

Article

A Deep Learning approach for Diagnosis of Mild Cognitive Impairment Based on MRI Images

Hamed Taheri Gorji and Naima Kaabouch *

Department of Electrical Engineering, University of North Dakota, Grand Forks, ND 58202-7165, USA

* Correspondence: naima.kaabouch@und.edu

Received: 12 July 2019; Accepted: 26 August 2019; Published: 28 August 2019



Abstract: Mild cognitive impairment (MCI) is an intermediary stage condition between healthy people and Alzheimer's disease (AD) patients and other dementias. AD is a progressive and irreversible neurodegenerative disorder, which is a significant threat to people, age 65 and older. Although MCI does not always lead to AD, an early diagnosis at the stage of MCI can be very helpful in identifying people who are at risk of AD. Moreover, the early diagnosis of MCI can lead to more effective treatment, or at least, significantly delay the disease's progress, and can lead to social and financial benefits. Magnetic resonance imaging (MRI), which has become a significant tool for the diagnosis of MCI and AD, can provide neuropsychological data for analyzing the variance in brain structure and function. MCI is divided into early and late MCI (EMCI and LMCI) and sadly, there is no clear differentiation between the brain structure of healthy people and MCI patients, especially in the EMCI stage. This paper aims to use a deep learning approach, which is one of the most powerful branches of machine learning, to discriminate between healthy people and the two types of MCI groups based on MRI results. The convolutional neural network (CNN) with an efficient architecture was used to extract high-quality features from MRIs to classify people into healthy, EMCI, or LMCI groups. The MRIs of 600 individuals used in this study included 200 control normal (CN) people, 200 EMCI patients, and 200 LMCI patients. This study randomly selected 70 percent of the data to train our model and 30 percent for the test set. The results showed the best overall classification between CN and LMCI groups in the sagittal view with an accuracy of 94.54 percent. In addition, 93.96 percent and 93.00 percent accuracy were reached for the pairs of EMCI/LMCI and CN/EMCI, respectively.

Keywords: deep learning; convolutional neural network; mild cognitive impairment; Alzheimer's disease

1. Introduction

Mild cognitive impairment (MCI), which is considered as a potential forerunner to Alzheimer's disease (AD) and other types of dementia, is a condition in which individuals have a slight but considerable and measurable decline in their mental and cognitive abilities. Although MCI does not always lead to dementia, people with MCI, mainly those who complain about memory problems, are more likely to develop AD [1,2]. The four-year conversion rate between all MCI subtypes to dementia and AD is 56 percent, and 46 percent, respectively [3], and the probability of conversion from MCI to AD is almost seven times higher than in other elderly people [4]. Therefore, MCI can be considered as a transitional state between being healthy and having Alzheimer's [5].

Currently, AD is the sixth leading cause of death in the United States [6]. Approximately 5.7 million Americans are living with AD [7] and this number is predicted to increase to 13.8 million in 2050 [8]. This rapid growth in AD, not only affect the patients and their families, but it is also one of the challenges facing governments. In 2019, AD and other types of dementia have been estimated to cost

the U.S. approximately \$290 billion, and based on this estimation, the cost is expected to increase to more than \$1.1 trillion by 2050 [9].

The ability to accurately and reliably diagnose MCI can lead to benefits to both individuals and governments. The persons diagnosed as having MCI are able to undergo additional technologically advanced and precise tests to find out their current stage, and if it is due to AD or not. With the diagnosis of MCI due to AD, they can be eligible for a number of clinical trials. Moreover, individuals and their families can have more time to plan for social, financial, and medical decisions. Further, it can lead to considerable savings in medical and long-term care costs for both the patients and governments. For example, in the U.S., if all the AD patients were diagnosed in the early stages (during MCI stage), it would save a total of \$7 trillion to \$7.9 trillion [7].

In recent years, based on the new criteria defined by the Alzheimer's Disease Neuroimaging Initiative (ADNI) (ADNI go, ADNI 2), MCI patients have been classified into two groups: Early MCI (EMCI) and late MCI (LMCI). These two groups are discriminated from each other based on the degree of memory impairment. In the EMCI patients, the decline in memory is approximately between 1.0–1.5 standard deviations (SD) below the normative mean, while in LMCI, the decline in memory is at least approximately 1.5 SD below the normative mean [10]. Due to the similarities between the normal aging and MCI patients' brain structures, a diagnosis of the MCI stage based on MRI and the discrimination between these two groups, mainly between EMCI and normal aging, is one of the most challenging parts of aging research.

Brain imaging methods, such as magnetic resonance imaging (MRI), functional magnetic resonance imaging (fMRI), positron emission tomography (PET), single-photon emission computed tomography (SPECT), and computed tomography (CT) are prevalent methods researchers and doctors use to diagnose MCI and AD in patients [11,12].

There are many features in brain images which can be used for a diagnosis and to discriminate between MCI, AD, and healthy people. The cerebral atrophy rate is one of the features that is significantly greater in MCI and AD patients than in normally aging people [13,14]. Ventricle enlargement [15], hippocampal atrophy, and the rate of change of brain atrophy are the other features that are usually used by scientists [16,17]. The changes in gray matter (GM), white matter (WM) and cerebrospinal fluid (CSF) are three other main features of MCI and AD patients [18–20]. Given that, the WM and CSF are considered as the two main features for the diagnosis of MCI, and researchers show that the GM atrophy is more associated with the cognitive performance in MCIs [21–23].

The statistical parametric method (SPM) opened a new window for measuring the spatial distribution of atrophy in brain aging research [24–26]. The SPM was created by Karl Friston and uses statistical techniques for analyzing brain activity in MRI, fMRI, or positron emission tomography (PET) scans [27].

Due to the engineering field of science, it was possible to create computer-aided diagnosis (CAD) systems, which play a crucial role in assisting the researchers and doctors in their interpretation of medical imaging. In recent years, the use of the machine learning approach, especially deep learning methods in CAD systems for the diagnosis and classification of healthy control normal (CN) people, MCI and AD patients, has exponentially increased [28–30]. Gorji et al. developed a Zernike moment-based method for the early diagnosis and classification between AD, MCI and normal control groups from structural MRIs, achieving an accuracy rate of 97.27%, 95.59% and 94.88% for AD/CN, MCI/CN, and AD/MCI respectively. They proposed a novel method that uses Pseudo Zernike moments (PZMs) to extract discriminative information from the MR images of the three mentioned groups. Furthermore, they used two types of artificial neural networks (pattern recognition and learning vector quantization (LVQ) networks), to classify the extracted information of the MRIs.

Ramirez et al. [31] proposed a CAD system for the early detection of AD based on a partial least squares (PLS) regression model and a random forest (RF) predictor using SPECT. They found that PLS outperformed PCA as a feature extraction method with the highest sensitivity, specificity, and accuracy values of 100%, 92.7%, and 96.9%, respectively. Furthermore, their proposed PLS-RF method

outperformed other CAD systems, such as a principal component analysis (PCA)-RF, Gaussian mixture models (GMM)-support vector machine (SVM) and voxel-as-features (VAF)-SVM methods.

In deep learning, convolutional neural networks (CNNs) are widely known for their abilities for image recognition, segmentation, and classification [32–34]. One of the main advantages of CNNs is that, unlike the conventional machine learning methods, there is no need for a manual feature extraction step, and CNNs are able to extract the efficient features automatically and then classify the images.

Current research is directed towards the development of a CNN-based method for the diagnosis and classification of the CN, EMCI and LMCI groups based on GM images extracted from MRIs by SPM. This can be very challenging, but also helpful for the early diagnosis of AD and bring other benefits as mentioned earlier.

2. Materials

2.1. Database

The data used in this study were acquired from the Alzheimer’s disease Neuroimaging Initiative (ADNI) database (<http://www.loni.ucla.edu/ADNI>). The ADNI was funded in 2004 by the National Institute on Aging (NIA), the National Institute of Biomedical Imaging and Bioengineering (NIBIB) and a contribution from many pharmaceutical companies and foundations. The measurement of the progression of the early AD and MCI was the primary goal of ADNI, which was investigated by a combination of clinical and neuropsychological assessments, serial MRI, PET, and other biological markers.

In this study, MRIs of a total of 600 subjects were used, which included 200 EMCI, 200 LMCI patients, and 200 CN individuals. The demographic information of all subjects in the three groups of cognitively normal (CN), EMCI, or LMCI are shown in Table 1.

Table 1. The subjects’ clinical and demographic characteristics. For each group, *N* represents the total number of subjects, *M* and *F* show number of males and females, along with the average age, standard deviation (SD) and average mini-mental state examination (MMSE) score.

	CN (<i>N</i> = 200; 112 F/88 M)		EMCI (<i>N</i> = 200;93 F/107 M)		LMCI (<i>N</i> = 200; 84 F/116 M)	
	Mean	SD	Mean	SD	Mean	SD
Age	74.2	6.1	68.2	6.9	71.1	7.2
MMSE	28.8	1.3	28.4	1.2	27.3	1.8

2.2. Convolutional Neural Network

The convolutional neural networks (CNNs) are a type of deep artificial neural network which was inspired by a hierarchical model of the visual cortex [35]. The main application of the CNNs are in image recognition and classification [33], video classification [36], and medical image and signal analyzing [37–39]. A CNN typically consists of one or more convolution layers, one or more optionally down-sampling layers, a non-linearity layer, and a fully connected layer.

In CNNs, for classifying the input images, each image passes through the convolutional layer. The convolutional layer is comprised of a collection of filters, which plays a crucial role in CNNs architecture. A series of learnable filters are applied to the input images and detect specific patterns and features. This set of filters slide over the input image by a certain amount called the stride value. The stride determines the number of pixels that the filter shifts across the input image matrix. By convoluting between the filters and each part of the input image, a dot product is computed to produce a set of 2D feature maps.

The non-linearity layer, which also is named the activation layer, is used to introduce non-linearity to the CNN because during the convolutional layer, only linear calculations are computed.

In recent years, rectified linear units (ReLU) [33] have been used more as a non-linear function than the other non-linear functions, like sigmoid or hyperbolic tangents, because by means of ReLU ($f(x) = \max(0, x)$) a CNN can train much faster due to the computational efficiency without a notable penalty to the generalization accuracy [33]. Moreover, using ReLU can be very helpful to decrease the vanishing gradient [40].

The pooling layer is the next layer of a CNN and is usually positioned between the convolutional layers. The pooling layer performs down-sampling and reduces the spatial dimension of the feature maps. Therefore, it cannot only reduce the network computational complexity, but also control overfitting. The function of the pooling layers are like the convolutional layers, but instead of doing the convolution operation, it moves over the feature map and replaces the value of each region with its maximum or average, using the defined pool size and stride.

The fully connected (FC) layer, which is the other name for the conventional neural network, computes the class score to generate the outputs equal to the number of classes desired by the CNN. The term, fully connected, implies that all neurons in a fully connected layer are connected to all the neurons in the previous layer. In the end, a sigmoid or softmax classification layer can be used in accordance with the binary or multi-classification task, respectively.

3. Methodology

3.1. Image Preprocessing

In this study, as previously mentioned, the MRIs from the ADNI database have been used. All the data were downloaded with the Neuroimaging Informatics Technology Initiative (NIfTI) format. SPM12 has been used for preprocessing. This study focused on GM because atrophy in GM can be considered as a prominent early AD biomarker [41] and, as it is mentioned earlier, GM is more associated with the cognitive performance in MCIs.

For preprocessing the neuroimages, three main steps were applied to the data which are explained in the following:

3.1.1. Segmentation

By means of SPM segmentation, the brain tissue can be automatically segmented into three main parts, which are GM, WM, and CSF. This paper set the bias regularization on the light regularization (0.001), the full width at half maximum (FWHM) of Gaussian smoothness of bias on the 60 mm cutoff, and affine regularization on the ICBM space template. Moreover, for the spatial normalization of the data to the Montreal Neurological Institute (MNI) spaces, the deformation field was set in the forwarding mode.

3.1.2. Normalization

After the segmentation of the data, GM images were considered for further analysis because, as mentioned before, GM is a significant early AD biomarker. For normalizing all GM images to MNI space, this paper set the written normalized images voxel size on (2 2 2) mm and, for sampling the images to MNI space, the 4th-degree B-Spline for interpolation was considered.

3.1.3. Smoothing

Finally, all normalized GM images were smoothed by a Gaussian kernel. The FWHM of the Gaussian smoothing kernel were set on (2 2 2) mm. The results of the preprocessing steps are shown in Figure 1.

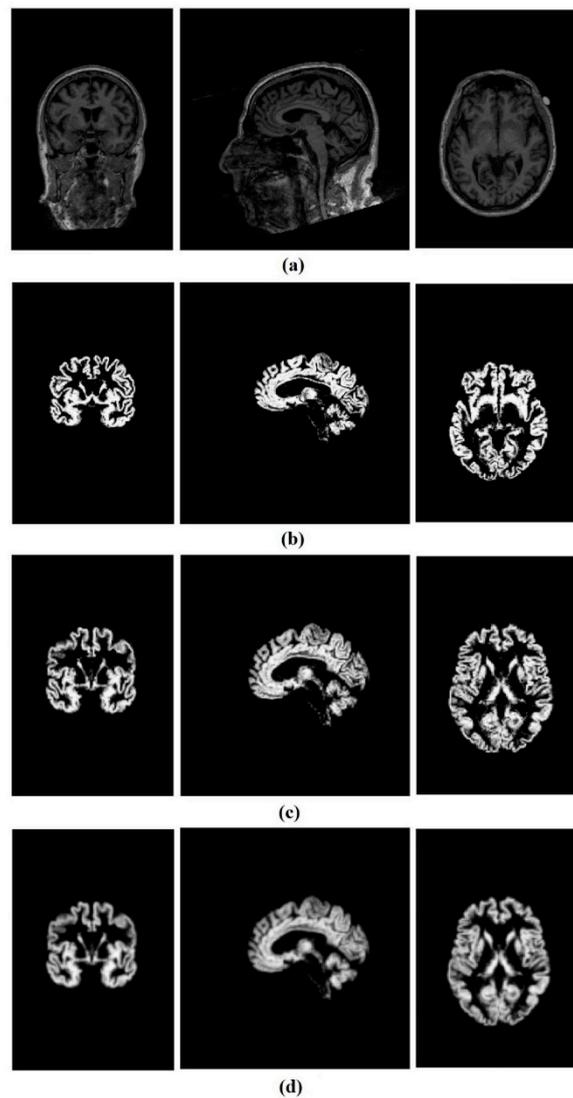


Figure 1. A control healthy subject's MRI, (a) from left to right sagittal, coronal and axial view, (b) gray matter, (c) gray matter after normalization, (d) gray matter after smoothing.

The original size of the data was $176 \times 240 \times 256$, and after the preprocessing steps, all the GM images were reduced to the size of $79 \times 95 \times 79$. For further analysis, 3D images were decomposed into 2D slices along the third direction, which named axial, coronal, and sagittal views. Then, all the 2D.nii GM files were converted to a portable network graphics (PNG) format and resized to 64×64 pixels by means of MATLAB (2018b) to be useable by our CNN. In each view of the MR images, there were some slices at the beginning and the end which did not contain any useful information, so they were removed, and consequently, 20 images of each view were considered for further processing (60 images per subject).

3.2. CNN architecture

The CNN architecture used in this study is composed of three convolutional layers which take an input image with a size of 64×64 (see Figure 2). All three convolutional layers were followed by a max-pooling layer. The 32 filters with a kernel size of 3×3 were considered for the first convolutional layer and the max-pooling layer kernel size was set on 2×2 . The second and third layers comprised of 128 and 512 filters respectively, with the same filter and max-pooling kernel sizes as the first convolution layer. It is worthwhile to mention that ReLU was used as the activation functions in all convolutional

layers. Moreover, the glorot uniform [42] kernel and the bias initializer were used for initializing the layer's weights.

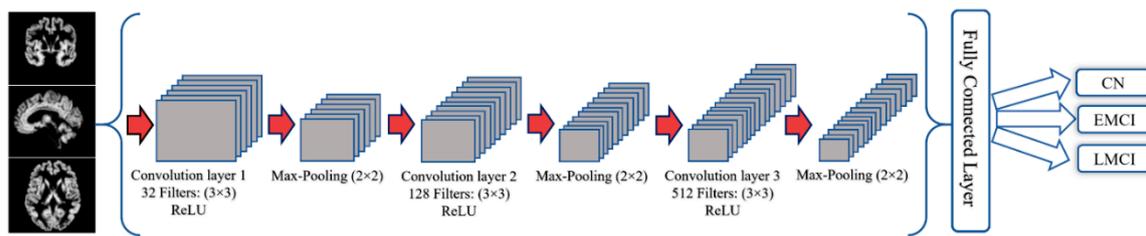


Figure 2. The architecture of the convolutional neural network.

Finally, a fully connected layer with 128 input neurons and a ReLU activation function were used. Further, one output neuron followed by a sigmoid activation function was used in order to facilitate the binary classification. To prevent over-fitting, kernel and bias regularizers and also, the dropout for each convolution layer were employed. The ridge regression regularization technique used for both the kernel and bias regularization and the parameters were set to 0.005 and 0.01, respectively. Moreover, the dropout rates for the three convolution layers were set on 0.3 and the dropout of the fully connected layer to 0.5.

Then, the model complied with the Adam optimizer, which is an algorithm for first-order gradient-based optimization of stochastic objective functions based on the adaptive estimates of lower-order moments [43]. Furthermore, the binary cross-entropy loss function was used to measure the performance of our classification model.

One of the essential conditions for achieving excellent performance in a deep neural network is a large amount of training data. Image augmentation is a technique used when the training set is not large enough. It uses some image processing methods, such as random rotations, horizontal/vertical flips or shifts, shear, changes the image brightness, etc. to artificially increase the amount of data to boost the performance of a deep neural network.

In the current study, 12,000 images were used from each of the three views: Sagittal, coronal and axial coming equally from each group (CN, EMCI and LMCI) for a total of 36,000 images used in our study.

These three mentioned groups were classified based on the three different views, and for each view there were just 4000 images per group. Therefore, this study decided to use image augmentation to generate additional training data by means of sheering, random rotation and zooming. In the following, 70 percent of the images for training and 30 percent for testing were randomly chosen. Further, 20 percent of the training set were selected for the validation set, which provides an unbiased evaluation of the proposed model during the training phase. The batch size was set to 512 images and used 300 epochs for our CNN.

One of the ambiguities of a CNN is how it processes the data, how it extracts the features, and what the features' map looks like. To answer these questions, this study illustrated the output of the first and second convolution layers and also the max-pooling layers as shown in Figure 3. The input image is an axial view of the MRI similar to what was showed above. In Figure 3a, the 32 filters from the first convolution layer are presented and one can see the first convolution layer retains the input picture shape and information completely. Moreover, the blank columns show the filters that are not activated. The role of the max-pooling layer for reducing the spatial dimension of the feature maps is presented in Figure 3b. In a CNN, the deeper that is gone, the lesser the output is interpretable, and more image class-related features can be extracted. As shown in Figure 3c, the output of the second convolution layer (128 filters) is less transparent than the first layer. Finally, Figure 3d shows the output of the second map-pooling layer.

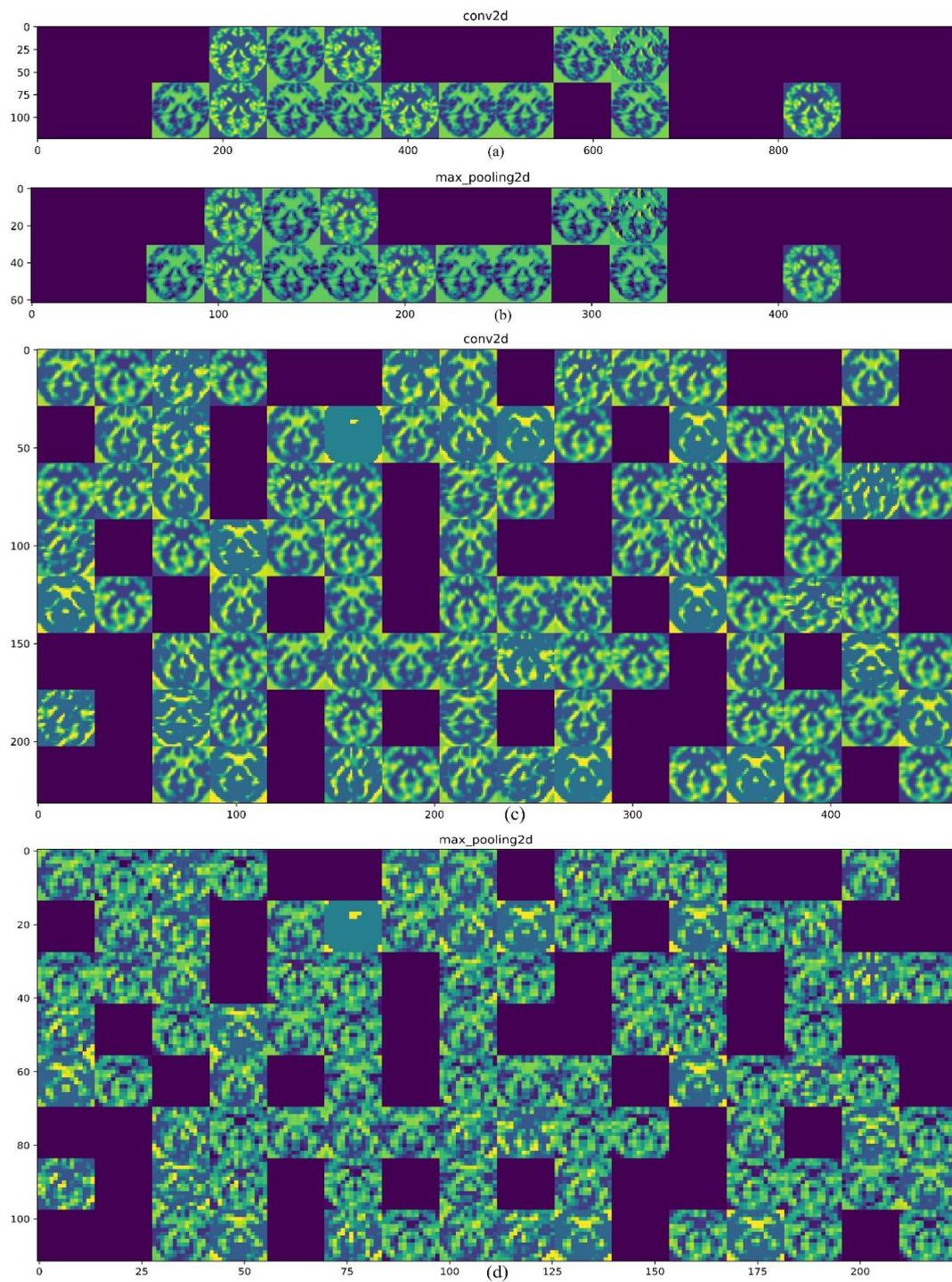


Figure 3. Illustration of the convolutional neural network (CNN) layers output. (a) First convolution layer output; (b) first max-pooling layer output; (c) second convolution layer output; (d) second max-pooling layer output.

The performance of the proposed CAD system was evaluated using five metrics which are accuracy, sensitivity, specificity, F-score, and receiver operating characteristic-area under the curve (AUC-ROC). The parameters definitions are as follows:

$$\text{Sensitivity} = \frac{\text{TP}}{\text{TP} + \text{FN}} \times 100 \quad (1)$$

$$\text{Specificity} = \frac{\text{TN}}{\text{TN} + \text{FP}} \times 100 \quad (2)$$

$$\text{Accuracy} = \frac{\text{TP} + \text{TN}}{\text{TP} + \text{FP} + \text{TN} + \text{FN}} \times 100 \quad (3)$$

$$\text{F-Score} = 2 \times \frac{\text{P} \times \text{R}}{\text{P} + \text{R}} \times 100 \quad (4)$$

In this study TP, TN, FP and FN denote the true positive (i.e., the number of EMCI or LMCI patients who were correctly classified), true negative (i.e., the number of NC, which were correctly classified), false positive (i.e., the number of EMCI or LMCI who were classified as NC) and false negative (i.e., the number of NC group, which were classified as EMCI or LMCI) respectively. Further, P and R in the F-score equation are denoted by the precision ($\frac{\text{TP}}{\text{TP} + \text{FP}}$) and recall ($\frac{\text{TP}}{\text{TP} + \text{FN}}$), respectively.

According to the definition above, sensitivity reflects how many EMCI or LMCI subjects were detected accurately. The higher the sensitivity, the fewer AD cases missed. The specificity reflects how many NC cases were detected accurately. The higher the specificity, the fewer normal subjects were misrecognized as EMCI or LMCI. The accuracy represents the ability of the designed system to differentiate between NC, EMCI, and LMCI groups correctly. In addition, the F-score which has a realistic accuracy measurement of the test ROC represents the probability curve, and AUC illustrates the measure of separability. Using these metrics together, the performance of the proposed method can be evaluated comprehensively.

4. Results

4.1. Classification of CN and LMCI

The classification results for pairs of CN/LMCI for sagittal, coronal, and axial views are shown in Table 2 and also ROC-AUC values are presented in Figure 4. The best classification results were attained from the sagittal view of GM images for the CN and LMCI groups with an accuracy of 94.54%, an F-score of 94.84%, and an AUC of 99.40% (sensitivity of 91.70%, specificity of 97.96). For the axial view, the proposed CAD system reached an accuracy of 93.18%, an F-score of 93.53%, and an AUC of 98.40% (sensitivity of 90.02%, specificity of 97.01%). Moreover, the results for the differentiation between CN/LMCI for the coronal view gave an accuracy of 91.65%, an F-score of 92.19%, and an AUC of 97.70% (sensitivity of 90.28%, specificity of 93.30%).

Table 2. The classification results of the control normal (CN) versus early mild cognitive impairment (EMCI), CN versus late mild cognitive impairment (LMCI) and EMCI versus LMCI.

	MRI Views	Sensitivity (%)	Specificity (%)	Accuracy (%)	F-Score (%)	AUC (%)
CN vs. LMCI	Sagittal	91.70	97.96	94.54	94.84	99.40
	Axial	90.02	97.01	93.18	93.53	98.40
	Coronal	90.28	93.30	91.65	92.19	97.70
CN vs. EMCI	Sagittal	90.46	98.19	93.96	94.25	98.80
	Axial	90.63	91.42	90.99	91.65	97.00
	Coronal	88.60	89.95	89.21	89.96	95.10
EMCI vs. LMCI	Sagittal	91.48	94.82	93.00	93.46	98.10
	Axial	87.01	94.57	90.45	90.86	96.70
	Coronal	85.44	92.07	88.45	88.98	93.60

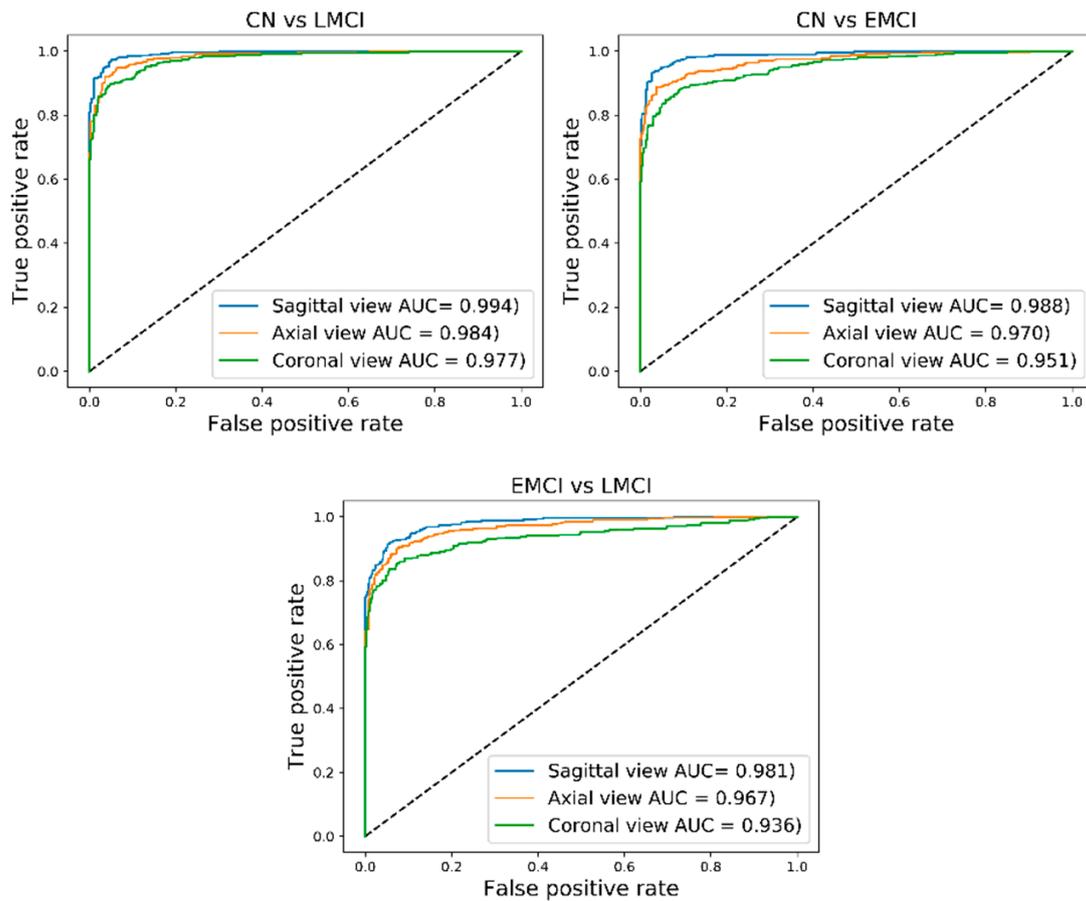


Figure 4. Receiver operating characteristic-area under the curve (ROC-AUC) results of the sagittal, coronal, and axial views.

4.2. Classification of CN and EMCI

For the pairs of CN/EMCI, the best classification results were achieved for the sagittal view with an accuracy of 93.96%, an F-score of 94.25%, and an AUC of 98.80% (sensitivity of 90.46%, specificity 98.19%). The views of the axial and coronal obtained lower results than the sagittal view. The achieved results for the axial view were an accuracy of 90.99%, an F-score of 91.65%, an AUC of 97% (sensitivity of 90.63%, specificity of 91.42%). For the coronal view, an accuracy 89.21%, an F-score of 89.96%, and an AUC of 95.10% (sensitivity of 88.60%, specificity 89.95%) was obtained. The classification results and ROC-AUC values are shown in Table 2 and Figure 4.

4.3. Classification of EMCI and LMCI

The highest overall classification results for the EMCI and LMCI groups which were obtained for the sagittal view are as follows: Sensitivity of 91.48%, specificity of 94.82%, an accuracy of 93%, an F-score 93.46%, and AUC of 98.10%. The classification results of the axial view reached a sensitivity of 87.01%, a specificity of 94.57%, an accuracy of 90.45%, an F-score of 90.86, and an AUC of 96.70%. The sensitivity, specificity, accuracy, F-score, and AUC of the coronal view were 85.44%, 92.07%, 88.45%, 88.98%, and 93.60%, respectively. The classification results of EMCI and LMCI groups are shown in Table 2, and the ROC-AUC values are illustrated in Figure 4.

5. Discussion

Researchers have recently conducted several studies looking into the early diagnosis of AD using deep learning techniques, such as convolutional neural networks (CNNs) [44–46]. This research has achieved high accuracy for the classification of CN individuals, the patients with MCI, and AD patients.

MCI status is critical in the early diagnosis of AD because patients with late MCI (LMCI) are classified as having a very high risk of conversion to AD [47], and, more importantly, EMCI can be considered to be the starting point of AD. An accurate and reliable diagnosis of MCI can result in identifying individuals who are at an increased risk of the progression to dementia, and opens doors for providing the potential and routine treatment and gives people the opportunity to plan for the future. Thus, designing an accurate and reliable CAD system for the classification of CN, EMCI, and LMCI patients could be a substantial step forward towards the aging research field. This study, therefore, focused on discriminating CN people from the patients with EMCI and LMCI.

In recent years, deep learning, mainly employing CNNs, has become popular for solving complex tasks, such as image and video processing. This fully trainable system outperforms the traditional methods of classification when using very large data, and unlike conventional machine learning methods, it does not require manual feature extraction steps. A CNN was therefore used in the current study to automatically extract the discriminative features for classifying CN, EMCI, and LMCI patients. The performance of the CAD system the authors have designed has been evaluated using the five parameters of accuracy, F-score, sensitivity, specificity and ROC-AUC.

Several CNN architectures were tried and tested to obtain a reliable and accurate CAD system for the classification of the three groups mentioned above, and the best results were achieved with a CNN consisting of three convolution layers of 32, 128 and 512 filters, respectively. Each convolution layer was also accompanied by a max-pooling layer for down-sampling the spatial dimensions of the input data. It is worth mentioning that to prevent the co-adaptation of feature detectors [48] and to solve the overfitting problem, a dropout was employed for each convolution layer. Finally, the network includes a fully connected layer that is connected to all the neurons in the last convolution layer and a classification layer with a sigmoid activation function that is used to classify the images.

The ADNI MRI data were randomly divided into two groups, one for training and one for testing, comprising 70% and 30% of the total dataset, respectively. The GM was extracted from MRI images of the sagittal, axial and coronal planes and fed into the CNN described above. As shown in Table 2, our proposed deep learning-based CAD system obtained high values for accuracy, F-score, ROC-AUC as well as sensitivity and specificity when attempting to classify the following pairs: CN versus EMCI, CN versus LMCI, and EMCI versus LMCI. The proposed method yielded the best overall results with CN versus LMCI pairs using a sagittal view (accuracy of 94.54%, F-score of 94.84%, and AUC of 99.40%). Moreover, our results show that the sagittal view also leads to better classification performance for CN versus EMCI pairs and EMCI versus LMCI pairs, (accuracy of 93.96%, F-score of 94.25%, AUC of 98.80%, F-score of 93.46%, and an AUC of 98.10%, respectively).

Recently, several studies using different approaches have investigated the early diagnosis of AD based on the classification of healthy normal people, patients with MCI, and AD patients [29,49–51]. However, to the best of the authors' knowledge, there are only three studies that have investigated the classification of healthy normal people, EMCI patients, and LMCI patients [52–54]. The highest accuracy achieved by Korolev et al. was 73% for LMCI versus NC, 67% for LMCI versus EMCI, and 63% for EMCI versus NC. Cabrera-León et al. reported an average accuracy for several different methods, of which the best was 57.6%, achieved using the random forest method. The highest F-scores reported by Singh et al. were 83.25% for cognitively unimpaired controls (CU) versus LMCI, 72% for CU versus EMCI, and 68.44% for LMCI versus EMCI.

Thus, comparing this study's results with the results reported by other studies, it can be concluded that our proposed CAD system is a more reliable and accurate approach.

6. Conclusions

Deep neural networks, especially convolutional neural networks, can provide meaningful information for the diagnosis and prognosis of MCI. In this paper, a CNN based method was proposed for extracting discriminative features from structural MRI, with the aim of the diagnosis of EMCI and LMCI and the classification between these two groups and healthy subjects. The proposed method can lead to several benefits for potential MCI individuals and also can lead to an early diagnosis of AD. The experimental results on the ADNI database for 600 subjects demonstrated that our proposed method for feature extraction and classification delivered a high accuracy for the EMCI, LMCI, and CN groups. The best results were achieved for the classification between CN and LMCI groups in the sagittal view and also, the pairs of EMCI/LMCI have achieved slightly better accuracy than CN/EMCI concerning all views of the MRI.

The proposed method yielded a 94.54% classification accuracy (94.84% F-score and 99.40% AUC) for CN versus LMCI, 93.96% classification accuracy for the pairs of CN/EMCI (94.25% F-score and 98.80% AUC), and 93.00% classification accuracy for the classification of the pairs of EMCI/LMCI (93.46% F-score and 98.10% AUC) which all of the above mentioned results achieved from the sagittal view. The above-mentioned results demonstrate higher reliability and precision of our proposed method for the diagnosis of MCI and the classification between the three groups of CN, EMCI, and LMCI.

Author Contributions: H.T.G. and N.K. contributed to the design and implementation of the research, to the analysis of the results and to the writing of the manuscript.

Funding: This research received no external funding.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Kantarci, K.; Weigand, S.; Przybelski, S.; Shiung, M.; Whitwell, J.L.; Negash, S.; Knopman, D.S.; Boeve, B.F.; O'Brien, P.; Petersen, R.C. Risk of dementia in MCI: Combined effect of cerebrovascular disease, volumetric MRI, and 1H MRS. *Neurology* **2009**, *72*, 1519–1525. [[CrossRef](#)] [[PubMed](#)]
2. Mitchell, A.J.; Shiri-Feshki, M. Rate of progression of mild cognitive impairment to dementia—Meta-analysis of 41 robust inception cohort studies. *Acta Psychiatr. Scand.* **2009**, *119*, 252–265. [[CrossRef](#)] [[PubMed](#)]
3. Rountree, S.; Waring, S.; Chan, W.; Lupo, P.; Darby, E.; Doody, R. Importance of subtle amnesic and nonamnesic deficits in mild cognitive impairment: Prognosis and conversion to dementia. *Dement. Geriatr. Cogn. Disord.* **2007**, *24*, 476–482. [[CrossRef](#)] [[PubMed](#)]
4. Boyle, P.; Wilson, R.; Aggarwal, N.; Tang, Y.; Bennett, D. Mild cognitive impairment Risk of Alzheimer disease and rate of cognitive decline. *Neurology* **2006**, *67*, 441–445. [[CrossRef](#)] [[PubMed](#)]
5. DeCarli, C. Mild cognitive impairment: Prevalence, prognosis, aetiology and treatment. *Lancet Neurol.* **2003**, *2*, 15–21. [[CrossRef](#)]
6. Kochanek, K.D.; Murphy, S.L.; Xu, J.; Tejada-Vera, B. Deaths: Final data for 2014. *Natl. Vital. Stat. Rep.* **2016**, *65*, 1–122. [[PubMed](#)]
7. Association, A.S. 2018 Alzheimer's disease facts and figures. *Alzheimer Dement.* **2018**, *14*, 367–429.
8. Hebert, L.E.; Weuve, J.; Scherr, P.A.; Evans, D.A. Alzheimer disease in the United States (2010–2050) estimated using the 2010 census. *Neurology* **2013**, *80*, 1778–1783. [[CrossRef](#)] [[PubMed](#)]
9. Gaugler, J.; James, B.; Johnson, T.; Marin, A.; Weuve, J. 2019 Alzheimer's disease facts and figures. *Alzheimer Dement. J. Alzheimer Assoc.* **2019**, *15*, 321–387. [[CrossRef](#)]
10. Aisen, P.S.; Petersen, R.C.; Donohue, M.C.; Gamst, A.; Raman, R.; Thomas, R.G.; Walter, S.; Trojanowski, J.Q.; Shaw, L.M.; Beckett, L.A. Clinical Core of the Alzheimer's Disease Neuroimaging Initiative: Progress and plans. *Alzheimer Dement.* **2010**, *6*, 239–246. [[CrossRef](#)]
11. Johnson, K.A.; Fox, N.C.; Sperling, R.A.; Klunk, W.E. Brain imaging in Alzheimer disease. *Cold Spring Harb. Perspect. Med.* **2012**, *2*, a006213. [[CrossRef](#)] [[PubMed](#)]
12. McGeer, P.L. Brain imaging in Alzheimer's disease. *Br. Med. Bull.* **1986**, *42*, 24–28. [[CrossRef](#)] [[PubMed](#)]

13. Fox, N.C.; Schott, J.M. Imaging cerebral atrophy: Normal ageing to Alzheimer's disease. *Lancet* **2004**, *363*, 392–394. [[CrossRef](#)]
14. Tabatabaei-Jafari, H.; Shaw, M.E.; Cherbuin, N. Cerebral atrophy in mild cognitive impairment: A systematic review with meta-analysis. *Alzheimer Dement. Diagn. Assess. Dis. Monit.* **2015**, *1*, 487–504. [[CrossRef](#)] [[PubMed](#)]
15. Nestor, S.M.; Rupsingh, R.; Borrie, M.; Smith, M.; Accomazzi, V.; Wells, J.L.; Fogarty, J.; Bartha, R.; Initiative, A.D.N. Ventricular enlargement as a possible measure of Alzheimer's disease progression validated using the Alzheimer's disease neuroimaging initiative database. *Brain* **2008**, *131*, 2443–2454. [[CrossRef](#)] [[PubMed](#)]
16. Prados, F.; Cardoso, M.J.; Leung, K.K.; Cash, D.M.; Modat, M.; Fox, N.C.; Wheeler-Kingshott, C.A.; Ourselin, S.; Initiative, A.D.N. Measuring brain atrophy with a generalized formulation of the boundary shift integral. *Neurobiol. Aging* **2015**, *36*, S81–S90. [[CrossRef](#)]
17. Henneman, W.; Sluimer, J.; Barnes, J.; Van Der Flier, W.; Sluimer, I.; Fox, N.; Scheltens, P.; Vrenken, H.; Barkhof, F. Hippocampal atrophy rates in Alzheimer disease: Added value over whole brain volume measures. *Neurology* **2009**, *72*, 999–1007. [[CrossRef](#)]
18. Wang, Y.; West, J.D.; Flashman, L.A.; Wishart, H.A.; Santulli, R.B.; Rabin, L.A.; Pare, N.; Arfanakis, K.; Saykin, A.J. Selective changes in white matter integrity in MCI and older adults with cognitive complaints. *Biochim. Biophys. Acta Mol. Basis Dis.* **2012**, *1822*, 423–430. [[CrossRef](#)]
19. Zhang, H.; Sachdev, P.S.; Wen, W.; Kochan, N.A.; Crawford, J.D.; Brodaty, H.; Slavin, M.J.; Reppermund, S.; Draper, B.; Zhu, W. Gray matter atrophy patterns of mild cognitive impairment subtypes. *J. Neurol. Sci.* **2012**, *315*, 26–32. [[CrossRef](#)]
20. Popp, J.; Wolfsgruber, S.; Heuser, I.; Peters, O.; Hüll, M.; Schröder, J.; Möller, H.-J.; Lewczuk, P.; Schneider, A.; Jahn, H. Cerebrospinal fluid cortisol and clinical disease progression in MCI and dementia of Alzheimer's type. *Neurobiol. Aging* **2015**, *36*, 601–607. [[CrossRef](#)]
21. Zhang, Y.; Schuff, N.; Camacho, M.; Chao, L.L.; Fletcher, T.P.; Yaffe, K.; Woolley, S.C.; Madison, C.; Rosen, H.J.; Miller, B.L. MRI markers for mild cognitive impairment: Comparisons between white matter integrity and gray matter volume measurements. *PLoS ONE* **2013**, *8*, e66367. [[CrossRef](#)] [[PubMed](#)]
22. Grundman, M.; Sencakova, D.; Jack, C.R.; Petersen, R.C.; Kim, H.T.; Schultz, A.; Weiner, M.F.; DeCarli, C.; DeKosky, S.T.; Van Dyck, C. Brain MRI hippocampal volume and prediction of clinical status in a mild cognitive impairment trial. *J. Mol. Neurosci.* **2002**, *19*, 23–27. [[CrossRef](#)] [[PubMed](#)]
23. Loewenstein, D.A.; Acevedo, A.; Potter, E.; Schinka, J.A.; Raj, A.; Greig, M.T.; Agron, J.; Barker, W.W.; Wu, Y.; Small, B. Severity of medial temporal atrophy and amnesic mild cognitive impairment: Selecting type and number of memory tests. *Am. J. Geriatr. Psychiatry* **2009**, *17*, 1050–1058. [[CrossRef](#)] [[PubMed](#)]
24. Chetelat, G.; Desgranges, B.; De La Sayette, V.; Viader, F.; Eustache, F.; Baron, J.-C. Mapping gray matter loss with voxel-based morphometry in mild cognitive impairment. *Neuroreport* **2002**, *13*, 1939–1943. [[CrossRef](#)] [[PubMed](#)]
25. Guo, X.; Wang, Z.; Li, K.; Li, Z.; Qi, Z.; Jin, Z.; Yao, L.; Chen, K. Voxel-based assessment of gray and white matter volumes in Alzheimer's disease. *Neurosci. Lett.* **2010**, *468*, 146–150. [[CrossRef](#)]
26. Davatzikos, C.; Bhatt, P.; Shaw, L.M.; Batmanghelich, K.N.; Trojanowski, J.Q. Prediction of MCI to AD conversion, via MRI, CSF biomarkers, and pattern classification. *Neurobiol. Aging* **2011**, *32*, 2322.e19–2322.e27. [[CrossRef](#)] [[PubMed](#)]
27. Friston, K.J.; Holmes, A.P.; Worsley, K.J.; Poline, J.P.; Frith, C.D.; Frackowiak, R.S. Statistical parametric maps in functional imaging: A general linear approach. *Hum. Brain Mapp.* **1994**, *2*, 189–210. [[CrossRef](#)]
28. Gorji, H.; Haddadnia, J. A novel method for early diagnosis of Alzheimer's disease based on pseudo Zernike moment from structural MRI. *Neuroscience* **2015**, *305*, 361–371. [[CrossRef](#)]
29. Liu, S.; Liu, S.; Cai, W.; Pujol, S.; Kikinis, R.; Feng, D. Early diagnosis of Alzheimer's disease with deep learning. In Proceedings of the 2014 IEEE 11th International Symposium on Biomedical Imaging (ISBI), Beijing, China, 29 April–2 May 2014; pp. 1015–1018.
30. Sarraf, S.; Tofighi, G. Classification of Alzheimer's disease using fmri data and deep learning convolutional neural networks. *arXiv* **2016**, arXiv:1603.08631.
31. Ramírez, J.; Górriz, J.; Segovia, F.; Chaves, R.; Salas-Gonzalez, D.; López, M.; Álvarez, I.; Padilla, P. Computer aided diagnosis system for the Alzheimer's disease based on partial least squares and random forest SPECT image classification. *Neurosci. Lett.* **2010**, *472*, 99–103. [[CrossRef](#)]

32. Traore, B.B.; Kamsu-Foguem, B.; Tangara, F. Deep convolution neural network for image recognition. *Ecol. Inform.* **2018**, *48*, 257–268. [[CrossRef](#)]
33. Krizhevsky, A.; Sutskever, I.; Hinton, G.E. Imagenet classification with deep convolutional neural networks. In Proceedings of the Advances in Neural Information Processing Systems, Lake Tahoe, CA, USA, 3 December 2012; pp. 1097–1105.
34. Ronneberger, O.; Fischer, P.; Brox, T. U-net: Convolutional networks for biomedical image segmentation. In *Lecture Notes in Computer Science, Proceedings of the International Conference on Medical Image Computing and Computer-Assisted Intervention, Munich, Germany, 5–9 October 2015*; Springer: Cham, Switzerland, 2015; pp. 234–241.
35. LeCun, Y.; Bengio, Y.; Hinton, G. Deep learning. *Nature* **2015**, *521*, 436. [[CrossRef](#)] [[PubMed](#)]
36. Karpathy, A.; Toderici, G.; Shetty, S.; Leung, T.; Sukthankar, R.; Fei-Fei, L. Large-scale video classification with convolutional neural networks. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Columbus, OH, USA, 24–27 June 2014; pp. 1725–1732.
37. Acharya, U.R.; Oh, S.L.; Hagiwara, Y.; Tan, J.H.; Adeli, H.; Subha, D.P. Automated EEG-based screening of depression using deep convolutional neural network. *Comput. Methods Programs Biomed.* **2018**, *161*, 103–113. [[CrossRef](#)] [[PubMed](#)]
38. Spasov, S.E.; Passamonti, L.; Duggento, A.; Liò, P.; Toschi, N. A Multi-modal Convolutional Neural Network Framework for the Prediction of Alzheimer’s Disease. In Proceedings of the 2018 40th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), Honolulu, HI, USA, 18–21 July 2018; pp. 1271–1274.
39. Payan, A.; Montana, G. Predicting Alzheimer’s disease: A neuroimaging study with 3D convolutional neural networks. *arXiv* **2015**, arXiv:1502.02506.
40. Maas, A.L.; Hannun, A.Y.; Ng, A.Y. Rectifier nonlinearities improve neural network acoustic models. In Proceedings of the Proc. Icml, Atlanta, GA, USA, 16–21 June 2013; p. 3.
41. Jacobs, H.I.; van Boxtel, M.P.; Gronenschild, E.H.; Uylings, H.B.; Jolles, J.; Verhey, F.R. Decreased gray matter diffusivity: A potential early Alzheimer’s disease biomarker? *Alzheimer Dement.* **2013**, *9*, 93–97. [[CrossRef](#)] [[PubMed](#)]
42. Glorot, X.; Bengio, Y. Understanding the difficulty of training deep feedforward neural networks. In Proceedings of the Thirteenth International Conference on Artificial Intelligence and Statistics, Sardinia, Italy, 13–15 May 2010; pp. 249–256.
43. Kingma, D.P.; Ba, J. Adam: A method for stochastic optimization. *arXiv* **2014**, arXiv:1412.6980.
44. Wang, H.; Shen, Y.; Wang, S.; Xiao, T.; Deng, L.; Wang, X.; Zhao, X. Ensemble of 3D densely connected convolutional network for diagnosis of mild cognitive impairment and Alzheimer’s disease. *Neurocomputing* **2019**, *333*, 145–156. [[CrossRef](#)]
45. Ju, R.; Hu, C.; Zhou, P.; Li, Q. Early diagnosis of Alzheimer’s disease based on resting-state brain networks and deep learning. *IEEE ACM Trans. Comput. Biol. Bioinform. (TCBB)* **2019**, *16*, 244–257. [[CrossRef](#)]
46. Huang, Y.; Xu, J.; Zhou, Y.; Tong, T.; Zhuang, X. Diagnosis of Alzheimer’s Disease via Multi-modality 3D Convolutional Neural Network. *arXiv* **2019**, arXiv:1902.09904. [[CrossRef](#)]
47. Jessen, F.; Wolfgruber, S.; Wiese, B.; Bickel, H.; Mösch, E.; Kaduszkiewicz, H.; Pentzek, M.; Riedel-Heller, S.G.; Luck, T.; Fuchs, A. AD dementia risk in late MCI, in early MCI, and in subjective memory impairment. *Alzheimer Dement.* **2014**, *10*, 76–83. [[CrossRef](#)]
48. Hinton, G.E.; Srivastava, N.; Krizhevsky, A.; Sutskever, I.; Salakhutdinov, R.R. Improving neural networks by preventing co-adaptation of feature detectors. *arXiv* **2012**, arXiv:1207.0580.
49. Ortiz, A.; Munilla, J.; Gorriz, J.M.; Ramirez, J. Ensembles of deep learning architectures for the early diagnosis of the Alzheimer’s disease. *Int. J. Neural Syst.* **2016**, *26*, 1650025. [[CrossRef](#)] [[PubMed](#)]
50. Suk, H.-I.; Shen, D. Deep learning-based feature representation for AD/MCI classification. In *Lecture Notes in Computer Science, Proceedings of the International Conference on Medical Image Computing and Computer-Assisted Intervention, Nagoya, Japan, 22–26 September 2013*; Springer: Berlin/Heidelberg, Germany, 2013; pp. 583–590.
51. Li, F.; Tran, L.; Thung, K.-H.; Ji, S.; Shen, D.; Li, J. Robust deep learning for improved classification of AD/MCI patients. In *Lecture Notes in Computer Science, Proceedings of the International Workshop on Machine Learning in Medical Imaging, Boston, MA, USA, 14 September 2014*; Springer: Cham, Switzerland, 2014; pp. 240–247.

52. Korolev, S.; Safiullin, A.; Belyaev, M.; Dodonova, Y. Residual and plain convolutional neural networks for 3D brain MRI classification. In Proceedings of the 2017 IEEE 14th International Symposium on Biomedical Imaging (ISBI 2017), Melbourne, VIC, Australia, 18–21 April 2017; pp. 835–838.
53. Cabrera-León, Y.; Báez, P.G.; Ruiz-Alzola, J.; Suárez-Araujo, C.P. Classification of Mild Cognitive Impairment Stages Using Machine Learning Methods. In Proceedings of the 2018 IEEE 22nd International Conference on Intelligent Engineering Systems (INES), Las Palmas de Gran Canaria, Spain, 21–23 June 2018; pp. 000067–000072.
54. Singh, S.; Srivastava, A.; Mi, L.; Caselli, R.J.; Chen, K.; Goradia, D.; Reiman, E.M.; Wang, Y. Deep-learning-based classification of FDG-PET data for Alzheimer’s disease categories. In Proceedings of the 13th International Conference on Medical Information Processing and Analysis, San Andres Island, Colombia, 17 November 2017; p. 105720.



© 2019 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).