On the Use of Kernel Fisher Discriminant Analysis as a Reduction Method for the Classification of EMG Signals

Ines Moudjari
Univ Rennes, Inserm, LTSI - UMR 1099, F-35000
Rennes, France
ines.moudjari@univ-rennes.fr

Caroline Pautard
Blueback
Cesson-Sévigné, France
cpautard@blueback.fr

Clément Jouanneau
Blueback
Cesson-Sévigné, France
cjouanneau@blueback.fr

Régine Le Bouquin Jeannès
Univ Rennes, Inserm, LTSI - UMR 1099, F-35000
Rennes, France
regine.le-bouquin-jeannes@univ-rennes.fr

***Abstract*—The muscular system is quite complex and can be divided into smaller systems. In this paper, we focus on the abdominal wall, in particular on the two deepest muscles that compose it, the transversus abdominis and the obliquus internus. Both muscles play an important role in many physiological phenomena, such as breathing. The purpose of this paper is to identify the co-contraction patterns of these two muscles. To this end, we use a combination of two well-known methods. Firstly, the kernel Fischer discriminant analysis (K-FDA) is used to transform the data extracted from surface electromyographic signals, acquired from one bipolar electrode, in order to build a representation of the data that facilitates the classification. Then, a support vector machine is used for the classification step. We tested four types of kernel for the K-FDA, namely linear, radial basis function, sigmoid and polynomial. Following a five-fold cross validation, we obtained an accuracy up to 100%.**

***Index Terms*—K-FDA, classification, support vector machine, deep abdominal muscles, physiotherapy.**

I. Introduction

The muscular system is made up of several muscles that work in synergy. The simplest and most everyday movements, such as breathing, require the activation and relaxation of several muscles in the body. This system can be divided into several sub-systems to simplify its study. For example, when studying a hand movement, it would be more interesting to focus on the forearm muscles. A deeper understanding of these synergistic and antagonistic systems could allow better rehabilitation of patients suffering from different pathologies. Due to its anatomical position, the abdominal wall has a central role in the transfer of forces from the trunk to the legs, which makes the muscles that compose it a very important part of the muscular system. The abdominal wall is composed of five muscles. In this paper, we only focus on two of these muscles, $i.e.$, the transversus abdominis (TRA) and the obliquus internus (OI), since they are the deepest muscles of the abdominal wall, making them difficult to access. The TRA has an important role in many physiological phenomena as breathing or delivery. It is also involved in some pathologies, as low back pain or urinary incontinence. More and more physiotherapists are devoting part of their rehabilitation to teaching their patients to voluntarily contract this muscle. On the other hand, the OI is a synergistic muscle of the TRA and is, for the most part, involved in the same physiological phenomena as the TRA. However, this muscle overlaps the TRA, which makes it difficult to acquire information on the TRA without interference from the OI. Diverse measuring devices are used to study the muscular system, such as electromyography (EMG) or ultrasounds. With a single surface EMG electrode, a mixture of TRA and OI signals can be easily extracted [1]. Now, our concern consists in differentiating between TRA and OI muscles using only one measuring electrode, which is a real challenge.

Linear discriminant analysis (LDA) has been widely used, as reduction method or as classification method; however, not all discrimination problems can be solved with linear method. To deal with this issue, two decades ago, the Kernel Fisher Discriminant Analysis (K-FDA) was developed. In this paper, we use K-FDA to transform our dataset before applying a classification algorithm. In the last decades, a number of classification algorithms have been developed to solve different issues [2]–[8]. In this work, we chose the Support Vector Machines (SVMs) because they are easy to use and have proven to be efficient in classifying EMG signals. Actually, several studies have applied these types of algorithms for different EMG signals classification purposes. Indeed, SVMs have been used to classify hand gestures [4], [8], arm or finger movements [2], [6], wrist-motion [3] or muscles and neuromuscular diseases [5], [7]. Now, these studies have considered surface muscles, as biceps brachii and medial vatus, whilst we are concerned by deep muscles. One team has worked with needle electrodes [5] whereas others [2]–[4], [6]–[8] used Ag/AgCl electrodes, while we used Datwyler softpulse™ dry electrodes.

The aim of our approach is to identify the co-contraction patterns of the two deep abdominal muscles mentioned above, namely the TRA and the OI. To this end, we first labelled the signals used in the following analyses according to four groups corresponding to the co-contraction pattern. The method presented in this paper consists in applying a K-FDA to the

initial data and then classifying the signals projected into the new basis into these four groups. In this contribution, Section II is devoted to the presentation of the method. Section III illustrates the results of classification and Section IV concludes this work and opens on some perspectives.

II. Materials and Methods

A. Acquisition and preprocessing

The data used in this paper come from an experiment conducted on two healthy volunteers (one female and one male). During the experiment the volunteers had to perform nine different exercises, which are presented in Table I. The exercises were performed in two positions, standing and supine.

TABLE I
List of exercises

Exercises in supine position	Exercises in standing position
Coughing	Coughing
Complete exhalation	Complete exhalation
Draw-in	Draw-in
Crunch hold 2s	Held arm curl with elastic
Linked crunch	Fast arm curl with elastic
Circle straight leg	Knee raise

To acquire the signals, we used a portable EMG device, called Blueback physio® and developed by the Blueback company (see Fig. 1). This device is by design composed of two bipolar electrodes which are able to acquire two distinct signals. In our studies, we kept the signals from each of the electrodes in order to double the number of usable signals per patient. Each of these signals was saved and preprocessed independently.

Fig. 1. The Blueback physio® device

The sampling rate of the device was equal to 500 Hz. The electrodes were placed 2 cm inward and 2 cm downward of the antero-superior illiac spine. The preprocessing is decomposed in 4 steps. First, we re-sampled the signals up to a sampling rate of 1000 Hz. Then, we applied a highpass filter with a cut-off frequency of 30 Hz to the re-sampled signals and we took the absolute value. Lastly, we segmented the signals into samples of 1 second and a total of 360 samples of 1 s were used. During the experiment, ultrasound images were obtained to get the ground-truth on the contraction of the two muscles of interest every second. We measured the thickness of each muscle on all images (see Fig. 2). A muscle was considered to be contracted as soon as its thickness exceeded a threshold fixed to the average thickness of this muscle for a given subject and body position (supine or standing).

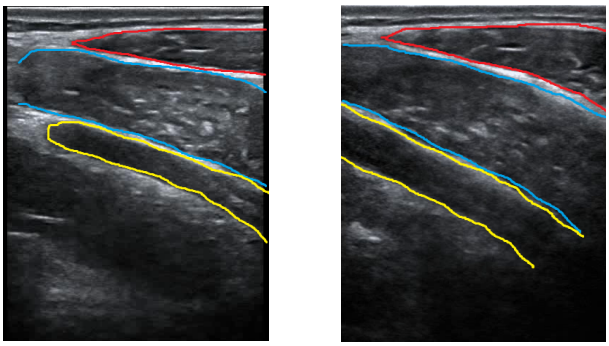

Fig. 2. Annotated ultrasound image, with the obliquus internus (in blue) and transversus abdominis (in yellow) muscles

B. Features

As detailed hereafter, we extracted 23 features of interest, including temporal, frequency and time-frequency features. The choice of these features comes from a bibliographical search in existing EMG classification studies as well as in other fields of research.

For each of the following features, let us assume that we have one signal x (corresponding to the signal collected on one electrode at the sampling frequency f_s) of N temporal points long, where N is equal to 1000. The periodogram and frequency are denoted by P and f respectively.

1) Temporal features: We have calculated 12 temporal features. The first temporal feature is the duration (expressed in second) of a contraction burst, noted WL. It corresponds to the difference between the time when the muscle contraction begins, $i.e.$, when the amplitude exceeds a threshold fixed at a quarter of the maximum value of the absolute value of the amplitude and the time when it ends, $i.e.$, when the amplitude falls below this threshold.

Then, we calculated the mean average value, MAV, and the root mean square, RMS, of the amplitude of the time signal.

The next three features are the Hjorth parameters, which are activity, mobility and complexity. The activity, A, represents the variance of the signal x. The mobility, mo, is the ratio between the variance of the first derivative of x and the variance of x,

$$mo = \sqrt{\frac{\sigma^2(x')}{\sigma^2(x)}} \tag{1}$$

where x' is the first derivative of x and σ^2 represents the variance of the signal. The complexity, c, allows to compare the signal to a sine wave:

$$c = \frac{mo(x')}{mo(x)} \tag{2}$$

where $mo(x)$ is the mobility of the signal x and $mo(x')$ is the mobility of the first derivative of x.

We also used static features such as the skewness, sk, which is the coefficient of asymmetry and the kurtosis, ku, which is the signal flattening coefficient respectively given by:

$$sk = \frac{\sum_{i=1}^{N}(x_i - \bar{x})^3}{(N-1) * \sigma^3} \tag{3}$$

$$ku = \frac{\sum_{i=1}^{N}(x_i - \bar{x})^4}{(N-1) * \sigma^4} \tag{4}$$

where $\bar{x}$ is the mean of x and σ is its standard deviation.

The last temporal features are classical entropy estimators such as sample entropy, $SampEn$, [9], approximate entropy, $ApEn$, multi-scale entropy, mse, and Shannon entropy, SEn. To calculate these features we used a library provided by Python.

2) Frequency features: We have considered 4 frequency features. The mean frequency, fm, is calculated on periodograms obtained with the Welch method (256-point Hanning window, 50% overlap, $NFFT$ frequency points) [10]:

$$fm = \frac{\sum_{i=1}^{NFFT} f_i * P(f_i)}{\sum_{i=1}^{NFFT} P(f_i)} \tag{5}$$

where $NFFT$ is the number of frequency points which is equal to N.

As for the median frequency, fmd, it is such as

$$\frac{\sum_{i=1}^{n_{fmd}} P(f_i)}{\sum_{i=n_{fmd}}^{NFFT} P(f_i)} = 1 \tag{6}$$

where n_{fmd} is the index of the fmd value in the frequency vector. The third feature is a ratio of power P_2 between the sum of the powers in the $[100; 200]$ Hz band and the sum of the total power:

$$P_2 = \frac{\sum_{i=100}^{200} P(f_i)}{\sum_{i=1}^{NFFT} P(f_i)} \tag{7}$$

Lastly, we computed a cut-off frequency, f_p, which has been determined by quantifying the power of the signal spectrum compared to a fixed threshold, th:

$$\frac{\sum_{i=1}^{n_{fp}} P(f_i)}{\sum_{i=1}^{NFFT} P(f_i)} \times 100 > th \tag{8}$$

where n_{fp} is the index of the fp value in the frequency vector.

3) Time-frequency features: We have calculated 7 time-frequency features. First, we applied the Empirical Mode Decomposition (EMD) [11] to decompose the signal into mono-components called Intrinsic Mode Functions (IMFs), where an IMF is an oscillating function with zero mean. The features extracted from this EMD are the median frequencies of the first 3 IMFs. In addition, a wavelet transform has been applied on the signals [12]. Three features were extracted from this representation, namely the maximum amplitude, m_{WT}, as well as the frequency, f_{WT}, and the time, t_{WT}, when this maximum m_{WT} occurs. Another feature, the coefficient of flatness, $TF_{flatness}$, was calcuted on this representation:

$$TF_{flatness} = \frac{\prod_{n=0}^{L-1}\prod_{m=0}^{M-1}\bar{S}[n,m]^{\frac{1}{ML}}}{\frac{1}{ML} \times \sum_{n=0}^{L-1}\sum_{m=0}^{M-1}\bar{S}[n,m]} \tag{9}$$

where L is the number of raws of the wavelet transform, M the number of columns and $\bar{S}$ is the absolute of the squared wavelet transform.

4) Reduced learning set: Given these 23 features, we used a wrapper algorithm [13], [14], [15] to find an optimal base containing 5 features. The 5 selected features are the frequency f_p, the ratio P_2, the coefficient of flatness $TF_{flatness}$ and two classical entropy estimators, which are the multi-scale entropy mse and the sample entropy $SampEn$.

C. Kernel Fisher Discriminant Analysis (K-FDA)

The K-FDA is an adaptation of the LDA, where the dot product is replaced by a kernel. The purpose of the K-FDA is the same as that of LDA, which is to maximize the between-class variance and to minimize the within-class variance. The optimization problem can be written as

$$\max_{\boldsymbol{\theta}} \frac{\boldsymbol{\theta}^T \boldsymbol{B} \boldsymbol{\theta}}{\boldsymbol{\theta}^T \boldsymbol{W} \boldsymbol{\theta}} \tag{10}$$

where $\boldsymbol{\theta} \in \mathbb{R}^{n_s \times 1}$ corresponds to the eigenvector with the maximum eigenvalue, also called Fisher axis or Fisher direction. n_s is the number of samples. $\boldsymbol{B} \in \mathbb{R}^{n_s \times n_s}$ corresponds to the between-scatter matrix, defined by:

$$\boldsymbol{B}(i,k) = \sum_{j=1}^{c}(\boldsymbol{m}_j(i) - \boldsymbol{m}_*(i))(\boldsymbol{m}_j(k) - \boldsymbol{m}_*(k))^T \tag{11}$$

where c is the number of classes, $\boldsymbol{m}_j \in \mathbb{R}^{n_s \times 1}$ is the mean vector of the j-th class, such as its i-th entry is

$$\boldsymbol{m}_j(i) = \frac{1}{n_j}\sum_{k=1}^{n_j}\mathcal{K}(\boldsymbol{w}(i), \boldsymbol{w}(k)) \tag{12}$$

where n_j is the number of samples of the j-th class, $\boldsymbol{w}$ is the features vector, $\mathcal{K}$ is the kernel function and $\boldsymbol{m}_* \in \mathbb{R}^{n_s \times 1}$ is such that its i-th entry is

$$\boldsymbol{m}_*(i) = \frac{1}{n_s}\sum_{k=1}^{n_s}\mathcal{K}(\boldsymbol{w}(i), \boldsymbol{w}(k)) \tag{13}$$

where $\boldsymbol{w}(d)$ $(d = i, k) \in \mathbb{R}^{n_f \times 1}$, n_f being the number of features. Finally, $\boldsymbol{W} \in \mathbb{R}^{n_s \times n_s}$ is the within-class scatter matrix, given by:

$$\boldsymbol{W}(i,k) = \sum_{j=1}^{c}\sum_{\eta=1}^{n_j}\boldsymbol{K}_j(i,\eta)\boldsymbol{H}_j(\eta,k)\boldsymbol{K}_j^T(\eta,k) \tag{14}$$

with $\boldsymbol{K}_j \in \mathbb{R}^{n_s \times n_j}$ such as:

$$\boldsymbol{K}_j(i,k) = \mathcal{K}(\boldsymbol{w}(i), \boldsymbol{w}_j(k)) \tag{15}$$

where $\boldsymbol{w}_j$ is a feature vector of the jth class. $\boldsymbol{H}_j \in \mathbb{R}^{n_j \times n_j}$ is the centering matrix, $\boldsymbol{H}_j = \boldsymbol{I} - \frac{1}{n_j}\boldsymbol{J}$, where $\boldsymbol{I} \in \mathbb{R}^{n_j \times n_j}$ is the identity matrix, $\boldsymbol{J} \in \mathbb{R}^{n_j \times n_j}$ is a matrix of all 1's. We use

four kernels for this step, which are linear (LINEAR), radial basis function (RBF), sigmoid (SIGMOID) and polynomial (POLY).

D. Classification

The classification algorithm used in this study is a kernel support vector machine (SVM). The principle of this algorithm is to project the feature vector into a higher-dimensional space in order to make the classes more separable. This problem consists in optimizing the following dual problem:

$$\min_{\boldsymbol{\alpha}} \frac{1}{2} \sum_{i=1}^{n_s} \sum_{j=1}^{n_s} \boldsymbol{y}(i)\boldsymbol{y}(j)\mathcal{K}(\boldsymbol{w}(i), \boldsymbol{w}(j))\boldsymbol{\alpha}(i)\boldsymbol{\alpha}(j) \tag{16}$$

where n_s is the number of samples, $\mathcal{K}(\boldsymbol{w}(i), \boldsymbol{w}(j))$ is the kernel between the vectors of features $\boldsymbol{w}(i)$ and $\boldsymbol{w}(j)$, $\boldsymbol{y} \in R^{n_s}$ corresponds to the vector of true classes and $\boldsymbol{\alpha} \in R^{n_s}$ is the vector of Lagrangian multipliers. The decision function, d, can be expressed as:

$$d(\boldsymbol{w}(p)) = sgn(\sum_{i=1}^{n_s} \boldsymbol{y}(i)\boldsymbol{\alpha}(i)\mathcal{K}(\boldsymbol{w}(i), \boldsymbol{w}(p)) + b) \tag{17}$$

where $w(p)$ is the vector of features to be classified and b is the threshold. We tested three different kernels for this step, namely linear (LINEAR), radial basis function (RBF) and polynomial (POLY) kernels.

III. Results

A. Classes

We considered four classes. Each class represents one scheme of co-contraction of the two muscles under study. In the OI class, the obliquus internus is the only contracted muscle. The R class corresponds to no contraction. In the third TRA class, the transversus abdominis is the only muscle contracted. Lastly, the group TRA-OI is constituted of signals where both muscles are contracted simultaneously.

B. Data projection following K-FDA

As indicated before, we tested four kernels in the K-FDA. The linear kernel discriminant analysis was not able to discriminate the different classes, hence it is not represented in this section. Since the polynomial and sigmoid kernels led to comparable performance, we represent on Fig. 3 the projection of the original data into the K-FDA database along the first three discriminant axes using the sigmoid kernel. From this figure, it is easy to discriminate the four classes. However, some samples of different classes are too close, which may create some confusion for a classification algorithm. The first discriminant axis allows us to make a first separation. As a matter of fact, on one side, we can distinguish TRA-OI and OI and, on the other side, R and TRA are grouped. Moreover, the second discriminant axis makes it possible to dissociate the remaining groups, in particular by separating TRA-OI from OI and R from TRA.

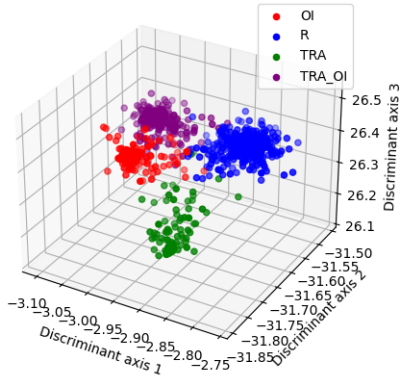

Fig. 3. Data projection along the first three discriminant axes of K-FDA using a sigmoid kernel

In a second step, we represent on Fig. 4 the projection using the RBF kernel, which allows to obtain the best trade-off between minimizing the within-class variance and maximizing the between-class variance. The four classes are well spaced out, the within-variance being very low and the between-variance being very high. The first discriminant axis separates two clusters. The first one is composed of TRA-OI and TRA classes, while the second cluster is formed by R and OI classes. In addition, the second discriminant axis allows the separation of the groups constituting the above-mentioned clusters. Indeed, it allows to dissociate TRA-OI from TRA and R from OI.

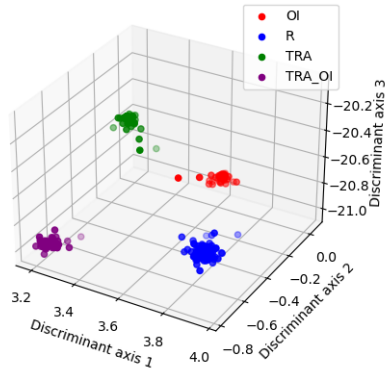

Fig. 4. Data projection along the first three discriminant axes of K-FDA using a RBF kernel

C. Performance

The performance of the classification algorithms was tested following a k-fold cross-validation. We used 5 equally distributed subsets. Then, we calculated the accuracy on the confusion matrix.

Table II reports the performance of the classification algorithms according to the kernel used in the discriminant analysis and the one used in the SVM classifier. On the one hand, the discriminant analysis with linear kernel leads to a very low classification performance, with an accuracy of 48.89%. On the other hand, we observe that, whatever the kernel used in

the classifier, the RBF kernel used in the discriminant analysis leads to 100% accuracy. This result is not surprising given the representation obtained from the RBF kernel discriminant analysis. Discriminant analysis with POLY and SIGMOID kernels also allow us to obtain satisfactory results with an accuracy equal to 99.72% and 97.78% respectively when coupled with a SVM classifier using a RBF kernel.

TABLE II
PERFORMANCE OF THE ALGORITHMS

Kernel of K-FDA	Kernel of classifier	Accuracy
RBF	**RBF**	**100%**
RBF	**POLY**	**100%**
RBF	**LINEAR**	**100%**
POLY	RBF	99.72%
POLY	LINEAR	99.72%
SIGMOID	RBF	97.78%
SIGMOID	POLY	97.22%
SIGMOID	LINEAR	97.78%
LINEAR	RBF	48.89%
LINEAR	POLY	34.72%

IV. CONCLUSION

The main results of this contribution are the great performance of classification obtained by the means of K-FDA. The RBF kernel discriminant analysis has achieved the best results with 100% of accuracy. Likewise, the accuracy of POLY and SIGMOID kernels was satisfying with a ratio of 99.72% and 97.78% respectively. These results show that this method is able to differentiate the co-contraction patterns of two deep muscles of the abdominal wall. We can compare our results with those of Alkan $et\ al.$ [2] and Castiblanco $et\ al.$ [8] who also used a SVM for the classification step. They achieved 99% and 82%, respectively. Our results are in the same range. However, they intended to classify hand and forearm movements, whereas we wanted to distinguish co-contraction patterns. Moreover, in [8], they used a total of six electrodes (four on the right arm and two on the left arm) to collect the signals. As for Subasi $et\ al.$ [7], they achieved 99% of accuracy to classify neuromuscular disorders. Unlike us, they used needle electrodes to acquire the signals, which is an invasive technique even if it eliminates crosstalk between muscles. The use of a needle EMG is not possible in our case since it can only acquire information on a single muscle. Now, none of the previous studies addressed the identification of co-contraction patterns with only one bipolar electrode for the acquisition, and, to the best of our knowledge, no such study has been proposed. We are also aware of the current limitations of our study. Indeed, we exploited signals from only two volunteers, which can lead to bias because EMG signals are highly variable between subjects. Moreover, both volunteers were in good physical condition and we know that physiological pathology can change co-contraction patterns, as well as intrinsic muscle functioning. However, this study proved that it was possible to identify the co-contraction patterns of two deep abdominal muscles from a single surface electrode. In a future work, we plan to acquire more signals to confirm our results on a larger database. It would be interesting to acquire these signals on healthy and pathological subjects. Finally, this type of algorithm can bring benefits in research and clinical practice in physiotherapy. Indeed, an algorithm capable of recognizing muscle patterns can be useful in the study of muscular system functioning as a set of synergistic and antagonistic muscles.

REFERENCES

[1] C. Crasto, A. M. Montes, P. Carvalho, and J. C. Carral, "Abdominal muscle activity and pelvic motion according to active straight leg raising test results in adults with and without chronic low back pain," *Musculoskeletal Science and Practice*, vol. 50, p. 102245, 2020.

[2] A. Alkan and M. Günay, "Identification of emg signals using discriminant analysis and svm classifier," *Expert Systems with Applications*, vol. 39, no. 1, pp. 44–47, 2012.

[3] K. S. Kim, H. H. Choi, C. S. Moon, and C. W. Mun, "Comparison of k-nearest neighbor, quadratic discriminant and linear discriminant analysis in classification of electromyogram signals based on the wrist-motion directions," *Current Applied Physics*, vol. 11, no. 3, pp. 740–745, 2011.

[4] M. Tavakoli, C. Benussi, P. Alhais Lopes, L. B. Osorio, and A. T. de Almeida, "Robust hand gesture recognition with a double channel surface emg wearable armband and svm classifier," *Biomedical Signal Processing and Control*, vol. 46, pp. 121–130, 2018.

[5] R. Dubey, M. Kumar, A. Upadhyay, and R. B. Pachori, "Automated diagnosis of muscle diseases from emg signals using empirical mode decomposition based method," *Biomedical Signal Processing and Control*, vol. 71, p. 103098, 2022.

[6] M. V. Arteaga, J. C. Castiblanco, I. F. Mondragon, J. D. Colorado, and C. Alvarado-Rojas, "Emg-driven hand model based on the classification of individual finger movements," *Biomedical Signal Processing and Control*, vol. 58, p. 101834, 2020.

[7] A. Subasi, E. Yaman, Y. Somaily, H. A. Alynabawi, F. Alobaidi, and S. Altheibani, "Automated emg signal classification for diagnosis of neuromuscular disorders using dwt and bagging," *Procedia Computer Science*, vol. 140, pp. 230–237, 2018. Cyber Physical Systems and Deep Learning Chicago, Illinois November 5-7, 2018.

[8] J. C. Castiblanco, S. Ortmann, I. F. Mondragon, C. Alvarado-Rojas, M. Jöbges, and J. D. Colorado, "Myoelectric pattern recognition of hand motions for stroke rehabilitation," *Biomedical Signal Processing and Control*, vol. 57, p. 101737, 2020.

[9] N. Gozzi, L. Malandri, F. Mercorio, and A. Pedrocchi, "Xai for myo-controlled prosthesis: Explaining emg data for hand gesture classification," *Knowledge-Based Systems*, vol. 240, p. 108053, 2022.

[10] C. Tepe and M. C. Demir, "The effects of the number of channels and gyroscopic data on the classification performance in emg data acquired by myo armband," *Journal of Computational Science*, vol. 51, p. 101348, 2021.

[11] V. K. Mishra, V. Bajaj, A. Kumar, D. Sharma, and G. Singh, "An efficient method for analysis of emg signals using improved empirical mode decomposition," *AEU - International Journal of Electronics and Communications*, vol. 72, pp. 200–209, 2017.

[12] J. M. Fajardo, O. Gomez, and F. Prieto, "Emg hand gesture classification using handcrafted and deep features," *Biomedical Signal Processing and Control*, vol. 63, p. 102210, 2021.

[13] M. R. Alnowami, F. A. Abolaban, and E. Taha, "A wrapper-based feature selection approach to investigate potential biomarkers for early detection of breast cancer," *Journal of Radiation Research and Applied Sciences*, vol. 15, no. 1, pp. 104–110, 2022.

[14] W. Liu and J. Wang, "Recursive elimination current algorithms and a distributed computing scheme to accelerate wrapper feature selection," *Information Sciences*, vol. 589, pp. 636–654, 2022.

[15] J. Linja, J. Hämäläinen, P. Nieminen, and T. Kärkkäinen, "Feature selection for distance-based regression: An umbrella review and a one-shot wrapper," *Neurocomputing*, vol. 518, pp. 344–359, 2023.