Automated machine learning approach for the prediction of early Multiple Sclerosis from resting state fMRI connectivity data

V. Saccà[1], A. Sarica[2], F. Novellino[2], S. Barone[3], T. Tallarico[3], E. Filippelli[3], A. Granata[3], S. Polidoro[3], S. Scannapieco[3], G. Demonte[3], P. Valentino[3], A. Quattrone[2,3]

1. Department of Medical and Surgical Sciences, University "Magna Graecia" of Catanzaro, Italy
2. Institute of Bioimaging and Molecular Physiology (IBFM), National Research Council, Catanzaro, Italy,
3. Institute of Neurology, University Magna Graecia of Catanzaro, Italy

Introduction

This work aimed at evaluating whether Machine Learning technique could support early diagnosis of Multiple Sclerosis (MS) from resting-state-functional (rs-fMRI) connectivity data. In particular, we explored the ability in distinguishing between controls and patients of mean signals extracted from ICA-components corresponding to 15 well-known networks, using Support Vector Machine (SVM).

Methods

Eighteen patients with early-MS (mean-age 37.42±8.11, 9 females) were recruited according to McDonald and Polman criteria [1], and matched for demographic variables with 19 healthy controls (mean-age 37.55±14.76, 10 females). All rsFMRI images were acquired and were pre-processed using the FMRIB's Software Library v5.0. Then, we applied independent component analysis (ICA) with the GIFT-toolbox and fifteen mean components were visually identified by two experts (Fig.1).
Machine Learning analysis were conducted on this dataset composed of 37 rows (subjects) and 15 features (mean signal in the network) with R-language. The dataset was randomly split into training (75%) and test-sets and Radial-SVM [2] was built.
The feature selection was performed with Recursive Feature Elimination (rfe) and feature weights were calculated. Thus, we trained a classifier on the most important features and we evaluated the accuracies (with and without feature selection) on test-set.

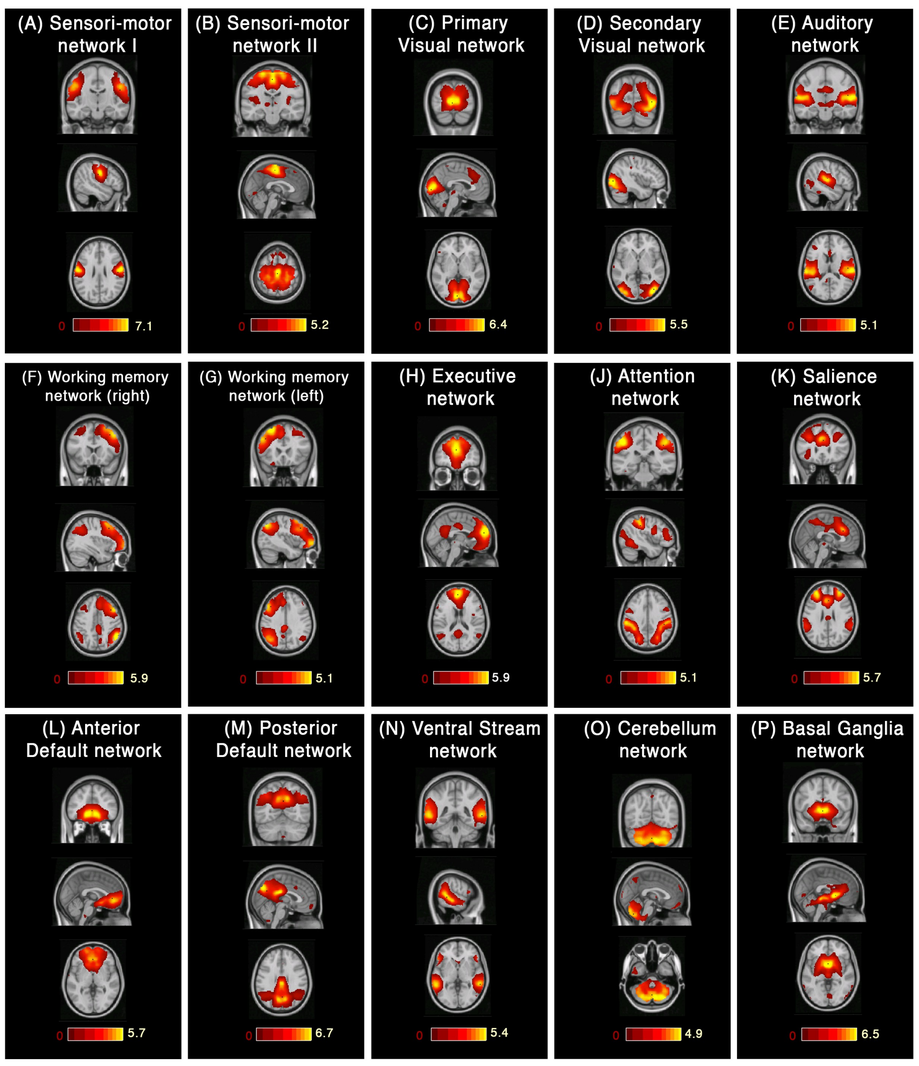

Fig.1: The 15 networks generated by GIFT. The color bars show z-score.

Results

The rfe showed that the most important variable was the sensorimotor I (Fig.2). With only this network, SVM-classifier reached an accuracy of 87.5% on test-set, while, trained on all the features, it presented poor accuracies (train:65.52% and test-sets=50%). The accuracies are reported in Fig.3. To confirm the SVM results, we have created also a Random-Forest classifier, which showed the same ranking of features importance and accuracy on the test-set.

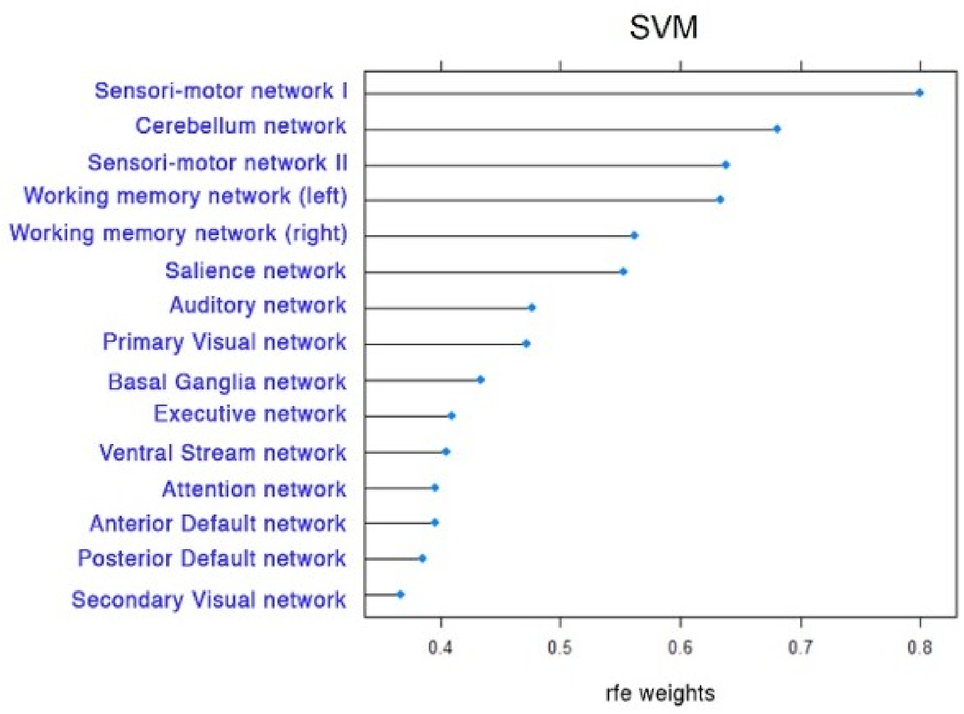

Fig.2: rfe applied on the SVM. The most important network is the sensori-motor I.

Conclusions

The use of resting-state fMRI as a biomarker of early disease recognition and as predictor of motor and cognitive function in MS has aroused growing interest in recent years.
Our results suggest the view that the study of brain functional connectivity changes by fMRI holds great promise to better recognize the networks reorganization also occurring at an early stage of the disease. According to early manifestation of motor/sensorial deficits in MS, in our study the sensorimotor I network was the more accurate variable for a correct recognition of early-MS subjects. These findings could represent an encouraging step toward the translation to rfMRI connectivity evaluation into clinical practice.

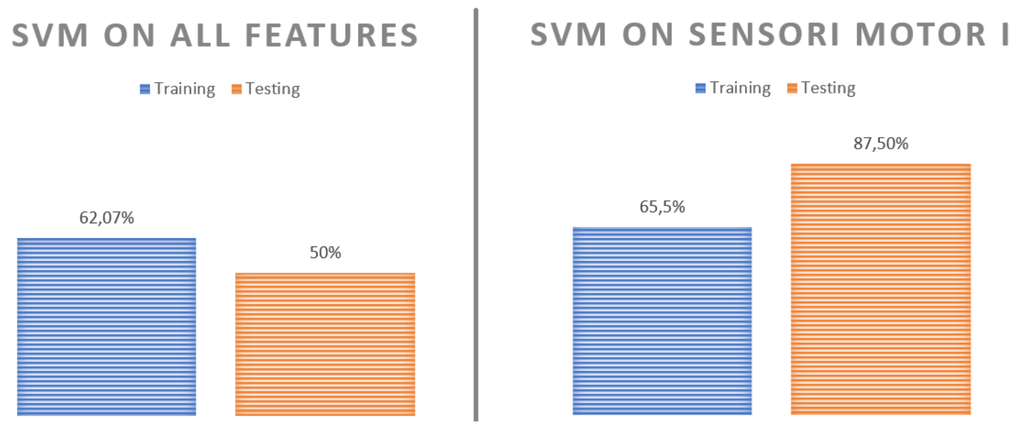

Fig.3: Accuracies obtained with all features and with only the sensori motor I..

Bibliography

[1] Polman, C. H., Reingold, S. C., Banwell, B., Clanet, M., Cohen, J. A., Filippi, M., et al., (2011). Diagnostic criteria for multiple sclerosis: 2010 revisions to the McDonald criteria. Annals of neurology, 69(2), 292-302.
[2] Bhuvaneswari, P and Kumar, J Satheesh, Support Vector Machine Technique for EEG Signals, International Journal of Computer Applications Volume 63 N.13, 2013