ALZHEIMER'S DIAGNOSIS USING PCA AND SUPPORT VECTOR MACHINES IN VOIS APPLIED ON THE ADNI DATABASE

Tayebeh Razmi

Instituto de Telecomunicações (IT), Instituto Superior Técnico, Lisbon, Portugal

t.razmi@lx.it.pt

Abstract

IN this work a computer aided diagnosis tool is presented for the the study of Alzheimer's Disease based on PET images. The proposed tool applies Principal Component analysis (PCA) to the voxels contained in small regions of interest selected by a physician, in order to reduce dimensionality. Subsequently, classification is performed using Support Vector Machines (SVM). With the combination of these methods we archived accuracy results of up to 99.33% for AD vs. Normal, 98.44% in AD vs. MCI and 98.93% for Normal vs. MCI group.

1. PCA& Eigenbrains

- Principal component analysis (PCA) maximizes the scatter of all the projected samples.
- Mean subtraction

  $y_i = X_i - \mu$
  $X = [X_1, X_2, ..., X_m]$: set of voxls vectors,
  m: number of patients,
  μ: average brain of the set.
- The covariance matrix of the normalized vectors set

$$C = \frac{1}{m}\sum_{i=1}^{m} y_i y_i^t = \frac{1}{m} Y Y^t \quad (1)$$
$$Y = [y_1, y_2, ..., y_m]$$

- The eigenvector (Ψ) and eigenvalue (Λ) matrices

$$C\Psi = \Lambda\Psi \quad (2)$$

- Because the sample size m is much smaller than the dimensionality n, so diagonalizing Y^tY instead of YY^t reduces the computational complexity [6]
- Denoting $\hat{v}_i$ as the set of eigenvectors with associated eigenvalues λ_i for matrix Y^tY.

$$(Y^tY)\hat{v}_i = \lambda_i \hat{v}_i \quad (3)$$

- $YY^t\Psi = \Lambda\Psi$, $\Lambda = diag\{\Lambda_1, \Lambda_2, ..., \Lambda_m\}$ are eigenvalues $\Psi = Y\hat{v}_i$, $\Psi = [\Psi_1, \Psi_2, ..., \Psi_m]$ are eigenvectors of YY^t [5].
- Derived from the *eigenface* concept[6], the *eigenbrains* correspond to the dominant eigenvectors of the brain covariance matrix.
- Only k leading eigenvectors are used, which define the matrix U_{reduce}

$$U_{reduce} = [\Psi_1, \Psi_2, ..., \Psi_k] \quad (4)$$

- For choosing k (number of principal components) we keep 99% of the variance

$$\frac{\sum_{i=1}^{k}\Lambda_i}{\sum_{i=1}^{m}\Lambda_i} \geq 0.99 \quad (5)$$

- Reconstruction of the original data let $Z = U_{reduce}^t Y$, where Z maps y_i to Z_i. The desired Y_{approx} which is the closest approximation to Y is:

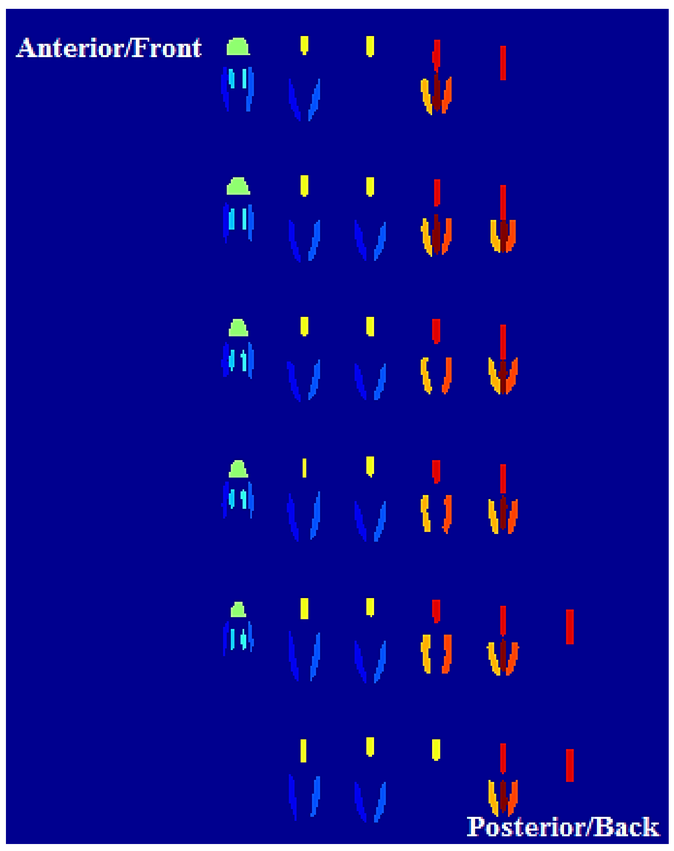
$$Y_{approx} = U_{reduce} Z \quad (6)$$

2. Volume of Interest (VOI)

- One of the approaches which is implemented here is the definition of Volume of interest (VOI).
- One physician with great experience in the clinical use of FDG-PET was asked to identify the regions.
- The physician identified seven distinct regions, with some of them having symmetrical counterparts.
- These regions are
  - Lateral temporal (right and left)
  - Mesial temporal (right and left)
  - Inferior frontal gyrus/Orbitofrontal
  - Inferior anterior cingulate
  - Superior anterior cingulate
  - Dorsolateral parietal (right and left)
  - Posterior cingulate and precuneus
- These regions were manually segmented in the original image volume reference space.
- The voxels inside the segmented VOIs as forming the set A, one ends up with an image volume where for each voxel v_{ijk} [1]:

$$v_{ijk} = \begin{cases} 1 & \text{if } v_{ijk} \in A \\ 0 & \text{otherwise} \end{cases}$$

- The voxels which are inside of the VOIs are considered for applying PCA.
- Only k leading eigenvectors, which we refered them as eigenlobes are used as VOIs descriptors.
- Defined region for different slices of brain.



3. Classification Using Linear SVM

- SVMs (Support Vector Machines) are useful technique for data classification.
- A classification task usually involves separating data into training and testing sets.
- Each instance in the training set contains one "target value" (i.e. the class labels) and several "attributes" (i.e. the features or observed variables).
- SVM finds a linear separating hyperplane with the maximal margin in this higher dimensional space. $C > 0$ is the penalty parameter of the error term.
- In this work we have used linear kernel function which is $K(x_i, x_j) = x_i^T x_j$.
- For this work, LIBSVM which is more Accurate and LIBLINEAR which is faster are used, libraries developed by C.Chang and C. Lin [2], [3].
- To assess the quality of the training result, we performed a 10-fold cross-validation on the training data.

4. ADNI Database

- The data herein utilized to test the several implemented classifiers were taken from the ADNI database.
- AD subjects were studied at months 0, 6, 12 and 24, MCI subjects at months 0, 6, 12, 18, 24 and 36, and CN patients were studied at months 0, 6, 12, 24 and 36.
- Every FDG-PET volume had been subject to a preprocessing procedure, with format, orientation and resolution uniformization purposes. At the end of this process, the volumes were $128 \times 128 \times 60$ voxels wide.
- The clinical dementia rating (CDR) score was used as eligibility criterion for patients' selection: 0 for CN, 0.5 for MCI and 0.5 or more for AD patients had to be met.
- Clinical profile of the groups studied. Values are presented according to (mean$\pm$ standard deviation), MMSE (Mini-Mental State Examination) and CDR.

Group	AD	MCI	Normal
Number of patients	95	207	104
Number of PET scans	314	971	435
CDR	0.5 or higher	0.5	0
MMSE	19.6$\pm$ 5.0	26.3$\pm$ 3.1	29.2$\pm$ 0.9

5. Results

For analysis and comparison, the data is arranged into three different groups:

- Group 1: Only AD (positive) and Normal controls(CN) (negative) patient images are considered.
- Group 2: Only AD (positive) and MCI (negative) patient images are considered. which is very relevant in clinical practice (MCI patients are at higher risk of converting to AD)
- Group 3: Only MCI (positive) and Normal controls (negative) patient images are considered. The most difficult classification task concerning ADNI database is to distinguish between NORMAL and MCI patients, due to the wide range spanned by the features extracted from MCI patients [4]

Results remarks:

- After PCA is applied over the database the eigenlobes are obtained. On average only 323 eigenlobes were necessary to explain the 99% of the variance retained.
- For each groups, 10 percent of the data are selected randomly as a test set and the remaining as a training set. Then we extracted the eigenlobes based on their training set
- Results obtained from LIBSVM on the three groups:

Group	# Eigenlobes	Accuracy (%)	F1-score (%)
AD vs. CN	323	99.33	99.30
AD vs. MCI	473	98.44	96.44
CN vs. MCI	522	98.93	99.22

- Results obtained from LIBLINEAR on the three groups:

Group	# Eigenlobes	Accuracy (%)	F1-score (%)
AD vs. CN	326	98.67	98.25
AD vs. MCI	476	94.55	88.71
CN vs. MCI	527	98.58	98.99

6. Conclusions

- We presented a methodology combining the use of PCA with SVM-based classification of PET images, using only a set of VOIs, as determined by an expert in the field.
- 99.33% accuracy in identifying AD vs. CN, 98.44% AD vs. MCI and 98.93% MCI vs. CN.
- Future work will include MCI to AD Conversion for Long-term applied on the ADNI Database.

References

[1] Eduardo Bicacro. Master thesis: Alzheimer's disease diagnosis using 3d brain images, November 2011.

[2] Chih-Chung Chang and Chih-Jen Lin. LIBSVM: A library for support vector machines. *ACM Transactions on Intelligent Systems and Technology*, 2:27:1–27:27, 2011. Software available at http://www.csie.ntu.edu.tw/~cjlin/libsvm.

[3] Rong-En Fan, Kai-Wei Chang, Cho-Jui Hsieh, Xiang-Rui Wang, and Chih-Jen Lin. LIBLINEAR: A library for large linear classification. *Journal of Machine Learning Research*, 9:1871–1874, 2008.

[4] M. López, J. Ramírez, J. M. Górriz, I. Álvarez, D. Salas-Gonzalez, F. Segovia, R. Chaves, P. Padilla, and M. Gómez-Río. Principal component analysis-based techniques and supervised classification schemes for the early detection of alzheimer's disease. *Neurocomput.*, 74:1260–1271, March 2011.

[5] Jonathon Shlens. A tutorial on principal component analysis. In *Systems Neurobiology Laboratory, Salk Institute for Biological Studies*, 2005.

[6] Matthew Turk and Alex Pentland. Eigenfaces for recognition. *J. Cognitive Neuroscience*, 3:71–86, January 1991.

The author benefited from a scholarship of Fundaçõ para a Ciência e Tecnologia (FCT).The author also would like to thank Dr. Durval C. Costa and Instituto de Telecomunicações (IT) for their consulting and collaborating.