Real Competition between feature selection methods for discriminating AD from MCI

A Sarica[1], A Cerasa[1], A Quattrone[2] and for the Alzheimer's Disease Neuroimaging Initiative***

[1] IBFM, Neuroimaging Unit, National Research Council, Catanzaro, Italy.
[2] Institute of Neurology, University *Magna Graecia*, Germaneto (CZ), Italy.

Objective

- A major challenge in Neuroscience is the discovery of the best subset of biomarkers that could improve the accuracy in discriminating Alzheimer's disease (AD) from the Mild Cognitive Impairment (MCI), independently from the acquisition methods, the scanners, the pre-processing techniques and the software used for this purpose[1].

> **This study proposes a fully automatic feature selection approach for reducing high dimensionality of structural features extracted from MRI[2,3].**

Materials

> The ADNI** dataset (Table 1) includes those subjects diagnosed AD and late MCI, which MRI:

- are acquired by *3T* scanners;
- *Non-Accelerated T1* scans related to the baseline visit;
- completely passed the quality test;

Table 1. Descriptive statistics of ADNI training set. Age values (years) are *mean±standard deviation* and include both female and male subjects.

Class	*Nr. of subjects*	*Female %*	*Age (mean ± std)*
AD	70	52.8%	74.15±8.07
MCI	70	45.7%	72.6±7.78

> Feature extracted from MRIs by FreeSurfer[4]:

- 45 volumes of subcortical structures;
- 34 mean thickness and 34 cortical volumes for each hemisphere;
- 8 hippocampus volume subfields for each hemisphere;

for a total of **200 attributes** including the diagnosis, gender and age.

Methods

> Feature selection[5] can mitigate computational performance issues and improve the classification accuracy.

> The adopted workflow is composed by four steps: (i) *IntraCranial Volume* normalization; (ii) *Feature Selection* with three techniques; (iii) *Z-Score normalization*; (iv) binary classification with Support Vector Machines.

> The core of the proposed method consists in sequentially applying a Correlation filter (0.90), a Random Forest (0.5) filter and a Support Vector Machines (SVM Radial kernel) wrapper on the training dataset, to identify a subset of features that provides the highest binary classification accuracy.

Results

> The classifier that showed the highest 10-fold cross validation accuracy, which training subsets have the lowest number of features and that have been generated with lowest steps, is that with Correlation filter+ Random Forest filter (63.6%).

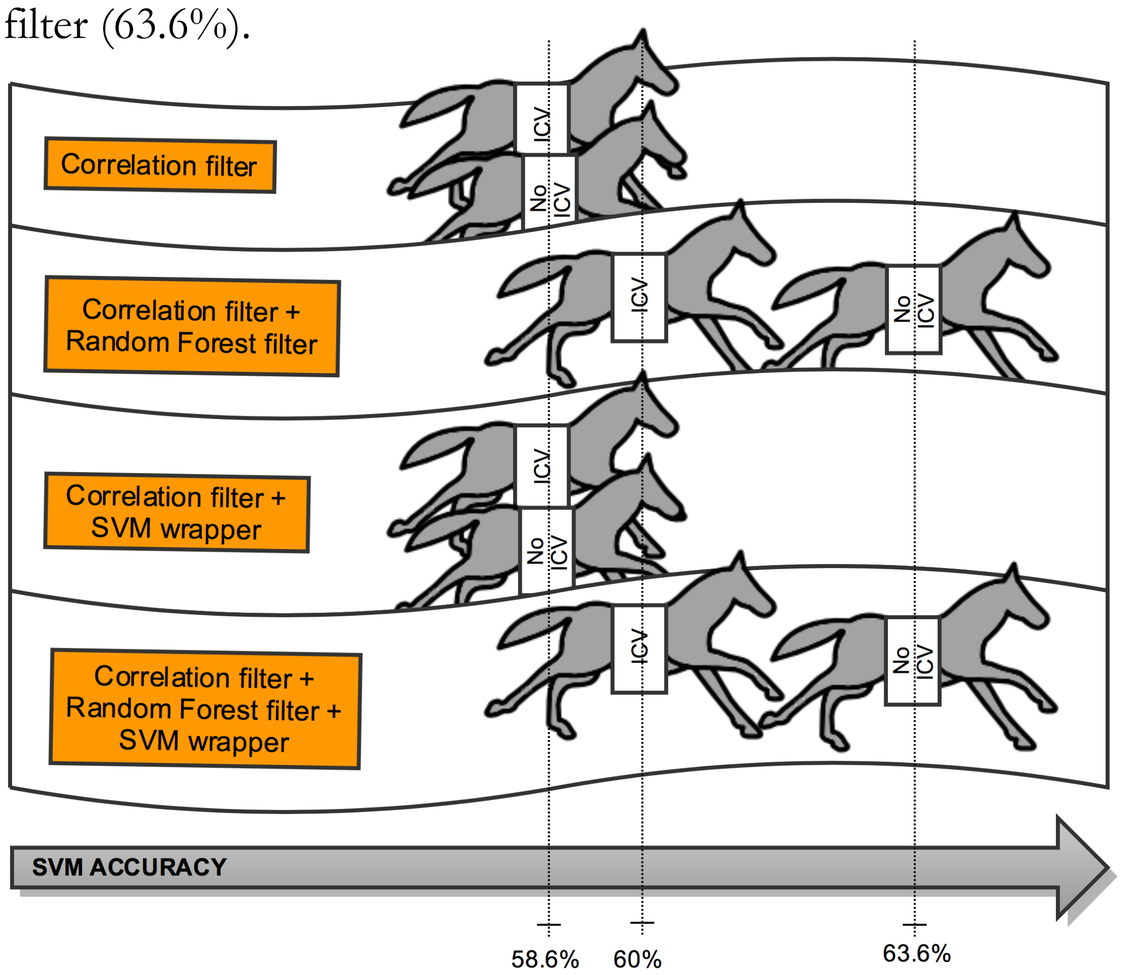

Fig. 1 Graphical representation of the results. Each feature selection technique is applied on datasets with (ICV) and without (No ICV) *IntraCranial Volume* normalization.

Conclusions

> We showed that our new advanced feature selection algorithm could support the study of Dementia by automatically finding the best subset of volumetric features that discriminate AD from MCI.

> The findings suggest that the ICV normalization worse accuracy and that the best feature selection approach is to sequentially use Correlation and Random Forest filters.

References

1. http://caddementia.grand-challenge.org
2. Sarica, A., Di Fatta, G., & Cannataro. Advanced Feature Selection in Multinominal Dementia Classification from Structural MRI Data. In: CADDementia workshop, Medical Image Computing and Computer Assisted Intervention (MICCAI) 2014 conference, 14-18 Sep 2014, Boston.
3. Bron, E. E., Smits., ..., Sarica, A. & Klein, S. (2015). Standardized evaluation of algorithms for computer-aided diagnosis of dementia based on structural MRI: The CADDementia challenge. NeuroImage, 111, 562-579.
4. Anders M Dale, Bruce Fischl, and Martin I Sereno. Cortical surface-based analysis: I. segmentation and surface reconstruction. Neuroimage, 9(2):179–194, 1999.
5. Isabelle Guyon and Andre Elisseeff. An introduction to variable and feature selection. J. Mach. Learn. Res., 3:1157–1182, March 2003.

* sarica@unicz.it

**Data used in preparation of this article were obtained from the Alzheimer's Disease Neuroimaging Initiative (ADNI) database (adni.loni.usc.edu). As such, the inves- tigators within the ADNI contributed to the design and implementation of ADNI and/or provided data but did not participate in analysis or writing of this report. A complete listing of ADNI investigators can be found at: http://adni.loni.usc.edu/wp- content/ uploads/how to apply/ADNI Acknowledgement List.pdf.

XLVI CONGRESSO NAZIONALE
10-13 OTTOBRE 2015 – GENOVA
Sin