SEGMENTATION OF NEURO MR IMAGES THROUGH SEMI-SUPERVISED LEARNING

Vladimir H. Kanchev[1], Ivo R. Draganov[2], Antoaneta A. Popova[3]*

[1,2,3]Radiocommunications and Videotechnologies Dept.
Technical University - Sofia
8 Kliment Ohridski Blvd., 1000 Sofia, Bulgaria
Tel.:+359 2 965 2274, E-mail: { v_kanchev, idraganov, antoaneta.popova}@tu-sofia.bg

Abstract

In this paper an algorithm is presented for segmentation of degraded and healthy tissue from neuro images of patients who suffered from AD, MCI, and healthy ones. We construct a model of 3D region mask and then imposed it on the test neuro images for segmentation. During training neuro images of AD and healthy patients are subjected to the following operations: pre-processing, brain extraction, image registration to an atlas, tissue segmentation and then classifier is applied in order to create a 3D mask. During segmentation white, grey matter and CSF are separated by Fuzzy C-means clustering. After that we perform feature extraction, estimation of statistics of 3D mask region in the test image.

Key-words: *segmentation of neuro images, semi-supervised learning, Alzheimer's disease, clustering*

1. Introduction

Alzheimer's disease (AD) is an incurable, degenerative and terminal disease that affects more than 5 million Americans now and it is estimated that 13 million elderly will be diagnosed by AD in the 2050 year in the US alone [1].

We should first distinguish AD, Mild Cognitive Impairment (MCI), and dementia. AD is a progressive neurodegenerative disorder associated with a disruption of neuronal function and a gradual deterioration in cognition, function, and behaviour [2], MCI describes and early, but abnormal state of cognitive impairment in which people continues to do well in their daily activities [3]. Dementia is a serious loss of cognitive ability in a person, beyond what is expected from their age.

AD is pathologically characterized by presence of amyloid deposition and neurofibrillary tangles, together with the loss of cortical neurons and synapses – it starts firstly in the entorhinal cortex and hippocampus, then to temporal lobe, parietal lobe and finally into parts of the frontal cortex and cingulated gyrus [4]. It must be noticed that cognitive impairment of normal ageing – deficits in memory as an example, is concerned with reversible synaptic alteration, not neuron death [5].

Often AD goes unrecognized because it can be definitely diagnosed after patient's death, when the brain can be closely examined for certain microscopic changes caused by the disease. For diagnose of AD, are used patient's history, additional information of relatives, clinical observation with neuro - psychological tests as Mini-Mental State Examination (MMSE).

Since MMSE will provide normal results for a patient at early stages of AD, medical visualization techniques, as MRI, fMRI and PET, are employed additionally. Cognitive test results depend on comfort of patient with testing, his or her general health, used medication, fatigue and so on.

There are two underlying approaches for visualisation: in first case during SPECT and PET neuro imaging pharmaceutical compound as carbon-11, fluorine-8 are used for tracing of A-beta deposits beta - amyloid deposits. In other case we measure volume of human brain to estimate atrophy of human brain during progress of AD. Volume shrinkage of brain happens far earlier before symptoms of AD and MCI appear. This is less expensive and faster method, as well.

Since properties of human brains are different, it is crucial to use registration to a MRI brain image template. There are three main methods for registration: voxel - based morphometric techniques [7], ROI method and computational anatomy methods.

Current problem of estimation of brain shrinkage during AD is developed mainly in [8], and algorithms of semi-supervised learning are presented in [9, 10]. In [11] is presented another method for detection of AD by other popular semi-supervised learning method.

By modelling we aim to extract regions of brain which are most discriminative for presence of traces of mentioned diseases and to reduce preliminary information we need and calculation complexity of algorithms. Main purpose of current paper is to compare and estimate different methods of calculation of statistics of extracted features from selected regions. At last we apply clustering algorithms as ISO – Data.

In Part 2 the main algorithm is given, then in Part 3 – some experimental results and finally a conclusion is made.

2. Main algorithm

Here are presented stages of two algorithms: first one is for calculation of 3D mask that comprises parts of human brain where presence of degenerative tissues is most probable. The mask is created by applying a SVM classifier or KL divergence criteria for determination of most distinctive regions in two pre-processed and registered images of two patients - AD and healthy patients (Fig. 1). In second algorithm (Fig. 2) a mask is imposed on diagnosed patients with AD, MCI and healthy one and statistics of extracted features are computed.

The algorithms should manage with following obstacles: non-identical properties of human brain tissues in MRI images according to employed scanner, different size of human brain and due to age and sex of patients, varied degradation of brain tissues in different stages of development of AD, presence of different artefacts and image noises.

2.1. Pre-processing

In order to subdue already mentioned differences in degenerated tissue the pre-processing of neural images consists of following operations:

1) Reorientation of human brain – change orientation of human brain when subsequent operations demand it.

2) Resampling of image files - perform resampling of image data to decrease resolution of MRI image. Thus we reduce necessary memory and following operations requires objects with smaller size. On the contrary we perform upsampling in case of application of higher resolution mask.

3) Skull stripping – algorithms of removing of extracranial tissue from human brain image are semi-automatical or automatical. Yet current software tools that execute this operation need human inspection to obtain accurate results.

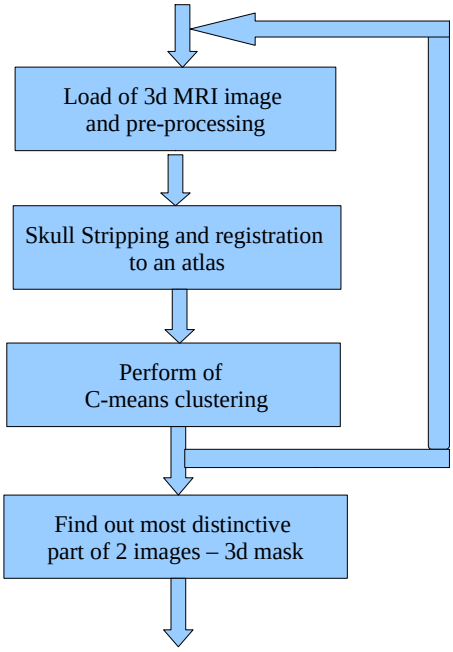

Fig. 1. 3D Mask construction by semi-supervised learning

4) Bias field correction – due to appeared artefacts, MRI images of human brain contain spatial variations. We apply bias correction in order to compensate them – it is implemented in software tools, as well.

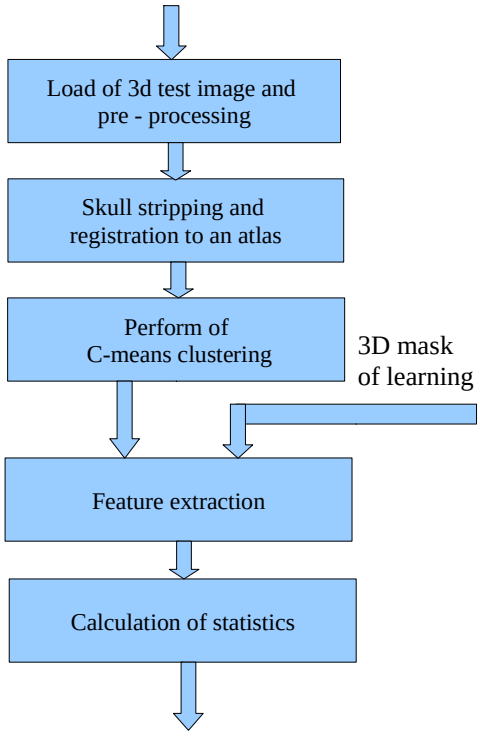

Fig. 2. Calculation of statistics from test image

5) Registration – image registration (warping) transforms image data to one coordinate system. In case of intra subject registration we register different images to a coordinate system of one image from image set and apply usually linear warping. In case of inter subject registration we use non-linear or linear registration to one coordinate system of external 3D model of human brain called atlas.

6) Resliceing – after registration we reslice preprocessed image. In this case transform matrix is applied to processed image, which is cut, analyzed and reconstructed with voxels whose size, volume and form is the same as template image.

7) Tissue segmentation – in order to extract basic anatomical entities in human brain, as white matter (WM), gray matter (GM) and cerebral spinal fluid (CSF) we apply C-means clustering method.

8) Resliceing – final operation is resliceing of preprocessed image. In this case transform matrix is applied to processed image, which is cut, analyzed and reconstructed with voxels whose size, volume and form are the same as template image.

After preprocessing of images we construct saturation map of segmented brain tissues in which different concentricity of intensity values account for healthy or non-degraded tissues.

2.2. Classifiers for construction of mask

In order to construct mask of most discriminating part of two groups of neuro images, we apply two methods: estimation with Kullback-Leibler (KL) divergence and probabilistic SVM.

In a certain point of distribution, KL divergence has the following value:

$$KL_{q,p}(u) = \int_{-\infty}^{+\infty} p(x)\log\frac{p(x)}{q(x)}dx\,, \qquad (1)$$

where x is a descriptor of image in position u, $p(x)$ and $q(x)$ are probability distributions of measured quantities. In our case x is a measure of intensity values from a corresponding tissue from MRI image. So output of this operation is a gray-scale mask whose maximum values indicate regions with higher difference between two groups of neuro images.

We select SVM because of its good generalization properties with small training sample. In current algorithm it is the probabilistic version of SVM described in [12]. The SVM training procedure consists of solving convex optimization problem. Let $\{x_i, y_i\}_{1\le i\le N}$ is a set of training samples, where each sample $x \in \Re^d$ (d is the dimension of input space) is labelled with $y \in \{-1,1\}$. If $K_{ij} = K(x_i,x_j)$ is the kernel matrix, where $K(x,y)$ is a Mercer kernel. Training SVM consists of determination the Lagrange multipliers α_i of the following optimization problem:

$$\min_{\alpha} \sum_{i=1}^{M}\sum_{j=1}^{M}\alpha_i\alpha_j y_i y_j K(x_i,x_j) - \sum_{i=1}^{M}\alpha_i$$
$$s.t.: \sum_{i=1}^{M}\alpha_i y_i = 0,\ \ C \ge \alpha_i \ge 0,\ \ i = 1,...,M \qquad (2)$$

where C > 0 is the upper bound determining the generalization properties of the SVM. For our classification task, we use the Gaussian kernel represented by:

$$K(x_i,x_j) = exp(-((x_i - x_j)^2)/2\sigma^2)\,. \qquad (3)$$

The class assignment for sample x with unknown class label, is given by:

$$y = \mathrm{sgn}\left(\sum \alpha_i y_i K(x_i,x) + b\right). \qquad (4)$$

2.3. Fuzzy C-means and ISO-data clustering

FCM is a clustering algorithm that performs fuzzy discrimination of input data through the following function for optimal discrimination:

$$J_{FCM}(P,U,\chi,c,m) = \sum_{i=1}^{c}\sum_{k=1}^{N}(u_{ik})^m . d_{ik}^2(x_k,p_i) \qquad (5)$$

on the condition that:

$$\sum_{i=1}^{c} u_{ik} = 1 \quad for \quad \forall k \in \{1...N\}, \qquad (6)$$

where P and U are variables whose optimal values we seek, and: χ is the number of clusters of input data, $m \ge 1$ is the fuzzy degree, and u_{ik} describes the level of memberships of feature vector x_k to cluster, represented by $U=[u_{ik}]$ with $c \times N$ discriminating fuzzy matrix. The total number of feature vectors is N. While d_{ik}^2 is the distance between feature vector x_k and the prototype p_j. The matrix d_{ik}^2 is determined from:

$$d_{ik}^2(x_k, p_i) = |x_k - p_i|_A^2 = (x_k - p_i)^T . A(x_k - p_i). \quad (7)$$

The matrix A can be every positive or negative matrix.

Let minimum of $J_{FCM}(P,U)$ be denoted with $(P*,U*)$ and the necessary conditions for it are:

$$p_i = \frac{\sum u_{ik}^m x_k}{\sum u_{ik}^m}, \quad (8)$$

$$u_{ik} = \frac{1}{\sum \left(\frac{d_{jk}^2}{d_{jk}^2} \right)^{1(m-1)}} . \quad (9)$$

ISODATA is a more complex algorithm, where the number of clusters is determined by repetetive calculations, division or merging of clusters, depending of standard deviations of samples from input data. Here the number of clusters, number of elements in separate clusters, thresholds for standard deviations of their elements and distances between elements are determined a priori. This algorithm is more flexible than standard k-means.

At the beginning is supposed presence of *c* fixed clusters in input data, their centroids are calculated, labels of separated samples are determined, and once again the centroids are calculated in order to decrease quadratic error of input data and current prototypes. Last two operations are repeated iteratively while the sum of distances between data points and prototypes is smaller than the selected threshold. Here minimization function is

$$J = \sum_{i=1}^{c} \left(\sum_{k, x_k \in S_i} d_{ik}^2 \right) = \sum_{i=1}^{c} \left(\sum_{k, x_k \in S_i} |x_k - p_i|^2 \right), \quad (10)$$

where S_i is a part of input data χ, corresponding to cluster *i*, d_{ik}^2 is the Euclidian distance metric between cluster prototypes and vectors of input data, belonging to them $x_k \in S_i$. With *J* is presented the total intra-cluster sum of quadratic error. Cluster prototypes are calculated in the following way:

$$p_i = \frac{\sum_{k=1}^{N} u_{ik} . x_k}{\sum_{k=1}^{N} u_{ik}}, \quad (11)$$

where u_{ik} take 0 or 1 according to memberships of x_k in S_i.

In the case of k-means clustering we have fixed, unchanging prototypes.

3. Experimental results

In order to construct 3D mask of probable degenerative tissue of human brain, we select two distinctive images from diagnosed AD and healthy patients.

We perform skull stripping automatically by BrainSuite [13], registration of neuro images and statistical estimation of extracted features – intensity values by BrainImage Suite [14]. Neuro images are selected from ADNI – Alzheimer's Disease Neuroimageing Initiative. We calculate mean value and standard deviation of intensity values of voxels from three tissues – white tissue, gray tissue and CSF from applied mask on neuro images of AD, MCI and healthy patients (Table 1).

	Mean intensity value of voxels from segmented regions			Standart intensi-ty deviation of voxels from seg-mented regions		
	CSF	GM	WM	CSF	GM	WM
1.Healthy patient	0.02	124.9	224,1	3	60	16.74
2. MCI patient	0.1	148.7	411.7	5.1	65.2	102.6
3. AD patient	0.05	87.5	217.8	2.7	40.7	53.5

Table 1. Statistical intensity estimation of voxels from segmented regions

4. Conclusion

We see that degeneration of brain tissue leads to change in intensity values of voxels in neuro images. In order to increase discrimination of healthy from degenerative tissues, we need to use more complex features or combination of features from preliminary selected regions of human brain.

5. Acknowledgment

This work is supported by Research and Development Sector (R&DS) with TU-Sofia under Grant №

102pd177-7 "Creation of models and object segmentation, applied to medicine and education".

References

[1] L. Hebert, P. Scherr, J. Bienias, D. Bennett, and D. Evans. Alzheimer Disease in the US Population: Prevalence Estimates Using the 2000 Census. Arch Neurol, Vol. 60, 2003, pp. 1119–1122.

[2] R. Petersen, Mild Cognitive Impairment. Continuum Lifelong Learning, Neurology, Vol. 13, 2007, pp. 13–36.

[3] R. Terry, E. Masliah, D. Salmon, et al. Physical Basis of Cognitive Alterations in Alzheimer's Disease: Synapse Loss is the Major Correlate of Cognitive Impairment. Ann Neurol., Vol. 30, 1991, pp. 572–580.

[4] R. Petersen, Mild Cognitive Impairment as a Diagnostic Entity. J Intern Med, Vol. 256, 2004, pp. 183-194.

[5] L. Apostolova, P. Thompson, A. Green, C. Jack, D. Harvey, R. Petersen, L. Thal, J. Cummings, C. DeCarli for the ADCS Group, 3D Analysis of Hippocampal Atrophy Progression in MCI subjects. In Proc. Of the 132nd Annual Meeting of the American Neurological Association, Featured in the "Dementia Walking Tour" ANA Session, Washington DC, October 2007.

[6] P. Hof, J. Morrison. The Aging Brain: Morphomolecular Senescence of Cortical Circuits, Trends Neurosci, Vol. 27, 2004, pp. 607–613.

[7] J. Ashburner, and K. Friston, Voxel-Based Morphometry — The Methods. Neuroimage Vol. 11, pp. 805–821.

[8] R. Filipovych, S. Resnick, and C. Davatzikos, Semi-Supervised Cluster Analysis of Imaging Data, Neuroimage, Vol. 54, Iss. 3, 2011, pp. 2185-2197.

[9] O. Chapelle, B. Schölkopf, and A. Zien, *Semi-Supervised Learning*. MIT Press, Cambridge, MA, 2006.

[10] X. Zhu, Semi-Supervised Learning Literature Survey, University of Wisconsin–Madison, Tech. Rep. 1530, 2006.

[11] R. Teramoto, Prediction of Alzheimer's Diagnosis using Semi-Supervised Distance Metric Learning with Label Propagation, Computational Biology and Chemistry, Vol. 32.6, Dec 2008, pp. 438-441.

[12] J. Platt, Probabilistic Outputs for Support Vector Machines and Comparison to Regularized Likelihood Methods, In: A. Smola, P. Bartlett, B. Scholkopf, and D. Schuurmans (Eds.): "Advances in Large Margin Classiffiers", Cambridge, MA, 2000.

[13] X. Papademetris, M. Jackowski, N. Rajeevan, R. Constable, and L. Staib, BioImage Suite: An Integrated Medical Image Analysis Suite, Section of Bioimaging Sciences, Dept. of Diagnostic Radiology, Yale School of Medicine, http://www.bioimagesuite. org.

[14] B. Dogdas, D. Shattuck, and R. Leahy, Segmentation of Skull and Scalp in 3D Human MRI Using Mathematical Morphology Human Brain Mapping, Vol. 26, No. 4, 2005, pp. 273-85.