Alzheimer's diagnosis using deep learning in segmenting and classifying 3D brain MR images

Tran Anh Tuan[a,*], Pham The Bao[b], Jin Young Kim[c], João Manuel R. S. Tavares[d]

[a] *Faculty of Mathematics and Computer Science, University of Science, Vietnam National University, Ho Chi Minh City, Vietnam*
[b] *Department of Computer Science, Sai Gon University, Vietnam*
[c] *Department of Electronic and Computer Engineering, Chonnam National University, South Korea*
[d] *Instituto de Ciência e Inovação em Engenharia Mecânica e Engenharia Industrial, Departamento de Engenharia Mecânica Faculdade de Engenharia, Universidade do Porto, Portugal*
* *Corresponding author: tatuan@hcmus.edu.vn , https://orcid.org/0000-0001-9994-8077*

Abstract

Dementia is one of the brain diseases with serious symptoms such as memory loss, and thinking problems. According to the World Alzheimer Report 2016, in the world, there are 47 million people having dementia and it can be 131 million by 2050. There is no standard method to diagnose dementia, and consequently unable to access the treatment effectively. Hence, the computational diagnosis of the disease from brain Magnetic Resonance Image (MRI) scans plays an important role in supporting the early diagnosis. Alzheimer's Disease (AD), a common type of Dementia, includes problems related to disorientation, mood swings, not managing self-care, and behavioral issues. In this article, a new computational method is presented to diagnose AD from 3D brain MRI. The efficient approach comprises two phases: I) segmentation and II) classification, both based on deep learning. After the brain tissues are segmented by a model that used Gaussian Mixture Model (GMM) as an additional input for Convolutional Neural Network (CNN), a new model for combining Extreme Gradient Boosting (XGBoost) and Support Vector Machine (SVM) is used to classify Alzheimer's disease based on the segmented tissues. For comparison, the new method was evaluated using the AD-86 and AD-126 datasets leading to Dice 0.96 for segmentation in both datasets and accuracies 0.88, and 0.80 for classification, respectively.

Keywords: Image analysis; Medical Imaging; CNN; SVM; XGBoost

1. Introduction

Dementia, one of the serious diseases affecting mental cognitive tasks, can lead to the death of many people [1]. Around the world, 9.9 million new cases of dementia were estimated in 2015, corresponding to one new case every 3 seconds, and the total estimated cost of dementia worldwide was US$ 818 billion. Now dementia is called a trillion-dollar disease rising to US$ 2 trillion by 2030 [2]. Alzheimer's Disease (AD), one of the common types of dementia, develops slowly and becomes severe with daily tasks related to memory, thinking, and behavior problems [3].

People developing the disease present seven different stages including three main stages: pre-clinical Alzheimer's disease stage, Mild Cognitive Impairment (MCI), and AD [4]. In the AD stage, the symptoms of patients are clear and they are difficult to function normally in daily life. Recently, many methods have been proposed for AD diagnosis from Magnetic Resonance Image (MRI) [4, 5, 6-14] and we divided into 3 groups (Figure 1):

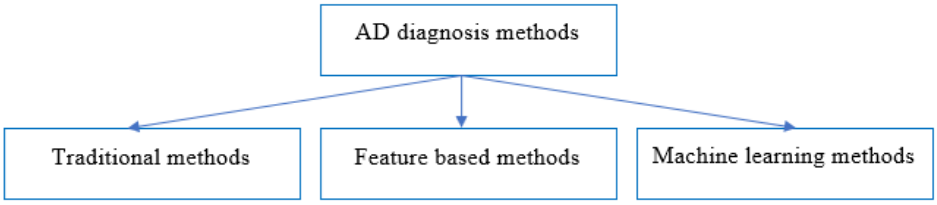

Fig. 1. Categories of current methods for AD diagnosis from 3D Brain MRI scans.

a) Traditional methods: A well-known approach is voxel-based morphometry (VBM) [15]. VBM uses a voxel-wise comparison of tissue densities. In brain MRI, VBM segments input data into many tissues such as white matter (WM), grey matter (GM), and Cerebrospinal Fluid (CSF) [6, 16]. Besides, there are other methods in this category to diagnosis the AD disease, such as the ones based on the measurement of the cortical thickness [17-18], or deformation assessment used as a biomarker [19-20]. Querbes O et al. [17] proposed for each subject, cortical thickness was measured on the MRIs volume. The resulting cortical thickness map was divided into 22 regions and a normalized thickness index was computed using the subset of regions (right medial temporal, left lateral temporal, right posterior cingulate) that distinguished stable mild cognitive impairment from progressive mild cognitive impairment. Hidetaka Arimura et al. [18] proposed an automated method for 3D measurement of cerebral cortical thicknesses based on fuzzy membership maps for evaluation of AD. The cerebral cortical thickness was measured by using a localized gradient vector trajectory in a fuzzy membership map. Alexandre Savio et al. [19] performs the non-linear registration of a subject's structural MRI volume to a reference template. Then, computes scalar measures of the registration's deformation field and performs across volume statistical group analysis of these scalar measures to detect effects. Long X et al. [20] proposed a deformation-based machine learning method for discrimination of AD. The method computed and analyzed the regional morphological differences of the brain between groups. Distance between each pair of subjects was quantified from asymmetric diffeomorphic registration, followed by an embedding algorithm and a learning approach for classification

b) Feature-based methods: Scale-Invariant Feature Transform (SIFT) [21] is an algorithm commonly used to detect local features in the input images. They are invariant in transforming the images such as scale, translation, and rotation because they are associated with key points of structures in the input

image. Thus, SIFT has been commonly used as a feature extractor for AD diagnosis [22-23]. Other local features that are commonly used for AD diagnosis is the histogram of oriented gradient (HOG) [24-25]. HOG decomposes an image into small squared cells, computes a histogram of the oriented gradient in each cell, normalizes the result, and returns a descriptor as a feature for each cell.

c) Machine learning methods: Support vector machine (SVM) is a classifier frequently used to diagnosis the disease from extracted features [12-13, 26]. Recently, many methods based on deep learning are proposed to diagnose AD from MRI scans [27-28].

Besides knowing the condition of a disease, the physician usually examines the tissues to attain more information about the size, location, and other characteristics of brain components. Therefore, in this study, we focus on tissue segmentation before classification to also assist the doctors in diagnosing brain-related diseases visually. The common methods for such a goal usually include four main steps:

- Preprocessing: skull stripping (brain segmentation) is the important step before tissue segmentation from 3D brain MRI [6,29];
- Segmentation: brain tissues such as WM, GM, CSF [16, 30], or Hippocampus [7,31] are used to segment from the images under analysis to predict the disease.
- Feature Extraction: different features are extracted from the segmented tissues by many techniques such as Principal component analysis (PCA) [8-9], and Gray-Level Invariant Features [10], to be used in the classification step;
- Classification: the diagnosis can be achieved using techniques of machine learning like SVM, which has been frequently used [11-13].

In this study, we propose a model that combines Gaussian Mixture Model (GMM) and Convolutional Neural Network (CNN) for WM, GM, and CSF tissues segmentation since these tissues are highly relevant on the disease diagnosis [14,32], and a new model that combines Extreme Gradient Boosting (XGBoost) and Support Vector Machine (SVM) to classify the patient Normal or AD.

This article is organized as follows: Section 2 reviews AD diagnosis methods and presents in detail the proposed method, section 3, presents the evaluation of the proposed method using two 3D brain MRI datasets, and the last section, provides the conclusions and perspectives of future works.

2. Methods

Our proposed method for tissue segmentation and AD diagnosis from MRI is depicted in Figure 2. In the preprocessing step, the 3D Brain MRI input is transformed into 2D slices and the skull stripping method is used [33]. Our approach comprises two main steps: segmentation and classification. In the segmentation step, a CNN model is used to segment the WM, GM, and CSF tissues in each slice. The model used 2D slices input and Gaussian Mixture Model (GMM) as an additional input for segmentation. In the classification step, 2D slices including the WM, GM, and CSF tissues are the input for another CNN. The model is trained to extract features to predict disease probability in each slice. From the prediction of each slice, XGBoost is used to predict a person into two classes: AD or Normal Control. After that, we select the important slices by using importance scores from XGBoosts and use SVM to predict person disease.

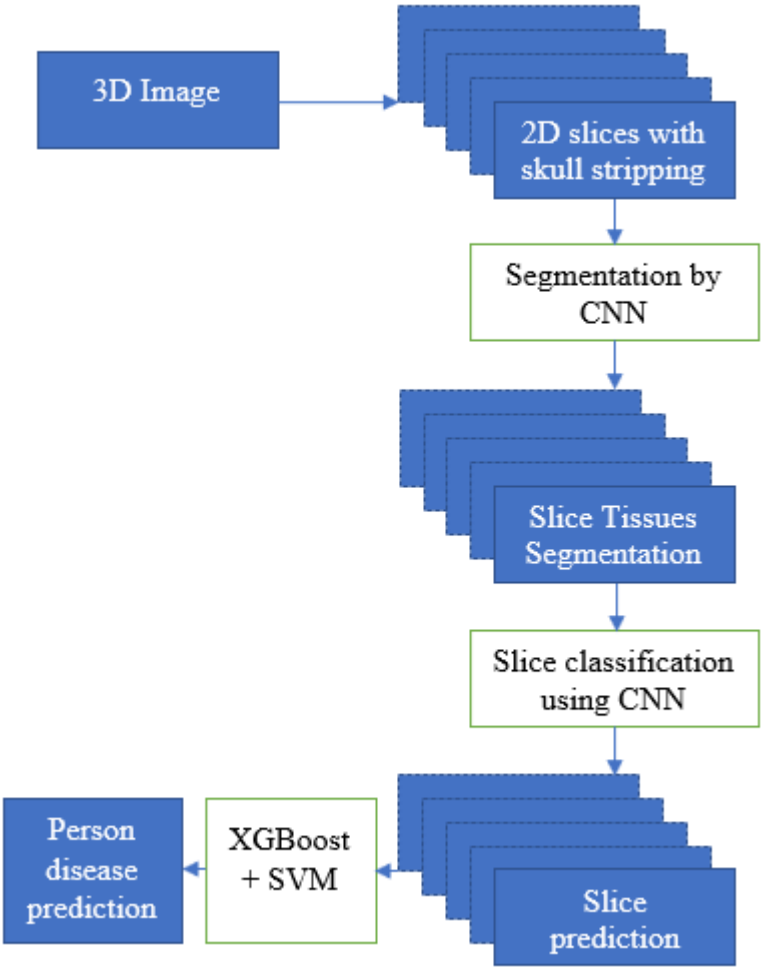

Fig. 2. Overview of the proposed method for AD diagnosis from 3D brain MRI scans.

2.1 Tissues Segmentation

Nowadays, many useful algorithms for localizing structures in MRI have been proposed [34]. Recently, the methods based on machine learning have become a prime trend in technology for image analysis nowadays. Based on CNN, there are several methods have been proposed to perform image segmentation such as Fully Convolutional Neural Networks [35] and Deep Convolutional Encoder-Decoder Neural Networks [36]. 2D CNN models which take each 2D slice as input gives the best result with few parameters and time-consuming in the model. However, because 2D CNNs gets one slice as input, they fail to take the context from adjacent slices. The information from adjacent slices is very useful for the prediction of any label. However, the disadvantage of using 3D CNN is the required power of the machine and the number of data. To overcome this disadvantage, we propose the use of the Relationship Probability of Pixel (RPP) as an additional input to the CNN model. Hence, the proposed RPP is computed from the n slices of the 3D MRI scan instead of being computed only from the slice to which the pixel under analysis belongs to. The illustration of the proposed tissue segmentation approach is presented in Figure 3.

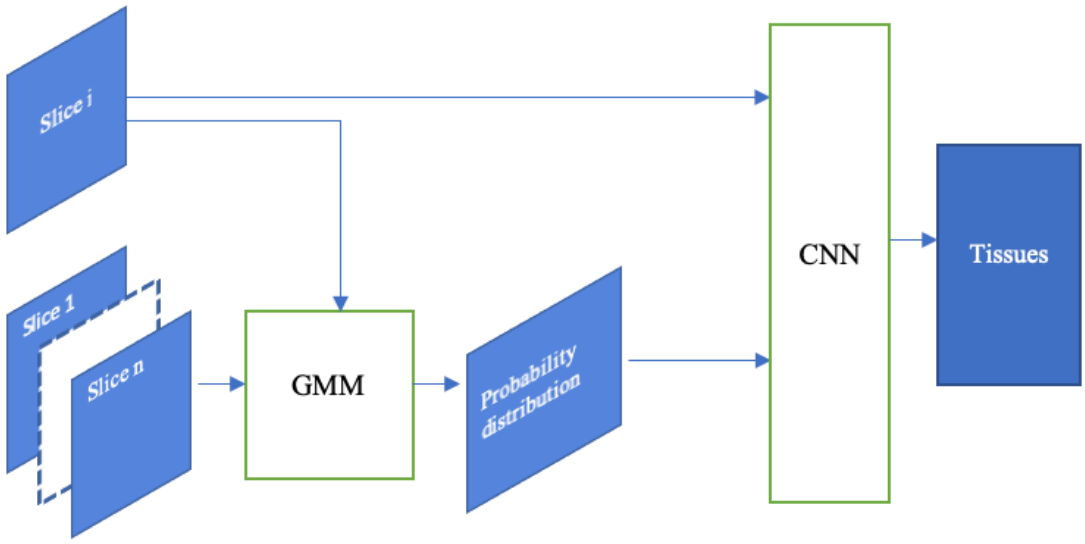

Fig. 3. A proposed method to segment brain tissues in brain MRI slices.

The CNN based segmentation of image X into image Y can be represented as a function $f: X \rightarrow Y$. Here, we define a new input $X' = \{(x_{0,0}, p_{0,0}), \dots, (x_{a,b}, p_{a,b})\}$, where *(a, b)* is the size of the original slice, to combine the information in the slice and the RPP of each of its pixels. RPP of each pixel in a slice is denoted as $P_{i,j}(p_1, \dots, p_m)$, where $i = \overline{1 \dots a}$, $j = \overline{1 \dots b}$, and m is the number of classes to segment and can be calculated by using the Gaussian Mixture Model on the intensity of the pixels of all slices of the 3D MRI scan. GMM is a probabilistic model that assumes that all the data points are generated from a mixture of a finite number of Gaussian distributions with unknown parameters [24]. It implements the expectation-maximization (EM) algorithm for fitting the mixture-of-Gaussian model.

2.2 Disease Classification

From the result of previous segmentation, we proposed using a CNN model for AD classification. With the model, we do not need RPP as an additional input because it inherits global features from previous segmentation results. If a 3D CNN model for 3D MRI classification is used, the result will not good because of the amount of data and the complexity of images, especially in the medical field. Instead of using a 3D CNN model for 3D MRI classification, we proposed using a typical CNN for each slice classification and then using classifier methods for 3D MRI classification from the slice prediction. Figure 4 shows a CNN model used for classification; this model as Convolution kernels of size 3x3, MaxPooling kernels of size 2x2, and Full Connected layers act as feature extractor and classifier.

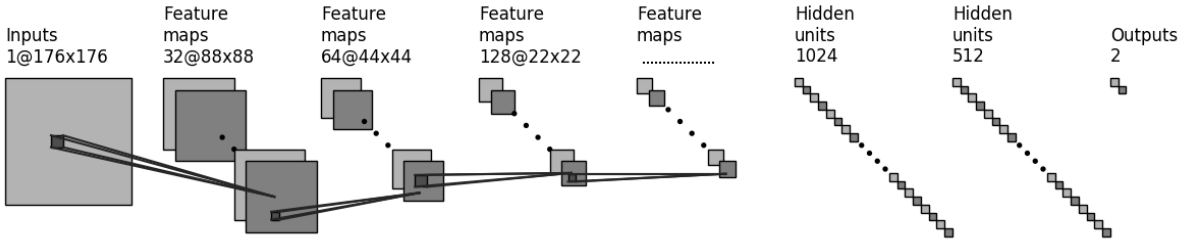

Fig. 4. A CNN model used for classification

There are many classifier methods such as SVMs, Decision tree, Random forest, XGboost. Like random forests, gradient boosting is a set of decision trees. However, Random forest builds each tree independently while gradient boosting builds one tree at a time. This additive model (ensemble) works in a forward stage-wise manner, introducing a weak learner to improve the shortcomings of existing weak learners. Random forests combine results at the end of the process while gradient boosting combines results along the way. Therefore, we used XGboost for the first prediction. XGBoost give a better result. Another reason is that each MRI has more than a hundred of 2D slices and there are many slices that do not necessary of disease prediction. The XGboost is used as a slice selection. It uses features of each slice to train and evaluate the importance score indicates how valuable the related feature was in the building of the boosted decision trees within the model. XGBoost [37] efficiently deals with sparse data and is suitable for large-scale datasets. It also implements distributed and parallel computing flexibility. XGBoost estimates the target feature by a series of decision trees and defining quantized weight for each leaf node. The prediction function is defined as:

$$\hat{y}_i = \sum_{k=1}^{K} f_k(x_i), \tag{1}$$

where $\hat{y}_i$ is the predicted class of the $i-th$ observation, x_i is the corresponding feature vector, K is the total number of decision trees. The function f_k is defined as:

$$f_k(x_i) = \omega_{qk(x_i)}, \tag{2}$$

where $qk(x_i)$ is the structure function of the $k-th$ decision tree that map x_i to the corresponding leaf node and ω is the vector of the quantized weight.

A proposed method for AD prediction by using XGBoost on WM, GM and CSF tissues segmented in 3D brain MR images are shown in Figure 5

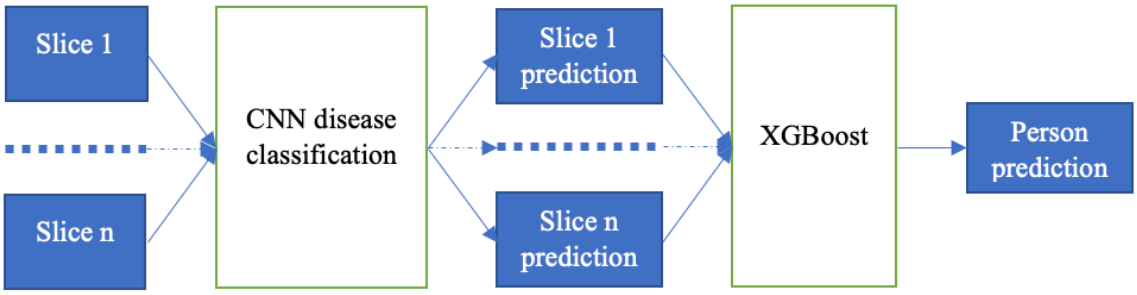

Fig. 5. A proposed method for AD prediction by using XGBoost on WM, GM, and CSF tissues segmented in 3D brain MR images.

SVMs are supervised learning models that analyze the data used for classification [38]. The use of SVMs combined with CNN has proved to be promising for classification. SVMs and boosting are two techniques for learning both having received considerable attention in recent years and many successful applications have been described in the literature. The objective of SVMs is to maximize the separation between the classes. By using a kernel trick to map the training samples from an input space to a high dimensional feature space, SVM finds an optimal separating hyperplane in the feature space and uses a regularization parameter to balance model complexity and training error. Therefore, with a suitable kernel, the SVM can be better. With our approach, we combine two classifiers, the classification accuracies obtained on the two dataset groups may not be the highest ones, but the obtained results are more robust for the two groups than the ones obtained by the related methods.

In this work, the disease probability of each slice prediction is used as the feature for disease prediction based on an SVM classifier. SVM computes the discriminant function as a linear combination of the similarity scores with learned weights α_i, where many of them may be zero:

$$f(x) = sgn(\sum_i y_i \alpha_i K(x_i, x) + b), \quad (3)$$

where the kernel function $K = (x_i, x)$ measures the similarity between the input pattern x and the training sample x_i, and the samples x_i for which the corresponding α_i are non-zero are the support vectors.

Instead of using all the features for classification, is common to use methods for feature selection to get better accuracy and performance, particularly on high-dimensional datasets. With the disease probability computed from each slice, we use an SVM on the features selected using XGBoost to classify the 3D brain MRI scan under analysis into two classes: AD or Normal Control (NC).

The advantage of using XGBoost is after the boosted trees are built, importance scores for each attribute can be calculated. The importance score indicates how valuable the related feature was in the building of the boosted decision trees within the model. The proposed classification method for the combination is shown in Figures 6.

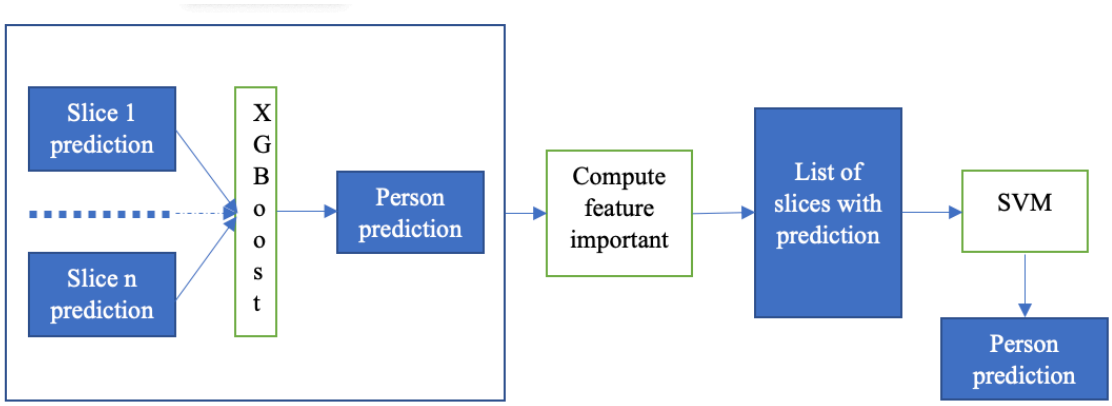

Fig. 6. A proposed method for AD prediction by using an SVM classifier from XGBoost selected features.

I: Let the segmentation result of a 3D brain MRI scan be $X = \{(x_1, l_1), (x_2, l_2), \dots, (x_n, l_n)\}$, where x_i denotes each 2D slice and l_i the brain tissue labels. Let Y= $\{y_1, y_2, \dots, y_n\}$ be the probability of disease for each 2D slice, CNN is a function $g: X \to Y$ used for prediction.

II: With n persons, each person has m features, each feature is the probability of disease in a 2D slice, $D = \{x_i, y_i\}$ $(|D| = n, x_i \in \mathbb{R}^m)$ is a tree ensemble model that uses K additive functions to predict the output:

$$\hat{y}_i = \phi(x_i) = \sum_{k=1}^{K} f_k(x_i), f_k \in \mathcal{F}, \quad (4)$$

where $\mathcal{F} = \{f(x) = \omega_{q(x)}\}(q: \mathbb{R}^m \to \mathrm{T}, \omega \in \mathbb{R}^m)$ is the space of the regression trees, q represents the structure of each tree that maps a sample to the leaf index, and ω is the weight vector of each leaf.

III: Using m features $\{y_1, y_2, \dots, y_m\}$ to build a tree ensemble model, feature importance $Y' = \{y'_1, y'_2, \dots, y'_n\}$ is calculated by counting the number of times each feature is used in all generated trees. Then, using an SVM classifier on features Y', the disease is predicted.

3. Results and discussion

There are many public image datasets available that could be used to evaluate the proposed method such as the Open Access Series of Imaging Studies (OASIS) and Alzheimer's disease Neuroimaging Initiative (http://adni.loni.usc.edu/) datasets. ADNI and OASIS both are well-known datasets about AD. ADNI is a longitudinal multicenter study designed to develop clinical, imaging, genetic, and biochemical biomarkers for the early detection and tracking of AD. Three overarching goals of the ADNI study are detecting AD at the earliest possible stage (pre-dementia) and identify ways to track the progression of the diseases with biomarkers; supporting advances in AD intervention, prevention, and treatment through the application of new diagnostic methods at the earliest possible stages (when intervention may be most effective); continually administer ADNIs innovative data-access policy, which provides all data without embargo to all scientists in the world. However, there is no WM, GM, and CSF ground truth for segmentation comparison in the ADNI dataset. Therefore, we choose the OASIS dataset to evaluate the segmentation method and classification method. We evaluate the segmentation method against the corresponding ground truth on OASIS dataset [39], which contains 3D brain

MRI scans of 98 normal control (NC) subjects and 99 probable AD subjects aged between 60 and 96 years. In this research, the subjects were divided into two subsets by their age and dementia status according to Clinical Dementia Rating (CDR) [40]. CDR is a numeric scale commonly used to quantify the severity of symptoms of dementia based on the patient's cognitive and functional performance in six areas: memory, orientation, judgment & problem solving, community affairs, home & hobbies, and personal care. Then, the two groups built were:

- AD-86: 86 subjects aged between 60 and 80 years, including 20 patients mild AD (CDR = 1) and 66 healthy subjects (CDR = 0);
- AD-126: 126 subjects aged between 60 and 96 years, including 28 patients mild AD (CDR = 1) and 98 healthy subjects (CDR = 0).

In the OASIS dataset, the 3D brain MRI scans were first averaged and gain-field corrected to improve the signal/noise ratio, then registered to the Talairach space [41] via an affine transform and finally, the skull was masked out.

3.1 Segmentation comparison

We evaluated the performance of the segmentations obtained by the proposed method using a Dice Similarity Coefficient (DSC):

$$S(X,Y) = \frac{2|X \cap Y|}{|X| + |Y|}, \quad (5)$$

where |X| and |Y| are the number of pixels in the sets under comparison, *X* and *Y*, respectively, and |X∩Y| is the number of pixels shared by the two sets.

We implemented the proposed method in Python and used the Keras [42] and scikit-learn [43] library toolkits for deep neural networks. We adopted UNET [44] as the original CNN model. Each 3D brain MRI scan, with the original size of 208x176x17, was converted to 208 2D images, i.e. slices, with the size of 176x176 pixels. Then, after applying the correspondent skull stripping mask included in the dataset, the input to the CNN is 2D slices, and then the three brain tissues: WM, GM, and CSF, were segmented as the output of the network. The optimizer used in the CNN was the 'Adam(lr=1e-4)' [45], and the loss was 'binary_crossentropy' [46]. We used model checkpoint, early stopping to avoid overfitting. We adopted 3-Fold validation [47]. The model is trained with 2D slices input and Gaussian Mixture Model (GMM) as an additional input. The output is 3 tissues segmentation ground truth. Because the developed implementation is an enhancement of the UNET method, we compared it against the original one. The brain tissue segmentation is presented in Table 1 and Figure 7. The main advantage of our proposed method is the use of the relationship probability of each pixel computed taken into account all slices include in the 3D MRI scan under study. However, these features do not lead to better results in all cases, particularly when the structures to be segmented are very small or when there are not enough data for training such was the case of the CSF tissue segmentation in the AD-86 dataset group (Table 1).

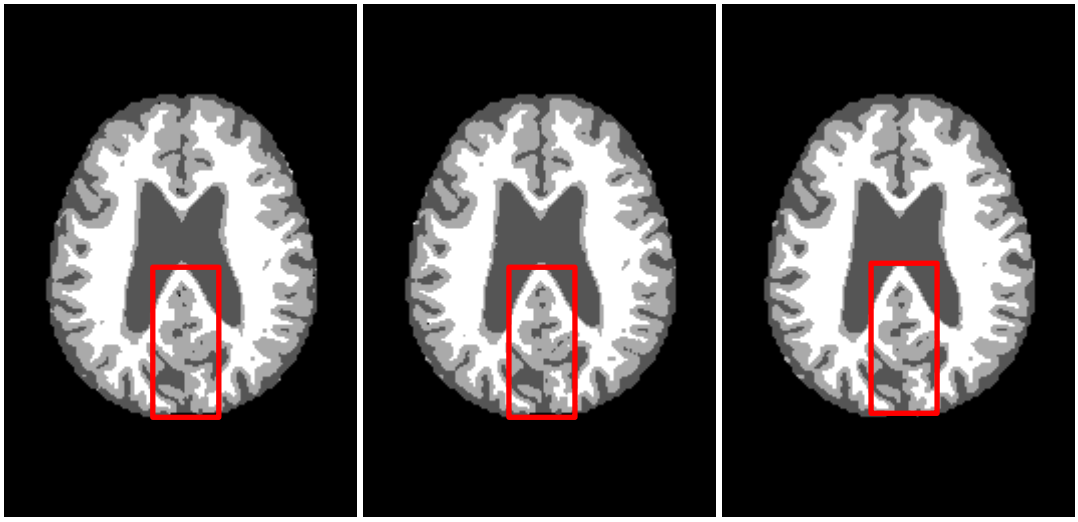

Fig. 7. An example of the segmentation results obtained by the proposed method. From the left to right are the results of UNET, of the proposed method, and the corresponding ground truth. The main visible difference is on the border of the tissues inside the red rectangle.

Table 1. Results of the proposed segmentation method and the ones obtained using the traditional UNET method in terms of DSC.

Dataset groups	Method	WM	GM	CSF	Average
AD-86	UNET	0.96	0.94	0.95	0.95
	Proposed method	**0.97**	**0.95**	**0.95**	**0.96**
AD-126	UNET	0.95	0.94	0.95	0.95
	Proposed method	**0.96**	**0.95**	**0.96**	**0.96**

3.2 Classification Result

We evaluated the proposed classification approach based on EER accuracy [16,48]. EER is calculated by choosing a threshold so that the false positives rate is equal to the false-negative rate and calculating the classification accuracy with the chosen threshold. We evaluate the method with 10-fold cross-validation [47].

In the classification, all slices that contained the brain tissues WM, GM, and CSF previously segmented were used for disease prediction. We implemented the proposed classification method using a typical CNN with 28 layers, including Convolution, Max-pooling, and Full Connected before using *softmax activation* for classification. are used to avoid overfitting. The slices which contain no brain tissues were not used. XGBoost was used with the parameters, i.e. *objective': 'binary:logistic', "max_depth": 6, "eta": 0.3*, "*seed":1, number of three: 1000*. The feature importance is the counting number of times that feature was used in all trees. The SVM classifier was used with the default parameters. In the CNN model, the input is 3 tissues of each slice (we get the feature from the segmentation step) and the output is AD or NC (0 or 1). The accuracies from our proposed classification method and the ones obtained by related methods for the two groups of 3D brain MRI scans, AD-86, and AD-126, are presented in Table 2.

Table 2. Classification accuracies were obtained by the proposed method and by common methods for AD classification in 3D brain MRI scans in terms of ACC.

Methods		**AD-86**	**AD-126**
Segmentation	**Classification**		
U-Net+GMM	**CNN + XGBoost + SVM**	**88%**	**80%**
U-Net	**CNN + XGBoost + SVM**	**88%**	**79%**
	CNN + XGBoost	**89%**	76%
	CNN + SVM	83%	**81%**
State of art [49] methods:			
- Toews et al.		80%	76%
- Daliri		86%	78%
- Chen et al.		83%	71%
- Wang et al.		80%	**79%**
- Tao Li et al.		**88%**	78%

When using only U-Net+GMM for segmentation and using CNN+XGBoost or CNN+SVM for classification, the final result is not better. However, when using U-Net+GMM for segmentation and using CNN+ XGBoost+SVM for classification, the accuracy is better. Although the U-Net+GMM segmentation result is better than U-Net 1%, there are many slices that play important roles in disease classification. XGboost evaluate the importance score of each slice and get the feature of important slices. After that, SVM is used to classify these slices. Therefore, if having more data and the images are complex, the improvement of segmentation is necessary before the disease classification when using the combination of classifiers.

According to the classification findings (Table 2), the XGBoost based approach achieved good results because its decision is based on several decision trees. However, in some cases, SVM with a suitable kernel can obtain better results, here the Radial Basis Function (RBF) kernel gave a better result than XGBoost as to the AD-126 dataset group. With our approach, we combine two classifiers, the classification accuracies obtained on the two dataset groups may not be the highest ones, but the obtained results are more robust for the two groups than the ones obtained by the related methods. This higher classification robustness is possible because, with the combination of the two classifiers, the limitations of each one are overcome. On the other hand, the AD-126 dataset group contains brain 3D MRI scans of older patients that are very difficult to classify: many images of old AD patients are complex to analyze because of the anatomical changes due to normal aging additionally to the anatomical abnormalities due to the disease. Therefore, the classification accuracy obtained for the AD-126 group was lower than for the AD-86 group. From the classification results presented in Table 2, one can confirm that the proposed classification method has a very promising performance.

4. Conclusion

In this article, a feature from the GMM method is used as an additional feature for the CNN model to segment the brain tissues, and then, the other CNN model is used to extract the features for classification. XGBoost is used as a feature selection for SVM to enhance the classification accuracy achieved. The proposed method was

evaluated using the public OASIS image dataset and show high accuracy. With deep learning, the diagnosis gives better results if the training dataset is enlarged or data augmentation is used in a suitable way.

The method should be improved by segmenting more brain tissues in the future such as the hippocampus. With more tissue segmentation, more features can lead to high classification accuracy. We can concentrate on the algorithm inside the model such as reducing loss information in down-sampling to enhance the architecture of the CNN model.

Acknowledgments

Tran Anh Tuan was supported by Erasmus Mundus Program IMPAKT Project Fellowship Program. João Manuel R.S. Tavares gratefully acknowledges the funding of Project NORTE-01-0145-FEDER-000022 - SciTech - Science and Technology for Competitive and Sustainable Industries, cofinanced by "Programa Operacional Regional do Norte" (NORTE2020), through "Fundo Europeu de Desenvolvimento Regional" (FEDER). Data were provided by the Open Access Series of Imaging Studies [OASIS: Cross-Sectional: Principal Investigators: D. Marcus, R, Buckner, J, Csernansky J. Morris; P50 AG05681, P01 AG03991, P01 AG026276, R01 AG021910, P20 MH071616, U24 RR021382]

Compliance with ethical standards

Funding

This study did not receive a specific grant.

Conflict of interest statement

The authors report no conflict of interest.

Ethical approval

This article does not contain any studies with human participants performed by any of the authors

Informed Consent

Not applicable

References

[1] M. Prince, R. Bryce, E. Albanese, A. Wimo, W. Ribeiro, and C. P. Ferri. (2013). The global prevalence of dementia: a systematic review and meta-analysis. Alzheimers Dement, 9(1):63-75. doi: 10.1016/j.jalz.2012.11.007.

[2] M. Prince, A. Wimo, M. Guerchet, G. Ali, Y. Wu, and M. Prina. (2015). World Alzheimer report 2015, The global impact of dementia, An analysis of prevalence, incidence, cost and trends. Alzheimer's Disease International, London.

[3] Tarawneh R, Holtzman DM. (2012). The Clinical Problem of Symptomatic Alzheimer Disease and Mild Cognitive Impairment. Cold Spring Harb Perspect Med, 2(5):a006148. doi: 10.1101/cshperspect.a006148.

[4] Barry Reisberg. (2019). Fisher Center for Alzheimer's Research Foundation. Clinical Stages of Alzheimer's. https://www.alzinfo.org/understand-alzheimers/clinical-stages-of-alzheimers/ Accessed March 13, 2019.

[5] Domingos Vieira, Ricardo Vardasca, João Manuel R.S. Tavares. (2017). PET/MRI technique role in Alzheimer disease, European Journal of Nuclear Medicine and Molecular Imaging. 44(2): S506.

doi:10.1007/s00259-017-3822-1

[6] P. Kalavathi and V. B. Surya Prasath, (2016). Methods on Skull Stripping of MRI Head Scan Images—a Review, J Digit Imaging. 29(3)(2016): 365–379. doi: 10.1007/s10278-015-9847-8

[7] Jonathan H. Morra; Zhuowen Tu; Liana G. Apostolova; Amity E. Green; Arthur W. Toga; Paul M. Thompson. (2010). Comparison of Ada-Boost and Support Vector Machines for Detecting Alzheimer's Disease Through Automated Hippocampal Segmentation, IEEE Transactions on Medical Imaging. 29(1): 30-43. doi: 10.1109/TMI.2009.2021941.

[8] Lama RK, Gwak J, Park JS, Lee SW. (2017). Diagnosis of Alzheimer's Disease Based on Structural MRI Images Using a Regularized Extreme Learning Machine and PCA Features, J Healthc Eng. 2017: 5485080. doi: 10.1155/2017/5485080

[9] Devvi Sarwinda, Aniati M. Arymurthy. (2013). Feature selection using kernel PCA for Alzheimer's disease detection with 3D MR Images of brain, Advanced Computer Science and Information Systems (ICACSIS). doi: 10.1109/ICACSIS.2013.6761597

[10] Li M, Oishi K, He X, Qin Y, Gao F, Mori S. (2014). An Efficient Approach for Differentiating Alzheimer's Disease from Normal Elderly Based on Multicenter MRI Using Gray-Level Invariant Features, PLoS One. 9(8): e105563. doi: 10.1371/journal.pone.0105563.

[11] P. Padilla, M. Lopez, J. M. Gorriz, J. Ramirez, D. Salas-Gonzalez, I. Alvarez. (2012). NMF-SVM based CAD tool applied to functional brain images for the diagnosis of Alzheimer's disease, IEEE Trans Med Imaging. 31(2): 207-16. doi: 10.1109/TMI.2011.2167628.

[12] Jun Zhang, Yue Gao, Yaozong Gao, Brent C. Munsell, Ding-gang Shen. (2016). Detecting Anatomical Landmarks for Fast Alzheimer's Disease Diagnosis, IEEE Transactions on Medical Imaging. 35(12): 2524-2533. doi: 10.1109/TMI.2016.2582386

[13] Esther E. Bron, Marion Smits, Wiro J. Niessen, Stefan Klein. (2015). Feature Selection Based on the SVM Weight Vector for Classification of Dementia, IEEE Journal of Biomedical and Health Informatics.19(5): 1617-1626. doi: 10.1109/JBHI.2015.2432832

[14] R Barber, P Scheltens, A Gholkar, C Ballard, I McKeith, P Ince, R Perry, J O'Brien, White matter lesions on magnetic resonance imaging in dementia with Lewy bodies, Alzheimer's disease, vascular dementia, and normal aging, J Neurol Neurosurg Psychiatry. 67(1)(1999): 66–72

[15] Ashburner J, Friston KJ. (2000). Voxel-based morphometry-the methods, Neuroimage. 11(6 Pt 1): 805-21. doi: 10.1006/nimg. .0582

[16] Hyemin Jang, Hunki Kwon, Jin-Ju Yang, Jinwoo Hong, Yeshin Kim, Ko Woon Kim, Jin San Lee, Young Kyoung Jang, Sung Tae Kim, Kyung Han Lee, Jae Hong Lee, Duk L. Na, Sang Won Seo, Hee Jin Kim & Jong-Min Lee. (2017). Correlations between Gray Matter and White Matter Degeneration in Pure Alzheimer's Disease, Pure Subcortical Vascular Dementia, and Mixed Dementia, Scientific Reports. 7: 9541. doi: 10.1038/s41598-017-10074-x

[17] Querbes O, Aubry F, Pariente J, Lotterie JA, Démonet JF, Duret V, Puel M, Berry I, Fort JC, Celsis P. (2009). Early diagnosis of Alzheimer's disease using cortical thickness: impact of cognitive reserve, Brain. 132(8): 2036-2047. doi: 10.1093/brain/awp105.

[18] Hidetaka Arimura, Chiaki Tokunaga, Takashi Yoshiura, Tomoyuki Ohara, Yasuo Yamashita, Fukai Toyofuku.(2013). Automated measurement of cerebral cortical thickness based on fuzzy membership map derived

from MR images for evaluation of Alzheimer's disease, International Journal of Computer Assisted Radiology & Surgery. 2013:7116-9 . doi: 10.1109/EMBC.2013.6611198.

[19] Alexandre Savio, Manuel Graña (2013). Deformation based feature selection for Computer Aided Diagnosis of Alzheimer's Disease. 40 (5): 1619-1628. doi: 10.1016/j.eswa.2012.09.009

[20] Long X, Chen L, Jiang C, Zhang L. (2017). Alzheimer's Disease Neuroimaging Initiative, Prediction and classification of Alzheimer disease based on quantification of MRI deformation, PLoS One. 12(3): e0173372. doi: 10.1371/journal.pone.0173372

[21] David G. Lowe. (1999). Object Recognition from Scale Invariant Features, International Conference of Computer Vision. 6365386 . DOI: 10.1109/ICCV.1999.790410

[22] Chen Y, Storrs J, Tan L, Mazlack LJ, Lee JH, Lu LJ. (2014). Detecting brain structural changes as biomarker from magnetic resonance images using a local feature based SVM approach, J Neurosci Methods. 221: 22-31. doi: 10.1016/j.jneumeth.2013.09.001

[23] Daliri MR. (2012). Automated diagnosis of Alzheimer disease using the scale-invariant feature transforms in magnetic resonance images, Med Syst. 36(2): 995-1000. doi: 10.1007/s10916-011-9738-6.

[24] Xiaofeng Zhu, Heung-Il Suk, Yonghua Zhu, Kim-Han Thung, Guorong Wu, Dinggang Shen. (2015). Multi-view Classification for Identification of Alzheimer's Disease, Mach Learn Med Imaging. 9352: 255–262. doi: 10.1007/978-3-319-24888-2_31

[25] Dalal, N. and B. Triggs. (2005). Histograms of Oriented Gradients for Human Detection, IEEE Computer Society Conference on Computer Vision and Pattern Recognition. 1: 886–893. doi: 10.1109/CVPR.2005.177

[26] Enrico Pellegrini, Lucia Ballerini, Maria del C. Valdes Hernandez, Francesca M.Chappell, Victor González-Castro, Devasuda Anblagan, Samuel Dansoa, Susana Muñoz Maniega, Dominic Job, Cyril Pernet, Grant Mair, Tom MacGillivray, Emanuele Trucco, Joanna Wardlaw. (2018). Machine learning of neuroimaging for assisted diagnosis of cognitive impairment and dementia: A systematic review, Alzheimer's & Dementia: Diagnosis, Assessment & Disease Monitoring. doi: 10.1016/j.dadm.2018.07.004

[27] Jun Shi, Xiao Zheng, Yan Li, Qi Zhang, Shihui Ying. (2018). Multimodal Neuroimaging Feature Learning With Multimodal Stacked Deep Polynomial Networks for Diagnosis of Alzheimer's Disease, IEEE J Biomed Health Inform. 22(1): 173-183. doi: 10.1109/JBHI.2017.2655720.

[28] McCrackin L. (2018). Early Detection of Alzheimer's Disease Using Deep Learning. Canadian AI 2018: Advances in Artificial Intelligence. 355-359. doi: 10.1007/978-3-319-89656-4_40

[29] Ahmed Serag, Manuel Blesa, Emma J. Moore, Rozalia Pataky, Sarah A. Sparrow, A. G. Wilkinson, Gillian Macnaught, Scott I. Semple & James P. Boardman. (2016). Accurate Learning with Few Atlases (ALFA): an algorithm for MRI neonatal brain extraction and comparison with 11 publicly available methods, Scientific Reports. 6: 23470. doi: 10.1038/srep23470.

[30] Arnaldo Mayer and Hayit Greenspan. (2009). An Adaptive Mean-Shift Framework for MRI Brain Segmentation, IEEE Transactions on Medical Imaging. 28(8): 1238-1250. doi: 10.1109/TMI.2009.2013850.

[31] Kim J, Valdes-Hernandez Mdel C, Royle NA, Park J. (2015). Hippocampal Shape modeling based on a Progressive Template Surface De-formation and its Verification, IEEE Trans Med Imaging. 34(6): 1242-61. doi: 10.1109/TMI.2014.2382581.

[32] - Gunning-Dixon FM, Raz N, The cognitive correlates of white matter abnormalities in normal aging: a quantitative review, Neuropsychology. 14(2)(2000): 224-232.

[33] Zeynettin Akkus, Alfiia Galimzianova, Assaf Hoogi, Daniel L. Ru-bin, Bradley J. Erickson. (2017). Deep Learning for Brain MRI Segmentation: State of the Art and Future Directions, Journal of Digital Imaging. 30(4): 449–459. doi: 10.1007/s10278-017-9983-4.

[34] Carlos A. S. J. Gulo, Antonio C. Sementille, João Manuel R. S. Tavares. (2017). Techniques of Medical Image Processing and Analysis accelerated by High-Performance Computing: A Systematic Literature Review, Journal of Real-Time Image Processing. 1–18. doi: 10.1007/s11554-017-0734-z

[35] Jonathan Long, Evan Shelhamer, Trevor Darrell. (2017). Fully convolutional networks for semantic segmentation, IEEE Transactions on Pattern Analysis and Machine Intelligence. 39(4): 640-651. doi: 10.1109/TPAMI.2016.2572683

[36] Vijay Badrinarayanan, Alex Kendall, Roberto Cipolla. (2017). SegNet: A Deep Convolutional Encoder-Decoder Architecture for Image Segmentation, IEEE Transactions on Pattern Analysis and Machine Intelligence. 39(12): 2481-2495. doi: 10.1109/TPAMI.2016.2644615

[37] Tianqi Chen, Carlos Guestrin. (2016). XGBoost: A Scalable Tree Boosting System, 22nd SIGKDD Conference on Knowledge Discovery and Data Mining, 785-794. doi: 10.1145/2939672.2939785

[38] Cristianini, John Shawe-Taylor. (2000). An Introduction to Support Vector Machines and Other Kernel-based Learning Methods, Cambridge: Cambridge University Press.

[39] Marcus DS, Wang TH, Parker J, Csernansky JG, Morris JC, Buckner RL. (2007). Open Access Series of Imaging Studies (OASIS): cross-sectional MRI data in young, middle aged, nondemented, and demented older adults, Journal of cognitive neuroscience. 19(9): 1498-1507. doi: 10.1162/jocn.2007.19.9.1498

[40] Hughes CP, Berg L, Danziger WL, Coben LA, Martin RL. (1982). A new clinical scale for the staging of dementia, Br J Psychiatry. 140(1982):566-72. doi: 10.1192/bjp.140.6.566

[41] J. Talairach, G. Szikla. (1980). Application of Stereotactic Concepts to the Surgery of Epilepsy, Advances in Stereotactic and Functional Neurosurgery. 30: 35-54. doi: 10.1007/978-3-7091-8592-6_5

[42] Chollet, Francois and others. (2015). Keras, https://keras.io, (accessed September 2018)

[43] Fabian Pedregosa, Gaël Varoquaux, Alexandre Gramfort, Vincent Michel, Bertrand Thirion, Olivier Grisel, Mathieu Blondel, Peter Prettenhofer, Ron Weiss, Vincent Dubourg, Jake Vanderplas, Alexandre Passos, David Cournapeau, Matthieu Brucher, Matthieu Perrot, Édouard Duchesnay. (2011). Scikit-learn: Machine Learning in Python, Journal of Machine Learning Research, 12: 2825-2830.

[44] Olaf Ronneberger, Philipp Fischer, Thomas Brox. (2015). U-Net: Convolutional Networks for Biomedical Image Segmentation, Medical Image Computing and Computer-Assisted Intervention (MICCAI), Springer, LNCS. 9351: 234--241. doi: 10.1007/978-3-319-24574-4_28

[45] D.P. Kingma, L.J.Ba, Adam: A Method for Stochastic Optimization, Proceedings of the 3rd International Conference on Learning Representations (ICLR 2015)

[46] Bishop, Christopher M. (2006). Pattern Recognition and Machine Learning, Springer-Verlag New York

[47] G. James, D. Witten, T. Hastie, R Tibshirani. (2013). An Introduction to Statistical Learning, Springer-Verlag New York

[48] Toews M, Wells W, Collins DL, Arbel T. (2010). Feature-based morphometry: Discovering group-related anatomical patterns, NeuroImage. 49(3): 2318-2327. doi: 10.1016/j.neuroimage.2009.10.032

[49] Li T, Li W, Yang Y, Zhang W. (2017). Classification of brain disease in magnetic resonance images using two-stage local feature fusion, PLoS ONE. 12(2): e0171749. doi: 10.1371/journal.pone.0171749