

Electromyographic Signal Dynamic Behavior in Neuropathies

Spectral Parameters Evaluation and Classification

Maria Marta Santos^{1,2}, Ana Luísa Gomes^{1,2}, Hugo Gamboa^{1,2}, Mamede de Carvalho³, Susana Pinto³
and Carla Quintão^{1,4}

¹*Departamento de Física, Faculdade de Ciências e Tecnologia, Universidade Nova de Lisboa, Lisboa, Portugal*

²*PLUX - Wireless Biosignals, Lisboa, Portugal*

³*Faculdade de Medicina, Instituto de Medicina Molecular, Universidade de Lisboa, Lisboa, Portugal*

⁴*Instituto de Biofísica e Engenharia Biomédica, Faculdade de Ciências, Universidade de Lisboa, Lisboa, Portugal*

Keywords: Amyotrophic Lateral Sclerosis (ALS), Coherence, Phase Locking Factor (PLF), Fractal Dimension (FD), Lempel-Ziv (LZ), Detrended Fluctuation Analysis (DFA), Multiscale Entropy (MSE), Surface Electromyography (sEMG), Ipsilateral, Classification.

Abstract: Amyotrophic Lateral Sclerosis (ALS) is a neurodegenerative disease characterized by motor neurons degeneration, which reduces muscular force, being very difficult to diagnose. Mathematical methods, such as Coherence, Phase Locking Factor (PLF), Fractal Dimension (FD), Lempel-Ziv (LZ) techniques, Detrended Fluctuation Analysis (DFA) and Multiscale Entropy (MSE) are used to analyze the surface electromyographic signal's chaotic behavior and evaluate different muscle groups' synchronization. Surface electromyographic signal acquisitions were performed in upper limb muscles, being the analysis executed for instants of contraction recorded from patients and control groups. Results from LZ, DFA and MSE analysis present capability to distinguish between the patient and the control groups, whereas coherence, PLF and FD algorithms present results very similar for both groups. LZ, DFA and MSE algorithms appear then to be a good measure of corticospinal pathways integrity. A classification algorithm was applied to the results in combination with extracted features from the surface electromyographic signal, with an accuracy percentage higher than 70% for 118 combinations for at least one classifier. The classification results demonstrate capability to distinguish both groups. These results can demonstrate a major importance in the disease diagnose, once surface electromyography (sEMG) may be used as an auxiliary diagnose method.

1 INTRODUCTION

ALS is a fatal and very progressive disease, characterized by both upper and lower motor neurons degeneration, involving brainstem and also multiple spinal cord innervation regions. This disorder is responsible for abnormal motor activity. ALS patients typically present fatigue, quickly progressive weakness and reduced exercise capacity with loss of voluntary movement, spasticity, fasciculations, dysphagia (difficulty in swallowing), dyspnea (difficulties in breathing) and dysarthria (difficulties in speaking). After the first symptoms, death may occur within 3 – 5 years for most of the patients. ALS is very difficult to diagnose, since there isn't available a reliable biomarker of disease activity and progression (Kiernan et al., 2011; Mitchell and Borasio, 2007).

Upper motor neuron integrity can be evaluated through the investigation of oscillatory activity propagation. The motor cortex activity can be recorded, and both alpha (8 – 12Hz) and beta (15 – 30Hz) frequency bands can be analyzed via coherence and PLF (Farmer et al., 2007).

This work explores the analysis of ipsilateral acquisitions, which was presented with promissory preliminary results in (Camara, 2013), using different approaches.

Motor unit recruitment patterns complexity can be quantified using FD. However, the strength of a muscle's contraction is better estimated based on Maximum Fractal Length (MFL), even for very small muscle contraction strength, rather than FD (Poosapadi Arjunan and Kumar, 2012).

The LZ measure is a well suited feature regard-

ing sEMG analysis, particularly during dynamic contractions (highly non-stationary signal), since this feature doesn't make any assumption of stationarity (Talebinejad et al., 2011).

The DFA method is proved to be efficient to describe upper-limbs movements. DFA outperforms other methods such as correlation dimension and Higuchi methods, being a stable technique to quantify fractality and to establish self-similarity, being a robust method in the presence of nonstationarity time series and trends (Phinyomark et al., 2011).

Entropy is a feature which can detect and quantify differences in the EMG signal amplitude distributions due to neuromuscular conditions (pathology) (Kaplanis et al., 2010). This feature has been successfully applied to physiological signals, in order to quantify their degree of complexity. Multiscale entropy (MSE) has been proved to be more effective than single-scale entropy in this quantification, since it considers multiple spatiotemporal scales (Zhang et al., 2013).

k-Nearest Neighbor is a relatively simple and fast algorithm, important characteristics in the classification process (Kim et al., 2011). Decision tree methods have also been used for dealing with classification problems in various domains, such as pattern recognition, data mining, web mining and signal processing, among others. However, standard decision tree algorithms can only handle discrete attributes (Wang et al., 2006). The decision tree algorithm has usually good performance for large data sets in a short time (Zhang et al., 2011). Random Forest algorithm has been tested with real and simulated data sets. The results have been proven to be very accurate. This algorithm is fast, versatile and can be applied to very large data sets. It has also been shown its robustness against noise in the outcome compared with several other methods (Roy and Larocque, 2012). Discrimination between subject or patient group concerning age, gender and injuries in athletes has been proven to be effective using a classification approach with generic features and AdaBoost (Eskofier et al., 2012). Naïve Bayes method has shown to be competitive among much sophisticated induction algorithms concerning experiments on real world data, despite the assumption of conditional independence (Wang et al., 2006).

2 METHODS

2.1 Coherence

Coherence function estimates values from 0 to 1, assuming the value 0 if there is no association between

two signals at a certain frequency, and the value 1 if there is a perfectly linear association between them (Farmer et al., 2007)

2.2 Phase Locking Factor (PLF)

PLF is a measure of synchronization between two signals, in which the frequencies of interest are isolated by the application of a narrow band-passed filter. Then, the relationship between the phases of the two signals, $\phi_j(t)$ and $\phi_k(t)$ are analyzed (Almeida et al., 2011):

$$\rho_{jk} \equiv \left| \frac{1}{T} \sum_{t=1}^T e^{i[\phi_j(t) - \phi_k(t)]} \right| = |\langle e^{i[\phi_j(t) - \phi_k(t)]} \rangle| \quad (1)$$

where ρ_{jk} is the PLF and T is the number of discrete samples.

PLF ranges from 0 to 1. While $\rho_{jk} = 0$ corresponds to asynchronous signals (their phases are not correlated), $\rho_{jk} = 1$ is attained if the two signals are in perfect synchronization (Almeida et al., 2011).

2.3 Fractal Dimension (FD)

FD is one of the most used measurements for the evaluation of the dynamics of complex systems. A non-integer FD usually indicates a chaotic behavior, and the smallest integer bigger than FD is considered the minimum number of independent variables capable to describe this behavior (West, 1994).

For physiological signals, FD can be estimated using Higuchi algorithm, since it is suitable for non-periodic and irregular time series.

This algorithm results on a plot of several curve lengths, $L(k)$, for a pre-determined k_{max} , being $k \in [1, k_{max}]$ (Phinyomark et al., 2011).

2.4 Lempel-Ziv (LZ)

LZ is another tool used to analyze the deterministic complexity of a highly non-linear chaotic setting. To compute LZ, it is necessary to convert the sEMG signal into a symbolic sequence, conventionally a binary sequence, by comparison to a threshold. Next, the number of distinct patterns within a signal is obtained, which is directly related to the complexity of the system (Talebinejad et al., 2011).

2.5 Detrended Fluctuation Analysis (DFA)

The DFA method can be applied to the study of electrophysiological signals, being a modified root mean

square (RMS) analysis of a random walk (Phinyomark et al., 2011).

The outputs of the DFA algorithm are scaling exponents extracted from a log-log graph (for details (Phinyomark et al., 2011)). These exponents assume values between 0 and 2, according to the time series behavior:

- If $0 < \alpha < \frac{1}{2}$ the time series is anti-correlated
- If $\alpha \cong \frac{1}{2}$ the time series is uncorrelated, or indicates White noise (the value at one instant cannot be correlated with any previous value)
- If $\frac{1}{2} < \alpha < 1$ the time series is correlated
- If $\alpha \cong 1$ indicates Pink noise ($\frac{1}{f}$ noise)
- If $1 < \alpha < \frac{3}{2}$ indicates nonstationary or random walk
- If $\alpha \cong \frac{3}{2}$ indicates Brownian noise (i. e. the integration of the White noise)

2.6 Multiscale Entropy (MSE)

Entropy has been often used to quantify complexity, since traditional entropy definitions (e. g. Shannon-entropy) are used to measure disorder and uncertainty, as well as to characterize a systems' gain of information. Approximate Entropy and its modification Sample Entropy are entropy-based complexity measures with a single scale which are widely used in short and noisy time series. Thus, in MSE one of this approaches is applied to different time scales, resulting on a plot of the entropy value as a function of the factor scale, τ (Zhang et al., 2013).

2.7 Classification

Classification methods are used to identify the belonging of a novel observation in a set of categories (sub-populations). These categories are obtained with based on a training group of observations, therefore, it is necessary to have previous knowledge of the category membership of each observation of the training group. These observations are then analyzed according to the extracted features. In this project, the used features are: kurtosis, maximum frequency, mean, median frequency, power band, spectral kurtosis, spectral skewness, spectral spread and correlation. There will be also analyzed the results of the previously referred implemented algorithms (coherence, PLF, FD, LZ, DFA and MSE). Classification can be implemented throughout a various number of algorithms, the classifiers (k - Nearest Neighbor, Decision Tree, Random Forest, AdaBoost and Naïve Bayes). Leave-one-out cross validation iterator is used to split

data in train/test sets. Hence, all samples except one are used as a train set, being this left out sample tested after (Pedregosa et al., 2011).

3 ACQUISITIONS

3.1 Subjects

Measurements were performed in two different groups of subjects: group of patients, with 21 members presenting ALS, and group of control, with 26 members which do not evidence ALS disease. Therefore, the patients group contains 21 subjects, varying this number according to the analyzed channel (left or right hand or forearm), since some patients presented inability of self-controlled movement for one arm. All participants from the patients group have been diagnosed within less than three years, except two members which have been diagnosed previously. All participants from both groups have ages between 23 and 77 years (mean of 59 years for the patient group and 45 years for the control group).

3.2 Acquisition Protocol

The acquisition protocol is identical to the one in (Camara, 2013). The performed task was repeated for 6 minutes or less according to maximum time borne by the patients. Subjects sat down and placed both hands and forearms on a desk in a parallel position, 10 cm away from each other with hand palms facing one another in 90 degrees flexion with the elbow. While listening to a programmed sound, which guided the movement, subjects were asked to coordinately elevate both index fingers vertically with maximum articular amplitude in an opposite direction from the other fingers position, hold that position for 3 seconds while maintaining a certain force and return to the original position, remaining in that position for 3 seconds while trying to relax the arms muscles as much as possible.

3.3 Recordings

Contralateral and ipsilateral acquisitions were recorded simultaneously, and 4 signals were acquired from each subject using EMG sensors attached to a *bioPlux* device. The sensors were placed on the first dorsal interosseus muscle for both left and right hand, and on the extensor digitorum communis muscle for both left and right forearm. Ground was placed on ulna bone inferior extremity, where no

muscle activity is present. Figure 1 shows the surface electrodes placements.

The *bioPlux* device has eight analog channels with 12-bit resolution and also an external channel to be used as a reference ground. The EMG sensors have second order band pass filter with *cutoff* frequencies of 25 and 450. EMG signals were recorded using a sampling frequency of 1000 Hz, being the recorded data transmitted via Bluetooth to a computer.

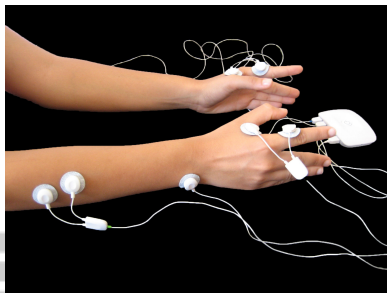


Figure 1: Simultaneous contralateral and ipsilateral experimental setup: *Bioflux* research device, placement of four EMG sensors and ground.

4 SIGNAL PROCESSING

All acquired signals were processed using python language. First, the signals' Direct Current (DC) component was removed and a third order butter band pass filter of 10 – 500 Hz was applied. From each pair of ipsilateral signals, intervals of common contraction were isolated from intervals of relaxation using the algorithm referred in (Camara, 2013).

All the previously referenced algorithms were applied to moments of contraction. While coherence and PLF were calculated twice for each subject, being calculated for a pair of signals, FD, LZ, DFA and MSE were calculated four times, being applied to each one of the four acquired signals individually.

Coherence and PLF were calculated for two common sections of data from an interval of one contraction. This was performed for all contractions, with posterior averaging of all epochs. FD, LZ, DFA and MSE calculus was applied to a concatenation of all contractions.

4.1 Coherence Processing

All EMG signals were full-wave rectified and then, for each contraction, coherence was calculated using python libraries (matplotlib.mlab.cohere tool). Coherence is then averaged for the moments of contraction. The used sampling frequency is 1000 Hz, the Nonequispaced fast Fourier transform (NFFT) is 512

and the value that dictates the dependency between FFT windows is $NFFT/2$. Coherence was averaged for each frequency among all the subjects within the same group.

4.2 PLF Processing

PLF algorithm was developed in (Camara, 2013). All signals were full-wave rectified and each signal was band pass filtered in order to remove all the other undesirable frequencies. The used filter was $[f - 2, f + 2]$, being f the analyzed frequency. Therefore, PLF calculus was performed as many times as the number of frequencies to analyze. PLF was obtained with resolution of 1Hz for all the frequencies within the beta band (15 – 30Hz). PLF was computed according to equation 1. Hence, a different PLF value is obtained for each contraction of the analyzed signals and, in order to attain a final value for each subject, PLF was averaged among all contractions within the same acquisition. Finally, for each analyzed frequency, PLF was averaged among all members within each group, for the patient and the control groups.

4.3 FD Processing

In the Higuchi fractal dimension algorithm k_{max} was defined as 128, as suggested in literature (Phinyomark et al., 2011). The FD coefficient was estimated and averaged among all subjects within each group.

4.4 LZ Processing

The LZ coefficient was calculated for a binary sequence obtained from the rectified filtered used signal with threshold defined as 0.4.

Since the LZ coefficient is greatly related to the number of different patterns within each signal, and the number of patterns is related to the length of the signal, all the signals were cut in accordance with the minimum reasonable signal length within both groups. An average of the LZ coefficient was calculated for all members within the patients and the control's groups.

4.5 DFA Processing

The DFA was applied among all subjects, obtaining two distinct scaling exponents, α_1 and α_2 . These scaling exponents were averaged for all members within each group.

4.6 MSE Processing

For MSE, the Sample Entropy was computed for 20 scales. These values were averaged among all subjects within each group, and plotted in a graph for both populations.

4.7 Classification

The feature extraction and the classification algorithms were developed in (Gomes, 2014).

The algorithms were adapted for the use of two channels, and the features extracted were selected according to this work objectives.

Features were extracted from a pair of signals (right hand and forearm). The sampling frequency was placed at 1000 Hz, being used a central window with 2000 points. The results obtained from the previously described algorithms were joined to these features.

Posteriorly several combinations of the extracted features, the developed algorithm results and of both of them were organized and classified.

5 RESULTS AND DISCUSSION

In this section, we show results obtained from right arm signals, being the most relevant conclusions also achieved for left arm analysis.

5.1 Coherence Results

Coherence mean values for the group of patients and the group of control are presented in Figure 2.

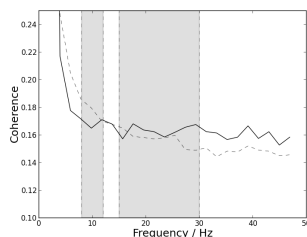


Figure 2: Mean coherence dependency on frequency for patients by the straight line and for controls by the dashed line with NFFT placed as 512. The first grey box delimitates the frequencies corresponding to the alpha band (8 – 12 Hz). The second grey box delimitates the frequencies corresponding to the beta band (15 – 30 Hz). These results are from the right arm.

By the observation of the graphics presented in Figure 2, coherence mean values are very similar

for both the patients and the control groups. Control group mean value for coherence is slightly higher within the alpha band. Coherence pooled value for patients is $0,23 \pm 0,21$ for the right arm. Coherence pooled value for the control group is $0,22 \pm 0,21$ for the right arm. These values are different from those found in literature (Fisher et al., 2012), since slightly higher values of coherence were expected for the control group within the beta band than for the patients group. These differences may be explained by the sampling frequency used, differences in the acquisition protocol, differences in the used algorithm or parameters and the tested subjects themselves (age, gender, lifestyle, etc.). The obtained results are also different from the expected for ipsilateral acquisitions in (Camara, 2013), since only preliminary results were obtained previously.

5.2 PLF Results

PLF mean values depending on frequency for both groups are presented in Figure 3.

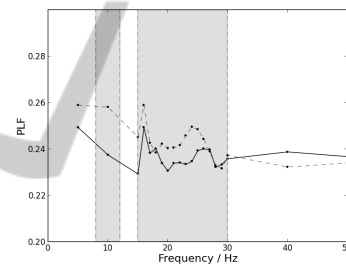


Figure 3: Mean PLF values dependency on frequency for patients by the straight line and for controls by the dashed line; The first grey box delimitates the frequencies corresponding to the alpha band (8 – 12 Hz). The second grey box delimitates the frequencies corresponding to the beta band (15 – 30 Hz). These results are from the right arm.

Observing the graphics presented in Figure 3, PLF values appear to be very similar for both groups. Alpha band appears to demonstrate a higher difference between both groups PLF values. However, since this work proceeds the investigation of PLF within the beta band (Camara, 2013), PLF was only computed for 10 Hz within the alpha band. PLF pooled value for the patients group is $0,237 \pm 0,006$ for the right arm. PLF pooled value for the control group is $0,242 \pm 0,009$ for the right arm. These results are different from the expected for ipsilateral acquisitions in (Camara, 2013), since only preliminary results were obtained previously. Differences in results may also be explained by the length of the used signal and the sampling frequency used, and characteristics such as age, gender and lifestyle may influence this analysis.

5.3 FD Results

FD coefficient mean and standard deviation values are presented in Table 1, for the group of patients and the group of control. Table 2 represents the MFL (the point of lowest scale) values for each group.

Table 1: Mean and standard deviation values of FD coefficient for patients and control group for the right arm.

	Patients group	Control group
Right hand	-1.985 ± 0.008	-1.983 ± 0.008
Right forearm	-1.983 ± 0.013	-1.984 ± 0.006

Table 2: MFL values for patients and control group.

	Patients group	Control group
Right hand	3.72 ± 0.39	3.99 ± 0.20
Right forearm	3.72 ± 0.47	3.80 ± 0.22

From the observation of Table 1 FD coefficient is almost identical for both groups. Therefore, FD coefficient is not a good measure of distinction between patients and control group. However, the obtained FD values are very similar to the results obtained from different upper limb activities in (Phinyomark et al., 2011). Observing Table 2, MFL is always slightly higher for the control group.

5.4 LZ Results

Table 3 shows the LZ coefficient mean value and standard deviation for a binary sequence.

Table 3: LZ coefficient for a binary sequence obtained from the rectified filtered used signal with threshold defined as 0.4.

	Patients group	Control group
Right hand	0.17 ± 0.13	0.27 ± 0.16
Right forearm	0.27 ± 0.29	0.15 ± 0.13

From the observation of Table 3, LZ coefficient presents higher values for the control group for the right hand, and higher values for the patients group for the right forearm. Tables 3 presents a good distinction between patient and control groups. Therefore, for a binary sequence obtained from the filtered used signal and for a binary sequence obtained from the rectified filtered used signal, LZ algorithm appears to distinguish subjects between both groups regarding pooled LZ coefficient values, however, with some variance among subjects within each group.

Table 4: DFA α_1 coefficient mean and standard deviation values for both groups for the right arm.

	Patients group	Control group
Right hand	1.23 ± 0.17	1.35 ± 0.14
Right forearm	1.39 ± 0.09	1.44 ± 0.11

Table 5: DFA α_2 coefficient mean and standard deviation values for both groups for the right arm.

	Patients group	Control group
Right hand	0.50 ± 0.12	0.52 ± 0.09
Right forearm	0.62 ± 0.16	0.59 ± 0.10

5.5 DFA Results

Tables 4 and 5 present the DFA coefficients, α_1 and α_2 mean values for both groups, respectively.

As observed in Tables 4 and 5, DFA coefficients generally present higher values for the control group. Therefore, DFA algorithm appears to distinguish both groups.

5.6 MSE Results

Figure 4 represents the Sample Entropy mean value for each scale for both patient and control groups.

Observing Figure 4, patients group exhibit higher entropies for the right hand, whereas control group exhibit higher entropies for the right forearm. Therefore, despite statistically the differences between patients and control's groups are very small, the MSE tendency appears to be distinct between both groups.

MSE pooled values for the patients and control groups are presented in Table 6.

Observing the results presented in Figure 4 and Table 6, MSE algorithm appears to be capable of distinguish both groups.

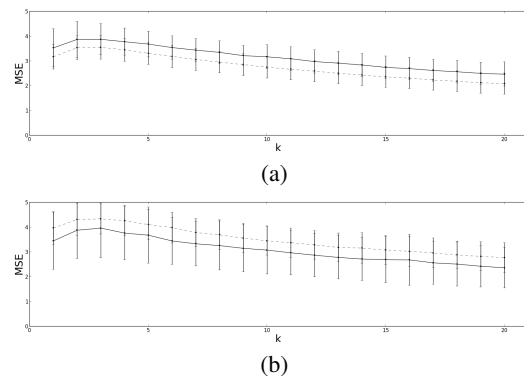


Figure 4: Sample Entropy mean value for each scale. The straight line represents the patient group, and the dashed line represents the control group. The errorbars represent the standard deviation. (a) Results for the right hand. (b) Results for the right forearm.

Table 6: MSE mean and standard deviation values, for both patients and control groups.

	Patients group	Control group
Right hand	3.15 ± 0.46	2.77 ± 0.48
Right forearm	3.08 ± 0.49	3.50 ± 0.52

5.7 Classification Results

Table 7 presents the 14 combinations and the 3 classifiers with better results for the right arm for a Leave-one-out cross validation.

Table 7: Classification results for the right arm for Decision Tree, Random Forest and AdaBoost Classifiers. 1) sum power band + MSE + DFA ($\alpha_1 + \alpha_2$); 2) mean + MSE + DFA ($\alpha_1 + \alpha_2$); 3) mean + MSE + DFA ($\alpha_1 + \alpha_2$) + LZ; 4) spectral skewness + MSE + DFA ($\alpha_1 + \alpha_2$) + LZ; 5) spectral spread + MSE + DFA ($\alpha_1 + \alpha_2$) + LZ; 6) spectral kurtosis + MSE + DFA ($\alpha_1 + \alpha_2$) + LZ; 7) sum power band + MSE + coherence + PLF + DFA ($\alpha_1 + \alpha_2$); 8) kurtosis + MSE + DFA ($\alpha_1 + \alpha_2$); 9) kurtosis + MSE + DFA ($\alpha_1 + \alpha_2$) + LZ; 10) kurtosis + MSE + coherence + PLF + DFA ($\alpha_1 + \alpha_2$); 11) maximum frequency + MSE + coherence + PLF + DFA ($\alpha_1 + \alpha_2$); 12) mean + MSE + DFA ($\alpha_1 + \alpha_2$); 13) spectral skewness + MSE + coherence + DFA ($\alpha_1 + \alpha_2$); 14) spectral spread + MSE + coherence + PLF + DFA ($\alpha_1 + \alpha_2$).

Combinations	Decision Tree	Random Forest	AdaBoost
Right Hand			
1)	0,806	0,694	0,722
2)	0,806	0,778	0,722
3)	0,833	0,611	0,750
4)	0,806	0,667	0,667
5)	0,806	0,639	0,722
Right Forearm			
6)	0,667	0,750	0,806
Right Arm			
7)	0,778	0,750	0,806
8)	0,556	0,750	0,806
9)	0,500	0,806	0,778
10)	0,694	0,833	0,694
11)	0,667	0,611	0,806
12)	0,556	0,806	0,639
13)	0,667	0,861	0,750
14)	0,694	0,833	0,694

Classification results demonstrate the algorithms distinction capability, since the presented results show the percentage of cases that the algorithm classified correctly (accuracy percentage).

Classification results were obtained for the right arm for a leave-one-out cross validation tested with k - Nearest Neighbor, Decision Tree, Random Forest, AdaBoost and Naïve Bayes classifiers. In spite of 201 combinations have been arranged, 118 combinations present values higher than 70.0% for at least one classifier, 41 combinations present results higher than 77,8% for at least one classifier, and 14 combina-

tions present results higher than 80.6% for one classifier (Decision Tree, Random Forest or AdaBoost).

The best results include MSE, DFA, LZ, coherence and PLF algorithm results, and also the extracted features mean, maximum frequency, spectral kurtosis, spectral skewness, spectral spread, sum power band and kurtosis, in several different combinations. The top 14 best combinations include the results of both DFA and MSE algorithms. The best obtained combination is spectral skewness + MSE + coherence + DFA (α_1 and α_2), with an accuracy percentage of 86,1% for a Random Forest Classifier. Therefore, it is proved that in spite of some algorithms may present slender differences between both control and patients groups, the combination of these algorithms with other measures can improve the distinction capability between members of patients and control groups.

6 CONCLUSIONS

In this work, methodology was developed to evaluate the complexity of sEMG signal acquired from different muscle groups of healthy subjects and patients with ALS. FD, LZ, DFA and MSE algorithms were implemented and all of these algorithms as well as coherence and PLF algorithms were applied to the acquired filtered signals.

Contrary to the results presented in literature, coherence analysis does not present significant differences between the group of patients and the group of control. PLF analysis also does not present any significant differences between both groups. Results from both algorithms appear to be slightly higher for the control group for the alpha band frequencies. For further work is suggested to compute PLF for all the frequencies within the alpha band with resolution of 1 Hz.

FD analysis results in FD coefficients very similar for all the signals for both patients and control groups. Therefore, this algorithm does not seem good to obtain a distinction between both groups. However, MFL presents slightly higher values regarding the control group, being this a better measure of distinction between both groups than the FD coefficient.

LZ analysis presents better results, being LZ coefficient higher for the control group for the right hand, and higher for the patients group for the right forearm. These results are for a binary sequence obtained from the rectified filtered used signal. Therefore, LZ coefficient appears to be a good reflection of neural degeneration. Since the used threshold to obtain the binary sequence was defined as 0.4, it is not adapted

to each individual signal. Therefore, it is suggested for further work a threshold obtained as a percentage of the standard deviation of each signal.

DFA analysis presents higher values for both α_1 and α_2 DFA coefficients for the control group. Then, this algorithm seems to be a good measure to reflect neural degeneration.

MSE analysis presents higher values for the right hand for the patients group, and higher values for the right forearm for the control group. This algorithm appears to be a good indicator of neural degeneration.

LZ, DFA and MSE analysis have then potential as a quantitative test for upper and lower neural integrity concerning ALS disease.

Classification results demonstrate to provide a good distinction of both groups, being the combination of various algorithms with features proved to be advantageous to improve both groups distinction capability.

REFERENCES

- Almeida, M., Vigário, R., and Bioucas-Dias, J. (2011). Phase locked matrix factorization. In *Proc. of the EU-SIPCO Conference*, pages 1728–1732.
- Camara, M. (2013). Coherence and phase locking disruption in electromyograms of patients with amyotrophic lateral sclerosis. Master's thesis, Faculdade de Ciências e Tecnologia da Universidade Nova de Lisboa.
- Eskofier, B. M., Kraus, M., Worobets, J. T., Stefanyshyn, D. J., and Nigg, B. M. (2012). Pattern classification of kinematic and kinetic running data to distinguish gender, shod/barefoot and injury groups with feature ranking. *Computer methods in biomechanics and biomedical engineering*, 15(5):467–474.
- Farmer, S. F., Gibbs, J., Halliday, D. M., Harrison, L. M., James, L. M., Mayston, M. J., and Stephens, J. A. (2007). Changes in emg coherence between long and short thumb abductor muscles during human development. *The Journal of physiology*, 579(2):389–402.
- Fisher, K. M., Zaaimi, B., Williams, T. L., Baker, S. N., and Baker, M. R. (2012). Beta-band intermuscular coherence: a novel biomarker of upper motor neuron dysfunction in motor neuron disease. *Brain*, 135(9):2849–2864.
- Gomes, A. L. (2014). Human activity recognition with accelerometry: Novel time and frequency features. Master's thesis, Faculdade de Ciências e Tecnologia da Universidade Nova de Lisboa.
- Kaplanis, P. A., Pattichis, C. S., Zazula, D., et al. (2010). Multiscale entropy-based approach to automated surface emg classification of neuromuscular disorders. *Medical & biological engineering & computing*, 48(8):773–781.
- Kiernan, M. C., Vucic, S., Cheah, B. C., Turner, M. R., Eisen, A., Hardiman, O., Burrell, J. R., and Zoing, M. C. (2011). Amyotrophic lateral sclerosis. *The Lancet*, 377(9769):942–955.
- Kim, K. S., Choi, H. H., Moon, C. S., and Mun, C. W. (2011). Comparison of i_k k_j/i_k -nearest neighbor, quadratic discriminant and linear discriminant analysis in classification of electromyogram signals based on the wrist-motion directions. *Current Applied Physics*, 11(3):740–745.
- Mitchell, J. D. and Borasio, G. D. (2007). Amyotrophic lateral sclerosis. *The lancet*, 369(9578):2031–2041.
- Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., Vanderplas, J., Passos, A., Cournapeau, D., Brucher, M., Perrot, M., and Duchesnay, E. (2011). Scikit-learn: Machine learning in Python. *Journal of Machine Learning Research*, 12:2825–2830.
- Phinyomark, A., Phukpattaranont, P., Limsakul, C., and Phothisonothai, M. (2011). Electromyography (emg) signal classification based on detrended fluctuation analysis. *Fluctuation and Noise Letters*, 10(03):281–301.
- Poosapadi Arjunan, S. and Kumar, D. K. (2012). Computation of fractal features based on the fractal analysis of surface electromyogram to estimate force of contraction of different muscles. *Computer Methods in Biomechanics and Biomedical Engineering*, (ahead-of-print):1–7.
- Roy, M.-H. and Larocque, D. (2012). Robustness of random forests for regression. *Journal of Nonparametric Statistics*, 24(4):993–1006.
- Talebinejad, M., Chan, A. D., and Miri, A. (2011). A lempel-ziv complexity measure for muscle fatigue estimation. *Journal of Electromyography and Kinesiology*, 21(2):236–241.
- Wang, L.-M., Li, X.-L., Cao, C.-H., and Yuan, S.-M. (2006). Combining decision tree and naive bayes for classification. *Knowledge-Based Systems*, 19(7):511–515.
- West, B. J. (1994). *Fractal physiology*, volume 2. Oxford University Press.
- Zhang, X., Chen, X., Barkhaus, P. E., and Zhou, P. (2013). Multiscale entropy analysis of different spontaneous motor unit discharge patterns. *Journal of Biomedical and Health Informatics*, 17(2).
- Zhang, X., Chen, X., Li, Y., Lantz, V., Wang, K., and Yang, J. (2011). A framework for hand gesture recognition based on accelerometer and emg sensors. *Systems, Man and Cybernetics, Part A: Systems and Humans, IEEE Transactions on*, 41(6):1064–1076.