MICCAI 2014 BOSTON - Challenge on Computer-Aided Diagnosis of Dementia based on structural MRI data

Advanced Feature Selection in Multinominal Dementia Classification from Structural MRI Data

Alessia Sarica[1§], Giuseppe Di Fatta[2], Garry Mark Smith[2,3], Mario Cannataro[1] and Douglas Saddy[3] and for the Alzheimer's Disease Neuroimaging Initiative[**]

[1]*Department of Surgical and Medical Sciences, Magna Graecia University of Catanzaro, Italy.* [2]*School of Systems Engineering, University of Reading, UK.* [3]*Centre for Integrative Neuroscience and Neurodynamics (CINN), University of Reading, UK.*

Introduction

❑ **A major challenge in Neuroscience[1] is the discovery of the best subset of biomarkers that could improve the accuracy in discriminating Alzheimer's disease from the Mild Cognitive Impairment.**

MRI Datasets and Feature Generation

❑ Feature extraction from MRIs performed by FreeSurfer[2]:

- 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.

❑ The ADNI dataset (D2, Table 1) includes those subjects:

- diagnosed as healthy control (HC), Alzheimer's disease (AD) and late Mild Cognitive Impairment (late MCI);
- that completely passed the quality test;
- that have *Non-Accelerated T1* scans related to the baseline visit.

Table 1. Descriptive statistics of the CADDementia and ADNI data sets. Age values (years) are mean±standard deviation and include both female and male subjects.

Description	*Name*	*Class*	*Nr. of subjects*	*Female%*	*Age (mean±std)*
CADDementia training set	D1	HC	12	25%	62.33±6.26
		AD	9	66.6%	66.11±5.21
		MCI	9	44.4%	68±8.54
ADNI	D2	HC	70	48.6%	73.6±5.49
		AD	70	52.8%	74.15±8.07
		MCI	70	45.7%	72.6±7.78

Feature Selection and Classification Model Inference

❑ **Feature selection[3] can reduce dimensionality, thus mitigating computational performance issues and improving the classification accuracy.**

❑ The adopted workflow is composed by five steps (Fig. 1): (i) *IntraCranial Volume* normalization; (ii) *Feature Selection* with three techniques; (iii) *Z-Score normalization*; (iv) binary classification and (v) multi-class classification (*one-versus-one*).

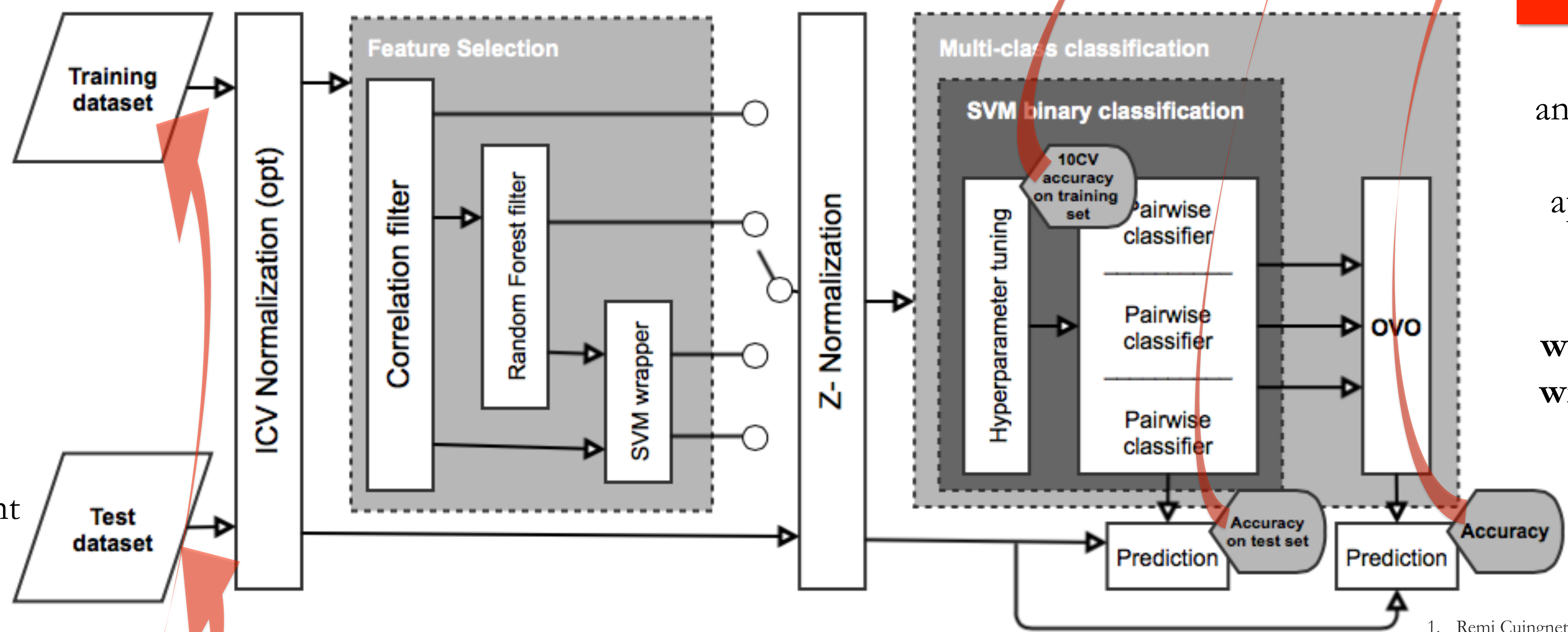

Fig. 1 Diagram of the adopted workflow based on a combination of feature selection techniques.

❑ **The core of the proposed method consists in sequentially applying a Correlation filter, a Random Forest (RF) filter and a Support Vector Machines (SVM) wrapper on the training dataset**, to identify a subset of features that provides the highest binary classification accuracy (Table 2 and 3).

❑ The three best binary classification models have been used for performing the multi-class classification.

Results

Table 2. Results of the binary classification: training performed on D2 and accuracy estimated on D2 (D2/D2) and D1 (D2/D1). Best results on D1 are indicated in bold.

	Feature selection method																							
	Corr. filter $|r|<0.90$						*Random Forest filter*						*SVM wrapper*						*Random Forest filter + SVM wrapper*					
	ICV			*No ICV*			*ICV*			*No ICV*			*ICV*			*No ICV*			*ICV*			*No ICV*		
	Nr. feat.	D2/D2 10xval	D2/D1 holdout	Nr. feat.	D2/D2 10xval	D2/D1 holdout	Nr. feat.	D2/D2 10xval	D2/D1 holdout	Nr. feat.	D2/D2 10xval	D2/D1 holdout	Nr. feat.	D2/D2 10xval	D2/D1 holdout	Nr. feat.	D2/D2 10xval	D2/D1 holdout	Nr. feat.	D2/D2 10xval	D2/D1 holdout	Nr. feat.	D2/D2 10xval	D2/D1 holdout
HCvsMCI	133	71.4%	42.9%	140	67.9%	**66.7%**	71	71.4%	42.9%	41	65.7%	52.4%	133	71.4%	42.9%	140	67.9%	66.7%	30	77.1%	42.9%	41	65.7%	52.4%
HCvsAD	133	90%	42.9%	140	81.4%	95.2%	109	86.4%	42.9%	95	82.1%	**95.2%**	133	90%	42.9%	140	81.4%	95.2%	109	86.4%	42.9%	95	82.1%	95.2%
ADvsMCI	133	58.6%	50%	139	58.6%	83.3%	47	60%	50%	44	62.1%	**88.9%**	133	58.6%	50%	139	58.6%	83.3%	47	60%	50%	44	62.1%	88.9%

Table 3. Results of multi-class classification. The VOTE strategy for the *one-versus-one* classification has been applied on datasets with and without IntraCranial Volume normalization.

	OVO accuracy
ICV normalization	30%
No ICV normalization	**70%**

Table 4. Mean computation time for each step of the proposed approach.

Step	*Time*
ROI feature extraction	5 hours per subject
Correlation filter	0.836 s
Random Forest filter	7.744 s
SVM wrapper	112.790 s
RF filter + SVM wrapper	91.790 s
OVO classification	0.496 s

Conclusions

❑ The present study was designed for exploring and evaluating alternative subsets of ROI features extracted by FreeSurfer from brain MRIs, by applying advanced methods for feature selection.

❑ **The findings suggest that (i) ICV normalization can lead to overfitting and worse accuracy and (ii) even though the SVM wrapper is more complex and slower than the Random Forest filter, it does not provide better accuracy.**

References

1. Remi Cuingnet, Emilie Gerardin, Jerome Tessieras, Guillaume Auzias, Stephane Lehericy, Marie-Odile Habert, Marie Chupin, Habib Benali, and Olivier Colliot. Automatic classification of patients with Alzheimer's disease from structural MRI: A comparison of ten methods using the ADNI database. NeuroImage, 56(2):766 – 781, 2011. Multivariate Decoding and Brain Reading.
2. Anders M Dale, Bruce Fischl, and Martin I Sereno. Cortical surface-based analysis: I. segmentation and surface reconstruction. Neuroimage, 9(2):179–194, 1999.
3. Isabelle Guyon and Andre Elisseeff. An introduction to variable and feature selection. J. Mach. Learn. Res., 3:1157–1182, March 2003.

Acknowledgements

§ sarica@unicz.it - This poster was co-funded by the European Commission, European Social Fund and "Regione Calabria". Authors had full control of the design of the study, methods used, outcome parameters and results, analysis of data and production of the manuscript.

** 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.

UMG Dulcium sapientiae initium · University of Reading · CINN Centre for Integrative Neuroscience and Neurodynamics