sEMG-Based Upper Limb Movement Classifier: Current Scenario and Upcoming Challenges

Maurício Cagliari Tosin  
(Corresponding author)  
MAURICIO.CTOSIN@GMAIL.COM

Juliano Costa Machado  
JULIANOC.MACHADO@GMAIL.COM

Alexandre Balbinot  
ALEXANDRE.BALBINOT@UFRGS.BR

Federal University of Rio Grande do Sul

Abstract

Despite achieving accuracies higher than 90% on recognizing upper-limb movements through sEMG (surface Electromyography) signal with the state of art classifiers in the laboratory environment, there are still issues to be addressed for a myo-controlled prosthesis achieve similar performance in real environment conditions. Thereby, the main goal of this review is to expose the latest researches in terms of strategies in each block of the system, giving a global view of the current state of academic research. A systematic review was conducted, and the retrieved papers were organized according to the system step related to the proposed method. Then, for each stage of the upper limb motion recognition system, the works were described and compared in terms of strategy, methodology and issue addressed. An additional section was destined for the description of works related to signal contamination that is often neglected in reviews focused on sEMG based motion classifiers. Therefore, this section is the main contribution of this paper. Deep learning methods are a current trend for classification stage, providing strategies based on time-series and transfer learning to address the issues related to limb position, temporal/inter-subject variation, and electrode displacement. Despite the promising strategies presented for contaminant detection, identification, and removal, there are still some factors to be considered, such as the occurrence of simultaneous contaminants. This review exposes the current scenario of the movement classification system, providing valuable information for new researchers and guiding future works towards myo-controlled devices.

1. Introduction

Electromyography (EMG) is the study of electrical manifestation resulting from neuromuscular activation in the occurrence of muscle contraction (De Luca 1979). Although the study of muscles has its importance recognized for centuries, the focus on researches regarding the electrical activity is recent. The first documented acquisition of surface electromyography signal dates back to 1849, being performed by the German physiologist Emil Heinrich Du Bois-Reymond through a primitive type of galvanometer (Pearce 2001; Pozzo et al. 2004). Since then, the study of electromyography has been important to several areas, from clinical medicine (Benazzouz and Slimane 2021; Parisi and RaviChandran 2020) and sports, for example, improving the performance of high-performance
athletes (Fronso et al. 2017; Judge et al. 2003) to biomedical engineering (Cene et al. 2019a; Schweisfurth et al. 2020; Trigili et al. 2019).

The raw surface electromyography (sEMG) signal has a strong stochastic nature as it is constituted by the synergy of several muscles (Anam et al. 2019; Reaz et al. 2006). Nevertheless, the sEMG signal provides valuable information related to upper limb kinematics, which can be extracted using appropriate processing techniques. Hence, the use of sEMG has intensified especially in biomedical engineering, with several researches related to the development of sEMG based Human Machine Interface (HMI) aiming the design of myo-controlled prosthesis (Atzori et al. 2012; Cene et al. 2019a; Krasoulis et al. 2020; Mantilla-Brito et al. 2020; Schweisfurth et al. 2020; Yamanoi et al. 2020), exoskeleton (Lu et al. 2019; Ma et al. 2019; Trigili et al. 2019), virtual reality environments for rehabilitation (Bouteraa et al. 2019; van Dijk et al. 2016; Yang et al. 2017), orthosis (Bos et al. 2020), teleoperation systems (Ye et al. 2019; Zhou et al. 2019), among others. All the research lines that are associated with the development of EMG based HMI systems are based on algorithms to translate the motion intention, usually using machine learning techniques to perform the device control.

Despite achieving hand-arm gesture recognition accuracy rates of over 90% in laboratory environments with state-of-the-art classifiers, there are still challenges for myoelectric devices to perform equally well in everyday environmental conditions. Currently, the difficulties that guide the search for solutions to increase the robustness of current systems in real situations can be categorized into five factors:

*Limb position:* it is associated with the muscle activation necessary to keep the limb at rest due to the action of gravitational force. Furthermore, depending on the arm spatial position, a displacement between the electrode-skin interface and muscle can occur. Both situations could cause artifacts in the sEMG signal, leading to misclassification due alteration in the signal pattern compared to the signal used to train the motion classifier (Campbell et al. 2020);

*Intensity of the muscle contraction:* the muscle contraction intensity is unconsciously controlled according to the expected effort to perform a specific action and it has been verified that there is a direct relationship between the intensity of contraction and the amplitude of the EMG signal (De Luca 1997). The change in muscle action intensity can even lead to changes in signal frequency characteristics (Campbell et al. 2020);

*Electrodes displacement:* consists of the electrode position displacement. When this occurs, the underlying musculature changes in relation to the sensor and even if the same muscle fibers are under the scope of the electrode, with the displacement there is a change in the biological tissue impedance, implying changes in the signal properties (Campbell et al. 2020);

*Temporal factor:* the sEMG signal is influenced by several biochemical, physiological or anatomical mechanisms that are time-varying (e.g. blood flow). The electrode placement variation due removal and replacement of the prosthesis between sessions can also be included, as well changes in the contraction intensity between different uses of the device (Campbell et al. 2020);

*Signal contamination:* there are intrinsic and extrinsic factors in the acquisition process that could change the characteristics of the sEMG signal. They can be related to motion artifact, environmental electromagnetic interference, electronic components noise, interference of other biological signals
as electrocardiography (ECG), among others. The contaminants might induce variations in time, frequency, morphological and statistical properties of the sEMG signal (Ijaz & Choi 2018; De Luca et al. 2010; McCool et al. 2014).

Considering this, there are currently no commercial applications utilizing interfaces based on sEMG, with the exception of neuroprosthesis (Kaczmarek et al. 2019). However, successful EMG-controlled prosthesis still employ only 2 bipolar electrodes located in the forearm extensor and flexor muscles (Wang et al. 2019), which limits the number of predicted movements, typically enabling sequential control of 2 or less degrees of freedom (DOF) (Dewald et al. 2019), and, consequently, the use of the artificial arm. In addition, sequential control strategies are still far from providing natural movements to the user, requiring a high level of practice and training (Atzori & Muller 2015), which configures one of the major obstacles to the full acceptance of the device.

However, researchers have been working in the search for solutions to obtain a system that is, at the same time, capable of recognizing a high number of movements and that is robust in relation to the aforementioned issues. Therefore, the problem could be analyzed from different points of view, and new approaches are proposed for each stage of the movement classification process. Thus, this paper aims to present a literature review focused on works related to the development of sEMG based upper limb movement classifiers. In this way, the main goal of this review is to expose the latest in terms of strategies in each block of the system, thus enabling a global view of the current state of academic research and guiding new researchers by presenting the roles that still need to be fulfilled in the area.

The review is structured into nine sections. The section “EMG based movement classification task” presents the block diagram of the motion classification system. The subsections “Signal Acquisition”, “Pre-Processing”, “Feature Extraction”, “Feature Selection”, “Classification / Regression” and “Robotic Arm” present a review of the most recent works proposed for the respective stages of the system, aiming to contribute with the description of innovative approaches for, but not limited to minimizing the effect of issues concerning Limb Position, Intensity of the Muscle Contraction, Electrode Displacement and Temporal Factor. The review of the works related to the effect of Signal Contamination is presented in a separated subsection named “Analysis of the presence of contaminants in the sEMG signal”, due to this factor being a generic problem, covering all sEMG applications, and, consequently, covering works associated with not just motion classification tasks. The contaminant effect is often neglected in reviews focused in sEMG based motion classifier and myo-controlled devices, so this subsection is a contribution of this paper to the academic research. Closing the review, a “Conclusion” section presents the conclusions and future trends into development of myo-controlled devices for upper limb assistance/rehabilitation/substitution.

2. EMG-based Movement Classification Task

The upper limb movement classification task can be divided into 6 blocks, as the diagram in Figure 1 show. In the first step, the EMG signal is acquired from the volunteer through a non-invasive method, with surface electrodes or invasively, through electrodes inserted percutaneously, that are
connect to a proper conditioning circuit which includes, at least, a filtering stage for noise removal, followed by an amplification stage for further digitization. Next, the digitized signal goes through a pre-processing stage in which additional filters, normalization, rectification, segmentation, among other procedures, could be applied. Once pre-processed, features can be extracted, in time, frequency or time-frequency domain, in order to highlight the useful information present in the signal. After this phase, the signal, now represented by the features calculated in the previous step, is submitted to a Feature Selection stage where the more relevant and informative for the classification task are selected, while the others are discarded. The remaining attributes from the Feature Selection step are then used to train a predictive model of motion, motion trajectory, or force depending on the application. Finally, the classifier output is used to control the robotic arm (prosthesis, orthosis, exoskeleton, virtual limb) making it perform the movement required by the user.

![Figure 1: Typical block diagram of an EMG-based motion identification system.](image)

In recent years, many works have been conducted in this area, exploring each of the steps described above and all with the common goal of advancing a step further towards the development of a robust and naturally controlled myoelectric system for an artificial limb. To present the current state of academic production, the next subsections will bring a description of the most recent works published, organized by stage of the classification task explored, according to Figure 1.

2.1. EMG Signal Acquisition

The EMG signal acquisition plays an important role in the movement classification process, since it is extremely important to perform the measurement in the proper limb location and with sufficient signal-to-noise ratio (SNR). Ensuring the quality of the EMG signal allows the necessary information to be extracted for model training and it is also necessary to have an adequate amount of signal samples, volunteers, movements, test conditions, among other controllable factors, and make the data publicly available. Thus, statistical robustness is obtained in the validation of the proposed algorithm, and, at the same time, allows others researches to use the same data to perform a fair comparison between methods.

In this context, many works have been conducted recently with this purpose. They can be divided into: creation and availability of sEMG database (Cene et al. 2019a; Cognolato et al. 2020; Du et al. 2017; Kaczmarek et al. 2019; Palermo et al. 2017; Pizzolato et al. 2017), development of hybrid...
acquisition systems (sEMG combined with a signal of another nature) (Jiang et al. 2020; Xia et al. 2019), proposal of low-cost acquisition systems (Besma et al. 2019; Islam et al. 2019; Prakash et al. 2019; Shaabana et al. 2019), development of wearable sEMG measurement devices (Abass et al. 2019; Lee et al. 2020) and related to the method used in the signal acquisition, which covers both the choice of the type of electrode (Dewald et al. 2019) and the way which they are fixed (Fonseca et al. 2019; Tamura et al. 2020).

Table 1 shows a description of the works that focused in provide publicly available sEMG datasets of hand gestures. The works conducted by Palermo et al. (2017), Du et al. (2017), Cene et al. (2019) and Kaczmarek et al. (2019) stand out for performing acquisitions with the same volunteer on different days. Thus, all temporal variability sources present in the sEMG signal, already mentioned, are included in the available data. Hence, it is possible for a confounding factor to be considered in the development and evaluation of classification systems designed by other researchers, enabling the creation of more robust algorithms.

<table>
<thead>
<tr>
<th>Authors</th>
<th>N° channels</th>
<th>Conf.</th>
<th>Additional input</th>
<th>#Mov.</th>
<th>Repetitions per session / sessions</th>
<th>Acquisitions in different days?</th>
<th>I. Vol.</th>
<th>A. Vol.</th>
<th>Issue considered</th>
</tr>
</thead>
<tbody>
<tr>
<td>(Cognolato et al. 2020)</td>
<td>12</td>
<td>LD</td>
<td>Accelerometer</td>
<td>10</td>
<td>8 / 1</td>
<td>No</td>
<td>3 (F)</td>
<td>2 (F)</td>
<td>Lumbar position</td>
</tr>
<tr>
<td>(Kaczmarek et al. 2019)</td>
<td>24</td>
<td>LD</td>
<td>No</td>
<td>8</td>
<td>20 / 2</td>
<td>Yes</td>
<td>8 (F)</td>
<td>36 (M)</td>
<td>Temporal variability</td>
</tr>
<tr>
<td>(Cene et al. 2019a)</td>
<td>12</td>
<td>LD + MT</td>
<td>No</td>
<td>17</td>
<td>6 / 6</td>
<td>Yes</td>
<td>1 (F)</td>
<td>3 (M)</td>
<td>Temporal variability</td>
</tr>
<tr>
<td>(Pizzolato et al. 2017)</td>
<td>12</td>
<td>LD + MT</td>
<td>No</td>
<td>52</td>
<td>6 / 1</td>
<td>No</td>
<td>4 (F)</td>
<td>6 (M)</td>
<td>Acquisition setup</td>
</tr>
<tr>
<td>(Pizzolato et al. 2017)</td>
<td>16</td>
<td>LD</td>
<td>No</td>
<td>52</td>
<td>6 / 1</td>
<td>No</td>
<td>2 (F)</td>
<td>8 (M)</td>
<td>Temporal variability</td>
</tr>
<tr>
<td>(Palermo et al. 2017)</td>
<td>14</td>
<td>LD</td>
<td>Accelerometer</td>
<td>7</td>
<td>12 / 10</td>
<td>Yes</td>
<td>3 (F)</td>
<td>7 (M)</td>
<td>Temporal variability</td>
</tr>
<tr>
<td>(Du et al. 2017)</td>
<td>128</td>
<td>HD</td>
<td>No</td>
<td>8</td>
<td>10 / 1</td>
<td>No</td>
<td>18</td>
<td>No</td>
<td>Temporal variability</td>
</tr>
<tr>
<td>(Du et al. 2017)</td>
<td>128</td>
<td>HD</td>
<td>No</td>
<td>8</td>
<td>10 / 2</td>
<td>Yes</td>
<td>10</td>
<td>No</td>
<td>Temporal variability</td>
</tr>
<tr>
<td>(Du et al. 2017)</td>
<td>128</td>
<td>HD</td>
<td>No</td>
<td>12</td>
<td>10 / 1</td>
<td>No</td>
<td>10</td>
<td>No</td>
<td>Temporal variability</td>
</tr>
</tbody>
</table>

In (Palermo et al. 2017), sEMG signals acquired from 10 intact subjects are made publicly available. The experiment consisted of performing seven squeezing movements on 14 different objects. Each volunteer performed 2 sessions per day, one in the morning and one in the afternoon, on five different days. A single session consists of 12 repetitions of each movement. These data form Dataset #6 of the NinaPro project. Du et al. (2017) recruited 23 non-amputated volunteers to perform trials of 8 or 12 basic hand movements, forming three different datasets. Ten of the 23 recruited subjects performed 2 sessions with 10 repetitions of eight distinct movements, on two different days spaced at least 1 week apart.
In contrast, in (Cene et al. 2019a) the objective was to create a dataset with a greater number of sessions per volunteer and with variations in their respective protocols. Four different methodologies were applied in the experiments: Type A - 6 repetitions and sequential execution order; Type B – 10 repetitions and sequential execution order; Type C - 6 repetitions and randomized order of each motion execution; Type D – 10 repetitions and random order of each motion execution. All protocols comprise the same 17 hand and forearm movements, each being performed 3 times by the 4 recruited subjects, i.e., a single volunteer runs 12 trials. During data collection, a maximum limit of four sections per day for the same volunteer was respected. Thus, it allows the analysis of both effects: the movement execution randomization that is something little explored in the available databases, and the temporal factor, since there are trials performed on different days by the same subject.

The confounding factor associated with limb position was considered in the database proposed by Cognolato et al. (2020). Each volunteer executed four of the eight repetitions while seated and four while standing, including the position limb variability in the acquired data.

Pizzolato et al. (2017) evaluated the influence of the acquisition set up on the data quality by creating two databases with two different acquisition hardware. In one of them, the device from Cometa manufacturer Mini Wave model (wireless sEMG measurement system) was used along with 12 gelled electrodes, while in the other, 2 Myo Armband units were considered, totaling 16 dry electrodes. Both datasets are constituted by 10 intact subjects performing the same 52 movements repeated 6 times, forming NinaPro dataset #4 (Cometa) and #5 (Myo Armband). The two acquisition systems were compared in terms of the accuracy of the movement classifier. Through the classification results obtained it was found that the accuracy achieved with Myo Armband hardware is comparable to that obtained with the Cometa (69.04 and 69.13% respectively), despite the price of the first being lower than 1/30 of the second.

Myo Armband is a sEMG acquisition device that also incorporates a hand movement classification system. Created by Thalmic Labs, it comprises eight differential dry electrodes, as well as inertial sensors such as 3-axes gyroscopes, magnetometers, and accelerometers (Zea & Benalcázar 2020). The favorable vote for the use of the Myo Armband acquisition system corroborates its intense use in recent researches in the area. In fact, numerous works have reported satisfactory results with the use of this device. To cite some examples, (Xu et al. 2020) obtained a 94.7% assertiveness rate in the identification of 6 movements with Artificial Neural Networks (ANN), (Zea and Benalcázar 2020) proposed a classification method based on an LSTM (Long Short-Term Memory) neural network, achieving an accuracy of 95.79% in the recognition of 5 movements and the work by Mantilla-Brito et al. (2020) which reached 88.02% of accuracy in the identification of 3 motions with Naive Bayes (NB) method in an online experiment.

However, despite the positive emphasis given to Myo Armband in (Pizzolato et al. 2017), the device suffers some criticism in (Shaabana et al. 2019). According to the authors, the 200 Hz sampling frequency is insufficient, since it limits the frequency band of the measured signal to a range from zero to 100 Hz and it is known that the sEMG has components between 6 and 500 Hz, leading information loss. Another negative point raised was the fact that it is not possible to place the electrodes in the ventral region of specific muscles, since the 8 available channels are equally spaced along the circumference of the band. Through an experiment, it was verified an increase of 4%, in
absolute value, in the classification accuracy of 6 finger movements with a method based on the Hidden Markov Model (HMM) when 8 electrodes were fixed in specific muscle regions of the forearm in relation to the use of the same 8 sensors placed equally spaced along its circumference.

As well as recommending the use of the Myo Armband system, the issue of the electrode fixation setup is also an open topic. According to (Wang et al. 2019) the electrode placement strategy can be divided into 3 groups: muscle-targeted layout with sensor placed in the ventral region, low density layout, that is typically formed by sets of 2 to 16 electrodes, and disregards the exact location of the muscular ventral region, being normally evenly distributed on the skin surface, and high density layout, that is a non-invasive technique that collects signals through several electrodes with inter-electrode distances smaller than 5 to 10 mm. This is observed in the variety of the channel number and placement configurations in the works listed in Table 1.

Despite the recommendation to measure the sEMG signal always in the ventral region of the muscle in order to ensure maximum amplitude (de Luca 1997), many researchers have adopted the high density configuration, using matrices with 128 (Chen et al. 2020b; Du et al. 2017; Jaber et al. 2020; Olsson et al. 2019b), 168 (Martinez et al. 2020) and up to 192 electrodes (Chen et al. 2020a). The main advantage of using this form of electrode placement is to obtain a greater spatial resolution of muscle activation, enabling the use of well-established techniques for image processing, as Convolutional Neural Networks (CNN) (Chen et al. 2020b; Olsson et al. 2020; Olsson et al. 2019b; Yang et al. 2019a). In this case, typically the sEMG signal is converted into a grayscale image, where the amplitude of each electrode in the array indicates the pixel intensity, before feed the network. Another advantage is that make the system less sensitive to the lack of an electrode, unlike what occurs in the muscle-targeted layout (Chen et al. 2020b). The main disadvantage of this strategy is the high cost and complexity, in addition to being interesting to reduce the number of electrodes in order to simplify the system and reduce the computational cost (Wang et al. 2019).

Finally, there is still a discussion regarding the nature of the signal to be used in the movements characterization, with researchers betting on hybrid systems, i.e., with other types of signal besides sEMG (DelPreto & Rus 2020; Jiang et al. 2020; Krasoulis et al. 2020; Xia et al. 2019). In the work by Krasoulis et al. (2020), promising success rates (86.5%) were achieved in the identification of 6 movements with a Regularized Discriminant Analysis (RDA) based classifier and sEMG signal acquired from only two channels combined with inertial measurements such as acceleration, angular velocity and orientation, performed by 3-axes accelerometers, magnetometers, and gyroscopes, respectively.

Jiang et al. (2020) propose a customized acquisition system, that is wearable and integrate sEMG and Force Myography (FMG), a pressure sensor that characterizes muscle strength. This system makes it possible to measure both sEMG and FMG at the same point, and a 91.6% accuracy result was achieved in 10 movement classification using Linear Discriminant Analysis (LDA) classifier.

In (Xia et al. 2019), a hybrid sEMG and ultrasonography acquisition system was developed, also in a compact and wearable form, enabling measurements of both signals in the same location of the limb. One of the advantages raised was the possibility to analyze the activity of deeper muscles located several centimeters under the skin, through ultrasound. Reported results showed a 20%
increase in the accuracy associated with the recognition of 20 wrist and finger movements when considering both signals (sEMG and ultrasound) than in relation to just sEMG (89.2% a 68.59%).

In this way, it was demonstrated that the information present in the sEMG signal can be complemented with the measurement of other parameters of the limb and muscles, to increase its ability to discriminate movements and improve the performance of the system. Thus, the acquisition of signals of another nature in addition to EMG appears as a promising alternative to increase the robustness of myoelectric control systems.

According to the data shown in Table 1, there are many sEMG databases with a large number of volunteers and movements covering different electrode arrangement configurations. All of them were generated under ideal laboratory conditions with the volunteer following a pre-defined protocol for executing movements and with minimal noise. Although these bases are essential for the development of motion classification algorithms, they do not provide subsidies to verify the performance of such systems in real situations. Considering this, there is a need to create EMG databases acquired under non-ideal conditions, closer to the user's reality. Here we can mention the execution of tests with the user in motion (walking on a treadmill for example) while performing the gestures, in environments contaminated with electromagnetic radiation, among other factors. In this way, it would make it possible for algorithms to be developed already taking into account these non-idealities, reducing the gap between the laboratory environment and the actual application of the device.

2.2. Pre-Processing

After the acquisition and conditioning of the EMG signal, a pre-processing stage can be applied before the feature extraction. At this stage, the signal is conditioned by applying additional filters. A fault detection stage can be used, to analyze if some of the channels are corrupted by intrinsic or extrinsic sources, and if possible, mitigate using signal reconstruction, contaminant removal, elimination of contaminated channel, among others strategies. Others pre-processing stages can be normalization, segmentation, transformation to highlight certain signal properties, to name a few. Thus, it is important to guarantee a proper signal conditioning to achieve maximum information related to the task in subsequent processing phases. Therefore, studies have been published in recent years with new proposals for pre-processing stage that appears as alternatives to improve the classifiers performance (Chen et al. 2020a; He et al. 2019; Machado et al. 2020; Sezgin 2019; Tam et al. 2020; Wahid et al. 2020; Xu et al. 2020; Zhou et al. 2020). Among the gains brought by these works, we can mention the improvement in the quality of the acquired signal, minimization of the effects of electrode displacement, temporal variation, and inter-subject variation in the characteristics of the EMG signal, to cite a few, which induce an increase in the performance of the movement classification system.

This subsection is dedicated to the analysis of the most recent works associated with pre-processing stage of the EMG-based movement classification process. However, those related to contaminant analysis will not be covered here, as they are described in a specific section. The others are summarized in Table 2.
The most frequently problem addressed from pre-processing stage is data quality improvement. The strategies range from digital filters to data transformation methods. Digital filtering of the sEMG signal is an almost mandatory step in the motion classification process, being used in most works. However, there is no definition of which filter topology is the most suitable, and to contribute to the clarification, Powar and Chemmangat (2019) performed an experiment comparing the performance of three filter topologies in terms of the accuracy of identification of 8 hand and wrist gestures with kNN (k-Nearest Neighbor): Butterworth, Spectrum Subtraction (SS) and Wiener. The first was marginally superior to the others with an assertive rate of 73.3%, against 67.2% (SS) and 65.0% (Wiener). However, SS provided lower computational cost, while Wiener was the most efficient in noise removal among the three. The last was the most indicated by the authors for muscle activation detection tasks.

Also exploring signal filtering, Phukan et al. (2019) proposed a methodology to remove high-frequency noise based on the Wavelet Transform. The algorithm is based on the maintenance of low-frequency coefficients and the application of a filter on the high-frequency components, in which all those that are less than a pre-defined threshold value are zeroed. The threshold is
determined from the statistical properties of the high-frequency coefficients. Afterward, the inverse transform procedure is applied, resulting the representation of the low frequency. Then, the resultant signal is summed to the modified high-frequency components to obtain the filtered signal. Promising results were achieved in the identification of 10 finger movements with only two measurement channels and SVM (Support Vector Machine) classifier (accuracy of 96.5%), confirming the proposed algorithm as a good strategy for suppression of unwanted artifacts in the sEMG signal.

Sheng et al. (2019) and Sezgin (2019) proposed new pre-processing strategies to address the issues related to the temporal and inter-subject variability of the sEMG signal. First, the authors presented an algorithm called Comom Spatial-Spectral Analysis (CSSA) to find what they defined as “common mode” in the sEMG signal. The method, which includes the application of the Wavelet Transform, is based on the premise that multivariate time series, as the sEMG signal, with spatial and temporal resolution, can be decomposed into a stationary and a non-stationary portion. Thus, the “common mode” of the signal is defined through its stationary portion. Sezgin (2019), in turn, considered bispectral analysis, which comprises a statistical method used to find non-linear relationships between signal components (Sezgin 2019). The representation of the signal bicoherence is determined based on the bispectral analysis, through bispectrum normalization, which is used in the feature extraction step.

Both pre-processing approaches mentioned above showed promising results in the classification of 13 movements with LDA (Sheng et al. 2019) and 5 gestures with ELM (Extreme Learning Machine) (Sezgin 2019). However, the CSSA method proved to be more efficient in minimizing the signal temporal variability factor, while the algorithm based on the bicoherence analysis obtained interesting hit rates in an experiment where the classifier was trained with data acquired from multiple volunteers. So, they propose an algorithm called Interpolated Peak Location (IPL) that consists of verifying the location of the muscle activation peak in the sensor array for a given reference movement, and then this value is compared with the one obtained after the electrodes replacement to evaluate a possible change in positioning.

The issue associated with electrode displacement was the focus of research by Xu et al. (2020) and He et al. (2019). In the first, it was proposed a technique to identify angular displacement in the position of electrodes arranged in armbands, an issue that typically occurs in wearable devices when the electrodes are evenly spaced along the circumference of the arm, and an algorithm based on the IPL method was used. In this way, it is possible that the placement of the electrodes can be corrected before using the device. With the same purpose, He et al. (2019) considered the verification of the similarity, calculated from the Mahalanobis distance, in relation to the activation pattern of the electrode set for a given reference movement, allowing to correct the electrode placement before using the device. Both approaches present themselves as interesting pre-processing methods to correct a problem that is very common in this type of procedure, which is electrode displacement.

Finally, the issue of the time window size used for signal segmentation and subsequent feature extraction is also an open point in the literature. According to the results of (Smith et al. 2011), the ideal window size varies between 150 and 250 ms. Furthermore, according to (Riillo et al. 2014), a maximum time segment of 300 ms is allowed between two consecutive classifier outputs to meet the requirements of a real-time system. Considering that very small windows hamper the extraction of useful information from the sEMG signal, choosing the ideal size consists of a compromise.
between the classifier’s hit rate and meeting the minimum requirements for real-time implementation.

Aiming to clarify this issue, Wahid et al. (2020) tested various combinations of temporal window sizes (50, 100, 150, 200, 250, 300, 400, and 500 ms) and percent overlap (0, 30, 50, 70, and 80%) over an extensive range database (Ninapro Database #2 which has 40 volunteers). For the first time, according to the authors’ knowledge, this issue was addressed by considering such an extensive database. The results presented indicated that both the size and the percentage of overlapping windows influence the classifier’s hit rates. The larger the signal and superposition segment, the greater the accuracy obtained. From this, it is suggested the application of overlap of at least 80% since it would allow the use of windows with a size greater than 300 ms, and that at the same time would meet the requirements of real-time application.

In summary, there are several recent works with promising strategies to address a variety of issues as automatic label assign (very important for database generation), the effect of window size in data segmentation (relevant for real-time applications as a myo-controlled prosthesis), and for general classifier performance improvement. Therefore, they complement each other and appear as options for the future of academic research supported by the promising results presented in Table 2.

2.3. Feature Extraction

The raw sEMG signal has a strong stochastic nature as it consists of the synergy of several muscles and it is not very informative in this form (Anam et al. 2019; Reaz et al. 2006). To extract useful information regarding muscle electrical activity, it is necessary to apply processing techniques that make it quantifiable and handleable. There are several methods for processing the sEMG signal, being the most used those based on the exploration of its statistical metrics. Currently, there are several classical metrics used for this purpose. They can be calculated in the time domain (Root Mean Square value, absolute mean, waveform length, standard deviation, power, etc.), frequency domain (average frequency, median, the wavelength of the frequency response, etc.), and time-frequency domain (Wavelet Transform, Short Time Fourier Transform – STFT, etc.), and are widely used (Arteaga et al. 2020; Batayneh et al. 2020; Cognolato et al. 2020; Wang et al. 2020; Yamanoi et al. 2020).

However, innovative approaches have been proposed as alternatives to classical features, that can be grouped according to the improvement they bring to the classification process: reduction of the impact associated with the muscle contraction intensity variation (Asogbon et al. 2020; Nougarou et al. 2019; Onay & Mert 2020; Tuncer et al. 2020), electrode displacement (Lv et al. 2018), inter-subject variation (Shivam et al. 2019; Tong et al. 2019), temporal variation (Jaber et al. 2020; Jaber et al. 2019), classifier general performance improvement in terms of computational cost minimization (Toledo-Perez et al. 2020) and accuracy increase (She et al. 2019; Sravani et al. 2020). Table 3 summarizes the most recent feature extraction strategies.

According to Table 3, several works explore muscle activation spatial resolution to design new features as an alternative to the classical features extracted on time and frequency domain. Onay and Mert (2020) used the arrangement of eight electrodes uniformly spaced along the limb circumference.
to propose a method of phasor representation from the features extracted in each channel, i.e., each sEMG measurement point has a specific phase angle. Thus, spatial resolution was introduced in the representation of two classical metrics (RMS and waveform length), and the approach effectiveness was verified through an experiment based on the recognition of six movements performed by nine volunteers with transradial amputation under three different strength levels. The classifier used was kNN (k-Nearest Neighbor), and the training was conducted with data from one strength intensity level and the tested with the others. The phasor-based feature extraction method showed an 8.3% improvement in the average accuracy compared to TD-PSD features (Time-Dependent Power Spectrum Descriptors), a strategy presented in (Khushaba 2014) aiming to improve the classifier performance in the presence of data extracted under strength variation.

<table>
<thead>
<tr>
<th>Authors</th>
<th>Domain</th>
<th>Method</th>
<th>#CH</th>
<th>#Mov.</th>
<th>I. Vol.</th>
<th>A. Vol.</th>
<th>Classifier</th>
<th>AvAcc [%]</th>
<th>Issue considered</th>
</tr>
</thead>
<tbody>
<tr>
<td>(Onay &amp; Mert 2020)</td>
<td>Time-spatial</td>
<td>Phasor representation</td>
<td>8</td>
<td>6</td>
<td>-</td>
<td>9</td>
<td>kNN</td>
<td>78.3</td>
<td>Intensity of the muscle contraction.</td>
</tr>
<tr>
<td>(Tuncer et al. 2020)</td>
<td>Time-frequency</td>
<td>Ternary Pattern-DWT (TP-DWT)</td>
<td>12</td>
<td>6</td>
<td>-</td>
<td>9</td>
<td>kNN</td>
<td>99.1</td>
<td>Intensity of the muscle contraction.</td>
</tr>
<tr>
<td>(Jaber et al. 2020)</td>
<td>Spatial</td>
<td>Average Intensity-HOG (AH)</td>
<td>128</td>
<td>8</td>
<td>18</td>
<td>-</td>
<td>SVM</td>
<td>96.4</td>
<td>Temporal variation.</td>
</tr>
<tr>
<td>(Asogbon et al. 2020)</td>
<td>Time</td>
<td>invTDD (invariant time-domain descriptor)</td>
<td>8</td>
<td>7</td>
<td>6</td>
<td>2</td>
<td>LDA</td>
<td>89.9 (I) 86.6 (A)</td>
<td>Intensity of the muscle contraction.</td>
</tr>
<tr>
<td>(Lv et al. 2018)</td>
<td>Time</td>
<td>Autoencoder Neural Network</td>
<td>64</td>
<td>10</td>
<td>9</td>
<td>1</td>
<td>ANN</td>
<td>90.3 (I) 85.4 (A)</td>
<td>Electrode shift.</td>
</tr>
<tr>
<td>(Li et al. 2019)</td>
<td>Spatial</td>
<td>Active Muscle Regions</td>
<td>16</td>
<td>4</td>
<td>9</td>
<td>-</td>
<td>SVM</td>
<td>87.0</td>
<td>General classifier performance improvement.</td>
</tr>
<tr>
<td>(Shivam et al. 2019)</td>
<td>Time-frequency</td>
<td>Stockwell Transform</td>
<td>2</td>
<td>6</td>
<td>5</td>
<td>3</td>
<td>ANN</td>
<td>98.4 (I) 97.7 (A)</td>
<td>General classifier performance improvement.</td>
</tr>
<tr>
<td>(Tong et al. 2019)</td>
<td>Time-spatial</td>
<td>LSTM-CNN</td>
<td>16</td>
<td>5</td>
<td>8</td>
<td>-</td>
<td>ANN</td>
<td>78.3</td>
<td>Inter subject variability.</td>
</tr>
<tr>
<td>(Nougarou et al. 2019)</td>
<td>Spatial</td>
<td>RMS HD-sEMG Maps</td>
<td>64</td>
<td>10</td>
<td>10</td>
<td>-</td>
<td>LDA</td>
<td>94.2</td>
<td>Intensity of the muscle contraction.</td>
</tr>
<tr>
<td>(Pancholi &amp; Joshi 2019)</td>
<td>Time</td>
<td>Time Derivative Moments</td>
<td>12</td>
<td>8</td>
<td>8</td>
<td>-</td>
<td>SVM</td>
<td>96.2</td>
<td>General classifier performance improvement.</td>
</tr>
<tr>
<td>(Toledo-Perez et al. 2020)</td>
<td>Time</td>
<td>Modified Zero-Crossing</td>
<td>2</td>
<td>6</td>
<td>5</td>
<td>-</td>
<td>SVM</td>
<td>94.8</td>
<td>Computational cost.</td>
</tr>
</tbody>
</table>

Table 3: A summary of works associated with feature extraction stage.
In the table: DWT is Discrete Wavelet Transform, HOG is Histogram of Oriented Gradients, and LSTM is Long Short-Term Memory.

Following in the same line, Nougarou et al. (2019) proposed a new way of extracting information of the sEMG signal, taking advantage of the spatial resolution provided by a matrix formed by 64 electrodes (8x8). Through the application of three spatial filters (Monopolar Filter, Bipolar Filter, and Inverse Binomial Filter) on the signal obtained by the sensor array, three new 6x6 dimension maps are obtained. Each map is divided into 9 sub-regions of which 3 features are extracted: center of gravity coordinates (indicating 2 attributes) and the influence percentage from the segment in relation to the entire image. Results showed the superiority of this signal representation in relation
to classical time-domain metrics, including in situations where there was variation in the contraction intensity applied in the movements execution.

Both strategies described above are promising alternatives for the features extraction that are more robust to the muscle strength variation. Thus, it is demonstrated the importance of considering the spatial resolution of the muscle activation pattern in the signal representation for the model training stage.

Aiming to provide a solution to the variability of the sEMG signal acquired from different subjects and allowing development of a multiuser system, Shivam et al. (2019) and Tong et al. (2019) presented two different feature extraction methods. The first extracts information from the synergy present in muscle activation through the application of Non-Negative Matrix Factorization (NMF). In the research by Tong et al. (2019), it was achieved both spatial and temporal information from the sEMG signal through the application of a hybrid algorithm relying on the combination of Convolutive Neural Networks and LSTM. Each of the networks generates a 256 elements vector in the output, containing both spatial representation, obtained with CNN, and temporal information, obtained with the LSTM network.

Both forms of transformation applied to the electromyography signal proved to be effective to train a multiuser classifier. In the strategy based on the synergy of the different channels, a 97% success rate was achieved in five fingers movement recognition, performed by five volunteers through an SVM classifier. The hybrid system CNN-LSTM provided a hit rate of 78.3% in five gestures identification performed by eight subjects. In both experiments, the leave-one-subject-out cross-validation technique was considered, that is, the results consist of n-classifiers hit rate average trained with data from n-1 volunteers and tested with the remaining subject, where n corresponds to the number of participants. Despite the lower accuracy, the algorithm structured in neural networks should be highlighted due to the greater number of volunteers and because the signal processing was conducted in 300 ms windows, allowing real-time applications.

Jaber et al. (2020) developed an innovative feature extraction methodology that is robust to the signal intrinsic temporal variations. First, the signal acquired through an array of 8x16 electrodes, was represented as an image, where each electrode is treated as a pixel, and then metrics were calculated based on the intensity of each pixel and on the Histogram of Oriented Gradients (HOG). Using the spatial information with a SVM classifier, it were achieved hit rates above 90% in an eight movements classification for 9 of the 10 volunteers considered. However, the data used were acquired in sections performed on different days, highlighting the capability of extracting time-invariant information of the proposed algorithm, showing its robustness.

Finally, the issue of electrode displacement was the topic addressed in (Lv et al. 2018) and (Tong et al. 2019), with similar methods where in the latter, neural networks were also used for the feature extraction. An Autoencoder topology was used, which consists of a neural network whose objective is to represent the input pattern in a coded form, and are structured in two parts, an encoder, and a decoder, consisted in three layers of neurons: input, hidden, and output. The encoder, composed of the hidden and input layers and whose neurons are fully connected, determines an encoding for the pattern inserted into the network input. This coded representation is given by the output of each one of the hidden layer nodes, which are connected to the output layer neurons that, in turn, decode the
signal to the original representation, i.e., they provide again the pattern inserted in the network input. Evidently, the number of nodes in the output layer is the same as in the input. Thus, the features considered by the presented strategy are the outputs of hidden layer neurons, that is, the encoder. The training of an LDA-based classifier fed by the features was carried out in order to validate the proposed algorithm. As a result, accuracy above 90% was achieved in the prediction of 10 movements considering displacements of -1 to 1 cm in the electrodes positioning, in relation to the reference. Supported on the above, the effectiveness of the proposed signal representation method was evidenced in presence of sensor positioning deviations.

2.4. Feature Selection

After the feature extraction step, the sEMG signal is represented by a vector formed by the number of channels used in the acquisition times the number of features considered. When considering the use of the low or high-density configuration, for example, the vector size easily goes beyond 100 elements. However, not all attributes provide useful or new information for the classification stage, and the greater the amount of data provided to train the model, the higher the computational cost and, consequently, the time required to learn and classify.

Therefore, the feature selection stage plays a fundamental role in the motion recognition system. Through it, the attributes formed by the channel/feature pairs are filtered, removing all those that are redundant and/or irrelevant to teach the model. Hence, a better performance of the classifier is achieved, both in terms of accuracy and processing time, as it has already been verified in previous works that a high number of features and the presence of redundant attributes can affect system performance (Jair et al. 2020; Tosin et al. 2020b).

Although still little explored by researchers compared to the classification stage, there is a considerable amount of work published in recent years focusing on the feature selection step. They can be grouped into the search for the best channels set (Batayneh et al. 2020; Hua et al. 2019; Krasoulis et al. 2020; Sun et al. 2020; Wang et al. 2019; Yu et al. 2019), features set (Luo et al. 2020; Phukan et al. 2019; Ye et al. 2019; Zhang et al. 2019; Zhou et al. 2019) and individual pairs formed by channels and features (Castiblanco et al. 2020; Jair et al. 2020; Jitaree & Phukpattaranont 2019; Tosin et al. 2020b; Tosin et al. 2020a; Tosin et al. 2017; Wu et al. 2019). Table 4 summarizes the most recent strategies for feature selection stage and the average accuracy presented refer to the best scenario reported by the authors in terms of feature selection method and classifier.

The feature selection methods can be divided according to the strategy adopted in wrapper, filter-based, and embedded. Filter-based methods perform the selection without the influence of a classifier. The feature relevance is evaluated through some statistical metrics, which might include distance index, dependence, consistency, information, and correlation (Tang et al. 2015). The wrapper strategy uses classification accuracy to assign importance to the features and is computationally more expensive than the filter-based one, as it performs the selection through multiple training and classifications, using different features sets. Filling the gap between the wrapper and filter-based algorithms, there are the embedded methods that combine the advantages of the wrapper and filter-based, as they also make use of a classifier model and are computationally
less costly than the wrapper. Such algorithms evaluate features during the classifier training stage (Tang et al. 2015) and usually, a rank is assigned to each feature, proportional to its relevance for the training process.

As shown in Table 4, the wrapper method was the most preferred, followed by filter and embedded, in this order. Typically, wrapper algorithms are associated with higher classifier accuracies (Aggarwal & Reddy 2013), which can explain the significant recent preference by researchers.

According to (Farina et al. 2014), the ability to achieve high performance in decoding movements with a minimum number of electrodes is the main challenge for the development of machine learning techniques based on myoelectric control. Given this, several researchers have dedicated their efforts to find a method for selecting the most relevant channels for the classification task.

In (Krasoulis et al. 2020; Wang et al. 2019; Yu et al. 2019) the wrapper strategy was used to select the ideal number of electrodes. Krasoulis et al. (2020) considered the accuracy of a LDA-based classifier to determine the most relevant channels in identifying six movements represented by seven time domain features. As a result, it was concluded that only two sEMG channels (of the 16 available) and inertial measurements (acceleration, angular velocity, and orientation) were sufficient to obtain 86.5% and 84.7% hit rate (for intact and amputee volunteers respectively) with RDA classifier.

To optimize the search for the best group of electrodes, Wang et al. (2019) proposed a genetic algorithm based on chromosome chains formed by binary words where each bit represents the use or not of a channel, 1 for yes and 0 for no. The fitness, in turn, was defined as the hit rate of an LDA-based classifier. Here, it was concluded that with 12 channels it is possible to reach 97% of the maximum accuracy (obtained using all 16 available signals) in the recognition of 13 movements with eight temporal features.

As an alternative to the wrapper methods, the works (Batayneh et al. 2020; Hua et al. 2019; Sun et al. 2020) invested in filter-based channel selection strategies, with Hua et al. (2019) and Sun et al. (2020) standing out. The first considered the signal-to-noise ratio analysis of each channel, eliminating those that are less informative, i.e., with smaller SNR. In the second, a similar analysis was performed, but considering the variance of the signal to estimate its activation level and, consequently, its relevance to the classification process. In both studies, 13 out of 16 available channels were selected, reaching accuracies of 95.5% (Sun et al. 2020) and 94.3% (Hua et al. 2019) in the recognition of 5 movements with SVM and ANN, respectively.

Some researchers adhered to the use of a smaller number of electrodes, focusing on finding the best set of features for the movement identification process (Luo et al. 2020; Phukan et al. 2019; Yang et al. 2020; Zhang et al. 2019). The work of Yang et al. (2020) and Phukan et al. (2019) stand out here, who achieved promising results in the 6 and 10 movements identification, respectively, with only 2 electrodes.
<table>
<thead>
<tr>
<th>Authors</th>
<th>Strategy</th>
<th>Method</th>
<th>#CH</th>
<th>#Feat.</th>
<th>N° Sel. F/C pairs</th>
<th>#Mov.</th>
<th>Vol.</th>
<th>Classifier</th>
<th>AvAcc [%]</th>
<th>Issue considered</th>
</tr>
</thead>
<tbody>
<tr>
<td>(Jair et al. 2020)</td>
<td>Wrapper/Filter</td>
<td>LMNN</td>
<td>8</td>
<td>21</td>
<td>TD, 9 FD</td>
<td>40</td>
<td>6</td>
<td>13 (I)</td>
<td>SVM</td>
<td>94.0</td>
</tr>
<tr>
<td>(Castiblanco et al. 2020)</td>
<td>Filter</td>
<td>Separability Index (BD)</td>
<td>8</td>
<td>10</td>
<td>TD, 7 FD</td>
<td>40 (I)</td>
<td>47 (S)</td>
<td>12</td>
<td>SVM</td>
<td>82.8 (I)</td>
</tr>
<tr>
<td>(Krasoulis et al. 2020)</td>
<td>Wrapper</td>
<td>Forward with LDA</td>
<td>16</td>
<td>12/13</td>
<td>(I)</td>
<td>7 TD</td>
<td>14</td>
<td>12 (I)</td>
<td>RDA</td>
<td>86.5 (I)</td>
</tr>
<tr>
<td>(Yang et al. 2020)</td>
<td>Filter</td>
<td>Separability Index (ED)</td>
<td>2</td>
<td>3</td>
<td>TD, 3 FD</td>
<td>4</td>
<td>6</td>
<td>9 (I)</td>
<td>ANN</td>
<td>95.5</td>
</tr>
<tr>
<td>(Wu et al. 2020)</td>
<td>Wrapper</td>
<td>Grid search with ANN</td>
<td>4</td>
<td>4</td>
<td>TD, 12</td>
<td>6</td>
<td>6</td>
<td>6 (I)</td>
<td>ANN</td>
<td>94.8</td>
</tr>
<tr>
<td>(Sun et al. 2020)</td>
<td>Filter</td>
<td>Variance Analysis</td>
<td>16</td>
<td>3</td>
<td>TD, 39</td>
<td>5</td>
<td>9</td>
<td>5 (I)</td>
<td>SVM</td>
<td>95.5</td>
</tr>
<tr>
<td>(Luo et al. 2020)</td>
<td>Wrapper</td>
<td>Grid search with LDA</td>
<td>8</td>
<td>3</td>
<td>TD, 8</td>
<td>4</td>
<td>2</td>
<td>2 (I)</td>
<td>LDA</td>
<td>79.3</td>
</tr>
<tr>
<td>(Yu et al. 2019)</td>
<td>Wrapper</td>
<td>Backward with ANN</td>
<td>8</td>
<td>5</td>
<td>TD, 35</td>
<td>11</td>
<td>5</td>
<td>5 (I)</td>
<td>ANN</td>
<td>90.5</td>
</tr>
<tr>
<td>(Wu et al. 2019)</td>
<td>Embedded</td>
<td>LDA</td>
<td>4</td>
<td>63*</td>
<td>TD, FD, TFD</td>
<td>5</td>
<td>6</td>
<td>15 (I)</td>
<td>DA</td>
<td>98.2</td>
</tr>
<tr>
<td>(Zhou et al. 2019)</td>
<td>Wrapper</td>
<td>Grid search with RF</td>
<td>12</td>
<td>9</td>
<td>TD, 36</td>
<td>12</td>
<td>10</td>
<td>10 (I)</td>
<td>RF</td>
<td>84.1</td>
</tr>
<tr>
<td>(Zhang et al. 2019)</td>
<td>Wrapper</td>
<td>Grid search with DBN</td>
<td>8</td>
<td>4</td>
<td>TD, 16</td>
<td>4</td>
<td>4</td>
<td>4 (I)</td>
<td>DBN</td>
<td>96.5</td>
</tr>
<tr>
<td>(Hua et al. 2019)</td>
<td>Filter</td>
<td>SNR Analysis</td>
<td>16</td>
<td>1</td>
<td>TD, 13</td>
<td>5</td>
<td>10</td>
<td>10 (I)</td>
<td>ANN</td>
<td>94.3</td>
</tr>
<tr>
<td>(Wang et al. 2019)</td>
<td>Wrapper</td>
<td>GA-based with LDA</td>
<td>16</td>
<td>8</td>
<td>TD, 96</td>
<td>13</td>
<td>6</td>
<td>6 (I)</td>
<td>LDA</td>
<td>76.2</td>
</tr>
<tr>
<td>(Ding et al. 2019)</td>
<td>Embedded</td>
<td>Generalized-MKL</td>
<td>4</td>
<td>12</td>
<td>TD, 2 FD</td>
<td>44</td>
<td>10</td>
<td>15 (I)</td>
<td>SVM</td>
<td>93.0</td>
</tr>
<tr>
<td>(Phukan et al. 2019)</td>
<td>Filter</td>
<td>Mutual Information</td>
<td>2</td>
<td>20</td>
<td>TD, 11 FD</td>
<td>10</td>
<td>10</td>
<td>4 (I)</td>
<td>SVM</td>
<td>96.5</td>
</tr>
<tr>
<td>(Yoo et al. 2019)</td>
<td>Wrapper</td>
<td>Grid search with DFDL</td>
<td>6</td>
<td>1</td>
<td>TFD, 2</td>
<td>4</td>
<td>2</td>
<td>22 (I)</td>
<td>DFDL</td>
<td>&gt;95.0</td>
</tr>
<tr>
<td>(Tosin et al. 2020b)</td>
<td>Wrapper/Embedded</td>
<td>SVM-RFE</td>
<td>12</td>
<td>11</td>
<td>TD, 2 FD</td>
<td>41</td>
<td>17</td>
<td>40 (I)</td>
<td>RELM</td>
<td>84.9 (I)</td>
</tr>
</tbody>
</table>

Table 4: A summary of works associated with feature selection stage.

In the table: the term “N° Sel. F/C pairs” refers to the number of selected feature/channel pairs, LMNN is Large Margin Nearest Neighbor, BD is Bhattacharyya Distance, RDA is Regularized Discriminant Analysis, ED is Euclidean Distance, DBN is Deep Belief Network, GA is Genetic Algorithm, MKL is Multiple Kernel Learning, DFDL is Discriminative Feature-Oriented Dictionary Learning, SVM-RFE is Support Vector Machine Recursive Feature Elimination, DA is Discriminant Analysis, RELM is Regularized Extreme Learning Machine, TD is Time Domain, FD is Frequency Domain, TFD is Time-Frequency Domain, (I) is intact, (S) is stroke, (A) is amputee.

*The authors did not specify the exact number of features extracted in each domain.

In (Yang et al. 2020) an algorithm for a 2 electrode system was proposed to assess the importance of each feature considered, among 6 different metrics calculated in the time and frequency domains, through the Euclidean distance, represented by a two-dimensional vector in a plane, where each axis represent an electrode. Thus, a numerical index was assigned to each of them according to the average distance between all samples associated with a specific gesture (direct relationship) and the average distance between the geometric centers of each movement (inverse relationship). Hence, the lower the value of this index, the more discriminating the feature. With the use of such methodology, 2 time-domain metrics were selected, reaching 95.5% of assertiveness with ANN.
Similarly, an approach based on the evaluation of mutual information for choosing the best group of features to represent the signal was presented in the research by (Phukan et al. 2019). Thus, 5 out of 31 different features were selected from the time and frequency domains, reaching a 96.5% hit rate in a SVM model.

Finally, there is still a research niche within the feature selection stage that considers both the channels and features simultaneously, analyzing the relevance of each pair formed by both individually (Castiblanco et al. 2020; Jitaree & Phukpattaranont 2019; Tosin et al. 2020b; Tosin et al. 2020a; Wu et al. 2020). Here, the researches by Tosin et al. (2020) and by Castiblanco et al. (2020) stand out. Their algorithms include the formation of a ranking of feature/channel pairs according to their importance for the classification stage, and then the best group is determined based on this ranking.

In (Castiblanco et al. 2020), 4 different methods were applied to sort the feature/channel pairs: t-test calculated between the distributions associated with each class (in pairs) for the same attribute (the lower the probability of two sets of samples from different classes being part of the same distribution, the higher the value assigned to the attribute relevance), Separability index determined by the distance function through Mahalanobis and Bhattacharyya methods, and the Davies-Boulding Index, commonly used to characterize the separation of clusters prioritizing greater concentration of data from the same group and greater spacing between geometric centers of different clusters. As a result, it was found (in most of the tests carried out) that less than 50 elements are needed for the sEMG signal representation out of the 136 (17 features x 8 channels) to obtain stability in the classifier accuracy. Also noteworthy is the promising assertiveness index in the classification of 12 movements with kNN (80.9% on average among the 4 selection methods and 4 volunteers) based on signals acquired from patients with motor dysfunction due to stroke.

The Support Vector Machine Recursive Feature Elimination (SVM-RFE), Monte Carlo Feature Selection, and Singular Value Decomposition (SVD) Entropy methods were applied in (Tosin et al. 2020b) to determine a ranking of the 156 attributes considered (12 channels x 13 features). Here, the ordering obtained was used as a guide for the execution of a wrapper-type algorithm with a Regularized Extreme Learning Machine (RELM) classifier. After implementing the proposed algorithm to identify 17 movements of the hand-arm segment, it was concluded that the inclusion of the feature selection step increased the hit rate from 80.6% to 84.3% compared to training considering all feature/channel pairs. The results correspond to the average of the 40 volunteers and the 3 strategies for creating the ranking). In addition to reducing the average number of attributes to 41, the authors also demonstrate that choosing the best set of values to represent the signal is more effective when considering, simultaneously, the channel and feature effect than when one is fixed.

However, considering that the approaches last described in this section generally comprise a broader field of search for the most relevant set of attributes, these tend to be more effective in terms of improving the assertiveness of the predictor model. Nonetheless, methods based on individual analysis of features or channels can be an interesting option when prioritizing the minimization of computational cost against the classifier accuracy.

2.5. Classification / Regression
The classification stage is the core of the human-machine interface system. All the acquired learning is concentrated in this step. Therefore, it is responsible for decoding the sEMG signal patterns (represented by the set of attributes determined in the feature selection stage or just by the pre-processed signal, depending on the learning strategy) into motion intentions. Therefore, naturally, this block concentrates the greatest amount of works promoting innovative and promising methodologies.

There are currently two main approaches to the HMI control task: the classic pattern recognition method and proportional control. The first is based on the recognition of a limited number of discrete movements, being associated with the vast majority of the strategies offered (Ameri et al. 2020; Cene et al. 2019b; Chen et al. 2020b; Zanghieri et al. 2020). However, such implementations do not provide smooth hand control (or at least not as much as the human hand) (Wang et al. 2020) and are not adequate to explore all trajectory possibilities offered by robotic prostheses, since only one gesture class is activated at a time (Anam et al. 2019). Therefore, the continuous movement estimation by sEMG has become a popular research field (Wang et al. 2020), emerging as an alternative to the previous one. In proportional control, individual predictions are required in each of the joints (Anam et al. 2019), which may be associated with the angular position (Yang et al. 2019b), force (Martinez et al. 2020), torque (Yu et al. 2020b), among other measurements. Thus, the aim is to make the myoelectric control more intuitive (Pan et al. 2019), facilitating user adaptation.

While the control strategy based on the identification of specific gestures depends on the training of a classifier, in proportional control the task consists of determining a regressor model, since, unlike the first one, here the output possibilities are unlimited. In this context, many works have been conducted in recent years with new and auspicious strategies for both functions (classification and regression). They can be divided into: propositions for proportional control (Anam et al. 2019; Belyea et al. 2019; Martinez et al. 2020; Wang et al. 2020; Yang et al. 2019b; Yu et al. 2020b), strategies designed around classical classification methodologies (Belyea et al. 2019; Cene and Balbinot 2019; BelRe & Rus 2020; Donati et al. 2019; Gong et al. 2019; Guo et al. 2019; Mantilla-Brito et al. 2020; Moin et al. 2019; Shin et al. 2020; Vasanthi & Jayasree 2020), deep learning methods (Ameri et al. 2020; Asif et al. 2020; Chen et al. 2020b; Côté-Allard et al. 2020; Huang & Chen 2019; Kim et al. 2020; Liu et al. 2019; Mukhopadhyay & Samui 2020; Olsson et al. 2019a; Olsson et al. 2020; Olsson et al. 2019b; Rahimian et al. 2019; Shao et al. 2020; Yamanoi et al. 2020; Yang et al. 2019a; Zanghieri et al. 2020) and research associated to the refinement of the classifier output also called post-processing (Ahmed et al. 2019; Cene et al. 2019b; Jafarzadeh et al. 2019; Yu et al. 2020a).

Table 5 summarizes the works which present new and promising strategies for proportional control. The results described corresponding to the best scenario reported by authors. Here, the metrics considered to evaluate the performance of the works are Normalized Root Mean Square Error, Root Mean Square Error, Relative Error, Pearson Correlation Index, and Determination Coefficient. The determination coefficient represents how explanatory a model is to data variance. It varies between 0 and 1, and the closer it is to 1, the more the regressor model fits the sample. According to Table 5, there is an evident preference by deep learning algorithms for regression, being considered in most of the works covered by this review through the methods CNN, DNN, and LSTM.
In general, aiming at the development of regression algorithms for the prediction of continuous hand movement, stand out the studies by Pan et al. (2019), Wang et al. (2020), and Yang et al. (2020). Wang et al. (2020) proposed an LSTM-based model to estimate the angular position in 20 hand fingers joints, during the execution of six grasping motions through sEMG acquired by 12 channels and represented by one TD feature. The data used in the training of the regressor were the inertial measurements performed in five volunteers using a CyberGlove. These data were extracted from NinaPro database 2 and were used as ground truth in the generation of the model. A Root Mean Square Error value of 5.89° was reported for all 20 predictors, which can be considered promising given the high number of estimated joins.

Angle prediction was also considered in (Yang et al. 2019b). The target was the characterization of three DOF of the wrist (flexion/extension, supination/pronation, radial/ulnar deviation). Therefore, a Convolutive Neural Network topology fed with raw 8-channel sEMG signal (no feature was extracted) was proposed for the training of a single regressor for the three DOF. Through experiments carried out with 8 volunteers, it was observed high adaptability of the system to variations caused by electrode displacement and differences between subjects, reaching coefficients of determination (R²) of 0.74 and 0.56, respectively, in each of the tests (with sensor deviations and multiuser).

<table>
<thead>
<tr>
<th>Authors</th>
<th>Control variable</th>
<th>N° DOF/joints</th>
<th>Joint or limb</th>
<th>#CH</th>
<th>#Vol.</th>
<th>Regression method</th>
<th>DC (R²)</th>
<th>Corr. (R)</th>
<th>Error – NRMSE</th>
<th>Issue considered</th>
</tr>
</thead>
<tbody>
<tr>
<td>(Wang et al. 2020)</td>
<td>Angular position</td>
<td>20/20</td>
<td>Fingers joints</td>
<td>12</td>
<td>5 (I)</td>
<td>LSTM</td>
<td>-</td>
<td>0.84</td>
<td>5.9° (RMSE)</td>
<td>Continuous grasp motion prediction.</td>
</tr>
<tr>
<td>(Yu et al. 2020b)</td>
<td>Torque</td>
<td>3/1</td>
<td>Wrist</td>
<td>16</td>
<td>8 (I)</td>
<td>SAE-DNN</td>
<td>0.74 (I)</td>
<td>0.68 (A)</td>
<td>-</td>
<td>Continuous wrist torque prediction.</td>
</tr>
<tr>
<td>(Martinez et al. 2020)</td>
<td>Force</td>
<td>4 (Force level)</td>
<td>Hand</td>
<td>168</td>
<td>12 (I)</td>
<td>ENR-RLR</td>
<td>0.82</td>
<td>-</td>
<td>-</td>
<td>Continuous grasp force prediction.</td>
</tr>
<tr>
<td>(Anam et al. 2019)</td>
<td>Angular position</td>
<td>8/8</td>
<td>Fingers joints</td>
<td>16</td>
<td>4 (I)</td>
<td>DNN</td>
<td>0.96</td>
<td>-</td>
<td>-</td>
<td>Continuous grasp motion prediction.</td>
</tr>
<tr>
<td>(Yang et al. 2019b)</td>
<td>Angular position</td>
<td>3/1</td>
<td>Wrist</td>
<td>8</td>
<td>8 (I)</td>
<td>CNN</td>
<td>0.74</td>
<td>-</td>
<td>-</td>
<td>Electrode displacement; Inter subject variability.</td>
</tr>
<tr>
<td>(Pan et al. 2019)</td>
<td>Angular position</td>
<td>2/2</td>
<td>Wrist Metac.</td>
<td>4*</td>
<td>6 (I)</td>
<td>MM</td>
<td>0.88 (I)</td>
<td>0.80 (A)</td>
<td>0.13 (I)</td>
<td>Limb position.</td>
</tr>
<tr>
<td>(Ameri et al. 2019)</td>
<td>Angular velocity</td>
<td>2/1</td>
<td>Wrist</td>
<td>8</td>
<td>10 (I)</td>
<td>CNN</td>
<td>-</td>
<td>-</td>
<td>5.8% (RE)</td>
<td>Continuous wrist motion prediction.</td>
</tr>
<tr>
<td>(Lei 2019)</td>
<td>Angular position</td>
<td>1/1</td>
<td>Elbow</td>
<td>1</td>
<td>4 (I)</td>
<td>ANN</td>
<td>-</td>
<td>-</td>
<td>8.4% (RE)</td>
<td>Real-time application; Embedded system.</td>
</tr>
</tbody>
</table>

Table 5: A summary of works associated with proportional control strategies.

In the table: the term “DC” is Determination Coefficient, “Corr.” refers to Pearson Correlation index, “Metac.” is Metacarpophalangeal, SAE-DNN is Stacked Autoencoder-based Deep Neural Network, ENR-RLR is Elastic Nets Ridge-Regularized Linear Regression, DNN is Deep Neural Network, MM is Musculoskeletal Model, NRMSE is Normalized Root Mean Square Error, RMSE is Root Mean Square Error, and RE is Relative Error. *For the amputee subject, the EMG acquisition was made percutaneously. For the intact ones, it was used surface electrodes.

However, Pan et al. (2019) addressed the issue of variability caused by limb positioning to propose a proportional control algorithm in two joints (wrist and metacarpophalangeal). It was developed through the creation of a Musculoskeletal Model (MM), obtained from the measurement
of the electrical activity of 4 muscles of the forearm (extensor digitorum, flexor digitorum, extensor carpi radialis longus, and flexor carpi radialis) performed superficially in 6 intact subjects and percutaneously in a volunteer with transradial amputation. The results demonstrated the superiority of the proposed strategy to classical regression methods (ANN and linear regression) both in terms of correlation (identification of movement variation tendency) and in the normalized RMS error (indication of strength) calculated from the ground truth. Likewise, through an experiment performed with nine different positions of the arm, MM demonstrated better performance to this confounding factor. According to the authors, this result can be explained due to the Musculoskeletal Model consider physiological properties of the limb, such as muscle limitations and detection of passive strength, which other methods neglect. Consequently, this presents itself as a good alternative for increasing robustness in HMI systems.

Despite the emerging works related to proportional control, the research around the classification of discrete movements is still a tendency. Table 6 summarizes the works that present new strategies for movement recognition using classical machine learning approaches. Here, the term “classical machine learning approaches” is related to those algorithms that are not associated with deep learning methods. These will be described posteriorly.

<table>
<thead>
<tr>
<th>Authors</th>
<th>Strategy</th>
<th>#CH</th>
<th>#Feat.</th>
<th>Emb.</th>
<th>#Mov.</th>
<th>Vol.</th>
<th>Classifier</th>
<th>AvAcc [%]</th>
<th>Issue considered</th>
</tr>
</thead>
<tbody>
<tr>
<td>(DelPreto &amp; Rus 2020)</td>
<td>Online clustering</td>
<td>10</td>
<td>No</td>
<td>No</td>
<td>8</td>
<td>6</td>
<td>GMM-ANN</td>
<td>97.6</td>
<td>Online training.</td>
</tr>
<tr>
<td>(Powar &amp; Chemmangat 2020)</td>
<td>Similarity measurement</td>
<td>6</td>
<td>1 TD</td>
<td>No</td>
<td>6</td>
<td>10 (I)</td>
<td>DTW</td>
<td>93.3</td>
<td>Wrist orientation variability.</td>
</tr>
<tr>
<td>(Mantilla-Brito et al. 2020)</td>
<td>Window-length reduction</td>
<td>8</td>
<td>4 TD</td>
<td>Yes</td>
<td>3</td>
<td>4 (I)</td>
<td>NB</td>
<td>88.0</td>
<td>Real-time application; Computational cost reduction.</td>
</tr>
<tr>
<td>(Marcheix et al. 2019)</td>
<td>Similarity measurement</td>
<td>8</td>
<td>No</td>
<td>No</td>
<td>5</td>
<td>5 (I)</td>
<td>Matching Score</td>
<td>98.8</td>
<td>Training time reduction.</td>
</tr>
<tr>
<td>(Gong et al. 2019)</td>
<td>Hyperparameter</td>
<td>16</td>
<td>4 TD</td>
<td>No</td>
<td>6</td>
<td>10 (I)</td>
<td>GA-SVM</td>
<td>94.4</td>
<td>General classifier performance improvement.</td>
</tr>
<tr>
<td>(Moin et al. 2019)</td>
<td>Similarity measurement</td>
<td>64</td>
<td>1 TD</td>
<td>No</td>
<td>9</td>
<td>5 (I)</td>
<td>Hyperdimensional</td>
<td>88.2</td>
<td>Intensity of the muscle contraction.</td>
</tr>
<tr>
<td>(Donati et al. 2019)</td>
<td>Spike train signal</td>
<td>4</td>
<td>No</td>
<td>Yes</td>
<td>3</td>
<td>10 (I)</td>
<td>SNN</td>
<td>74.0</td>
<td>Low-power consumption.</td>
</tr>
<tr>
<td>(Gao et al. 2019)</td>
<td>Wavelet-based</td>
<td>4</td>
<td>1 TD</td>
<td>No</td>
<td>6</td>
<td>10 (I)</td>
<td>WNN</td>
<td>93.7</td>
<td>General classifier performance improvement.</td>
</tr>
<tr>
<td>(Cene &amp; Balbinot 2019)</td>
<td>Non-iterative</td>
<td>12</td>
<td>1 TD</td>
<td>Yes</td>
<td>17</td>
<td>10 (I)</td>
<td>ELM</td>
<td>77.2 (A)</td>
<td>Real-time application; Computational cost reduction.</td>
</tr>
<tr>
<td>(Hameed et al. 2018)</td>
<td>Amp. independent</td>
<td>1</td>
<td>2 TD</td>
<td>No</td>
<td>2</td>
<td>1 (I)</td>
<td>FLA and MSE</td>
<td>95.0</td>
<td>Weak sEMG recording (arm impairment volunteers)</td>
</tr>
</tbody>
</table>

Table 6 A summary of works associated with classical classification approaches. In the table: the term “Emb.” is embedded, “Amp.” is amplitude, GMM is Gaussian Mixture Model, DTW is Dynamic Time Warping, GA-SVM is Genetic Algorithm Support Vector Machine, SNN is Spiking Neural Network, WNN is Wavelet Neural Network, FLA is First Lag Autocorrelation, and MSE is Modified Sample Entropy.

Among approaches developed around classical classification methodologies, stand out the researches (Cene & Balbinot 2019; Donati et al. 2019; Mantilla-Brito et al. 2020; Moin et al. 2019; Powar & Chemmangat 2020). The issue of the variability of muscle activity intensity was addressed.
in (Moin et al. 2019), where an algorithm based on Hyperdimensional Computation was developed in which the signal, acquired by 64 channels, is represented by hypervectors with 10000 dimensions. Thus, the class is chosen according to the similarity between the tested sample and a hypervector that represents each of the movements. An assertiveness index greater than 88.2% in identifying 9 movements performed at 3 different strength levels was achieved. The relative simplicity and computational efficiency and its stability against variations in the intensity of muscle contraction configure the main contribution of the method.

Powar and Chemmangat (2020) proposed a classification methodology based on Dynamic Time Warping (DTW) to increase the performance of recognizing six hand movements when performed with the wrist under different orientations. The algorithm of the DTW consists of obtaining the optimal nonlinear alignment of two-time series and, based on it, determining the distance between them. Thus, for each class, there is a set of training data (signal segment used as training templates associated with each movement) that is applied in the computation of the distance to a given test sample. Therefore, the winning motion for a test template corresponds to the one associated with the training template with the shortest distance. A 60% accuracy was achieved in an experiment where signals measured under a pulse orientation and tested under two other configurations were considered, and 93.3% when part of the data from all three positions of the joint was included in the training. Here, the low computational cost of the method also stands out.

However, the search for an algorithm that is computationally efficient is justified not only by the reduction in processing time, which impacts the system response delay, but also by the reduction in the amount of memory and energy consumption required from the available hardware. Considering that, in a real application, the controller must be embedded in a portable device, the works of Mantilla-Brito et al. (2020) and Donati et al. (2019) present classification strategies implemented in an STM32 microcontroller (STMicroelectronics) and FPGA (Field Programmable Gate Array), respectively. In the first, the NB method was adopted for the recognition of three gestures (including opening and closing the hand) using four different time-domain features extracted from eight channels. In an online trial, an 88.0% hit rate and a processing time of only 39.9 ms were achieved.

In (Donati et al. 2019) an energy-efficient classifier strategy based on the neuromorphic implementation (circuits that mimic the functioning of the nervous system) of a Spiking Neural Network (SNN) was proposed. The algorithm consists of decoding the sEMG signal in a train of impulses for submission to the SNN, which, in turn, stimulates the synaptic connections of the brain (excitation or inhibition of the neuron). Therefore, neurons will only transmit information when activated, unlike classical Artificial Neural Networks. The aforementioned decoding was performed in an FPGA while the neural network was implemented in a Dynamic Neuromorphic Asynchronous Processor (DYNAP). It was achieved a 74% accuracy in the identification of three movements consuming 0.05 mW on average.

Still treating the type of training required by the classifier, Cene and Balbinot (2019) compared the performance of iterative (SVM and Regularized Logistic Regression – RLR) and non-iterative methods (ELM, Random Vector Functional-Link Networks – RVFL, and its regularized versions RELM and R-RVFL) in motion recognition. The non-iterative methods showed superiority in general, both in classification and testing time. Thus, it demonstrated that non-iterative algorithms
are good options among the classical classification approaches for obtaining systems with high assertiveness indices and low computational cost.

However, despite the new and promising proposals for classifiers following the classical learning methodology, several researchers have focused their efforts on deep learning, as shown in Table 7. The ability to extract deeper features from the EMG signal allows the exploration of information that is practically unreachable by classical classification methods. This is reflected in the superiority of those to these in terms of success rates, verified in some researches (Du et al. 2017; Jia et al. 2020). However, the main disadvantage of this approach lies in the high computational cost commonly required in training such models, which can take a few hours (Olsson et al. 2019b) since the number of weights to be learned can easily exceed 1 million (Chen et al. 2020b). In contrast, the deep representations of the EMG signal can dispense the feature extraction step (as observed in most works listed in Table 7), which generates a significant reduction in processing time, especially when considering those extracted in the frequency or time-frequency domain. Furthermore, it also reduces the complexity of the system since the feature selection significantly influences the performance of the predictor model (Phinyomark et al. 2013; Tosin et al. 2020b).

Deep learning networks require a considerable amount of time for training, which can make its application in a real system unfeasible when recalibrations are necessary for each new use (both in the case of different days and for new volunteers). Therefore, Ameri et al. (2020) and Kim et al. (2020) proposed strategies based on transfer learning to minimize system recalibration disorders. Both considered Convolutional Neural Networks as a classifier.

In (Ameri et al. 2020), the aim was to minimize the effect of electrode displacement. To do so, they proposed a methodology for recalibrating the system that requires little additional training data. The algorithm uses the parameters of the already trained classifier as an initial approximation for retraining the model with a small number of new data. The efficiency of the method was verified in an experiment where the system was pre-trained with signals from five electrodes. Then, 25% (one of the 4 repetitions) of the data acquired through the five channels displaced to the original position were used to refine the model already obtained (simulating a recalibration). A 93.4% accuracy was achieved in identifying nine movements, higher than that achieved when considering only 25% of the data with the electrodes displaced in the training of a new classifier through CNN and SVM.

Kim et al. (2020) presented a transfer learning strategy to address the inter-volunteer variability issue. Firstly, a CNN model is trained with data from 10 subjects. Then, the sEMG signal of one repetition of each movement of the target volunteer is used to refine the classifier parameters. The results showed the superiority of this method to traditional training (with the same number of samples, i.e., only one repetition of each gesture and without transfer learning) and other forms of transfer learning.

The confounding factor of inter-subject variability was also the theme of the work by Côté-Allard et al. (2020), where a new method based on deep learning called Adaptive Domain Adversarial Neural Network (ADANN) was presented. Roughly speaking, it consists of forcing the network to learn domain-independent (domain here refers to volunteer) features by inserting data associated with a single subject at each training period. The response of the proposed methodology in multiuser
experiments was promising, indicating an increase of 19.4% in the average hit rate in the classification of 11 movements compared to a typical CNN.

According to Table 7, several researchers have adopted strategies based on time-series approaches, taking advantage of the properties of classifiers such as LSTM (Huang & Chen 2019; Zea & Benalcázar 2020) and TCN (Tsinganos et al. 2019b; Zanghieri et al. 2020) or features as cTDD (Mukhopadhyay & Samui 2020) to include temporal resolution of the sEMG signal in the learning of the model. Within this context, it is worth citing Zanghieri et al. (2020), that developed an innovative classification algorithm, robust to the variation of the sEMG signal over time, called TEMPONet (Temporal Embedded Muscular Processing Online Network). This method is structured in a topology based on Temporal Convolutive Networks, which consist of convolutional neural networks of only one dimension where the convolution process is causal (the filter response depends only on the present and past values of the signal) and it maintains the output of the same length as the sequence inserted in the input (Yan et al. 2020). An accuracy of 93.7% was achieved with data from different sections in the recognition of eight gestures. It is noteworthy that TEMPONet's training and testing were performed with data from different days.

The temporal variation issue was also considered in the CNN topology proposed in (Chen et al. 2020b). In the presented Convolutive Neural Network, the input is represented in three dimensions (two of them associated with the location of 128 electrodes, organized in a matrix form (8x16), and the third indicating the time). As a result, both spatial and temporal resolution was achieved in the insertion of the signal into the network, including the time factor in learning the model's parameters. Three-dimensional CNN proved superior to a typical two-dimensional implementation in identifying eight movements, hitting 98.6% of the samples tested against 96.8% of its opponent.

Finally, there is still a niche of researchers who invest in post-processing algorithms to improve classifier accuracy. The purpose is to refine the classifier output to increase the robustness of the system. In this context, Cene et al. (2019b) proposed a filter called Exponential Smoothing Filter at the end of the classification stage of an ELM. Results showed efficiency in reducing ripple in recognizing a specific movement, making the system control more stable.

Also aiming to reduce the incidence of false positives, Yu et al. (2020) presented a strategy for refining the classifier output based on the similarity between two consecutive windows of the signal (here represented by a 6x8 matrix, respecting the electrode placement). The goal is to detect changes in the pattern of muscle activation in subsequent windows by using the Pattern Distance Index (PDI) metric. Hence, a success rate of 86.6% was achieved in the recognition of 12 gestures with LDA, which means an increase of 8.2% to the use of the same classifier without the post-processing step, demonstrating the effectiveness of the proposed algorithm.
Table 7: A summary of works associated with deep learning approaches.

<table>
<thead>
<tr>
<th>Authors</th>
<th>Strategy</th>
<th>#CH</th>
<th>FE</th>
<th>#Mov.</th>
<th>Vol.</th>
<th>Classifier</th>
<th>AvAcc [%]</th>
<th>Issue considered</th>
</tr>
</thead>
<tbody>
<tr>
<td>(Olsson et al. 2020)</td>
<td>EA-based topology optimization</td>
<td>128</td>
<td>No</td>
<td>N/A</td>
<td>8</td>
<td>CNN</td>
<td>99.3</td>
<td>Hardware requirement; Embedded system.</td>
</tr>
<tr>
<td>(Zanghieri et al. 2020)</td>
<td>Time series approach</td>
<td>8</td>
<td>No</td>
<td>N/A</td>
<td>3</td>
<td>TEMPONet</td>
<td>93.7</td>
<td>Temporal variation.</td>
</tr>
<tr>
<td>(Côté-Allard et al. 2020)</td>
<td>Multi-Domain Learning approach</td>
<td>10</td>
<td>No</td>
<td>N/A</td>
<td>11</td>
<td>ADANN</td>
<td>84.4</td>
<td>Inter subject variability.</td>
</tr>
<tr>
<td>(Asif et al. 2020)</td>
<td>Hyperparameter evaluation</td>
<td>6</td>
<td>No</td>
<td>N/A</td>
<td>10</td>
<td>CNN</td>
<td>92.0</td>
<td>General classifier performance improvement.</td>
</tr>
<tr>
<td>(Shao et al. 2020)</td>
<td>Wavelet-based activation function</td>
<td>6</td>
<td>No*</td>
<td>N/A</td>
<td>12</td>
<td>SVD-WDBN</td>
<td>97.1</td>
<td>General classifier performance improvement.</td>
</tr>
<tr>
<td>(Chen et al. 2020b)</td>
<td>3D-input (spatial and temporal resolution)</td>
<td>128</td>
<td>No</td>
<td>N/A</td>
<td>18</td>
<td>3D-CNN</td>
<td>98.6</td>
<td>General classifier performance improvement.</td>
</tr>
<tr>
<td>(Ameri et al. 2020)</td>
<td>Transfer learning</td>
<td>5</td>
<td>No</td>
<td>N/A</td>
<td>9</td>
<td>CNN-TL</td>
<td>93.4</td>
<td>Electrode shift.</td>
</tr>
<tr>
<td>(Zea &amp; Benalcázar 2020)</td>
<td>Time series approach</td>
<td>8</td>
<td>Yes</td>
<td>5</td>
<td>TD</td>
<td>LSTM</td>
<td>95.8</td>
<td>Real-time application.</td>
</tr>
<tr>
<td>(Mukhopadhyay &amp; Samui 2020)</td>
<td>Time series approach</td>
<td>7</td>
<td>Yes</td>
<td>6</td>
<td>cTDD</td>
<td>DNN</td>
<td>98.9</td>
<td>Limb position.</td>
</tr>
<tr>
<td>(Yamanoi et al. 2020)</td>
<td>Time-frequency mapping</td>
<td>5</td>
<td>Yes</td>
<td>1 TFD</td>
<td>25</td>
<td>CNN</td>
<td>84.0 (I)</td>
<td>Temporal variation.</td>
</tr>
<tr>
<td>(Kim et al. 2020)</td>
<td>Transfer learning</td>
<td>12</td>
<td>Yes</td>
<td>1 TFD</td>
<td>50</td>
<td>CNN</td>
<td>52.5 (I)</td>
<td>Inter subject variability.</td>
</tr>
<tr>
<td>(Liu et al. 2019)</td>
<td>Topology evaluation</td>
<td>8</td>
<td>No</td>
<td>N/A</td>
<td>10</td>
<td>CNN</td>
<td>93.8</td>
<td>General classifier performance improvement.</td>
</tr>
<tr>
<td>(Olsson et al. 2019b)</td>
<td>Multi-labelled movement approach</td>
<td>128</td>
<td>No</td>
<td>N/A</td>
<td>65</td>
<td>CNN</td>
<td>78.7</td>
<td>General classifier performance improvement.</td>
</tr>
<tr>
<td>(Rahimian et al. 2019)</td>
<td>Time series approach</td>
<td>12</td>
<td>No</td>
<td>N/A</td>
<td>17</td>
<td>DCNN</td>
<td>92.5</td>
<td>General classifier performance improvement.</td>
</tr>
<tr>
<td>(Huang &amp; Chen 2019)</td>
<td>Time series approach</td>
<td>12</td>
<td>Yes</td>
<td>1 TFD</td>
<td>49</td>
<td>CNN-LSTM</td>
<td>79.3</td>
<td>General classifier performance improvement.</td>
</tr>
<tr>
<td>(Tsinganos et al. 2019b)</td>
<td>Divide-and-conquer (input segmentation)</td>
<td>10</td>
<td>No</td>
<td>N/A</td>
<td>53</td>
<td>TCN</td>
<td>89.8</td>
<td>General classifier performance improvement.</td>
</tr>
<tr>
<td>(Wei et al. 2019)</td>
<td>Divide-and-conquer (input segmentation)</td>
<td>128</td>
<td>No</td>
<td>N/A</td>
<td>8</td>
<td>MS-CNN</td>
<td>99.7</td>
<td>General classifier performance improvement.</td>
</tr>
</tbody>
</table>

Table 7: A summary of works associated with deep learning approaches.

In the table: FE is Feature Extraction, EA is Evolutionary Algorithm, TEMPONet is Temporal Embedded Muscular Processing Online Network, ADANN is Adaptive Domain Adversarial Neural Network, WDBN is Wavelet Deep Belief Networks, CNN-TL is Convolutional Neural Network Transfer Learning, DNN is Deep Neural Network, DCNN is Dilated Convolutional Neural Network, TCN is Temporal Convolutional Network, cTDD is correlated Time Domain Descriptors, and MS-CNN is Multi-Stream Convolutional Neural Network.

*The model input was the FFT components of the sEMG signal.

2.6. Robotic Arm

Closing the loop of the movement identification system is the robotic arm that will perform the gesture predicted by the trained model. This step has considerable relevance, as it allows the execution of online tests of the HMI and the evaluation of the applicability of the designed algorithm in a situation close to the real one. Recently, some researches have been conducted aiming at the proposition of an artificial limb. Here, are included the development of robotic hand prototypes.
(Feng et al. 2019; Sánchez-Velasco et al. 2020) and exoskeletons for rehabilitation purposes (Bouteraa et al. 2019; Ma et al. 2019).

In (Sánchez-Velasco et al. 2020) a low-cost prototype of a robotic hand with six degrees of freedom was presented. It has eleven joints, which guarantees the possibility of a great variability of gestures, with emphasis on the inclusion of a junction between the proximal and distal phalanx of the thumb, increasing the naturalness of its movement. The gears were made of acrylic as an alternative to metal, which considerably reduced member costs.

In the work by (Feng et al. 2019), it was proposed a prototype robotic manipulator constructed from polymer with fiberglass reinforcement. With pneumatic control of the joints, it was developed to perform different gripping movements. The artificial hand was validated through the relationship between the angles obtained through simulation and those derived from a model of the distribution of static forces that act on it.

Finally, the development of an exoskeleton from 3D printing was the work by Ma et al. (2019), where they presented a control system based on ADRC (Active Disturbance Rejection Controller) which proved to be superior to a typical PID (Proportional Derivative Integrator) in reference tracking and disturbance rejection. The proposed control method rejected the disturbance faster than the PID and presented a lower Root Mean Square Error (RMSE) in the reference tracking test (0.992 against 1.2409 for the PID).

2.7. Analysis of the Presence of Contaminants in the sEMG Signal

There are intrinsic and extrinsic factors in the acquisition process that influence the proprieties of the EMG signal, manifesting in the form of contaminants. According to literature, these can be categorized into seven categories.

*Inherent noise from electronic equipment and instrumentation:* all electronic devices generate noise that cannot be eliminated, only minimized through the use of high-quality components (Reaz et al. 2006). Here, are included the thermal noise present in the electronic elements of the amplification system, being present in the form of baseline noise (De Luca et al. 2010), operational-amplifier saturation, quantization error, weak contact between the electrode and the skin including its detachment (Fraser et al. 2014), as examples;

*Environmental noise:* this factor is associated with electromagnetic interference. The human body is constantly submitted to electromagnetic radiation, being practically impossible to avoid such exposure (Reaz et al. 2006). In this context, contamination by radio frequency (RF) and that associated with the power line (Fraser et al. 2014) are included, manifesting in high frequencies (RF) and the frequencies of 50 and 60Hz (in power line case, depending on the region) and its harmonics;

*Motion artifact:* it manifests itself through the distortion of information, causing irregularities in the measured signal (Reaz et al. 2006). This contaminant can occur under two circumstances: due to the movement of the electrode cabling and due to the interface between the electrode and the skin (Ijaz & Choi 2018; De Luca et al. 2010). The last, in turn, presents itself in two different situations. One of them is due to the displacement of the muscle to the skin, causing a variation in the load distribution and, consequently, inducing a change in the potential difference of the electrode-skin
interface (De Luca et al. 2010; Pozzo et al. 2004). It can also appear when a force impulse travels through the muscle to the skin, causing a displacement in the sensor-skin interface (De Luca et al. 2010). The motion artifact is present in the low-frequency region, with a spectral range that generally extends from the DC level to 20 Hz (Ijaz & Choi 2018; Pozzo et al. 2004);

*Inherent instability of the EMG signal:* signal amplitude is stochastic in nature and depends on the firing rate of the motor units, which typically ranges from 0 to 20Hz (Reaz et al. 2006). Thus, the energy between the frequencies of 10 and 20 Hz contains peaks whose amplitude depends on the activation rate of the motor units. These values fluctuate mainly in the occurrence of small magnitude contractions, making the energy in this low-frequency range unstable and, consequently, not providing reliable information about muscle activity (De Luca et al. 2010);

*Temporal anomalous muscle activity and inactivity:* the first occurs in the presence of muscle action at a given moment of unexpected time and the second, on the other hand, occurs when muscle inactivity is detected at a given time when an action was expected (Ijaz & Choi 2018);

*Interference by other unwanted biological signals:* are included in this factor the cross talk (measured signal from an inactive muscle and generated by an active one (Pozzo et al. 2004)), interference by electrocardiography (ECG) signal, among others;

*Noise related to the electrode-skin interface:* this interference is also named electrochemical noise. It forms the baseline noise with the thermal noise present in the electronic components of the amplification system. It is detected whenever a sensor is connected to the skin (De Luca et al. 2010).

The presence of these factors in the EMG signal is unwanted and may make it impossible to extract the information depending on the level of contamination. Thereby, many researchers have been working on methods for detection, EMG signal recovery, identification of the type of interference, as well as classification strategies robust to the presence of contaminants (Favieiro & Balbinot 2019; Fraser et al. 2014; Fraser et al. 2012a; Ijaz & Choi 2018; De Luca et al. 2010; Machado et al. 2020; Machado et al. 2019; McCool et al. 2014; De Moura & Balbinot 2018; Stachaczzyk et al. 2020). Table 8 summarizes some of these works, where the abbreviations “Cont.” is Contaminant, “Mov.” is Movements, “Feat.” is features, MI is Movement Identification, CI is Contaminant Identification, CD is Contaminant Detection, “Class.” is classifier, TVARMA is Time-Varying Autoregressive Moving Average, rPCA is robust Principal Component Analysis, SOM is Self-Organizing Map, LSAA is Least Squares Adaptive Algorithm, CMMV is Consecutive Minimum or Maximum Values, SQNR is Signal-to-Quantization Noise Ratio, PCC is Pearson Correlation Coefficient, TS is Template Subtraction, ATS is Adaptive Template Subtraction, MBF is Model Based Filtering, WD is Wavelet Denoising, EMD is Empirical Mode Decomposition, MYOPm is Modified Myopulse Percentage rate, IMCRA is Improved Minima Controlled Recursive Averaging, VAE is Variational Autoencoder, MA is Motion Artifact, AS is Amplifier Saturation, AWGN is Additive White Gaussian Noise, PLI is Power Line, CN is Correlated Noise, MixA is Mix Anomalies, QN is Quantization Noise, ADC-C is Analog-to-Digital Converter Clipping.
### Table 8: A summary of works associated with the treatment of contaminants in EMG recordings.

* The result is the average correlation coefficient between predicted and actual SNR. **The result is the lowest median RMSE of the predicted elbow angle.

<table>
<thead>
<tr>
<th>Authors</th>
<th>Stage</th>
<th>Strategy</th>
<th>Method</th>
<th>Feat.</th>
<th>Mov.</th>
<th>Cont. type</th>
<th>Cont. level [dB]</th>
<th>Class</th>
<th>AvAcc [%]</th>
<th>Issue considered</th>
</tr>
</thead>
<tbody>
<tr>
<td>(De Moura &amp; Balbinot 2018)</td>
<td>Pre-processing</td>
<td>Virtual sensor estimation</td>
<td>TVARMA</td>
<td>5 TD</td>
<td></td>
<td>MA AS AWGN PLI ECG</td>
<td>Not informed</td>
<td>SVM</td>
<td>69.2 (I) 47.5 (A)</td>
<td>EMG signal recovering.</td>
</tr>
<tr>
<td>(McCool et al. 2014)</td>
<td>Pre-processing</td>
<td>Handcrafted features analysis</td>
<td>SVM</td>
<td>3 TD</td>
<td>4 FD</td>
<td>MA AS AWGN PLI ECG</td>
<td>-20 to 20</td>
<td>N/A</td>
<td>99.6 to 20.4 (I)</td>
<td>Contaminant type identification.</td>
</tr>
<tr>
<td>(Ijaz &amp; Choi 2018)</td>
<td>Pre-processing</td>
<td>Clustering</td>
<td>rPCA-SOM</td>
<td>1 TD</td>
<td></td>
<td>N/A</td>
<td>Not informed</td>
<td>N/A</td>
<td>88.2 (I)</td>
<td>Contaminant type identification.</td>
</tr>
<tr>
<td>(Fraser et al. 2014)</td>
<td>Pre-processing</td>
<td>Handcrafted features analysis</td>
<td>SVM</td>
<td>4 TD</td>
<td>2 FD</td>
<td>MA AS AWGN PLI ECG</td>
<td>-20 to 45</td>
<td>N/A</td>
<td>N/A</td>
<td>Contaminant detection.</td>
</tr>
<tr>
<td>(Fraser et al. 2011)</td>
<td>Pre-processing</td>
<td>Contaminant signal estimation</td>
<td>LSAAS</td>
<td>No</td>
<td>N/A</td>
<td>N/A</td>
<td>-10 to 40</td>
<td>N/A</td>
<td>N/A</td>
<td>Contaminant removal.</td>
</tr>
<tr>
<td>(Fraser et al. 2012a)</td>
<td>Pre-processing</td>
<td>Contaminant signal estimation</td>
<td>Moving Average Filter</td>
<td>No</td>
<td>N/A</td>
<td>N/A</td>
<td>-8 to 8</td>
<td>N/A</td>
<td>N/A</td>
<td>Contaminant removal.</td>
</tr>
<tr>
<td>(Fraser et al. 2012b)</td>
<td>Pre-processing</td>
<td>Handcrafted features analysis</td>
<td>SVM</td>
<td>3 TD</td>
<td>4 FD</td>
<td>MA AS AWGN PLI ECG</td>
<td>Not informed</td>
<td>N/A</td>
<td>N/A</td>
<td>Contaminant detection.</td>
</tr>
<tr>
<td>(Machado et al. 2019)</td>
<td>Pre-processing</td>
<td>Virtual sensor estimation</td>
<td>RNN-LSTM</td>
<td>1 TD</td>
<td></td>
<td>AS</td>
<td>Not informed</td>
<td>SVM</td>
<td>66.5 (MI)</td>
<td>EMG signal recovery.</td>
</tr>
<tr>
<td>(Machado et al. 2020)</td>
<td>Pre-processing</td>
<td>Time series approach</td>
<td>RNN-LSTM</td>
<td>No</td>
<td>N/A</td>
<td>MA AWGN PLI ECG</td>
<td>-20 to 20</td>
<td>N/A</td>
<td>87.8 to 20.4 (I)</td>
<td>Contaminant detection; Contaminant type identification</td>
</tr>
<tr>
<td>(De Luca et al. 2010)</td>
<td>Pre-processing</td>
<td>Cutoff frequency determination</td>
<td>Butterworth HP-filter</td>
<td>No</td>
<td>N/A</td>
<td>MA</td>
<td>Not informed</td>
<td>N/A</td>
<td>N/A</td>
<td>Contaminant removal.</td>
</tr>
<tr>
<td>(Stachaczyn &amp; al. 2020)</td>
<td>Pre-processing</td>
<td>Spectro-temporal similarity</td>
<td>Adaptive Spatial Filter</td>
<td>No</td>
<td>4</td>
<td>AWGN PLI</td>
<td>Not informed</td>
<td>LDA</td>
<td>84.5 (MI)</td>
<td>Contaminant detection and attenuation.</td>
</tr>
<tr>
<td>(Oo and Phukpattaranont 2020)</td>
<td>Pre-processing</td>
<td>ANN-based regressor</td>
<td>ANN</td>
<td>1 TD</td>
<td></td>
<td>ECG</td>
<td>-20 to 0</td>
<td>N/A</td>
<td>0.966*</td>
<td>SNR estimation of the contaminated sEMG.</td>
</tr>
<tr>
<td>(Petersen et al. 2020)</td>
<td>Pre-processing</td>
<td>Several</td>
<td>TS, ATS, MBF, WD, EMD, HP-Filter</td>
<td>No</td>
<td>N/A</td>
<td>ECG</td>
<td>-42.6 to 9.7</td>
<td>N/A</td>
<td>N/A</td>
<td>Contaminant removal.</td>
</tr>
<tr>
<td>(Triviyanto et al. 2018)</td>
<td>Feature extraction</td>
<td>Adaptive feature thresholding</td>
<td>MYOPm</td>
<td>1 TD</td>
<td>1 DOF</td>
<td>AWGN</td>
<td>0 to 60</td>
<td>Kalman Filter</td>
<td>9**</td>
<td>Improvement of the system robustness to contamination.</td>
</tr>
<tr>
<td>(Thompson ja et al. 2016)</td>
<td>Pre-processing</td>
<td>Statistical metrics analysis</td>
<td>IMCRA-based Spectral Enhancement</td>
<td>7 TD</td>
<td>7</td>
<td>AWGN</td>
<td>-20 to 20</td>
<td>N/A</td>
<td>N/A</td>
<td>Contaminant detection.</td>
</tr>
<tr>
<td>(McCool et al. 2015)</td>
<td>Pre-processing</td>
<td>Spectral subtraction</td>
<td>IMCRA-based Spectral Enhancement</td>
<td>5 TD</td>
<td>7</td>
<td>AWGN</td>
<td>-10 to 0</td>
<td>SVM</td>
<td>59.9 (I) 82.4 (MI)</td>
<td>Improvement of the system robustness to contamination</td>
</tr>
<tr>
<td>(Fraser et al. 2013)</td>
<td>Pre-processing</td>
<td>Statistical metrics analysis</td>
<td>PCC</td>
<td>No</td>
<td>N/A</td>
<td>MA PLI ECG</td>
<td>-20 to 30</td>
<td>N/A</td>
<td>N/A</td>
<td>Contaminant detection.</td>
</tr>
<tr>
<td>(Teh &amp; Hargrove 2021)</td>
<td>Feature extraction</td>
<td>EMG signal projection</td>
<td>VAE</td>
<td>4 TD</td>
<td></td>
<td>Flatlined signal</td>
<td>Not informed</td>
<td>LDA</td>
<td>&gt;90.0 (MI)</td>
<td>Improvement of the system robustness to contamination</td>
</tr>
</tbody>
</table>

* The result is the average correlation coefficient between predicted and actual SNR. **The result is the lowest median RMSE of the predicted elbow angle.
It is noteworthy that, despite the majority of works listed on Table 8 was not performed aiming the motion recognition task, all apply to it. Thereby, most of them are related to pre-processing stage and propose methods for contaminant detection (it identifies the interference but not the type) or removal. The strategies adopted to contaminant detection generally relies on extracting statistical metrics (Fraser et al. 2013; Thongpanja et al. 2016) and handcrafted features of sEMG signal (Fraser et al. 2014; Fraser et al. 2012b). For contaminant removal, it is commonly employed filters (De Luca et al. 2010; Petersen et al. 2020) or techniques to estimate the interference parameters (Fraser et al. 2012a; Fraser et al. 2011).

However, despite the importance of contamination detection for data quality assessment, knowing the type of interference is also desirable for eliminating its source by correcting the acquisition procedure. Within this context are the works (Ijaz & Choi 2018; Machado et al. 2020; McCool et al. 2014). McCool et al. (2014) implemented an SVM-based classifier to recognize five different sources of contamination (ECG, motion artifact, amplifier saturation, power line interference, and baseline noise, represented by additive white Gaussian noise). Through an experiment with signal contamination ranging from -20 to 20 dB, hit rates of 100% and 97.79% for -20 and -10 dB were achieved. However, as the interference intensity decreased, the algorithm performance was reduced, registering an accuracy of only 20% to 20 dB.

In contrast, Ijaz and Choi (2018) presented an unsupervised algorithm for detecting and identifying the contaminant type. It includes a pre-processing stage composed of filtering (Wavelet Transform) and dimensionality reduction (rPCA). Thus, the principal components are considered for determining a Self-Organizing Map (SOM) that clusters the data. By implementing this methodology, it was achieved success rates of 88.2% in the separation of five different anomalies (baseline noise, ECG, correlated noise, electrical network, and mixture of different contaminants) and clean signal. The method's main contribution is to perform the recognition in an unsupervised way. However, the SNR of the contamination was not reported.

Dispensing with a pre-processing and feature extraction step was the innovation introduced by Machado et al. (2020). In their work, the identification of four contaminant types (power line, motion artifact, white noise characterizing baseline noise, and ECG) plus clean signal was assigned to a classifier based on Recurrent Neural Networks (RNR) and LSTM. An accuracy of 97.72% was achieved by processing data contaminated with -20 dB SNR.

Aiming at recovering the contaminated sEMG signal, Fraser et al. (2011, 2012a) and De Luca et al. (2010) proposed strategies for removing power line, ECG, and low-frequency noise caused by motion artifact, respectively. Fraser et al. presented two algorithms to estimate the parameters of interference by power line and ECG. In the first work by Fraser et al. (2011), a Least Squares Adaptive Algorithm (LSAA) estimated the amplitude, phase, and frequency of a sinusoidal signal superimposed to the sEMG (power line interference). In the more recent work, Fraser et al. (2012a) used a moving-average filter for obtaining the ECG signal from sEMG recording. In both methods, the obtained contaminant estimation was subtracted from the sEMG signal. The LSAA proved to be efficient to remove power line interference with SNR of 15 dB or less. The RMSE of the Moving-Average filter was compared with that of a classic method (measurement of a standard ECG signal for later use in removing the interference present in the signal). The new approach outperformed the
classic one in more severe contaminations. It was reported a 15.5% reduction in the RMSE for the interference with SNR of 2 dB.

In (De Luca et al. 2010), a study was carried out on the best choice for the cutoff frequency of a high-pass filter in removing noise caused by motion artifact. By performing some tests ranging the cutoff frequency from 1 to 30 Hz of a Butterworth filter with a transition zone inclination of 12 dB/octave, it was concluded that the choice of the value depends on the type of muscle and of the intended analysis. For the study of isometric movements or common and natural gestures, 20 Hz was recommended. However, if the goal is to evaluate more vigorous movements, such as those associated with sports, it is suggested to increase the frequency beyond 20 Hz, for greater attenuation of movement artifact interference, but at the cost of inducing deformations and additional attenuation in the EMG signal spectrum.

However, as an alternative to the pre-processing algorithms for contaminant type identification and interference removal, some researchers aim to propose techniques for EMG signal recovering (Machado et al. 2019; De Moura & Balbinot 2018) and system robustness improvement to contaminants (McCool et al. 2015; Teh & Hargrove 2021; Triwiyanto et al. 2018). Machado et al. (2019) and De Moura and Balbinot (2018) presented virtual sensor-based strategies to recover the information of a contaminated channel. De Moura and Balbinot (2018) used cross-correlation between acquisition channels (considering multichannel configuration) to estimate a model for the output of a contaminated electrode. Hence, the generated data is independent of the physically measured signal. For that, they proposed a strategy based on the Time-Varying Autoregressive Moving Average (TVARMA) and another based on the Time-Varying Kalman filter (TVK). To detect the presence of contamination in the signal and enable the use of the sensor, a one-class SVM-based model was used. The method performance assessment was made by recognizing 17 movements through 12 sEMG channels and an SVM classifier. The contaminants considered were ECG, white Gaussian noise, motion artifact, power line, and amplifier saturation. The algorithm efficiency was verified by comparing the classifier performance with and without contamination. Applying the virtual sensor model developed from TVARMA, it was obtained a reduction (on average) of only 5.6% in the accuracy of the classifier to the test without contamination, outperforming its opponent (TVK).

In contrast, the proposal by Machado et al. (2019) is based on the estimation of the signal of a corrupted channel from a regression model obtained from the other (clean) channels. A hybrid system was implemented for the regression step, comprising a Recurrent Neural Network and LSTM. Reported results showed an increase from 9 to 66% in the classifier's assertiveness by reconstructing a channel contaminated with noise from amplifier saturation compared to using the corrupted signal (without applying the method). Here, SVM was considered for identifying 17 movements of the hand-arm segment.

The promising results of the virtual sensor strategy indicate it as a good option for EMG signal recovery. However, if a great number of channels are contaminated, the data reconstruction will be impaired. Thereby, McCool et al. (2015) presented a pre-processing algorithm called Improved Minima Controlled Recursive Averaging (IMCRA)-based Spectral Enhancement to attenuate the white Gaussian noise in the sEMG signal. The algorithm computes the STFT to estimate the noise present in the signal recursively in each time/frequency segment. After noise estimation, it is
subtracted from the contaminated signal. Average accuracies of 59.4 and 82.4 were achieved in classifying seven movements through data contaminated with SNR of -10 and 0 dB, respectively, demonstrating the method efficiency.

Therefore, contaminants in the sEMG signal are one of the factors that lead to the reduced performance of classification systems when tested outside the laboratory. Consequently, for a myo-controlled device to be effective also in non-ideal situations, it is essential that it includes the treatment of contaminants in its control system, either by a specific method in the signal preprocessing step or built into the motion classifier. This section summarized some promising and innovative strategies to solve this issue. However, most of them were not considered in the motion recognition context hence their applicability in the real-time scenario is not guaranteed. Therefore, there are still some steps to be addressed to improve current and future myo-controlled devices. Besides the real-time issue, the inherent temporal variability of the sEMG signal properties and the occurrence of simultaneous contaminants should also be considered in the algorithms for detecting and removing interferences. These factors will probably guide the research's next steps.

Furthermore, all the works cited in Table 8 used artificially contaminated EMG signals to validate the proposed models. Although this procedure is acceptable in the literature and widely used, it can be a problem when applied to motion classification systems since it may not faithfully represent the way it is presented in practice during the use of the myo-controlled device. Therefore, the need to generate a database with EMG signals in non-ideal situations is identified, reproducing possible contamination, similar to those considered in current research. Thus, for example, tests with volunteers performing movements with poorly glued electrodes, with movement artifacts, in environments with excessive electromagnetic interference, and with muscle fatigue, among other factors, would be fundamental to observe how the contaminants would appear in practice. In this way, they will guide the development of algorithms for the treatment of contaminants that will have a better performance during the use of the device by the user and possibly will help to increase the accuracy of the classifiers outside the laboratory environment.

3. Conclusion

This paper presents a review of the most recent works related to sEMG movement recognition. Thus, the current scenario and trends associated with each stage of the system were summarized. For signal acquisition, there are some open topics such as the electrodes number and placement setup, acquisition hardware, mixed signals, to name a few. Moreover, some databases are indicated to support future researches to address these issues. Among the pre-processing strategies listed in the respective section, the majority are mainly based on data transformation for inter-subject/temporal variability addressing, data quality improvement, and general classifier performance increasing. Aiming to propose new and innovative approaches to extract sEMG signal information, several authors presented techniques exploring the spatial domain enabled by electrodes arrangement as an alternative to the traditional features computed on time and frequency domains. The promising results showed in Table 3 indicate it as a relevant option to the feature extraction stage for solving
the confounding factors related to temporal/inter-subject variability and changing in the intensity of muscle activation.

In the sequence, it was presented several works aiming at finding the best set of features. The approaches are mainly related to wrapper strategies (according to Table 4) and can be grouped in feature, channel, or combined feature/channel pairs selection. Using the last form may lead the classifier to reach higher accuracies by expanding the search field. However, individual analysis of features or channels is an interesting option when computational cost reduction is a priority.

The motion recognition step can be executed by a classifier (control based on a finite set of movements) or by a regressor (proportional control). Aiming at proportional control, the works covered by this review are often related to strategies for angular position prediction of finger or wrist joints. The methods adopted ranging from classical machine learning algorithms (ANN, RLR) to deep learning (CNN, DNN, LSTM), both with promising results. The same distinction can be made between the propositions for the movement classification stage. Despite the increasing tendency of using deep learning techniques (Table 7), the approaches around classical machine learning theory are still quite considered (Table 6). Among the deep learning strategies, it is noteworthy the time-series approaches. The promising results (Table 7) indicate the inclusion of temporal resolution in the model training as a good option to address the issues related to limb position and temporal variation. Furthermore, the transfer-learning technique appears as a promising solution for electrode displacement and inter-subject variability.

Finally, the contaminant issue was considered in a separate section in the paper. The approaches are mainly related to pre-processing step and are not directed to the motion recognition task (in the most part). Despite the promising strategies presented for interference detection, identification, and removal, there are still some factors to be considered, such as the application in real-time systems, the inherent temporal variability of the sEMG signal properties, and the occurrence of simultaneous contaminants.

In summary, this review exposes the current scenario of the movement classification system, providing valuable information for new researchers and guide future works towards myo-controlled devices.

Acknowledgements

This study was financed in part by the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (CAPES) - Finance Code 001 and by the Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq).

The contributions of each author are: **Maurício Cagliari Tosin**: Conceptualization, Methodology, Formal analysis, Investigation, Writing – Original Draft. **Juliano Costa Machado**: Conceptualization, Formal analysis, Writing - Review & Editing. **Alexandre Balbinot**: Conceptualization, Writing - Review & Editing, Supervision, Project administration.

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.
Conflicts of Interest

The authors declare no conflicts of interest.

References


Asogbon, M. G., Samuel, O. W., Geng, Y., Oluwagbemi, O., Ning, J., Chen, S., ... Li, G. (2020). Towards resolving the co-existing impacts of multiple dynamic factors on the performance of EMG-pattern recognition based prostheses. Computer Methods and Programs


TOSIN, MACHADO, & BALBINOT


Measurement and Control in Robotics (ISMCR) (pp. A1-4-1-A1-4-11). Houston, TX, USA : IEEE. https://doi.org/10.1109/ISMCR47492.2019.8955725


