



Multiplane Convolutional Neural Network (Mp-CNN) for Alzheimer's Disease Classification

Cucun Very Angkoso^{1,2}Hapsari Peni Agustin Tjahyaningtjas³
Mauridhi Hery Purnomo^{1,4*}I Ketut Eddy Purnama^{1,4}¹*Department of Electrical Engineering, Institut Teknologi Sepuluh Nopember, Indonesia*²*Informatics Department, University of Trunojoyo Madura, Indonesia*³*Department of Electrical Engineering, Universitas Negeri Surabaya, Indonesia*⁴*Departement of Computer Engineering, UCE AIHeS, Institut Teknologi Sepuluh Nopember Surabaya, Indonesia**Corresponding author's Email: cucun.17071@mhs.its.ac.id, hery@ee.its.ac.id

Abstract: One of the objectives of medical imaging research is to develop an effective and reliable clinical support tool for the early detection of various neurological conditions in patients such as Alzheimer's Disease (AD). Classification based on three-dimensional (3D) MRI (Magnetic-Resonance-Imaging) images have shown a good performance. However, solving 3D object classification as a 3D object using classical machine learning or deep learning is computationally high. 3D images from MRI can be reconstructed into three depth directions: axial, coronal, and sagittal. However, the multi-view-based method cannot explore the content of each image slice for all planes in 3D-MRI. Our study proposes a multiplane method by using three two-dimensional (2D) CNNs to capture discriminatory information on all available planes. Our method is a new fusion strategy to deal with the disadvantages of shape-based multi-view techniques. Our framework selects three slices: the three largest 3-planes to represent whole 3D-MRI objects as multi-2DCNN inputs. Extensive experiments have shown that the proposed method can outperform the 2DCNN method, which uses only one plane. More importantly, by taking full advantage of 2DCNN, we offer a new method for identifying 3D objects that is both easy and efficient. We called this new architecture Multiplane Convolutional Neural Network (Mp-CNN) since it used multiple inputs in its design. We evaluated the proposed method using T1-weighted structural MRI data consisting of 500 AD, 500 Mild Cognitive Impairment (MCI), and 500 Normal Cognition (NC) subjects collected from the MRI database of Alzheimer's Disease Neuroimaging Initiative (ADNI). From the performed experiments, the proposed method achieves 93% accuracy for multiclass AD-MCI-NC and good precision AD 93%, MCI 91%, and NC 95%, respectively. Our study also demonstrates that the proposed multiplane approach outperforms the single-plane approach.

Keywords: Alzheimer's disease neuroimaging initiative, Multiplane, Convolutional neural network, Alzheimer's disease.

1. Introduction

Dementia is a weakening of the human brain's ability, characterized by changes in personality, often forgetting something, and emotional unpredictability. As a common form of dementia, Alzheimer's disease (AD) is a neurodegenerative disease that attacks the structures of the brain throughout its duration. Brain chemicals and structures change during the illness, resulting in the death of brain cells and the patient's ability to perform daily tasks [1]. AD must be realized

as a disease and not a regular occurrence to everyone, so research studies for early detection, treatment, and medication are necessary. Early treatment of the patient is critical for preventing the progression of AD by initiating therapy and prevention as soon as the disease is identified. Early detection and research activities that better understand the disease progression and new treatments are expected to lower healthcare costs [2, 3].

Mild cognitive impairment (MCI) is a classification label for patients who have AD but are

still in the early stages. However, it should be noted that not all patients with MCI at this stage will later change their AD. MCI is a stage between normal and AD in which a person's cognitive abilities decrease slightly, but they can still do daily tasks. MCI affects about 15-20 % of adults over 65, and 30-40 % of MCI will convert AD in the next five years [4].

Machine learning has been widely used for medical imaging analysis in recent years. Machine learning frameworks evolve rapidly, giving rise to a new deep learning technique based on neural networks. Deep neural networks (deep learning) are recognized as powerful and popular algorithms in analyzing medical images. The convolutional neural network (CNN) is one of the most popular types of neural networks within deep neural networks and deep learning, especially when dealing with high-dimensional data such as images and videos. Furthermore, CNN has recently performed well in medical image analysis to detect various diseases [5]. CNN's concept is similar to multilayer perceptron (MLP). Still, every neuron on CNN is represented in two dimensions to store spatial information from image data, different from MLP, which considers each pixel an independent feature.

Magnetic resonance imaging (MRI) scan tool is commonly used to observe the human brain. Magnetic resonance (MR) brain imaging is a non-invasive approach for quantifying neurological diseases such as AD. Even before clinical symptoms occur or nerve damage has occurred, MR imaging can provide useful indicators. Studies suggest that the use of MRI generally carries a lesser health risk than other modalities such as computed tomography (CT) and positron emission tomography (PET) [6]. In addition, MRI is preferred for analyzing the brain's anatomical structure because of its high spatial resolution and ability to distinguish soft tissues.

Image processing technologies and artificial intelligence in computer systems can help enhance disease detection accuracy and consistency in hospitals. Advances in medical imaging analysis have resulted in sophisticated tools for diagnosing neurodegeneration, and interest in using imaging data to diagnose disease increases. Furthermore, computers have recently been as accurate as radiologists in their assessments [7].

Several studies involving various binary and multiclass classification methods, feature extraction algorithms, and feature selection strategies have been conducted to diagnose AD using machine learning [8–12]. Many methods and evaluations of machine learning techniques indicate that typical machine learning approaches are insufficient to deal with complex challenges like AD classification [13]. The

difficulty in classifying AD comes from the fact that the features and image patterns in the brain look similar, so addressing the problem requires the utilization of advanced methods. Deep learning networks are known for having various advantages over conventional machine learning approaches [14].

Deep learning requires huge amounts of data to perform better than other methods and is very expensive to train due to complex data models. However, thanks to advances in the graphics processing unit (GPU), the benefits of this method can now be maximized even with the prevailing computational complexity. Deep learning models, particularly convolutional neural networks (CNNs), have been shown to be effective in medical imaging research for organ segmentation and disease identification in recent years [5].

CNN is a great deep learning algorithm for classifying MRI images. Lin et al. [15] used CNN and freesurfer to extract the feature. The extracted feature then becomes an input for the extreme learning model. Although it has implemented CNN for feature extraction, this method cannot automatically classify the alzheimer's disease neuroimaging initiative (ADNI) dataset. Lee et al. [16] proposed deep CNN to classify AD using alexnet architecture using multiple slices selected based on the image entropy using a histogram. Kumar et al. also used entropy slicing to choose the most informative MRI slices during training using visual geometry group 16 (VGG16) [17]. However, the problem appeared when the slice was a noisy pixel, then the accuracy of AD classification decreased. Using the CNN method, pan et al. proposed eight layers of training, validation, and testing by taking a set of 3D images from an MR scan consisting of three sets of 2D image orientations: sagittal, coronal, or axial [18].

The evaluation to classify AD as in [1] [19] shows an extensive process in feature extraction techniques.

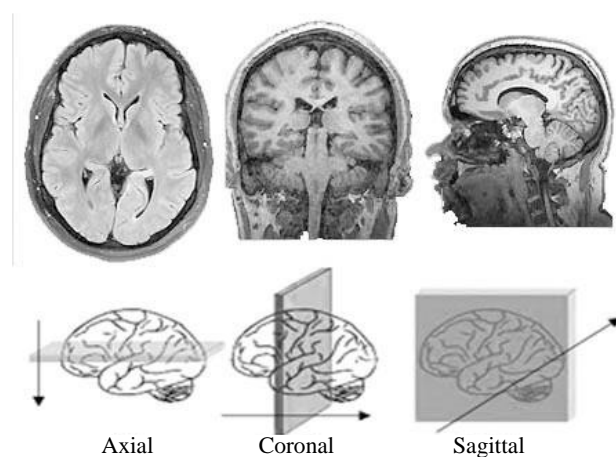


Figure 1. MRI anatomical plane

When using voxelwise feature vectors produced from noisy brains MRI for machine learning classification, adequate preprocessing is required to reduce feature selection faults. Compared to other region-of-interests (ROIs), such as ventricular volume, the cortical and hippocampal thickness contour caused by AD shows less variation when considered part of the overall brain structure.

Although it is possible to perform the AD classification method by segmenting brain images into white matter (WM), cerebrospinal fluid (CSF), and gray matter (GM) [20] [21], this method has the potential of increasing the computational process and the amount of data processed for each plane in the MRI. Furthermore, the quality of the chosen features during the feature extraction process is highly dependent on the image preprocessing due to registration errors and noise. Thus, this feature engineering stage requires domain expertise.

Recognition of 3D objects is challenging for many reasons, including the high cost of computational resources compared to 2D image analysis. As a simple comparison, 2D-CNN uses a 2D kernel, but 3D-CNN uses a 3D kernel, which results in more parameters. This approach increases the computational cost required by the CNN as the number of parameters evaluated increases. To alleviate the high computation cost for 3D images, we propose an AD classification approach that simultaneously uses multiple 2D images as input from the axial, coronal, and sagittal planes with a new approach named multiplane convolutional neural network (Mp-CNN).

The image of the human body's internal tissues and organs generated from an MRI scan is a collection of 2-dimensional sliced images that form a 3-dimensional object of the scanned organ. Since the disease is in the brain tissue, and the study does not focus on the brain's shape, the shape-based object recognition method is unsuitable for our study.

Fig. 1 shows a 3D image from the MRI that can be reconstructed in three depth directions (direction of the arrows are the plane direction of the MRI image), i.e., axial, coronal, and sagittal planes. These three image acquisition planes are used as inputs for the AD detection system proposed in this study. Furthermore, by analyzing the images as a whole plane, it is expected that all information from the object of MRI images can be more detailed so that better detection accuracy is obtained.

The contribution of this study is as follows:

- i. We propose a new network architecture design based on sequential CNN, called multiplane convolutional neural network

(Mp-CNN), that enables simultaneous processing of three MRI input planes.

- ii. Propose removing non-brain tissue before classification to prevent non-brain areas that are irrelevant to the disease from being included in the classification process.
- iii. Propose a method of selecting significant MRI slices for the system classification input. Thus, it is unnecessary to process the entire slice for each patient to provide computational simplicity.
- iv. Develop a multiclass AD classification system based on deep learning techniques on AD to help minimize visual inspection and manual assessment by medical staff so that patient care can be carried out earlier.

The remainder of this study is organized as follows: first, we discuss AD, previous studies, and related work. Next, we discuss the employed dataset description and preprocessing. Further, we present our proposed method to answer the existing problem. Then, we present the results of the experiments and hold an open discussion. Finally, we provide our findings as a conclusion in the last section.

2. Dataset and their preparation

We used images from ADNI to evaluate our proposed method [22]. The ADNI has well-organized and processed data freely available on the internet (<http://www.loni.ucla.edu/ADNI>). T1 and T2 scans from 1.5T and 3T MRI systems were used to create the database images. All participants with sufficient follow-up information were chosen from the ADNI-1 baseline magnetization prepared - rapid gradient echo (MP-RAGE) T1-weighted sequence at 1.5T tesla, usually 256 x 256 x 170 voxels with a voxel size of around 1mm, 1mm, 1.2mm. Our main goal is to categorize each person in the database as having AD, MCI, or normal cognition (NC). From primary T1-weighted MRI data, we randomly chose 449 subjects, including 500 AD, 500 MCI, and 500 NC images data.

A comprehensive psychiatric evaluation is required to diagnose AD, including patient history, a mini-mental state examination (MMSE), a clinical dementia level (CDR), and a physical and neurobiological assessment.

Table 1 shows the demographic distribution data of our ADNI collection, and more detailed information on the procedures for the acquisition of images can be found on the ADNI website. Next, by paying attention to the image size format, i.e., total slice, width, and height, all the data images were re-

Table 1. Demographic data of ADNI dataset

Data	AD	MCI	NC
Number of instances	345	605	991
Gender	186M/159F	301M/304F	646M/345F
MMSE +- SD	27.03+-2.60	21.88+-2.15	35.22+-1.24

sampled to 62 x 96 x 96 voxels in size 1 x 1 x 1mm³.

The axial, coronal, and sagittal planes are the three types of orientations an MRI scanner can generate. An X-Y plane parallel to the ground separates the head from the feet in the X-Y-Z coordinate system. A coronal is a front-to-back X-Z plane. A Y-Z plane that divides left and right is called a sagittal plane. The 3D Cartesian coordinate system (three principal planes) utilized for MRI scanning is depicted in Fig. 2. By evaluating all three MRI planes, we were able to extract more comprehensive information from the observed object. Therefore, we employed all three planes as input to our suggested system.

As seen in Fig. 3, AD patients' hippocampus and cerebral cortex shrink while the brain's ventricles grow. In addition, many studies have found that as the disease progresses, some parts of the brain, i.e., the hippocampus, amygdala, entorhinal, and parahippocampal cortices of the medial temporal lobe (MTL), shrink significantly [23]–[25]. These abnormalities in the cerebrum and hippocampus impact memory, planning, thinking, and judgment.

3. The proposed method

The proposed method includes the following steps: collecting data, transformation into three-plane, selecting only brain area using BET2 (brain extraction tool), selecting the three images with the largest area for each plane, and the final step is

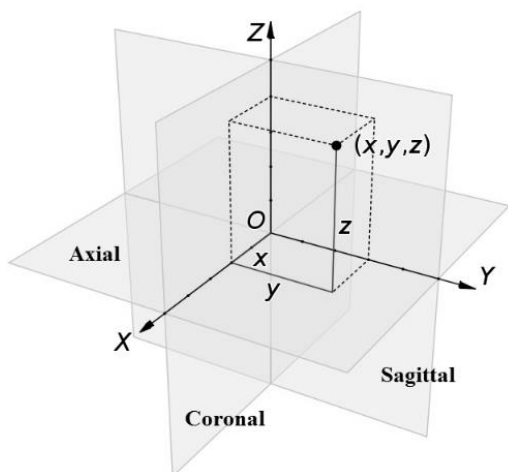


Figure. 2 The anatomical plane on MRI

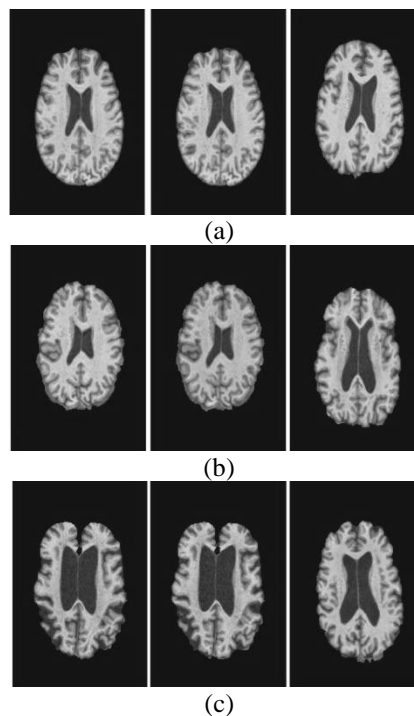


Figure. 3 The MRI's image samples; (a) normal, (b) MCI, (c) AD

multiplane 2DCNN with three outputs fully connected (FC) layers AD, MCI and, NC.

Our proposed approach begins with preprocessing the MRI images, as illustrated in Fig. 4. First, we have to exclude non-brain areas such as the skull and neck voxels from the MRI images as clinically evidence shows that the disease features reside in these areas. Since the abnormality affects the brain tissue only thus, we must concentrate solely on the brain tissue area. In this study, we employed functional magnetic resonance imaging of the brain (FMRIB) software library (FSL) to extract the brain area. The images were stripped using BET to remove non-brain tissue from the head's entire image [26] (<https://fsl.fmrib.ox.ac.uk/>).

The inner and outer skull surfaces and the outer scalp surface can also be estimated and affined to align all MR images to the template image using FMRIB software library (FSL) 5.0.

Spatial normalization is the next step in our preprocessing. Our dataset comprises high-resolution (T1-weighting (T1w) MRI images), but we need the same settings to examine data from various subjects.

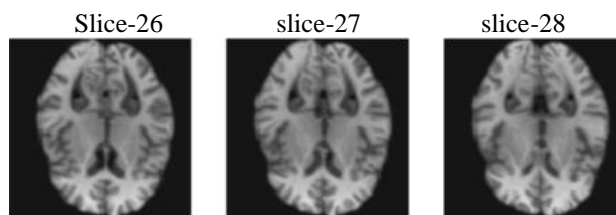


Figure. 4 Sample of the 3-top largest are of MRI image

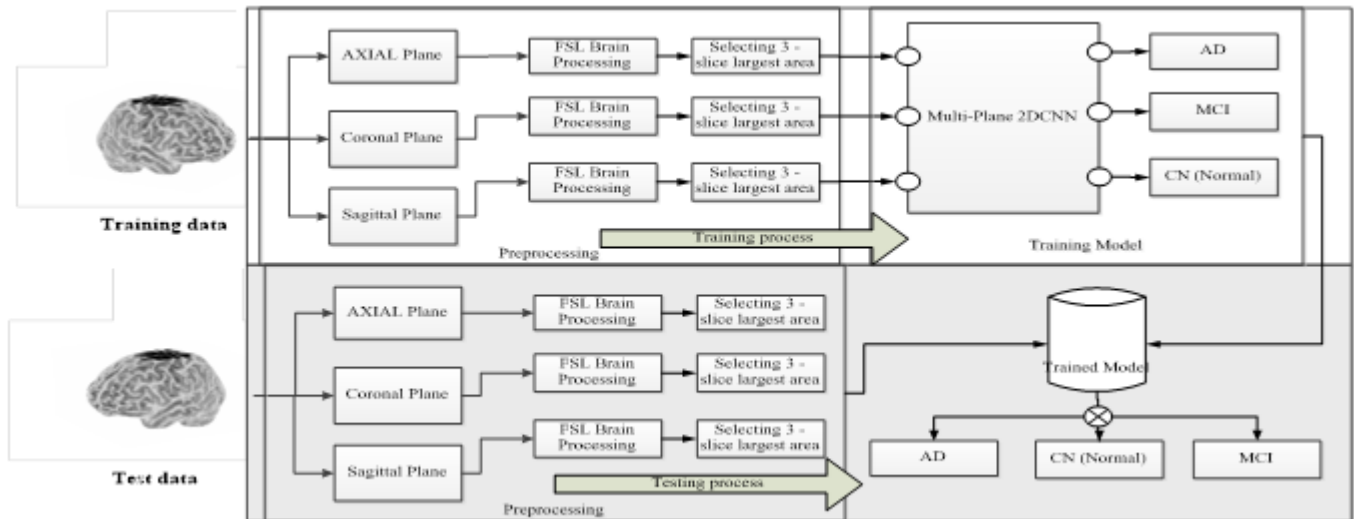


Figure. 5 Block diagram of the proposed method

We can only compare changes in specific voxels in specific locations for all patients with corresponding brains in the standard spatial domain. We utilized FMRIB’s linear image registration tool (FLIRT) (<https://fsl.fmrib.ox.ac.uk/fsl/fslwiki/FLIRT>) for spatial normalization. The images were registered with the montreal neurological institute 152 (MNI152) template data containing 152 registered standard images [27] using a linear transformation with 7 degrees of freedom (7 DOF). Translation, rotation, zoom, and sheer are also used to modify the shape and size of the object. As a result of these preprocessing steps, we obtained neuroimaging informatics technology initiative (NifTI) 4D image files with dimensions of 62 x 96 x 96 each.

3.1 MRI images selection

It is defined that a single MRI image is referred to as a slice. Since an MRI image comprises numerous slices, selecting just a few makes the computational process easier than processing the complete slice on a computer. We can determine the influenced slice of the MR image using the Eq. (1) by ascending the value of L , and then take one slice above it and one slice below it. Thus, we acquired three slices on each plane for each observed object.

$$L(x) = \frac{\sum p_i}{\sum_s^1 (\sum p_i)} \quad (1)$$

Where,

- L = Percentage of the image area on a spesific slice
- s = Total number of slices
- i = The slice number ($i = 1 \dots s$)
- p = The non-zero pixel

Algorithm	: Pseudocode for image slice selection
Input	: MRI brain images which have already excluded its non-brain areas
Output	: The three slices of brain-imaging with the largest area
<ol style="list-style-type: none"> 1. for (slice =1) to (slice \leftarrow endSlice): collectArea \leftarrow countNonZeroArea(slice) 2. TotalArea $\leftarrow \sum$ (collectArea) 3. for (slice =1) to (slice \leftarrow endSlice): percentageArea(slice) ... \leftarrow (collectArea(slice)/TotalArea) 4. LargestArea \leftarrow find(max(percentageArea(slice))) 5. Top3Slice \leftarrow [(SliceNumberLargestArea - 1), (SliceNumberLargestArea), (SliceNumberLargestArea + 1)] 6. Output \leftarrow MRI_Images(Top3Slice) 	

3.2 Convolutional neural network

In general, CNN architecture comprises three layers: convolutional, sub-sampling, and FC, each of which can be adjusted on many layers. The convolutional layer is the initial layer of a CNN, is prone to detecting local features taken from input image locations. Following that, those layers create a linking layer that translates inputted images using feature transformations. Unlike the feature engineering used in machine learning, CNN, as one-of the deep learning methods, uses supervised feature learning where the feature extraction process is done automatically and adaptively by the model.

Feature learning is used because feature engineering is very limited. After all, each case of different data requires a different feature extraction method. As a result, the feature engineering method cannot generalize to various types of data required in object classification in the image. Furthermore, in complex data processing, feature extraction is very

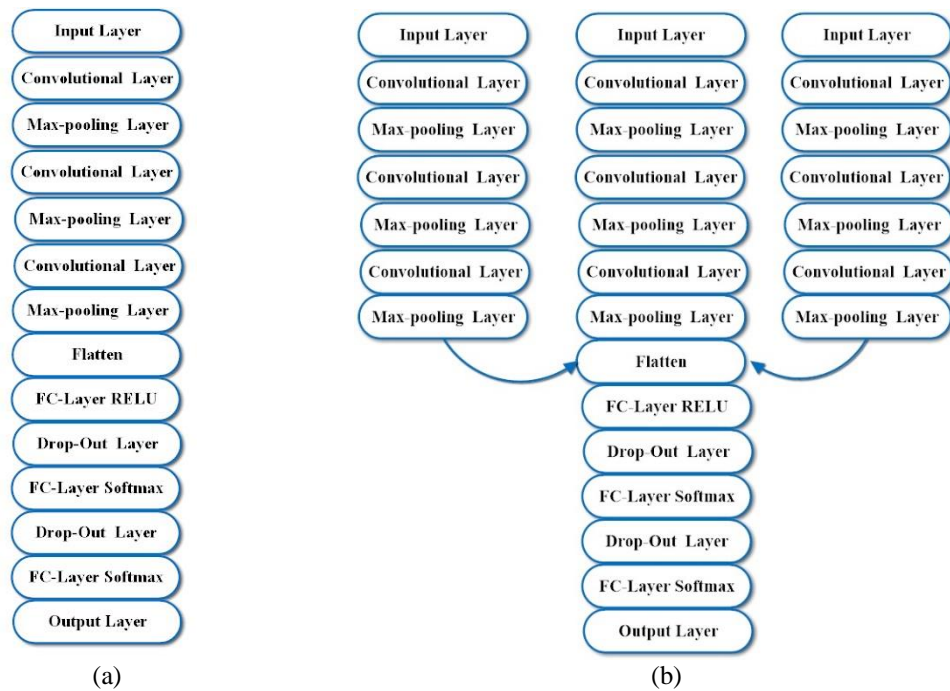


Figure. 6 The proposed architecture (a) for single plane MRI, (b) for multiplane CNN (Mp-CNN)

time-consuming and tends to be less able to describe the overall information value of the data. Feature learning overcomes this by creating an adaptive feature extraction process that can automatically adjust the data used. As a result, CNN is noted for its advantages in overcoming minor data variations, requiring minimal preprocessing, and requiring no manual feature extraction outside the CNN system.

3.3 Multiplane convolutional neural network (Mp-CNN)

We propose a novel architecture based on a sequential convolutional neural network layer to characterize all the spatial structure information simultaneously. In the field of image recognition, this sequential convolutional neural network layer has had much success. LeNet-5, AlexNet, and VGG16 are three successful examples of CNN architectures for this type.

Since each MR image has three standard planes, axial, coronal, and sagittal, it is important to process all of the image data at the same time. Our architecture addresses some of the incompleteness of earlier research by forcing all accessible visual data to be processed simultaneously, whereas existing studies have focused on only one plane. Our proposed CNN network for a single plane is shown in Fig. 6. Each section of our solution has 13 layers, resulting in a final architecture consisting of three networks for three planes. This network’s input is an MR-images that has been preprocessed as described previously.

Our architecture is based on the common

architecture in a convolution neural network which consists of a layered stack of convolution layers. This architecture model is also known as a sequential convolutional neural network layer, with well-known architectural examples such as LeNet-5, AlexNet, and VGG-net. The number of filters arranged increases over time; namely, the number of filters in the previous layer must be less than the next layer. Additionally, all of the filters are 3 x 3 in size.

The detail of the proposed network consists of fourteen layers, including input and output layers. The first layer is the input layer, where the input image is a hyperplane image of multimodal axial, coronal and sagittal. Furthermore, there are three layers of convolutional layers followed by the max pooling layer.

The activation function helps in solving complex problems by applying non-linear input transformations. In the proposed framework, the activation function rectified linear unit (ReLU) was examined. This ReLU approach is more comparable to how biological neurons perform. ReLU was chosen because it is a non-linear function with no backpropagation errors, unlike the sigmoid function. Additionally, for bigger neural networks, the speed of constructing models using ReLU is much faster than using sigmoids. As shown in Eq. 2, if the input is less than 0, the ReLU will produce a 0 output; if the input is more than 0, the output will be the same as the input.

$$ReLU, f(x) = \max(x, 0) \tag{2}$$

The softmax activation function transforms a list of numbers into a list of probabilities, with each value's probability proportional to the vector's relative scale. The softmax not only maps our output to a [0, 1] range, but it also maps each output so that the entire sum equals 1, making it suitable for predicting a multinomial probability distribution. Thus, softmax is excellent to be used as an activation function for problems requiring class membership on more than two class labels. Since our architecture is built for multiclass classification, we use softmax as an activation function in the output layer.

$$\text{Softmax, } f(x)_j = \frac{e^{x_j}}{\sum_1^k e^{x_k}} \tag{3}$$

Where,

- x = a vector of inputs to the output layer
- j = a number between 1 and k
- k = represents the number of output class

The activation function, learning rate, loss function, optimization function, and sample size all have an impact on the model. Therefore, adjustments of these parameters are used in the experiments. A flatten layer is the eighth layer. It converts the feature map into a 1-D vector that should be utilized on the FC Layer. Next, the training process begins in order to obtain the best CNN model. Finally, an optimization function should be applied to minimize the loss function and obtain an accurate model.

In neural networks, the loss function plays an important part in the training procedure. It is used to determine the discrepancy between the expected and actual results. The challenge is to make the loss function's value smaller. The cross-entropy loss function was chosen in the proposed framework since it performed better than mean squared error (MSE). Moreover, since the output is a probability value between 0 and 1, the cross-entropy loss is performed to analyze the classification model's performance. As a result, we should use this loss function for our

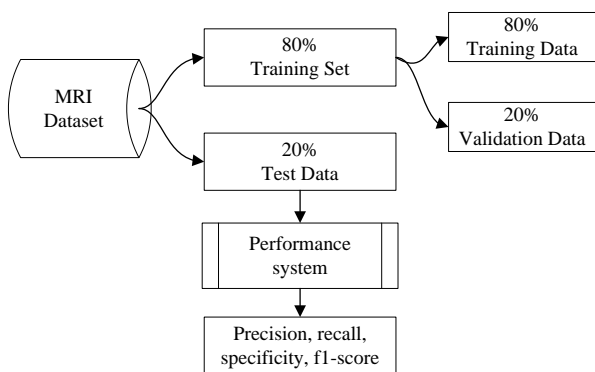


Figure. 7 The dataset distribution setting

classifications system. In regression neural networks, on the other hand, MSE can be implemented.

4. Experiment and result

This section contains comprehensive information about model training and parameter setting during training as well as the results of our experiments. We employed 5-fold cross-validation to evaluate the network's performance, which divided the MRI dataset into training and testing sets in a 5-fold manner. The training and testing sets were split 80 % : 20 %, as shown in Fig. 7. Each of the four partitions was utilized for training the model, with the remaining serving as test data. As a result, the model was trained five times, and the testing accuracy for each training was calculated. The average performance was then calculated to obtain a five times cross-validation accuracy.

Table 2 shows the image distribution of the dataset we used, the same as the number of processed 2D images for a single-plane CNN. The total number of 2D MRI images used by Mp-CNN, on the other hand, is substantially larger. For example, the number of images processed by Mp-CNN first, second, and third-largest slices is 3x more than the CNN-single slice, while the number of images processed by Mp-CNN (largest slice, +1,-1) is 9x more than CNN-single slice. This is because CNN (largest slice, +1,-1) uses MRI as input by taking an image from the largest slice, one slice above it, and one slice below it.

The preprocessed images have final axial dimensions of 62 x 96 x 96, with 62 denoting the number of image slices used as input data. The input images are convolutionally calculated and downsampled three times, with the number of convolutional kernels (3 3) in each convolutional layer 16, 32, and 64, respectively, stride 1 and padding 1. The maximum pooling approach with a size of 2 is used in the down-sampling layer. The result is the maximum value of neurons for each region, depending on the results of convolution calculations. The last Conv-maxpooling network's process is continued by transforming the output to a long-vector (flatten) and feeding it into the FC. Altogether, we utilized three layers of FC layers. The first FC layer has 4096 neurons, whereas the last FC

Table 2. Dataset distribution for training and testing

Type	Number of MRI images	Training	Testing
NC	500	394	106
MCI	500	403	97
AD	500	403	97
Total	1500	1200	300

layer has three class categories that indicate NC, MCI, and AD. The dropout value utilized is 0.5, which successfully prevents the overfitting phenomena caused by the network’s depth. Dropout is a technique that successfully terminates the output of some neurons based on a set of probabilities.

We evaluated the accuracy and other model performance metrics, including precision, recall or sensitivity, and f1-score. The true positive, true-negative, false-positive, and false-negative values must first be determined in order to assess the model’s performance. True positive and negative values relate to instances where the model properly identified positive and negative predictions. On the other hand, false positives and negatives are instances of the model classifying negative and positive predictions into positive and negative classifications.

$$Accuracy = \frac{True\ Negative + True\ Positive}{Total\ Data} \quad (4)$$

$$Precision = \frac{True\ Positive}{True\ Positive + False\ Positive} \quad (5)$$

$$Recall\ (Sensitivity) = \frac{True\ Positive}{True\ Positive + False\ Negative} \quad (6)$$

$$Specificity = \frac{True\ Negative}{True\ Negative + False\ Positive} \quad (7)$$

$$F1 = \frac{2x\ Precision\ x\ recall}{Precision + recall} \quad (8)$$

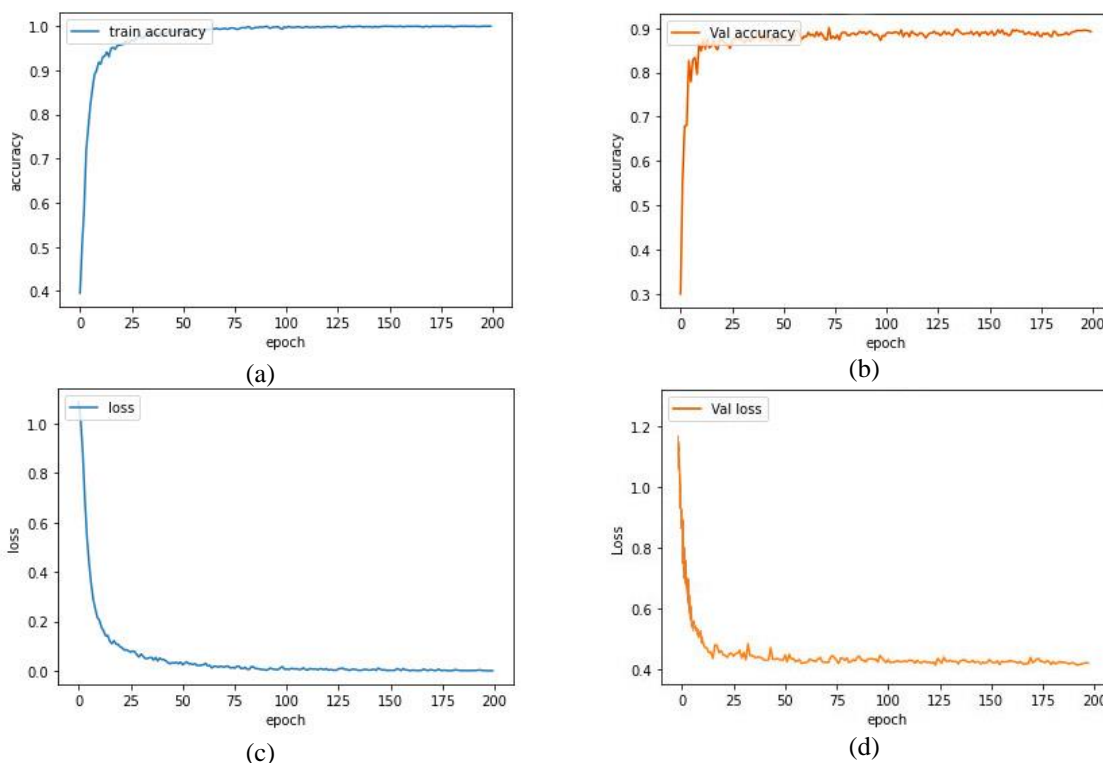


Figure. 8 The curves of (a) training model accuracy, (b) validation accuracy, (c) training model loss, and (d) validation loss

Table 3. The accuracy and precision

Model		Accuracy	Precision		
			NC	MCI	AD
CNN (Single-plane)	Axial	0.86	0.92	0.86	0.80
	Coronal	0.85	0.91	0.81	0.81
	Sagittal	0.85	0.89	0.83	0.83
Mp-CNN	The largest slice	0.92	0.98	0.89	0.89
	2nd largest slice	0.92	0.99	0.87	0.90
	3rd largest slice	0.91	0.97	0.88	0.87
Mp-CNN	Largest slice, +1,-1	0.93	0.95	0.91	0.93
TL-Resnet50	The largest slice	0.88	0.92	0.81	0.90
	Combine largest slice, +1,-1	0.89	0.95	0.81	0.92

Table 4. Recall and F1 score

Model		Recall			F1-score		
		NC	MCI	AD	NC	MCI	AD
CNN (Single-plane)	Axial	0.97	0.75	0.85	0.94	0.80	0.82
	Coronal	0.95	0.79	0.78	0.93	0.80	0.80
	Sagittal	0.92	0.75	0.89	0.90	0.79	0.86
Mp-CNN	The largest slice	0.98	0.89	0.89	0.98	0.89	0.89
	2 nd largest slice	0.98	0.89	0.89	0.99	0.88	0.89
	3 rd largest slice	0.95	0.84	0.93	0.96	0.86	0.90
Mp-CNN	Largest slice, +1,-1	0.98	0.89	0.92	0.97	0.90	0.92
TL-Resnet50	The largest slice	0.94	0.83	0.85	0.93	0.82	0.87
	Combine largest slice, +1,-1	0.95	0.88	0.95	0.95	0.84	0.88

Fig. 4(a) and 4(b) show that training and validation accuracy increased significantly until the 20th iteration, then reached saturation above 125, with the final value remaining stable until the 200th iteration. Finally, when the iterations exceed 175, as illustrated in Fig. 4(c) and 4(d), the loss in training is saturated. Hence, 200 iterations are sufficient to represent the model’s performance results.

In comparison to the results of other experiments, the proposed method consistently achieves the highest accuracy, precision, recall (sensitivity), and f1-score. We can observe from the experimental data in Tables 3 and 4 that the proposed model outperforms previous experiments that were also conducted. The proposed model outperforms single-plane and the Resnet50 network-based transfer learning method. All studies demonstrate that employing multiplane can increase model performance, and using more data also offers better results, as seen when Mp-CNN only utilizes the three largest slices, which can be beaten by the performance of Mp-CNN when using nine slices, each consisting of the largest slices and two slices around it

Discussion

Many studies have been conducted to aid clinicians in defining the characteristics of AD. In

addition, a digital image-based study using MRI to diagnose AD using several advanced machine learning methodologies and data modalities has also been proposed. Table 5 compares the proposed approach to previous studies, where the proposed method is more precise since it has higher accuracy than the others. These results indicate that, in general, CNN is superior to SVM (support vector machine) and LR (linear regression) for classification.

This section explores previous studies on AD classifications and discusses our proposed method, which is more effective for distinguishing AD, MCI, and NC. Pan et al. [18] proposed a classification method using ensemble learning CNN, using the MRI slice selection method based on reslicing from 3D images to three 2D images. The preprocessing detail of this method is complex with a lengthy process, so we propose a method for selecting the three slices, i.e., the largest slice, one slice above, and below the largest slice so that the preprocessing computation could be minimized. Altaf et al.[1], Zhang et al. [28] and Minhas et al. [29] proposed a machine learning method using SVM or LR as a classifier. Lu et al. [30] used multiscale deep neural network (MDNN) as an AD classifier, taking the multiscale patch-wise metabolism features as input. In their study, they used hand-crafted features.

In contrast, the classification approach followed in

Table 5. The proposed method and prior comparable studies

Study	Dataset	Modality	Classification Type	Algorithm	ACC	SEN	SPE
[1]	ADNI	MRI + clinical	AD Vs MCI Vs NC	SVM	0.61	0.57	0.89
[28]	ADNI	ROI MRI	MCI Vs NC	SVM	0.76	0.86	0.60
[29]	ADNI	MRI + clinical	sMCI Vs pMCI	LR	0.90	0.88	0.92
[30]	ADNI	FDG-PET	sMCI Vs pMCI	MDNN	0.83	0.81	0.83
[31]	ADNI	MRI	AD Vs NC	CNN	0.90	0.01	0.01
[32]	ADNI	MRI	AD Vs NC	CNN	0.87	0.89	0.85
Our work	ADNI	MRI	AD Vs MCI Vs NC	Mp-CNN	0.93	0.92	0.93

FDG-PET: Fludeoxyglucose positron emission tomography, progressive mild cognitive impairment (pMCI), stable mild cognitive impairment (sMCI), ACC: Accuracy, SEN: Sensitivity, SPE: Specificity.

our study is a deep learning method in which mp-CNN extracts features automatically. The results of the experiments indicate that the proposed mp-CNN-based architecture can support achieving satisfactory diagnostic decisions.

Furthermore, Lian et al. [31] proposed hierarchical fully convolutional network based on a 3D patch, while Oh et al. [32] proposed a 3DCNN method to classify AD. However, both used a 3D model, which results in high computational costs. Therefore, we propose a 2D approach to reduce the computational costs that may occur during the classification process. Overall, the results of our experiments indicate that the proposed Mp-CNN-based architecture can effectively support the AD classification.

The scientific contribution of this study will be explained in complete detail below to help describe our proposed approach. We proposed a multiclass AD classification system based on deep learning techniques to aid medical professionals in minimizing visual inspection and manual assessment, allowing them to begin patient care earlier. Our architecture allows us to process axial, coronal, and sagittal images at the same time. Under the name Mp-CNN, we present a new network architecture based on a sequential type of network. We use all three MRI planes as input to the AD recognition system to better understand the object instead of just one. Non-brain tissue was removed before the classification process to ensure that non-brain areas irrelevant to AD disease were not included in the classification process. This approach is needed for the recognition system to focus on disease and avoid being distracted by non-brain objects' features. Finally, we present an MRI slice selection approach for our system's classification input. As a result, each MRI's slice does not have to be processed as its whole, benefitting in computing efficiency.

Conclusions

This article presents the multiplane convolutional neural network (Mp-CNN), a new deep learning network architecture for AD classification that allows simultaneous processing of three MRI planes: sagittal, coronal, and transverse (axial). Mp-CNN was applied to classify patients with Alzheimer's disease, mild cognitive impairment, or normal cognition.

The following are the three main conclusions to be derived from our study. First, the brain damage associated with AD is not only noticeable in one plane, such as axial, coronal, or sagittal, but also all three. The result shows that multiplane recognition system outcomes extend beyond single planes.

Second, the results show that the performance of the Mp-CNN delivers good results. Our proposed method has a 93 % accuracy for AD-MCI-NC classification and precision of AD 93 %, 91 % for MCI, and 95 % for NC. Lastly, our study also proves that the proposed Mp-CNN method outperforms other alternative methods we investigated.

Conflicts of interest

The authors declare no conflict of interest.

Author contributions

Conceptualization, I Ketut Eddy Purnama, Mauridhi H. Purnomo; methodology, Cucun Very Angkoso; software, Cucun Very Angkoso; validation, I Ketut Eddy Purnama; formal analysis, Cucun Very Angkoso, I Ketut Eddy Purnama and Mauridhi H. Purnomo; investigation, Cucun Very Angkoso; resources, Cucun Very Angkoso; data curation, Hapsari Peni Agustin Tjahyaningtjas; writing original draft preparation, Cucun Very Angkoso, I Ketut Eddy Purnama, and Hapsari Peni Agustin Tjahyaningtjas; review and editing, Mauridhi H. Purnomo and I Ketut Eddy Purnama; visualization, Cucun Very Angkoso; supervision, I Ketut Eddy Purnama, Mauridhi H. Purnomo; project administration, Cucun Very Angkoso; funding acquisition, I Ketut Eddy Purnama.

Acknowledgments

Our study was supported by BPPDN scholarship and Enhancing International Publication Program (EIP) from the Directorate General of Higher Education of Ministry of Education and Culture - Republic of Indonesia, and University Center of Excellence on Artificial Intelligence for Healthcare and Society (UCE AIHeS). In addition, this study was partially funded by the Education Fund Management Institute (LPDP) under the Innovative Productive Research Grant (RISPRO) scheme - Invitation 2019, contract number: PRJ-41/LPDP/2019. Finally, we would like to extend our gratitude to ADNI for supplying the dataset for the study.

References

- [1] T. Altaf, S. Muhammad, N. Gul, M. Nadeem, and M. Majid, "Multi-class Alzheimer's disease classification using image and clinical features", *Biomed. Signal Process. Control*, Vol. 43, pp. 64–74, 2018.
- [2] A. Association, "2016 Alzheimer's disease facts and figures", *Alzheimer's Dement.*, Vol. 12, No. 4, pp. 459–509, 2016.

- [3] S. Belleville, C. Fouquet, S. Duchesne, D. L. Collins, and C. Hudon, "Detecting early preclinical alzheimer's disease via cognition, neuropsychiatry, and neuroimaging: Qualitative review and recommendations for testing", *J. Alzheimer's Dis.*, Vol. 42, pp. S375–S382, 2014.
- [4] A. Association, "2018 Alzheimer's disease facts and figures", *Alzheimer's Dement.*, Vol. 14, No. 3, pp. 367–429, 2018.
- [5] G. Litjens, T. Kooi, B. E. Bejnordi, A. A. A. Setio, F. Ciompi, M. Ghafoorian, J. A. W. M. V. D. Laak, V. Ginneken, and C. I. Sánchez, "A survey on deep learning in medical image analysis", *Med. Image Anal.*, Vol. 42, pp. 60–88, 2017.
- [6] Y. Zhang, S. Wang, P. Phillips, Z. Dong, and G. Ji, "Detection of Alzheimer's disease and mild cognitive impairment based on structural volumetric MR images using 3D-DWT and WTA-KSVM trained by PSOTVAC", *Biomed. Signal Process. Control*, Vol. 21, pp. 58–73, 2015.
- [7] S. Klöppel, "Accuracy of dementia diagnosis—A direct comparison between radiologists and a computerized method", *Brain*, Vol. 131, No. 11, pp. 2969–2974, 2008.
- [8] F. Falahati, K. Institutet, E. Westman, K. Institutet, and A. Simmons, "Multivariate Data Analysis and Machine Learning in Alzheimer's Disease with a Focus on Structural Magnetic Resonance Imaging", *Journal of Alzheimer's Disease*, Vol. 41, No. 3. pp. 685–708, 2014
- [9] S. Leandrou, S. Petroudi, P. A. Kyriacou, C. C. R. Aldasoro, and C. S. Pattichis, "Quantitative MRI Brain Studies in Mild Cognitive Impairment and Alzheimer's Disease: A Methodological Review.", *IEEE Rev. Biomed. Eng.*, Vol. 11, pp. 97–111, 2018.
- [10] A. Khan and M. Usman, "Early diagnosis of Alzheimer's disease using machine learning techniques: A review paper", *2015 7th Int. Jt. Conf. Knowl. Discov. Knowl. Eng. Knowl. Manag.*, Vol. 1, pp. 380–387, 2015.
- [11] C. Zheng, Y. Xia, Y. Pan, and J. Chen, "Automated identification of dementia using medical imaging: a survey from a pattern classification perspective", *Brain informatics*, Vol. 3, No. 1, pp. 17–27, Mar. 2016.
- [12] R. Cuingnet, E. Gerardin, J. Tessieras, G. Auzias, S. Lehéricy, M. O. Habert, M. Chupin, H. Benali, and O. Colliot, "Automatic classification of patients with Alzheimer's disease from structural MRI: a comparison of ten methods using the ADNI database", *Neuroimage*, Vol. 56, No. 2, pp. 766–781, May 2011.
- [13] M. I. Razzak, S. Naz, and A. Zaib, "Deep learning for medical image processing: Overview, challenges and the future", *Lecture Notes in Computational Vision and Biomechanics*, Vol. 26, pp. 323–350, 2018.
- [14] N. Amoroso, M. L. Rocca, L. Bellantuono, D. Diacono, A. Fanizzi, E. Lella, A. Lombardi, and T. Maggipinto, "Deep Learning and Multiplex Networks for Accurate Modeling of Brain Age", *Front. Aging Neurosci.*, Vol. 11, No. May, 2019.
- [15] W. Lin, T. Tong, Q. Gao, D. Guo, X. Du, Y. Yang, G. Guo, M. Xiao, M. Du, and X. Qu, "Convolutional neural networks-based MRI image analysis for the Alzheimer's disease prediction from mild cognitive impairment", *Front. Neurosci.*, Vol. 12, No. Nov., pp. 1–13, 2018.
- [16] B. Lee, W. Ellahi, and J. Y. Choi, "Using deep CNN with data permutation scheme for classification of Alzheimer's disease in structural magnetic resonance imaging (SMRI)", *IEICE Trans. Inf. Syst.*, Vol. E102D, No. 7, pp. 1384–1395, 2019.
- [17] S. S. Kumar and M. Nandhini, "Entropy Slicing Extraction and Transfer Learning Classification for Early Diagnosis of Alzheimer Diseases with sMRI", *ACM Trans. Multimed. Comput. Commun. Appl.*, Vol. 17, No. 2, 2021.
- [18] D. Pan, A. Zeng, L. Jia, Y. Huang, T. Frizzell, and X. Song, "Early Detection of Alzheimer's Disease Using Magnetic Resonance Imaging: A Novel Approach Combining Convolutional Neural Networks and Ensemble Learning", *Front. Neurosci.*, Vol. 14, No. May, pp. 1–19, 2020.
- [19] Y. Zhang, S. Wang, K. Xia, Y. Jiang, and P. Qian, "Alzheimer's disease multiclass diagnosis via multimodal neuroimaging embedding feature selection and fusion", *Inf. Fusion*, Vol. 66, No. September 2020, pp. 170–183, 2021.
- [20] C. V. Angkoso, I. K. E. Purnama, and M. H. Purnomo, "Analysis of Brain Tissue and Cerebrospinal Fluid Feature for Alzheimer's Disease Detection", *2018 Int. Conf. Comput. Eng. Netw. Intell. Multimedia, CENIM 2018*, pp. 285–288, 2018.
- [21] H. S. Suresha and S. S. Parthasarathy, "Diagnosis of Alzheimer disease using fast independent component analysis and Otsu multi-level thresholding", *International Journal of Intelligent Engineering and Systems*, Vol. 11, No. 5, pp. 74–83, 2018.
- [22] M. W. Weiner, D. P. Veitch, P. S. Aisen, L. A. Beckett, N. J. Cairns, J. Cedarbaum, R. C. Green, D. Harvey, C. R. Jack, W. Jagust, J. Luthman, J.

- C. Morris, R. C. Petersen, A. J. Saykin, L. Shaw, L. Shen, A. Schwarz, A. W. Toga, and J. Q. Trojanowski, “2014 Update of the Alzheimer’s Disease Neuroimaging Initiative: A review of papers published since its inception”, *Alzheimer’s Dement.*, Vol. 11, No. 6, pp. e1–e120, 2015.
- [23] B. C. Dickerson, I. Goncharova, M. P. Sullivan, C. Forchetti, R. S. Wilson, D. A. Bennett, L. A. Beckett, and L. D. Morrell, “MRI-derived entorhinal and hippocampal atrophy in incipient and very mild Alzheimer’s disease”, *Neurobiol. Aging*, Vol. 22, No. 5, pp. 747–754, 2001.
- [24] Y. K. Koerkamp, R. A. Heckemann, K. T. Ramdeen, O. Moreaud, S. Keignart, A. Krainik, A. Hammers, M. Baciú, P. Hot, and A. D. N. Initiative, “Amygdalar Atrophy in Early Alzheimer’s Disease”, *Current Alzheimer Research*, Vol. 11, No. 3, pp. 239–252, 2014.
- [25] R. A. Sperling, P. S. Aisen, L. A. Beckett, D. A. Bennett, S. Craft, A. M. Fagan, T. Iwatsubo, C. R. Jack, J. Kaye, T. J. Montine, D. C. Park, E. M. Reiman, C. C. Rowe, E. Siemers, Y. Stern, K. Yaffe, M. C. Carrillo, B. Thies, M. M. Bogorad, M. V. Wagster, and C. H. Phelps, “Toward defining the preclinical stages of Alzheimer’s disease: Recommendations from the National Institute on Aging-Alzheimer’s Association workgroups on diagnostic guidelines for Alzheimer’s disease”, *Alzheimer’s Dement.*, Vol. 7, No. 3, pp. 280–292, 2011.
- [26] S. M. Smith, “Fast robust automated brain extraction”, *Human Brain Mapping.*, Vol. 17, No. 3, pp. 143–155, 2002.
- [27] S. Sarraf and J. Sun, “Functional Brain Imaging: A Comprehensive Survey”, *arXiv preprint arXiv:1602.02225*, 2016.
- [28] J. Zhang, M. Liu, L. An, Y. Gao, and D. Shen, “Alzheimer’s disease diagnosis using landmark-based features from longitudinal structural MR images”, *IEEE J. Biomed. Heal. Informatics*, Vol. 21, No. 6, pp. 1607–1616, 2017.
- [29] S. Minhas, A. Khanum, F. Riaz, A. Alvi, and S. A. Khan, “A Nonparametric Approach for Mild Cognitive Impairment to AD Conversion Prediction: Results on Longitudinal Data”, *IEEE J. Biomed. Heal. Informatics*, Vol. 21, No. 5, pp. 1403–1410, 2017.
- [30] D. Lu, K. Popuri, G. W. Ding, and R. Balachandar, “Multiscale deep neural network based analysis of FDG-PET images for the early diagnosis of Alzheimer’s disease”, *Med. Image Anal.*, Vol. 46, pp. 26–34, 2018.
- [31] C. Lian, M. Liu, J. Zhang, and D. Shen, “Hierarchical Fully Convolutional Network for Joint Atrophy Localization and Alzheimer’s Disease Diagnosis Using Structural MRI”, *IEEE Trans. Pattern Anal. Mach. Intell.*, Vol. 42, No. 4, pp. 880–893, Apr. 2020.
- [32] K. Oh, Y. C. Chung, K. W. Kim, W. S. Kim, and I. S. Oh, “Classification and Visualization of Alzheimer’s Disease using Volumetric Convolutional Neural Network and Transfer Learning”, *Sci. Reports 2019 91*, Vol. 9, No. 1, pp. 1–16, Dec. 2019.