Journal of Neuroscience Methods 337 (2020) 108669

Contents lists available at ScienceDirect

Journal of Neuroscience Methods

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

Hippocampal atrophy based Alzheimer's disease diagnosis via machine learning methods

Gokce Uysal[a], Mahmut Ozturk[b,*]

[a] Department of Biomedical Engineering, Institute of Graduate Studies, Istanbul University-Cerrahpasa, Avcilar, 34320, Istanbul, Turkey
[b] Department of Electrical and Electronics Engineering, Istanbul University-Cerrahpasa, Avcilar, 34320, Istanbul, Turkey

ARTICLE INFO

Keywords:
Alzheimer's disease
Hippocampal atrophy
Magnetic resonance imaging
Semi-automatic segmentation
Machine learning

ABSTRACT

Alzheimer's disease is the most common form of dementia and is a serious health problem. The disease is expected to increase further in the upcoming years with the increase of the elderly population. Developing new treatments and diagnostic methods is getting more important. In this study, we focused on the early diagnosis of dementia in Alzheimer's disease via analysis of neuroimages. We analyzed the data diagnosed by the Alzheimer's Disease Neuroimaging Initiative (ADNI) protocol. The analyzed data were T1-weighted magnetic resonance images of 159 patients with Alzheimer's disease, 217 patients with mild cognitive impairment and 109 cognitively healthy older people. In this study, we propose that the volumetric reduction in the hippocampus is the most important indicator of Alzheimer's disease. There is not much research about the relationship between the volumetric reduction in the hippocampus and Alzheimer's disease. This volume information was calculated through semi-automatic segmentation software ITK-SNAP and a data set was created based on age, gender, diagnosis, and right and left hippocampal volume values. The diagnosis via hippocampal volume information was made by using machine learning techniques. By using this approach, we conclude that brain MRIs can be used to distinguish the patients with Alzheimer's Disease (AD), Mild Cognitive Impairment (MCI) and Cognitive Normal (CN) from each other; while most of the studies were only able to distinguish AD from CN. Our results have revealed that our approach improves the performance of the computer-aided diagnosis of Alzheimer's disease.

1. Introduction

In recent years, the percentage of the older people in the human population has been increasing rapidly. The data of the Turkish Statistical Agency shows that 8,7 % of the population was older than 65 age in 2018 while it was 4,7 % in 1980. The Statistical Agency predicts that the percentage of the people over 65 age will be the quarter of the total population in 2080. Higher percentages of older people can be seen in almost all European countries also. As a global problem, the aging of the human population causes an increase in health problems that depend on aging. The growth of the number of patients with Alzheimer's Disease (AD) as the most common form of dementia is a direct result of the aging of the population. According to the annual reports of Alzheimer's Disease International (ADI), approximately 50 million people across the globe suffered from Alzheimer's disease in 2018. Moreover, they predict that the numbers will be twice in every 20 years. As a result of the growth of this problem, developing new and effective methods for early diagnosis and treatment of AD is getting more important and popular.

With the rapid developments of the neuroimaging techniques, it is getting possible to diagnose AD by using the neuroimages like some other disorders at the brain activity. Diagnosis of the disease at the early stages can give a chance to control and slow down the results of it.

In recent years, a lot of new studies have been done to develop effective methods for the diagnosis of AD from neuroimages. By analyzing the Magnetic Resonance (MR) images of the patients' brains, researchers specified that the volumetric reduction in some parts of the brain can be used as a biomarker for AD. Most of the researchers have worked on the volumetric reduction at the White Matter (WM) of the brain. Balthazar et al. observed that the WM volume of the patients with AD is very lower than the Cognitive Normal (CN) older people's (Balthazar et al., 2009). Frings et al. showed that decreasing of the temporal WM volume in patients with AD is faster than CN older people's (Frings et al., 2014). Migliaccio et al. showed that volumetric damage in the WM in Early Age Onset AD and its atypical variants is consistent with Grey Matter (GM) reduction, cognitive decline and

* Corresponding author.
E-mail address: mahmutoz@istanbul.edu.tr (M. Ozturk).

https://doi.org/10.1016/j.jneumeth.2020.108669
Received 15 November 2019; Received in revised form 27 February 2020; Accepted 28 February 2020
Available online 29 February 2020

psychological symptoms (Migliaccio et al., 2015). As a result of (Balthazar et al., 2009; Frings et al., 2014; Migliaccio et al., 2015) and some other works, researchers concluded that volumetric reduction of the WM can be observed via MR images and it can be used as a good biomarker to diagnose AD instead of the classical neuropsychological tests.

Some researchers focused on the atrophies at the other parts of the brain. Neuronal losses and the atrophy at big cortical neurons have been seen in the brains of AD patients. Approximately 50 % of cells and synapses at the cortex and hippocampus have been lost (Oguten, 2012; Eagleman, 2018). It is concluded at the neuropathological investigations of the brain tissue that the cerebral cortex is thinner than the normal situation (Arslan, 2012). By using the volumetric atrophy measurements it is specified that the hippocampus is the most affected anatomical region.

Risacher et al. claimed that the hippocampal volume reduction is the most effective biomarker at the process of transformation from Mild Cognitive Impairment (MCI) to Alzheimer's Disease (AD) (Risacher et al., 2009). They investigated that the biomarkers of the neuroimages of MCI-stable and MCI-transformed groups to measure the effects of anatomical regions. The percentage of influence of the anatomical regions can be seen in Fig. 1.

As it can be seen with a rapid literature survey, most of the studies which try to diagnose AD by using image processing techniques, focus on the volumetric reduction of WM. On contrary to this trend, in this study, we investigated the volumetric changes in the hippocampus. We analyzed the right and the left hippocampus separately and together.

In the following parts of this study, firstly we are describing how we calculate the volume of the hippocampus from MR images. Then we are going to compare the hippocampus volumes of patients with AD, MCI, and CN. Distinguishing the patients with MCI from AD or CN is the hardest part of the diagnosing work because their bioindicators can be very near to the AD group or CN group. So, distinguishing these three groups from each other is a challenging issue at the AD researches. Some researchers have used EEG signals to distinguish AD, MCI and CN from each other (Dauwell et al., 2010; Mizuno et al., 2010; Adeli et al., 2005; Jiang et al., 2019; Kang et al., 2018; Urigüen and Garcia-Zapirain, 2015). There are some well-performed algorithms proposed to diagnose AD using EEG signals. Despite the possible measurement problems (weak amplitudes, noise addition… etc), EEG is an effective and important data for diagnosing neurological disorders. Nowadays, MRI is getting more common, highly resolute, cheaper and more accessible. So, we preferred to focus on the MRI results like most of the recent researches in the area of neurological disorders. We claim that using the hippocampus volume to distinguish stages of AD via machine learning is a novel idea and its diagnosis performance at our work looks highly satisfactory. We think that using MRI and EEG data of the same patient simultaneously will show more successful diagnosing results and it can be thought of as a future work. After that, you can see the results of our simulations for the diagnosis of AD using machine learning techniques. This work will be completed with conclusions.

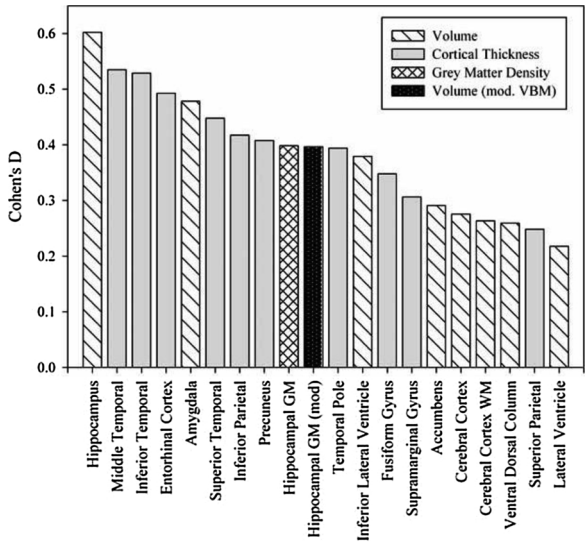

Fig. 1. Effect sizes of the comparison between MCI-Stable and MCI-Converter groups evaluated for selected imaging biomarkers (Risacher et al., 2009).

2. Methodology

The main purpose of this work is to diagnose the AD clinically as early as possible, to make the diagnosis process faster, and to get minimize the human factor on the diagnosis process. At this work, we propose that brain MRIs can be used to distinguish the patients with AD, MCI, and CN from each other. We propose to focus on the hippocampal atrophy on the brain. By using the volume information of the hippocampus with the ages and the genders of the patients, we have shown that the diagnosis of AD from MRIs is highly possible.

A lot of studies have shown that the tau increase and the amyloid plaque deposition, which are accepted as the biomarkers of the neuronal damage, are strictly related to hippocampal atrophy. Early level dementia studies show that the hippocampal atrophy prevalence increases with age and is widely transformed into AD from MCI condition (Jack et al., 2009; Visser et al., 1999; Sontheimer, 2015; Leon et al., 1997; Leon et al., 2004; Zandifar et al., 2017; Parker et al., 2018).

2.1. Calculation of hippocampus volume

In this stage, we used ITK-SNAP open program to calculate the hippocampus volume. This program was produced with the protocol between Pennsylvania University – Image Calculation and Science Laboratory and the University of Utah – Scientific Calculation and Imaging Institute. ITK-SNAP provides semi-automatic segmentation to analyze medical images and obtain three-dimensional models of them. Because the ITK-SNAP software is semi-automatic, it is not possible to calculate hippocampus volume directly. For each MRI, we determined and labeled the borders of the left and right hippocampus separately. Moreover, this assignment done on the coronal, sagittal and axial planes one by one. The determination and labeling process for each MRI separately has taken a very long time and much work power. It is possible to make segmentation full-automatically for some parts of the brain. Because of the place of the hippocampus, it is very hard to distinguish it full-automatically. There are some new researches to distinguish the hippocampus full-automatically, but it is still a challenging problem and needs much more effort. We think that developing a full-automatically segmentation algorithm for the left and right hippocampus can be thought of as future work in this area. In this work, three-dimensional models of image data obtained and their volumes calculated using some tools of ITK-SNAP.

Volumetric images can be modeled as a series of three-dimensional densities. The voxels, that is formed with the depth knowledge of the pixels of a two-dimensional image, is expressed with three coordinates. At ITK-SNAP, x and y coordinates show the position of the pixel, and the z coordinate shows the number of the slice which the pixel belongs to. By matching the patient's anatomical coordinates with the coordinates of the segmentation tool, the anatomical localization is defined.

The MR image data which is selected from ADNI based on imaging protocols, age, gender, and diagnostic criteria is transferred to the ITK-SNAP software. Segmentation labeling has been performed on slices of MR images taken primarily in the coronal, sagittal and axial plane during image pre-processing step. With the addition of closed curves of each area in three-planes utilizing ITK-SNAP, a three-dimensional hippocampus model is created and hippocampal volume information calculated by the software is read. Three-dimensional model of the left

hippocampus can be seen in Fig. 2. For this sample, the volume of the left hippocampus was calculated as 251,330 mm^3.

2.2. Using machine learning for diagnosis of AD

Machine learning algorithms are divided into three categories according to learning methods: supervised, unsupervised and semi-supervised algorithms. The main purpose of supervised learning is to derive a learning model from the tagged educational data that allows us to make predictions about unknown or future data. At this work, the supervised data corresponds to a group of samples already labeled for estimation. Classification or regression models are formed according to the data class used. If the output data is a continuous value numerical data, regression models are used; and if they are included in the categorical data group, classification models are used (Raschka, 2016).

Within the scope of this study, while analyzing the volume, age and gender data with the machine learning algorithms, the data analyzed were divided into 66,6% as education and 33,3% as test data like generally accepted in the literature and classification models were formed.

In this work, we analyzed the data by using the Logistic Regression (LR), K-Nearest Neighbors (KNN), Support Vector Machines (SVM), Decision Tree (DT), Random Forest (RF), and Gaussian Naive Bayes (GNB).

3. Results

Image data analyzed in this study consist of T1-weighted MR images in NIFTI format of 485 patients diagnosed as AD (n = 159, 31,8 %), MCI (n = 217, 44,7 %), CN (n = 109, 22,5 %) of these data. In all groups, 189 (38,97 %) were female and 296 (61,03 %) were male; in the AD subgroup 94 (19,38 %) were female and 65 (13,40 %) were male; in the MCI subgroup 62 (12,78 %) were female and 155 (31,96 %) were male; in the CN subgroup 33 (6,80 %) were female and 76 (15,67 %) were male. Although all of the available data were included in the study, there was not enough data in the age group of 90 and over in all categories and in the 55–69 age group in the CN category.

Right and left hippocampus volumes of all patients included in this work are calculated using ITK-SNAP semi-automatic segmentation software and significant atrophy differences were observed. For example, in Fig. 3 right and left hippocampus slices of a 73 years old Cognitively Normal woman can be seen on the coronal, sagittal and axial planes[1] . The left hippocampus volume of this patient was calculated as 285,874 mm^3, and the right one was calculated as 307,768 mm^3.

By using same calculation techniques, left and right hippocampus volumes of an 87 years old woman[2] with AD were calculated 84,612 mm^3 and 59,283 mm^3 respectively. Hippocampal atrophy can be observed clearly at the MR images of this patient showed in Fig. 4 on the coronal, sagittal and axial planes.

In the two samples examined, it is seen that the right and left hippocampus volumes of the patient with AD have decreased approximately 2000 mm^3 compared to the healthy elderly.

For all data included in this work, averages of the right and left hippocampus volumes were obtained separately considering the age period and gender can be seen in Table 1.

When the data of the cognitively healthy elderly in Table 1 are examined, the approximate value of the hippocampal atrophy is 150 mm^3 for male patients in the age range of 55–95 years, while this value is approximately 75 mm^3 for female patients in the age range of 66–95. Although direct mathematical inference can't be made with limited data, it is seen that the amount of hippocampal atrophy in the aging process of cognitively healthy elder people is much lower than the other groups.

When the average volume data of the patients with Mild Cognitive Impairment (MCI) are examined, while the approximate value of hippocampal atrophy is 700 mm^3 for male patients in the age range of 55–95, it is approximately 350 mm^3 for female patients in the same age period. It can be seen that the hippocampus volume of male patients decreases very steeply depending on the aging, but it can be said that the hippocampus volume of female patients does not decrease with the same speed. Moreover, it is seen that the hippocampal atrophy at male patients with MCI is approximately twice of female patients'.

Although the main effect of age on hippocampal atrophy is well known, and parallel with our findings, there are some studies suggesting that males' atrophy rate is higher than females' (Raschka, 2016), it is thought that the reason of hippocampal atrophy in male patients is twice as high as in female patients might be related to distribution of sample data.

When the volume data of patients with AD given in Table 1 are analyzed, it is seen that the hippocampal atrophy for male patients in the age range of 55–95 is approximately 400 mm^3, while it is approximately 550 mm^3 for female patients in the same age period. Despite the more cognitive impairment, the lesser estimate of the hippocampal atrophy compared to the MCI group is due to the assessment within the diagnostic group itself. Because, even when the hippocampus volume values of the 55–65-year-old AD and MCI diagnostic groups are compared, a volume difference of approximately 500 mm^3 is seen. Since Alzheimer's disease is insidious and early onset, hippocampal volume values are very low compared to CN and MCI even in patients diagnosed at a young age. To clarify the findings, the changes in the hippocampus volume at different levels and velocity depending on age in each diagnostic group are given in Fig. 5 for the left hippocampus and Fig. 6 for the right hippocampus. Although the main effect of age on regional gray matter volume, especially on the hippocampus, is known, the volume differences at a baseline level of the diagnoses are significant, as long as they are evaluated.

3.1. Comparative diagnosis performances of classification algorithms

The right hippocampus volume and left hippocampus volume calculated from the image data of each diagnostic group within the scope of the study were made by taking into consideration the age and gender criteria of the individuals and diagnosis estimation was made through classification algorithms. In each classification model, double and triple diagnostic groups were analyzed within single and multiple parameters. The predictive performances of the classification models were measured by looking at the accuracy, sensitivity and specificity values read over the confusion matrix.

In Table 2, only the data of AD and CN diagnostic groups were evaluated and the effect sizes of different combinations of right hippocampus volume, left hippocampus volume, age, and gender on classification algorithms were measured. For each model, 268 data were divided into 89 (33.3 %) test data and 179 (66.7 %) training data. In the data set where Alzheimer's and healthy elderly are evaluated jointly, classification is expected to be more pronounced than other diagnostic groups. The current findings were in line with the literature studies and

[1] MRI protocols: Acquisition Plane = SAGITTAL; Acquisition Type = 3D; Coil = HEAD; Field Strength = 3.0 tesla; Flip Angle = 8.0 degree; Manufacturer = GE MEDICAL SYSTEMS; Matrix X = 256.0 pixels; Matrix Y = 256.0 pixels; Matrix Z = 166.0 ; Mfg Model = GENESIS_SIGNA; Pixel Spacing X = 1.0156199932098389 mm; Pixel Spacing Y = 1.0156199932098389 mm; Pulse Sequence = RM; Slice Thickness = 1.2000000476837158 mm; TE = 3.052000045776367 ms; TI = 900.0 ms; TR = 7.5 ms; Weighting = T1.

[2] MRI protocols: Acquisition Plane = SAGITTAL; Acquisition Type = 3D; Coil = PA; Field Strength = 1.49399995803833 tesla; Flip Angle = 8.0 degree; Manufacturer = SIEMENS; Matrix X = 192.0 pixels; Matrix Y = 192.0 pixels; Matrix Z = 160.0 ; Mfg Model = Symphony; Pixel Spacing X = 1.25 mm; Pixel Spacing Y = 1.25 mm; Pulse Sequence = IR/GR; Slice Thickness = 1.2000000476837158 mm; TE = 3.7899999618530273 ms; TI = 1000.0 ms; TR = 3000.0 ms; Weighting = T1.

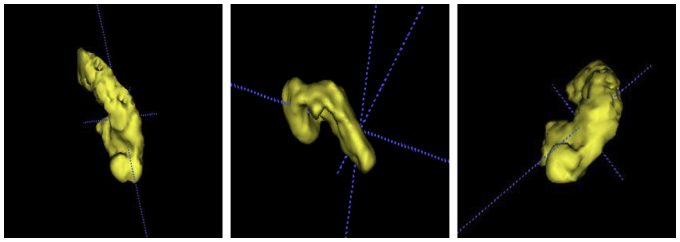

Fig. 2. 3D model of the left hippocampus.

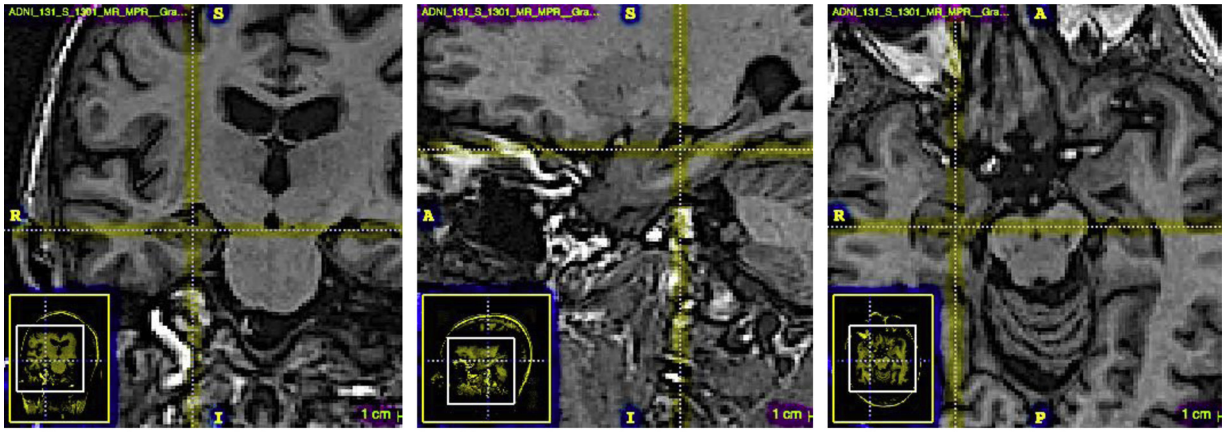

Fig. 3. Hippocampus slices on three planes of the cognitively normal woman.

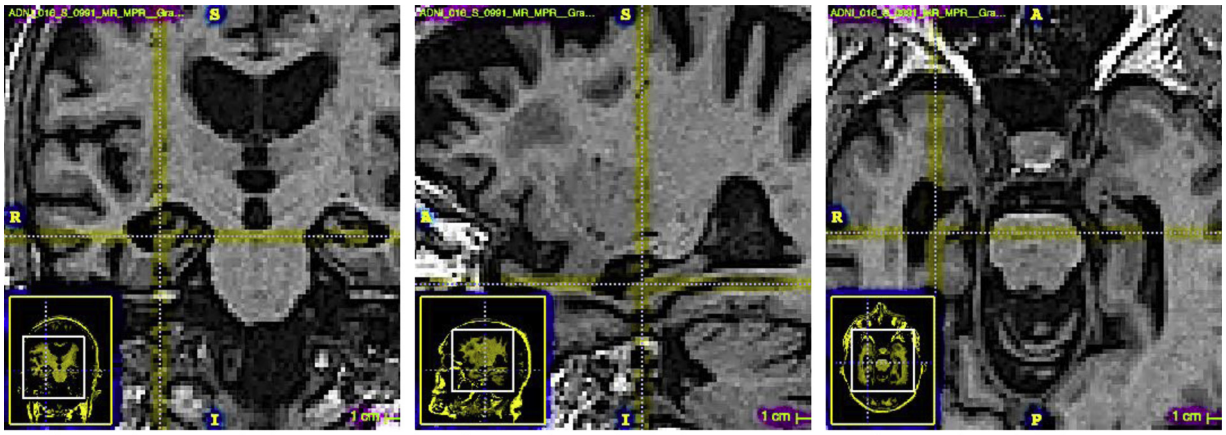

Fig. 4. Hippocampus slices on three planes of the patient with Alzheimer's disease.

Table 1
Average volume values of the right and left hippocampus for all examined samples.

	The average volume of the right hippocampus (mm^3)		The average volume of the left hippocampus (mm^3)	
n:109 CN	Female	Male	Female	Male
55–65	–	2582,91	–	2648,66
66–79	2518,00	2570,75	2447,81	2521,80
80–95	2423,80	2481.96	2392,47	2414.13
n:217 MCI	Female	Male	Female	Male
55–65	2211,92	2414,78	2197,13	2392,48
66–79	1927,75	2005,49	1887,26	1946,36
80–95	1863.26	1726,41	1833,51	1731.57
n = 159 AD	Female	Male	Female	Male
55–65	1871,75	1774,66	1763,64	1741,10
66–79	1474.19	1506,26	1390.07	1378,68
80–95	1291,49	1361,68	1212,56	1339,09

the most successful estimation results were realized in this data set in all classification models.

Increasing the effect parameters in the LR classification model caused almost no change in the predicted value and increased the specificity of the algorithm. This is considered as an indication that LR can produce the same regression model even with limited data. The KNN classification model was the algorithm for making the most accurate estimates for the current data set, with the inclusion of gender information increasing the percentage of accuracy, sensitivity, and specificity, resulting in an overall improvement in achievement performance. In the SVM classification model, the left hippocampus value provided a more efficient algorithm model than the right hippocampus. This can be explained by the fact that the right hippocampus values in the data set were calculated closer together. The increase in the parameters examined in the GNB classification model does not change the success performance of the algorithm model. And this is thought to be

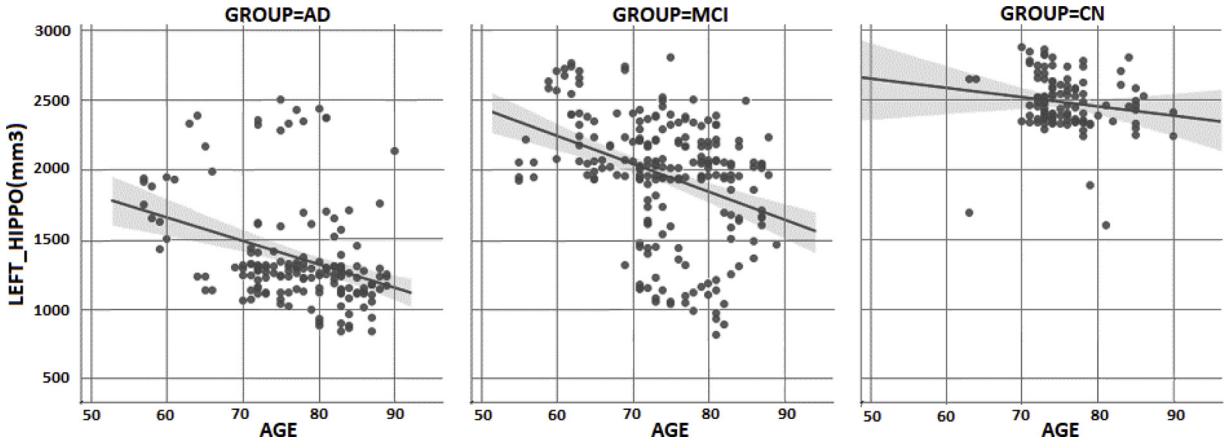

Fig. 5. Age-related change of left hippocampus volume for each diagnostic group.

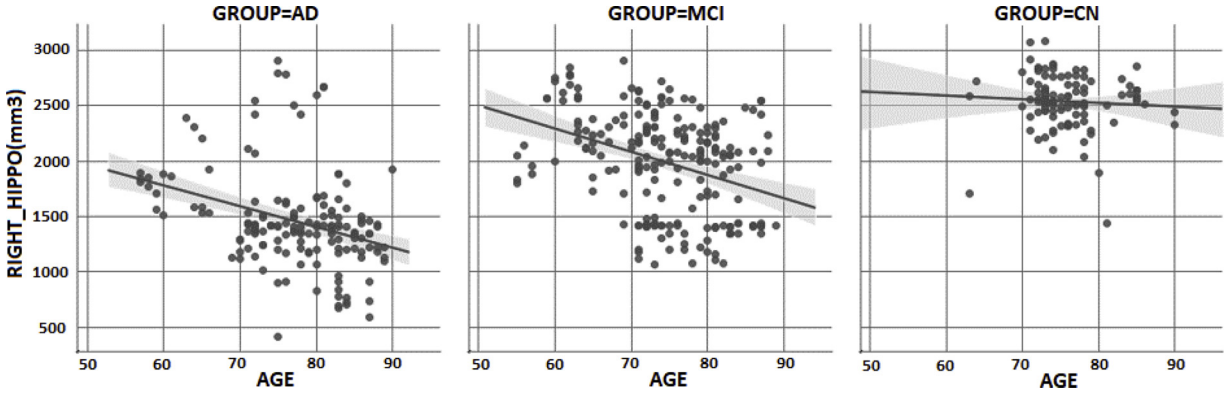

Fig. 6. Age-related change of right hippocampus volume for each diagnostic group.

Table 2
Diagnosis performance of classification algorithms for the AD-CN diagnostic group.

Diagnosis Performances	Effect sizes of investigated parameters				
	RHV	LHV	RHV + LHV	RHV + LHV + A	RHV + LHV + A + G
LR					
%Acc.	0.92	0.93	0.94	0.93	0.93
%Sen.	0.92	0.90	0.92	0.90	0.90
%Spe.	0.92	0.97	0.97	0.97	0.97
KNN					
%Acc.	0.90	0.90	0.92	0.94	0.98
%Sen.	0.96	0.90	0.96	0.96	0.98
%Spe.	0.82	0.90	0.86	0.92	0.97
SVM					
%Acc.	0.83	0.94	0.92	0.92	0.93
%Sen.	0.92	0.92	0.92	0.90	0.90
%Spe.	0.71	0.97	0.92	0.95	0.97
GNB					
%Acc.	0.92	0.93	0.94	0.94	0.94
%Sen.	0.92	0.90	0.92	0.92	0.92
%Spe.	0.92	0.97	0.97	0.97	0.97
DT					
%Acc.	0.90	0.90	0.94	0.87	0.91
%Sen.	0.96	0.90	0.96	0.96	0.98
%Spe.	0.82	0.90	0.92	0.73	0.82
RF					
%Acc.	0.90	0.91	0.93	0.96	0.93
%Sen.	0.92	0.92	0.94	0.96	0.94
%Spe.	0.87	0.90	0.92	0.95	0.92

(Abbreviations are: Acc, Accuracy Value; Sen, Sensitivity; Spe, Specificity; RHV, Right Hippocampus Volume (mm^3); LHV, Left Hippocampus Volume (mm^3); A, Age; G, Gender.).

Table 3
Diagnosis performance of classification algorithms for the AD-MCI diagnostic group.

Diagnosis Performances	Effect sizes of investigated parameters				
	RHV	LHV	RHV + LHV	RHV + LHV + A	RHV + LHV + A + G
LR					
%Acc.	0.73	0.83	0.81	0.82	0.83
%Sen.	0.76	0.91	0.91	0.91	0.91
%Spe.	0.71	0.78	0.74	0.77	0.78
KNN					
%Acc.	0.65	0.75	0.74	0.83	0.87
%Sen.	0.67	0.74	0.78	0.85	0.87
%Spe.	0.63	0.76	0.72	0.82	0.87
SVM					
%Acc.	0.69	0.82	0.82	0.70	0.81
%Sen.	0.33	0.80	0.76	0.63	0.89
%Spe.	0.91	0.83	0.85	0.75	0.76
GNB					
%Acc.	0.74	0.82	0.79	0.78	0.77
%Sen.	0.80	0.91	0.91	0.91	0.89
%Spe.	0.71	0.77	0.71	0.70	0.70
DT					
%Acc.	0.65	0.75	0.79	0.81	0.81
%Sen.	0.67	0.74	0.83	0.87	0.87
%Spe.	0.63	0.76	0.77	0.77	0.78
RF					
%Acc.	0.66	0.75	0.80	0.81	0.82
%Sen.	0.65	0.74	0.87	0.87	0.91
%Spe.	0.64	0.76	0.77	0.78	0.75

due to the fact that this model evaluates each assumption gracefully. In the DT classification model, the success performance of the algorithm decreased with the inclusion of age information. Since the knowledge of gender does not have a significant effect on the formation of diagnostic groups, linearity in the decision tree flowchart is disturbed. The fact that the same situation increases the success of RF is due to the fact that random trees are less susceptible to outliers or low values compared to decision trees.

Table 3 evaluates the performance of prediction algorithms established between AD-MCI diagnostic groups. For each model, 373 data were divided into 124 (33.2 %) test data and 249 (66.7 %) training data. While the left hippocampus volume value in LR increased the success of the algorithm higher than the right hippocampus volume value, age and gender parameters caused lower performance increase. KNN is the most successful prediction model for this data group, and the effect size of the parameters is almost parallel to LR, where the effect size of age and gender is greater.

The left hippocampus volume value and gender algorithm in the SVM classification model significantly increases success. Unlike other classification models, the age parameter decreases performance and is caused by the fact that the effect of the left hippocampus is much greater than age. While left hippocampus volume value increased predictive success in GNB, gender had a low negative effect, unlike other algorithms. While increasing the left hippocampus in DT, gender had no effect on the algorithm but only a low rate of specificity. Although the left hippocampus value in RF is a more effective parameter than the right hippocampus value, the evaluation of the hippocampus information together has created an effect that strengthens the algorithm model.

In Table 4, the classification performances for MCI-CN diagnostic groups were measured. For each model, 324 data were divided into 107 (33 %) test data and 217 (67 %) training data. The most successful forecasting models are GNB and RF. In both classification models, data details are analyzed extensively, showing the strongest effect even with the lowest effect parameter. Especially in the MCI-CN group, where data points are very close to each other, it was seen that GNB and RF were successful at the points where the data were intertwined.

In Table 5, all diagnostic groups were included in the common data set and the diagnosis performances of classification algorithms were measured. For each model, 482 data were divided into 160 (33.2 %) test data and 322 (66.8 %) training data. Estimation performances were evaluated by classification among all diagnostic groups. Since the left hippocampus volume value yielded more successful estimation results than the right hippocampus in each model, a new domain was examined by adding age and gender parameters to which the right hippocampus value was not included. In this situation, success rate increased. In this case, the most successful prediction algorithm was KNN with 80 % accuracy when all parameters were evaluated, and GNB with 82 % accuracy with the exclusion of right hippocampus volume value. Age and gender have positive effects for each algorithm. When all the findings are evaluated, it is seen that a significant diagnosis can be made with hippocampus volume information and hippocampal volume differences between ages and genders are effective in the diagnosis decision.

4. Discussions

The main focus of this work is to use MR imaging technique for early diagnosis of Alzheimer's disease, and then to predict the diagnosis with machine learning techniques. Since atrophy caused by the disease is an important biomarker, MRI is used as the main technique in many imaging interventions by providing structural imaging and continuous volumetric analysis. MRI has also been the most widely used technique in the Alzheimer's Disease Neuroimaging Initiative (ADNI). The data obtained with the ADNI protocol consists of T1-weighted MR data.

The second focus of this work is the selection of the main anatomic localization affected by the disease. Today, a holistic approach to the brain is getting effective in neurodegenerative diseases. Especially when memory is evaluated, it is seen that it creates an effective area in the whole brain. These levels of effect vary depending on the type of dementia, age, life, and many other parameters. Moreover, the increase in age-related plasticity allows new neuronal pathways to be made, thus preventing the loss of function due to an age-related increase in the individual. This situation means that biomarkers of AD have been observed but there is no cognitive loss. Although the effect level of this situation cannot be directly evaluated and explained directly on an

Table 4
Diagnosis performance of classification algorithms for the MCI-CN diagnostic group.

Diagnosis Performances	Effect sizes of investigated parameters				
	RHV	LHV	RHV + LHV	RHV + LHV + A	RHV + LHV + A+G
LR					
%Acc.	0.84	0.84	0.84	0.89	0.90
%Sen.	0.65	0.60	0.63	0.74	0.74
%Spe.	0.93	0.96	0.94	0.96	0.97
KNN					
%Acc.	0.77	0.80	0.83	0.87	0.89
%Sen.	0.60	0.60	0.69	0.69	0.72
%Spe.	0.86	0.90	0.90	0.96	0.97
SVM					
%Acc.	0.75	0.75	0.75	0.81	0.86
%Sen.	0.31	0.28	0.31	0.43	0.60
%Spe.	0.96	0.97	0.95	100	0.99
GNB					
%Acc.	0.83	0.90	0.90	0.95	0.95
%Sen.	0.74	0.92	0.94	0.97	0.97
%Spe.	0.88	0.89	0.89	0.94	0.94
DT					
%Acc.	0.78	0.80	0.79	0.88	0.90
%Sen.	0.60	0.60	0.57	0.69	0.74
%Spe.	0.86	0.90	0.89	0.97	0.97
RF					
%Acc.	0.78	0.79	0.85	0.92	0.94
%Sen.	0.57	0.57	0.71	0.77	0.82
%Spe.	0.88	0.90	0.93	0.99	100

individual basis in dementia studies, it is thought that the stress level of the individual's living conditions and even the nutritional style will affect its creation. Whether the increase in plasticity is not at the same level in every one or the reasons for its positive effect is investigated, the main focus of cognitive loss in the literature is the limbic system. The hippocampus, which is an important part of the limbic system, is seen as one of the most affected structures in many types of dementia and the most changed in the disease process. In the scope of our work, the effect of hippocampal atrophy for dementia in Alzheimer's disease specificity was analyzed and it was emphasized that the possibility of early diagnosis could be possible. On the other hand, the findings of this work confirmed this approach and clearly demonstrated the increase in age-related atrophy. By making the three-dimensional modeling of the right and left hippocampus on the MR images obtained by ADNI protocol, volume information was obtained. At this stage, the use of a semi-automatic segmentation tool, that is, the hippocampus labeling and thresholding sections are made manually or the use of a tool where the hippocampal frame is manually determined, allows the control of the

Table 5
Diagnosis performance of classification algorithms for the AD, MCI, and CN diagnostic group.

Diagnosis Performances	Effect sizes of investigated parameters					
	RHV	LHV	RHV + LHV	RHV + LHV + A	RHV + LHV + A+G	LHV + A+G
LR						
%Acc.	0.67	0.71	0.71	0.72	0.74	0.78
%Sen.	0.85	0.87	0.87	0.87	0.86	0.87
%Spe.	0.67	0.51	0.62	0.67	0.71	0.74
KNN[a]						
%Acc.	0.56	0.70	0.69	0.77	0.80	0.78
%Sen.	0.63	0.72	0.80	0.80	0.83	0.80
%Spe.	0.51	0.77	0.59	0.77	0.77	0.77
SVM						
%Acc.	0.58	0.69	0.69	0.70	0.76	0.79
%Sen.	0.35	0.76	0.65	0.59	0.72	0.80
%Spe.	0.54	0.44	0.62	0.67	0.74	0.79
GNB						
%Acc.	0.65	0.77	0.74	0.74	0.74	0.82
%Sen.	0.77	0.87	0.89	0.89	0.87	0.87
%Spe.	0.74	0.90	0.85	0.87	0.87	0.95
DT[b]						
%Acc.	0.56	0.69	0.68	0.76	0.74	0.80
%Sen.	0.63	0.72	0.76	0.78	0.74	0.87
%Spe.	0.51	0.77	0.56	0.67	0.79	0.77
RF						
%Acc.	0.57	0.71	0.71	0.78	0.78	0.78
%Sen.	0.61	0.74	0.80	0.87	0.80	0.85
%Spe.	0.56	0.77	0.69	0.74	0.77	0.69

[a] Minkowski was used as metric.
[b] Entropy was used.

desired area to obtain the volume, but it may also create a margin of error. Therefore, the analyses were repeated many times and very close results were obtained. The parallelism of the hippocampal atrophy comparisons with the older studies has also provided continuity of this work. As a result of the findings, it has been focused on the prediction of diagnosis by using machine learning techniques, considering that the interpretation of atrophy should contribute to the clinician in practical life especially for the rapidity of diagnosis.

For all diagnostic groups, we have obtained the best classification results when we use RHV and LHV together. Adding Age (A) and Gender (G) information has made minor improvements. By checking the diagnosis results, we see that the LHV is more effective and significant on the progression of AD.

As can be seen from Tables 2–4, diagnosis performances of classifications algorithms in this work are satisfactory, but still there are some false diagnosis results. The false diagnosis results can be explained related to not fully understood reasons of AD. There are still some important discussions about the reasons and biomarkers of AD. As can be seen in Figs. 5 and 6, volume reduction at the brain is not an exact indicator for some cases. It is possible to see some old and healthy people with very low hippocampus and brain volumes. Reversely, in some AD cases, it is specified that the age of the patient is relatively low and the hippocampus and brain volumes are high. The percentage of these unidentified anomalies is low. Because of this kind of cases, the reason of AD is still a lively discussion between neurologists.

At the first and second columns of the tables, we used RHV or LHV information uniquely for classification. So, low classification performances are expected and normal. The best classification performances of our work can be seen at the fourth and fifth columns of the tables.

In our work, we have investigated the hippocampus volume. A lot of researches in this area look at the volumes of white and gray matters. Calculation of hippocampus volume is much more difficult, requires much more effort and the possibility of false calculation is higher. Despite these challenges, we propose that the hippocampus volume is a more significative indicator for the diagnosis of AD, and focusing on it is an important and relatively new idea.

We see that there are also some disadvantages of increasing the number of parameters while diagnosing the stage of the AD. Firstly, increasing the MCI data made it difficult to distinguish all the groups because the hippocampal volume data of the MCI group are among the volume values of the AD and CN groups. Secondly, as can be seen from the results, the left hippocampus volume (LHV) is a more significant and effective indicator to observe the stage of AD. So, adding the right hippocampus volume (RHV) parameter reduces the estimation value by a small amount. Moreover, as we mentioned before, the reason of the some patients' hippocampal atrophy is brain plasticity. Generally, they don't have cognitive impairment. If the number of data of these patients increases, the percentage of the successful diagnosis decreases. Also, another disadvantage of increasing the number of parameters is the processing load of the computer. The machine learning methods use some complex algorithms and cause a quite high processing load. With every new parameter addition, the processing time of the machine learning algorithms has increased considerably.

5. Conclusions

The application of machine learning techniques to radiological images is a field of research that is expected to grow dramatically in the upcoming years. Recent advances in machine learning techniques have the potential to recognize and classify complex patterns from different radiological imaging methods such as x-ray, computed tomography, magnetic resonance imaging, and positron emission tomography imaging. In many studies in the literature, systems based on machine learning are compared with the human decision-making process and performance evaluation is performed. Machine learning practices are seen as key components of future clinical decision-making and monitoring systems. In this way, it is aimed for clinicians to diagnose faster.

In this work, classification models were used for the estimation of diagnosis within the parameters of right and left hippocampal volume, age, and gender on the data obtained from ADNI. According to the performance evaluations made with the confusion matrix, significant success was achieved. In the presence of the obtained results, it was seen that the parameters evaluated, ie combinations of biomarkers, provided a more successful estimation than the individual evaluation of biomarkers. Even the gender criterion which we ignore in the hippocampal volume calculation has been shown to play an active role in increasing success. In this case, it has been shown that not only atrophy values but also gender can affect the diagnosis decision, especially in the detection of Alzheimer's disease. Also, the left hippocampus volume which is directly related to the analytical thinking skills has been found to cause more accurate estimates in many classification algorithms. Whether the left side of the brain is a decisive criterion in the presence of Alzheimer's disease is a new research topic and it can not be possible to find enough research about the relationship between the cognitive loss at the left side of the brain and Alzheimer's disease.

CRediT authorship contribution statement

Gokce Uysal: Conceptualization, Methodology, Software, Investigation, Writing - original draft, Writing - review & editing. **Mahmut Ozturk:** Conceptualization, Methodology, Software, Investigation, Writing - original draft, Writing - review & editing.

Acknowledgements

This paper was partially supported by the Research Fund of Istanbul University – Cerrahpasa with the project no 39227, and FYL-2018-31431. MRI data used in this work were obtained from the Alzheimer's Disease Neuroimaging Initiative (ADNI) database.

References

Adeli, H., Ghosh-Dastider, S., Dadmehr, N., 2005. Alzheimer's disease: models of computation and analysis of EEGs. Clin. EEG Neurosci. 36, 131–140.
Arslan, E., 2012. Importance of PET/CT on Differential Diagnosis of Dementia: Contribution of 3-dimensional Stereotactic Surface Projection (3D-SSP) Software Analysis. Specialization in Medicine Thesis. Istanbul University, pp. 3–4.
Balthazar, M.L.F., Yasuda, C.L., Pereira, F.R., Pedro, T., Damasceno, B.P., Cendes, F., 2009. Differences in grey and white matter atrophy in amnestic mild cognitive impairment and mild Alzheimer's disease. Eur. J. Neurol. 16, 468–474.
Dauwell, J., Vialatte, F., Musha, T., Cichocki, A., 2010. A comparative study of synchrony measures for the early diagnosis of Alzheimer's disease based on EEG. Neuroimage 49, 668–693.
Eagleman, D., 2018. The Brain: The Story of You. Istanbul. Domingo. pp. 11.
Frings, L., Yew, B., Flanagan, E., Lam, B.Y.K., Hull, M., Huppertz, H.J., 2014. Longitudinal grey and white matter changes in frontotemporal dementia and Alzheimer's disease. PLoS One 9.
Jack, C.R., Petersen, R.C., Xu, Y.C., 2009. Prediction of AD with MRI-based hippocampal volume in mild cognitive impairement. Neurology 52, 1397–1403.
Jiang, X., Bian, G.-B., Tian, Z., 2019. Removal of artifacts from EEG signals: a review. Sensors 19 (5), 987–1005.
Kang, G., Jin, S.H., Keun Kim, D., Kang, S.W., 2018. T59. EEG artifacts removal using machine learning algorithms and independent component analysis. Clin. Neurophysiol. 129, 24.
Leon, M.J., George, A.E., Golomb, J., Tarshish, C., Convit, A., Kluger, A., DeSanti, S., McRae, T., Ferris, S.H., Reisberg, B., Ince, C., Rusinek, H., Bobinski, M., Quinn, B., Miller, D.C., Wisniewski, H.M., 1997. Frequency of hippocampal formation atrophy in normal aging and Alzheimer's disease. Neurobiol. Aging 18, 1–11.
Leon, M.J., DeSanti, S., Zinkowski, R., Mehta, P.D., Pratico, D., Segal, S., Clark, C., Kerkman, D., Debernardis, J., Li, J., Lair, L., Reisberg, B., Tsui, W., Rusinek, H., 2004. MRI and CSF studies in the early diagnosis of Alzheimer's disease. J. Intern. Med. 256 (3), 205–223.
Migliaccio, R., Agosta, F., Possin, K.L., Rabinovici, G.D., Miller, B.L., Gorno-Tempini, M.L., 2015. White matter atrophy in Alzheimer's disease variants. Alzheimer's Dementia 301, 553–562.
Mizuno, T., Takahashi, T., Cho, R.Y., Kikuchi, M., Murata, T., Takahashi, K., Wada, Y., 2010. Assessment of EEG dynamical complexity in Alzheimer's disease using multiscale entropy. Clin. Neurophysiol. 121, 1438–1466.
Oguten, E.G., 2012. The Relationship of Psychotic Symptoms With Theory of Mind Skills

and Functionality in Alzheimer Patients. Specialization in Medicine Thesis. Istanbul University, pp. 14–15.
Parker, T.D., Slattery, C.F., Yong, K.X.X., Nicholas, J.M., Paterson, R.W., Foulkes, A.J.M., Malone, I.B., Thomas, D.L., Casha, D.M., Crutch, S.J., Fox, N.C., Schott, J.M., 2018. Differences in hippocampal subfield volume are seen in phenotypic variants of early onset Alzheimer's disease. NeuroImage Clinical; 2213-1582.
Raschka, S., 2016. Giving computers the ability to learn from data, python machine learning. In: Olson, R.S. (Ed.), Chapter 1. Packt Publishing Ltd., Birmingham. UK, pp. 3–10.
Risacher, S.L., Saykin, A.J., West, J.D., Shen, L., Firpi, H.A., McDonald, B.C., 2009. Baseline MRI predictors of conversion from MCI to probable AD in the ADNI cohort. Curr. Alzheimer Res. 6, 347–361.
Sontheimer, H., 2015. Aging, dementia, and Alzheimer disease. Dis. Nerv. Syst. 4, 99–131.
Urigüen, J.A., Garcia-Zapirain, B., 2015. EEG artifact removal—state-of-the-art and guidelines. J. Neural Eng. 12 (3), 031001.
Visser, P.J., Scheltens, P., Verhey, F.R., Schmand, B., Launer, L.J., Jolles, J., Jonker, C., 1999. Medial temporal lobe atrophy and memory dysfunction as predictors for dementia in subjects with mild cognitive impairment. J. Neurol. 246 (6), 477–485.
Zandifar, A., Fonov, V., Coupé, P., Pruessner, J., Collins, D.L., 2017. A comparison of accurate automatic hippocampal segmentation methods. NeuroImage 155, 383–393.