KJES Kerbala Journal for Engineering Science
https://kjes.uokerbala.edu.iq/

# Classification of Alzheimer's disease based on several features extracted from MRI T1-weighted brain images

Zahraa Sh. Aaraji [a]*, Hawraa H. Abbas [a]

Electrical and Electronic department, Collage of Engineering, University of Kerbala, Kerbala, Iraq
* Corresponding author, Email: zahraa.sh@s.uokerbala.edu.iq

**Received: 23 August 2021; Revised: 24 December 2021; Accepted: 25 December 2021**

***Abstract:***

Alzheimer's disease (AD) diagnosis at an early stage plays a significant role in reducing its symptoms and decelerating cognitive deterioration. Hence the use of computer-aided systems for early and accurate AD diagnosis is critical. The proposed diagnostic tool depends on classifying features extracted from brain Magnetic Resonance Imaging (MRI). These Features must accurately capture main AD-related variations of the anatomical brain structures, such as hippocampi region atrophy, lateral ventricle enlargement, cortical thickness, brain volume, etc. In this work, T1-weighted structural MRIs were employed for extracting these AD-related features. The images resulting from MRI scans are interpreted to high-intensity visible features, making preprocessing and segmentation less complex. This work has proposed a software pipeline consisting of several preprocessing steps, a segmentation method for segmenting brain tissues, and Convolutional Neural Networks (CNN) for Regions of Interest (ROIs) Parcellation that is AD-related. Features extracted from these segmented tissues and ROIs are utilized for the final AD classification using a Support Vector Machine (SVM) classifier. The results show that the proposed approach has reached 89.1% accuracy in the binary classification of AD vs. CN (Cognitively Normal), Demonstrating promising classification performance.

**Keywords:** Alzheimer's disease, Neuroimaging, Structural Brain MRI, Deep Learning, 3D CNN, U-Net, SVM.

## 1. Introduction:

In healthy aging, a brain shrinks to some degree, whilst in AD the shrinkage is more noticeable, it destroys memory and cogitation skills, and, lastly, the person's capability to do out simple daily task. AD is the most common cause of dementia, accounting for upwards of 80% of reported cases of all dementia cases [1].

AD is caused by beta amyloids and tau tangles abnormal protein deposits in the brain that damage brain cells in the memory and mental functions areas in the brain. When more neurons die, entire brain areas shrink, which leads to cognitive function problems that are the main AD symptoms. In the more developed stage, damages become diffused and the brain undergoes a huge amount of atrophy. Because beta-amyloid grows gradually over time, it will take over 10 years before a patient starts to see any obvious signs of the disease [1] Figure 1 explain protein deposits in AD patient neurons.

According to scientists, the causes seem to be related to a combination of environmental and genetic factors. The most common factors associated with the risk of developing AD are age and some environmental risk factors including smoking, strokes, heart disease, depression, arthritis, and diabetes. In 2018, more than 122,000 people died from AD, an increase of 146% from the year 2000. An estimated 5.8 million adults over 65 years are living with AD, with the number expected to more than double by the year 2050 to approximately 14 million individuals [2].

AD is an irreversible and progressive disease, meaning that no known cure is currently available for stopping or reversing the disease, therefore the medications are focused on relieving the patients from the disease symptoms. Nevertheless, few precautionary steps can be taken to reduce the

risk factors and decelerate the retrogressive growth. It has been shown that early detection and intervention of AD are effective in maintaining patients' quality of life.

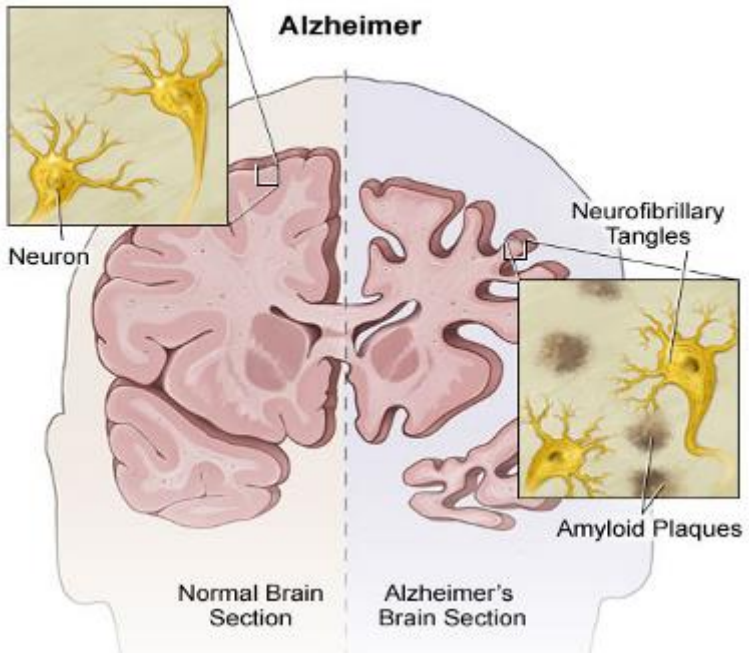



**Figure 1 Normal brain vs AD brain**

One main center of creating new memories in the brain is the Hippocampi located in the temporal lobe at both brain sides. Hippocampus atrophy is a significant biomarker that occurs in an early stage of the disease; therefore, it is considered an effective region to efficiently diagnose this type of dementia [3].

Developments in neuroimaging techniques, such as the MRI and Positron Emission Tomography (PET) scan techniques, coupled with advanced computational methods, have led to accurate prediction of the presence of the disease. MRI is a technique that creates a detailed 3D image of the brain employing magnetic fields and radio waves [4]. In this study, T1-weighted structural MRI (sMRI) was used for the classification of AD. Figure 2 shows the 3 MRI planes (taken from the ADNI* datasets).

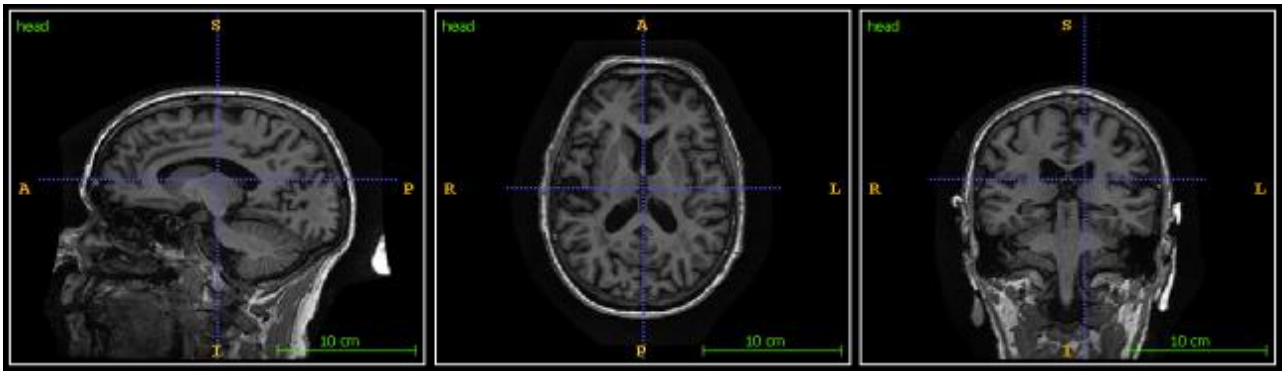



**Figure 2 T1 weighted MRI from ADNI dataset. a) Sagittal view. b) Axial view. c) Coronal view.**

This paper proposes an AD diagnosis framework that depends on machine learning models. First, several segmentation techniques were used for segmenting the brain tissues, White Matter (WM), Gray Matter (GM), and Cerebrospinal Fluid (CSF). GM volume is one of AD essential biomarkers. A convolutional neural network is utilized for the parcellation of ROIs that is affected by Alzheimer's disease, such as the hippocampus and lateral ventricles ROIs. Even though, the proposed network learns from a small dataset, it demonstrates superior performance regarding to the affected regions parcellation. Afterwards, features from both tissue segmentation and ROIs parcellation are extracted, such as volume, thickness, and shape description. Finally, the classification is performed based on extracted features using an SVM classifier. Figure 3 shows the process flowchart of the proposed framework.

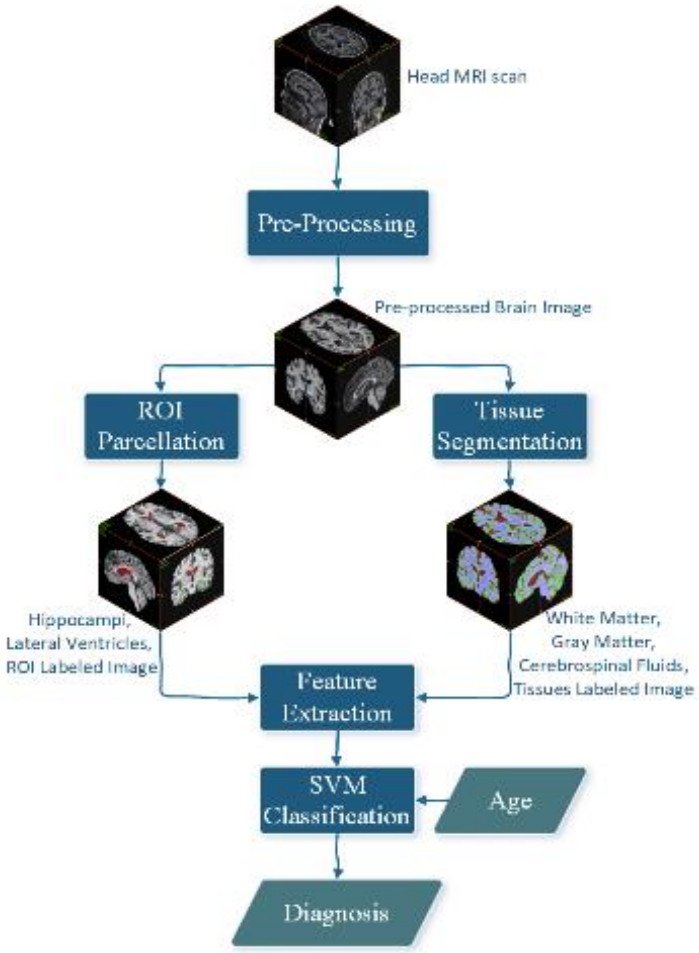



**Figure 3 Flowchart of the proposed software pipeline**

This article is organized as follows. Section 2 offers a brief review of different classification methods that have been reported in the literature for this problem. The data used in this paper is described in Section 3 and the proposed approach is introduced in Section 4. In Section 5, the experimental results are presented, and in section 6 the results are discussed. Finally, in section 7 the conclusion and future work are presented.

## 2. Related Works

Machine Learning has been used extensively in medical image analysis. Several models have been proposed to perform multiple operations on MRI images such as segmentation, registration, and classification.

The classifiers were built by using brain imaging data and clinical information for AD diagnosis. Several studies have acknowledged the importance of structural variation in the regions such as the hippocampi, cerebral cortex, and the lateral ventricles in discriminating CN brains from AD brains. In recent years, many studies focused only on extracting a single feature from sMRI images for the classification of AD and CN subjects. The accuracy achieved in their classification is relatively

low. However, the latest studies have shown that combining several features from different sMRI analysis techniques improves the accuracy of the classification [4][6].

Baglat et al. [7] applied different machine learning techniques such as Logistic Regression, Random Forest, Decision Tree, AdaBoost, and SVM for the earlier diagnosis and Classification of Alzheimer's disease using Open Access Series of Imaging Studies (OASIS) dataset, in which a significant performance was achieved using Random Forest classifier.

Vemuri et al. [8] down-sampled GM density maps from an isotropic voxel size of (1mm3) to (8mm3) by simple averaging. This step reduced the GM maps by (8) times, lowering its voxel count. Vemuri et al. performed an additional feature selection step by ensuring a condition of large margins when using a linear SVM classifier. This study was performed on (190) AD patients and (190) CN subjects. It concluded that the use of SVM with feature reduction and selection can generalize well to new data.

Ben Ahmed et al. [9] used two biomarkers: the visual feature that were extracted from the most Common region affected by AD (hippocampal area) and the surrounding CSF amount. They proposed a late fusion scheme for the classification of both biomarkers. First, the authors evaluated the approach on the baseline MR images from the (ADNI) dataset and then tested it on a 3T weighted contrast MRI obtained from Bordeaux dataset, which is a subsample of a large French epidemiological study. The experimental results Show that classification of subjects with AD vs. CN subjects achieves 87% and 85% accuracies for ADNI subset and Bordeaux dataset respectively.

Shen et al. [10] investigated the hippocampal shape variation using several Statistical Shape Models (SSM). SSMs' dimensionality was reduced using PCA, and their discriminative ability was tested using SVM classifiers. They concluded that while volume alone provides significant discrimination ability, the shape of the hippocampus can provide valuable information for the diagnosis of AD.

Toshkhujaev et al. [11] used multi-atlas label propagation with expectation–maximization-based refinement segmentation method to segment the brain images into 138 anatomical morphometry regions (in which 40 features belonged to subcortical volumes and the residual 98 features belonged to cortical thickness). The entire dataset was split into a 70:30 ratio (training stage and testing stage) before classifying the data. Principal component analysis was used for dimensionality reduction. Finally, an SVM classifier with a radial basis function was used for the classification of the two groups, AD vs. CN on the ADNI dataset.

Khedher et al. [12] extracted the density maps of GM, WM, and CSF using SPM8. Then, they exploited two feature reduction methods: Principal Component Analysis (PCA) and Partial Least

Squares (PLS), to reduce the dimensionality of the density maps. Khedher et al. tested the two feature reduction methods with two SVM classifiers: linear and Radial Basis Function (RBF). The PLS method reached a peak accuracy rate and outperformed the PCA method.

Daliri et al. [13] , used the SIFT descriptors to extract features from the whole brain of the subject for the classification of brains with and without AD. Chupin et al., Colliot et al., Liua et al. and Shen et al**.** segmented the hippocampus automatically and use their volume for the classification, In addition to volumetric methods, several surface-based shape description approaches have been proposed to comprehend the development of AD [14][17] .

In several studies, the CNN is used as a feature extractor, and the classification is performed using a SVM with linear or polynomial kernels and logistic regression Çitak-ER et al. [18] , random forest Chaddad et al. [19] , extreme ML Lin et al. [20] , SVM with different kernels Liua et al. [16], or logistic regression and XGBoost (decision trees) Shen et al. [21] . Only [21] compared the CNN classification results with those obtained with other classifiers based on CNN features extraction, and concluded that the last is more efficient.

In the above cited studies various brain regions were used: the whole brain, various cortical and subcortical regions such as hippocampus, GM, CSF etc. In our work we focus on the more significant regions, lateral ventricles and the hippocampus plus features extracted from different tissue types.

## 3. Materials

The MRIs and clinical data used in this paper were taken from Alzheimer Disease Neuroimaging Initiative (ADNI) dataset. ADNI dataset is open access and freely available to researchers. The dataset goal is to track and diagnose Alzheimer's disease in its early stages. ADNI has been running since 2004 and is currently funded until 2021. ADNI is the result of the efforts of many co-investigators from a broad range of academic institutions and private corporations. In the ADNI1 dataset, there are three diagnostic groups: AD, CN, and MCI. AD group refers to the patients with Alzheimer's diagnosis, CN group refers to normal cognitive status subjects that show no sign of AD. MCI group references patients that can take care of their daily activities with mild damage in other cognitive areas. The images used in this paper are 3-dimensional T1-weighted structural MRI taken from baseline and screening ADNI1 group. The number of subjects participating is 835 divided into three groups:

- 228 CN subjects.
- 407 Mild Cognitive Impairment (MCI) subjects.

- 200 AD subjects.

In this work, 200 subjects are chosen randomly from each group to provide the MR images and sustain the balance between classes classification bias aiming to classify AD vs. CN.

## 4. Methodology

In this section, the proposed classification pipeline will be described in detail. As shown in figure 3, the procedure used for the MRI classification, includes four main steps pre-processing, post-processing, feature extraction, and classification. All work was implemented using MATLAB version; 2021a, on a single PC with the following specifications:

- Intel 10th Gen Core i7 with 8C/16T Processor.
- Nvidia RTX 2070 with 8 Gigabyte of VRAM.
- 32 Gigabyte of system RAM.

### 4.1 Pre-Processing Stage

The image data downloaded from ADNI has already undergone several pre-processing steps. This was done to limit the MRI disparities from different scanners, for better comparability across subjects. The pre-processing steps from ADNI include distortion correction, B1 non-uniformity bias field correction. Further pre-processing steps in this paper are performed using MATLAB which Consists of:

- Brain extraction (skull-stripping).
- Intensity normalization.
- Registration

#### 4.1.1 Brain Extraction

Brain extraction is an essential step in medical image analysis to remove non-brain tissues like the skull, air, facial tissues, fat, dura matter, etc. An accurate brain extraction method should exclude non-brain tissues without removing any part of the brain. ADNI Provides a brain tissue mask named (MIDAS) that is created with a semi-automatic segmentation method. These masks accurately define the main brain tissues: WM and GM but they exclude all interior and exterior CSF. Therefore, we used voxel-based morphological operations (dilation followed by erosion) to include the CSF. Figure 4 shows the adopted mask used for extracting brain tissues and the surrounding CSF.

*(a)* *(b)*

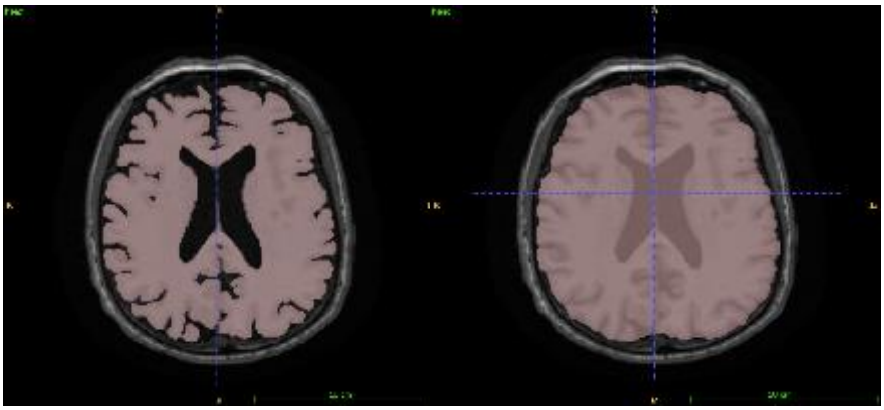


**Figure 4 the difference between a) MIDAS mask and b) Adopted mask**

### 4.1.2 Intensity Normalization

Different MRI scanners and parameters result in large intensity variation between MRIs, this would greatly undermine the quality and performance of the following processes, especially in the post-processing stage, lowering the final classification accuracies. The simple Intensity scaling technique is chosen based on three intensity regions:

- CSF intensity (low intensity region).
- GM intensity (medium intensity region).
- WM intensity (high-intensity region).

The three intensity regions are segmented using K-means clustering algorithm, the computed means is shifted to match the Montreal Neurological Institute (MNI) ICBM152 template.

### 4.1.3 Registration

Another issue of using different scanners with different parameters is the spatial variation between MRIs. This means that a subject's head is not located at the same coordinates, or has the same scale and angle concerning another subject's head.

To find correspondence between subjects, image registration is a necessity, in which a brain region in one subject MRI is correspondent to the same region in all other subjects MRIs. This helps to reduce the complexity of the following post-processing stage and increase ROIs parcellation accuracy and robustness.

Affine registration is adopted for the brain alignment task. Affine registration is a rigid-body linear registration that tends to preserve parallelism, i.e., does not deform brain structures, but rather manipulate linear parameters:

- Translation (head location).

- Rotation (head angle).
- Scaling (voxel size).
- Shearing (planar direction).

All subjects are registered to the ICBM152 2009a template, which is a nonlinear symmetric T1 weighted brain MRI with voxel size of $1\times1\times1$ $mm^3$.

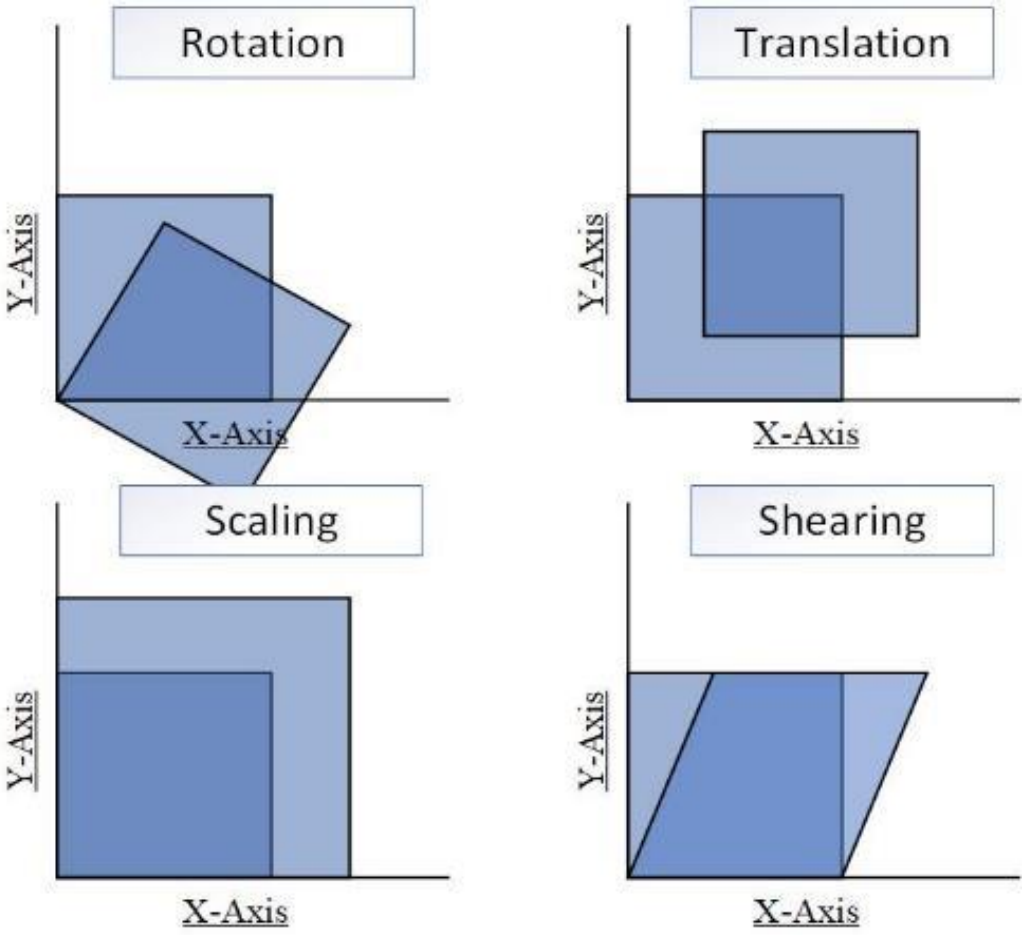



**Figure 5 2D representation of Image Registration Parameters: Rotation, Translation, Scaling and Shearing.**

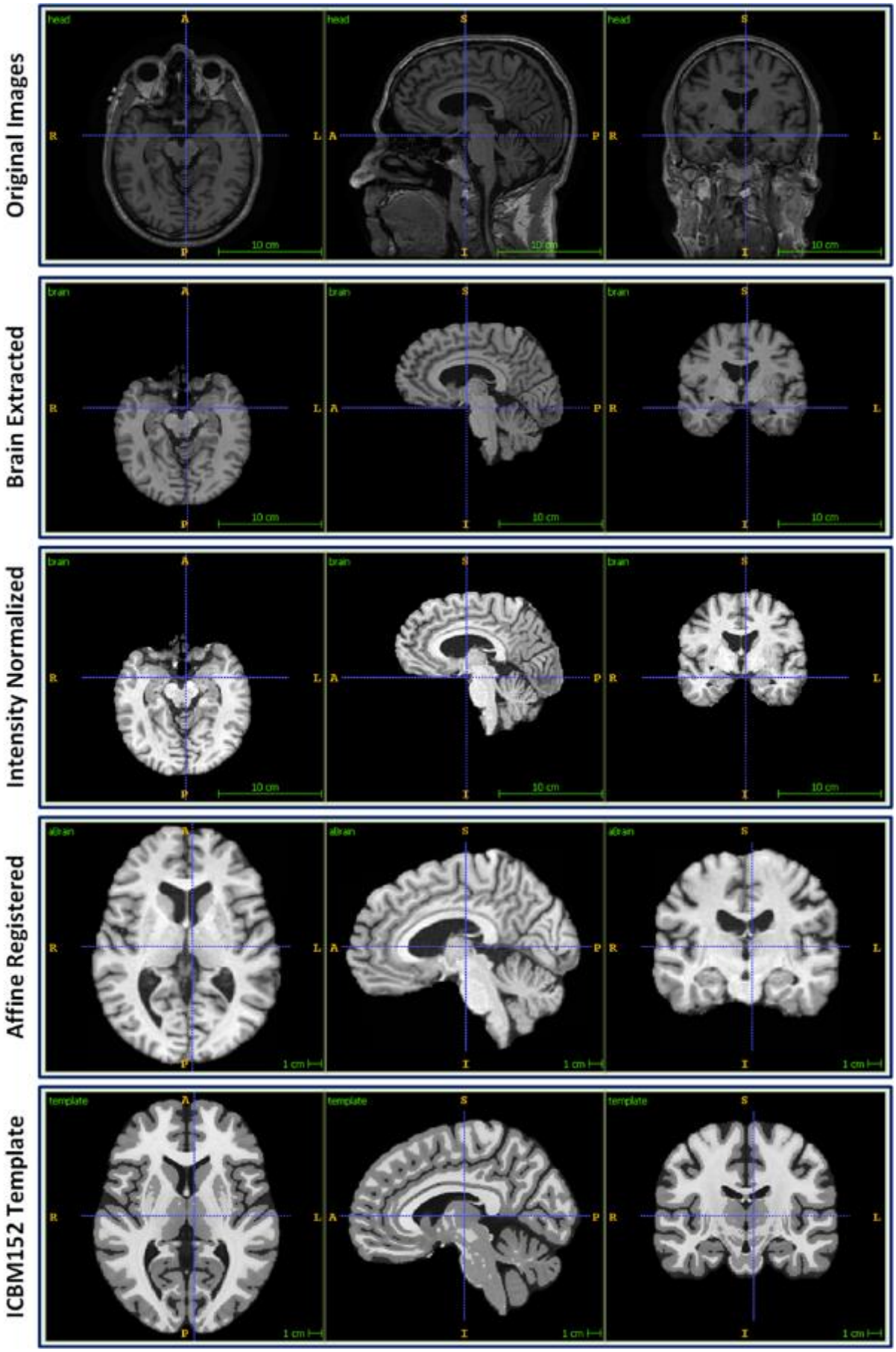



**Figure 6 Pre-processed MR images and ICBM152 template**

### 4.2 Post-Processing Stage

The pre-processed MRI images will be further processed to segment brain tissues and extract the AD-related ROIs.

#### 4.2.1 Brain Tissue Segmentation

Image segmentation is the process of labeling every pixel in an image with the purpose that the pixels with the same label share certain characteristics. This process aims to reduce the complexity of the image, and thus analyzing the image becomes Simpler. In image processing there are various techniques available for the segmentation process, in this study, segmentation methods used in previous works are utilized and evaluated against each other to find the best segmentation method. References [22][26] explored the statistical approaches which segment an image based on its voxels' intensity such as K-means, Otsu, and Fuzzy C-means. The Genetic Algorithm (GA) and Particle Swarm Optimization (PSO) methods are compared against Hidden Markov Random Field (HMRF) statistical model for segmentation refinement. HMRF is a probabilistic model that can improve the segmentation by using other image characteristics such as voxel neighborhood [27] . Evaluation of these methods is performed upon two factors; time of execution and quality of segmentation.

Brain segmentation aims to extract the GM and CSF as a biomarker for AD diagnosis. Table1 shows the average evaluation accuracies for seven different segmentation algorithms applied to brain MRI images.

**Table 1 Evaluation of Segmentation methods useg for segmenting brain tissues**

| Method | Accuracy | Sensitivity | Specificity | Dice |
|---|---|---|---|---|
| K-means | 92.585% | 89.639% | 94.388% | 89.148% |
| OTSU | 90.174% | 86.410% | 92.644% | 85.597% |
| FC-means | 90.686% | 87.012% | 92.973% | 86.351% |
| **K-means + HMRF-EM** | **96.522%** | **94.943%** | **97.317%** | **94.920%** |
| OTSU + HMRF-EM | 95.587% | 93.793% | 96.661% | 93.527% |
| K-means + PSO | 92.588% | 89.639% | 94.388% | 89.148% |
| Genetic Algorithm | 92.590% | 89.647% | 94.392% | 89.155% |

As shown in Table 1, the K-means clustering technique followed by Hidden Markov Random Fields (HMRF) technique gives the best accuracy amongst other segmentation methods. This is because the segmentation method depends on two features instead of one, the first segmentation feature

is the voxel intensity, and the second feature is the voxel neighborhood. K-means alone is weak against noise since it does not take voxel spatial information into consideration. K-means provides the initial segmentation for the HMRF algorithm depending only on voxel intensity, whilst HMRF will further optimize the segmentation depending on both voxel intensity and neighborhood

### 4.2.2 ROIs parcelation

Hippocampus is a complex brain structure located in the depth of the temporal lobe and it has a major role in forming new memories in human brains. References [28], [29] showed that the loss in hippocampal volume helped to distinguish very mild AD from healthy Aging.

The ventricles in the other hand, are two interconnected cavities distributed throughout the brain. Lateral ventricular enlargement is an important and key abnormality biomarker used in early detection of AD. Not only volume dilation but also the average change in ventricular volume has been studied as a potential surrogate marker to predict the improvement and progression of this type of dementia [6]. Figure 7 shows the Hippocampus atrophy and Lateral ventricles enlargement in AD patient compared to NC subject both downloaded from ADNI dataset.

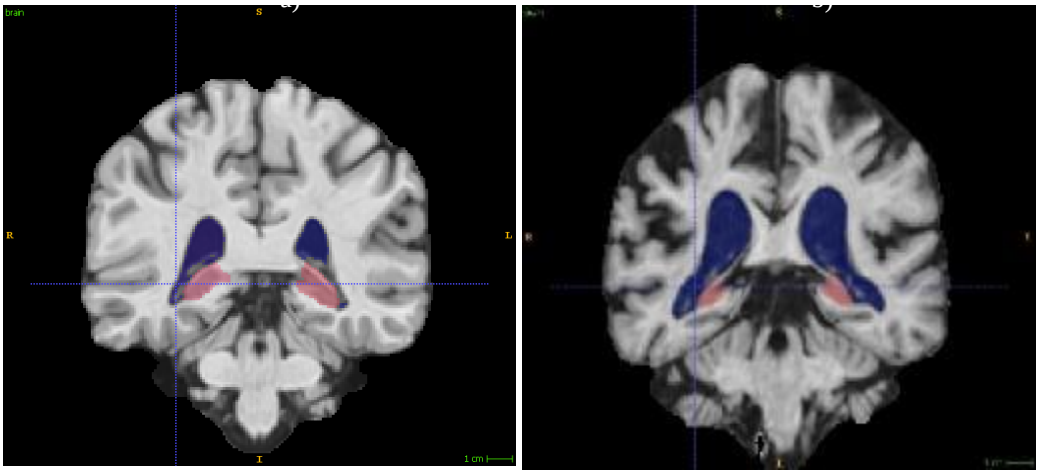


**Figure 7 The difference in the hippocampus (colored in red) and lateral ventricles (colored in blue) between CN brain (left image) and AD brain (right image). The images are taken from the coronal view.**

Since both the hippocampus and the lateral ventricles have common patterns across all human brain, both can be parceled by the use of pattern recognition, in which a CNN dominates. CNN is one of many deep learning classes best used for analyzing visual imagery, which suits the ROI parcellation task by means of voxel classification [30] .

Usually, CNN includes one or several convolution layers, one or several optionally pooling layers, non-linearity layers, and input and output layers. Each image passes through all CNN layers.

A set of 3D filters are found in each convolutional layer, these 3D filters play an essential role in CNNs architecture, when images are inputted to a CNN, A series of learnable filters are applied upon the image to detect specific patterns and features, this application of filters is a convolution operation, hence the layer name. Each filter slides over the input image to compute the feature maps. Rectified linear units (ReLU) is a type of activation function, which introduces non-linearity to the CNN by removing all the negative values and keeping the positive values, this concept helps the CNN to converge faster to the optimal solution. The pooling layer performs down-sizing (down-sampling) and reduces the feature maps spatial dimensionality. Therefore, reduce network computational complexity. Deconvolution layers, on the other hand, is a mathematical operation that works the opposite way of a convolutional layer and max-pooling layer. Each voxel of the input feature map slides over the filter, as a result, the input feature map is up-sampled. The classification layer is the network's last layer, this layer outputs probability maps (heat maps) for voxel classification. In the forward learning stage, this layer calculates the output loss (misclassification rate).

The architecture used in the parcellation process is the U-net [31] . It is composed of two stages, encoder, decoder, and a bridge between them. Each stage includes four blocks, and each block has seven layers. Each block in the encoder stage consists of two convolutional layers, two normalization layers, two activation layers and max-pooling at the beginning.

In the decoder stage, there are seven layers in each block, however there is only one deconvolution layer at the end of each block and no max-pooling layer at the beginning. The Bridge that separates the two stages consists of a max polling layer, a deconvolution layer in addition to the main six layers. Figure 8 shows the architecture adopted in this experiment.

The U-net input is a 3D MRI brain image and the output is the labeled image. The network can learn the patterns to classify voxel if provided with the desired output; in this case, a ground-truth image. In the forward learning, the final layer; the classification layer, computes the dice [32] coefficient which represents the output loss. In the backpropagation, the network computes the gradients of all the learnable parameters to the output loss. These gradients can be used to update all learnable parameters in the right direction and by the right amount.

Input images are cropped to contain regions of the lateral ventricles only and discard other unneeded brain regions. The cropped images can reduce U-net complexity and speed up the training phase since they have smaller dimensions than whole-brain MR images.

Ground-truth images for the hippocampus were provided by ADNI, while ground-truth images for the lateral ventricles were created using ITK-SNAP software. ITK-SNAP is a software application used to segment structures in 3D medical images, and it has various segmentation tools for this task.

Figure 9 shows all resulting images from the necessary processes on the subject downloaded from ADNI datasets

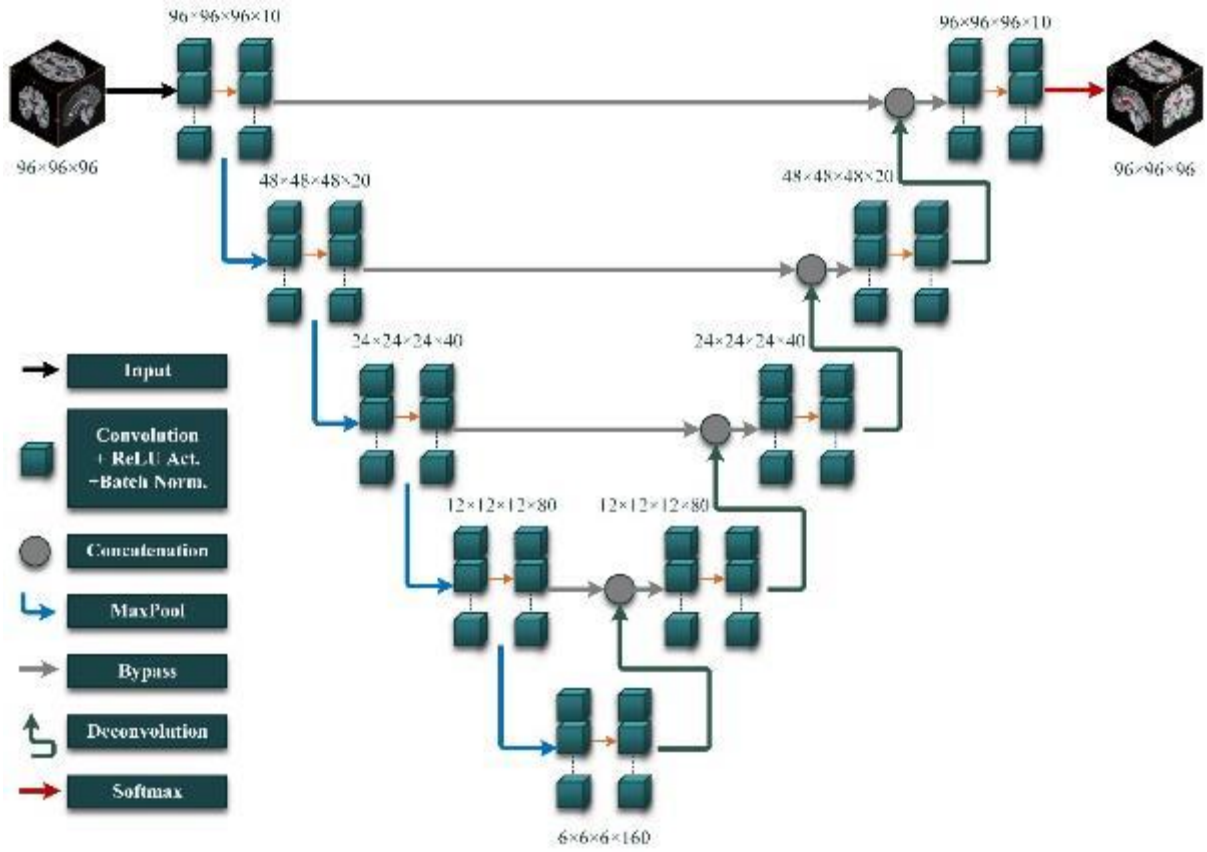



**Figure 8 U-net architecture used in ROI parcellation**

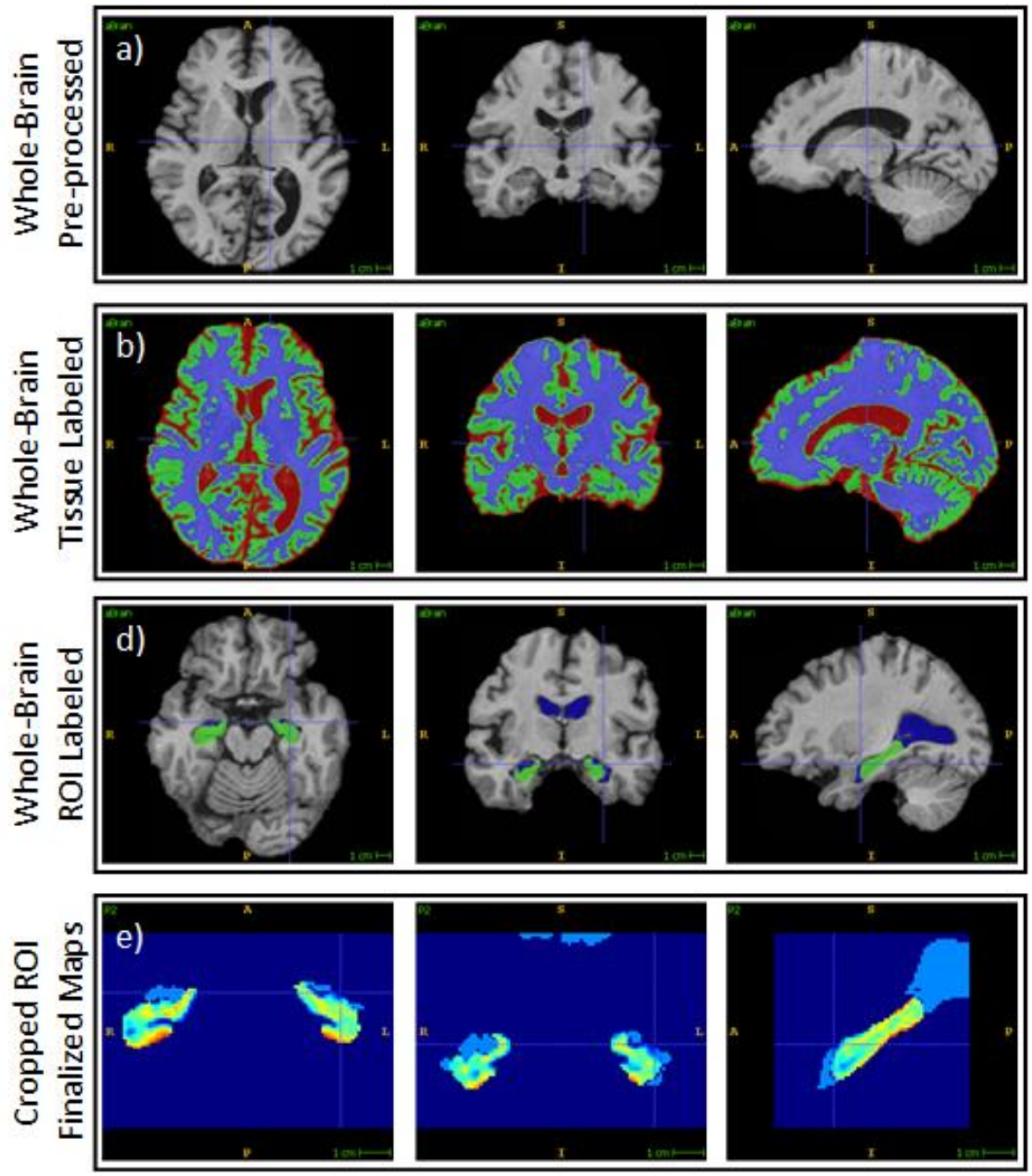



**Figure 9 Axial, sagittal, and coronal view respectively left to right in a head scan from a single-subject MRI session along with all resulting images from the necessary processes.**

### 4.3 Feature Extraction:

After brain tissue segmentation and parcellation of ROIs that are most affected by Alzheimer's dementia, the next step is feature extraction from both segmentation and parcellation label images. These features represent information about brain structures such as GM thickness and volume, CSF amount, brain volume, hippocampus volume, lateral ventricles volume, etc.

Since all MRI brain images are spatially normalized and affine registered to the MNI template, the number of voxels assigned to a label represents the normalized volume of that label. This means

that all subjects have roughly the same mask volume. To compute the brain volume, only WM, and GM voxels are calculated.GM thickness is computed using distance transform across voxels with GM label resulted in the segmentation process. Distance transform calculates the lowest distance between a GM label voxel and a non-GM label voxel. This procedure has additional benefit of reducing the segmentation noise that is not removed in the HMRF part of the segmentation result, i.e., a voxel with no neighbors that has a distance value of zero.

Age is an important risk factor of AD. Subjects with age beyond 65 years old are more susceptible to AD while younger adults are highly unlikely to have AD. Hence, all features mentioned above should be corrected for the subject's age by adding the age as another feature [33].

**4.4 Features Classification:**

In machine learning the classification is the process of recognizing, understanding, and grouping objects into preset categories. The classifier is an algorithm that sorts unlabeled data into labeled classes or categories of information. Many classification algorithms exist, supervised and unsupervised. It is not possible to conclude which one is superior to the others since it depends on the application and the nature of the available data set. In this work, several machine learning models had been implemented such as logistic regression, decision tree, Fisher's linear discriminant, SVM, naïve Bayes classifier, and Nearest neighbor classifiers. Table 2 shows the accuracy of each mode to compare between classifiers. The best result was achieved using an SVM classifier for binary classification of AD vs CN.

**Table 2 shows the comparison between classical machine learning techniques in classifying the extracted features.**

| Methods | Accuracy |
|---|---|
| **SVM** | **93.3%** |
| Logistic Regression | 91.7% |
| Naïve Bayes Classifier | 89.0% |
| Fisher's Linear Discriminant | 92.7% |
| Decision Trees | 90.7% |
| Neural Networks | 92.3% |
| K-Nearest Neighbors | 87.1% |

## 5. Experimental Results:

This section provides all results for the parcellation and classification processes.

### 5.1 U-net Training Results:

In this paper, images used to train the U-net architecture are 3D T1-weighted structural brain MRI downloaded from the ADNI1 dataset. Table 3 shows the number of ground-truths for both the hippocampus and the lateral ventricles.

**Table 3 Number of ground-truths used in the ROIs parcellation network training. The number of validation sets are 30% of the training sets and the test sets**

| | **Hippocampus GTs** | **Lateral Ventricles GTs** |
|---|---|---|
| **All** | 487 | 180 |
| **Training** | 300 | 110 |
| **Validation** | 100 | 35 |
| **Test** | 87 | 35 |

Prior to the training, ground-truths are separated randomly into 3 sets. The validation set is used to observe network generalization with respect to the whole dataset patterns. Network training was optimized using the Stochastic Gradient Descent with Momentum (SGDM). The network was trained several times to fine-tune the hyperparameters of the optimization algorithm. This prevents the network from overfitting or underfitting and converges to a global optimum.

The training of lateral ventricles parcellation network converged to the global optimum with no over or underfitting since the patterns were simple to recognize and the intensity differences of the boundaries separating the ventricles from its surrounding GM was clear. Whilst in the hippocampus parcellation network, an overfit of 7% was produced due to MRI resolution being relatively low. Hippocampus size is very small and its boundaries were indistinct. This problem can be solved using a higher resolution MRI. The overfitting was too small to produce high parcellation error and the network did generalize well enough to the whole dataset. Figure 8 shows the chart of parcellation dice coefficients. Figure 9 shows the MATLAB training plot for ROIs parcellation U-net training.

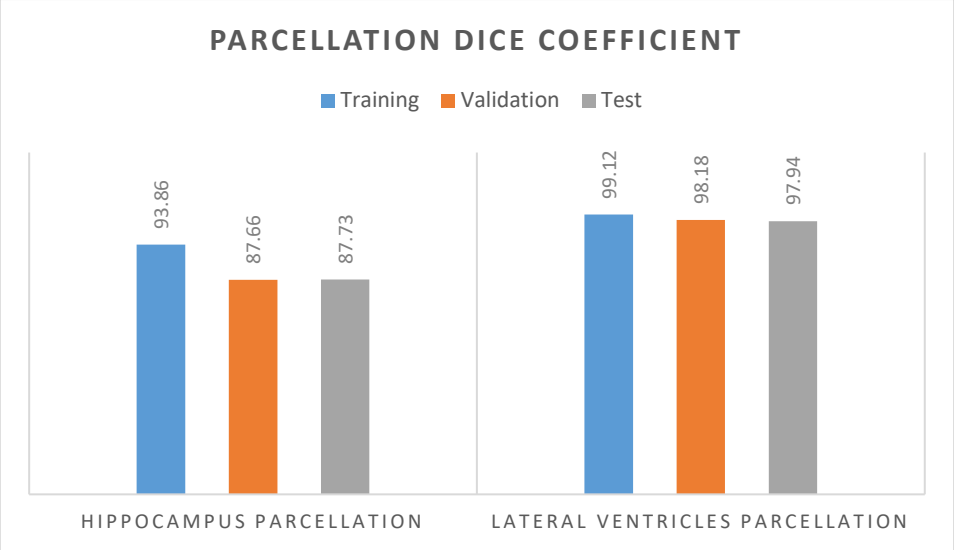
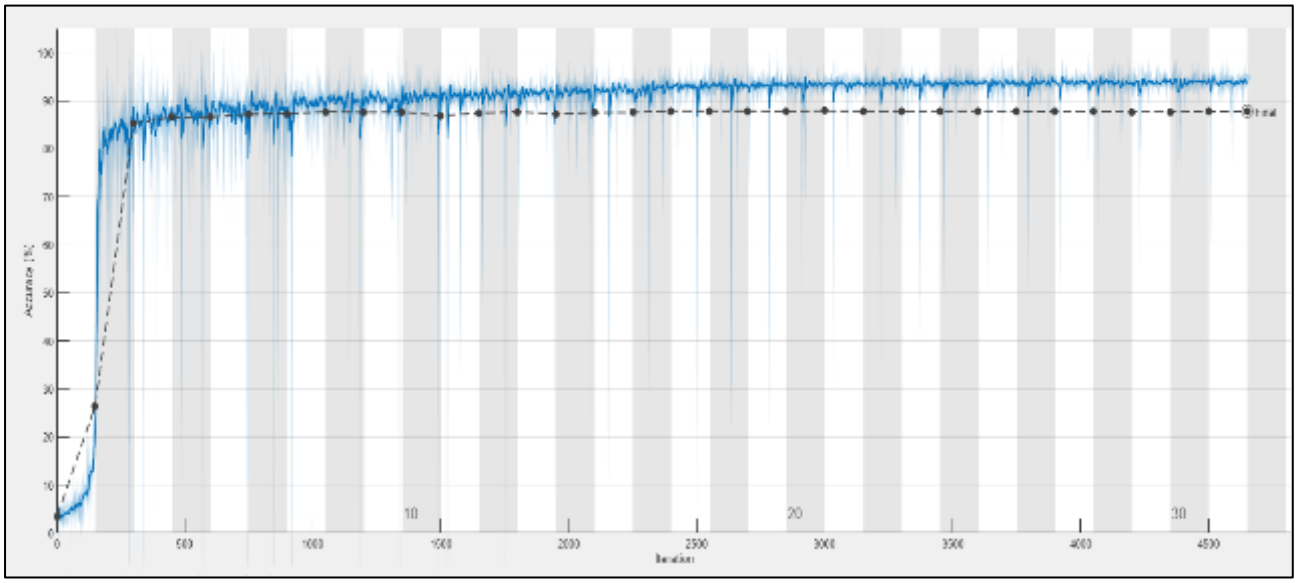



**Figure 10 Dice coefficients for training, validation and test sets ROIs parcellation U-net.**

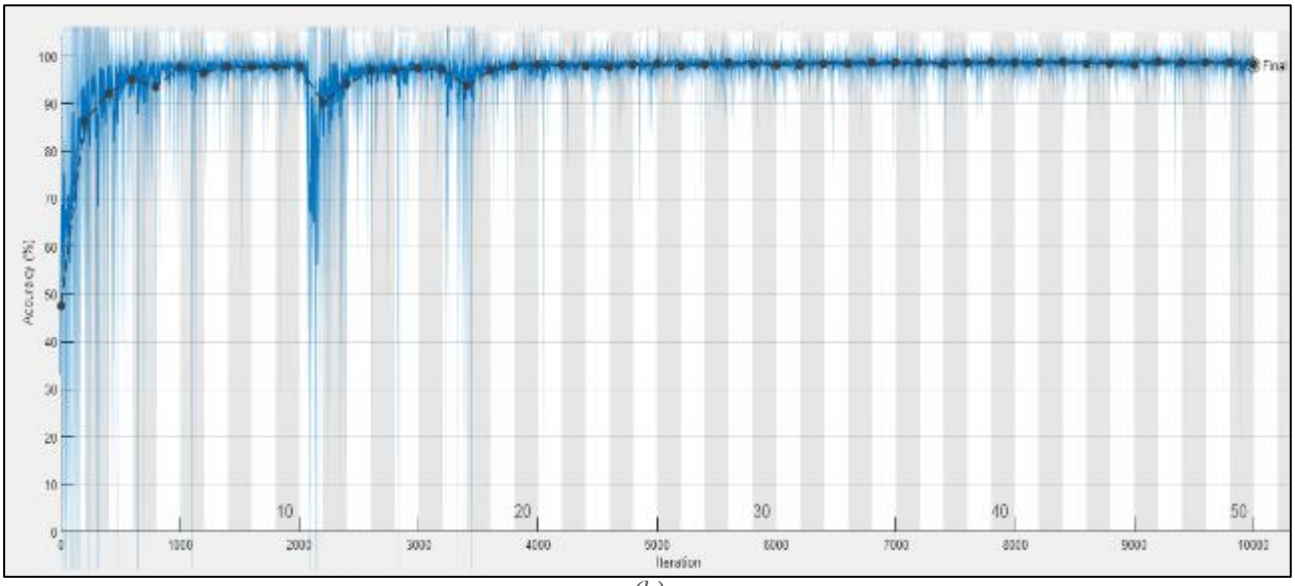
### 5.2 SVM Classification Results:

After segmentation and parcellation, hippocampus volume, lateral ventricles volume, GM thickness, and CSF amount will be computed to be used in subjects classification. Two other features will be added to the classifier to improve upon classification accuracy, brain atrophy (whole brain volume) and the subject's age. Brain atrophy is computed by counting voxels with GM and WM labels from the segmented images, i.e., CSF voxels excluded.

Two classes are used to train the SVM classifier in a binary classification way, AD vs CN. The response data used for the SVM training is the diagnosis for each MRI session downloaded from the ADNI website.



*(a)*



*(b)*

**Figure 11 Training plots for the U-nets used for ROIs parcellation. a) Hippocampus Parcellation U-net, b) Lateral Ventricles Parcellation U-net**.

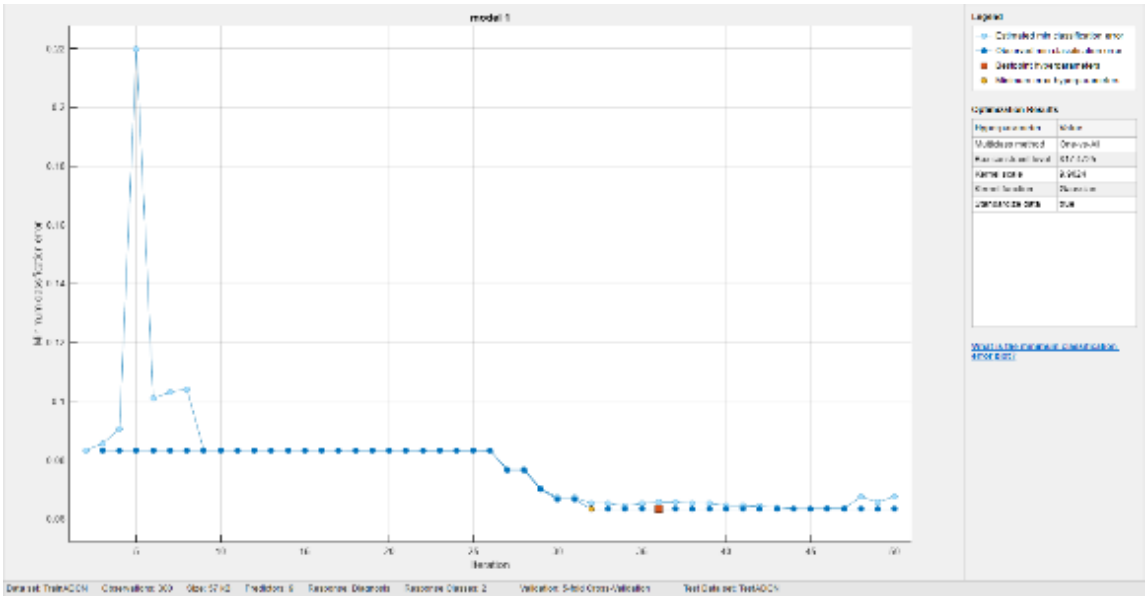


**Figure 12 SVM minimum classification error optimization plot to fine tune hyperparameters and choose the best kernel function**

The data used to train SVM consists of 200 AD MRI brain images and 228 CN MRI brain images. Data is divided into 180 training and validation sets for each class with a 5-fold cross-validation training. The rest is used as test sets.

SVM training is optimized to choose the best hyperparameters; specifically the kernel which determines the shape of the hyperplane that separates data into classes. This optimization depends on the validation set classification error to define the best hyperparameters.

The best results were achieved using the Gaussian kernel function. Figure 10 shows the SVM optimization plot from the MATLAB classification learner toolbox. Figure 11 shows the final validation and test accuracies for the binary classification of AD vs. CN.

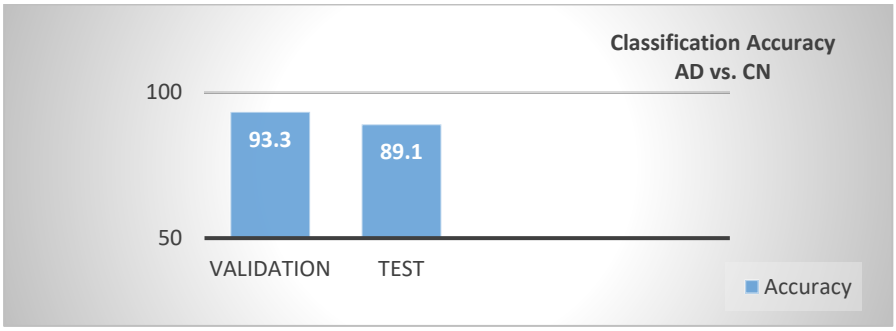


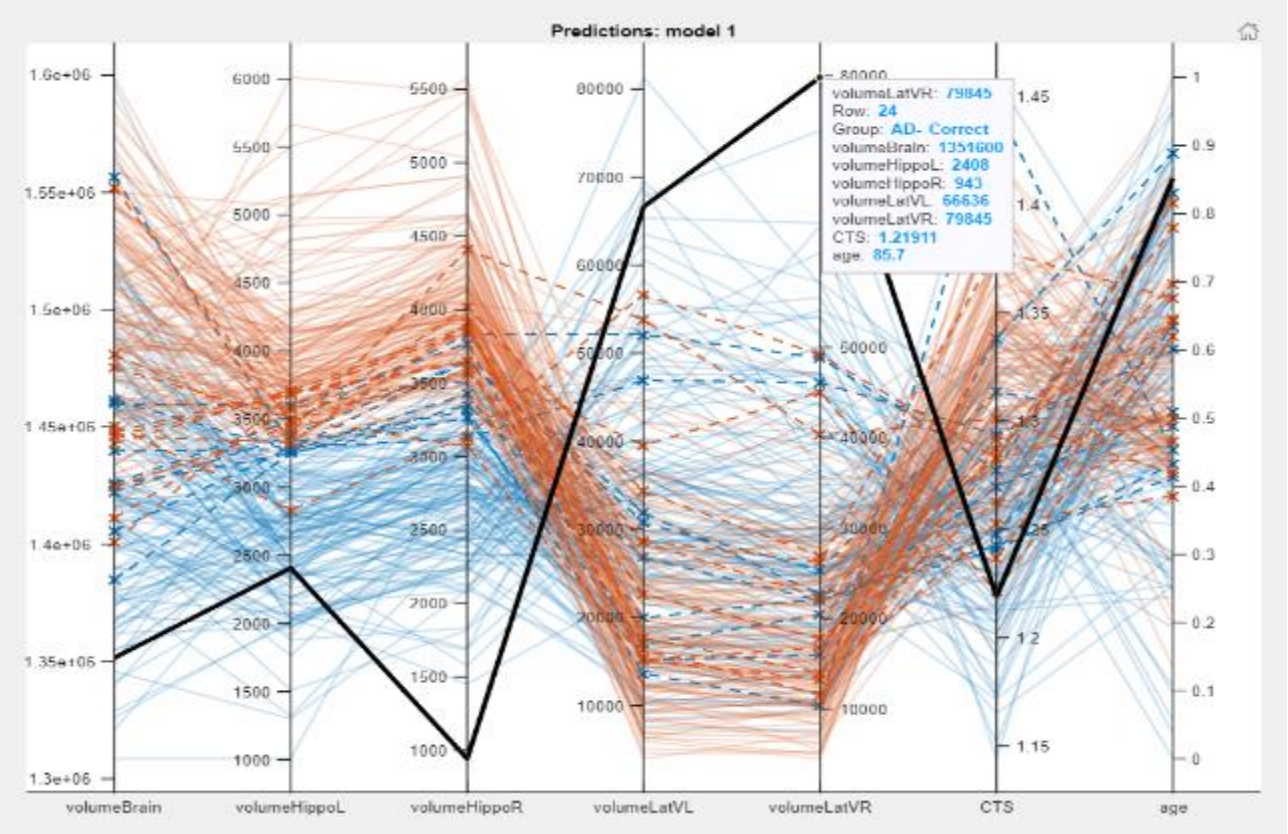

**Figure 13: The final classification accuracy of AD vs. CN using SVM.**

### 5.3 Final pridection model

Figure 12 shows the final prediction model for AD vs CN. The red line represents CN subjects' classification and the blue lines represent AD subjects' classification. The black line in Figure 12a

shows an example of a correct prediction of a subject affected by AD, volumes of his brain and that both (right and left) hippocampi are lower than normal stating atrophy. While volumes of both (left and right) lateral ventricles are higher than normal stating enlargement. The GM thickness score is also lower than normal stating GM atrophy. The black line in Figure 12b shows a correct prediction for a healthy subject. This line is opposite in direction to the black line plotted in Figure 12a, meaning that the subject has no atrophy relative to the red lines in the brain and both hippocampus volume, has no enlargement in the ventricles and the GM thickness is relatively high.

*(a)*




*(b)*

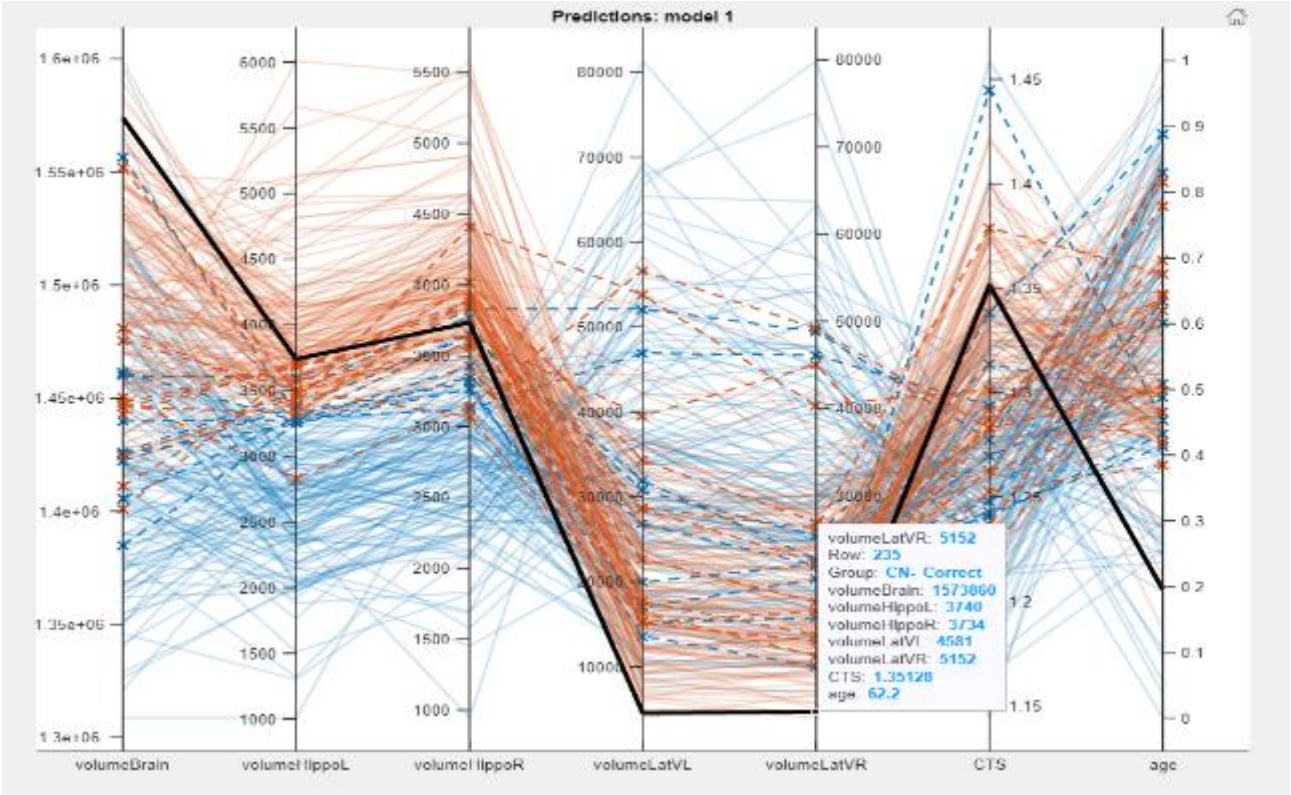



**Figure 14 The SVM prediction model for AD vs. CN with features on the x-axis. a) Black line shows a correct classified AD patient, b) black line shows a healthy subject.**

## 6. Discussion

In this study, we proposed a classification software pipeline that combines CNN and classical machine learning techniques to be applied to the most commonly acquired anatomical MRI of the brain. The work aimed to accomplish two goals: classification of AD vs. CN, and identification of the complex brain structural variations related to AD.

Table 4 shows the comparison with some of the previous method reported in the literature, Reference [7] Achieved 86.84% in AD discriminant problem by using random forest classifier. The implementation of this work is not explicit, no description of the methodology is provided on how the data is fed in to the classifiers.

References [8], and [12] used the density maps resulted from segmenting the brain tissues. Their feature selection methods were voxel-wise that cannot represent fully detailed patterns. Their methods also suffered from:

1. They require intensive processing steps before manual feature extraction and selection.
2. Highly prone to the processing errors such as registration, normalization, segmentation, etc.

Reference [9] used combination of two types of features e.g., Visual feature obtained from the hippocampus and accumulated surrounding CSF; then adding supplementary voxels from the lateral ventricles to enhanced the performance. The final result achieved are 87% and 85% accuracies for ADNI subset and Bordeaux dataset respectively, which is lower than results obtained in the present work.

Reference [11] achieved an accuracy of 91.57% on ADNI dataset. The author used MALPEM tool for the segmentation of the 138 ROIs. The time period for segmenting one subject by this tool is between 8 and 10 hours. The author also supplemented these features by age, gender and education.

These feature selection methods were voxel-wise that could not represent fully detailed patterns. Except [8] who did average down-sampling to GM density maps, the rest suffered from the same issue of being prone to registration error.

**Table** Error! No text of specified style in document. **Comparison of proposed ML method with state-of-art methods**.

| Method | AD vs. CN |
|---|---|
| Baglat et al. [7] | 86.8% |
| Ahmed et al. [9] | 87.00% |
| Shen et al. [10] | 88.00% |
| Khedher et al. [12] | 88.49% |
| Vemuri et al [8]. | 89.30% |
| Toshkhujaev et al. [11] | 91.57% |
| Proposed ML method | 89.10% |

In this work, to discriminate AD from CN, brain volume, hippocampi volumes, lateral ventricles volumes, GM thickness score, and age are adopted. Features extraction was automated using two segmentation methods; HMRF for tissue segmentation and U-net architecture for ROIs parcellation. Age is another important factor to enhance the accuracy of AD classification. The ML model Training required less than 5 seconds to train the model and less than 2 minutes to optimize it since only 7 features (7-dimensions) are employed to train the model. The ML model achieved an accuracy of 93.3% in validation and 89.1% in testing. Table 4 illustrates that the proposed method achieved a classification result in line with the state-of-art methods.

## 7. Conclusion:

Early and efficient diagnosis of AD helps maintain patient quality of life and to adopt a different lifestyle to slow down the disease progression. This was a challenging task which many authors are focusing on; they had developed many computer-aided diagnosis (CAD) systems to perform the diagnosis of AD. This paper explains an automatic AD diagnosis system that is based on deep learning on a 3D brain MRI.

The main contribution of this paper is the utilization of the U-net architecture and the HMRF probabilistic model for the segmentation of brain tissues and AD-related ROIs. Both methods achieved accurate image segmentation compared to other methods while being highly efficient.

The HMRF is used as a method for three tissue segmentation: WM, GM, and CSF which achieved the highest accuracy than another segmentation method. For hippocampus and ventricles parcellation, a main advantage of CNNs compared to other classic parcellation methods is that patterns can be automatically recognized from raw data without any expert supervision.

This paper utilized U-net architecture which was invented for medical image analysis, to parcellate both ROIs. The U-net achieved 97.94% and 87.73% test parcellation dice coefficient for hippocampus and lateral ventricles respectively.

Features extracted from the segmentation and parcellation are supplemented with age scores to improve the classification. In the binary features' classification step of AD vs CN, the SVM classifier achieved the highest accuracy than other classification methods which reached 93.3 %, 89.1% validation, and testing accuracies respectively.

The rate of classification error is due to the ground truths for the hippocampus were not accurate enough in addition to the noise in the MRI images that have low resolution. The experimental results on the ADNI dataset demonstrated that the proposed model gives an accurate prediction of the binary classification. Lastly, it can be shown that for the AD classification task, our proposed model achieves results within the same rate as state-of-the-art models.

In the future, efforts can be made to improve these results by combining multi-modality clinical data different features can be obtained from different brain neuroimaging techniques which can be integrated to enhance the capability to recognize the images patterns.

**8. Refrences :**


[1] J. Weller and A. Budson, "Current understanding of Alzheimer's disease diagnosis and treatment," *F1000Research*, vol. 7, p. F1000 Faculty Rev-1161, Jul. 2018, doi: 10.12688/f1000research.14506.1.

[2] "Dementia," *World Health Organization*, Sep. 02, 2021. https://www.who.int/en/news-room/fact-sheets/detail/dementia

[3] D. Amaral and P. Lavenex, "Hippocampal neuroanatomy." in *The hippocampus book.* New York, NY, US: Oxford University Press, 2007, pp. 37–114.

[4] P. Vemuri and C. R. Jack Jr, "Role of structural MRI in Alzheimer's disease," *Alzheimer's research & therapy*, vol. 2, no. 4, p. 23, Aug. 2010, doi: 10.1186/alzrt47.

[5] G. B. Frisoni, N. C. Fox, C. R. Jack Jr, P. Scheltens, and P. M. Thompson, "The clinical use of structural MRI in Alzheimer disease," *Nature reviews. Neurology*, vol. 6, no. 2, pp. 67–77, Feb. 2010, doi: 10.1038/nrneurol.2009.215.

[6] S. M. Nestor *et al.*, "Ventricular enlargement as a possible measure of Alzheimer's disease progression validated using the Alzheimer's disease neuroimaging initiative database," *Brain : a journal of neurology*, vol. 131, no. Pt 9, pp. 2443–2454, Sep. 2008, doi: 10.1093/brain/awn146.

[7] P. Baglat, A. W. Salehi, A. Gupta, and G. Gupta, “Multiple Machine Learning Models for Detection of Alzheimer’s Disease Using OASIS Dataset,” 2020, doi: 10.1007/978-3-030-64849-7_54.

[8] P. Vemuri *et al.*, “Alzheimer’s disease diagnosis in individual subjects using structural MR images: Validation studies,” *NeuroImage*, vol. 39, no. 3, Feb. 2008, doi: 10.1016/j.neuroimage.2007.09.073.

[9] O. ben Ahmed, J. Benois-Pineau, M. Allard, C. ben Amar, and G. Catheline, “Classification of Alzheimer’s disease subjects from MRI using hippocampal visual features,” *Multimedia Tools and Applications*, vol. 74, no. 4, Feb. 2015, doi: 10.1007/s11042-014-2123-y.

[10] K. Shen, J. Fripp, F. Mériaudeau, G. Chételat, O. Salvado, and P. Bourgeat, “Detecting global and local hippocampal shape changes in Alzheimer’s disease using statistical shape models,” *NeuroImage*, vol. 59, no. 3, Feb. 2012, doi: 10.1016/j.neuroimage.2011.10.014.

[11] S. Toshkhujaev *et al.*, “Classification of Alzheimer’s Disease and Mild Cognitive Impairment Based on Cortical and Subcortical Features from MRI T1 Brain Images Utilizing Four Different Types of Datasets,” *Journal of Healthcare Engineering*, vol. 2020, p. 3743171, 2020, doi: 10.1155/2020/3743171.

[12] L. Khedher, J. Ramírez, J. M. Górriz, A. Brahim, and F. Segovia, “Early diagnosis of Alzheimer′s disease based on partial least squares, principal component analysis and support vector machine using segmented MRI images,” *Neurocomputing*, vol. 151, Mar. 2015, doi: 10.1016/j.neucom.2014.09.072.

[13] M. R. Daliri, “Automated Diagnosis of Alzheimer Disease using the Scale-Invariant Feature Transforms in Magnetic Resonance Images,” *Journal of Medical Systems*, vol. 36, no. 2, pp. 995–1000, 2012, doi: 10.1007/s10916-011-9738-6.

[14] M. Chupin *et al.*, “Fully automatic hippocampus segmentation and classification in Alzheimer’s disease and mild cognitive impairment applied on data from ADNI,” *Hippocampus*, vol. 19, no. 6, pp. 579–587, Jun. 2009, doi: 10.1002/hipo.20626.

[15] O. Colliot *et al.*, “Discrimination between Alzheimer Disease, Mild Cognitive Impairment, and Normal Aging by Using Automated Segmentation of the Hippocampus,” *Radiology*, vol. 248, no. 1, Jul. 2008, doi: 10.1148/radiol.2481070876.

[16] Y. Liu *et al.*, “Combination analysis of neuropsychological tests and structural MRI measures in differentiating AD, MCI and control groups—The AddNeuroMed study,” *Neurobiology of Aging*, vol. 32, no. 7, Jul. 2011, doi: 10.1016/j.neurobiolaging.2009.07.008.

[17] Kai-kai SHEN, “Automatic segmentation and shape analysis of human hippocampus in Alzheimer’s disease,” Sep. 2011.

[18] F. Citak-Er, D. Goularas, B. Ormeci, and T. Initiative, “A novel convolutional neural network model based on voxel-based morphometry of imaging data in predicting the prognosis of patients with mild cognitive impairment,” *Journal of Neurological Sciences*, vol. 34, pp. 52–69, Jan. 2017.

[19] A. Chaddad, C. Desrosiers, and T. Niazi, “Deep Radiomic Analysis of MRI Related to Alzheimer’s Disease,” *IEEE Access*, vol. 6, 2018, doi: 10.1109/ACCESS.2018.2871977.

[20] W. Lin *et al.*, “Convolutional Neural Networks-Based MRI Image Analysis for the Alzheimer’s Disease Prediction From Mild Cognitive Impairment,” *Frontiers in Neuroscience*, vol. 12, Nov. 2018, doi: 10.3389/fnins.2018.00777.

[21] T. Shen, J. Jiang, Y. Li, P. Wu, C. Zuo, and Z. Yan, “Decision Supporting Model for One-year Conversion Probability from MCI to AD using CNN and SVM,” Jul. 2018. doi: 10.1109/EMBC.2018.8512398.

[22] E. Ambar Pambudi, P. Nurtantio Andono, and R. Anggi Pramunendar, “IMAGE SEGMENTATION ANALYSIS BASED ON K-MEANS PSO BY USING THREE DISTANCE MEASURES,” *ICTACT Journal on Image and Video Processing*, vol. 9, no. 1, Aug. 2018, doi: 10.21917/ijivp.2018.0256.

[23] A. Sheta, M. S. Braik, and S. Aljahdali, “Genetic Algorithms: A tool for image segmentation,” May 2012. doi: 10.1109/ICMCS.2012.6320144.

[24] U. Maulik and S. Bandyopadhyay, “Genetic algorithm-based clustering technique,” *Pattern Recognition*, vol. 33, no. 9, Sep. 2000, doi: 10.1016/S0031-3203(99)00137-5.

[25] Y. K. Dubey and M. M. Mushrif, “FCM Clustering Algorithms for Segmentation of Brain MR Images,” *Advances in Fuzzy Systems*, vol. 2016, 2016, doi: 10.1155/2016/3406406.

[26] P. Kalavathi, “Brain tissue segmentation in MR brain images using multiple Otsu’s thresholding technique,” in *2013 8th International Conference on Computer Science & Education*, 2013, pp. 639–642. doi: 10.1109/ICCSE.2013.6553987.

[27] Y. Zhang, M. Brady, and S. Smith, “Segmentation of brain MR images through a hidden Markov random field model and the expectation-maximization algorithm,” *IEEE Transactions on Medical Imaging*, vol. 20, no. 1, 2001, doi: 10.1109/42.906424.

[28] A. I. Scher *et al.*, “Hippocampal shape analysis in Alzheimer’s disease: A population-based study,” *NeuroImage*, vol. 36, no. 1, pp. 8–18, 2007, doi: https://doi.org/10.1016/j.neuroimage.2006.12.036.

[29] B. Gutman, Y. Wang, J. Morra, A. W. Toga, and P. M. Thompson, *Disease classification with hippocampal shape invariants*, vol. 19. 2009. doi: 10.1002/hipo.20627.

[30] X. Liu, L. Song, S. Liu, and Y. Zhang, “A Review of Deep-Learning-Based Medical Image Segmentation Methods,” *Sustainability*, vol. 13, no. 3, Jan. 2021, doi: 10.3390/su13031224.

[31] M. Shao *et al.*, “Brain ventricle parcellation using a deep neural network: Application to patients with ventriculomegaly,” *NeuroImage: Clinical*, vol. 23, 2019, doi: 10.1016/j.nicl.2019.101871.

[32] K. H. Zou *et al.*, “Statistical validation of image segmentation quality based on a spatial overlap index,” *Academic radiology*, vol. 11, no. 2, pp. 178–189, Feb. 2004, doi: 10.1016/s1076-6332(03)00671-8.

[33] F. Falahati *et al.*, “The Effect of Age Correction on Multivariate Classification in Alzheimer’s Disease, with a Focus on the Characteristics of Incorrectly and Correctly Classified Subjects,” *Brain topography*, vol. 29, no. 2, pp. 296–307, Mar. 2016, doi: 10.1007/s10548-015-0455-1.

**تصنيف مرض الزهايمر بالإعتماد على عدد من الصفات المستخرجة من صور الرنين المغناطيسى (T1) للدماغ**

***الخلاصة:*** *تشخيص مرض الزهايمر فى مراحله الأولى يشغل حيزا مهماً فى الحد من أعراض المرض التقليل من التدهور الإدراكى الذى يسببه المرض, ولذلك أصبحت أنضمة الحاسوب المساعدة (Computer-Aided-Systems) حاسمةً فى التشخيص الدقيق والمبكر للمرض. إنّ النظام المقترح فى هذه الدراسة يعتمد على فرز الصفات المستخرجة من صور الرنين المغناطيسى للدماغ, وهذه الصفات يجب أن تكون دقيقة ومتقّنة فى تصوير التغييرات التشريحة الحاصلة فى هيكلية الدماغ المصاب بمرض الزهايمر, كتآكل منطقة الحصين (Hippocampus) وتوسع التجاويف وسمك القشرة المخية وحجم انسجة الدماغ وغيرها من الصفات. وفى هذه الدراسة, تم إستخدام صور الرنين المغناطيسى ذو وزن (T1) والمتضمنة على هيكلية الدماغ, ويعتبر هذا النوع من الصور ذى صفاتٍ مرئيةٍ عالية الدقة مما يجعل خطوات معالجتها وفرزها أقل تعقيداً. النظام المقترح فى هذهِ الدراسة يتكون من عدة خطوات تبدأ بالمعالجة الأولية للصور, تتبعها طريقةً لتقسيم أنسجة الدماغ فى هذه الصور ومن ثُمّ تقطيع الدماغ الى المناطق المهمة والمتضمنة على صفات المرض. الصفات المستخرجة من انسجة الدماغ ومناطقهِ يتم إستغلالها فى تشخيص المرض من خلال فرز وتصنيف هذه الصفات بإستخدام أحد أكثر طرق الذكاء الإصطناعى التقليدية شيوعاً (Support Vector Machine) كآخر مرحلة فى هذا النظام المقترح. وتؤكد النتائج على أداءٍ عالٍ ودقةٍ واعدة فى تشخيص المرض بإستخدام المنهجية المقترحة فى هذه الدراسة, إذ كانت النتيجة (89.1%) فى الفرز بين الأشخاص المصابين بالزهايمر عن ذوى الإدراك السليم (Alzheimer’s Demented vs. Cognitively Normal).*