Radiation Physics and Chemistry ∎ (∎∎∎∎) ∎∎∎–∎∎∎

Contents lists available at ScienceDirect

Radiation Physics and Chemistry

journal homepage: www.elsevier.com/locate/radphyschem

^{18}F-FDG PET brain images as features for Alzheimer classification

M.H. Azmi [a], M.I. Saripan [a,*], A.J. Nordin [b], F.F. Ahmad Saad [b], S.A. Abdul Aziz [a], W.A. Wan Adnan [a], the Alzheimer's Disease Neuroimaging Initiative [1]

[a] *Faculty of Engineering, Universiti Putra Malaysia, UPM Serdang, Selangor, Malaysia*
[b] *Center for Diagnostic Nuclear Imaging, Universiti Putra Malaysia, UPM Serdang, Selangor, Malaysia*

HIGHLIGHTS

- Investigate global and slice-based z-score and FDR as features to classify Alzheimer disease.
- Investigate optimal threshold of z-score and FDR to select significant voxels to discriminate Alzheimer disease.
- Investigate the potential of neural network as classifier in Alzheimer disease classification.

ARTICLE INFO

Article history:
Received 17 January 2016
Received in revised form
25 August 2016
Accepted 29 August 2016

Keywords:
PET
Alzheimer
Automatic classification
Z-score
Image processing

ABSTRACT

2-Deoxy-2-[fluorine-18] fluoro-D-glucose (^{18}F-FDG) Positron Emission Tomography (PET) imaging offers meaningful information for various types of diseases diagnosis. In Alzheimer's disease (AD), the hypometabolism of glucose which observed on the low intensity voxel in PET image may relate to the onset of the disease. The importance of early detection of AD is inevitable because the resultant brain damage is irreversible. Several statistical analysis and machine learning algorithm have been proposed to investigate the rate and the pattern of the hypometabolism. This study focus on the same aim with further investigation was performed on several hypometabolism pattern. Some pre-processing steps were implemented to standardize the data in order to minimize the effect of resolution and anatomical differences. The features used are the mean voxel intensity within the AD pattern mask, which derived from several z-score and FDR threshold values. The global mean voxel (GMV) and slice-based mean voxel (SbMV) intensity were observed and used as input to the neural network. Several neural network architectures were tested and compared to the nearest neighbour method. The highest accuracy equals to 0.9 and recorded at z-score ≤ -1.3 with 1 node neural network architecture (sensitivity$=$0.81 and specificity$=$0.95) and at z-score ≤ -0.7 with 10 nodes neural network (sensitivity$=$0.83 and specificity$=$0.94).

1. Introduction

Alzheimer's disease (AD) is among the critical illness that shows significant rise of cases being reported in recent years. What is more worrying, while the forecast number of AD cases is estimated to be triple in the next five years, the causes and progression of the disease is still vague. Lots of efforts have been made to find the best treatment to at least stop the disease from advancing and investigation on this matter is still ongoing. Apart from clinical examination and neuropsychological test, neuroimaging is well-accepted as one of the modalities to provide crucial information for the diagnosis. 2-deoxy-2-[fluorine-18] fluoro-D-glucose (^{18}F-FDG) is one of the radiopharmaceutical that is able to reveal the rate of glucose metabolism at a specific area of the body through Positron Emission Tomography (PET) imaging. Currently, it is widely used in oncology field where its information is vital for the diagnosis, staging, prognosis and therapy responses assessment (Tixier et al., 2016). In addition, we have explored several potential of ^{18}F-FDG on the segmentation of abnormal lung region of tuberculosis (Avazpour et al., 2009) and lung cancer cells (I. et al., 2009).

The application of ^{18}F-FDG PET imaging in AD diagnosis is not something peculiar nowadays. Its capability to provide the

* Corresponding author.
E-mail address: iqbal@upm.edu.my (M.I. Saripan).

[1] 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 investigators 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.

http://dx.doi.org/10.1016/j.radphyschem.2016.08.028
0969-806X/© 2016 Published by Elsevier Ltd.

functional information of a specific brain region, increases the potential for early detection of the disease compared to the other structural imaging modalities such as Computed Tomography (CT) and Magnetic Resonance Imaging (MRI). In AD, the decline of human cognitive ability relates to the unusual activity of a specific brain area responsible for that specific task, which is usually referred to the hypometabolism (Hoffman et al., 2000; Kato et al., 2016; Mosconi, 2005). The affected brain area radiates less photons compared to the normal functioning regions. PET imaging projects this information as a stack of images for a subject, which each slice represent an axial view at a specific location of the brain. The interpretation of these images may somehow vary among different radiologists, depending on their expertise, experience and the reported cases. Furthermore, similar signs can also be detected in normal aging person and people affected with other cognitive-related diseases such as Lewy-bodies dementia (DLB) and fronto-temporal dementia (FTD), which make the diagnosis more difficult. To cope with these problems, several algorithms to automatically classify AD subjects have been proposed in recent years (Illán et al., 2011; Martínez-Murcia et al., 2012; Ramírez et al., 2010; Segovia et al., 2013). However, the rate of hypometabolism and the pattern of where the hypometabolism occur that highly correlate to AD are the two main issues that still need to be investigated.

Previous works investigate AD features based on several statistical measures such as *t*-test and z-score (Ishii et al., 2006; Mosconi et al., 2008). These features are measured either on individual voxel or on a specific region. Classification was performed by a fixed threshold that best discriminates the normal and AD group. However, most of the time, the threshold values are not clearly defined, and AD abnormalities are not necessary appear within all area of the defined pattern. In Herholz et al. (2002), the AD mask was constructed based on the correlation of voxels with the Mini Mental State Examination (MMSE) scores of the probable AD patient. Whereas works related to z-score did not focus on different level of threshold to establish the abnormal area. In machine learning approach, selecting only the voxels with the highest discriminative power does not assure that the classification accuracy will be high. Instead, it produces a very rigid model that has low generalization capability to cope with certain cases. Furthermore, these voxels may appear due to the anatomical differences between subjects. Therefore, the aim of this paper is to further investigate these two issues and use the findings for the development of AD classification algorithm.

There are also works that used machine learning approach to classify AD. For instance, in Ramírez et al. (2013), statistical measures with highest FDR at specific slice were selected as features into support vector machine (SVM). FDR was also applied in several other works (López et al., 2009; Padilla et al., 2012) with the same goal, which is to select the features that have the highest discriminative power. PCA is also one of the popular method used to construct a feature dimension (Salmon et al., 2008). However, PCA may remove lots of information, which some of them may as well bring useful information to classify AD. The images used in this study were obtained from the Alzheimer Disease Neuroimaging Initiative (ADNI) database.

Firstly, all of the PET images will be standardized based on the co-registration to their baseline scan and resampled, followed by the high-frequency correction to achieve more uniform spatial resolution across images obtained from different scanners and centers (Joshi et al., 2009). Then, the images were spatially normalized to PET template using Statistical Parametric Mapping (SPM12) procedure. Next, the normal template was derived based on the simple averaging operation. Further analysis was then narrowed to the brain tissue area of normal template which was yielded from Fuzzy C-Means (FCM) segmentation. This is important to remove the insignificant voxels thus reduce the number of studied voxels and decrease the computational burden. The selection of voxels for the derivation of AD pattern mask was performed based on the *z*-score and FDR value which were calculated on voxel-by-voxel basis from AD subjects to the normal template.

The mean voxel intensity within the constructed AD mask will be calculated for each subject. In addition to that, the contribution of slice-specific mean voxel intensity was also investigated due to the fact that AD-related hypometabolism only occurs in certain location (Grey et al., 2012, 2011; Herholz et al., 2002; Ishii et al., 2006; Mosconi et al., 2008). These steps will be repeated for different AD mask obtained from different z-score and FDR threshold. The classification of AD is performed using neural network (NN) due to its capability to combine multiple features and compensate with non-linear input data. For assessment, 10-fold cross-validation approach was implemented and the performance of the algorithm is presented by the computed sensitivity, specificity, accuracy and also supported by the likelihood ratios.

Next section describes the data used in this study, including a brief explanation on ADNI database. Details of methods and algorithm proposed are elaborated in Section 3, followed by the results and discussion in Section 4. Summary and recommendation for future works is explained in the final section.

2. Data collection

All data were obtained from the ADNI database (adni.loni.usc.edu) which was launched in 2003 as a public-private partnership, led by Principal Investigator Michael W. Weiner, MD. The primary goal of ADNI was to test whether serial magnetic resonance imaging (MRI), positron emission tomography (PET), other biological markers, and clinical and neuropsychological assessment can be combined to measure the progression of mild cognitive impairment (MCI) and early Alzheimer's disease (AD). It is a multicenter and longitudinal type of study, with more than 50 centers have participated in the investigation.

This study considers subjects with age between 70 and 80 years old. Among them, 219 are normal subjects (112 male and 107 female) and 126 are probable AD patients (69 male and 57 female). This work focused only on the subjects between 70 and 80 years of age because the highest AD prevalence lies within this range, and also the highest number of data available in the database. The normal volunteer should have MMSE score between 24 and 30, CDR=0, which shows that no sign of depression, not diagnosed with MCI and confirmed to have non-demented brain. AD subjects should meet the MMSE score between 20 and 26, CDR=0.5 or 1 and meet the NINCDS/ADRDA (McKhann et al., 1984) criteria for probable AD.

3. Material and method

In general, the flow of the method is shown in Fig. 1. All images have passed through the image standardization procedure and spatially normalized to compensate the resolution differences and anatomical variations. Normal template was derived based on the standardized and normalized image. Further investigation was limited to the brain area of the template to reduce the insignificant data. AD mask was created based on several different thresholds of two measures; z-score and FDR. Then, the mean voxel value within the AD mask was considered as feature. The classification was performed based on the global mean voxel (GMV) and the slice-based mean voxel (SbMV) features using neural network.

In this study, 10-fold cross-validation method has been implemented. The available data were divided into 10 different training