Automatic tool for Alzheimer’s disease diagnosis using PCA and Bayesian classification rules


An automatic tool to assist the interpretation of single photon emission computed tomography (SPECT) and positron emission tomography (PET) images for the diagnosis of the Alzheimer’s disease (AD) is demonstrated. The main problem to be handled is the so-called small size sample, which consists of having a small number of available images compared to the large number of features. This problem is faced by intensively reducing the dimension of the feature space by means of principal component analysis (PCA). Our approach is based on Bayesian classifiers, which uses a posteriori information to determine in which class the subject belongs, yielding 88.6 and 98.3% accuracy values for SPECT and PET images, respectively. These results mean an improvement over the accuracy values reached by other existing techniques.

Introduction: Single photon emission computed tomography (SPECT) and positron emission tomography (PET) provide valuable clinical information relevant to Alzheimer’s disease (AD) regarding regional cerebral blood flow or metabolic activity in the brain rather than imaging anatomical structures. The evaluation of these images is usually done through visual ratings performed by experts and other subjective steps which are time-consuming and prone to error.

A fully computer-aided diagnosis (CAD) system for the early detection of AD is shown in this Letter. For the classification task, a previous approach considered the use of the intensity values of the whole brain image as features (also known as voxels-as-features (VAF)) to apply statistical learning methods [1], giving rise to high-dimensional feature vectors. In our approach, this feature space dimension is reduced by applying the principal component analysis (PCA) transformation. The resultant data are used to make up a Bayesian classifier which makes use of the a posteriori information to classify the coming images of a new patient.

PCA and eigenbrains: A reconstructed and spatially normalised 3D voxel representation of the brain of each subject is rearranged in a vector form. Let \( Y = \{Y_1, Y_2, \ldots, Y_n\} \), where \( Y_i = (y_{i1}, y_{i2}, \ldots, y_{iN}) \), \( i = 1, 2, \ldots, n \) where \( n \) is the number of patients, be the sample of these vectors after normalising them to unity norm and subtracting the grand mean. The covariance matrix of the normalised vectors set is defined as:

\[
\Sigma_Y = \frac{1}{n} \sum_{i=1}^{n} Y_i Y'_i = YY'
\]  

(1)

and the eigenvector and eigenvalue matrices \( \Phi, \Delta \), are computed as:

\[
\Sigma_Y \Phi = \Phi \Delta
\]  

(2)

Note that \( YY' \) is an \( N \times N \) matrix while \( Y'Y \) is an \( n \times n \) matrix. If the sample size \( n \) is much smaller than the dimensionality \( N \), then diagonalising \( YY' \) instead of \( Y'Y \) reduces the computational complexity [2]:

\[
(Y'Y) \Psi = \Psi \Delta_1
\]  

(3)

\[
T = Y' \Psi
\]  

(4)

where \( \Delta_1 = \text{diag}(\lambda_1, \lambda_2, \ldots, \lambda_m) \) and \( T = [\Phi_1, \Phi_2, \ldots, \Phi_m] \). Derived from the eigenface concept [2], the eigenvectors correspond to the dominant eigenvectors of the brain covariance matrix. In this approach, only \( m \) leading eigenvectors are used, which define the matrix \( P \):

\[
P(\Phi_1, \Phi_2, \ldots, \Phi_m)
\]  

(5)

The criterion to choose the most discriminant eigenbrains is set by their separation ability, which is measured by their Fisher discriminant ratio (FDR), defined as:

\[
FDR = \frac{(\mu_1 - \mu_2)^2}{\sigma_1^2 + \sigma_2^2}
\]  

(6)

where \( \mu_i \) and \( \sigma_i^2 \) denote the \( i \)th class mean value and variance, respectively. Fig. 1 represents the aspect of the three main eigenbrains of three slices along the transaxial axis. For the classification task we project each patient vector into the previously defined eigenbrain space, producing a vector of weights. For the whole database, a matrix of weights can be constructed, given by

\[
Z = P^T Y
\]  

(7)


Fig. 1 Three representative transaxial slices for first three eigenbrains, ranked by their eigenvalues

Bayes classifier: The Bayes classifier evaluates the a posteriori probability function [3]. Let \( \omega_i \) and \( \omega_2 \) denote the object classes (AD and NORMAL), and \( Z \) a patient voxels vector in the reduced PCA subspace. The a posteriori probability function of \( \omega_i \) given \( Z \) is defined as

\[
P(\omega_i|Z) = \frac{p(Z|\omega_i)P(\omega_i)}{p(Z)} \quad i = 1, 2
\]  

(8)

where \( P(\omega_i) \) is a priori probability, \( p(Z|\omega_i) \) the conditional probability density function of \( Z \) given \( \omega_i \) and \( p(Z) \) is the mixture density. The maximum a posteriori (MAP) decision rule for the Bayes classifier is defined as

\[
p(Z|\omega_i)P(\omega_i) = \max_j [p(Z|\omega_j)P(\omega_j)] \quad Z \in \omega_i
\]  

(9)

The brain projected data \( Z \) is classified to \( \omega_i \) of which the a posteriori probability given \( Z \) is the largest between the classes. The within-class densities are usually modelled as normal distributions

\[
P(\omega_i|Z) = \frac{1}{(2\pi)^{N/2} |\Sigma_i|^{1/2}} \exp \left\{ -\frac{1}{2} (Z - M_i)^T \Sigma_i^{-1} (Z - M_i) \right\}
\]  

(10)

where \( M_i \) and \( \Sigma_i \) are the mean and covariance matrix of class \( \omega_i \), respectively, and their expressions are

\[
M_i = \frac{1}{N_i} \sum_{j=1}^{N_i} Z_i^{(j)}, \quad i = 1, 2
\]  

(11)

where \( Z_i^{(j)}, j = 1, 2, \ldots, N_i \) represents the sample voxels vector for the \( \omega_i \) class. Under the probabilistic reasoning model 1 (PMR-1) derived in [4] the covariance matrices are identical, diagonal, and each diagonal element is estimated by the sample variance in the one dimensional PCA subspace:

\[
\Sigma_i = \text{diag}(\sigma_1^2, \sigma_2^2, \ldots, \sigma_m^2)
\]  

(12)

\[
\sigma_i^2 = \frac{1}{N_i - 1} \sum_{j=1}^{N_i} (Z_i^{(j)} - M_i)^2
\]  

(13)

where \( Z_i^{(j)} \) is the \( j \)th element of the sample \( Z_i^{(j)} \), \( m_i \) the \( j \)th element of \( M_i \), and \( L \) the number of classes (two in this approach).

Experiments: The experiments were performed over two databases consisting of a set of SPECT and PET images, taken with a PRISM 3000 machine and a SIEMENS ECAT 47, respectively. Initially they were
labelled by experienced clinicians of the ‘Virgen de las Nieves’ hospital (Granada, Spain) and ‘Clinica PET Cartuja’ (Seville, Spain), respectively. The SPECT database consists of 79 patients (41 labelled as NORMAL and 38 labelled as AD). Regarding PET patients, 60 images were provided (18 NORMAL and 42 AD). Both the SPECT and PET images are spatially normalised using the SPM software [5]. Initially, the original brain image 79 × 95 × 69 voxel size is reduced by averaging over subsets of 4 × 4 × 4 voxels. For each database, a mask is applied in order to select the voxels of interest and discard the rest, as proposed in [6]. After that, PCA is applied over the remaining voxels of the training set and eigenbrains are obtained. The new patient to be classified is projected into the eigenbrain space, and the MAP rule in (9) is applied. Fig. 2 represents the three main PCA coefficients as 3D points for NORMAL and AD patients. Fig. 3 shows the accuracy values obtained for SPECT and PET images when the number of considered principal components $m$ increases. These values were obtained when the leave-one-out cross-validation strategy was applied.

Fig. 2 For PET images, distributions of three first principal coefficients for AD and NORMAL patients and decision surface

Fig. 3 Accuracy for SPECT and PET images using Bayesian classifiers when number $m$ of considered principal components increases

Conclusions: A computer-aided diagnosis system for the early detection of Alzheimer’s disease is presented. The system was developed by performing PCA to a subset of remaining voxels after some preprocessing steps, which drastically reduces the feature space dimension. The most important components in terms of ability to separate are chosen to make up a Bayesian classifier that classifies a new patient using the a posteriori information. With this approach, 88.6 and 98.3% accuracy values are obtained for SPECT and PET images, respectively. These accuracy values exceed the results obtained by VAF (78.5 and 96.7% for SPECT and PET images, respectively), which was taken as reference work.

Acknowledgment: This work was partly supported by the Spanish Government under the PETRI DENCLASES (PET2006-0253), TEC2008-02113, NAPOLÉON (TEC2007-68030-C02-01) projects and the Consejería de Innovación, Ciencia y Empresa (Junta de Andalucía, Spain) under the Excellence Project (TIC-02566).

© The Institution of Engineering and Technology 2009
19 January 2009
doi: 10.1049/el.2009.0176

M. López, J. Ramírez, J.M. Górriz, D. Salas-Gonzalez, I. Álvarez and F. Segovia (Departamento Teoría de la Señal, Telemática y Comunicaciones, Universidad Granada, Spain)
E-mail: miriamlp@ugr.es
C.G. Puntonet (Departamento de Arquitectura y Tecnología de Computadores, Universidad Granada, Spain)

References