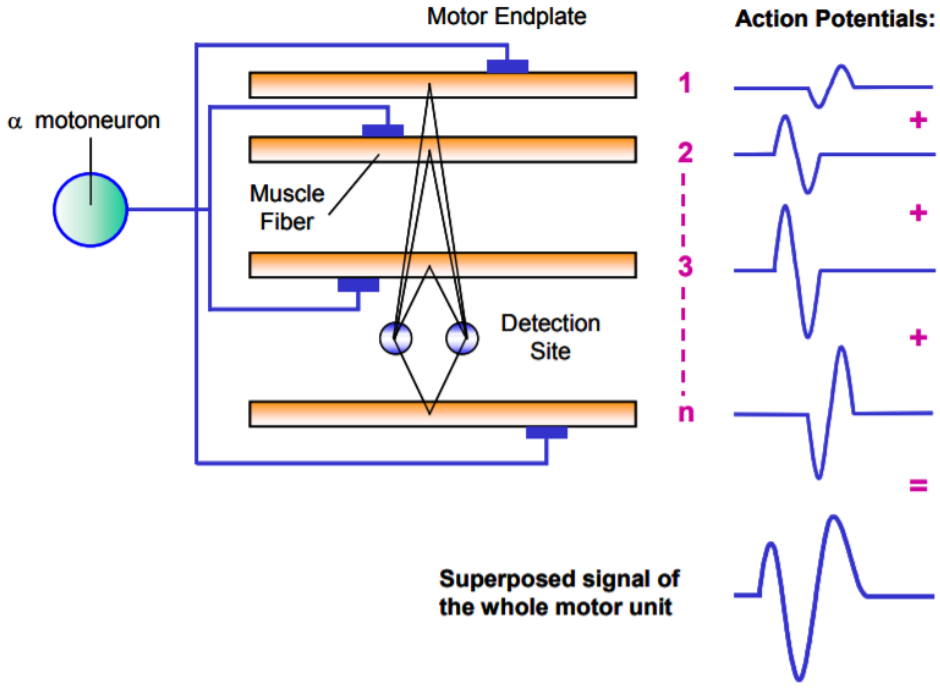


Figure 6: Superposition of n-action potentials. Note that, the greater the distance between detection site and muscle fiber, the weaker is the amplitude of the action potential. All the action potentials of muscle fiber are summed up to obtain the contribute of the entire motor unit. Figure taken from [35].

2.1.2 EMG signal acquisition

Acquisition of biological signals is performed by electrodes. Basically, electrodes are special probes able to transduce the ionic current, generated during the contraction of the muscle fibers, into an electric current. Such devices provide internally an electronic architecture through which the signal, besides being captured by terminal plates, are amplified, filtered and converted into digital format. At the end of this chain a raw EMG signal, ready for processing, is available. An organic representation of this architecture is depicted in Figure 7.

There are two main families of electrodes: intramuscular electrodes and surface electrodes. Due to their non-invasive nature, hence minimal discomfort created by the wiring, surface electrodes (Figure 8) are particularly addicted for dynamic tasks [34]. Anyway, besides the benefit of easy handling, only surface muscles can be detected. For deeper muscles, like those ones covered by superficial muscles or bones, intramuscular electrodes are inevitable. For surface electrodes, silver/silver chloride pre-gelled electrodes are the most often used electrodes and recommended for the general use. The conductive area of these electrodes is typically of reduced dimensions, usually sized to 1 cm or smaller [39]. Commercial disposable electrodes are manufactured as wet gel electrodes or adhesive gel electrodes. Wet-gel electrodes have better conduction and impedance conditions than adhesive gel electrodes. Intramuscular electrodes represent an invasive approach of acquisition, since a needle, or a thin wire, (Figure 9) is inserted into the skin and directly placed over the muscle of interest. Its small insertion area enables the electrode to detect individual MUAPs during relatively low force contractions. Moreover, the electrode tip (a thin canula) can be re-positioned within the muscle (after insertion) when new neighbour tissue has to be investigated [39]. However, it is an invasive technique that may be stressful for the subject that undergoes such type of inspection.

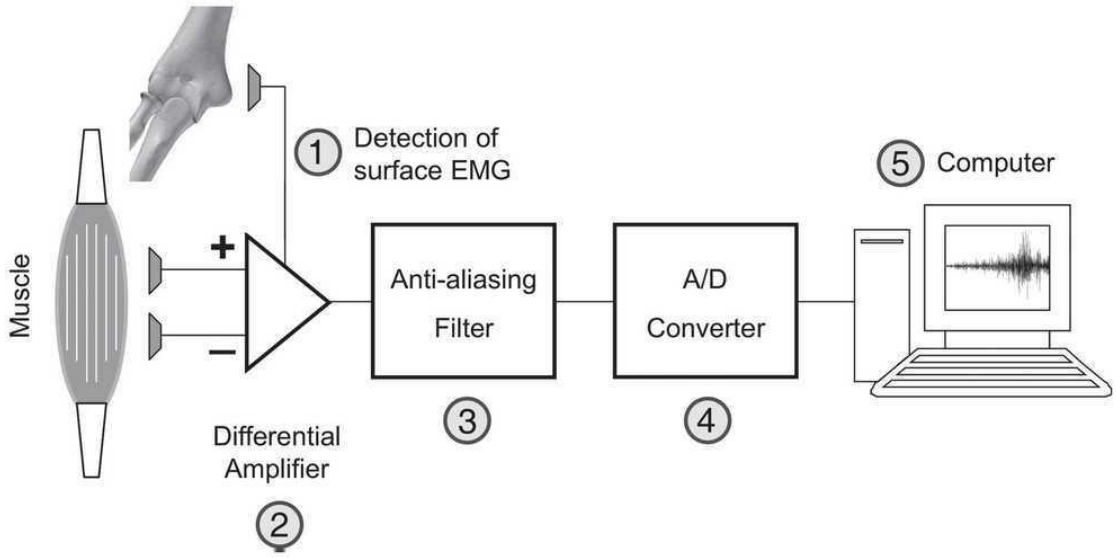


Figure 7: Block diagram of the acquisition process of EMG signal: (1) the detection of myoelectric potentials with surface probes and a reference electrode, schematically illustrated on a bone; (2) the amplification of such potentials with differential amplifiers allows to eliminate the potentially much greater noise signal from power line sources. This differential detecting configuration is able to discard any signal that originates far away from the detection sites while it amplify signals that are in the immediate neighborhood of the detection surfaces. (3) analog filtering of the amplified potentials to avoid aliasing and, finally (4) the sampling of the surface electromyogram into digital voltage values to be stored on a computer (5). Figure adopted from [40]

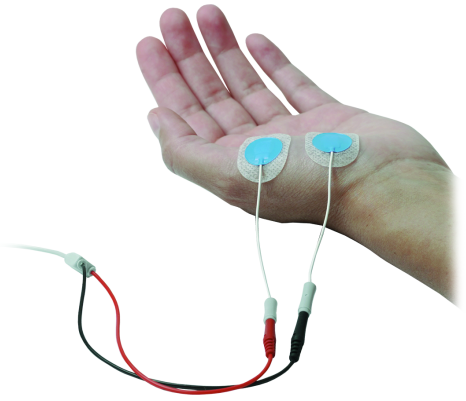


Figure 8: Surface electrodes. Thanks to the adhesive foil, probes can be solidly positioned.

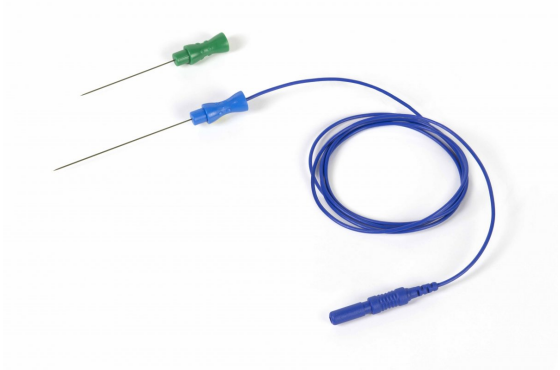


Figure 9: Needle electrode. Tips can be of many dimensions according to the depth of the muscle to be inspected.

2.1.3 Common EMG artifacts

The quality of the acquired signal depends on many factors. Location of sensors must be chosen according to some specific criteria. Indeed, most recommendations for electrode application work with specific anatomical landmark system, based on dominant bone areas or other structures that can easily be palpated [33, 35]. Moreover, the electrode should be placed upon the middle of muscle belly. Proper functioning of sensor and a correct electrode-skin interface (good skin preparation) are fundamental factor to obtain a good quality raw EMG signal [33]. In general, due to its sensitive nature (signal range starts from a few millivolts) the EMG signal can be easily influenced by external noise or other artifact sources. Among the most common artifacts there are power line noise, baseline offset, movements artifacts, ECG interference and muscle cross-talk.

- **Powerline noise** The main source of this noise is the electromagnetic radiation. Power line noise comes from the power line and is transmitted by electrical devices (i.e. amplifiers or computer) placed near the EMG data acquisition device. In fact, an EMG amplifier can catch ground noise from the power net thus resulting in an EMG signal with increased baseline noise (typical frequency of power line noise is 50/60 Hz) (Figure 10). Changing the power plug and avoiding multiple plug connectors and cable drums for the EMG amplifier can overcome this problem.

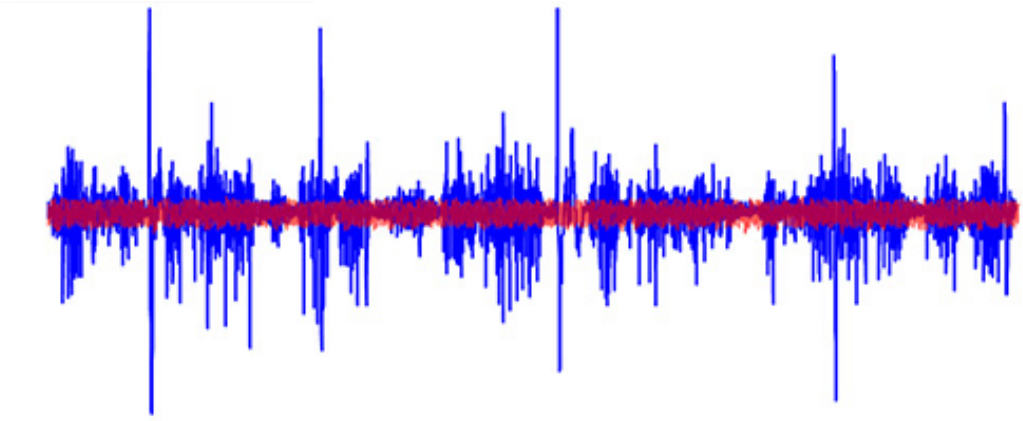


Figure 10: Powerline noise. The baseline of the signal is remarkbly widen by the noise.

- **Baseline offset** Baseline offset is another common artifact that occurs frequently during recordings. It is caused by the difference in the impedance between the skin and the electrodes. Furthermore, it may occur if the subject did not relax at the beginning of measurement. It adds an offset to the raw signal which is normally centred on 0 (Figure 11). A proper skin preparation and robust placement of electrodes on the skin generally prevent the problem.

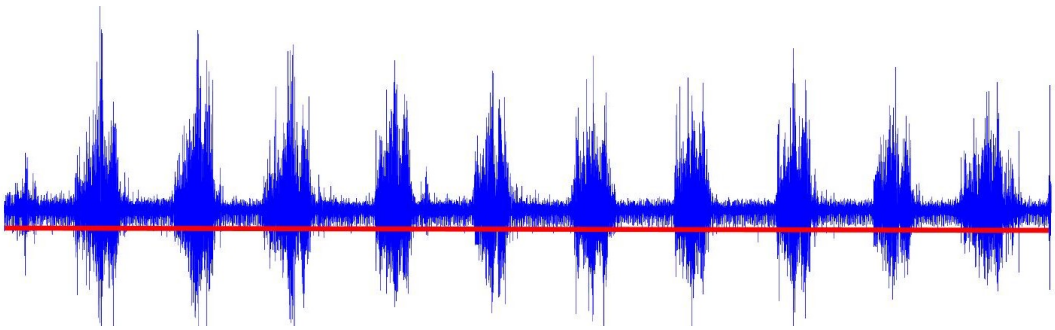


Figure 11: Baseline offset. Raw signal is not centred in 0 baseline (red line) but is slightly upper shifted.

- **Movement artifact** Movement artifacts occur during patient movements, the electrodes can move, or the cables be pulled, which may create artifacts in the EMG signal. It can be avoided ensuring a proper fixation of probes. It causes a random and abnormal trend of baseline (Figure 12).

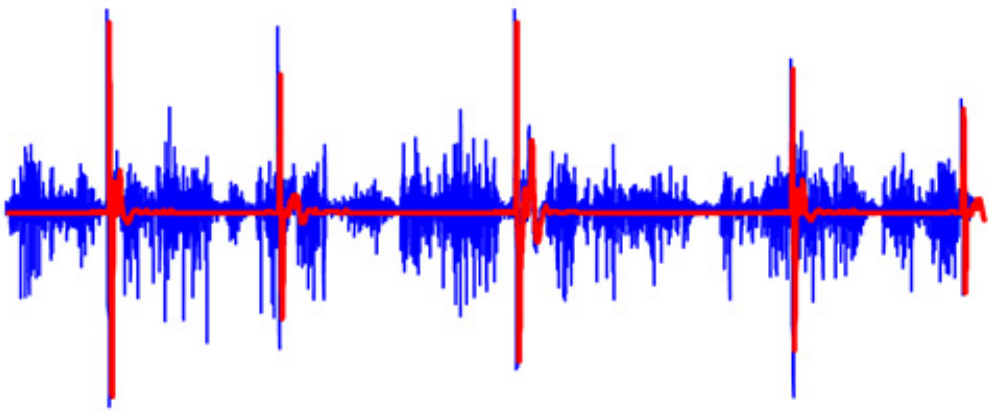


Figure 12: Motion artifact. EMG track is affected by motion interference because of sudden movement of patient during acquisition.

- **ECG interference** Whenever a measure near the heart (shoulder and trunk muscles on the left side) is conducted, ECG bursts may contaminate the EMG recording (Figure 13). This is a biological artifact that often cannot be avoided. It can be reduced by very good skin preparation and modified position of the ground electrode. Proper low pass filtering can delete ECG interference without corrupting the regular EMG characteristics.

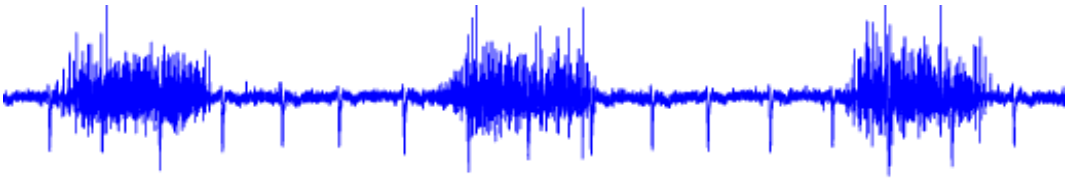


Figure 13: ECG artifact. Note the presence of QRS complex spikes due to EMG recordings.

- **Crosstalk** Crosstalk is another common unwanted phenomenon. Neighbouring muscles may produce a significant amount of EMG that is detected by the local electrode site (figure 14). Even if crosstalk does not exceed 10%-15% of the overall signal contents (or it is not available at all), care must be taken for arrangements in a narrow space within muscle groups. Indeed, crosstalk can be avoided by choosing the appropriate inter-electrode distance, at least 2 cm.

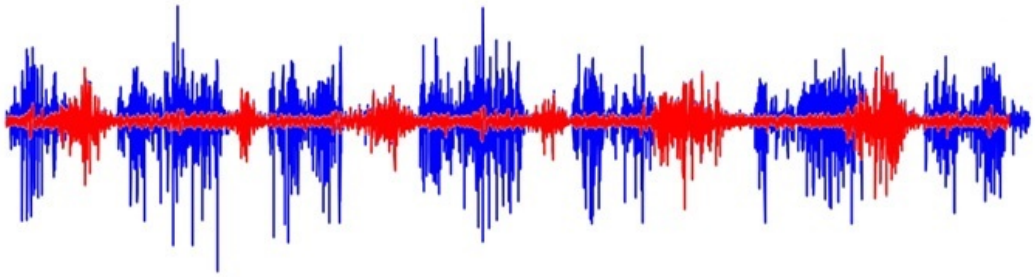


Figure 14: Crosstalk of muscles. Blue track and red track are the EMG signals derived from two different muscles activation. Crosstalk phenomenon is characterized by a superimposition of both recordings, due to the neighbor electrodes placement.

2.1.4 Classical signal processing steps

The raw EMG recording already contains very important information and may serve as a first objective information and documentation of the muscle innervation. A qualitative assessment and a first understanding of neuro-muscular control can be directly derived from original signal [34]. If a quantitative analysis is required, it is common practice to perform some EMG specific signal processing steps in order to increase the reliability and validity of findings. Generally, this signal processing phase is composed by 4 steps but, of course none is mandatory or strictly necessary. According to the purpose of the study, an EMG signal can undergo or not the following classical steps.

- **Rectification** In this step, a full wave rectified EMG is obtained by computing the absolute value of the raw signal. At the end of this process, all negative amplitudes are converted to positive values. Since raw EMG signal has a zero-mean, rectification is usually performed, besides to inspecting the signal more easily, to apply to the curve standard amplitude parameters, like mean amplitude, maximum peak value or area under a curve.

- **Smoothing or linear envelope** Pattern of EMG is of random nature. This is due to the fact that the actual set of recruited motor units constantly changes within the diameter of available motor units and the superposition of MUAPs is arbitrary. As a consequence, activation bursts cannot be reproduced a second time exactly as it is by its precise shape. Digital smoothing algorithms, directly applied to raw data, outline the mean trend of signal development, in order to minimize its non-reproducibility. In practice, steep amplitude spikes are cut away and the signal receives a linear envelope.

- **Digital filtering** With the exception of amplifier bandpass filtering, additional filtering is not usually needed in dynamic EMG studies (performed with modern amplifier technology). Rather scientific recommendations for research studies (i.e. European Recommendations for Surface Electromyography SENIAM) do not suggest any narrower band setting and the target is to measure the EMG in the full band length of 10 to 500 Hz. In particular, the application of notch filters, to cancel out 50 or 60 Hz noise, is not always recommended because it can compromise too much EMG signal power information.

- **Amplitude normalization** EMG signal amplitude is strongly influenced by detection conditions. Amplitude can strongly vary among different subjects (and within same subject) according to electrode positioning sites. To overcome this uncertainty and to standardize as much as possible scaled parameters, amplitude normalization is performed. One very common technique is the normalization with respect to maximum voluntary contraction. In practical terms, before acquiring data, the subject is asked to contract as much as possible the muscle of interest (or a group of muscles) to obtain the maximum amplitude value. Raw data are divided by this value to be normalized. New normalized data will now lay in the range -1 to 1. The main effect of all normalization methods is that the influence of the given detection condition is eliminated, and data are rescaled from millivolt to a percent of selected reference value. Amplitude normalization performs a sole rescaling, without changing the shape of signal.

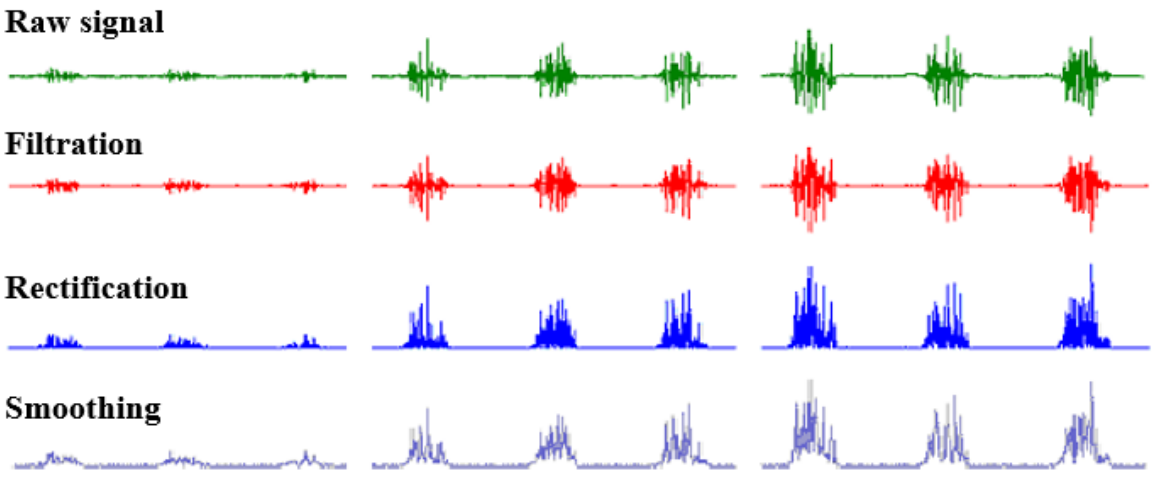


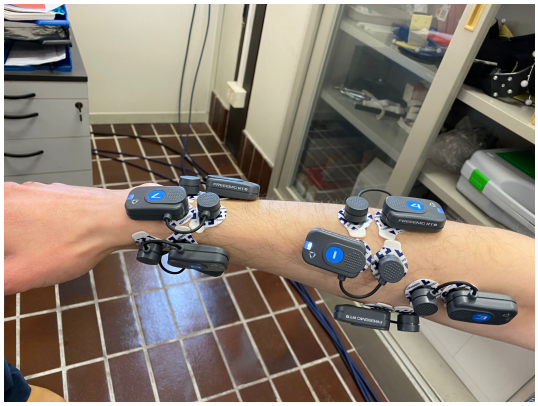
Figure 15: Main steps of EMG signal processing. Figure adopted from [41]

2.2 Hardware and setup

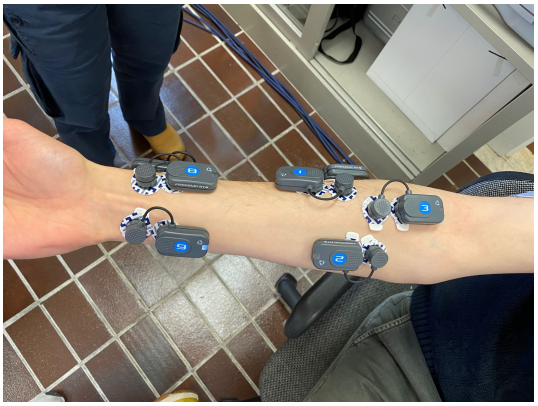
Before starting the handwriting task, each subject's wrist and forearm are armed with EMG electrodes. For this type of experiment, BTS Bioengineering FREEEMG surface electrodes are used (Figure 16,17). Such probes transduce myo-electric signal generated by muscle contraction into electric digital signal that can be directly transferred to a computer through Wi-Fi technology. In this way, connection cables are avoided, hence handwriting can be performed in a more comfortable way. Their small and compact shape guarantees a precise placement over muscles of interest. During positioning phase, it is important to ensure that every electrode is robustly attached to the skin, in order to avoid any undesired movement during dynamic task that introduces random noise to the signal. Furthermore, electrodes must be positioned as closest as possible to the centre of the muscle, and, at the same time, they must have an inter-distance one each other to avoid cross-talking phenomena. These electrodes enable to acquire EMG data with a sampling frequency of 1000 Hz. A unique positioning probes protocol has been followed for all the subjects. For this specific task, the choice of setup was inspired by the one proposed for a hand gesture recognition study by Botros et al [11]. We performed a 8 channel EMG analysis. We divide up the positioning of electrodes in such a manner that 4 probes are located in the superior part of the forearm, surrounding its whole circumference, and the other 4 probes are located at the level of the wrist. Figure 18 depicts the placement of electrodes on wrist and forearm of one of the subjects and Figure 19 shows the anatomical landmarks over which probes are positioned. Forearm electrodes 1,2,3,4 are positioned circumferentially around the proximal forearm. In detail, electrodes 1 and 2 are placed over the extensor digitorum (ED) and flexor carpi radialis (FCR) respectively, proximal to the elbow joint, consistent with the locations of maximum EMG signals from these muscles. The other two electrodes, 3 and 4, were equidistantly located between 2 and 4 on opposite side of the forearm, covering brachioradialis and flexor carpi ulnaris respectively, in order to achieve the widest spatial coverage. This forearm electrode configuration was selected to match the number and distribution of wrist electrodes and to be in line with common practices in literature [42, 43, 44]. Wrist electrodes 5,6,7,8 are positioned in such a way that electrodes 5 and 7 cover extensor digitorum minimi and extensor digiti respectively, on the posterior side of the wrist proximally to the ulnar styloid process, and electrodes 6 and 8 were placed equally but on the opposite anterior side of the wrist, over distal ending of flexor carpi radialis in correspondence of the deep flexor pollicis longus and over flexor digitorum superficialis respectively. Such wrist configuration facilitates the recording of EMG signals from bellies of deep layer muscles, involved in the control of fine finger motion, coming to the surface at the wrist level [45, 46].

Figure 16: BTS surface electrode. Its small dimension and the absence of wires makes it suitable for handwriting task.

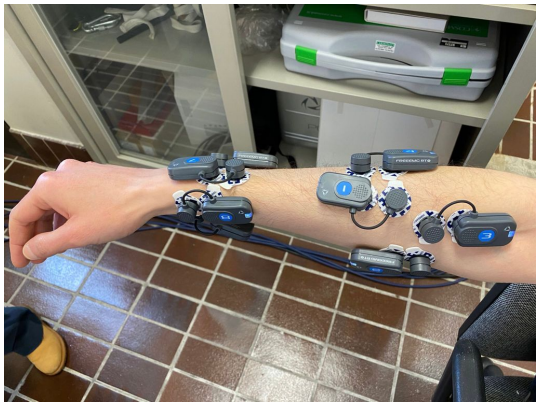
Figure 17: Charging box BTS electrodes. When batteries are completely charged, electrodes guarantee up to 6 hours of nonstop acquisition.



(a) Posterior view of right arm.

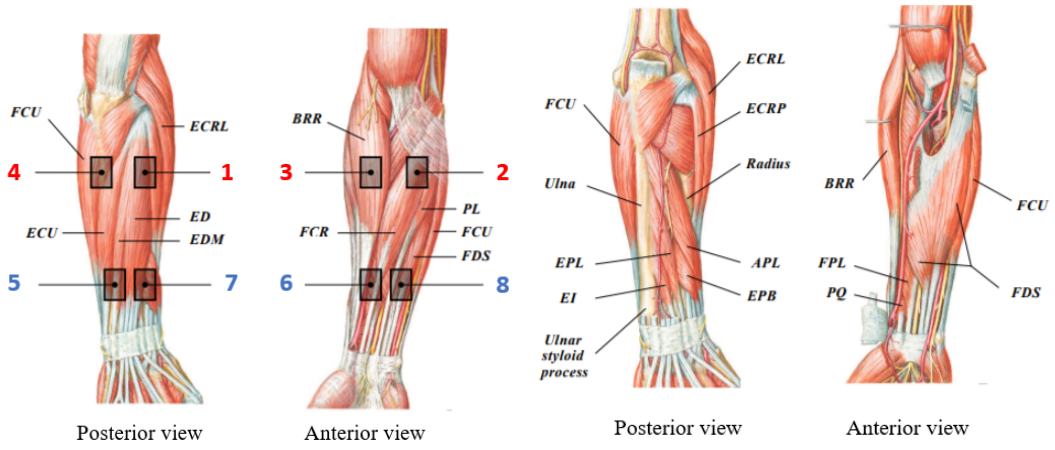


(b) Anterior view of right arm.



(c) Lateral view of right.

Figure 18: Setup of the electrodes adopted for handwriting recognition task.



(a) Superficial muscles.

(b) Deep muscles.

Figure 19: Locations of forearm (1,2,3,4) and wrist (5,6,7,8) electrodes with anatomical references of superficial (a) and deep (b) muscles. *APL: abductor pollicis longus; BRR: brachioradialis; ECRB: extensor carpi radialis brevis; ECRL: extensor carpi radialis longus; ECU: extensor carpi ulnaris; ED: extensor digitorum; EDM: extensor digiti minimi; EI: extensor indicis; EPB: extensor pollicis brevis; EPL: extensor pollicis longus; FCR: flexor carpi radialis; FCU: flexor carpi ulnaris; FDS: flexor digitorum superficialis; FPL: flexor pollicis longus; PL: palmaris longus; PQ: pronator quadratus.* Figure adopted by [11].

2.3 Definition of handwriting task

11 healthy subjects, 6 men and 5 women aged between 21 and 50, took part to this experiment. Two of them are left-handed, all the others are right-handed. All participants are asked to write numbers from 0 to 9 with their dominant hand, sitting in a comfortable position and leaning the arm against desk to not feel fatigue during the task. In order to standardize the writing pattern of every digit and facilitate the writer, a template of numbers (Figure 20) is followed. As estabilished by pre-graphism rules, each number is written 10 times. Subjects are asked to alternate a writing phase of 3 seconds and a resting phase of 5 seconds for every number repetition. Approximately 80 seconds are necessary to write down one single number, for a total duration of the task of about 15 minutes.

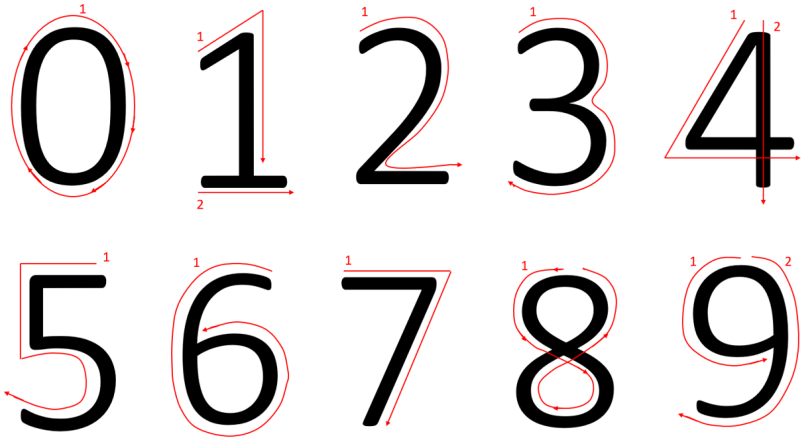


Figure 20: Number template. Red arrows indicate writing direction, small red numbers indicate strokes, hence digit 1, digit 4 and digit 9 are double stroke (pen lifts once), all other digits are single stroke (pen keeps attached to the sheet for all the writing phase).

2.4 Raw EMG signal

Raw EMG (Figure 21) signals are the outcome of the electromyographic acquisition. When the muscle is relaxed, a more or less flat EMG baseline can be seen. The raw EMG baseline noise depends on many factors, especially the quality of the EMG amplifier, the environment noise and the quality of the given detection condition. When the muscle is contracted, a burst of spikes occurs. Duration of each burst coincides with the time required by the subject to write down the number. On average, writing phase is 3000 samples long (3 seconds) while resting phase is 5000 samples long (5 seconds). By its nature, raw EMG spikes are of random shape, which means one raw recording burst cannot be precisely reproduced in exact shape. From a single multi-channel acquisition, 8 tracks are recorded and can be visualized, as shown in Figure 22.

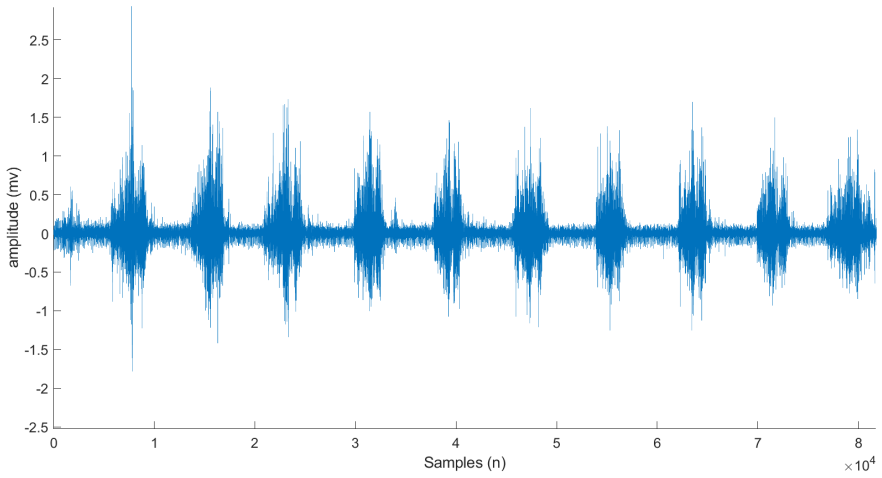


Figure 21: Raw EMG signal.

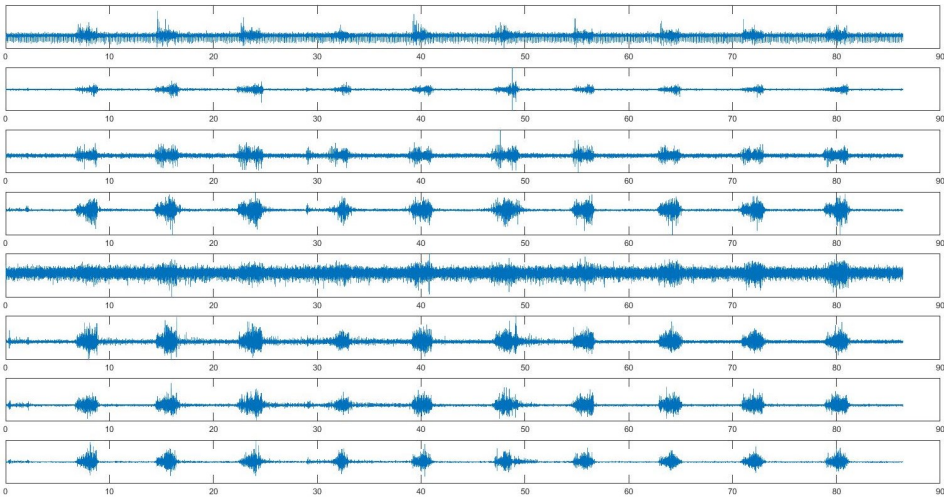


Figure 22: 8 channels recording of a single writing task. One single acquisition includes 8 channels recording, as shown in figure. Each channel refers to an electrode sensing electrical activity of a different muscle or a group of muscles.

In order to check the validity of raw EMG signal, total power spectrum of the EMG signal (Figure 23), which shows the frequency power distribution (Y-axis) in ratio to the frequency band (X-axis), is analysed. Fast Fourier Transform (FFT) is used to pass from time domain to frequency domain. We can observe that most of the surface EMG frequency power is located between 10 and 250 Hz. Moreover, the whole frequency content ranges between 10 and 450 Hz. This information can be exploited to set cut off frequencies for band pass filtering in signal processing phase. Generally, the precise shape of the total power spectrum can vary widely, depending on the FFT settings and the measurement conditions (anatomical muscle properties or tissues/skin attenuation effects).

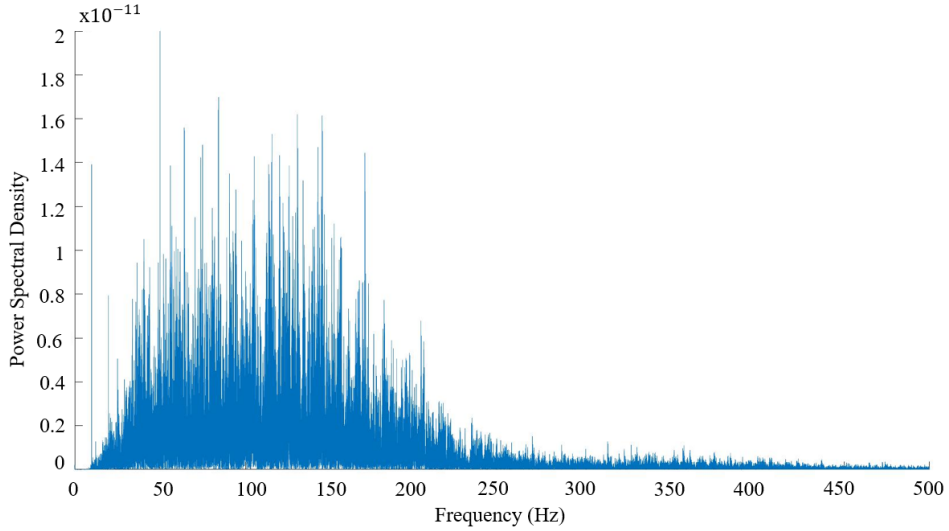


Figure 23: Power Spectral Density of raw EMG signal.

2.5 EMG signal processing

As first step, raw signal is normalized. Signal amplitude is normalized with respect to the maximum peak value, hence all data are divided by the maximum amplitude value detected during period of contraction. The new signal has a shape identical to the original one, but resulting in an amplitude scaled to a reduced range of 0 and 1. This method, namely *peak dynamic method*, is largely adopted, especially in gait analysis researches [47, 48]. Indeed, this method helps to reduce as much as possible the variability between subjects and facilitate the comparison between electrical patterns of different groups of muscles within subject. Signals are not rectified, since we need to calculate features that take into account values below the baseline, like zero crossing. Moreover, it is applied a second order bandpass Butterworth filter with 30 Hz lower cut-off frequency and 450 Hz higher cut-off frequency.

At this point, we need to know instants of activation of each group of muscles. For this reason, an ON-OFF detection is performed in order to access information contained in the contraction phase (writing phase). In order to accomplish this step, we performed a manual detection of activation samples. For every number writing task, a total of 8 raw recordings is observed and, among them, the cleanest signal is chosen. From this signal, we extracted 10 successive ON values (corresponding to the sample around which signal amplitude starts to increase as shown in Figure 24 and we added 3000 samples to each one, to obtain OFF values. In this way, ten 3000 samples long windows are extracted from a single signal (Figure 25). The previously defined ON OFF signals are used to cut the remaining 7 signals (Figure 26), in order to obtain activation windows for all the remaining channels. At the end of this process, all windows are organized in a 10x8 array (10 activations for 8 recordings).

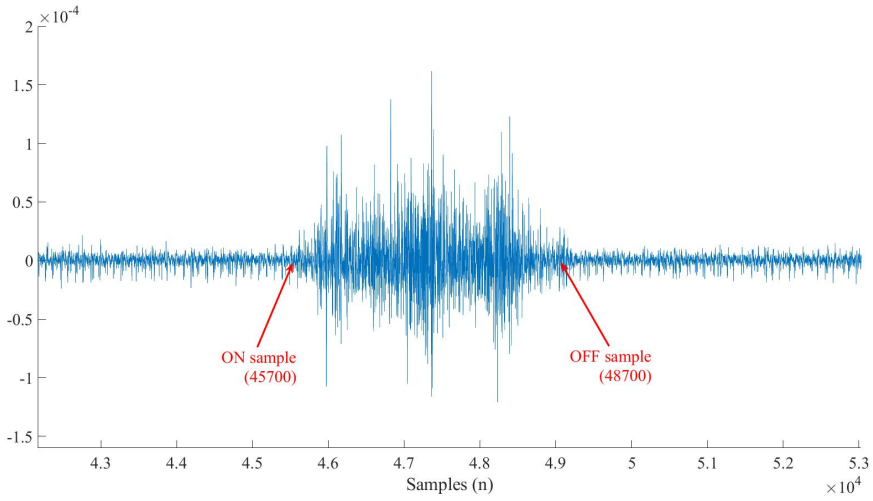


Figure 24: Selection of ON and OFF values (in samples) indicating the beginning and the end of the muscular contraction phase.

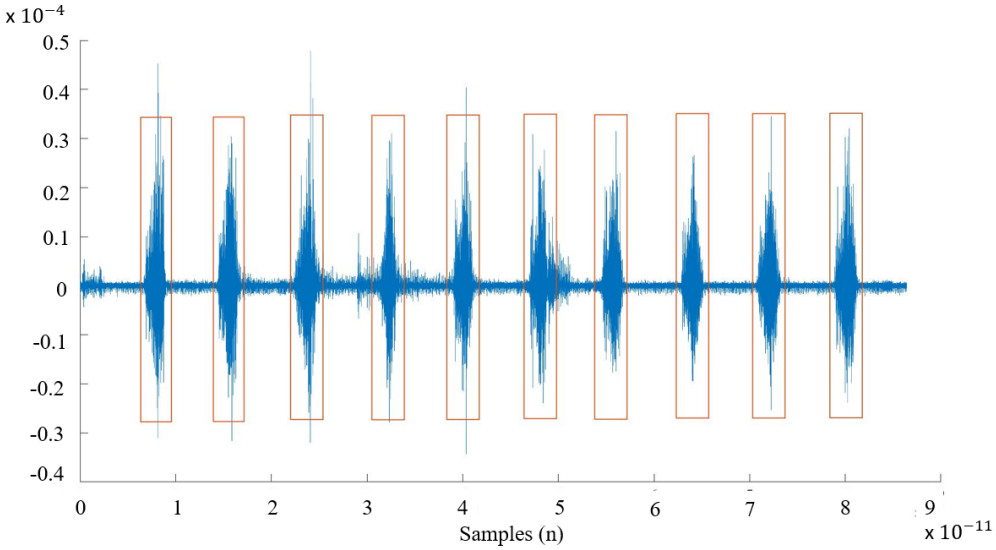


Figure 25: Selection of all writing phases in one single channel. 10 activation windows are identified.

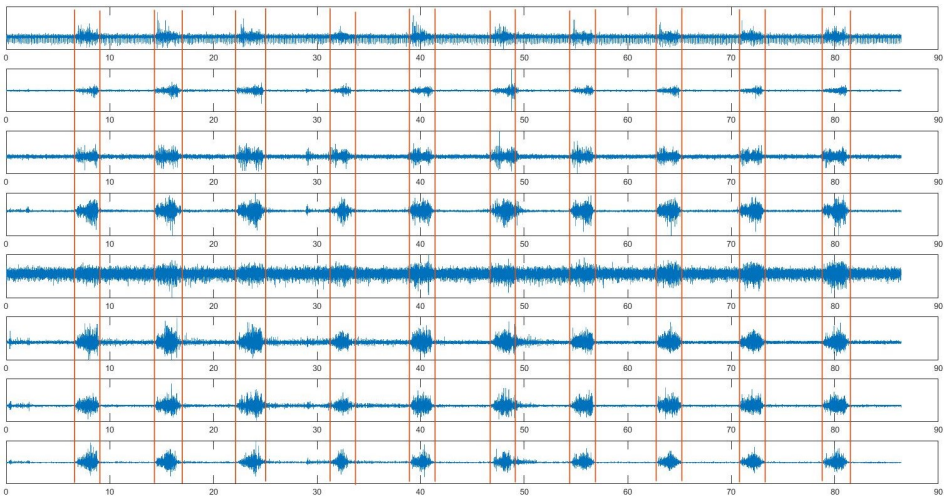


Figure 26: Each activation of all 8 EMG signals are detected.

2.6 Segmentation of signal

EMG signal is highly variable in nature, and a segmentation approach is required to analyse its random and dynamic pattern. Pre-processed EMG signal is not regarded as a useful input to pattern recognition and classification techniques [49], therefore windowing methods are largely adopted to segment and prepare data to further classification. Basically, windowing means sub-dividing the entire signal into small parts, with a proper length, from which features will be extracted. There are mainly two types of windowing methods, which are mostly used in pattern recognition systems: adjacent windowing and overlapping windowing. In adjacent windowing, a predefined length of windows is taken, and the next window starts at the end of preceding one. Features are extracted from these windows. However, this approach of windowing does not generate a dense array of signals and the whole information

may be underutilized [50]. In overlapping window, a predefined window length is chosen, which slides along the whole the signal with an increment size that is less than the original window size, usually expressed in a percentage of window length. In this way, information contained in one window is shared with the adjacent one, creating a that makes full use of the available processing power [51]. The choice of window length is a delicate step since the size of the window has a direct relationship with overall recognition performance. In general, the greater the window size, the higher is the classification accuracy [43]. Indeed, a larger amount of data will result in features with lower statistical variance and, therefore, greater classification accuracy [43]. On the other hand, large window sizes are much more expensive in terms of computational burden (more power is required to extract features from large windows), hence this choice should be a trade off between a real-time delay and classification accuracy. Typically, window sizes should be studied according the specific task that is investigated. However as a general rule in pattern recognition problems, due to real-time constraints, a segment length plus the processing time of generating classified control commands should be equal or less than 300 ms [52]. Moreover, more recent studies suggest that the window size should optimally be kept between 100-250ms [42, 53]. In the present work, we set the window size to 150 ms with an overlap of 75 ms (50% of the window dimension) (Figure 27). Segmentation is performed over the activation phases, where myoelectric information is contained.

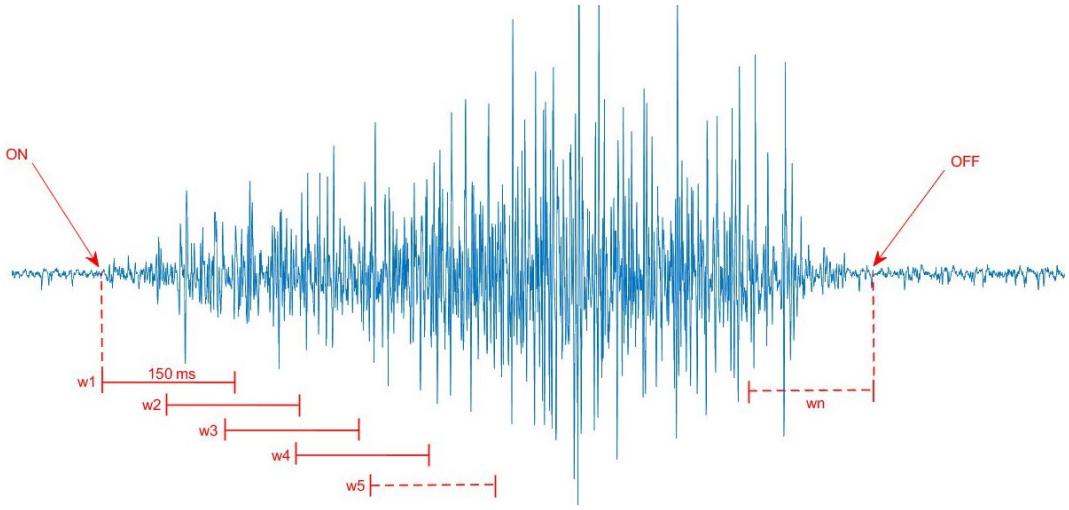


Figure 27: Overlapping windows (150 ms) sliding along the activation burst with an overlapping of 50%.

2.7 Feature extraction

Selection and extraction of highly effective features is one of the most critical stages in pattern recognition and myoelectric control design. Pattern recognition problems depends almost entirely on the selection of proper features [54]. Feeding a classifier with a raw myoelectric signal is impractical due to randomness of the signal. Therefore, the sequence must be mapped into a smaller dimension vector, that is a feature vector [54]. Data features computed over a time window are an approximation of the

true value of a feature, with an associated bias and variance. Indeed, in general features are strongly dependent on the window length and the method of feature extraction. The two most employed characteristics in feature extraction are amplitude and power spectrum. While amplitude and its related features are often investigated in time-domain analysis, power spectrum is studied in a frequency domain analysis. Furthermore, there is another group that mixes time and frequency domains, namely time-scale representation. In the present study a total of 27 features are extracted from each signal segment: 16 time-domain features, 10 frequency-domain features and 1 time-scale domain.

2.7.1 Time-domain features

Time-domain features are the most popular in myoelectric classification because of their computational simplicity, fast implementation, no transformation required and classification performances in low noise environments. On the other side, time-domain features have the disadvantage that amplitude of signals is strongly influenced by electrode location, the thickness of tissues, the distribution of motor units in muscle fiber, muscle conduction velocities, and the detection system used to acquire the signal. Amplitude fluctuations and, more in general, a non-stationarity and randomness of the signal, result in changing of statistical properties over time [13]. Moreover, due to their calculations based on EMG signal amplitude values, high interference levels that affect recording acquisition, results in an important drawback, especially for features that are extracted from energy property [55]. Signal-to-Noise Ratio (SNR) is often adopted as a quality meter of time domain features. Higher values of SNR yield better features [54].

Time-domain features that we have extracted are listed below.

Integrated EMG (IEMG) is defined as a summation of absolute values of the EMG signal amplitude, which can be expressed as:

$$IEMG = \sum_{i=1}^{N} |x_i| \tag{1}$$

where x_i represents the EMG signal samples in a segment i and N denotes length of the EMG signal.

Mean Absolute Value (MAV) is one of the most used feature in EMG signal analysis. This feature is an average of absolute value of the EMG signal amplitude in a segment, which can be defined as:

$$MAV = \frac{1}{N} \sum_{i=1}^{N} |x_i| \tag{2}$$

Variance of EMG (VAR) is defined as an average of square values of the deviation of that variable. However, mean value of EMG signal is close to zero and variance of the EMG signal can also be defined as:

$$VAR = \frac{1}{N-1} \sum_{i=1}^{N} x_i^2 \tag{3}$$

Root Mean Square (RMS) is another frequently used feature. It is modeled as amplitude modulated Gaussian random process whose relates to constant force and non-fatiguing contraction [13]. RMS is

defined as:

$$RMS = \sqrt{\frac{1}{N}\sum_{i=1}^{N} x_i^2} \tag{4}$$

Waveform Length (WL) is defined as cumulative length of the EMG waveform over the time window. It is computed as:

$$WL = \sum_{i=1}^{N-1} |x_{i+1} - x_i| \tag{5}$$

Difference Absolute Mean Value (DAMV) is defined as WL divided by length N minus one. Mathematichally:

$$DAMV = \frac{1}{N-1}\sum_{i=1}^{N-1} |x_{i+1} - x_i| \tag{6}$$

Difference Absolute Standard Deviation Value (DASDV) is similar to RMS feature, it is a standard deviation value of the wavelength [56]. It can be defined as:

$$DASDV = \sqrt{\frac{1}{N-1}\sum_{i}^{N-1} (x_{i+1} - x_i)^2} \tag{7}$$

Zero Crossing (ZC) describes the number of times that amplitude values of the EMG signal cross zero amplitude level.

$$\begin{gathered} ZC = \sum_{i=1}^{N-1} [sgn(x_i \cdot x_{i+1}) \cap |x_i - x_{i+1}| \geq threshold], \\ sgn(x) = \begin{cases} 1, & \text{if} \quad x \geq threshold \\ 0, & otherwise \end{cases} \end{gathered} \tag{8}$$

where threshold condition is implemented to avoid low voltage fluctuations or background noises. We set threshold value for ZC and the three following features (MYOP WAMP and SSC) to 0.001.

Myopulse percentage rate (MYOP) is an average value of myopulse output which is defined as one when absolute value of the EMG signal exceeds a pre-defined threshold value [57].

$$\begin{gathered} MYOP = \frac{1}{N}\sum_{i=1}^{N} [f(x_i)], \\ f(x) = \begin{cases} 1, & \text{if} \quad x \geq threshold \\ 0, & otherwise \end{cases} \end{gathered} \tag{9}$$

Wilson amplitude (WAMP) describes number of times resulting from difference between the EMG signal amplitude among two adjoining segments that exceeds a pre-defined threshold. Moreover, it is strictly related to the firing of MUAPs and muscle contraction force [13]. It is defined as:

$$\begin{gathered} WAMP = \sum_{1}^{N-1} [f(|x_n - x_{n-1}|)]; \\ f(x) = \begin{cases} 1, & \text{if} \quad x \geq threshold \\ 0, & otherwise \end{cases} \end{gathered} \tag{10}$$

Slope Sign Change (SSC) It is a number of times that slope of the EMG signal changes sign. The number of changes between the positive and negative slopes among three sequential segments is performed with a threshold function [13].It is expressed as:

$$SSC = \sum_{2}^{N-1} [f[(x_i - x_{i-1}) \times (x_i - x_{i+1})]];$$
$$f(x) = \begin{cases} 1, & \text{if} \quad x \geq threshold \\ 0, & otherwise \end{cases} \tag{11}$$

Fuzzy Entropy (FuzzyEN) was proposed by [10] as a measure of time-series complexity, providing the conditional probability that two vectors recognized as similar for m samples remain similar for the next $m+1$ samples as well [58]. It is defined as:

$$FuzzyEN(m, n, r) = \ln \phi_m(r, n) - \ln \phi_{m+1}(r, n) \tag{12}$$

where $\phi_m(r, n)$ is a function that takes into account he degree of similarity between embedding vector z_i and $z - j$.

Weighted Permutation Entropy (WPermEN) was proposed by [59]. It is calculated as:

$$WPermEN = -\sum_{i=1}^{d!} p_w(\pi_i) \ln (p_w(\pi_i)) \tag{13}$$

where $(p_w(\pi_i))$. The relative frequencies of all the possible permutations π are computed, thereby producing an ordinal pattern probability distribution.

$$P = \{p(\pi_i)\} \quad i = 1, 2, ..., d! \tag{14}$$

Histogram of EMG (HIST) divides elements in the EMG signal into N equally spaced segments and returns number of signal elements for each segment. We used 9 bin levels, as recommended by [60]. Hence, this feature would account for 9 different numbers.

Auto-regressive coefficients (AR) models individual EMG signals as a linear auto-regressive time series and provides information about the muscle's contraction state [61]. In EMG signal classification, coefficients of the AR model a_p, have been used as a feature vector. The model is a 4th order model, as recommended in literature [62], hence four AR values are calculated for each signal window and used as features. The equation model is:

$$x_i = \sum_{p=1}^{P} a_p x_{i-p} + w_i \tag{15}$$

where P is the order of the model and w_i is a term that takes into account white noise.

Cepstrum Coefficients (CC). A cepstrum of a signal is the result of taking the Fourier transform of the decibel spectrum as if it were a signal. This measure provides information about the rate of change in different frequency spectrum bands of a signal [62] [61]. Cepstrum coefficients are derived from the autoregressive model and calculated as

$$\begin{aligned} c_1 &= -a_1 \\ c_i &= -a_i - \sum_{n=1}^{i-1}(1 - \frac{n}{i})a_n c_{i-1} \end{aligned} \tag{16}$$

where a_i is the ith AR coefficient, c_i is the ith Cepstrum coefficient and i is the dimensionality of the model. A great advantage of CC is that this feature is computed very fast, since it does not require Fourier transform, hence it is still considered a time-domain feature.

2.7.2 Frequency-domain features

Frequency domain analysis is particularly addicted to study muscle fatigue, and to evaluate changes in motor units recruitment [63]. Since the spectrum of EMG signal contains useful information about contraction force (in a time-invariant manner), it can be exploited for pattern recognition problems. Indeed, spectrum is influenced mainly by two factors: the firing rate of a recruited MU in the low-frequency range (below 40 Hz), and the morphology of the action potential travelling along a muscle fiber in a high-frequency range (above 40 Hz) [64]. Power spectral density (PSD) is a major analysis tool in frequency domain. PSD of a signal gives an analysis of the distribution of power over the entire frequency range. The main objective of using this method is to obtain the spectral density estimation from the given data. To calculate PSD, there are 2 mainly approaches. The first method, that is the most commonly used, is the periodogram. It is defined as the square of the absolute value of the Fourier transform of the signal divided by the signal length. In this way the standard deviation of the PSD estimate at each frequency is equal to 100% of the expected value and may be reduced by averaging spectra of nearby windows under the assumption of stationarity over these windows [65]. The second one is a parametric, or model-based approach. It makes use of a mathematical model, under the assumption that signal is a realization of a stochastic process. It is assumed that modeling of the signal is done by considering it stationary in current time windows. Auto-regressive model has been by far the most widely used for spectral estimation. Yule−Walker auto-regressive parametric models are used in this study. Usually a model-order comprised between 4 and 11 is chosen [65]. We used a tenth-order AR model for obtaining the model coefficients. Thus, the corresponding estimate of PSD is obtained by the equation:

$$PSD(nT) = \frac{E_p}{\left|1 + \sum_{n=1}^{p} a_p(k) e^{-i2fnT2\pi}\right|^2} \tag{17}$$

where E_p is the mean square error coefficient, a_p are the model parameters, f is the sampling frequency, and p is the order of the model.

Frequency domain features that we have calculated are listed below.

Mean Frequency (MNF) is an average frequency which is calculated as sum of product of the EMG power spectrum and the frequency divided by total sum of the spectrum intensity. It is calculated as:

$$MNF = \frac{\sum_{j=1}^{M} f_j P_j}{\sum_{j=1}^{M} P_j} \tag{18}$$

where f_j is frequency of the spectrum at frequency bin j, P_j is the EMG power spectrum at frequency bin j, and M is length of the frequency bin.

Median Frequency (MDF) is a frequency at which the spectrum is divided into two regions with equal amplitude. It is defined as:

$$MDF = \frac{1}{2} \sum_{j=1}^{M} P_j \tag{19}$$

Peak Frequency (PKF) is a frequency value at which the maximum power occurs. It can be expressed as:

$$PKF = max(P_j), \text{ where } j = 1, ..., M. \tag{20}$$

Total Power (TTP) is defined as a simply summation of powers of every frequency bin. It is calculated as:

$$TTP = \sum_{j=1} MP_j \tag{21}$$

1*st,*2*nd,*3*rd Spectral Moment (SM)* is an alternative possibility to extract feature from a PSD estimate. Du et al [66] stated that the most important spectral moments are the first three moments, hence this feature contain 3 sub-features, that are SM1, SM2 and SM3. We can define them as :

$$SM1 = \sum_{j=1}^{M} P_j f_j \tag{22}$$

$$SM2 = \sum_{j=1}^{M} P_j f_j^2 \tag{23}$$

$$SM3 = \sum_{j=1}^{M} P_j f_j^3 \tag{24}$$

Frequency Ratio (FR) aims to distinguish between contraction and relaxation of muscle using ratio between the low frequency components and the high frequency components of the EMG signal. It is calculated as:

$$FR = \frac{\sum_{j=LLC}^{ULC} P_j}{\sum_{j=LHC}^{UHC} P_j} \tag{25}$$

where ULC and LLC are upper and lower cutoff frequency of the low frequency band and UHC and LHC are the upper- and lower-cutoff frequency of the high frequency band, respectively. High frequency band and low frequency band are defined by using as a threshold the value of MNF, as proposed by [67].

Power Spectrum Ratio (PSR) is defined as ratio between the energy P0 which is nearby the maximum value of the EMG power spectrum and the energy P which is the whole energy of the EMG power spectrum. It is simply calculated as:

$$PSR = \frac{P_0}{P} = \frac{\sum_{j=f_0-n}^{f_0+n} P_j}{\sum_{j=-\infty}^{\infty} P_j} \tag{26}$$

Variance of Central Frequency (VCF) is derived from spectral moments, and can be calculated as

$$VCF = \frac{SM1}{SM0} - (\frac{SM1}{SM0})^2 \tag{27}$$

2.7.3 Time-scale features

In spectral analysis, a Fourier transform misses time domain information, and it is impossible to localize the correct instant when a particular event happens. Now, if the signal is stationary, not changing properties over time, it is acceptable. However, myoelectric signals contain numerous non-stationary or transitory characteristics. Moreover, a short-time Fourier transform (STFT), mapping data into a two-dimensional function of time and frequency merely obtains this information with a precision that is too much dependant on the chosen window length. To propose another time-frequency method, we though interesting to apply a wavelet transform (WT), because it enables local analysis to be performed. Wavelet analysis, in some circumstances, may reveal data aspects that other techniques miss, such as trends, breakdown points and discontinuities in higher derivatives [54]. WT iteratively transforms the signal of interest into multi-resolution subsets of coefficients by scaling and translating a wavelet function (in our case scaling function is Coiflet 5, shown in Figure 31) over the whole signal, as shown in figure 30. Then the original EMG signal is passed through high-pass and low-pass filters, to carry out both a detailed coefficient subset (cD1) and an approximation coefficient subset (cA1) at the first level. In order to perform a multi-level analysis, repetitive transformations are performed until 3 levels of decomposition are performed (Figure 29). If the decomposition level is set at 3, WT generates respectively the coefficient subsets at the third level approximation (cA3) and the first to the third level details (cD1, cD2 and cD3), as depicted in Figure 28. These 4 level values are used as feature and adapted in a single feature vector.

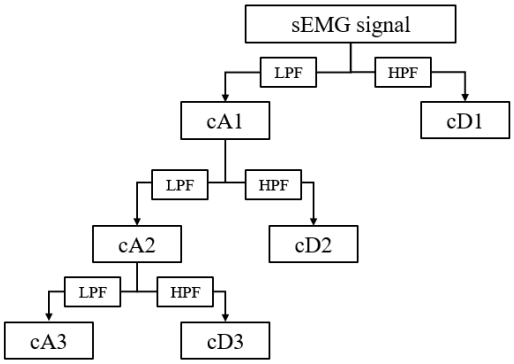


Figure 28: Wavelet decomposition tree. Starting from the sEMG signal, a low pass filter (LPF) is applied to obtain approximate coefficients, while high pass filter (HPF) is applied to obtain detail coefficients. Each split of the three halves the cut off frequency of the filter. Figure adapted from [68].

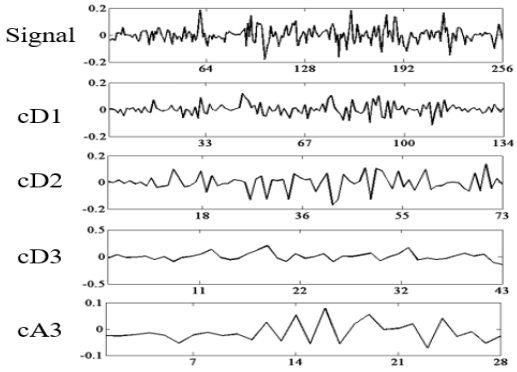


Figure 29: Example of the sEMG signal using wavelet multi-resolution analysis with coiflet wavelet and 3-level decomposition. Starting from the sEMG data (top), signal is decomposed level by level until arriving at the last one in which signal has been low pass filtered, halving three times the cut off frequency.

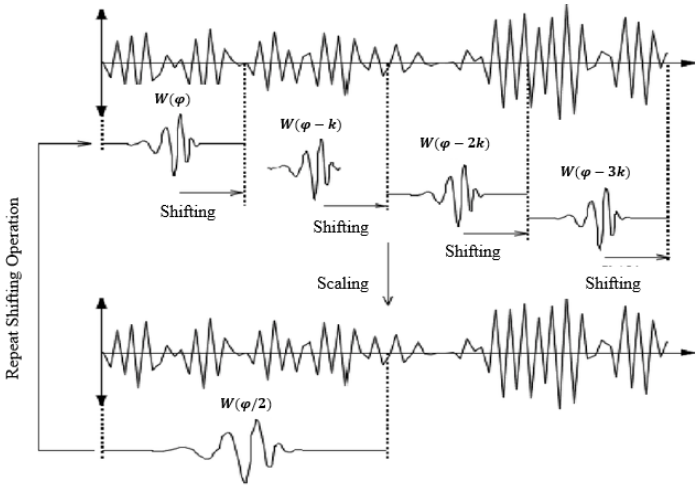


Figure 30: Coiflet wavelet function is shifted all over the signal length, and scaled to adapt its shape to the sEMG pattern. Figure adapted from [69].

Figure 31: Coiflet 5 function. Figure taken from [70].

2.8 Feature selection, aggregation and reduction

Once time-domain, frequency-domain and time-frequency domain features are extracted, they are arranged in a matrix whose number of rows coincides with the total number of segmented windows, and columns contain the value of the features related to all 8 channels. Moreover, a numeric label (going from 1 to 10) is associated to each class. It will be used as target for further classification step, and it is appended in the last column of such matrix. Now that all data are organized, aggregation of features is performed. Aggregating features in a feature set is essential, since the latter represents the actual input to the classifier. A robust EMG feature set should exhibit minimal impact from undesired disturbances, yet remain sensitive to the user's intended gestures [61]. It should guarantee a high-degree of classification accuracy, should be composed by a combination of features that provide a good separation among data (i.e. the distance between different classes in the feature space), while redundancy should be avoided. We calculated six classical feature sets found in literature, listed below.

- **Hudgin's Feature Set (Hudgins) [71]**: Mean Absolute Value (MAV), Waveform Length (WL), Slope Sign Change (SSC) and Zero Crossing (ZC).
- **Du's Feature Set (Du) [72]**: Integrated EMG (IEMG), Variance (VAR), Wilson Amplitude (WAMP),WL, SSC, ZC.
- **Phinyomark's Feature Set 1 (Phinyomark 1) [13]**: MAV, WL, WAMP, ZC, Auto-regressive Coefficients (ARs), Mean Frequency (MNF) and Power Spectrum Ratio (PSR).
- **Phinyomark's Feature Set 2 (Phinyomark 2) [14]**: Weighted permanent entropy (WPermEn), Cepstrum Coefficients (CC), RMS and WL.
- **Time Domain features combined with Autoregressive model coefficients (TDAR) [73]**: MAV, SSC, WL, VAR, WAMP, ARs, and ZC.
- **Discrete Wavelet Transform Coefficients (DWTC) [74]**: cA3, cD1, cD2, cD3.

Robustness of a feature set can be evaluated according some parameters. To asses the goodness of feature sets, we used three different feature space metrics. Moreover, such metrics are exploited to evaluate also the discriminating behaviour of single features. Beyond it, for all 6 feature sets, we calculated the following three indexes.

- **Separability Index (SI)**

  Separability index was suggested to reflect distances between different gesture classes in the EMG feature space [44]. In other words, it evaluates features based on their discriminative power using their content, hence in a manner that is totally independent of the classifier type. We used an adapted version of the separability index found in [75]. The adapted SI was evaluated and defined for each feature set. It is computed as:

$$SI = \frac{1}{10}\sum_{j=1}^{10}(\frac{1}{2}\sqrt{(\mu_j - \mu c_j)^T S^{-1}(\mu_j - \mu c_j)}) \tag{28}$$

where μ_j is the centroid of class j, μ_{Cj} is the centroid of the most conflicting class (with respect to class j), and S is defined as:

$$S = \frac{S_j + S_{Cj}}{2} \tag{29}$$

where S_j is the covariance of class j and S_{Cj} is the covariance of the most conflicting class (with respect to class j) [76]. Larger SI values indicates more distinct classes.

- **Mean-Semi-Principal Axis (MSA)**

  The MSA was proposed as a measure for intra-class variability [44]. MSA considers the feature vectors of each class as a cluster in the shape of a hyper-ellipsoid. Despite SI by itself indicates the distinctness of classes, it may happen that, an increasing in distinctness may result from hyperellipsoids that are smaller or from hyperellipsoids that are farther apart. Thus, the size of the cluster is approximated through singular value decomposition of the feature vector and subsequently calculating the geometric mean of the singular values.

$$MSA = \frac{1}{10}\sum_{j=1}^{10}((\prod_{k=1}^{n} a_{jk})^{\frac{1}{n}}) \tag{30}$$

  where a a_{jk} is the k^{th} of n singular values of class j and n is equal to the number of feature space dimensions. n varies according the number of features contained in each feature set (i.e. for Hudgins set, n is equal to 32, because we have 4 features by 8 EMG channels). Reducing the variability among data results in a lower MSA.

- **Davies Bouldin Index (DBI)**

  Davies Bouldin index is another measure of separability of classes. It describes the goodness and validity of cluster separability in feature space, taking into account two properties, that are compactness (how close data inside clusters are) and isolation (how separate are cluster among them). It is calculated as

$$DB = \frac{1}{K}\sum_{i=1}^{K} max(R_{ij}) \tag{31}$$

  where R_{ij} is cluster similarity and it is computed as

$$R_{ij} = \frac{S_i + S_j}{D_{ij}} \tag{32}$$

  where S_i and S_j are the dispersions of the i^{th} and j^{th} clusters respectively, and D_{ij} is the distance between their mean value.

  In other words, the DB index describes how badly the clusters overlap their closest neighbors [60]. Thus, lower values of the DB index stands for a higher degree of cluster separability. DB strictly depends on geometrical distances between clusters.

2.8.1 Dimensionality reduction

Dimensionality reduction is a data processing step that is mostly applied when dealing with array made up of big dimension feature space. It is used to reduce the dimension of an original feature vector, while preserving the most discriminative information and removing the remaining irrelevant. Several sEMG studies rely on the PCA to interpret muscle activation patterns and better understand the complex coordination of muscle activities in the human body. In particular, PCA technique is largely adopted when hand gesture recognition problems are to be faced [77, 78, 79, 80]. Indeed, beside reducing the dimensionality of feature space, PCA allows a quantative assessment to measure redundancies of feature sets and can identify patterns of coordination [81]. Despite feature sets in our work have modest dimension, we applied Principal Component Analysis (PCA) with 95% of explained variance to reduce the dimensionality of feature sets and decreasing the computational time required by machine processor to train classifiers. The PCA method is a linear technique that projects the data into a lower dimensional space while preserving the variance. This linear transformation converts a set of possibly correlated variables to a set of uncorrelated variables called principal components. The principal components are sorted in order of decreasing variance. Hence, the first principal component has the largest variance and accounts for most of the variability in the data. The principal components (S) are computed as:

$$S = WX \tag{33}$$

where X is an mxn data matrix, in which m is the number of features and n is the number of samples. W is an m x m matrix that contains the eigenvectors of the covariance matrix of the centered X. These eigenvectors are sorted in a descending order of the corresponding eigenvalues. When fetaure space is reduced to size k, the first k rows of S are maintained. The advantage of the PCA is that it is simple, intuitive and computationally fast since the extension to new points is done straight forward, by projecting them into the low-dimensional space.
Before being fed to classification algorithms, feature sets are normalized in the range of 0 and 1.

2.9 Machine learning classification algorithms

Once feature sets are crafted, classifier's input is ready. All six feature sets are used to train and test classification models. We adopted a supervised machine learning approach to perform pattern recognition and classification. In particular, three well known classification algorithms are used: Linear Discriminant Analysis (LDA), Support Vector Machine (SVM) and Random Forest (RF). We performed a within subject study, then these algorithms are trained and validated separately subject by subject, with all 6 feature sets. A brief description of such architectures is reported below.

2.9.1 Linear Discriminant Analysis

Linear Discriminant Analysis is widely adopted as a tool both for dimensionality reduction and multiclass data classification. We will only focus on LDA as a classifier technique. LDA is based upon the

concept of searching for a linear combination of variables (predictors) that best separates two or more classes (targets). The main aim of this algorithm is to maximize the inter-class distance and minimize the within class variance [82]. LDA classifier has been adopted in our study because of the good performance in classification of the EMG signal, the robustness in a long-term effect usage,low computational cost, short training time and its intuitive structure [83]. LDA makes simplifying assumptions about the dataset: as first, it assumes that all variables are distributed in a Gaussian distribution. As second assumption, it considers that each variable varies around the mean by the same amount, on average.

In this way, the LDA model estimates the mean and variance from data for each class. The mean (m) value of each input (x) for each class (k) can be estimated in the normal way by dividing the sum of values by the total number of values:

$$m_k = \frac{1}{n_k} \sum_{i=1}^{n_k} x_i \tag{34}$$

where m_k is the mean value of x for the class k, n_k is the number of instances with class k. The variance is calculated across all classes as the average squared difference of each value from the mean.

$$\sigma^2 = \frac{1}{n-k} \sum_{i=1}^{n_k} (x_i - m_k)^2 \tag{35}$$

where σ^2 is the variance across all inputs x, n is the number of instances, k is the number of classes and m is the mean for input x. In general terms, LDA makes predictions by estimating the probability that a new set of inputs belongs to each class. The class that gets the highest probability is the output class and a prediction is made [84]. Bayes theorem is used to estimate the probability of the output class k, given the input x using the probability of each class and the probability of the data belonging to each class:

$$P(Y = x|X = x) = \frac{(\pi_k \ f_k(x))}{\sum(\pi_l f_l(x))} \tag{36}$$

Where π_k refers to the base probability of each class k observed in training data. In Bayes theorem this is called the prior probability and it is defined as:

$$\pi_k = \frac{n_k}{n} \tag{37}$$

The $f(x)$ above is the estimated probability of x belonging to the class. A Gaussian distribution function is used for $f(x)$. Discriminant function is finally determined as:

$$D_k(x) = x \ (\frac{m_k}{\sigma^2}) - (\frac{m_k^2}{2\sigma^2}) + ln(\pi_k) \tag{38}$$

Where $D_k(x)$ is the discriminate function for class k given input x, the m_k, σ^2 and π_k are all estimated from data.

2.9.2 Support Vector Machine

Support Vector Machine is a supervised machine learning algorithm that is largely employed when classification and regression problems are to be faced. It was adopted since it is able to solve linear and non-linear problems and work well for many multi-class pattern recognition problems. The basic concept behind SVM is simple: the algorithm creates a line (if feature space has 2 dimensions), a plane (if feature space has 3 dimensions) or a hyperplane (if feature space dimension is greater than 3) which separates the data into classes [85]. In the case of linear models, if more than one hyperplane exists, SVM looks for the one that provides the largest margin, i.e. maximizing the distance between the hyperplane and points belonging to different classes that are closest to hyperplane (that takes the name of Support Vectors), hence, in terms of classification, a reduction in mis-classification error. The farther from hyperplane data are, the highest the probability to correctly classify that point. When new test data are added, the model decides what class is to be assigned. Mathematically, we can define an hyperplane as:

$$\vec{w}x - b = 0 \tag{39}$$

where w is a weight vector, x is input feature vector and b is bias. Considering a feature space dimension equal to n, we can write hyperplane equation as:

$$b + w_1x_1 + w_2x_2 + ... + w_nx_n = 0 \tag{40}$$

In other words, an hyperplane is a linear combination of all dimensions equal to zero. All the points that are above and below the hyperplane can be defined as:

$$\begin{aligned} b + w_1x_1 + w_2x_2 + ... + w_nx_n > 0 \\ b + w_1x_1 + w_2x_2 + ... + w_nx_n < 0 \end{aligned} \tag{41}$$

If we consider a binary classification, and including margin borders conditions, we can regulate w_1 and w_2 as follows:

$$\begin{aligned} b + w_1x_1 + w_2x_2 \geq 0 \;\; for \;\; y = +1 \\ b + w_1x_1 + w_2x_2 \leq 0 \;\; for \;\; y = -1 \end{aligned} \tag{42}$$

where y is the class label that can assume only +1 and -1. Now if w represents weights vector and $\|w\|$ is its length, we have that the maximal dimension of margin is:

$$\frac{1}{|w|} + \frac{1}{|w|} = \frac{2}{|w|} \tag{43}$$

Minimizing the weight vector w, we have maximal margin hence determining the optimal hyperplane.

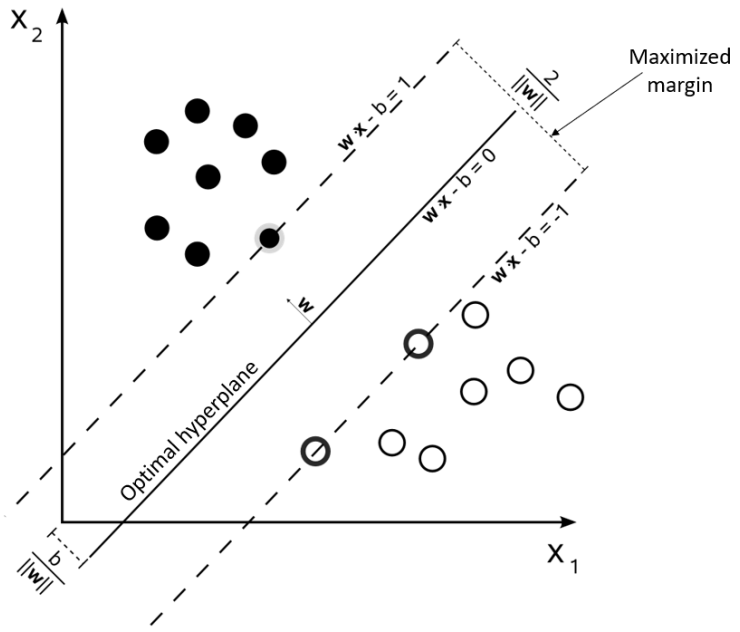


Figure 32: Example of linear separation. Optimal hyperplane is identified when w is minimized.

If the algorithms can't find such hyperplane, a non-linear mapping is used to transform training data in a superior dimension. Thus, when data are not linearly separable, a kernel method is adopted. Such technique basically aims to convert original data into a linearly separable data in higher dimension. By using a non-linear kernel, a non-linear classifier can be obtained without completely transforming data. There exist many types of kernels; the choice of kernel function may heavily influence the performance of SVM model. Anyway, the best way to choose the proper kernel is empirical, and the one that we adopted for our pattern recognition problem is a polynomial kernel of second order, as reported in the equation:

$$K(x_i, y_j) = (x_i + y_j + c)^d, \quad where \;\; d = 2. \tag{44}$$

2.9.3 Random Forest

Random forest is a supervised machine learning algorithm that is constructed from decision tree algorithms. It utilizes ensemble learning, which is a technique that exploits the combination of many classifiers to provide preciser solutions to complex problems. A random forest algorithm is basically an ensemble of many decision trees. It predicts by taking the average of the output from various trees that compounds the forest [86].

A decision tree consists of split nodes and leaf nodes, as shown in Figure 33. Each split node performs a split decision and routes a data sample x to the left child node or to the right child node. To let the tree know where to address the decision, a set of rules is necessary. Rules are under the form of numerical thresholds. The data sample x is routed to the left child node if the value of feature

of x is smaller than a threshold and to the right child node otherwise. All leaf nodes store votes for the classes. Starting at the root node, the data is recursively split into subsets. In each step the best split is determined according specific criteria, like Gini index or entropy level, determining the information gain reached at the leaves of the tree [86]. Classification in random forests employs an ensemble methodology to attain the outcome. The training data is fed to train various decision trees. This dataset consists of observations and features that will be selected randomly during the splitting of nodes. The selection of the final output follows the majority-voting system. In this case, the output chosen by the majority of the decision trees becomes the final output of the model.

RF does not require any assumption of linearity of the data, thus resulting in a good approach when facing non-linear features classification problem [87].

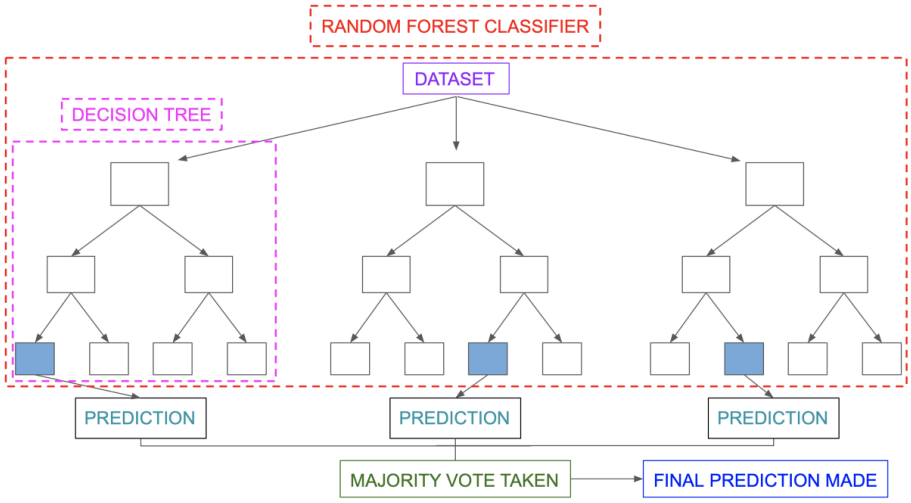


Figure 33: Block scheme of Random Forest algorithm.

Feature sets are under the form of matrices that contain all the numeric values of features (that takes part of that set), appended with a target categorical array. Target indicates to which class each data belongs to. These labels are necessary to the algorithms during training phase, when the machine has to learn relationships between available data and actual data. Moreover, target values are necessary to check the validity of degree of learning of the machine in validation phase, and to test the final classification performance. In machine learning, classification refers to a predictive modeling problem where a class label is predicted for a given example of input data. In our study, we deal with a multi-class classification, since 10 different classes are to be recognized and distinguished. We seek a model that learns the relationship between the input and output variables using the training dataset. Algorithm should learn a relationship that correctly generalizes to new examples beyond the training dataset. This purpose motivates the use of k-fold cross-validation to estimate the performance of the model when making predictions on data not used during training. All the training models (obtained by LDA, SVM nad RF) are validated through a 5-fold cross validation. This latter is a very popular technique used to validate model predictions. In detail, implied algorithm should learn a relationship

that correctly generalizes to new examples. K-fold cross validation randomly divides the dataset into k groups, or folds with almost equal dimensions. First fold is conserved for testing, while model is trained on k-1 folds. This process is repeated K times and, for validation, different folds or different group of data are used at every time. This means that every data point must be in a test set exactly once and in a training test k-1 times. The optimal K fold is the one that better predict over new data, thus reducing overfitting.

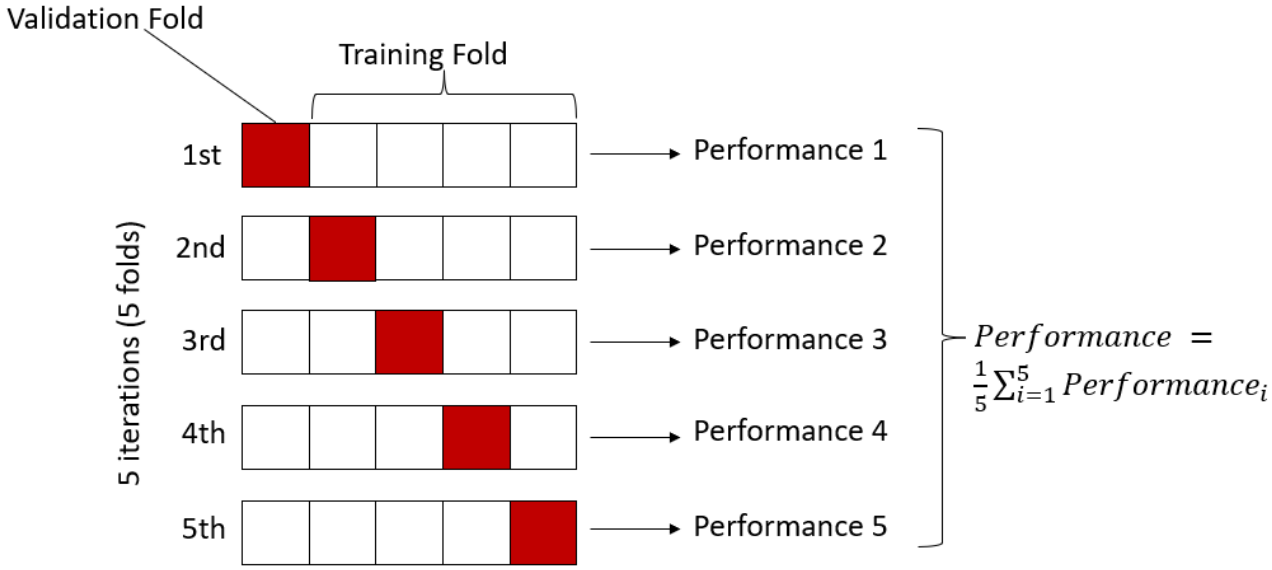


Figure 34: 5-fold cross validation. Dataset is split into k folds. Each fold is then used once as a validation while the k-1 remaining folds form the training set. Validation set moves until the whole length of dataset is covered. Final performance is weighted over single fold performances.

Thus, when making predictions on data not used during testing, K-fold cross validation check out the model validity: it helps us to understand if the model is working properly or if it is overfitting. Overfitting refers to a model that predicts over the training data too well. This phenomenon happens when a model learns the detail and noise in the training data to the extent that it negatively impacts the performance of the model on new data. This means that the noise or random fluctuations in the training data is picked up and learned as right concept by the model. The problem is that these relationships do not apply to new data and deny the trained model to generalize over new inputs. A drawback of K-fold cross validation is that the training algorithm must be re-started from zero for k times, hence spending more time to perform an evaluation.

2.9.4 Classifier performance metrics

In order to evaluate the performance of machine learning classifiers, we evaluated 5 classical meters:accuracy, precision, recall, specificity and F1 score. These parameters are calculated for the three classifiers separately trained with all 6 different feature sets.

- Accuracy, in classification problems, is the number of correct predictions made by the model over all kinds predictions made. It is a measure that should be adopted when the target variable classes in the data are nearly balanced. In our case, there are equal number of samples belonging to each class, there is not a predominance of one class on the other, hence accuracy is a reliable metric. It is defined as:

$$Accuracy = \frac{Number\ of\ correct\ predictions}{Total\ number\ of\ predictions\ made} = \frac{TP + TN}{TP + TN + FP + FN} \tag{45}$$

  where TP (true positive) is an outcome where the model correctly predicts the positive class, TN (true negative) is an outcome where the model correctly predicts the negative class, FP (false positive) is an outcome where the model incorrectly predicts the positive class and FN (false negative) is an outcome where the model incorrectly predicts the negative class.

- Precision, in classification problems, quantifies the number of positive class predictions that actually belong to the positive class. It is defined as:

$$Precision = \frac{TP}{TP + FP} \tag{46}$$

- Specificity, or true negative rate, corresponds to the proportion of negative data points that are correctly considered as negative, with respect to all negative data points, it is defined as:

$$Specificity = \frac{TN}{TN + FP} \tag{47}$$

- Recall, or sensitivity, is the proportion of positive data points that are correctly considered as positive, with respect to all positive data points. It can be expressed as:

$$Recall = \frac{TP}{FN + TP} \tag{48}$$

- F1 score is the harmonic mean between precision and recall. This metric describes how precise the classifier is (how many instances it classifies correctly), as well as how robust it is (it does not miss a significant number of instances). It is one of the most significative parameters, the greater the F1 Score, the better is the performance of our model. Mathematically, it can be expressed as:

$$F1 = 2\ \frac{precision\ \ recall}{precision + recall} \tag{49}$$

3 Results

This section contains all the results carried out from our experiments. We can divide this macro-section into four subsections, each of which reports results corresponding to a specific goal. First subsection describes the results obtained when setup of electrodes is reduced to only wrist or forearm channels. Second one shows the results related to single features. Third one is dedicated to the results related to feature sets and fourth one reports the effect of PCA when it is used to reduce the dimensionality of feature sets.

3.1 Effect of reduced electrodes setup over global performance

This section aims to explore which are the effects over the SVM, LDA and RF classifiers when EMG recording is performed using only wrist channels or forearm channels. Table 1 reports precision values when only forearm channels are used to acquire the EMG signal of each subject, while Table 2 when only wrist channels are considered. Histograms in Figure 35 and Figure 36 allow a visual inspection of these concepts. Notice that overall performance of all six feature sets is unsatisfying, whichever architecture is adopted. When only forearm or wrist channels are considered, a minor amount of information is retrieved, thus translating into poor level of classification precision (at most 64.7% in forearm channels and 70.8% in wrist channel case).

	QSVM	LDA	RF
Hudgins	59.1 ± 3.1	52.3 ± 4.2	58.0 ± 6.2
Du	60.2 ± 4.2	56.2 ± 6.2	58.8 ± 4.3
Phinyomark 1	64.3 ± 2.5	57.1 ± 5.0	56.6 ± 3.9
Phinyomark 2	63.2 ± 3.4	57.1 ± 3.6	56.8 ± 3.7
TDAR	64.7 ± 5.7	60.8 ± 6.5	57.4 ± 6.1
DWTC	45.8 ± 0.3	40.7 ± 2.8	43.9 ± 1.7

Table 1: Mean and standard deviation of precision values in the case in which EMG signal is acquired by only four forearm channels.

	QSVM	LDA	RF
Hudgins	67.2 ± 3.2	60.2 ± 6.5	64.0 ± 5.6
Du	68.2 ± 2.4	64.3 ± 5.6	64.6 ± 7.0
Phinyomark 1	69.4 ± 4.3	63.5 ± 4.7	63.5 ± 6.5
Phinyomark 2	70.5 ± 2.4	64.3 ± 4.0	63.4 ± 7.1
TDAR	70.8 ± 5.7	67.2 ± 5.9	63.9 ± 6.9
DWTC	49.0 ± 1.6	44.9 ± 2.4	49.4 ± 2.3

Table 2: Mean and standard deviation of precision values in the case in which EMG signal is acquired by only four wrist channels.

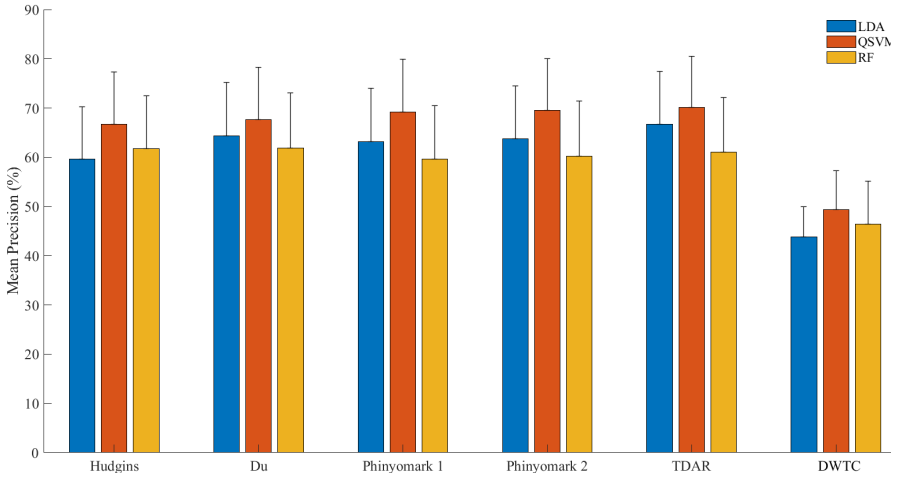


Figure 35: Histogram of average precision obtained by four forearm channels.

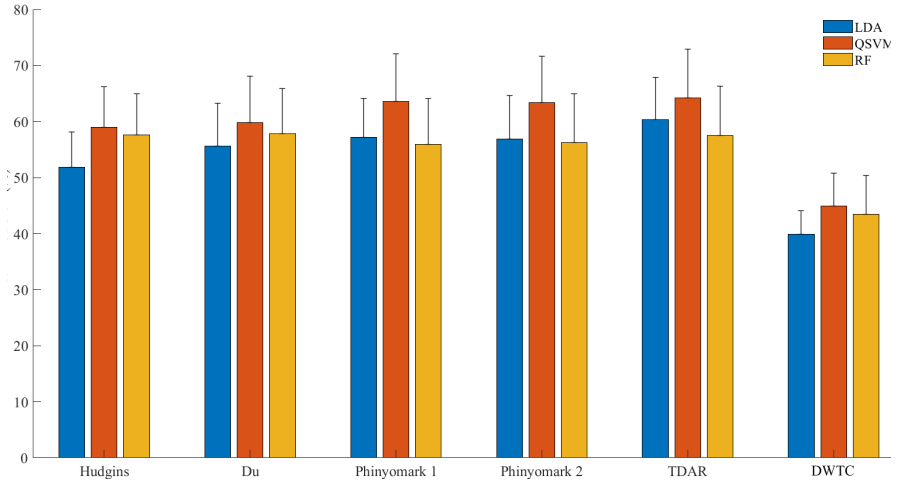


Figure 36: Histogram of average precision obtained by four wrist channels.

3.2 Single feature analysis

This section reports all the findings relative to single feature. Results are reported for time-domain features, frequency-domain features and time-frequency domain features. For each of these group, we report a table relative to the classification precision, namely the outcome obtained when such features are used singularly to feed a QSVM classifier, a table relative to feature metrics, containing MSA, SI and DBI values, and some scatter plots of the most representative features of each group that are useful to understand the discriminating behaviour of single features.

3.2.1 Time-domain features

Table 3 shows all the classification precision values obtained with time-domain features. In bold are the values that achieve highest degree of precision. Notice that MAV, RMS, WL, DAMV, DASDV and IEMG are the ones that, on average, could better predict over ten classes (around 70% if used alone). If used singularly, all the other features can not achieve classification precision higher than 30%. Only HIST(All), AR(All) and CC(All) could rise slightly upper. However they do not account for single feature but they are calculated averaging 9 values for HIST and 4 values of AR and CC.

	Subjects											
	1	2	3	4	5	6	7	8	9	10	11	Mean + Std
MAV	56.3	62.8	58.7	67.8	65.9	55.6	74.5	77.2	69.4	79.4	74.3	**67.4 ±8.4**
VAR	48.7	55.8	47.6	59.6	51.7	47.8	65.2	66.1	52.9	65.7	62.1	56.7 ±7.4
RMS	57.1	61.3	55.6	66.5	61.6	53.9	72.4	74.6	66.1	74.4	70.7	**64.9±7.6**
WL	57.8	70.8	60.2	70.7	70.0	56.6	78.9	81.0	71.9	84.5	81.7	**71.3 ±9.8**
DAMV	57.8	70.5	60.9	69.9	70.9	56.3	79.0	81.2	71.8	84.6	81.6	**71.4 ±9.5**
DASDV	57.4	70.2	58.5	68.3	67.3	54.4	76.8	78.2	71.0	79.3	78.7	**69.1 ±9.0**
IEMG	56.0	63.2	58.7	68.0	65.6	55.5	74.3	76.9	69.5	79.8	74.5	**67.5 ±8.5**
ZC	40.3	23.8	22.9	21.3	29.8	24.3	21.7	25.9	25.8	22.6	15.7	24.9 ±6.2
MYOP	29.4	14.9	21.9	15.5	15.3	15.0	16.4	17.2	20.3	21.0	14.7	18.3 ±4.5
WILSON	29.7	14.9	19.9	15.9	16.8	17.2	17.3	18.5	20.5	21.4	14.6	18.8 ±4.2
SSC	39.2	22.7	22.2	19.0	27.2	22.1	23.2	24.2	25.2	20.3	16.9	23.8 ±5.8
WPERMEN	37.3	22.6	21.9	18.4	26.9	22.0	22.1	22.3	24.8	20.0	15.7	23.1 ±5.6
FUZZYEN	37.7	31.4	24.9	25.3	29.4	25.8	29.1	26.0	28.0	23.5	18.8	27.3 ±4.8
HIST (ALL)	59.0	52.0	41.4	49.1	46.3	47.9	54.9	52.4	50.5	42.2	50.5	49.7 ±5.2
AR (ALL)	50.0	36.0	39.0	31.0	41.0	37.0	37.9	39.7	43.6	35.2	24.2	37.7 ±6.6
AR1	40.9	30.3	28.4	22.2	30.9	23.8	27.2	25.3	28.5	23.6	19.5	27.3 ±5.7
AR2	37	24.3	21.2	16.5	21.5	18.3	22.2	24.0	24.5	20.3	16.0	22.3 ±5.7
AR3	33.9	18.4	20.7	15.1	17.4	16.2	17.7	20.4	20.5	15.3	14.0	19.1 ±5.4
AR4	32.2	15.8	19	16.1	17.8	17	18.3	19.2	21.1	17.1	13.8	18.9 ±4.8
CC (ALL)	49.0	34.0	39.0	29.0	39.0	36.0	37.5	38.1	42.7	33.8	24.8	36.6 ±6.5
CC1	40.7	29.9	28.6	21.6	31.3	23.6	27.1	25.7	28.4	24.1	19.3	27.3 ±5.7
CC2	37.9	25.1	22.6	16.7	22.8	18.7	22.5	23.5	25.1	21.3	16.4	23.0 ±5.8
CC3	30.9	15.7	21.1	15.4	17.0	19.0	16.6	17.9	21.2	15.9	13.1	18.9 ±4.8
CC4	31.5	16.0	18.7	16.6	18.6	17.4	18.9	19.1	20.5	17.1	13.5	18.8 ±4.6

Table 3: Precision of single time-domain features reported for each subject and their mean value and standard deviation.

Table 4 shows feature metrics calculated for each time-domain feature as average of all 11 subjects. In bold are reported the most representative SI values. We highlight the ability of some time-domain features to separate data inside each feature space. Notice that MAV, RMS, WL, DAMV, DASDV and IEMG show SI values higher than 1, thus indicating a great separation among classes. However, a deeper analysis is conducted even over MSA and DBI in chapter 4.

	MSA	SI	DB
MAV	0.0074 ± 0.0014	$\mathbf{1.06 \pm 0.15}$	0.72 ± 0.16
VAR	0.0058 ± 0.0017	0.78 ± 0.11	0.58 ± 0.12
RMS	0.0084 ± 0.0014	$\mathbf{0.99 \pm 0.13}$	0.73 ± 0.16
WL	0.0067 ± 0.0014	$\mathbf{1.15 \pm 0.19}$	0.73 ± 0.15
DAMV	0.0067 ± 0.0014	$\mathbf{1.15 \pm 0.18}$	0.73 ± 0.16
DASDV	0.0075 ± 0.0014	$\mathbf{1.09 \pm 0.16}$	0.73 ± 0.15
IEMG	0.0074 ± 0.0014	$\mathbf{1.06 \pm 0.15}$	0.72 ± 0.16
ZC	0.0143 ± 0.0025	0.32 ± 0.14	0.71 ± 0.06
MYOP	0.0077 ± 0.0056	0.25 ± 0.15	0.65 ± 0.09
WILSON	0.0078 ± 0.0062	0.26 ± 0.16	0.63 ± 0.11
SSC	0.0119 ± 0.0041	0.30 ± 0.15	0.67 ± 0.10
WPERMEN	0.0110 ± 0.0046	0.30 ± 0.16	0.66 ± 0.11
FUZZYEN	0.0113 ± 0.0021	0.35 ± 0.13	0.67 ± 0.08
HIST (All)	0.0121 ± 0.0022	0.88 ± 0.21	1.22 ± 0.16
AR (All)	0.0062 ± 0.0017	0.62 ± 0.18	1.36 ± 0.11
CC (All)	0.0068 ± 0.0017	0.63 ± 0.24	1.33 ± 0.12

Table 4: Mean and standard deviation of MSA, SI and DB value of 11 subjects calculated for time-domain features singularly.

Figures 37,38,39,40 report scatter plots of time-domain features, MAV, WL, DAMV and DASDV respectively. Different colours represents data belonging to different classes. Channel A (abscissa) and Channel B (ordinate) represent 2 random channels inside the same subject EMG recording. All features depicted in scatter plots are extracted from EMG signals belonging to the same subject. Notice distribution of data: all ten classes are distinctly separate in all four charts. SI indexes, previously reported in Table 4, confirm this trend for these four features.

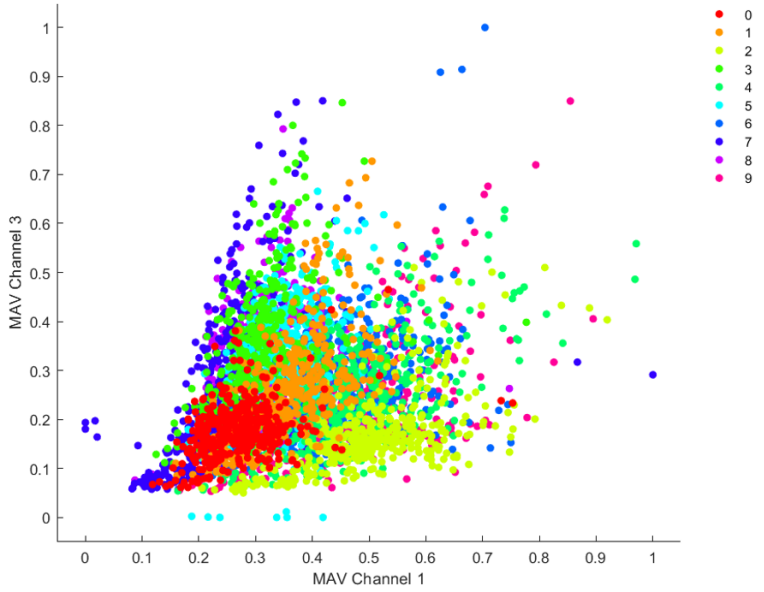


Figure 37: Scatter plot of MAV feature for 2 channels and 10 classes of a single subject

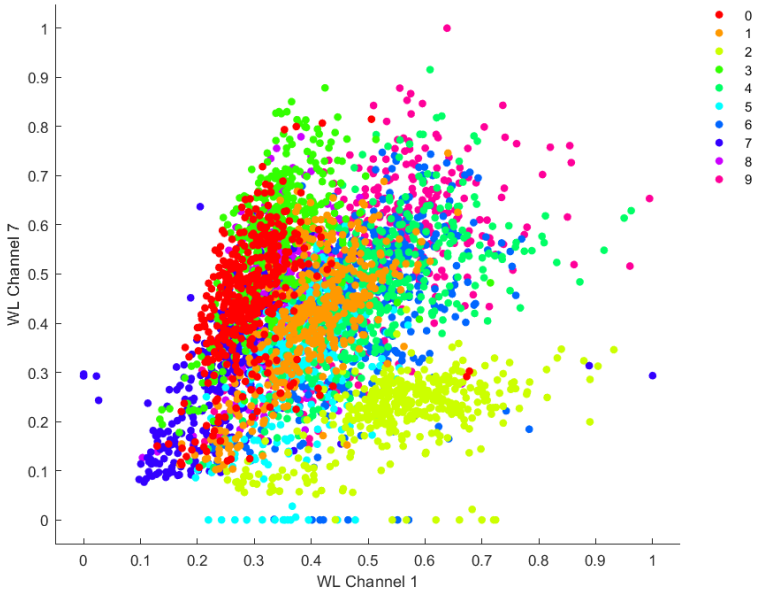


Figure 38: Scatter plot of WL feature for 2 channels and 10 classes of a single subject

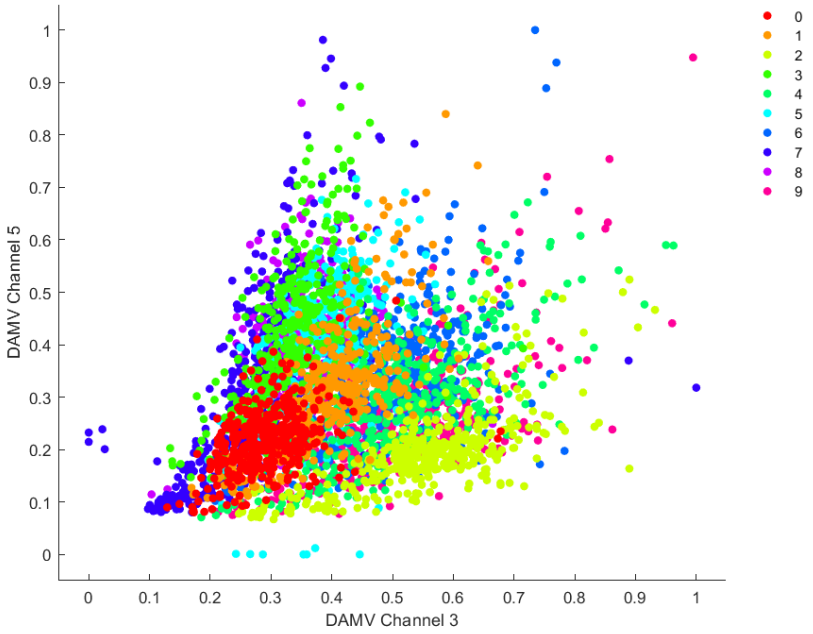


Figure 39: Scatter plot of DAMV feature for 2 channels and 10 classes of a single subject

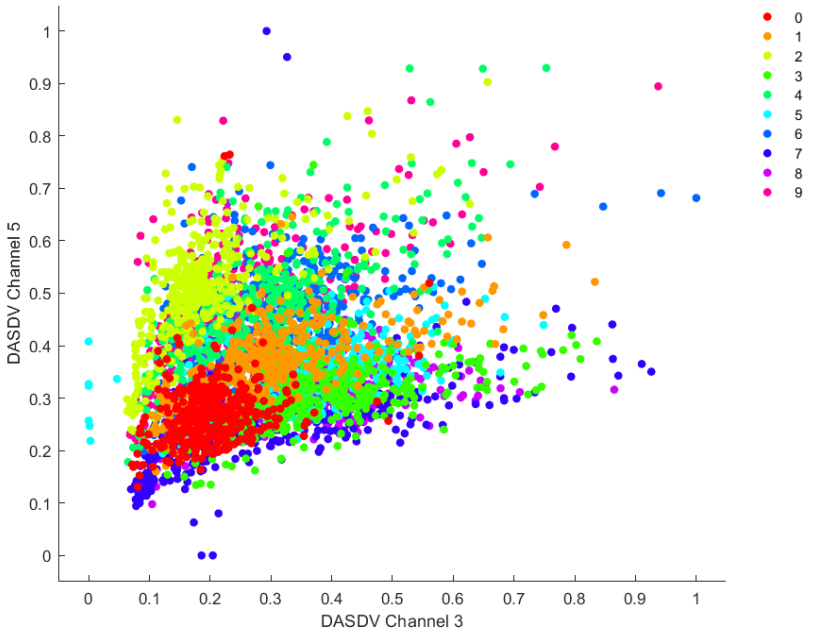


Figure 40: Scatter plot of DASDV feature for 2 channels and 10 classes of a single subject

3.2.2 Frequency-domain features

Table 5 shows all the classification precision values obtained with frequency-domain features. In bold are the values that achieve highest degree of precision. Notice that, with respect to time-domain features, frequency domain features achieves lower level of accuracy. At most, SM3 could predict over classes with a 63% precision. All the other features couldn't provide satisfying results.

	Subjects											
	1	2	3	4	5	6	7	8	9	10	11	Mean+Std
MNF	41.7	28.2	28.8	23.8	36.0	31.2	27.8	28.9	30.8	24.9	19.6	29.2 ± 6.2
MDF	36.1	22.3	25.0	19.0	30.1	27.5	25.5	26.7	25.2	24.1	17.7	25.4 ± 5.0
PKF	43.4	37.8	36.8	45.6	43.3	37.6	49.3	50.0	38.0	45.9	41.7	43.1 ± 6.5
TTP	48.4	56.1	48.4	59.4	51.5	48.0	65.1	66.2	52.8	65.6	62.4	57.0 ± 6.2
SM1	49.3	60.1	49.4	59.9	55.6	48.8	67.3	69.3	56.5	67.8	67.0	59.4 ± 5.3
SM2	49.9	63.4	50.8	61.5	60.6	50.8	69.9	72.0	59.2	69.2	70.8	62.1 ± 6.8
SM3	50.1	65.7	52.3	62.3	62.4	49.9	71.4	73.2	59.2	69	73.1	$\mathbf{63.4 \pm 9.1}$
FR	29.5	15.6	18.0	16.8	24.3	17.2	17.6	17.7	21.8	18.7	15.1	19.5 ± 4.2
PSR	32.1	16.4	17.6	17.0	24.2	18.6	18.4	20.4	19.9	18.9	14.4	21.0 ± 5.3
VCF	37.8	26.2	24.0	22.4	27.2	21.2	29.4	24.9	29.8	23.9	17.1	26.8 ± 4.8

Table 5: Precision of single frequency-domain features reported for each subject and their mean value and standard deviation.

In Table 6 all feature metrics calculated for each frequency-domain feature are reported. In bold are reported the most representative SI values. Notice that SM feature achieve the highest SI value, thus confirming the highest precision values reported in Table 5.

	MSA	SI	DBI
MNF	0.0120 ± 0.0030	0.39 ± 0.12	0.68 ± 0.15
MDF	0.0155 ± 0.0031	0.34 ± 0.18	0.75 ± 0.19
PKF	0.0044 ± 0.0016	0.56 ± 0.21	0.43 ± 0.08
TTP	0.0058 ± 0.0017	0.78 ± 0.14	0.57 ± 0.19
SM1	0.0056 ± 0.0016	$\mathbf{0.83 \pm 0.16}$	0.59 ± 0.15
SM2	0.0051 ± 0.0013	$\mathbf{0.89 \pm 0.19}$	0.58 ± 0.11
SM3	0.0047 ± 0.0012	$\mathbf{0.92 \pm 0.22}$	0.56 ± 0.16
FR	0.0087 ± 0.0035	0.23 ± 0.21	0.58 ± 0.09
PSR	0.0069 ± 0.0044	0.34 ± 0.14	0.51 ± 0.08
VCF	0.0092 ± 0.0017	0.41 ± 0.29	0.64 ± 0.05

Table 6: Mean and standard deviation of MSA, SI and DB value of 11 subjects calculated for frequency-domain features singularly.

Figure 41,42,43, depicts scatter plots of PSR, MNF and SM1 respectively. Notice that, for PSR and MNF data between classes are not well separated, resulting in an overlapping and random distribution. This indicates a low level of separability of data inside feature space that results in a less precise prediction of classifier when it is trained with PSR and MNF features. In the case of SM1, classes result slightly more separated among each other: SI index of this latter one is indeed the highest among all frequency-domain features.

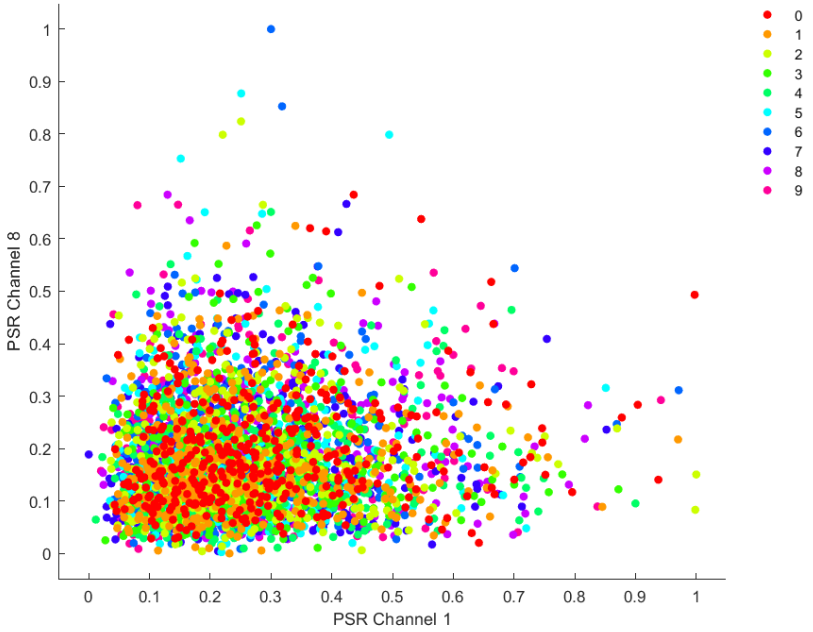


Figure 41: Scatter plot of PSR feature for 2 channels and 10 classes of a single subject.

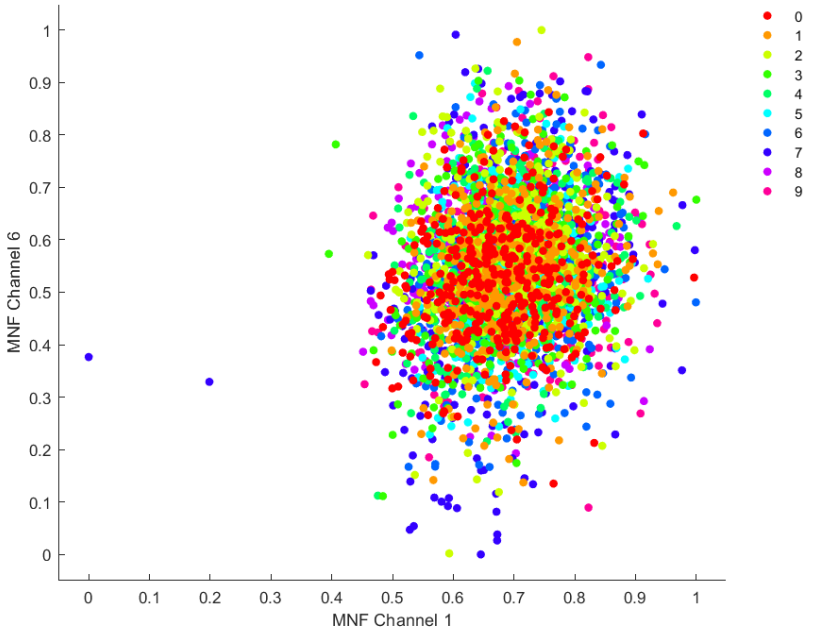


Figure 42: Scatter plot of MNF feature for 2 channels and 10 classes of a single subject.

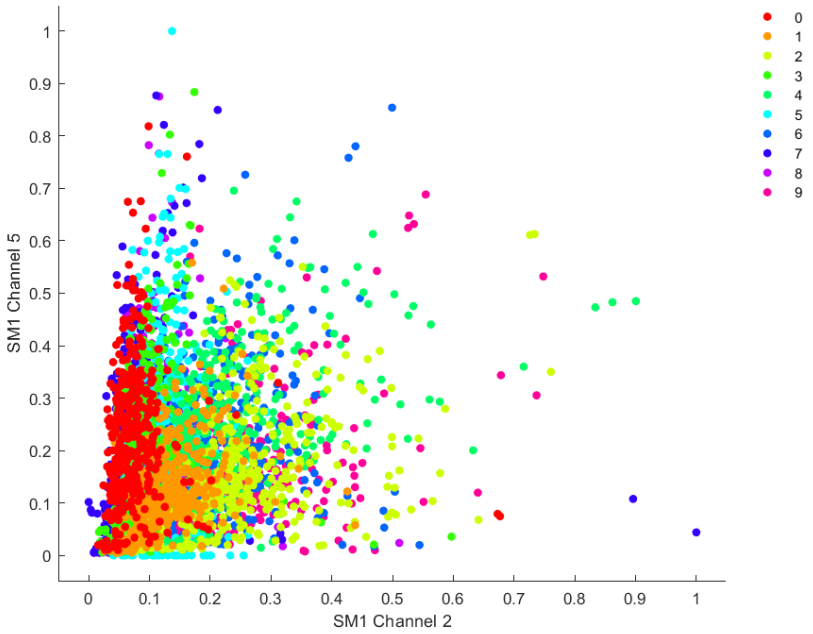


Figure 43: Scatter plot of SM1 feature for 2 channels and 10 classes of a single subject.

3.2.3 Time-frequency domain features

Table 7 shows all the classification precision values obtained when discrete wavelet coefficient features are used to train and test a LDA mode. When evaluated singularly, such features show poor results. Notice that, in terms of classification precision, DWT coefficients achieve, at most, 49.8% for D1.

	Subjects											
	1	2	3	4	5	6	7	8	9	10	11	Mean + Std
A3	46.2	33.8	34.3	40.9	39.5	31.6	41.2	45	36.1	43.3	39.4	39.2 ±5.1
D1	51.2	48.2	45.2	52.5	49.5	38.8	55.4	54.2	51.2	51.2	52.4	**49.8 ±5.3**
D2	49.1	46.3	41.3	45.1	47.6	39	51.2	51.1	47.4	51	48	47.0 ±4.5
D3	50.1	37.7	38	42.7	43.4	36.7	47.4	48.8	41.4	44.3	43.7	43.1 ±4.3

Table 7: Precision of discrete wavelet coefficient (time-scale) features reported for each subject and their mean value and standard deviation.

In table 8, all feature metrics calculated for each time-frequency domain feature are reported. Notice that, SI value calculated on D1 feature, i.e. the best condition, achieve at most 0.70 that is much lower with respect to the best condition of time-domain feature (1.15 for WL in Table4) and frequency-domain feature (0.92 for SM3 in Table 6.

	MSA	SI	DB
A3	0.0083 ± 0.0015	0.53 ± 0.11	0.55 ± 0.12
D1	0.0084 ± 0.0009	$\mathbf{0.70 \pm 0.13}$	0.61 ± 0.10
D2	0.0092 ± 0.0016	0.64 ± 0.19	0.63 ± 0.09
D3	0.0099 ± 0.0021	0.59 ± 0.09	0.64 ± 0.08

Table 8: Mean and standard deviation of MSA, SI and DB value of 11 subjects calculated for discrete wavelet coefficients (time-scale) singularly.

Also in this case, scatter plot of D1 coefficient provide a low degree of separation as one can observe from Figure 44.

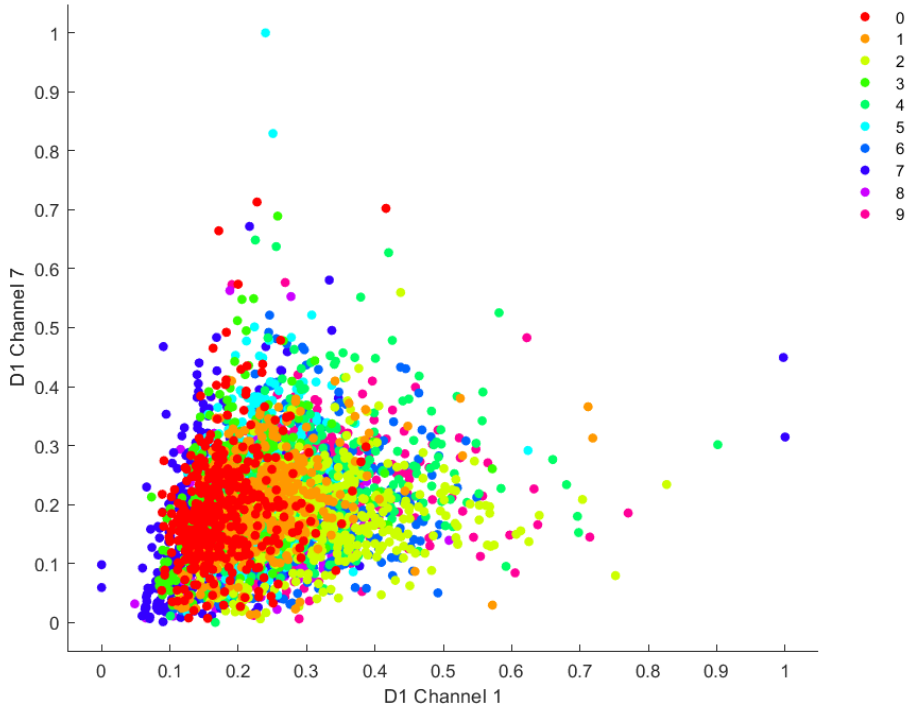


Figure 44: Scatter plot of first detail wavelet coefficient D1 feature for 2 channels and 10 classes of a single subject.

3.3 Feature set analysis

This section summarizes all the results obtained when feature sets are used to train and validate classifier algorithms. Further to this point, Table 9 contains all the values of performance metrics accuracy, precision, recall, specificity and F1 score when each of the six feature set is used as input to QSVM, LDA and RF. Note that promising results are obtained when all 8 channels are taken into examination. Indeed, if compared with those cases in which only forearm or wrist channels were adopted to record EMG input signal, these results are remarkably better. Moreover, in terms of overall performance, feature sets show increasing classification accuracies with respect to single feature cases. However, a detailed examination and comparison of these findings is held in chapter 4. Histogram depicted in Figure 45 graphically summarizes the values of classification precision. It is highlighted that TDAR is the feature set that, among all, performs better.

		Hudgins	Du	Phinyomark 1	Phinyomark 2	**TDAR**	DWTC
QSVM	Accuracy	96.9 ± 1.6	97.0 ± 1.6	97.0 ± 1.5	97.1 ± 1.5	$\mathbf{97.2 \pm 1.5}$	94.1 ± 1.3
	Precision	84.0 ± 7.7	85.0 ± 8.2	84.6 ± 0.7	84.9 ± 7.5	$\mathbf{85.3 \pm 7.6}$	69.2 ± 6.8
	Recall	83.8 ± 7.8	84.7 ± 8.2	84.3 ± 7.7	84.6 ± 7.6	85.6 ± 7.2	69.3 ± 6.8
	Specificity	98.3 ± 0.8	98.3 ± 0.9	98.4 ± 0.9	98.4 ± 0.9	98.2 ± 0.8	97.2 ± 0.7
	F1	83.9 ± 7.8	84.8 ± 8.1	84.4 ± 7.5	84.8 ± 7.6	84.9 ± 7.6	69.3 ± 6.8
LDA	Accuracy	96.0 ± 1.5	96.7 ± 1.5	96.6 ± 1.4	96.5 ± 1.5	$\mathbf{97.3 \pm 1.4}$	92.6 ± 1.1
	Precision	80.2 ± 7.8	83.6 ± 7.8	83.2 ± 7.4	82.7 ± 7.9	$\mathbf{86.2 \pm 7.3}$	61.5 ± 5.8
	Recall	80.2 ± 7.7	83.6 ± 7.7	83.1 ± 7.4	82.7 ± 7.8	86.7 ± 7.3	62.4 ± 5.7
	Specificity	97.8 ± 0.8	98.2 ± 0.9	98.1 ± 0.8	98.1 ± 0.9	98.2 ± 0.8	96.5 ± 0.6
	F1	80.2 ± 7.8	83.6 ± 7.8	83.2 ± 7.4	82.7 ± 7.8	86.5 ± 7.3	62.8 ± 5.7
RF	Accuracy	95.6 ± 1.5	95.7 ± 1.6	94.8 ± 1.6	95.2 ± 1.7	$\mathbf{95.3 \pm 1.5}$	92.9 ± 1.4
	Precision	79.8 ± 9.5	78.6 ± 8.0	74.7 ± 8.0	76.5 ± 8.3	$\mathbf{78.7 \pm 7.6}$	62.8 ± 7.5
	Recall	78.0 ± 7.6	78.5 ± 8.0	74.4 ± 8.1	76.3 ± 8.5	77.1 ± 7.7	62.8 ± 7.3
	Specificity	97.5 ± 0.8	97.6 ± 0.9	97.1 ± 0.9	97.3 ± 0.9	97.5 ± 0.8	96.7 ± 0.7
	F1	83.2 ± 7.6	83.5 ± 8.0	80.8 ± 8.0	81.9 ± 8.4	83.1 ± 7.6	71.4 ± 7.4

Table 9: Performance parameters. Accuracy, precision, recall, specificity and F1 score calculated as mean and standard deviation of all 11 subjects. Values are reported for 6 all feature sets and 3 all classifiers.

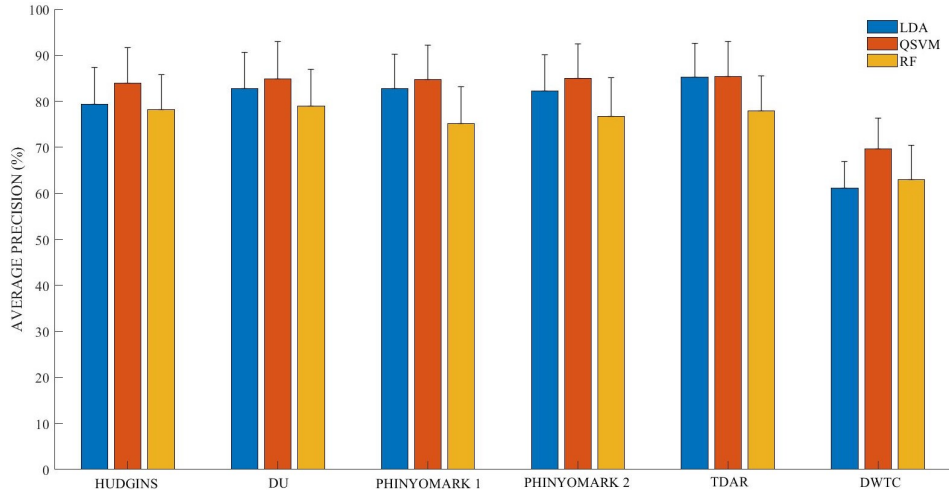


Figure 45: Histogram of average precision of LDA, QSVM and RF grouped by feature sets.

Table 10 shows feature metrics calculated for each feature set. Note that these metrics can be adopted as comparison metrics only if two feature set with equal dimensions are to be compared. It is thus meaningful to compare TDAR with Pninyomark 2 set, Hudgins with DWTC set and Du with Phinyomark 1 set.

	MSA	SI	DB
Hudgins	0.0039 ± 0.0006	$\mathbf{1.60 \pm 0.28}$	1.37 ± 0.20
Du	0.0027 ± 0.0007	1.88 ± 0.38	1.57 ± 0.22
Phinyomark 1	0.0025 ± 0.0006	1.89 ± 0.29	2.00 ± 0.23
Phinyomark 2	0.0042 ± 0.0008	1.90 ± 0.31	1.89 ± 0.21
TDAR	0.0029 ± 0.0007	$\mathbf{2.15 \pm 0.24}$	1.87 ± 0.23
DWTC	0.0059 ± 0.0009	1.02 ± 0.19	1.20 ± 0.18

Table 10: Mean and standard deviation of MSA, SI and DB value of 11 subjects calculated for all 6 feature sets.

Below there are listed six confusion matrices obtained as outcome of a QSVM model. These misclassification matrices are representative if one wants to explore some single case subjects scenarios. In particular, the best and the worst subject case are reported for 3 representative feature sets, Hudgins (Figure 46,47), TDAR (Figure 48, 49), and DWTC (Figure 50, 51). Notice in Figures 47,49,51, one can appreciate the results of a misclassification between task 0 and task 9. Such phenomenon keeps repeating whichever feature sets are used to train the QSVM classifier.

Actual \ Predicted	Task 0	Task 1	Task 2	Task 3	Task 4	Task 5	Task 6	Task 7	Task 8	Task 9
Task 0	97.4	0	0.5	0	0.7	0.5	1.8	0	0.7	0
Task 1	0	96.2	0.7	0.5	6.3	0	0	0	0.5	0
Task 2	0.3	0.3	91.2	0.8	1.7	1.3	0.5	0	0	0
Task 3	0	0.3	1.5	97.6	2.2	0	0.3	0	0	0
Task 4	0.3	3	2.5	1	85.1	1	0	0	1.5	0
Task 5	0.5	0	3.2	0	1	96.1	0.8	0	0	0
Task 6	0.8	0	0.2	0	0	1	94.4	0	2.5	0
Task 7	0.5	0	0.2	0	1.2	0	0	100	0	0
Task 8	0.3	0	0	0	1.7	0	0.5	0	92.2	1.1
Task 9	0	0.3	0	0	0.2	0	1.8	0	2.7	98.9

Figure 46: Classification accuracy of Hudgins feature set for the best subject case.

Actual \ Predicted	Task 0	Task 1	Task 2	Task 3	Task 4	Task 5	Task 6	Task 7	Task 8	Task 9
Task 0	17.4	0	0.3	0	0	0.3	0.3	0	0.3	83.6
Task 1	0.5	94.5	0.8	0.5	2.2	0.5	0.8	0	1.3	0
Task 2	0	0.3	93.7	0.5	1	1.3	1	0	1.3	0
Task 3	0.5	0.5	0	93	3.5	1	0.3	0	1.3	0.5
Task 4	0.5	2.3	0.3	2.8	86.1	1.8	1	0	1.3	1.3
Task 5	0.5	0.3	0.3	1	4.5	89.8	0.5	0	5	0.3
Task 6	1.2	1.1	3	0.5	1	1.6	92.4	0	0.3	0.5
Task 7	0	0	0	0	0.7	0	1	100	0	0
Task 8	0.5	1	1.5	1.6	1	3.4	2.3	0	89	1.3
Task 9	79.1	0	0.3	0	0	0.3	0.3	0	0.3	12.4

Figure 47: Classification accuracy of Hudgins feature set for the worst subject case.

Actual \ Predicted	Task 0	Task 1	Task 2	Task 3	Task 4	Task 5	Task 6	Task 7	Task 8	Task 9
Task 0	97.7	0	1	0	0.5	0	0.8	0	0	0
Task 1	0	98.3	0.5	0.3	6.5	0	0	0	0.5	0
Task 2	0.8	0	92.6	0.8	0.9	0.5	0	0 0	0	
Task 3	0	0	2.5	98.4	2.1	0.3	0	0	0	0
Task 4	0	1.7	1.2	0.3	85.9	0.8	0.3	0	0.3	0.3
Task 5	0	0	2	0	1.2	97.7	0	0	0.3	0
Task 6	0.3	0	0	0	0	0.8	96.7	0	0.8	0
Task 7	1	0	0.2	0	1.2	0	0	100	0	0
Task 8	0.3	0	0	0	1.6	0	0.5	0	96.4	1.1
Task 9	0	0	0	0.3	0.2	0	1.8	0	1.8	98.7

Figure 48: Classification accuracy of TDAR feature set for the best subject case.

Actual \ Predicted	Task 0	Task 1	Task 2	Task 3	Task 4	Task 5	Task 6	Task 7	Task 8	Task 9
Task 0	11.4	0	0.8	0	0	0	0	0	0	88.3
Task 1	0.2	90.6	0	0.5	2.5	0.5	0	1	0.3	0.5
Task 2	1.2	0	97.7	0	0	0.3	0	0	0	1.3
Task 3	0	2.5	0	96.8	3.2	0.3	0	1	0	0
Task 4	0	2	0	1.9	92.5	0.3	0	0	0.3	0.3
Task 5	0	1.7	0	0	1	95.6	0	0	2.1	0
Task 6	0	0.2	0.5	0	0	0.3	99.5	0	0.3	0
Task 7	0.2	2.2	0.3	0.5	0.5	0	0	97.4	0.8	0
Task 8	0	0.7	0	0.3	0.2	2.8	0.5	0.5	96.4	0
Task 9	86.8	0	0.8	0	0	0	0	0	0	9.6

Figure 49: Classification accuracy of TDAR feature set for the worst subject case.

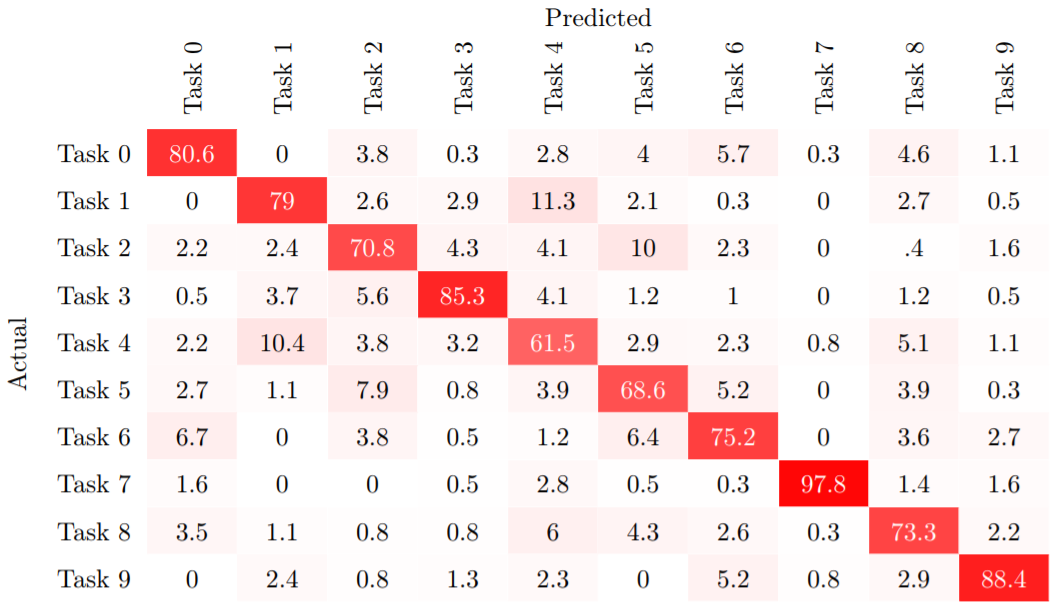


Figure 50: Classification accuracy of DWTC feature set for the best subject case.

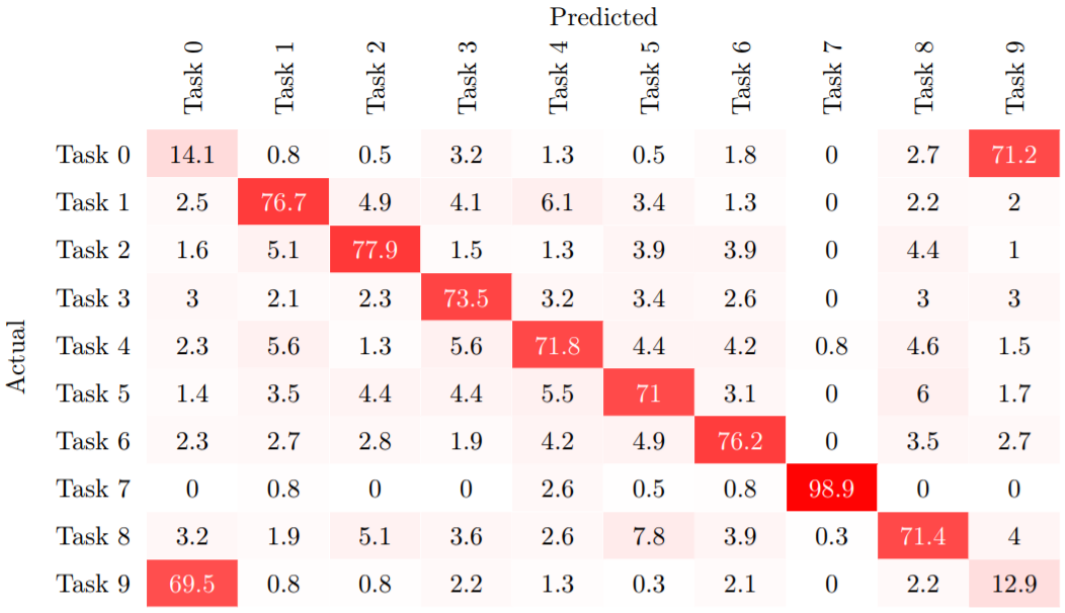


Figure 51: Classification accuracy of DWTC feature set for the worst subject case.

3.4 Effect of PCA dimensionality reduction over global performance

One last objective that we want to explore in this thesis is how the overall handwriting recognition performance is influenced when a dimensionality reduction technique is applied to all six feature sets in question. Further to this point, effects of application of PCA can be appreciated in Table 11. These table reports performance metrics of all six feature sets (exracted from the worst subject case EMG recording) when PCA is applied and when is not applied. By comparing F1 scores reported in the two cases, it is clear that applying this reduction technique decreases drastically classification performances, expecially when LDA and RF are used as classification architecture.

		Hudgins		Du		Phinyomark 1		Phinyomark 2		TDAR		DWTC	
		no PCA	PCA	no PCA	PCA	no PCA	PCA	no PCA	PCA	no PCA	PCA	no PCA	PCA
QSVM	Accuracy	95.3	93.2	95.1	94.5	95.1	94.2	95.3	94.4	95.4	94.6	92.7	92.2
	Precision	76.8	66.8	76.0	73.1	76.4	72.1	77.3	72.8	77.4	73.7	64.4	62.0
	Recall	76.4	65.8	75.3	72.5	75.6	71.1	76.7	72.0	76.8	72.8	63.5	60.9
	Specificity	97.4	96.2	97.3	96.9	97.3	96.8	97.4	96.9	97.4	97.0	95.9	95.1
	F1	76.6	66.3	75.6	72.8	76.0	71.6	77.0	72.4	77.1	73.2	64.0	61.5
LDA	Accuracy	94.4	92.2	95.3	92.2	95.1	93.4	94.9	93.5	95.7	93.4	91.1	90.1
	Precision	72.8	62.0	76.5	62.0	76.0	66.8	74.5	66.9	78.9	66.9	56.1	53.1
	Recall	72.2	60.9	76.4	60.9	75.4	67.1	74.4	67.1	78.6	67.1	55.6	52.7
	Specificity	96.9	95.7	97.4	95.7	97.3	96.3	97.2	96.3	97.6	96.3	95.1	94.2
	F1	72.5	61.5	76.5	61.5	75.7	67.0	74.5	67.0	78.8	67.0	55.8	52.4
RF	Accuracy	94.4	91.5	94.2	91.7	93.4	90.6	93.9	90.5	94.2	90.6	91.8	90.7
	Precision	72.0	57.5	71.0	59.2	67.8	53.1	69.9	53.2	71.0	53.8	60.0	55.2
	Recall	72.1	57.4	70.9	58.6	66.8	52.8	69.5	52.4	70.9	52.8	59.2	54.0
	Specificity	96.9	95.3	96.8	95.4	96.3	94.8	96.6	94.7	96.8	94.8	95.5	94.8
	F1	72.1	57.5	71.0	58.9	67.3	53.0	69.7	52.8	70.9	53.3	59.6	54.6

Table 11: Performance meters for all feature sets when QSVM, LDA and RF are used. For each feature set, there are reported the two cases in which PCA is applied and when it is not applied. This table refers to a single subject case, that is the worst one.

4 Discussion and conclusion

Results are discussed starting from single features analysis and then generalizing over larger feature sets, as reported in next sections.

First of all, for this specific handwriting recognition problem, the optimal setup has shown to be the one that includes the electrodes of the wrist and forearm together. These two sets of probes allow to combine the myoelectric information brought by both anatomical regions contemporaneously. If used separately, wrist channels and forearm channels can't provide satisfactory results, as one can notice from Table 1 and and Table 2 (subsection 3.1). Average precision of three classifiers for all feature set don't rise over 64.7% in the case of forearm channels and 70.8% for wrist channels. As worst case, if predictions are made with a linear classifier (LDA) trained with DWTC dataset, precision falls down to 40.7 %, thus demonstrating the unfeasability of this approach. In contrast to gesture recognition problems where smaller setups are sufficient for a correct gesture classification [11], in handwriting recognition problem, due to due to complexity and fineness of the task, larger setups are required [28].

4.1 Single feature analysis.

When a feature set is to be created, a proper selection of features is crucial. In order to observe and explain the behaviour of each single feature, we evaluated their precision when they are used singularly to train a QSVM classifier. Despite LDA and RF would result in faster training times and less computational cost, QSVM has been chosen as reference classifier since it could achieve higher discrimination accuracy values when it is implied in predicting value over 10 classes, even with single features. Tables 3, 5 and 7 (subsection 3.2) report single subject precision values for each time-domain, frequency-domain and time-scale features respectively. The study carried out that, for handwriting pattern recognition problem, time-domain features could discriminate better between classes. Among time-domain features, MAV, RMS, WL, DAMV, DASDV and IEMG are the ones that can better predict over all 10 classes. Their mean classification precision is around 70% as standing alone. Notice that precision values widely ranges among subjects: in the best subject case, namely subject 10 in Table 3 (subsection 3.3), such features could reach up to 84.6% of classification precision while in worst subject case, i.e. subject 6, decrease down to 56.3%. Moreover, notice that DAMV and WL precision values are strictly correlated: this is due to the fact that DAMV is calculated exactly as WL but averaging it by the number of window samples (Eq. 5,6). It should be noted that WL would be preferred instead of DAMV when a feature set is to be created, since WL requires less computational cost than DAMV [14]. The same considerations can be done for DASDV and RMS. Thus, using DAMV and WL (or DASDV and RMS) together in the same feature set is not properly recommended since it would increase redundancy without supplying additional information. On the contrary, using WL and RMS inside a single feature set could be a good choice since their combined effect provide energy (RMS) and complexity (WL) information contemporaneously [13, 14]. EMG histogram feature is not regarded as a stable feature. The value reported in Table 3 takes into account the average value of 9

singular histogram bins calculated for each segment window. Although the EMG histogram feature is frequently adopted in feature sets that are robust with respect to electrode location stability [76], for this type of application this feature alone does not elicit any augment in classification accuracy when compared with the performance of other time-domain features. Thus, the EMG histogram feature is not included in any feature set that we have analyzed.

Therefore, our study indicates that the EMG histogram feature is not robust when pattern recognition problems are to be faced. Such considerations about histogram feature are concordant with results found in [13]. One interesting trend in our results is the performance of AR and CC features. If AR and CC are singularly evaluated, their performance is poor with an average precision that does not rise upper than 38% and 37% respectively, as reported in Table 3 (subsection 3.2.1). However, inclusion of these two features inside a feature set, adds neural information useful to supply a contribution to increase overall recognition accuracy. Indeed, TDAR feature set combines AR coefficients with 6 other time-domain features while Phyniomark 2 feature set mixes cepstral coefficient with 3 time-domain features more. The effect of adding such feature significantly improves discrimination between classes, as one can observe in Table 9 (subsection 3.3). However, a detailed feature set analysis will be conducted in next paragraph.

Nevertheless, AR and CC features are multi-dimensional, hence inclusion of these features in a feature set increases the dimensionality of feature vectors, which enhances the complexity of the classifier and the computational burden for real-time applications. As shown in Table 5 (subsection 3.2.2), performance of single frequency-domain features is worse than single time-domain features. SM3 is the most reliable among all, with a mean precision of 63.4%. Creating a feature set of only frequency-domain features may not provide good results. At the contrary, combining FD features to TD features can supply additional EMG spectral information. Moreover, DWTC features have a similar behaviour. As reported in Table 7 (subsection 3.2.3), singular discrete wavelet coefficients present a low degree of classification precision, first detail coefficient results at most in a 49.8%.

Validity of the above discussed results is confirmed by feature metrics, SI, MSA and DBI, reported in Tables 4, 6, 8. It has been previously stated that time domain features could provide higher quality results in predicting classes. In particular MAV, RMS, WL, DAMV, DASDV and IEMG were the ones that, in terms of overall precision, showed higher level. If we observe Table 4 and we focus on such features, we realize that MSA SI and DBI reflects the good performance stated by precision levels. In particular, an evident correlation between the separability index and precision is found, as discussed in [76]. As stated in subsection 2.8, the higher the SI value, the more distinct are the classes and the more separate are data clusters. This confirms the highest value of SI for above mentioned features, they are the only ones whose SI rises above 1, up to 1.30 in the best subject case. MSA values are relatively low (0.00067 - 0.00074), indicating a moderate variability in features distribution. DB indexes are almost constant around 0.72, indicating a good degree of cluster separability. Graphically, this can be confirmed by MAV, WL, DAMV and DASDV scatter plots reported in Figure 37 - 40. Features data of all 10 classes are grouped in well separate clusters. Furthermore, it is demonstrated that,

frequency-domain features, in class separability viewpoint, are not suitable for the EMG recognition system. Separability indexes for most of the frequency-domain features ranges between 0.23-0.56. Slightly higher values are those related to SM reaching at most 0.92, confirming scatter distribution depicted in Figure 43. However, scarce classes separability is observable by scatter plots reported in Figures 41,42, in line with results obtained by Phinyomark et al. in a hand gesture recognition problem [13]. There is no evident separation between data and feature clusters are overlapping. The same considerations can be expanded for time-frequnecy domain features. D1 scatter plot, in Figure 44, shows a slightly better ability in separating classes relative to writing 0 and 1.

4.2 Feature set and classification analysis.

It is noticeable from Table 9 (subsection 3.3) that all three machine learning algorithms present a very high degree of accuracy when trained with all six feature sets. An average accuracy of 92.6% is obtained in the worst performing classifier-feature set combination (RF with DWTC) while 97.3% in the best case (QSVM-TDAR). If we evaluate the performance in terms of precision, a decreasing in its average value can be observed for all feature sets. The same holds for recall and F1 score. These three parameters are strictly correlated each other, thus to make a performance analysis, besides accuracy, we can rely on precision value.

TDAR is the feature set that presents the higher degree of precision. When this set is fed to LDA classifier, it carries out an average precision of 86.2%. However, QSVM classifier could correctly predict as well, with an overall precision of 85.3%. It is observable that performance of RF is lower, with an overall precision of 78.7%. Despite the solid performance and reliable result obtained using TDAR, it must be noted that this set has the largest feature space (along with Phinyomark 1 set) among all the others. Indeed, TDAR is composed by 9 features (considering multi-dimension characteristic of AR cofficients array), hence its significant dimensionality is a drawback when real-time experiments are to be conducted. Dealing with a feature set whose dimension is relatively high results in a computational burden and stretched computational times during classification phase. However, from our point of view, TDAR is one of the most suitable set to adopt when an offline handwriting recognition problem is to be faced. TDAR can guarantee a very high class separation ability, SI index is 2.15, much higher with respect to Phinyomark 1 set, that is 1.90 (Table 10).

Hudgins feature set has been largely adopted in many recent EMG classification studies [88, 89]. It's widespread usage derives from its robustness since it represents the best trade off between discriminating performance and low feature space dimensionality. Indeed, the employ of only 4 time domain features ensures a widesprad application when dealing with online pattern recognition problems. In our study, Hudgins feature set gives best classification precision when is fed to QSVM algorithm (84%); lower values are registered when LDA (80.2%) and RF (79.8%) are utilized, as it can be observed in Table 9. It should be noted that, among features that compose this set, there is MAV and WL, which they have been shown to be reliable standing alone features in terms of accuracy (Table 4). The adding of SSC and ZC enhances classification precision, even though disregarding the latter two feature may

not influence too much on performance, as previously explored by [90].

DWTC feature set is the worst performing group of features. Even if accuracy value for this set is satisfying (QSVM 94.1%, LDA 92.6%, RF 92.9%), in terms of precision and recall DWTC has serious deficits. This means that a classifier that is trained with discrete wavelet coefficient values can perform correctly when true positive and true negative are to be identifyied contemporaneously, but deficits when true positive value for classes must be predicted. DWTC achieves higher precision (and recall) when non-linear classification algorithms are used for predicting data, in particular QSVM classifer could achieve at most 69.2% of overall precision with this set. It is still a too poor result to rely on, but it is higher with respect to LDA (61.5%) and RF (62.8%) that are not suitable at all when DWTC is to be used. This confirms what has been demostrated by Lorrain et al in [91], namely that time-frequency domain features can achieve higher accuracy for a properly tuned SVM classifier in a pattern recognition contest [13, 91]. The high non-linear relationships among data points inside feature space of DWTC set does not affect too much SVM algorithm. LDA and RF are instead not useful in this context. Other type of classifiers that are less sensitive to non-linearity could perform better, as demonstrated by Cote et al. [92], where a deep learning convolutional neural network approach was adopted to solve hand gesture recognition problem. Our results suggest that discrete wavelet coefficients are not recommended in combination with the pattern recognition architectures used in this work. Indeed, it is common practice to supply DWT coefficients by additional time domain feature for a major robustness of feature set [93]. Performance metrics confirm the poor suitability of DWTC. Mean SI index is 1.02, namely much lower than Hudgins set (1.60). MSA index underlines a high variability of data while DB index value has a low value (1.20), smaller than Hudgins (1.60). This latter result might be misleading, as one should expect that DBI would increase in DWTC set. This is due to the fact that DBI index takes into account only geometrical distances between clusters, i.e. it relies on euclidean distances between centroids of each class group, without considering any statistical difference among them. Indeed, euclidean distances can distort the metrics to the scale of the larger magnitude features [76]. In this case a more reliable result is provided by SI index that calculates distinctness of classes passing through calculation of covariance matrix of all 10 class data.

Du feature set performance is similar to Hudgins (Table 9). Also in this case, the involvement of a QSVM classifier guarantees higher degree of precision (85%). Anyway, even LDA and RF are suitable, achieving 83.6% and 78.6% respectively. Thus, it can be noted a slight increase in overall performance with respect to Hudgins. However, Du has a feature space dimensionality equal to 6, hence it is computationally more expensive. QSVM predictions are more precise even in the case the classifier is fed with Phinyomark 1 and Phinyomark 2 sets. 84.6% and 84.9% are the overall precision they achieve respectively.

In Figures 46-51, we can observe the classification performances of QSVM classifier trained with best and worse subject case for Hudgins, TDAR and DWTC respectively. In Figure 46, Hudgins feature set is shown to be excellent in terms of accuracy: it distinguishes correctly all 10 classes of handwriting task. Indeed, values in the diagonal of matrix represents the percentage accuracy of correct associa-

tions between target values (vertical axis) and predicted values (horizontal axis). In the worst subject case (Figure 47) a strong misclassification occurred between first and last class. Classifier couldn't apparently distinguish task 0 from task 9. This is due to the fact that number 0 and number 9 share a similar writing pattern. Despite of that, all the other classes are brilliantly predicted. This concept can be extended to all other classifiers. It should be noted that TDAR confirms its discirminating power (Figure 48), and QSVM deficits in correct classification when DWTC is used as training input, even in the best subject case (Figure 50). The misclassification between 0 and 9 keeps repeating when TDAR and DWTC are used to predict classes over the worst case subject (Figures 49,51).

For handwriting recognition, application of Principal Component Analysis has proved to not be a reliable method when feature space dimensionality reduction is desired. As one can observe from Table 11 (subsection 3.4), overall performances of each classifier for each feature set are negatively influenced by application of PCA. If we refer to F1 score to make comparison between original dimension feature set and reduced dimension feature set, we can notice that: for QSVM classifier, Hudgins F1 score decreases of 10.3 %, Du 2.8%, Phinyomark 1 3.4 %, Phinyomark 2 4.6%, TDAR 3.9% and DWTC 2.5%. For LDA classifier, Hudgins F1 score decreases of 10 %, Du 15%, Phinyomark 1 8.7%, Phinyomark 2 7.5%, TDAR 11.8% and DWTC 2.8%. For RF classifier, Hudgins F1 score decreases of 14.6 %, Du 12.1%, Phinyomark 1 14.3%, Phinyomark 2 14.9%, TDAR 16.9% and DWTC 5%. From the previous results, it comes out that PCA is particularly affecting performance of LDA and RF, with minor effect over SVM, though its accuracy decreases. Indeed, on average, F1 score of all feature sets has decreased of 4.58% for SVM classifier. LDA and RF lose an average of 9.3% and 12.97% in F1 score respectively. Even though the feature space dimension is considerabily reduced in every feature sets (on average, Hudgins feature set dimension is reduced from 32 to 17, Du from 48 to 22, Phinyomark 1 from 80 to 39, Phinyomark 2 from 64 to 33, TDAR from 80 to 37 and DWTC from 32 to 24), too relevant decreases in terms of classification accuracy are registered. This diminishing in classifier performance is due to the fact that, when data has in it non-linear relationships, then projecting it onto a linear subspace results in an unfaithful model [94]. Moreover, PCA technique only discovers correlation among patterns and their elements, as well as ordered intrinsic directions where the data patterns change most (with maximum variance), without considering the interaction between the features to discriminate the classes [73]. In this respect, these findings about handwriting recognition differs from results achieved in hand gesture recognition studies, where the application of PCA could increase the multi-class recognition accuracy [78, 80, 95].

In conclusion, we showed that the analysis of myoelectric signal is a reliable method to perform offline handwriting recognition. We demonstrated that, in contrast to hand gesture recognition problems, a larger electrode setup is required. Due to high fineness and complexity of the task, a spatial coverage of forearm and wrist with surface probes contemporaneously is essential to achieve the optimal classification accuracy.

We performed a within subject study where TDAR is shown to be the optimal feature set in terms of

accuracy and precision, while Hudgins in terms of fast computation and real time applicability. Single feature analysis carried out that time domain features are the most suitable in this application. The comparison between three different classification algorithm SVM, LDA and RF highlithed that QSVM is the most appropriate in this pattern recognition problem. However, its better performance is payed with training times that are greater than LDA and RF. Also LDA could discriminate correctly over a multi-class environment while RF suffered more the non-linear distribution of data in feature spaces, hence showing slight decreasing performances. Furthermore, dimensionality reduction didn't provide satisfactory results when applied to all six feature sets.

Further research may be addressed to the analysis of other non-linear approaches to dimensionality reduction in order to keep accuracy and precision as high as possible, and meanwhile, reduce feature sets dimension, thus leading to faster computation of classification outputs. The analysis of single feature and feature metrics held in this research may represent the basis for creation of further robust feature sets, paving the way to new approaches even in online handwriting recognition. Future studies may deal with combination of different handwriting recognition methods, for example exploiting contemporaneously EMG data and kinematic data, using IMU devices in conjunction with EMG electrodes. Moreover, further research could focus on using other types of classifiers to expand the single character recognition to more complex words or sentences analysis and increase even more discriminating performance in myoelectric based handwriting recognition.

References

[1] Samuel Planton, Mélanie Jucla, Franck-Emmanuel Roux, and Jean-François Démonet. The "handwriting brain": a meta-analysis of neuroimaging studies of motor versus orthographic processes. *Cortex*, 49(10):2772–2787, 2013.

[2] Marieke Longcamp, Jean-luc Velay, Virginia Wise Berninger, and Todd Richards. Neuroanatomy of handwriting and related reading and writing skills in adults and children with and without learning disabilities: French-american connections. *Pratiques. Linguistique, littérature, didactique*, (171-172), 2016.

[3] Lambert Schomaker. *Handling within-writer variability and between-writer variation in the recognition of on-line handwriting.* PhD thesis, Citeseer, 1995.

[4] Moussa Djioua and Réjean Plamondon. Studying the variability of handwriting patterns using the kinematic theory. *Human movement science*, 28(5):588–601, 2009.

[5] MONIKA Saini and A Kapoor. Variability in handwriting patterns among ethnic groups of india. *International Journal of Humanities and Social Sciences. 2014b; 1 (3): 49*, 60, 2014.

[6] Charles C. Tappert, Ching Y. Suen, and Toru Wakahara. The state of the art in online handwriting recognition. *IEEE Transactions on pattern analysis and machine intelligence*, 12(8):787–808, 1990.

[7] Jose Guadalupe Beltran-Hernandez, Jose Ruiz-Pinales, Pedro Lopez-Rodriguez, Jose Luis Lopez-Ramirez, and Juan Gabriel Avina-Cervantes. Multi-stroke handwriting character recognition based on semg using convolutional-recurrent neural networks. *Mathematical Biosciences and Engineering*, 17(5):5432–5448, 2020.

[8] Alan M Wing. cd variability in handwritten characters. *Visible Language*, 13(3):283–298, 1979.

[9] Mary Benbow. *Neurokinesthetic Approach to Hand Function and Handwriting: Part 11 & 12.* Clinician's view, 2008.

[10] Weiting Chen, Zhizhong Wang, Hongbo Xie, and Wangxin Yu. Characterization of surface emg signal based on fuzzy entropy. *IEEE Transactions on neural systems and rehabilitation engineering*, 15(2):266–272, 2007.

[11] Fady Botros, Angkoon Phinyomark, and Erik Scheme. Emg-based gesture recognition: Is it time to change focus from the forearm to the wrist? *IEEE Transactions on Industrial Informatics*, 2020.

[12] Jian Wu, Lu Sun, and Roozbeh Jafari. A wearable system for recognizing american sign language in real-time using imu and surface emg sensors. *IEEE journal of biomedical and health informatics*, 20(5):1281–1290, 2016.

[13] Angkoon Phinyomark, Pornchai Phukpattaranont, and Chusak Limsakul. Feature reduction and selection for emg signal classification. *Expert systems with applications*, 39(8):7420–7431, 2012.

[14] Angkoon Phinyomark, Franck Quaine, Sylvie Charbonnier, Christine Serviere, Franck Tarpin-Bernard, and Yann Laurillau. Emg feature evaluation for improving myoelectric pattern recognition robustness. *Expert Systems with applications*, 40(12):4832–4840, 2013.

[15] Martin Petkov. the elbow: joint structure, movements and muscles. `https://www.martinpetkov.com/your-opportunity/the-elbow-joint-structure-movements-and-muscles`, May 2013.

[16] V Hiroshi Tanaka V Naomi Iwayama and V Katsuhiko Akiyama. Online handwriting recognition technology and its applications. *FUJITSU Sci. Tech. J*, 40(1):170–178, 2004.

[17] Charles C Tappert and Sung-Hyuk Cha. English language handwriting recognition interfaces. *Text entry systems: Mobility, accessibility, universality*, pages 123–137, 2007.

[18] Rohini Salunke, Dipali Badhe, Vanita Doke, Yogeshwari Raykar, and Bhushan S Thakare. The state of the art in text recognition techniques.

[19] Ayushman Dash, Amit Sahu, Rajveer Shringi, John Gamboa, Muhammad Zeshan Afzal, Muhammad Imran Malik, Andreas Dengel, and Sheraz Ahmed. Airscript-creating documents in air. In *2017 14th IAPR International Conference on Document Analysis and Recognition (ICDAR)*, volume 1, pages 908–913. IEEE, 2017.

[20] Prasun Roy, Subhankar Ghosh, and Umapada Pal. A cnn based framework for unistroke numeral recognition in air-writing. In *2018 16th international conference on frontiers in handwriting recognition (ICFHR)*, pages 404–409. IEEE, 2018.

[21] Andreas Schröter, Roland Mergl, Katharina Bürger, Harald Hampel, H-J Möller, and Ulrich Hegerl. Kinematic analysis of handwriting movements in patients with alzheimer's disease, mild cognitive impairment, depression and healthy subjects. *Dementia and geriatric cognitive disorders*, 15(3):132–142, 2003.

[22] Peyvand Ghaderyan, Ataollah Abbasi, and Sajad Saber. A new algorithm for kinematic analysis of handwriting data; towards a reliable handwriting-based tool for early detection of alzheimer's disease. *Expert Systems with Applications*, 114:428–440, 2018.

[23] Md Rezwanul Ahsan, Muhammad Ibn Ibrahimy, and Othman O Khalifa. Electromygraphy (emg) signal based hand gesture recognition using artificial neural network (ann). In *2011 4th international conference on mechatronics (ICOM)*, pages 1–6. IEEE, 2011.

[24] Silvestro Micera, Jacopo Carpaneto, and Stanisa Raspopovic. Control of hand prostheses using peripheral information. *IEEE reviews in biomedical engineering*, 3:48–68, 2010.

[25] Miguel Simão, Nuno Mendes, Olivier Gibaru, and Pedro Neto. A review on electromyography decoding and pattern recognition for human-machine interaction. *IEEE Access*, 7:39564–39582, 2019.

[26] Yuri Gloumakov, Joao Bimbo, and Aaron Dollar. Trajectory control for a myoelectric prosthetic wrist. In *MEC Symposium Conference*, 2020.

[27] Reza Boostani and Mohammad Hassan Moradi. Evaluation of the forearm emg signal features for the control of a prosthetic hand. *Physiological measurement*, 24(2):309, 2003.

[28] Michael Linderman, Mikhail A Lebedev, and Joseph S Erlichman. Recognition of handwriting from electromyography. *PLoS One*, 4(8):e6791, 2009.

[29] Gan Huang, Dingguo Zhang, Xidian Zheng, and Xiangyang Zhu. An emg-based handwriting recognition through dynamic time warping. In *2010 Annual International Conference of the IEEE Engineering in Medicine and Biology*, pages 4902–4905. IEEE, 2010.

[30] Chengzhang Li, Zheren Ma, Lin Yao, and Dingguo Zhang. Improvements on emg-based handwriting recognition with dtw algorithm. In *2013 35th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, pages 2144–2147. IEEE, 2013.

[31] Wentao Wei, Qingfeng Dai, Yongkang Wong, Yu Hu, Mohan Kankanhalli, and Weidong Geng. Surface-electromyography-based gesture recognition by multi-view deep learning. *IEEE Transactions on Biomedical Engineering*, 66(10):2964–2973, 2019.

[32] JV Basmajian and CJ Deluca. Muscles alive, 125-167, 1985.

[33] Carlo J De Luca. Physiology and mathematics of myoelectric signals. *IEEE Transactions on Biomedical Engineering*, (6):313–325, 1979.

[34] Roberto Merletti and Philip J Parker. *Electromyography: physiology, engineering, and non-invasive applications*, volume 11. John Wiley & Sons, 2004.

[35] Peter Konrad. The abc of emg. *A practical introduction to kinesiological electromyography*, 1(2005):30–5, 2005.

[36] Steven M Chrysafides, Stephen Bordes, and Sandeep Sharma. Physiology, resting potential. 2019.

[37] Mariam Al Harrach. *Modeling of the sEMG/Force relationship by data analysis of high resolution sensor network*. PhD thesis, Compiègne, 2016.

[38] Wikipedia contributors. Action potential — Wikipedia, the free encyclopedia, 2021. [Online; accessed 8-October-2021].

[39] Muhammad Zahak Jamal. Signal acquisition using surface emg and circuit design considerations for robotic prosthesis. *Computational Intelligence in Electromyography Analysis-A Perspective on Current Applications and Future Challenges*, 18:427–448, 2012.

[40] MA Cavalcanti Garcia and TMM Vieira. Surface electromyography: Why, when and how to use it. *Revista andaluza de medicina del deporte*, 4(1):17–28, 2011.

[41] Leandro Ricardo Altimari, José Luiz Dantas, Marcelo Bigliassi, Thiago Ferreira Dias Kanthack, AC Moraes, and Taufik Abrão. Influence of different strategies of treatment muscle contraction and relaxation phases on emg signal processing and analysis during cyclic exercise. *Computational intelligence in electromyography analysis—a perspective on current applications and future challenges*, pages 97–116, 2012.

[42] Lauren H Smith, Levi J Hargrove, Blair A Lock, and Todd A Kuiken. Determining the optimal window length for pattern recognition-based myoelectric control: balancing the competing effects of classification error and controller delay. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, 19(2):186–192, 2010.

[43] Kevin Englehart and Bernard Hudgins. A robust, real-time control scheme for multifunction myoelectric control. *IEEE transactions on biomedical engineering*, 50(7):848–854, 2003.

[44] Nathan E Bunderson and Todd A Kuiken. Quantification of feature space changes with experience during electromyogram pattern recognition control. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, 20(3):239–246, 2012.

[45] E Van Oudenaarde and RA Oostendorp. Functional relationship between the abductor pollicis longus and abductor pollicis brevis muscles: an emg analysis. *Journal of anatomy*, 186(Pt 3):509, 1995.

[46] Marc A Maier and Marie-Claude Hepp-Reymond. Emg activation patterns during force production in precision grip. *Experimental Brain Research*, 103(1):108–122, 1995.

[47] Lars Arendt-Nielsen, Thomas Graven-Nielsen, Heine Svarrer, and Peter Svensson. The influence of low back pain on muscle activity and coordination during gait: a clinical and experimental study. *Pain*, 64(2):231–240, 1996.

[48] HJA Van Hedel, L Tomatis, and R Müller. Modulation of leg muscle activity and gait kinematics by walking speed and bodyweight unloading. *Gait & posture*, 24(1):35–45, 2006.

[49] Nawadita Parajuli, Neethu Sreenivasan, Paolo Bifulco, Mario Cesarelli, Sergio Savino, Vincenzo Niola, Daniele Esposito, Tara J Hamilton, Ganesh R Naik, Upul Gunawardana, et al. Real-time emg based pattern recognition control for hand prostheses: a review on existing methods, challenges and future implementation. *Sensors*, 19(20):4596, 2019.

[50] Asad Ullah, Sarwan Ali, Imdadullah Khan, Muhammad Asad Khan, and Safiullah Faizullah. Effect of analysis window and feature selection on classification of hand movements using emg signal. In *Proceedings of SAI Intelligent Systems Conference*, pages 400–415. Springer, 2020.

[51] Oluwarotimi Williams Samuel, Mojisola Grace Asogbon, Yanjuan Geng, Ali H Al-Timemy, Sandeep Pirbhulal, Ning Ji, Shixiong Chen, Peng Fang, and Guanglin Li. Intelligent emg pattern recognition control method for upper-limb multifunctional prostheses: advances, current challenges, and future prospects. *Ieee Access*, 7:10150–10165, 2019.

[52] Erik Scheme and Kevin Englehart. Electromyogram pattern recognition for control of powered upper-limb prostheses: state of the art and challenges for clinical use. *Journal of Rehabilitation Research & Development*, 48(6), 2011.

[53] Todd R Farrell and Richard F Weir. The optimal controller delay for myoelectric prostheses. *IEEE Transactions on neural systems and rehabilitation engineering*, 15(1):111–118, 2007.

[54] Mohammadreza Asghari Oskoei and Huosheng Hu. Myoelectric control systems—a survey. *Biomedical signal processing and control*, 2(4):275–294, 2007.

[55] Angkoon Phinyomark, Chusak Limsakul, and Pornchai Phukpattaranont. Emg feature extraction for tolerance of 50 hz interference. In *Proc. of PSU-UNS Inter. Conf. on Engineering Technologies, ICET*, pages 289–293, 2009.

[56] Kang Soo Kim, Heung Ho Choi, Chang Soo Moon, and Chi Woong Mun. Comparison of k-nearest neighbor, quadratic discriminant and linear discriminant analysis in classification of electromyogram signals based on the wrist-motion directions. *Current applied physics*, 11(3):740–745, 2011.

[57] Anders Lyngvi Fougner. Proportional myoelectric control of a multifunction upper-limb prosthesis. Master's thesis, Institutt for teknisk kybernetikk, 2007.

[58] Alessandro Mengarelli, Andrea Tigrini, Sandro Fioretti, Stefano Cardarelli, and Federica Verdini. On the use of fuzzy and permutation entropy in hand gesture characterization from emg signals: Parameters selection and comparison. *Applied Sciences*, 10(20):7144, 2020.

[59] Bilal Fadlallah, Badong Chen, Andreas Keil, and Jose Principe. Weighted-permutation entropy: A complexity measure for time series incorporating amplitude information. *Physical Review E*, 87(2):022911, 2013.

[60] Mahyar Zardoshti-Kermani, Bruce C Wheeler, Kambiz Badie, and Reza M Hashemi. Emg feature evaluation for movement control of upper extremity prostheses. *IEEE Transactions on Rehabilitation Engineering*, 3(4):324–333, 1995.

[61] Dennis Tkach, He Huang, and Todd A Kuiken. Study of stability of time-domain features for electromyographic pattern recognition. *Journal of neuroengineering and rehabilitation*, 7(1):1–13, 2010.

[62] Omry Paiss and Gideon F Inbar. Autoregressive modeling of surface emg and its spectrum with application to fatigue. *IEEE transactions on biomedical engineering*, (10):761–770, 1987.

[63] Mamun Bin Ibne Reaz, M Sazzad Hussain, and Faisal Mohd-Yasin. Techniques of emg signal analysis: detection, processing, classification and applications. *Biological procedures online*, 8(1):11–35, 2006.

[64] Stefan Karlsson, Jun Yu, and Metin Akay. Enhancement of spectral analysis of myoelectric signals during static contractions using wavelet methods. *IEEE Transactions on Biomedical Engineering*, 46(6):670–684, 1999.

[65] Dario Farina and Roberto Merletti. Comparison of algorithms for estimation of emg variables during voluntary isometric contractions. *Journal of Electromyography and Kinesiology*, 10(5):337–349, 2000.

[66] Sijiang Du and Marko Vuskovic. Temporal vs. spectral approach to feature extraction from prehensile emg signals. In *Proceedings of the 2004 IEEE International Conference on Information Reuse and Integration, 2004. IRI 2004.*, pages 344–350. IEEE, 2004.

[67] Mohammadreza Asghari Oskoei and Huosheng Hu. Ga-based feature subset selection for myoelectric classification. In *2006 IEEE international conference on robotics and biomimetics*, pages 1465–1470. IEEE, 2006.

[68] Cheol-Sun Park, Jun-Ho Choi, Sun-Phil Nah, Won Jang, and Dae Young Kim. Automatic modulation recognition of digital signals using wavelet features and svm. In *2008 10th International Conference on Advanced Communication Technology*, volume 1, pages 387–390. IEEE, 2008.

[69] Dimitrios Moshou, Ivo Hostens, George Papaioannou, and Herman Ramon. Wavelets and self-organising maps in electromyogram (emg) analysis. In *Proceedings of the ESIT*, pages 14–15. Citeseer, 2000.

[70] TM Inbamalar and R Sivakumar. Improved algorithm for analysis of dna sequences using multiresolution transformation. *The Scientific World Journal*, 2015, 2015.

[71] Bernard Hudgins, Philip Parker, and Robert N Scott. A new strategy for multifunction myoelectric control. *IEEE transactions on biomedical engineering*, 40(1):82–94, 1993.

[72] Yi-Chun Du, Chia-Hung Lin, Liang-Yu Shyu, and Tainsong Chen. Portable hand motion classifier for multi-channel surface electromyography recognition using grey relational analysis. *Expert Systems with Applications*, 37(6):4283–4291, 2010.

[73] Hu Huang, Hong-Bo Xie, Jing-Yi Guo, and Hui-Juan Chen. Ant colony optimization-based feature selection method for surface electromyography signals classification. *Computers in biology and medicine*, 42(1):30–38, 2012.

[74] Feng Duan, Lili Dai, Wennan Chang, Zengqiang Chen, Chi Zhu, and Wei Li. semg-based identification of hand motion commands using wavelet neural network combined with discrete wavelet transform. *IEEE Transactions on Industrial Electronics*, 63(3):1923–1934, 2015.

[75] Niclas Nilsson and Max Ortiz-Catalan. Estimates of classification complexity for myoelectric pattern recognition. In *2016 23rd International Conference on Pattern Recognition (ICPR)*, pages 2682–2687. IEEE, 2016.

[76] Andreas W Franzke, Morten B Kristoffersen, Vinay Jayaram, Corry K van der Sluis, Alessio Murgia, and Raoul M Bongers. Exploring the relationship between emg feature space characteristics and control performance in machine learning myoelectric control. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, 29:21–30, 2020.

[77] Shunchong Li, Xingyu Chen, Xinjun Sheng, and Xiangyang Zhu. Preliminary study on proportional and simultaneous estimation of hand posture using surface emg based on synergy concept. In *2013 35th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, pages 6199–6202. IEEE, 2013.

[78] Daohui Zhang, Anbin Xiong, Xingang Zhao, and Jianda Han. Pca and lda for emg-based control of bionic mechanical hand. In *2012 IEEE International Conference on Information and Automation*, pages 960–965. IEEE, 2012.

[79] Giulia C Matrone, Christian Cipriani, Emanuele L Secco, Giovanni Magenes, and Maria Chiara Carrozza. Principal components analysis based control of a multi-dof underactuated prosthetic hand. *Journal of neuroengineering and rehabilitation*, 7(1):1–13, 2010.

[80] Mustafa Sezer Erkilinc and Ferat Sahin. Camera control with emg signals using principal component analysis and support vector machines. In *2011 IEEE International Systems Conference*, pages 417–421. IEEE, 2011.

[81] Ganesh R Naik, Suviseshamuthu Easter Selvan, Massimiliano Gobbo, Amit Acharyya, and Hung T Nguyen. Principal component analysis applied to surface electromyography: a comprehensive review. *IEEE Access*, 4:4025–4037, 2016.

[82] Christopher Castaldello, Alessio Gubert, Federico Galvanin, Alessandra Casonato, Roberto Padrini, Massimiliano Barolo, and Fabrizio Bezzo. A model-based support for diagnosing von willebrand disease. In *Computer Aided Chemical Engineering*, volume 40, pages 2779–2784. Elsevier, 2017.

[83] Kevin Englehart, B Hudgin, and Philip A Parker. A wavelet-based continuous classification scheme for multifunction myoelectric control. *IEEE Transactions on Biomedical Engineering*, 48(3):302–311, 2001.

[84] Jason Brownlee. Linear discriminant analysis for machine learning. *Machine learning mastery*, 6, 2016.

[85] SF Ding, BJ Qi, and HY Tan. An overview on theory and algorithm of support vector machines. *Journal of University of Electronic Science and Technology of China*, 40(1):2–10, 2011.

[86] Anne-Laure Boulesteix, Silke Janitza, Jochen Kruppa, and Inke R König. Overview of random forest methodology and practical guidance with emphasis on computational biology and bioinformatics. *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery*, 2(6):493–507, 2012.

[87] Vijay Kotu and Bala Deshpande. Chapter 4 - classification. In Vijay Kotu and Bala Deshpande, editors, *Predictive Analytics and Data Mining*, pages 63–163. Morgan Kaufmann, Boston, 2015.

[88] Guanglin Li, Aimee E Schultz, and Todd A Kuiken. Quantifying pattern recognition—based myoelectric control of multifunctional transradial prostheses. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, 18(2):185–192, 2010.

[89] Guanglin Li, Yaonan Li, Long Yu, and Yanjuan Geng. Conditioning and sampling issues of emg signals in motion recognition of multifunctional myoelectric prostheses. *Annals of biomedical engineering*, 39(6):1779–1787, 2011.

[90] Xin Liu, Rui Zhou, Licai Yang, and Guanglin Li. Performance of various emg features in identifying arm movements for control of multifunctional prostheses. In *2009 IEEE Youth Conference on Information, Computing and Telecommunication*, pages 287–290. IEEE, 2009.

[91] Thomas Lorrain, Ning Jiang, and Dario Farina. Influence of the training set on the accuracy of surface emg classification in dynamic contractions for the control of multifunction prostheses. *Journal of neuroengineering and rehabilitation*, 8(1):1–9, 2011.

[92] Ulysse Côté-Allard, Cheikh Latyr Fall, Alexandre Drouin, Alexandre Campeau-Lecours, Clément Gosselin, Kyrre Glette, François Laviolette, and Benoit Gosselin. Deep learning for electromyographic hand gesture signal classification using transfer learning. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, 27(4):760–771, 2019.

[93] M Hariharan, CY Fook, R Sindhu, Bukhari Ilias, and Sazali Yaacob. A comparative study of wavelet families for classification of wrist motions. *Computers & Electrical Engineering*, 38(6):1798–1807, 2012.

[94] Neta Rabin, Maayan Kahlon, Sarit Malayev, and Anat Ratnovsky. Classification of human hand movements based on emg signals using nonlinear dimensionality reduction and data fusion techniques. *Expert Systems with Applications*, 149:113281, 2020.

[95] YC Du, WC Hu, and LY Shyu. The effect of data reduction by independent component analysis and principal component analysis in hand motion identification. In *The 26th Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, volume 1, pages 84–86. IEEE, 2004.