Thought Recognition through EEG signals

Ciprian Matteo, Corniani Giulia, Mariotti Ettore
Department of Information Engineering, Università degli Studi di Padova, Italy

***Abstract*—In this paper, we propose a method to classify yes/no answers to simple questions, based on electroencephalographic (EEG) signals. We used the signals acquired in several sessions on a single patient in Completely Locked In State due to ALS disease and therefore unable to communicate. After a signals processing phase for the extraction of features, SVM classifiers were built achieving good performances in the distinction of the patient's thought. The low computational weight and the low number of signals initially required for the construction of the classifiers make this method suitable for online application with possible great benefits in the quality of life of patients affected by this disease.**

I. Introduction

Amyotrophic Lateral Sclerosis (ALS) is a neurological degenerative disease that mainly involves the nerve cells. Both the upper motor neurons and the lower motor neurons degenerate or die, and stop sending messages to the muscles. In this condition, muscles gradually weaken, start to twitch, and waste away until the brain loses its ability to initiate and control voluntary movements. People affected by neurodegenerative disease such as ALS end up in a completed locked in state (CLIS) in which movements and communication become impossible. Nonetheless, cognitive abilities are intact and different studies prove that quality of life in these people can be good and is strongly positively correlated with the possibility of communicating.

Previous works, such as [1], showed the possibility of designing a Brain-Computer Interface that allows basic communication, distinguishing yes and no thoughts of the patient. This kind of systems can use both non-invasive or invasive data acquisition instrumentation.

In our case, only non-invasive techniques were used: electroencephalography (EEG), electrooculography (EOG), electromyography (EMG) and functional near-infrared spectroscopy (fNIRS).
In a session, the patient answers 20 questions (10 yes + 10 no answer in a random sequence). The questions allow only a yes or no response and the answers are assumed to be known. The questions are sampled from a set built with the help of the patient family. Each session is approximately 10 minutes long and 28 signals are simultaneous acquired using different techniques, in particular:

- 15 fNIRS signals
- 5 EEG signals
- 4 EOG signals
- 4 EMG signals

In our work, we analyze the data acquired during different sessions and different days of acquisition.

The state of the art results of such a system uses a support vector machine (SVM) on the fNIRS dataset while using EEG just for asserting a vigilant state. The average performances of SVM+fNIRS achieved a 70% of classification accuracy. The non-excellent efficiency of this system allows to ask the patient only general question, mainly about his daily routine preferences but is still inappropriate for sensitive issues, such as decisions on the end of life, a significant topic for people in this situation.

In our work, we decided to use only EEG signals. EMG and EOG signals were discarded for the short-term results they can ensure. Indeed, due to the quick degeneration of the disease, the patient is destined to lose the use of the jaw muscles and the eyes move, activities which were recorded with these techniques. Consequently, classifiers built using these signals would lose accuracy with the evolution of the disease. No techniques have been implemented to integrate the information coming from the EEG and fNIRS signals, therefore the fNIRS signal was not used because alone it results less informative than the other.

Starting from the extraction and selection of features, we set an SVM classifier to discriminate between yes and no signals.

II. Methods

A. Dataset

The data analyzed in this work were acquired over seven days. Every day, a number ranging from 2 to 4 sessions, each composed of 20 question (10yes/10no), were performed.

The EEG electrodes were placed in the position C5, C6, FC5, FC6, and Cz of the international 10-20 system for electroencephalography electrode placement. The sampling frequency is 500Hz.

For further information about the data acquisition see [1]

B. Signal Filtering

First of all the drift of the signal was removed using the Matlab function `detrend` which removes a linear trend from a vector. The signal was processed using a passband Chebyshev Type II filter. In particular, after the visual inspection of the power spectrum and according to the physiological characteristics of EEG signals, the bandpass was set from 0.1 Hz to 30 Hz with 150 dB attenuation in the stop-bands. The filter was initially designed using the Matlab function `cheb2ord`, which set the filter order, and `cheby2`. Finally, with the function `zp2sos`, the zero-pole model of the filter was converted into a second-section form to improve the stability.

The filtering procedure was applied to the entire signals which were then segmented and labeled into yes/no instances using a trigger signal.

C. Feature Extraction Approach

The purpose of this phase is to calculate a number of features able to well characterize the signals under examination. For this goal, we used a MATLAB library developed for the analysis of EEG in neonatal intensive care, described in [2]. Despite the different application area, it resulted suitable for the extraction of useful features to describe our signals. The features extracted can be divided into four main categories: range, amplitude, spectral and connectivity.

- *Amplitude features:* characterization of the signals in the time domain. Amplitude is quantified by signal power and signal envelope. Also, the Gaussianity of the process is evaluated with skewness and kurtosis.
- *Range features:* range EEG (rEEG) is an alternative representation of EEG signals in the time domain. In particular, rEEG estimates a peak-to-peak measure of voltage. Range features, extracted from rEEG, summarizes their trends with different measures.
- *Spectral features:* quantification of the spectral characteristic of the signals starting from the Power Spectral Density (PSD).
- *Connectivity features:* measures of the connectivity and symmetry between the two hemispheres. These features are estimated using the symmetric channels C5, C6 and FC5, FC6.

All the features, apart from the spectral difference and the spectral edge frequency, are estimated within four different frequency bands of the EEG: [0.5-4; 4-7; 7-13; 13-30] Hz.
It is also fair to mention that in the three first categories (Range, Amplitude, Spectral) there are Single-Channel-Features which are calculated independently for each of the five acquisition channels. Conversely, the connectivity features are Multiple-Channel-Features since they are calculated considering the signals of the channels C5, C5, FC5 and FC6 jointly. This determined the structure of the features matrix.

D. Feature Selection

As described in the introduction each dataset is composed of n binary question with known answer (which we call instances). For each instance we have a total of 386 of different features all concatenated in a single row, thus our features dataset is represented by a $n \times 386$ matrix which we call *full features dataset*. Not all the features carry useful information for our classification goal thus it is important to develop a computationally feasible automatic algorithm able to understand what is the subset that yields the best performances. A brute force approach is indeed impractical because it would require to test $\sum_{k=1}^{386} \binom{386}{k} \approx 1.57 * 10^{116}$ different combinations.

In order to reduce this number, a score was given to each feature using ANOVA F-test. Using this, only the top k features with the highest score were selected. It was then defined a new *k-reduced dataset* of size $n \times k$.

For assessing the best value of k it is necessary to better define what we mean for "best performances". Given our specific problem, it is important to maximize the generalization accuracy of our classifier with the least number of train instances. Thus we choose to keep the smallest k that, given a particular classifier, would yield the highest validation accuracy on the test set while training on few instances.We proceeded with our work so that we could best mimic the real experimental scenario. The dataset is divided into two groups: training (X_{tr}) and testing (X_{te}) sets. This first partitioning is made keeping track of the historical sequence of the questions. Every parameter is computed on X_{tr}, leaving untouched every data point of the X_{te}. For the evaluation of the various parameters the set X_{tr} is randomly splitted in a sub-training set (X_{sub-tr}) and a validation set (X_{val}). Given the scarcity of data, in order to have stable estimates

of the scores we perform many (500) of these random sub-training/validation partitions. The procedure can be summarized as follows:

1) Split the *full features dataset* in X_{tr} and X_{te} sets
2) Generate 500 random partition of X_{tr} as X^i_{sub-tr} and X^i_{val}, for $i = 1, ..., 500$
3) For each k going from 1 to the maximum number of features do:
4) Compute the ANOVA F-test score statistics on X^i_{sub-tr}
5) Reduce X^i_{sub-tr} and X^i_{val} to have k features and call them $X^{i,k}_{sub-tr}$ and $X^{i,k}_{val}$, (using the reducer parameters computed on point 4)
6) Fit the classifier on the train set $X^{i,k}_{sub-tr}$ and estimate the validation performances on $X^{i,k}_{val}$. Call the validation score α^k_i
7) Return α^k as the average of the scores for each i, in other words $\alpha^k = 1/500 \sum_i \alpha^k_i$.
8) return $k_{best} = \min(\arg\max_k \alpha^k)$

E. Classifier

We used a Support Vector Machine (SVM) with linear kernel and penalty term C=1 as classifier. For numerical and technical reasons, before using the SVM, we standardized the data centering each feature to have null mean and unit variance. The parameters for the standardization were computed on X_{tr} and propagated to the X_{te}.

III. Results

In this section, the results are presented.
We can divide our analysis into two phases:

- Single-Day Dataset (Dataset zero): we worked on a single-day dataset: considering the experimental setup explained above, we tested our model splitting the dataset into different training/test portions and assessing the evolution of accuracy.
- Six-Days Dataset: using the dataset acquired in six consecutive days, we both analyzed the generalization properties of the built classifier, and built a different classifier for each day.

A. Single-Day Dataset Analysis

1) Feature selected: The optimal number of feature for our dataset resulted in $k = 60$; in particular, the features selected are reported in the in the appendix A.

As we can deduce from histogram reported in Figure 1, most of the feature selected (43 of 60) are calculated in the two lower frequencies bands ([0.5-4; 4-7]Hz). Regarding the contribution of the different channels to representative features, as shown in Figure 2, channel C5, C6, FC5, and FC6 contribute quite equally while Cz seems to be the least informative channels with 7 of 60 features brought.

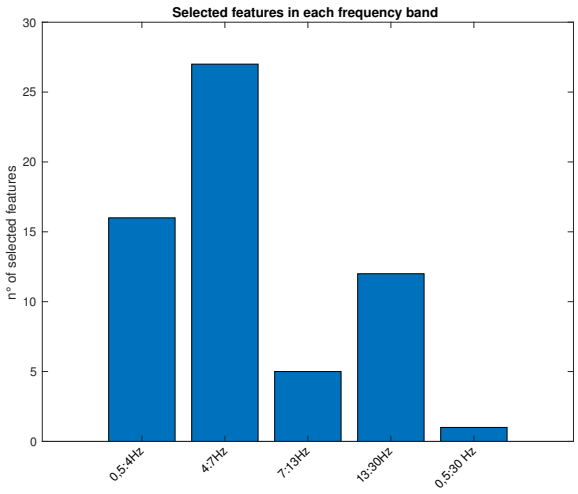

Fig. 1: Histogram of the selected features belonging to each sub-band

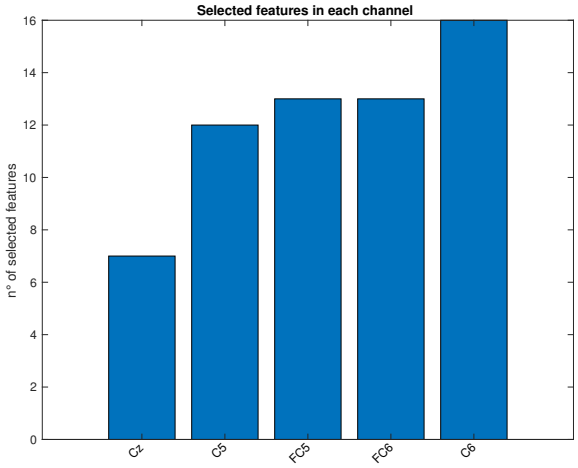

Fig. 2: Histogram of the selected features belonging to each channel

None of the connectivity features was selected for the classification.

2) Classifier performances: In Figure 3 are presented the accuracy results for different train/test partition size. For each train/test fraction, the accuracy of 10'000 randomly sampled datasets were computed: each point is the mean and each error bar represent the standard deviation of the distribution.

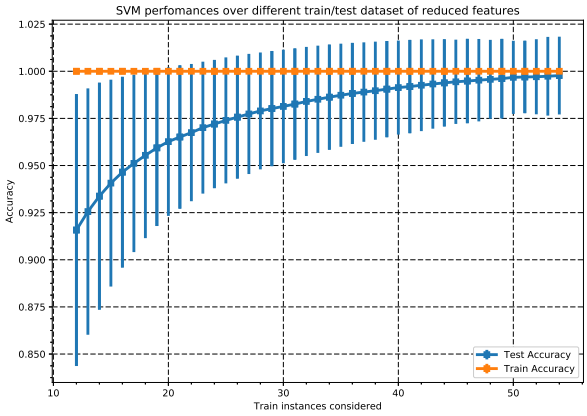

Fig. 3: Performances of the SVM with $k = 60$ with different train/test size on Dataset zero. Each point represent the mean of the results on 10'000 random partition and the error bar represent its standard deviation

Even though the SVM is always in over-fitting, it does indeed present very good generalization performances on our dataset with few training samples. This is a very desirable property in a real-application scenario.

Moreover, the SVM is extremely fast in both training and computing predictions and is memory efficient as it only requires a subset of the training points in the decision function (also called support vectors), so the hardware requirement for the classification are minimal.

B. Six-Days Dataset Analysis

1) Classifiers performance: For each day of acquisition, a different classifier was built trying to reproduce the experimental conditions. For this reason, the dataset of each day was divided into training and test sets according to the chronological order of signal acquisition. All the useful parameters for the construction of the classifier were then calculated using only the signals contained in the training set, which correspond to the first signals acquired each day, as explained in detail in section II.

In Table I the results obtained and the selected k with two different training and test partitions are reported. It can be observed that the classifiers achieve very different performances on different days. For example, the classifier built with data of the day 5 using only 2 sessions for the train, is able to correctly classify 100% of the signals, while the classifier built with the same procedure on the data of day 2 reaches much lower performance. This is probably due to different psychophysical states in which the patient was on the days of acquisition and which obviously cannot be modified to improve the success of the experiment.

TABLE I: Results of the whole procedure made in an experimental-like setting: Fixed the training dataset X_{tr} keeping the chronological order of the questions, the best k is searched and the performances are computed on the untouched test-set X_{te}

Dataset	Size	Score (tr/te 50%)	k_{opt}	Score (tr/te 75%)	k_{opt}
Day 0	60	100%	66	100%	52
Day 1	80	62.5%	9	85%	72
Day 2	80	52.5%	30	50%	35
Day 3	40	50%	21	60%	33
Day 4	80	67.5%	22	85%	41
Day 5	60	95%	94	100%	63
Day 6	80	85%	42	85%	33

2) Classifier generalization ability: Initially, the classifier built on the Single-Day dataset was used to classify the Six-Days dataset signals. The results obtained were very poor, with a classification accuracy not exceeding 50%. This is justifiable as the brain morphology and activation could have changed in the period after the acquisition of the first dataset, which dates back to a period much earlier than the second one.

Subsequently, the same approach was used with the Six-Days dataset signals: for example, with the data of day 1 a classifier was built and then used to classify the data of day 2. Although in this case the signals were acquired at a short time distance (min. 24 hours, max 5 days), the accuracy of the classification was always lower than 70%. This means that the method implemented in this work requires that in each session the patient is initially asked a series of questions with a known answer with which to create a new classifier. Only after this phase, it would be possible to proceed with open questions.

IV. Conclusions

In general, the proposed procedure leads to good results in the classification of signals analyzed. The main problem can be identified in the great variability of performance achieved using the datasets of different days. indeed, while on the one hand the results obtained can be considered excellent, on the other some classifiers are unusable to proceed with open questions, the ultimate goal of the system. We were not able to explain this difference in performance either for lack of time or for lack of information. In fact, the data were provided to

us and we were not the first person to acquire them with the possibility to assess, for example, the status of the patient. Further work on this procedure could lead to improving what we propose as discussed in the next section with huge benefits for the patients and his family.

V. Future works

The continuation of our work consists in modifying the implemented algorithms to make them usable in online applications: indeed, we have worked remotely only on signals acquired by others.

In this work, we have extracted different characteristics and selected some of them to build the classifiers without any consideration of their physical and physiological meaning. To obtain more stable results in the various days of classification, it is important to analyze the meaning of the most significant features and find a correlation between these and the patient's psychophysical state. The latter could be analyzed through an analysis of EEG signals or other physiological parameters that can highlight for example if the patient is in a state of sleep or if his attention is disturbed by pain.

It could be interesting analyzing the features selected to built each classifier and studying their distribution among different channels and different sub-bands. This could give important information about the evolution of the brain functionality through the time and the disease progress. In addition, a set of common features used in the different high-performance classifiers would be identified and used to build a single general classifier. This, although with suboptimal classification performances, would avoid the phase of training with known-answer questions in each session.

It would also be interesting to extend the study of this procedure to other patients in similar conditions and to healthy subjects. In this way, the differences in brain function induced by the disease could be highlighted and useful generalizations could be reached in the construction of future classifiers.

References

[1] U. Chaudhary, B. Xia, S. Silvoni, L. Cohen, and N. Birbaumer, "BrainComputer InterfaceBased Communication in the Completely Locked-In State," *PLoS Biol*, 2017.

[2] J. M. O. Toole and G. B. Boylan, "NEURAL: quantitative features for newborn EEG using Matlab," *eprint arXiv:1704.05694*, 2017.

APPENDIX A
FEATURES SELECTED

Feature index	*Feature Name*	*Band Frequency*	*Channel*
10	Range Lower Margin	7:13Hz	FC5
13	Range Upper Margin	4:7Hz	FC5
17	Range Width	4:7Hz	FC5
21	Range SD	4:7Hz	FC5
38	Amplitude SD	7:13Hz	FC5
44	Kurtosis	0,5:4Hz	FC5
45	Kurtosis	4:7Hz	FC5
57	Power	4:7Hz	FC5
65	Flatness	4:7Hz	FC5
66	Flatness	7:13Hz	FC5
69	Entropy	4:7Hz	FC5
70	Entropy	7:13Hz	FC5
73	Edge Frequency	0,5:30 Hz	FC5
81	Range Median	13:30Hz	FC6
92	Range Width	7:13Hz	FC6
105	Range Assimetry	13:30Hz	FC6
109	Amplitude Power	13:30Hz	FC6
114	Skew	0,5:4Hz	FC6
117	Skew	13:30Hz	FC6
118	Kurtosis	0,5:4Hz	FC6
119	Kurtosis	4:7Hz	FC6
121	Kurtosis	13:30Hz	FC6
127	Envelope SD	4:7Hz	FC6
135	Relative power	4:7Hz	FC6
138	Flatness	0,5:4Hz	FC6
142	Entropy	0,5:4Hz	FC6
157	Range Lower Margin	4:7Hz	C5
160	Range Upper Margin	0,5:4Hz	C5
164	Range Width	0,5:4Hz	C5
168	Range SD	0,5:4Hz	C5
172	Range CV	0,5:4Hz	C5
191	Skew	13:30Hz	C5
192	Kurtosis	0,5:4Hz	C5
200	Envelope SD	0,5:4Hz	C5
212	Flatness	0,5:4Hz	C5
213	Flatness	4:7Hz	C5
216	Entropy	0,5:4Hz	C5
217	Entropy	4:7Hz	C5
227	Range Median	4:7Hz	C6
235	Range Upper Margin	4:7Hz	C6
239	Range Width	4:7Hz	C6
243	Range SD	4:7Hz	C6
247	Range CV	4:7Hz	C6
255	Amplitude Power	4:7Hz	C6
259	Amplitude SD	4:7Hz	C6
261a	Amplitude SD	13:30Hz	C6
262	Skew	0,5:4Hz	C6
271	Envelope Mean	4:7Hz	C6
275	Envelope SD	4:7Hz	C6
283	Relative power	4:7Hz	C6
286	Flatness	0,5:4Hz	C6
287	Flatness	4:7Hz	C6
290	Entropy	0,5:4Hz	C6
291	Entropy	4:7Hz	C6
303	Range Median	13:30Hz	Cz
311	Range Upper Margin	13:30Hz	Cz
315	Range Width	13:30Hz	Cz
319	Range SD	13:30Hz	Cz
339	Skew	13:30Hz	Cz
353	Power	4:7Hz	Cz
357	Relative power	4:7Hz	Cz

APPENDIX B
FEATURES LIST

Amplitude Features

Feature Name	*Description*	*FB*
Amplitude Total Power	time-domain signal: total power	yes
Amplitude SD	time-domain signal: standard deviation	yes
Skew	time-domain signal: skewness	yes
Kurtosis	time-domain signal: kurtosis	yes
Envelope Mean	envelope: mean value	yes
Envelope SD	envelope: standard deviation (SD)	yes

Range Features

Feature Name	*Description*	*FB*
Mean	range EEG: mean	yes
Median	range EEG: median	yes
Lower Margin	range EEG: lower margin (5th percentile)	yes
Upper Margin	range EEG: upper margin (95th percentile)	yes
Width	range EEG: upper margin - lower margin	yes
SD	range EEG: standard deviation	yes
CV	range EEG: coefficient of variation	yes
Asymmetry	range EEG: measure of skew about median	yes

Spectral Features

Feature Name	*Description*	*FB*
Spectral Power	spectral power: absolute	yes
Spectral Relative Power	spectral power: relative (normalised to total spectral power)	yes
Flatness	spectral entropy: Wiener (measure of spectral flatness)	yes
Entropy	spectral entropy: Shannon	yes
Difference	difference between consecutive short-time spectral estimates	no
Edge Frequency	cut-off frequency (fc): 95% of spectral power contained between 0.5 and fc Hz	no
FD	fractal dimension	yes

Connectivity Features

Feature Name	*Description*	*FB*
BSI	brain symmetry index (see Van Putten 2007)	yes
Correlation	correlation (Spearman) between envelopes of hemisphere-paired channels	yes
Coherence Mean	coherence: mean value	yes
Coherence Max	coherence: maximum value	yes

FB: features generated for each frequency band (FB)