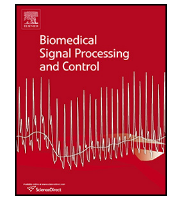
Biomedical Signal Processing and Control 59 (2020) 101920

Contents lists available at ScienceDirect

Biomedical Signal Processing and Control

journal homepage: www.elsevier.com/locate/bspc



Feature selection and dimensionality reduction: An extensive comparison in hand gesture classification by sEMG in eight channels armband approach

José Jair A. Mendes Junior [a,b], Melissa L.B. Freitas [a,c], Hugo V. Siqueira [a,d], André E. Lazzaretti [a,b], Sergio F. Pichorim [a,b], Sergio L. Stevan Jr. [a,c,*]

[a] *UTFPR - Federal University of Technology, Paraná, R. Dr. Washington Subtil Chueire, 330, Ponta Grossa/PR, 84017-220, Brazil*
[b] *Graduate Program in Electrical and Computer Engineering (CPGEI) - Federal University of Technology, Paraná (UTFPR) - Curitiba, Brazil*
[c] *Graduate Program in Electric Engineering (PPGEE) - Federal University of Technology, Paraná (UTFPR) - Ponta Grossa, Brazil*
[d] *Graduate Program in Computer Science (PPGCC) - Federal University of Technology, Paraná (UTFPR) - Ponta Grossa, Brazil*

ARTICLE INFO

Article history:
Received 14 November 2019
Received in revised form 18 February 2020
Accepted 26 February 2020

Keywords:
sEMG acquisition
Gestures recognition
Armband
Dimensionality reduction
Feature selection

ABSTRACT

Gesture recognition by surface electromyography (sEMG) signals is used for several applications as prosthesis control and human-machines interfaces. One of trending approaches to sEMG acquisition is the multiple-channel armband with equidistant electrodes. Several efforts are made to improve the performance of these devices in gestures recognition applications, especially in feature selection. It is necessary choose one approach for feature selection due to the great number of electromyography features available to use. In this work, an extensive comparison of feature reduction techniques and their influences in the classification process is presented. Unlike other works, we presents the comparison between methods of feature selection and classification; this main contribution is show how the feature reduction process can aid and increase the performance of classification of sEMG in armband acquisition approach. Two general methods were employed, feature selection by wrapper forward stepwise and dimensionality reduction, resulting in eight different techniques. The following dimensionality reduction techniques were used: Principal Component Analysis, Linear Discriminant Analysis, Isomap, Manifold Charting, Autoencoder, t-distributed Stochastic Neighbor Embedding, and Large Margin Nearest Neighbor (LMNN). Seven classifiers were used, aiming at recognize six gestures acquired from an 8-channel armband of 13 subjects. An average accuracy of 89.4 % was obtained with 5 features and an Extreme Learning Machine classifier, in the feature selection approach. On other hand, considering 40 dimensions, an average accuracy of 94 % was obtained, regarding a combination between Support Vector Machine with Gaussian kernel and a LMNN technique. These results showed significant differences in statistical tests.

1. Introduction

Gesture recognition presents a growing development of applications involving interactions between man and devices. Several sensors and techniques are part of this bridge, which links human gestures and movements with machines, like cameras [1], joysticks [2], and signals from human body, as movements (e.g. inertial sensors and flexible sensors) and electrophysiological signals [3].

Among the electrophysiological signals, surface electromyography (sEMG) is highlighted in these applications. The sEMG is a relevant alternative as input signal in human-machine interfaces, because it is a method to understand the human body's behaviours, being one of the most useful electrophysiological signal [4,5]. Thus, sEMG is applied in industrial area and generic human machine interfaces. In those cases, the recognition of sEMG signals, especially hand gestures, includes interaction on software [6], videogames [7,8], robotics, prosthesis control [9,10], and even sign language recognition [11,12].

For the hand gestures recognition, the multiple channel acquisition approach is useful. Devices as armbands applied in sEMG processing aid the placement the electrodes on users [13]. Armbands do not require the specific placement of electrodes on muscles: they use the signals from muscle groups in the forearm to gesture recognition [14]. Due to their facility of place in the upper limbs, armbands are being used in human-machines interfaces, and being a new trending of sEMG acquisition.

However, this advantage costs a decrease in accuracy compared to the traditional method of specific electrode placement. Usu-

* Corresponding author.
E-mail address: sstevanjr@utfpr.edu.br (S.L. Stevan Jr.).

https://doi.org/10.1016/j.bspc.2020.101920
1746-8094/

ally, armbands integrate acquisition and processing in the same device [15]. Thus, processing techniques are explored to avoid low accuracies and improve the classification time and performance in sEMG armbands. Besides that, a high number of inputs in the pattern recognition step may hinder classifier performance on sEMG processing [50].

Furthermore, one of the important steps in sEMG processing is the feature extraction [16]. Choosing the best feature set is essential to achieve a high classifier performance and each problem has its best feature set Methodologies to feature selection in classification are needed to evaluate which feature set should be used in the specific classification process. In this work, we compare two approaches: finding the best feature set for each classifier (feature selection - FS) and using dimensionality reduction (DR) techniques before classification.

Combinatorial analysis of features is a solution for feature selection. However, due to the number of available sEMG features, the full combinatorial analysis of all features is a task that requires high computational effort. On one hand, algorithms and techniques are developed to find a solution that approximates the optimal approach, normally defined as wrappers. On the other hand, the second explored method is the data projection by dimensionality reduction. These methods transform the dataset in a new dimensional space, which could aid in the class separation with a reduced number of instances in the feature space.

The comparison between feature selection and dimensionality reduction is a relevant theme in machine learning [17–19]. However, few works explore the influences of these two techniques in sEMG signals. Liu [20] compared two techniques of RD and FS and achieved accuracies above 95 % for three classifiers. Nonetheless, this work used six features and eight channels placed on specific muscles (not in armband approach), which demands proper knowledge for the user. Alam and Arefin [21] compared feature combination and Principal Component Analysis (PCA), achieving the same range of accuracies. Even so, signals were acquired from one subject. Abbaspour et al. [22] analysed several features for decoding hand gestures and used both approaches (feature selection and dimensionality reduction), but no comparison between techniques was performed. A dimensionality reduction technique was applied after feature selection, which did not present relevant difference in classifier accuracy and a fair comparison between the approaches. Others works as compared some RD and FS techniques but outnumbered or make qualitative analysis of feature combination [16,23–26].

In that context, this work compares the two approaches (feature selection by wrapper combination and dimensionality reduction) in feature reduction for sEMG acquired from forearm with multiple channel armband. In this case, six gestures that recruits similar muscle groups were acquired to analyse the performance of the two approaches. Unlike the other presented works, feature sets and their combination with different classifiers were investigated. Seven different classifiers were used and seven dimensionality reduction techniques were applied to the feature sets. In the comparison of feature selection, a wrapper forward stepwise approach was used in the feature combination, applied for each classifier. The main contribution of this work discusses how these two approaches affect the process of sEMG classification using an armband, aiming to provide alternatives in the processing steps in these devices. Statistical tests were performed in each result. In our previous work, we used an 8-channel armband for classification of the same 6 gestures used in this work. With four features, we achieved 90 % of accuracy with an Artificial Neural Network (ANN) and none dimensionality reduction or feature selection technique was used [14].

The paper is organized as follow: Section 2 presents the materials and methods used, as the signal acquisition, feature selection and classification approaches; Section 3 shows the numerical performance achieved by the models and a discussion regarding the results; Section 4 has the main conclusions.

2. Materials and methods

The following section presents the used material for acquisition, the methodology of processing techniques for feature extraction, selection, and dimensionality reduction, and the applied classifiers. In this section, the general process of signal acquisition is presented. Fig. 1 shows the process adopted in this work to acquire and process the sEMG signals. LabVIEW™ was used in the acquisition and the processing steps were performed in Matlab®.

2.1. sEMG acquisition

The data were obtained using an 8-bipolar channel armband, with 2000 Hz sampling rate, 16-bit Analog-Digital resolution, and 10−500 Hz band-pass filter (Fig. 1a). Thus, 16 Silver/Silver Chloride (Ag/AgCl) electrodes are responsible for acquiring the signal on skin surface. The electrodes were positioned in the forearm, and a reference electrode was placed near of ankle. In all subjects, the first channel of armband was positioned on *flexor carpi ulnaris* muscles on right forearm. The other channels were positioned equidistantly on the forearm. Signals of 13 people were collected, being eight men and five women. Ethical Committee in Human Research of Federal University of Technology - Paraná CAAE 89,638,918.0.0000.5547 supported all the process.

The acquired gestures are presented in Fig. 2 with their respective signals obtained for the 8 channels for a subject. An interval of 1.33 s (equivalent to 45 beats per minute) is used. The gestures were performed in the following sequence: wrist flexion, wrist extension, wrist flexion for the right, wrist extension for the left, forearm supination, and forearm pronation. These gestures were used in other works [14,20]. They represent similar groups, enabling analyse the performance of classification process to nature and similar gestures. A rest moment preceded gesture sequences and 50 repetitions for each gesture were collected, being acquired 300 gestures by subject. In total, 13 voluntaries resulted in a dataset with 3900 gestures.

Two filters were applied to signal conditioning: a 4th-order band-pass filter, with cut-off frequencies of 10−450 Hz, and a 6th-order notch filter (60 Hz), both with Butterworth approximation. Band-pass filter was used to delimit the sEMG signal bandwidth and to attenuate the noise from movement artefact and electromagnetic interference [27,28]. The notch filter attenuates the electrical grid interference (60 Hz). Onset technique was addressed to detect the periods with muscle activation to segmentation step [29]. The time instants were detected automatically by means of threshold value set as 20 % of peak value of normalized signal to a number of samples above of 500. Signals were segmented in windows of 2000 samples, equivalent to 1 s.

2.2. Feature extraction

From the segments described in the last subsection, 30 features were extracted from sEMG signal in time and frequency domains (Fig. 1b). These features were chosen based on features used in other datasets [30–33]. Table 1 presents the 21 time-domain features, the 9 frequency-domain features, and the parameters related to their extraction. Frequency features were obtained from Fast Fourier Transform for each signal segment with 2000 data points of resolution.

Each segmented signal is represented by a 2000×8 matrix 2000 samples $\times$ 8 channels. After extracting features, this matrix is converted to a vector with 8 indexes 1 feature $\times$ 1 channel. Specific features have more than one value in their outputs, e.g. 4^{th}-order

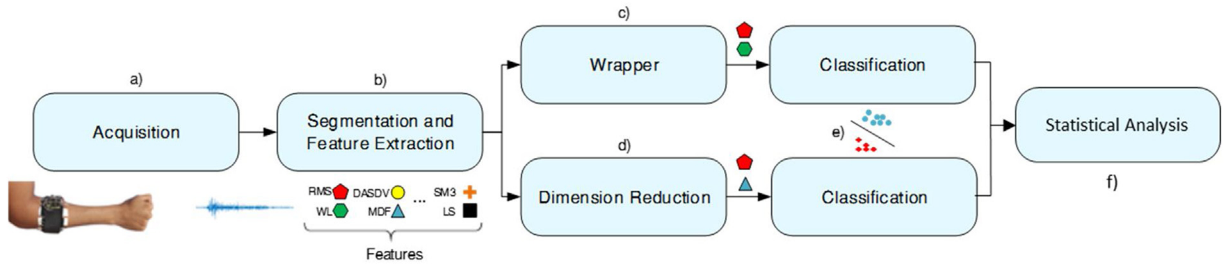


Fig. 1. Signal processing steps. a) The data are acquired by 8-channel armband. b) sEMG signals are conditioned and sent to segmentation and feature extraction. sEMG features like Root Mean Square (RMS), Waveform Length (WL), and others (defined in Section 2.2) are extracted and sorted in a vector. This vector is simultaneously sent to c) wrapper and d) dimensionality reduction steps. e) Data is classified in their respective classes and f) a statistical analysis from results is performed.

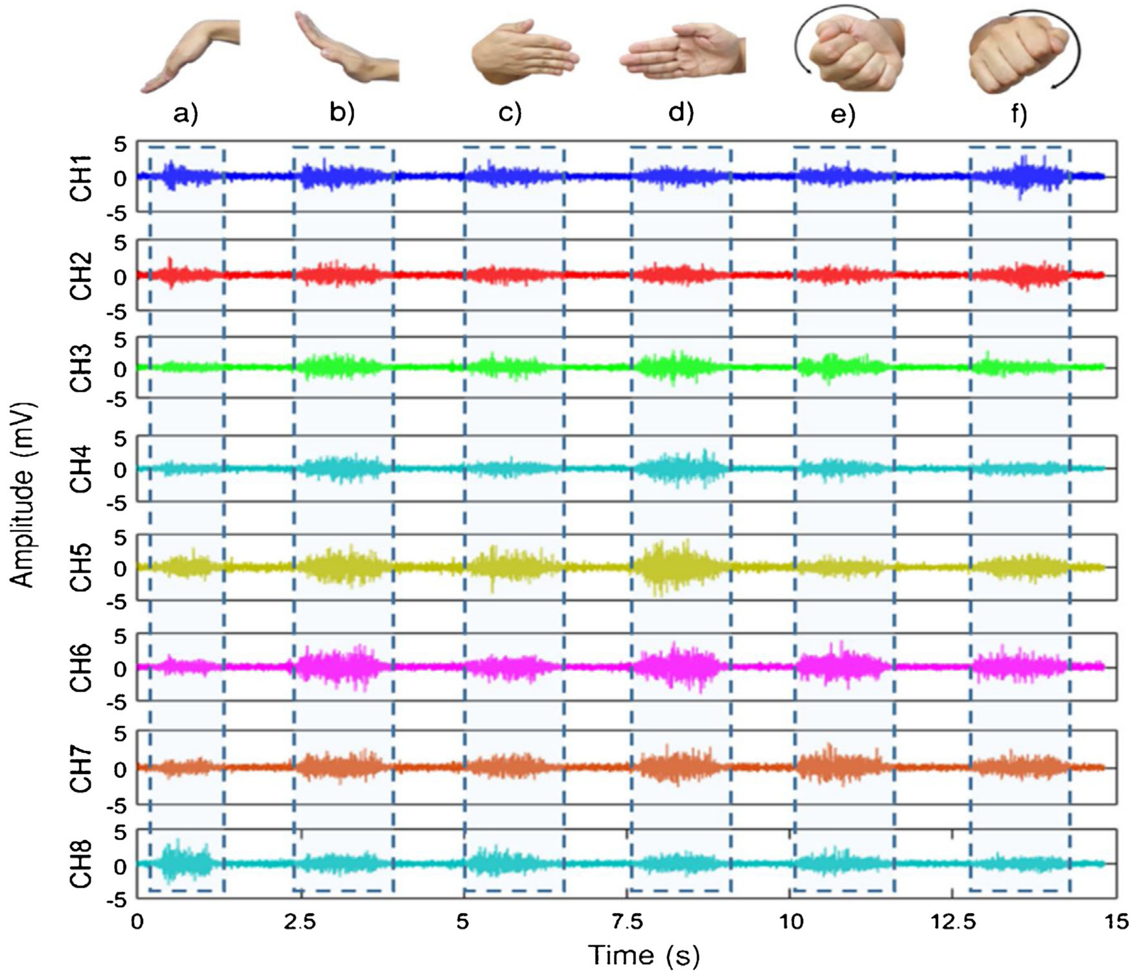


Fig. 2. Gestures developed to data acquisition and signal example acquired in each armband channel. These gestures are: a) wrist flexion, b) wrist extension, c) wrist flexion to right, d) wrist extension to left, e) forearm supination, and f) forearm pronation.

autoregressive coefficients, and 4th-order cepstral coefficients have four outputs instead of one. These features generate a vector with 32 indexes 4 coefficients × 8 channels. Therefore, the feature matrix is composed of a set with 288 attributes by 3900 samples 300 gestures × 13 subjects. Each feature vector was normalized in the range [-1, 1].

2.3. Feature selection approach (wrapper)

Finding the best feature set is a searching problem, being the wrapper forward stepwise one of the most used methods to solve it [51]. The wrapper approach (Fig. 1c), used in this work, selects sequentially the best feature set from an empty one [34]. At the first execution, the individual performance of each feature is analysed, and it is selected the one that presents the best accuracy. The next execution measures the performance for the combination among the winner previous feature with all others. Thus, the feature set grows as the combination among the attributes contributes to classification [35].

Wrapper is an alternative but it demands a large computational cost. Moreover, the final output approximates the optimal solution, and its results are directly related with the classifier used during the forward steps [34].

In this work, the wrapper algorithm was executed 30 times for each classifier. K-fold (k = 10) cross validation was used for each classifier and their respective topologies. Thus, the minimum

Table 1
sEMG features used in this work. Their names, initials, and the parameters used in their extraction are indicated.

Domain	Feature	Feature name	Parameters
Time	WL	Waveform Length	–
	SSC	Sign Slope Change	Slope threshold = 10^{-4}
	MAV	Mean Absolute Value	–
	DASDV	Difference Absolute Standard Deviation Value	–
	RMS	Root Mean Square	–
	IEMG	Integral of EMG	–
	ZC	Zero Crossing	Amplitude threshold = 10^{-2}
	SampEn	Sample Entropy	Dimension = 2 r = 0,2σ
	AR4	4th-order Autoregressive Coefficients	–
	CEPS	Cepstral Coefficients	–
	VAREMG	Variance	–
	WAMP	Williamson Amplitude	Threshold = 10^{-2}
	MAV1 and MAV2	Modified Mean Absolute Value	–
	VORDER	V-Order	3 order
	MYOP	Myopulse value	Threshold = 10^{-2}
	LOGDEC	Log Detector	–
	TM3, TM4, and TM5	Absolute Value of 3rd, 4th and 5th Moments	–
	LS	L-Scale	2 Moment
Frequency	MNP	Mean Power Spectrum	–
	MNF	Mean Frequency	–
	MDF	Median Frequency	–
	FR	Frequency Ratio	Low frequencies = 30–200 Hz High frequencies = 201–450 Hz
	PKF	Peak Frequency	–
	SM1, SM2, and SM3	Spectral Momentum	–
	TTP	Total Power Spectrum	–

Table 2
Parameterization for dimensionality reduction techniques used in this work.

Technique	Parameter
PCA	N° dim = 30
LDA	N° dim = 20
Autoencoder	N° dim=60 1 to 60 neurons Cross validation for each topology (k-fold, k = 10)
LMNN	N° dim=60 40 nearest neighbours
Isomap	N° dim = 30 12 nearest neighbours
Manifold Charting	N° dim = 20 2 neighbours
t-SNE	N° dim = 40 Perplexity = 40

number of features with best performances was analysed, and then their respective frequency of occurrence is obtained. The features with more repetition during this process were verified. For example, if a classifier presented the maximum accuracy with five features, the best set is composed of the most five repeated features in the iterations.

2.4. Dimensionality reduction approach

Seven dimensionality reduction techniques were chosen for this work (Fig. 1d). Table 2 presents them, together with the best parameterization values found for this application.

The first applied technique was the classical Principal Component Analysis (PCA), an unsupervised statistical approach to transform data with high dimensionality in a smaller set of variables that represents the samples in another space [36]. PCA results in a new set where the dimensions are orthogonal and the axis are sorted according to the variance of data along them in decreasing order. The axis are named as Principal Components and the new dimension set is formed by the axis with more variance [37]. Even though being one of classical techniques used in dimensionality reduction, PCA cannot be a useful tool when there are a very high dimensional space to compute the eigenvectors and the variance of attributes is not a discriminant factor among the classes, besides discarding the information of the labels of each class [38].

Linear Discriminant Analysis (LDA) is a supervised technique to find a sub-space that supports the separability among classes, aiming to reducing its dimensionality. It is based on Fisher's linear discriminant, being its goal to maximize the between-class variance whilst to minimize the within-class variance [39]. LDA can be used as both dimensionality reduction technique and classifier. However, LDA could not be accurate in class discrimination, especially when the samples are in the hyper-plane of separation [40].

Autoencoder model is a feedforward artificial neural network (ANN) used to reconstruct signals. The signal inserted into the input is reconstructed in the output, using the synaptic weights and hidden neurons [41]. For example, considering an ANN with three layers (input, hidden, and output layers), n is the number of neurons in hidden layer and m is the number of neurons in input and output layers (they are equivalent). Aiming dimensionality reduction, m should be greater than n. In this case, there are two kinds of synaptic weights: $\mathbf{W}_1$ (between input and hidden layer) and $\mathbf{W}_2$ (between hidden and output layer). When the neural network training finishes using backpropagation with the reconstruction error (L2 norm) used as loss function, $\mathbf{W}_2$ is discarded. Thus, the n-dimensional representation refers to the data previously presented with low dimensionality. Autoencoders allow several non-linear activation functions, and consequently, several mapping can be constructed [38]. Nonetheless, the main limitation of autoencoders is their training time, which can be slow, since it depends on the topology and the convergence of the backpropagation.

Large Margin Nearest Neighbour (LMNN) was initially proposed to improve the performance of classifiers based on distances as k-nearest-neighbours [42]. LMNN searches a metric more appropriate to the data distribution, which is based on variations of the Mahalanobis distance. In this technique, a sample $\mathbf{x}$ has neighbours with the same class, named target neighbours, whilst the samples with different classes are called impostors. LMNN minimizes distances of target neighbours and maximizes distances of the impostors [43]. This process is made by minimization of a cost function, but it is sensitive to overfitting and Euclidean distance performance [44].

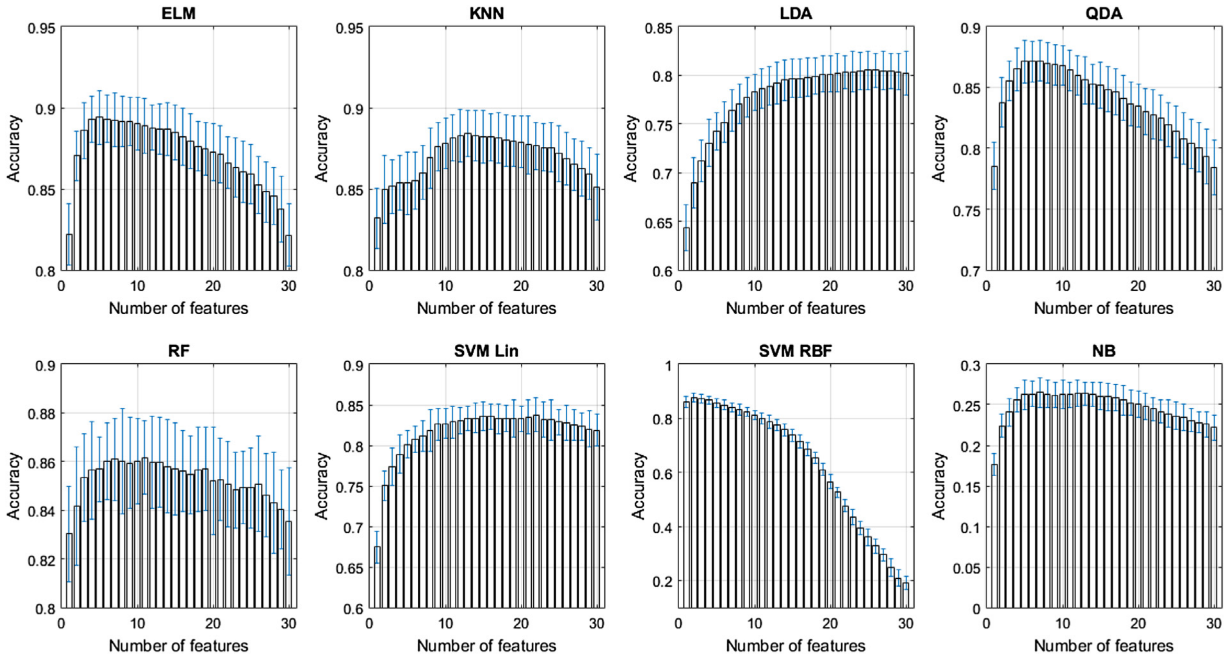


Fig. 3. Accuracies and standard deviations for the classifiers in wrapper feature selection method.

The next technique, Isomap (Isometric feature mapping), is based on classical scaling, which changes the distribution of calculated Euclidian distance to geodesic distance [45]. This process is indicated when the data distribution has curved characteristics, as biomedical signal (e. g. sEMG) [38]. Geodesic distances are calculated from a graph, in which each vertex is connected to its nearest neighbour. The closest path between two vertexes is the geodesic distance, being built a matrix for every sample [45]. This scaling is applied to calculate a low dimensionality space. Even though it is recommended for biomedical signals, Isomap generates the output without an explicit mapping function (as PCA), and is highly dependent on the data model. Moreover, it presents another limitation: wrong links on graph construction can degrade the performance of the model [38].

Another analysis with high dimension data is the manifold charting. Manifolds are based on data partition with neighbours with local linear low dimensionality [38]. This process is based on minimizing a convex cost function, calculating eigenvalues and eigenvectors. Even though it is less susceptive to noise than Isomap, Manifold Charting is highly dependent of resultant model.

At least, the t-distributed Stochastic Neighbour Embedding (t-SNE) technique for dimensionality reduction was used. The t-SNE separates high dimensionality data in set of manifolds that have relation between them, retaining their properties in low dimension [46]. This method constructs a probability distribution to represent the probability of similar and dissimilar points being picked. In other words, similar points have a high probability of being picked, whilst the dissimilar ones have a small probability of being selected, both on high and low-dimensional spaces. To do so, t-SNE minimizes the Kullback–Leibler divergence between the two distributions on high and low dimensional spaces [46]. In this work, the number of dimensions of t-SNE was extrapolated to verify its performance as dimensionality reducer, even knowing that it converges up to three dimensions.

2.5. Classifiers

Different classifiers were selected to perform the pattern recognition stage (Fig. 1e), representing several natures: probabilistic (Naïve Bayes, NB), distances and neighbourhood (k-Nearest Neighbour, KNN), artificial neural networks (Extreme Learning Machine, ELM), decision and ensemble (Random Forest, RF), discriminants (Linear and Quadratic Discriminant Analysis, LDA and QDA), and Support Vector Machines (Linear and Non-Linear with Gaussian Kernel type Radial Basis Function, SVM Lin and SVM RBF). All classifiers were built in Matlab® software, and SVMs were based on using the library LibSVM [47]. Table 3 presents the parameters used for each classifier, selected after individual analysis with a cross-validation procedure with the k-fold method ($k = 10$). From these results, statistical analysis were performed to evaluate the obtained data (Fig. 1f).

Table 3
Parameterization for classifiers.

Classifier	Parameter
ELM	800 neurons in the hidden layer Hyperbolical tangent as activation function
KNN	1 nearest neighbour
LDA	–
QDA	–
NB	Gaussian Distribution
SVM Lin	$C = 7$
SVM RBF	$C = 10$ $\gamma = 0.67$
RF	30 trees

3. Results and discussion

The results obtained for feature selection and dimensionality reduction are presented with the comparison obtained for each approach.

3.1. Results from feature selection

Fig. 3 presents the relation between the accuracies and number of features for the classifiers. As one can observe, increasing the number of features, the classifiers present two tendencies: accuracy decreasing (ELM, QDA, and SVM RBF) or stabilized in a determinate value without significantly variation (other techniques). Table 4

Table 4
Relation of number of features obtained in the classifiers. Also, the set of more repeated feature in each classifier and the best obtained accuracy are presented.

Classifier	Number of Features	Attributes in classifier input	Feature Set	Best Accuracy
ELM	5	40	WAMP,MYOP, WL, SM1 and SM2	89.4 %
KNN	13	104	WL, DASDV, MYOP, SM1, SM2, SSC, MAV, ZC, SampEn, WAMP, MAV1, LOGDEC, and IEMG	88.4 %
SVM RBF	3	24	DASDV, WL, and LS	87.7 %
QDA	7	88	SampEn, AR4, WAMP, MYOP, SM2, ZC, and LOGDEC	87.2 %
RF	7	56	WAMP, SM1, SM2, WL, TM5, DASDV, and SSC	86.1 %
SVM Lin	13	128	SSC, RMS, AR4, VAREMG, WAMP, MYOP, SM1, MAV, DASDV, CEPS, MAV2, MAV1, and WL	83.3 %
LDA	15	168	ZC, SampEn, AR4, CEPS, MYOP, WL, MAV2, SSC, DASDV, VAREMG, WAMP, RMS, IEMG, MAV1, and LOGDEC	80.1 %
NB	5	40	DASDV, WL, MAV, IEMG, and VAREMG	26.3 %

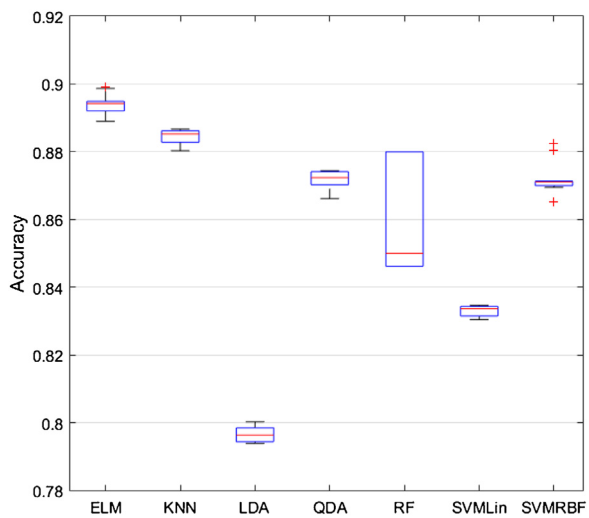


Fig. 4. Boxplot of distributions of best results in wrapper approach.

summarizes the obtained results and the more repeated feature in each case for the 30 wrapper execution.

Among the techniques, NB presented the worst performance in this application with maximum accuracy of 26 % with five features. In the NB formulation, all the features are supposed to be independent among them; however, it is not often true due to the contribution that the sEMG channels affect each other. Other techniques have accuracies up to 80 %. The best classifiers, that combine high accuracies and few features, are ELM (89 % of accuracy and 5 features) and SVM RBF (87 % and 3 features). Indeed, KNN classifier achieved 88 % of accuracy, but 13 features were necessary (104 input attributes for classifier input), which do not represent relevant gain in classification process. The other classifiers do not present relevant accuracy for the same number of features.

Fig. 4 shows the dispersion of results based on the boxplot graphic for the relevant classifiers. The classifier with low accuracies (excepting NB) on wrapper approach was LDA, close to 80 %. Even RF had 86.1 % of accuracy, RF presented high dispersion compared with the other classifiers (between 88 % and 84.5 %). The other classifiers presented low dispersion.

The Friedman test was used to analyse if the null hypothesis that there is not difference between the best response for the classification techniques with interval of confidence of 5% (p-value<0.05). The null hypothesis was rejected, since the obtained p-value was $8.2\ 10^{-14}$. Posteriorly, Tukey *post hoc* test was performed to evaluate the groups that have relevant differences among them. Table 5 presents these results and the groups with significant differences are highlighted. Significant differences were not found among the three best cases (ELM, KNN, and SVM RBF). Thus, the use of SVM RBF with 3 features is not statistically different from ELM or KNN with

Table 5
Tukey *post hoc* for Friedman test to significant difference analysis for the outputs obtained in the classifiers by wrapper approach. Highlighted values show significant difference values between the techniques.

	KNN	LDA	QDA	RF	SVM Lin	SVM RBF
ELM	0.974	**0.000**	0.172	0.053	**0.000**	0.172
KNN	-	**0.000**	0.779	0.476	**0.009**	0.779
LDA	-	-	0.068	0.211	0.985	0.068
QDA	-	-	-	1	0.476	1
RF	-	-	-	-	0.779	1
SVM Lin	-	-	-	-	-	0.476

13 features. However, ELM presents as an advantage in the training time: average of 0.75 s for ELM versus 2.1 s for SVM RBF (using a Dual Core i5-4210U processor). ELM parameterization requires the number of neurons in hidden layer whilst in SVM RBF requires parameter C, Gaussian size (γ), and error tolerance.

The features that are more frequent are shown in Fig. 5 (see Table 1). From the histogram (Fig. 5a), it can be noted that the time-domain features have more frequency of occurrence. SM1 and SM2 are the frequency-domain features that contribute positively for classification process, being appropriated for this kind of application. The features WL, DASD, WAMP, and MYOP were the most used; thus, they were addressed in the classifiers another time.

Fig. 5b presents the results with these four features (WL, DASDV, WAMP, and MYOP) for all the classifiers. It is noticeable that the accuracies shown less alteration in classification performance for the most repeated features in the previous test. Considering classifiers, the winner continues being ELM, which presents 89 % of success rate. In addition, classifier NB shows a considerable increasing (from 22.6–45%) in accuracy, in other words, wrapper approach cannot be a good indicative to choose the sEMG features for this classifier.

3.2. Results for dimensionality reduction

Fig. 6 presents the results for dimensionality reduction techniques and Table 6 shows the best accuracies regarding the

Table 6
Relation between the dimensionality reduction technique and best classifier obtained by its accuracy value. The number of dimensions is the number of attributes in the classifier inputs - ; it does not represent significant change in the accuracy.

Dimensionality Reduction Technique	Classifier	Number of Dimensions	Accuracy
LMNN	SVM RBF	40	94 %
Autoencoder	SVM RBF	18	89.9 %
PCA	ELM	21	89.3 %
Manifold Charting	ELM	16	88.5 %
t-SNE	ELM	10	87.7 %
LDA	LDA	12	86.6 %
Isomap	SVM RBF	16	83.1 %

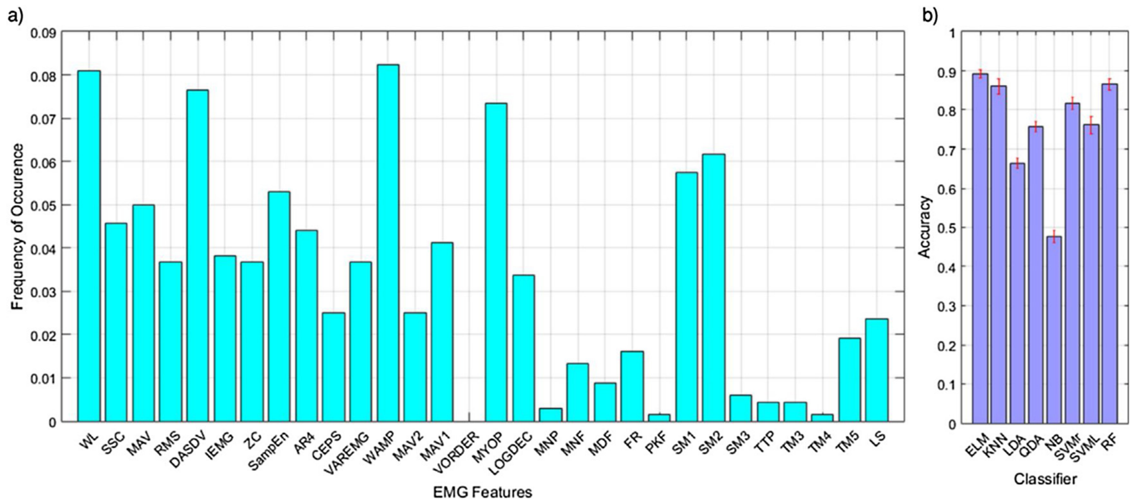


Fig. 5. Analysis made by the more repeated features in wrapper method: a) presents the frequency of occurrence for the features in all classifiers and b) presents the accuracy with more frequency of occurrence (WL, RMS, WAMP and MYOP). Even with one less feature, ELM presents the best accuracy.

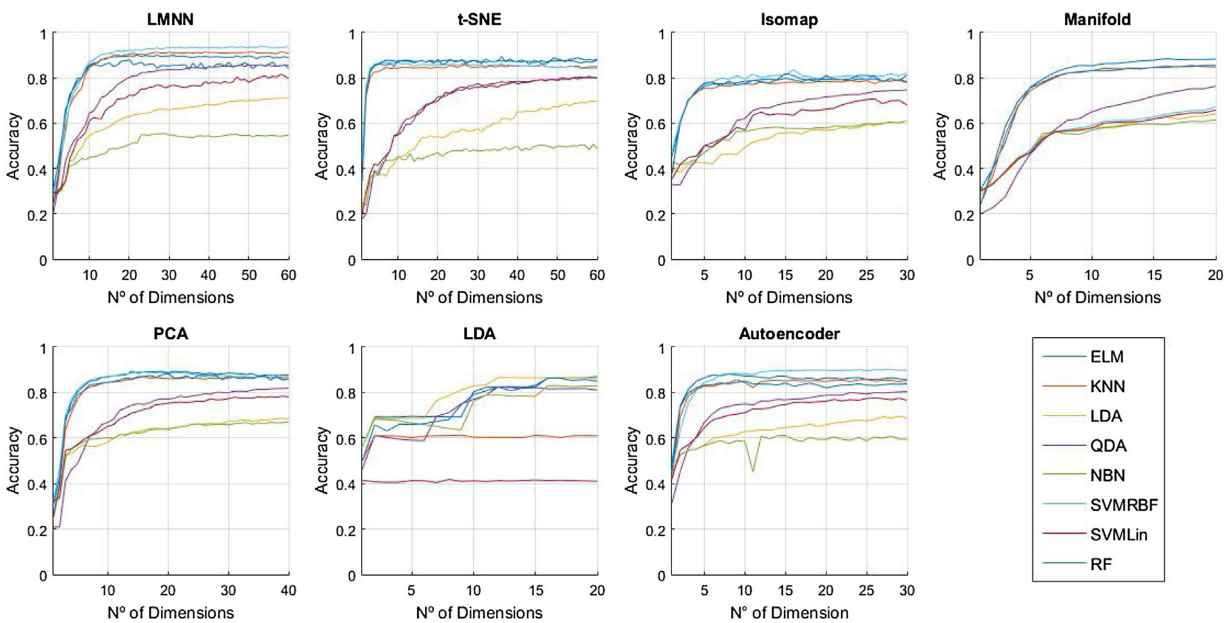


Fig. 6. Accuracies obtained combining each classifier with the dimensionality reduction techniques. It can be noted that the accuracy increases with the number of dimensions in each technique. However, this gain reaches a certain value and above this value, it cannot represent significant change.

technique/classifier. Among all techniques, the best accuracy was obtained combining LMNN with SVM RBF, reaching 94 % with 40 dimensions. Classifiers as ELM, KNN, and RF presented high performances using LMNN. This is due to the ability of non-linear mapping with proximity in LMNN that there is a combination between its behaviour and the KNN classifier [42]. KNN shows satisfactory results (previously showed on wrapper approach). However, LMNN is able to unite close and near samples, and it is improved with high dimensionality mapping made by SVM RBF for the sEMG signals. Excepting Isomap, other techniques present similar results with success rates above 80 %.

Regarding to classifiers, ELM and SVM RBF have similar performances. Regarding number of dimensions, it can be noticeable that 20 or less attributes can reach the maximum accuracies in each combination. An increase in the dimensions causes stagnation in accuracy, in other words, it does not represent significant differences in success rate.

The dispersions for the best results are presented on Fig. 7. Even though the combination of LMNN and SVM RBF shows best performance, it is necessary to evaluate if these results are statistically distinct. The Friedman test was used again, and the null hypothesis was rejected, showing that the methods had significant differences (p-value = $7.12\,10^{-9}$). After that, Tukey *post hoc* was applied to evaluate the differences among the techniques. Table 7 shows these results and the distributions that have significant differences is highlighted. The best combination (LMNN and SVMRBF) showed independence for the following combinations: t-SNE and ELM, Isomap and SVM RBF, and LDA and LDA. Moreover, Autoencoder and tSNE presented combinations significant similar than the others.

3.3. Classification among classes

Tables 8 and 9 present the confusion matrices for the best cases obtained by the wrapper and dimensionality reduction methods,

Table 7
Tukey *post hoc* for Friedman test to significant difference analysis for the outputs obtained in the classifiers by the best combinations in classifiers and dimensionality reduction approach. Highlighted values show significant difference values between the techniques.

	t-SNE + ELM	Isomap + SVM RBF	Manifold + ELM	PCA + ELM	LDA + LDA	Autoencoder + SVM RBF
LMNN + SVM RBF	**0.001**	**0.000**	0.064	0.468	**0.000**	0.740
TNSE + ELM	–	0.513	0.998	0.697	1.000	0.424
Isomap + SVM RBF	–	–	0.023	**0.001**	0.819	**0.000**
Manifold + ELM	–	–	–	1.000	0.950	0.999
PCA + ELM	–	–	–	–	0.382	1.000
LDA + LDA	–	–	–	–	–	0.183

Table 8
Confusion Matrix for the winner case in wrapper: ELM with 5 features. Null cells represent 0%.

Gesture		Target					
		a	b	c	d	e	f
Output	**a**	90.8 %		6.2 %	1.5 %		1.5 %
	b		87.7 %		12.3 %		
	c	6.2 %		92.3 %			1.5 %
	d		13.8 %	1.5 %	84.6 %		
	e		1.5 %	1.5 %	1.5 %	92.3 %	3.1 %
	f					3.1 %	96.9 %

Table 9
Confusion Matrix for the winner case in dimensionality reduction: LMNN and SVM RBF combination. Null cells represent 0%.

Gesture		Target					
		a	b	c	d	e	f
Output	**a**	93.8 %		3.1 %			3.1 %
	b		93.8 %		6.2 %		
	c	3.1 %		96.9 %			
	d		6.2 %		92.3 %	1.5 %	
	e			1.5 %		93.8 %	4.6 %
	f					1.5 %	98.5 %

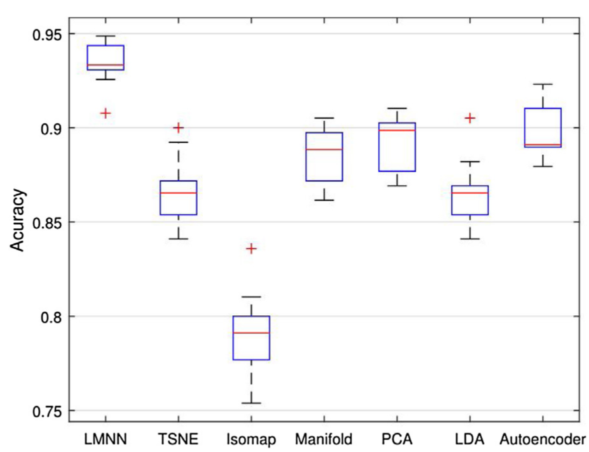


Fig. 7. Distribution of best results obtained for each dimensionality reduction technique.

respectively. As aforementioned, the gesture has similar muscle group recruitment (especially the pair a–c, b–d, and e–f seen in Fig. 2) to focusing the performance of classification processes. From confusion matrices, it can be noticeable that classes have more errors between their pairs. Thus, this misclassification is less for the dimensionality dimension approach. The gesture with the highest success rate is forearm supination (f) and the one with the highest error was the wrist extension for left (d). It occurred in both approaches.

Considering the other cited works, Table 10 presents the comparison between the results. Other investigations do not use the armband approach to compare the feature reduction methods; and when there is the application of two techniques, it is not always a comparative analysis [22]. Regarding the most similar work [20], the authors used 8 channels of specific electrode placement for well separable gestures, whilst we achieved 90 % and 94 % of accuracy in the best cases for similar gestures. However, using armband approach decrease between 5 and 10 % the success rate whilst ease the placement of channels in the subjects. It is noted that the number of techniques compared in this study, especially in dimensionality reduction, is higher than the others. The results considering the LMNN and SVM RBF seems to be a good combination for gestures classification. Comparing the results with our previously work [14], we obtained a gain both in terms of number of features and accuracy, because none feature reduction or selection technique were applied.

4. Conclusion

The comparison of feature reduction for sEMG classification using armband approach was proposed in this work. Six hand gestures were acquired from 13 subjects, feature selection, and dimensionality reduction techniques approaches were compared for seven different classifiers. Unlike the similar works previously mentioned, this work presented an extensive combination between classification and feature selection techniques for hand gestures recognition. Besides that, the combination between the methods (specially dimensionality reduction and classification) increasing the accuracy for the same dataset using armband approach.

By wrapper method, accuracies about 87 and 90 % were obtained using 5 and 3 features for ELM and SVM RBF classifiers, respectively. On the other hand, dimensionality reduction presented 94 % of accuracy by combination of LMNN and SVM with Gaussian kernel. Moreover, the other techniques present sim-

Table 10
Related works with analysis of dimensionality reduction and feature selection approach in sEMG classification for gesture recognition.

Author	Acquisition Mode	Gestures	Features	Techniques	Classifier	Accuracy
Camacho Navarro et al. [48]	2 channels / 5 subjects	6	Discrete Wavelet Transform	RD: DTW and PCA	SVM	≈ 95 %
Liu [20]	8 channels / 20 subjects	7	6	RD: PCA ULDA (uncorrelated linear discriminant analysis) Feature Selection: MRF (Markov Random Field) FOS-MOD (forward orthogonal search selector)	KNN, LDA, SVM	FS and RD: Above 95 %
Phinyomark et al. [16]	5 channels / 20 subjects	6	37	FS: Feature selection and combination	LDA	Relevance Analysis
Zhang et al. [26]	4 channels / 1 subject	9	14	RD: LDA and PCA	Minimum distance classifier (MDC)	97.5 %
Geng et al. [23]	56 channels / 8 subjects	22	4	RD: Uncorrelated Principal Component Analysis (uPCA) and Common Spatial Filter (CSP)	LDA	94.3 %
Adewuyi et al. [49]	4 channels / 20 subjects	4	25	FS: wrapper	LDA, QDA, and Multi-Layer Perceptron (MLP)	Relevance Analysis
Alam and Arefin [21]	8 channels (7 on forearm and 1 in biceps) / 1 subject	7	7	RD: PCA FS: Feature combination	LDA	RD: Above 95 % FS: 97.5 %
Too et al. [25]	Nina Pro database (DB7) / 2 subjects	17	3	FS: Competitive Binary Grey Wolf Optimization and Modified Binary Tree Growth Algorithm	KNN	Above 85 %
Abbaspour et al., [22]	4 channels / 20 subjects	11	44	FS: Feature Selection and Combination RD: PCA (after Feature Selection)	LDA, KNN, DT, Maximum Likelihood Estimation, SVM, and MLP	Above 97 %
Rabin et al., [24]	2 channels / 5 subjects	6	STFT	RD: PCA and diffusion maps	KNN	Multi-subject: ≈ 78 %
La Banca Freitas et al. [14]	8-channel armband / 3 subjects	6	4	–	KNN, DT, QDA, SVM, NB, SVM	90 %
This work	8-channel armband / 13 subjects	6	31	FS: wrapper forward stepwise RD: PCA, LDA, Autoenconder, Manifold Charting, Isomap, LMNN, t-SNE	KNN, ELM, LDA, QDA, NB, SVM (linear and non-linear by Radial Basis Function), RF	FS: 89.4 % RD: 94 %

ilar results above 80 % with a reduced number of dimensions. Statistical analysis showed that these results have significant differences.

At least, it can be noted that the feature reduction increases the performance of hand gesture recognition problem by sEMG armband acquisition. Nevertheless, it demands some computational effort to find the best feature set for this problem.

Future works in this area are focuses in increases the number of gestures with similar groups and the increase of wrapper methods, especially the algorithms based on optimization as genetic and particle swarm optimization.

CRediT authorship contribution statement

José Jair A. Mendes Junior: Conceptualization, Formal analysis, Writing - original draft. **Melissa L.B. Freitas:** Methodology. **Hugo V. Siqueira:** Software, Validation. **André E. Lazzaretti:** Validation, Writing - review & editing. **Sergio F. Pichorim:** Writing - review & editing, Supervision. **Sergio L. Stevan:** Conceptualization, Supervision, Project administration.

Acknowledgement

This study was financed in part by the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (CAPES) - Finance Code 001.

Declaration of Competing Interest

None.

References

[1] Y. Wu, T.S. Huang, Vision-based gesture recognition: a review, in: A. Braffort, R. Gherbi, S. Gibet, D. Teil, J. Richardson (Eds.), Gesture-Based Communication in Human-Computer Interaction, Lecture Notes in Computer Science, Springer, Berlin Heidelberg, 1999, pp. 103–115.
[2] A. Takemoto, K. yano, T. Miyoshi, K. Terashima, Operation assist control system of rotary crane using proposed haptic joystick as man-machine interface, in: RO-MAN 2004. 13th IEEE International Workshop on Robot and Human Interactive Communication (IEEE Catalog No.04TH8759). Presented at the RO-MAN 2004, 13th IEEE International Workshop on Robot and Human Interactive Communication (IEEE Catalog No.04TH8759), 2004, pp. 533–538, http://dx.doi.org/10.1109/ROMAN.2004.1374817.
[3] J. Ribeiro, F. Mota, T. Cavalcante, I. Nogueira, V. Gondim, V. Albuquerque, A. Alexandria, Analysis of man-machine interfaces in Upper-Limb prosthesis: a review, Robotics 8 (2019) 16, http://dx.doi.org/10.3390/robotics8010016.
[4] R.H. Chowdhury, M.B.I. Reaz, M.A.B.M. Ali, A.A.A. Bakar, K. Chellappan, T.G. Chang, Surface electromyography signal processing and classification techniques, Sensors 13 (2013) 12431–12466, http://dx.doi.org/10.3390/s130912431.
[5] D. Farina, M. Sartori, Surface Electromyography for MAN-Machine Interfacing in Rehabilitation Technologies [WWW Document]., 2016, http://dx.doi.org/10.1002/9781119082934.ch20.
[6] J.E.E. Goh, M.L.I. Goh, J.S. Estrada, N.C. Lindog, J.C.M. Tabulog, N.E.C. Talavera, Presentation-aid armband with IMU, EMG sensor and bluetooth for free-hand writing and hand gesture recognition, Int. J. Comput. Sci. Res. 1 (2018) 65–77.
[7] T.V.V. Batista, L.S. Machado, A.M.G. Valença, Surface electromyography for game-based hand motor rehabilitation, in: 2016 XVIII symposium on virtual and augmented reality (SVR), Presented at the 2016 XVIII Symposium on Virtual and Augmented Reality (SVR) (2016) 140–144, http://dx.doi.org/10.1109/SVR.2016.32.
[8] X. Zhang, X. Chen, W. Wang, J. Yang, V. Lantz, K. Wang, Hand gesture recognition and virtual game control based on 3D accelerometer and EMG sensors, in: Proceedingsc of the 13th International Conference on Intelligent User Interfaces - IUI' 09, Presented at the Proceedingsc of the 13th International Conference, ACM Press, Sanibel Island, Florida, USA, 2008, p. 401, http://dx.doi.org/10.1145/1502650.1502708.
[9] S. Bisi, L. De Luca, B. Shrestha, Z. Yang, V. Gandhi, Development of an EMG-Controlled mobile robot, Robotics 7 (2018) 36, http://dx.doi.org/10.3390/robotics7030036.
[10] D. Yang, J. Zhao, Y. Gu, L. Jiang, H. Liu, EMG pattern recognition and grasping force estimation: improvement to the myocontrol of multi-DOF prosthetic hands, in: 2009 IEEE/RSJ International Conference on Intelligent Robots and Systems, Presented at the 2009 IEEE/RSJ International Conference on Intelligent Robots and Systems, 2009, pp. 516–521, http://dx.doi.org/10.1109/IROS.2009.5354544.
[11] V.E. Kosmidou, L.J. Hadjileontiadis, Sign language recognition using intrinsic-mode sample entropy on sEMG and accelerometer data, IEEE Trans. Biomed. Eng. 56 (2009) 2879–2890, http://dx.doi.org/10.1109/TBME.2009.2013200.
[12] J.J.A. Mendes Junior, M.L.B. Freitas, S.L. Stevan, S.F. Pichorim, Recognition of libras static alphabet with MyoTM and multi-layer perceptron, in: R. Costa-Felix, J.C. Machado, A.V. Alvarenga (Eds.), XXVI Brazilian Congress on Biomedical Engineering, IFMBE Proceedings. Springer Singapore, 2019, pp. 413–419.
[13] E. Costanza, A. Perdomo, S.A. Inverso, R. Allen, EMG as a subtle input interface for Mobile computing, in: S. Brewster, M. Dunlop (Eds.), Mobile Human-Computer Interaction - MobileHCI 2004, Lecture Notes in Computer Science, Springer, Berlin Heidelberg, 2004, pp. 426–430, http://dx.doi.org/10.1007/978-3-540-28637-0_50.
[14] M. La Banca Freitas, J.J.A. Mendes, D.P. Campos, S.L. Stevan, Hand gestures classification using multichannel sEMG armband, in: R. Costa-Felix, J.C. Machado, A.V. Alvarenga (Eds.), XXVI Brazilian Congress on Biomedical Engineering, IFMBE Proceedings., Springer Singapore, 2019, pp. 239–246.
[15] M. Hakonen, H. Piitulainen, A. Visala, Current state of digital signal processing in myoelectric interfaces and related applications, Biomed. Signal Process. Control 18 (2015) 334–359, http://dx.doi.org/10.1016/j.bspc.2015.02.009.
[16] A. Phinyomark, F. Quaine, S. Charbonnier, C. Serviere, F. Tarpin-Bernard, Y. Laurillau, EMG feature evaluation for improving myoelectric pattern recognition robustness, Expert Syst. Appl. 40 (2013) 4832–4840, http://dx.doi.org/10.1016/j.eswa.2013.02.023.
[17] T. Lan, D. Erdogmus, L. Black, J. Van Santen, A comparison of different dimensionality reduction and feature selection methods for single trial ERP detection, Conf. Proc. Annu. Int. Conf. IEEE Eng. Med. Biol. Soc. IEEE Eng. Med. Biol. Soc. Annu. Conf. 2010 (2010) 6329–6332, http://dx.doi.org/10.1109/IEMBS.2010.5627642.
[18] W. Nick, J. Shelton, G. Bullock, A. Esterline, K. Asamene, Comparing dimensionality reduction techniques, in: SoutheastCon 2015, 2015, pp. 1–2, http://dx.doi.org/10.1109/SECON.2015.7132997, Presented at the SoutheastCon 2015.
[19] R. Ramlee, A.K. Muda, N.A. Emran, Comparison of feature dimension reduction approach for writer verification, Proc. First Int. Conf. Adv. Data Inf. Eng. DaEng-2013 (2014) 95–102, http://dx.doi.org/10.1007/978-981-4585-18-7_11.
[20] J. Liu, Feature dimensionality reduction for myoelectric pattern recognition: a comparison study of feature selection and feature projection methods, Med. Eng. Phys. 36 (2014) 1716–1720, http://dx.doi.org/10.1016/j.medengphy.2014.09.011.
[21] M.S. Alam, A.S. Arefin, Real-time classification of multi-channel forearm EMG to recognize hand movements using effective feature combination and LDA classifier, Bangladesh J. Med. Phys. 10 (2017) 25–39, http://dx.doi.org/10.3329/bjmp.v10i1.39148.
[22] S. Abbaspour, M. Lindén, H. Gholamhosseini, A. Naber, M. Ortiz-Catalan, Evaluation of surface EMG-based recognition algorithms for decoding hand movements, Med. Biol. Eng. Comput. (2019).
[23] Y. Geng, X. Kuang, M. Zhu, Y. Zhang, G. Li, Y.-T. Zhang, Exploration of data dimensionality reduction methods for improving classification performance of voluntary movements, in: Y.-T. Zhang (Ed.), The International Conference on Health Informatics, IFMBE Proceedings, Springer International Publishing, 2014, pp. 126–129.
[24] N. Rabin, M. Kahlon, A. SaritMalayev Ratnovsky, Classification of human hand movements based on EMG signals using nonlinear dimensionality reduction and data fusion techniques, Expert Syst. Appl. (2020) 113281, http://dx.doi.org/10.1016/j.eswa.2020.113281.
[25] J. Too, A.R. Abdullah, N.M. Saad, Study of EMG feature selection for hand motions classification, Int. J. Hum. Technol. Interact. IJHaTI 3 (2019), 19-24–24.
[26] D. Zhang, X. Zhao, J. Han, Y. Zhao, A comparative study on PCA and LDA based EMG pattern recognition for anthropomorphic robotic hand, in: 2014 IEEE International Conference on Robotics and Automation (ICRA), Presented at the 2014 IEEE International Conference on Robotics and Automation (ICRA), 2014, pp. 4850–4855, http://dx.doi.org/10.1109/ICRA.2014.6907569.
[27] R. Merletti, P.J. Parker, Electromyography: Physiology, Engineering, and Non-Invasive Applications, John Wiley & Sons, 2004.
[28] S.L. Stevan Jr, JoséJair Alves Mendes Jr, F.C. Janzen, Eloi Agostini Jr, S. Okida, Circuit design for surface electromyography data acquisition, IEEE Lat. Am. Trans. 13 (2015) 3193–3200, http://dx.doi.org/10.1109/TLA.2015.7387221.
[29] G. Kamen, D. Gabriel, Essentials of Electromyography, Human Kinetics, 2010.
[30] B. Hudgins, P. Parker, R.N. Scott, A new strategy for multifunction myoelectric control, IEEE Trans. Biomed. Eng. 40 (1993) 82–94, http://dx.doi.org/10.1109/10.204774.
[31] N. Nazmi, M.A. Abdul Rahman, S.-I. Yamamoto, S.A. Ahmad, H. Zamzuri, S.A. Mazlan, A review of classification techniques of EMG signals during isotonic and isometric contractions, Sensors 16 (2016) 1304, http://dx.doi.org/10.3390/s16081304.
[32] A. Phinyomark, N. Khushaba R, E. Scheme, Feature extraction and selection for myoelectric control based on wearable EMG sensors, Sensors 18 (2018) 1615, http://dx.doi.org/10.3390/s18051615.

[33] A. Phinyomark, P. Phukpattaranont, C. Limsakul, Feature reduction and selection for EMG signal classification, Expert Syst. Appl. 39 (2012) 7420–7431, http://dx.doi.org/10.1016/j.eswa.2012.01.102.
[34] N.E. Aboudi, L. Benhlima, Review on wrapper feature selection approaches, in: 2016 International Conference on engineering MIS (ICEMIS), PreSented at the 2016 International Conference on Engineering MIS (ICEMIS) (2016) 1–5, http://dx.doi.org/10.1109/ICEMIS.2016.7745366.
[35] R. Panthong, A. Srivihok, Wrapper feature subset selection for dimension reduction based on ensemble learning algorithm, Procedia Comput. Sci., The Third Information Systems International Conference 2015 72 (2015) 162–169, http://dx.doi.org/10.1016/j.procs.2015.12.117.
[36] G.H. Dunteman, Principal Components Analysis, SAGE, 1989.
[37] S. Theodoridis, A. Pikrakis, K. Koutroumbas, D. Cavouras, Introduction to Pattern Recognition: A Matlab Approach, 1 edition, Academic Press, Burlington, MA, 2010.
[38] L.V.D. Maaten, E. Postma, J.V.D. Herik, Dimensionality Reduction: a Comparative Review, 2009.
[39] A. Tharwat, T. Gaber, A. Ibrahim, A.E. Hassanien, Linear discriminant analysis: a detailed tutorial, J. Eng. Sustain. Community Dev. 30 (2017) 169–190, http://dx.doi.org/10.3233/AIC-170729.
[40] S. Shinmura, Four serious problems and New facts of the discriminant analysis, in: E. Pinson, F. Valente, B. Vitoriano (Eds.), Operations Research and Enterprise Systems, Communications in Computer and Information Science., Springer International Publishing, 2015, pp. 15–30.
[41] G.E. Hinton, R.R. Salakhutdinov, Reducing the dimensionality of data with neural networks, Science 313 (2006) 504–507, http://dx.doi.org/10.1126/science.1127647.
[42] K.Q. Weinberger, L.K. Saul, Distance metric learning for large margin nearest neighbor classification, J. Mach. Learn. Res. 10 (2009) 207–244.
[43] K.Q. Weinberger, L.K. Saul, Fast solvers and efficient implementations for distance metric learning, in: Proceedings of the 25th International Conference on Machine Learning, ICML' 08. ACM, New York, NY, USA, 2008, pp. 1160–1167, http://dx.doi.org/10.1145/1390156.1390302.
[44] A. Bellet, A. Habrard, M. Sebban, A Survey on Metric Learning for Feature Vectors and Structured Data, ArXiv13066709 Cs Stat, 2013.
[45] J.B. Tenenbaum, V. de Silva, J.C. Langford, A global geometric framework for nonlinear dimensionality reduction, Science 290 (2000) 2319–2323, http://dx.doi.org/10.1126/science.290.5500.2319.
[46] L. van der Maaten, G. Hinton, Visualizing data using t-SNE, J. Mach. Learn. Res. 9 (2008) 2579–2605.
[47] C.C. Chang, C.J. Lin, LIBSVM: a library for support vector machines, ACM Trans. Intell. Syst. Technol. 2 (2011) 1–27.
[48] J. Camacho Navarro, F. León-Vargas, J. Barrero Pérez, EMG-based system for habic hand movement recognition, DYNA 79 (2012) 41–49.
[49] A.A. Adewuyi, L.J. Hargrove, T.A. Kuiken, Evaluating EMG feature and classifier selection for application to partial-hand prosthesis control, Front. Neurorobotics 10 (2016), http://dx.doi.org/10.3389/fnbot.2016.00015.
[50] M. Simão, N. Mendes, O. Gibaru, P. Neto, A review on electromyography decoding and pattern recognition for human-machine interaction, IEEE Access 7 (2019) 39564–39582, http://dx.doi.org/10.1109/ACCESS.2019.2906584.
[51] H. Siqueira, L. Boccato, I. Luna, R. Attux, C. Lyra, Performance analysis of unorganized machines in streamflow forecasting of Brazilian plants, Appl. Soft Comput. 68 (2018) 494–506.