Automatic Localization and Discrete Volume Measurements of Hippocampi From MRI Data Using a Convolutional Neural Network

ABOL BASHER [ID]1, **BYEONG C. KIM**2,3, **KUN HO LEE** [ID]2,4, **AND HO YUB JUNG** [ID]1

1Department of Computer Engineering, Chosun University, Gwangju 61452, South Korea
2National Research Center for Dementia, Chosun University, Gwangju 61452, South Korea
3Department of Neurology, Chonnam National University Medical School, Gwangju 61469, South Korea
4Department of Biomedical Science, Chosun University, Gwangju 61452, South Korea

Corresponding author: Ho Yub Jung (hoyub@chosun.ac.kr)

ABSTRACT Automatic hippocampal volume measurement from brain magnetic resonance imaging (MRI) is a crucial task and an important research area, especially in the study of neurodegenerative diseases; hippocampal volume atrophy is known to be connected with Alzheimer's disease. In this research work, we propose a deep learning-based method to automatically measure the discrete hippocampal volume without prior segmentation of the volumetric MRI scans. We constructed a 2-D convolutional neural network (CNN) model that uses 3-channel 2-D patches to predict the number of voxels attributed to the hippocampus; the number of estimated hippocampal voxels is multiplied by the voxel volume to measure the discrete volume of the hippocampus. In addition, we demonstrate a preprocessing scheme to prepare the data using a relatively small number of MRI scans. The average errors in the measured volumes of the proposed approach and the compared atlas-based system were 4.3173 ± 3.5436 (avg. error% $\pm$ STD) and 4.1562 ± 3.5262 (avg. error % $\pm$ STD) for the left and right hippocampi, respectively. The correlation coefficients of the proposed approach with atlas-based volume measurement were statistically significant (p-value < 0.01, $R^2 = 0.834$ (left hippocampus), and $R^2 = 0.848$ (right hippocampus) based on 0.05 significance level), which suggests that the proposed approach can be used as a proxy method for the atlas-based system. Furthermore, the proposed approach is computationally efficient and requires less than 2 seconds to calculate the number of voxels for an MRI scan. Moreover, our method outperforms the state-of-the-art deep learning approach, such as 2-D U-Net and SegNet in the context of voxel/volume estimation errors% for the left and right hippocampi.

INDEX TERMS MRI, hippocampus, patch, Hough-CNN, localization, CNN, discrete volume.

I. INTRODUCTION

The hippocampus is a widely studied structure in the context of learning, memory, stress and neurological disorders. Hippocampal atrophy is known to be linked to various serious brain dysfunctions, such as Alzheimer's disease [1], [2], schizophrenia [3], and depression [1]. Hippocampal volume measurement is a crucial task to perform automatically, and it is relatively time consuming.

The hippocampus consists of distinct, interacting subregions with a complex, heterogeneous structure [4]. Therefore, it is highly difficult to perform critical analysis on its

The associate editor coordinating the review of this manuscript and approving it for publication was Wei Wei [ID].

subregions and measure the exact volume. Because of the increasing resolution of MRI scans, automatic segmentation of the hippocampus with its subfields becomes possible. To segment and measure the volume of hippocampi with its subfields, several methods [5]–[9] have been developed. Moreover, a few software packages, such as, FreeSurfer, FIRST, SPM, and Neuro I, and online platforms, such as VolBrain (https://volbrain.upv.es/) are available to estimate the volume from MRI scans. The gold standard method to measure hippocampal volume is the manual delineation of brain MRI scans. However, this method requires careful work by trained operators and is often impractically labor intensive. Therefore, considerable attention from the research communities is being paid for developing an automatic system to
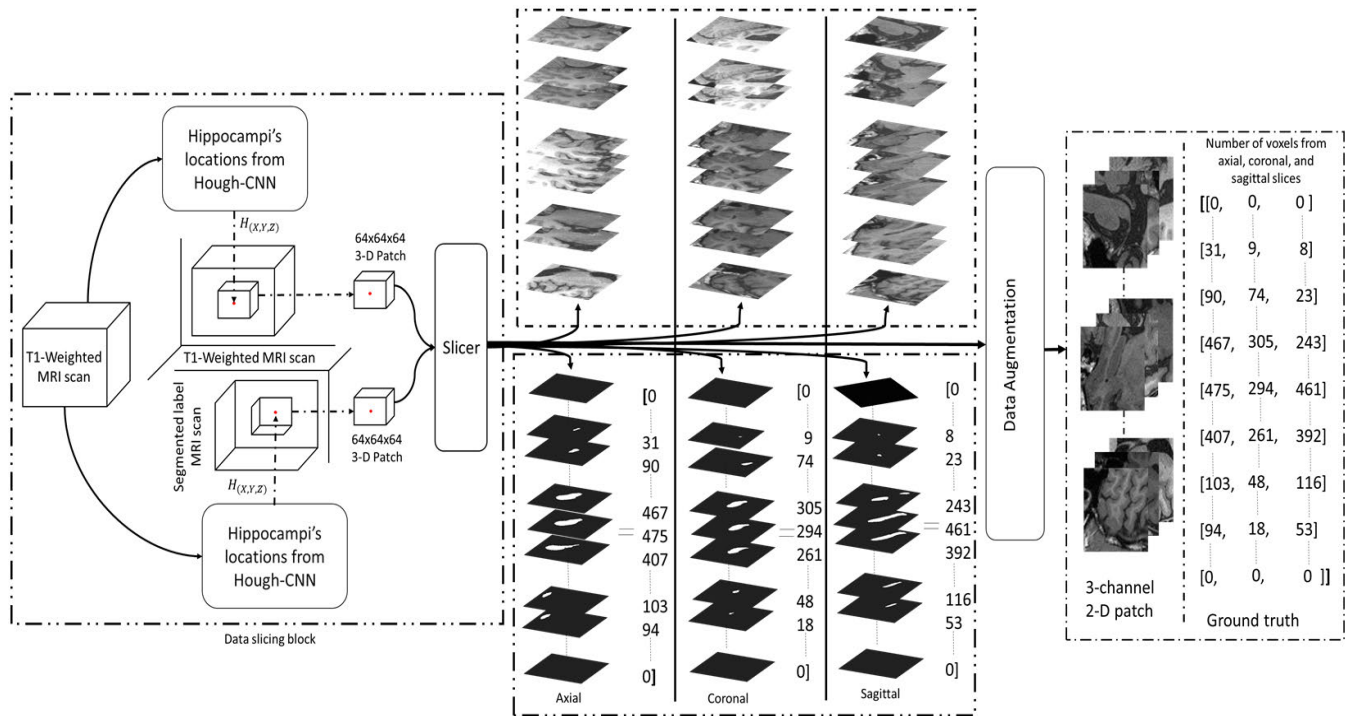
FIGURE 1. This figure describes data generation process for the proposed method. Three-channel 2-D patches with their corresponding segmented labels are generated for the volume measurement. The left and right hippocampal voxels' locations from T1-weighted MRI scans were estimated using two-stage ensemble Hough-CNN. The estimated voxel location is used to extract 3-D patches with the size of 64 × 64 × 64. The extracted 3-D patches were then fed to the slicer to separate the slices for the axial, coronal and sagittal views. The axial, coronal and sagittal slices were normalized separately, rotated by Θ-degree with a factor of n-times (in our case, n=10) and reshape into a size of 32 × 32 × 1. Similarly, from segmented label MRI scans, using the same voxel location and corresponding 3-D patches (size: 64 × 64 × 64) were generated. Slicer counts the number of pixels/voxels attributed in each slice for the axial, coronal and sagittal views. The 2-D slices from T1-weighted MRI scans were concatenated (along axis=2 (0-based axis)) to construct 3-channel 2-D patches with the size of 32 × 32 × 3. The 2-D patches with their corresponding labels (number of voxels/pixels) were fed to train the proposed CNN model. In the test phase, only the 2-D patches from the T1-weighted MRI scans are fed to the trained model to estimate the number of voxels/pixels attributed to each slice from the corresponding MRI scans.

analyze the hippocampus with its important subfields, such as the right parasubiculum, left and right presubiculum, right subiculum, left dentate gyrus, left CA4, left HATA and right tail. These subregions are believed to be correlated with normal aging and Alzheimer's disease. Moreover, they are known to be more sensitive biomarkers of AD and neurological disorders [4].

On the other hand, data acquisition plays a significant role in the analysis process. Cross sectional data provides less diversities than their longitudinal MRI data counterparts. Therefore, several large-scale investigations on MRI data are being carried out, and different organizations, including the Alzheimer's Disease Neuroimaging Initiative (ADNI), are collecting longitudinal MRI data along with cross-sectional scans. In the longitudinal data acquisition process, the confounding between-subject variability is removed which enables accurate quantification of within-subject neuroanatomical changes and deliver high sensitivity [4], [10]–[12]. Using longitudinal high resolution 3T MRI scans, it is possible to generate ex vivo atlas that allows us to distinguish the multiple subregions of the hippocampus. However, both MRI acquisition techniques [13] are being used to provide a broad range of methods for analyzing different neurological disorders effectively.

There are many different MRI analysis and clinical motivations to pursue various methods for segmentation, such as manual delineation, atlas/statistical-atlas-based methods [14]–[18], statistical parametric approach and/or statistical shape models [6], [19], [20], deformable morphometry-based approaches [21], Bayesian approaches [4], [10], patch-based methods [22], [23], machine learning-based approaches, and deep learning-based approaches [3], [8], [9], [24]–[27]. Segmentation of brain regions from MRI scans using any of these approaches does not strictly rely on the intensity information, rather, the intensity distribution of different subfields has a considerably overlapping intensity values. Furthermore, all the edge boundaries are not properly visible in MRI scans. For example, it is often seen that the white matter of the hippocampus is not well resolved and its boundary can overlap with that of the amygdala, in other words, the boundary can be invisible [28]. Therefore, prior knowledge of hippocampal boundaries is crucially important to trace it properly.

The most common method for segmenting and measuring the volume is an atlas-based system. In an atlas-based system, a reference image is used to coregister the target images where the regions of interest are manually traced onto reference images by an expert radiologist [14]. FreeSurfer is a software package for subcortical segmentation and cortical

parcellation, a popular example for atlas-based systems. It uses atlas images to register/segment the regions of interest from MRI scans. It offers the easiest way to perform such automated operations but it is time consuming.

In this research project, we considered the T1-weighted Gwangju Alzheimer's and Related Dementia (GARD) cohort dataset, consisting of 326 MRI scans, analyzed using the ANT algorithm [7], which is included in the Neuro I software package (http://www.infomeditech. com/). We explored a deep learning algorithm to analyze MRI scans to localize the hippocampi and measure the discrete volumes. The hippocampi were localized automatically using two stage ensemble Hough convolutional neural network (Hough-CNN) model [29], and the voxel positions of hippocampi were traced. Using those voxel locations, 2-D patches of the left and right hippocampi were extracted to train deep learning models. Utilizing the deep learning models, we quantitatively estimated the left and right hippocampal volumes automatically. To the best of our knowledge, this is the first attempt to measure discrete volumes without prior segmentation of hippocampi from the MRI scans.

This paper is organized using the following hierarchy. We illustrate the methodology in section II. In the same section, we described the necessary preprocessing steps to train, validate and test the individual model with their corresponding loss functions. In section III, we describe about the dataset and error estimation procedures. Discrete volume measurement procedure on the test phase and the comparative analysis with other state-of-the-art deep learning methods are explained in the same section. In section IV, we provide a detail overview of our method and it's limitations. Finally, a summary of the entire process is illustrated briefly in section V.

A. PRIOR WORKS
Hippocampal shape, size, partial volume, contrast and resolution constraints of MRI scans have led to the researchers to develop several methods to facilitate the recognition of its structure. To localize the hippocampus and measure the volume accurately and visualize the subfields successfully, several scanning processes and scanners are being built. Although the manual delineation of MRI scans is still considered the gold standard, many automatic methods are being proposed to perform segmentation and localization.

Depending on variations in head size, the total intracranial volume(ICV) is calculated, which is also known as the total cranial volume (TIV). Variation are observed because of different sexes, ages or races [6]. In [6], Ian B. Malone *et al.* used Statistical Parametric Mapping 12 (SPM 12) to automate the segmentation for TIV measurement. The hippocampal volume changes were measured by [5]. Mulder *et al.* [5] measured hippocampal atrophy rates in healthy aging, MCI and AD patients using the automated software package FreeSurfer (longitudinal processing stream) and manual delineation. Their study was to observe the atrophy measurements in between baseline scans and the 12-Month follow-

up visit in the Alzheimer's Disease Neuroimaging Initiative (ADNI) dataset of 80 subjects. On the other hand, the hippocampal subfields were studied using high-resolution 3T MRI scans in [1]. Winterburn *et al.* [1] measured the whole hippocampal volume along with the subfields volume using ex vivo specimens in a 9.4T small-bore scanner. In [30], [31] and [32], hippocampal and amygdala volume measurements were performed to analyze different types of neurological disorders, such as Alzheimer's disease and temporal lobe epilepsy. On the other hand, the relationship between memory/learning and hippocampal volume was studied in 7-years-old children by a group of researchers in [33] to understand whether there is any viable objective that exists with hippocampal volume in an early age. To predict Alzheimer's disease (AD) in patients with mild cognitive impairment (MCI), the hippocampal volume was studied by different groups of researchers. In [34], MCI following AD was observed using hippocampal volume changes.

FreeSurfer 5.3.0 uses it's own atlas scaling factor derived from the registration images of an average template to measure the TIV by utilizing a full affine transformation [6]. However, the latest version of FreeSurfer 6.0 has upgraded functionality to segment and measure the volume of hippocampal subfields using a computational atlas [35]. It requires a large amount of time (several hours for a single MRI scan using a recon-all pipeline with an additional flag) to register and/or segment a single MRI scan and measure the volume of the whole hippocampus with its subfields. On the other hand, FSL FIRST is a Bayesian model-based subcortical brain segmentation/registration tool that uses shapes/appearances from manually segmented images [36]. Similarly, Neuro I is another software package that was built on an open source algorithm provided by ANTs (Advanced Normalization Tools, (https://github.com/ANTsX/ANTs) [7] to reconstruct MRI images into a three-dimensional model and measure the thickness of the cerebral substructure volume and cerebral cortex. In this study, we used the ANT method to analyze data as a standard for the discrete volume measurement to train, validate and test our proposed approach.

Deep learning is a powerful tool for solving numerous complex problems in various disciplines, such as pattern recognition, speech recognition and medical imaging. Various complicated medical imaging problems have already been addressed using deep learning-based algorithms [9], [37]–[41]. Wachinger *et al.* [9] proposed the deep learning-based DeepNAT method to segment the neuroanatomy. BrainNetCNN was proposed by Kawahara *et al.* [24], and was used to predict the neurodevelopment of the brain using Diffusion MRI scans of preterm infants. Automatic hippocampal subfield segmentation is a crucial task for large-scale studies. Dolze *et al.* [3] proposed a 3-D fully convolutional network for subcortical segmentation of MRI scans. They tested their proposed network on two publicly available datasets (the ISBR dataset and the ABIDE dataset). In [42], the authors proposed U-Net, a deep learning-based method, to segment biomedical images. Algorithms based on Hough-CNNs [8],
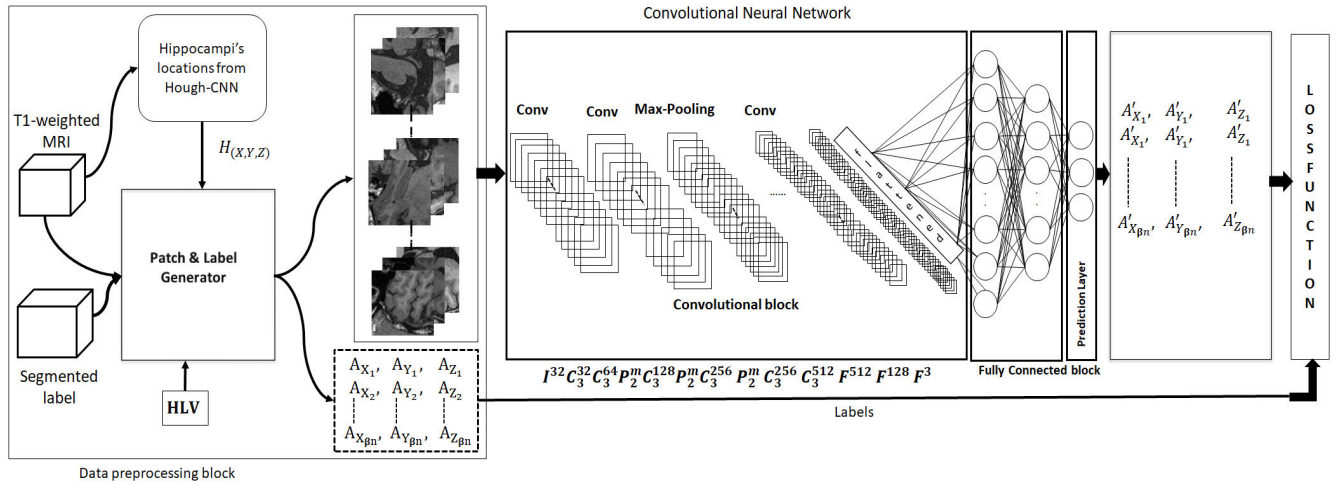
(a) Discrete volume measurement: training system
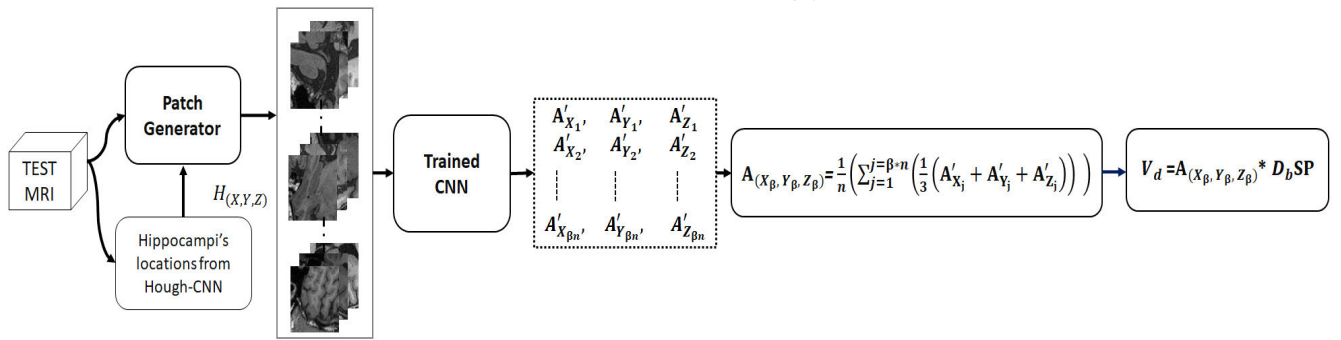
(b) Discrete volume measurement: testing system

FIGURE 2. Discrete volume measurement using the proposed approach.(a) Patch generator used in the single two stage Hough CNN model to locate the hippocampus. Then, using that voxel location, 3-D patches (size: 64 × 64 × 64) are extracted, and then the 3-D patches are split into axial, coronal and sagittal slices. The slices are resized into 32 × 32 2-D patches. The data augmentation operation is performed on the slices. Similarly, using the same localized voxel position, a 3-D patch (size: 64 × 64 × 64) is extracted from the labeled T1-weighted MRI data, and then, the patches are split into axial,coronal and sagittal slices. The number of voxels that were assigned to represent the left and right hippocampi were counted from each slice. The total number of voxels attributed to the hippocmapus in any slice is considered the slice area of that particular slice. The HLV denotes the hippocampus label value. In our case, 17 and 53 were assigned the hippocampus label values for the left and right hippocampi, respectively. The calculated total number of voxels from any slice is used as the ground truth for that particular slice. (b) In the test phase, we only generated the patches. The generated patches are passed through the trained network to determine the number of voxels in the slices contributing to the hippocampus. Then, the areas of the slices in the MRI scan are summed to measure the total area of an MRI scan and multiplied by the slice placement distance ($D_b SP$) to calculate the volume.

[43] [29] are used to detect and localize objects from images of different modalities. Also, a two phase, multimodel automatic brain tumour diagnosis system was developed using CNN in [44]. The ensemble system of a deep convolutional neural network (CNN) and transfer learning-based approach were proposed by a group of researchers in [27], [45], [46] to diagnose Alzheimer's disease from MRI and fMRI scans.

B. CONTRIBUTIONS

This study investigates the possibility of measuring the hard segmented discrete volume of hippocampi from MRI scans using a CNN without performing prior segmentation. Measuring the volume by segmenting the region of interest (ROI) is a conventional method. We tried to eliminate this conventional methodology and propose a direct approach to measure the hippocampal discrete volume automatically using deep learning. The proposed method is illustrated in Fig. 1 and

Fig. 2 for discrete volume measurement. The analysis of MRI scans using a CNN after performing several preprocessing steps offers us a quantitatively efficient deep learning-based model to estimate the volume. Our proposed approach requires a few seconds to estimate the hippocampal volume, which allows for the on-site analysis of a patient's condition. Moreover, this method can be extended to measure any ROI volume from 3-D cross-sectional/longitudinal MRI scans.

II. METHODS AND PREPROCESSING

Hippocampal volumes were automatically estimated from MRI scans using the deep learning model. From the 3-D MRI scans, the corresponding slices attributed to the hippocampi were separated, and the number of voxels was calculated. Then, the total number of voxels was multiplied by the voxel volume associated with the left and right hippocampi; thus, the hippocampal discrete volume was measured.

An algorithm was developed using a CNN to estimate the number of voxels associated with each slice. Prior to measuring the number of voxels from each slice, it is necessary to locate the left and right hippocampi inside an MRI scan. The locations of the left and right hippocampi were estimated using a two-stage Hough-CNN model similar to that in [29]. Using these estimated locations, 3-D patches were extracted, and then the slices were separated into 2-D patches for further preprocessing. The 2-D CNN models were trained using the 2-D patches. These trained models were used to calculate the number of voxels contributing to the hippocampus in each slice of the respective hemisphere in the test phase.

A. CONVOLUTIONAL NEURAL NETWORK

The proposed convolutional neural network (CNN) models [47], [48] consist of a number of layers that conduct operations on the input data (I_{size}). The convolutional layers ($C_{kernel}^{\#filter}$) perform convolution operations on the input images I_{size} with a number of preset kernels. They are usually followed by a batch normalization layer and activation function, where the normalization layer normalizes the convolution results and the activation function rescales the batch-normalized convolution outputs in a non linear manner. Pooling layers (P_{stride}^{Type}) are used to reduce the dimensions of the outcomes produced by the previous layers through downsampling. The type of pooling can be max-pooling or average pooling. Finally, to extract the high-level features, the fully connected layers ($F^{\#filter}$) are employed. The weights are optimized during training through backpropagation [49]. CNNs are good at extracting features without requiring any assistant from the user. The data are processed through the layers in a feed-forward manner, the results of the network are compared with the ground truth through a loss function, and the error is backpropagated to update the weights of all the layers. The training process continues until the model converges. After completing the training, predictions can be made by using the trained CNN model in a feed-forward manner, and the results are reported from the outputs of the last layer.

B. PATCH AND LABEL GENERATION FOR LOCALIZATION

The CNN models were designed in a similar way as those in [29], [50] to localize the left and right hippocampi. The network architecture has been changed slightly for this research work. Instead of using three models for each stage, one model was used for each stage; i.e., for the global phase, one global model was used, and for the local phase, one local model was used. For the global model, the patches were extracted from the whole MRI scan except the boundary region. The uniformly distributed random sample points were used to extract the global patches. The 96×96 patches were extracted with their corresponding displacement vectors and then resized into 32×32 2-D patches. The global patches were used to train the global model to estimate the global position of the hippocampus. For the local model, 32×32-sized patches were generated with corresponding displacement vectors. The patches were normalized considering a standard deviation of 1 and a mean of zero. Using these patches, a local model was trained and validated to locate the hippocampal positions from both sides of the hemispheres.

C. PATCH AND LABEL GENERATION FOR HARD SEGMENTED DISCRETE VOLUME MEASUREMENTS

The 326 MRI samples were separated into 5 roughly equal sections. The cross-validation were preformed by leaving one fold out for testing. For comparison with deep learning segmentation methods such as U-net and SegNet, only the results from the fifth fold is compared.

Using two-stage Hough-CNN model, we located the left and right hippocampi from the MRI scans. These located coordinates are used to extract $64 \times 64 \times 64$, 3-D patches in the vicinity of the left and right hippocampi. From these 3-D patches, the axial, coronal, and sagittal slices were extracted. In the axial, coronal and sagittal planes, 64 slices were collected for each view from each MRI scan. Then, the slices were normalized separately. Using data augmentation technique; the slices are rotated by Θ degrees and the dataset was expanded 10 times. The Θ ($-90 < \Theta < 90$) is chosen from a random integer generator. The augmented slices were then resized and reshaped into $32 \times 32 \times 1$. At the end, the reshaped axial, coronal and sagittal slices were concatenated along axis=2 (0-based axis) to construct 3-channel 2-D patches of size $32 \times 32 \times 3$. From 1-channel gray 3-D MRI images, we have generated 2-D 3-Channel input data for our proposed model.

Similarly, the $64 \times 64 \times 64$ 3-D patches were extracted from the segmented label MRI scans and were separated into slices. Then, the numbers of voxels assigned to the left and right hippocampi were counted. The number of voxels counted from each slice is considered as the ground truth for the corresponding slice in the original MRI. The complete network detail and step by step processes for training and testing are illustrated in Fig. 2. The patch generation and label counting procedures for each MRI scan are explained in Fig. 1.

D. LOCALIZATION NETWORK ARCHITECTURE

The two-stage ensemble Hough CNN [29] was implemented with a minor change. The network sizes were kept the same for the global Hough CNN (GH-CNN) and the local Hough CNN (LH-CNN). The network consists of 6 convolutional layers followed by 3 fully connected layers. Three max pooling layers are used after 2nd, 3rd, and 4th convolutional layers. Batch normalization [51] and ReLU [52] activation layers are used after each of the convolutional layer. Similarly, the fully connected layers are followed by ReLU activation and batch normalization layers, except for the last layer. The last layer is followed by a batch normalization layer. After the first and second fully connected layers, 25% and 35% dropout are used, respectively. The Adam optimizer [53] was used with its default parameter settings except the learning rate, which
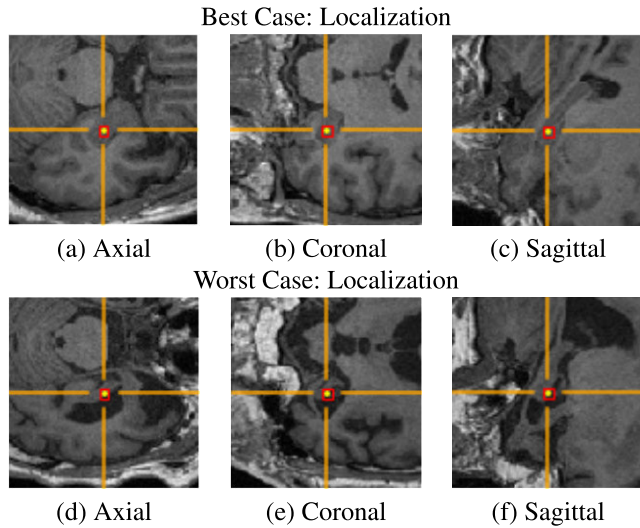
Best Case: Localization

(a) Axial   (b) Coronal   (c) Sagittal

Worst Case: Localization

(d) Axial   (e) Coronal   (f) Sagittal

FIGURE 3. The representative localization by the two-stage ensemble Hough-CNN of right hippocampus is shown for best case (a, b, c) and worst case (d, e, f) of the axial, coronal and sagittal views of the test MRI scans of GARD data set.

(a)   (b)

(c)   (d)

FIGURE 4. Training and testing loss (GARD cohort dataset): (a, c) and (b, d) are the training loss and testing loss on fifth fold for the left and right hippocampi, respectively.

is set to 1e-4. The network architecture details are illustrated in Table 1.

E. LOCALIZATION

The global and local models were trained in exactly the same way as in [29]. The global model learned the features from the whole MRI scan and predicted the global position of the hippocampus. Similarly, the local patches were extracted to train the local model to predict the exact location of the hippocampus. In each phase, the models were trained using the Hough voting strategy described in [8], [29], [43]. In the test phase, the positions estimated by the global model were used to extract the local patches to predict the exact position of the hippocampus. The Localized coordinates of the hippocampi in the test MRI scans were used to extract 3-D patches to estimate the volume of the hippocampi in the corresponding MRI scans. The representative localization of right hippocampus by the two-stage Hough-CNN is shown in Fig. 3.

F. DISCRETE VOLUME MEASUREMENT NETWORK ARCHITECTURE

For the left and right hippocampi's discrete volume measurements, we designed identical network structures. Six convolutional layers were used to form the CNN structures followed by a batch normalization layer and ReLU activation function. A Max pooling layer are used after 2nd, 3rd, and 4th convolutional layers. The fully connected layers are concatenated with the batch normalization layer and ReLU activation function. After the last fully connected layer, we used a batch normalization layer. The Adam optimizer is used with the default parameter settings with a learning rate of 1e-4. The detailed network architecture is shown in Table 2 and Fig. 2(a).
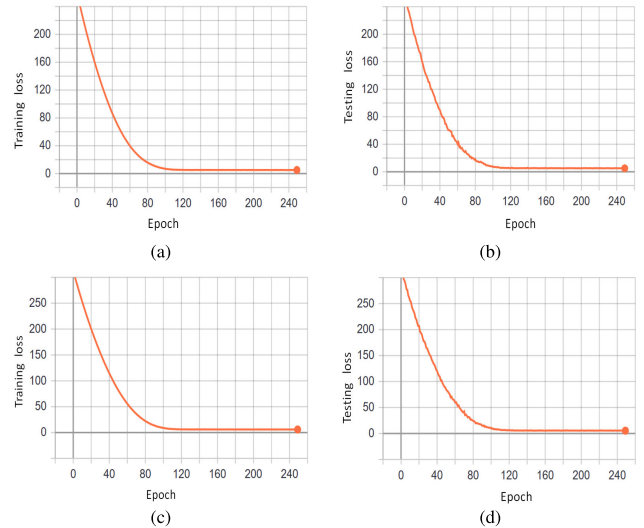
G. LOSS FUNCTION

To train the two-stage ensemble localization network model, the mean squared error is considered as the loss function.

$$MSE_{Hippocampus\ localization}$$
$$= \frac{1}{\alpha * q} \left(\sum_{j=1}^{j=\alpha*q} \right)$$
$$\times \left(\frac{1}{3} \left((X_j - X_j')^2 + (Y_j - Y_j')^2 + (Z_j - Z_j')^2 \right) \right) \quad (1)$$

where α is the number of patches generated from each MRI and q is the total number of MRI scans used for training. (X_j, Y_j, Z_j) are the target displacement vectors and (X_j', Y_j', Z_j') are the predicted displacement vectors.

For hard segmented volume measurements, we used the mean squared error as a loss function as well. The squared differences between the predicted number of voxels of any particular slice and the true number of voxels of the same slice were used to train the network.

$$MSE_{Discrete\ volume}$$
$$= \frac{1}{n} \left(\sum_{j=1}^{j=\beta*n} \right)$$
$$* \frac{1}{3} \left((A_{X_j} - A_{X_j}')^2 + (A_{Y_j} - A_{Y_j}')^2 + (A_{Z_j} - A_{Z_j}')^2 \right) \quad (2)$$

where, $A_{X_j}, A_{Y_j}, A_{Z_j}$ are the true numbers of voxels attributed to the corresponding j^{th} slices of the axial, coronal and sagittal views, respectively, calculated by the ANT method, whereas, $A_{X_j}', A_{Y_j}', A_{Z_j}'$ are the corresponding predicted numbers of voxels of the same j^{th} slice using our proposed approach. Here, β is the number of slices contributing to the hippocampus and n is the factor by which the slices are augmented to increase the dataset size to train the proposed neural network.

TABLE 1. GH-CNN and LH-CNN network architectures used with the GARD cohort dataset.

Model Name		Network Architecture	Activation Function	Batch Normalization	Optimizer
GH-CNN	GM_{lh}	$I_{32}, C_3^{16}, C_3^{32}, P_2^m, C_3^{64}, P_2^m, C_3^{128}, P_2^m, C_3^{256}, C_3^{512}, F^{512}, F^{128}, F^3$	ReLU	All Layers	Adam
	GM_{rh}	$I_{32}, C_3^{16}, C_3^{32}, P_2^m, C_3^{64}, P_2^m, C_3^{128}, P_2^m, C_3^{256}, C_3^{512}, F^{512}, F^{128}, F^3$			
LH-CNN	LM_{lh}	$I_{32}, C_3^{16}, C_3^{32}, P_2^m, C_3^{64}, P_2^m, C_3^{128}, P_2^m, C_3^{256}, C_3^{512}, F^{512}, F^{128}, F^3$	ReLU	All Layers	Adam
	LM_{rh}	$I_{32}, C_3^{16}, C_3^{32}, P_2^m, C_3^{64}, P_2^m, C_3^{128}, P_2^m, C_3^{256}, C_3^{512}, F^{512}, F^{128}, F^3$			

$I_{sample\ size}$ = Network input, $C_{(kernal\ size)}^{(\#filter)}$ = Convolutional layer, P_2^m = Max Pooling with stride 2, $F^{(\#filter)}$ = Fully connected layer.

The models have dropout layer after first fully connected layer (F^{512})(25 %) and second fully connected layer (F^{128})(35 %).

TABLE 2. LHM and RHM network architectures used with the GARD cohort dataset to measure the discrete volume of the left and right Hippocampi.

Model Name	Network Architecture	Activation Function	Batch Normalization	Optimizer
LHM*	$I_{32}, C_3^{32}, C_3^{64}, P_2^m, C_3^{128}, P_2^m, C_3^{256}, P_2^m, C_3^{256}, C_3^{512}, F^{512}, F^{128}, F^3$	ReLU	All Layers	Adam
RHM*	$I_{32}, C_3^{32}, C_3^{64}, P_2^m, C_3^{128}, P_2^m, C_3^{256}, P_2^m, C_3^{256}, C_3^{512}, F^{512}, F^{128}, F^3$	ReLU	All Layers	Adam

$I_{sample\ size}$ = Network input, $C_{(kernal\ size)}^{(\#filter)}$ = Convolutional layer, P_2^m = Max Pooling with stride 2, $F^{(\#filter)}$ = Fully connected layer.

*LHM and RHM stand for left hippocampal model and right hippocampal model, respectively. Both LHM and RHM have dropout layer after first fully connected layer (F^{512})(25 %) and second fully connected layer(F^{128})(35 %).

TABLE 3. Left and Right hippocampi: 5-Fold cross-validation results for the discrete volume measurements.

Model	Fold 1 (Avg. Error% ± STD)	Fold 2 (Avg. Error% ± STD)	Fold 3 (Avg. Error% ± STD)	Fold 4 (Avg. Error% ± STD)	Fold 5 (Avg. Error% ± STD)
LHM	4.5917±3.3713	4.1754±4.0106	4.0221±3.2328	4.3943 ±3.6070	4.4027±3.4963
RHM	2.9874±2.4385	4.5710±4.8443	3.7339±2.7142	4.9675±3.8832	4.5211±3.7510

The representative training and testing curves are shown in Fig. 4. We validated the proposed models by performing 5 folds cross validation. The cross-validated results are shown in Table 3. We used a HP workstation Intel Xeon Processor (3.10 GHz) with 32GB RAM along with INVIDIA Quadro MD4000 GPU (8GB) to conduct the training, validation and testing operation.

III. EXPERIMENTAL RESULTS

A. DATASET

We used the GARD cohort dataset to verify our proposed approach. The GARD cohort dataset consists of 326 MRI scans of 326 patients. The patient age range is from 49 to 87 years, and the average age is 70.0184±6.074 (avg. age ± STD). This dataset is divided into four classes: Alzheimer's disease dementia (ADD), asymptomatic Alzheimer's disease (aAD), mild Alzheimer's diseases (MCI), and normal control (NC). The GARD cohort dataset was analysed using the ANT algorithm (Neuro I software package). Using ANT, the regions of interest (ROIs) were segmented, and the volumes were measured. The segmented MRI ROIs were marked using specific label values. The left and right hippocampi were identified using label values of 17 and 53, respectively. T1-weighted 326 MRI scans with corresponding 326 segmented label MRI scans are considered in this research work.

The MRI scans have $0.512mm^3$ sized voxels with the dimensions of $320 \times 212 \times 240$.

B. DISCRETE VOLUME MEASUREMENT

In an MRI M_{XYZ}, if any voxel in location (X,Y,Z) contributes to the hippocampus, it is assigned as a target voxel to the corresponding slice. The total number of voxels attributed to the hippocampus in any particular slice are considered for the area measurement of that particular slice. In this way, we calculated the total number of voxels in different slices from an MRI scan. The slice area is multiplied by the slice placement distance D_bSP to calculate the total discrete volume.

Let us consider that for the axial, coronal, and sagittal views of any particular MRI scan, if the numbers of hippocampal voxels in that MRI scan are A_X, A_Y, and A_Z for each view, which can be separately considered as the total area of hippocampus in that MRI scan for each view, then the total number of voxels attributed to the hippocampus can be expressed using the following equation.

$$A_{(X,Y,Z)} = \frac{1}{3}(A_X + A_Y + A_Z) \qquad (3)$$

Next, for the β number of slices attributed to the hippocampus from any MRI scan with an augmentation factor of n, we can derive a general formula to calculate the total
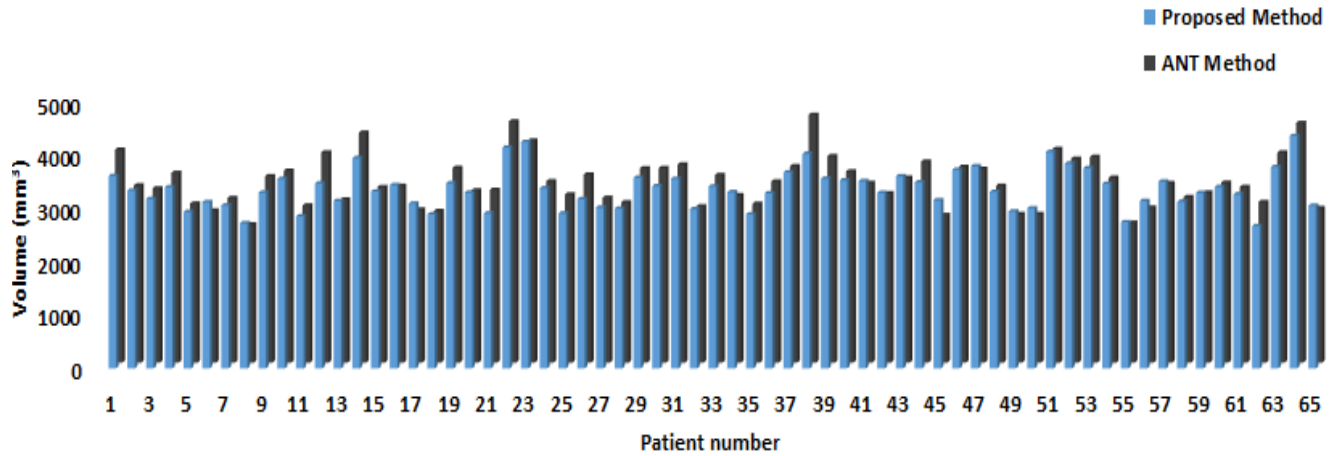
FIGURE 5. Measured left hippocampal volumes using our proposed approach and the ANT method on fifth fold.
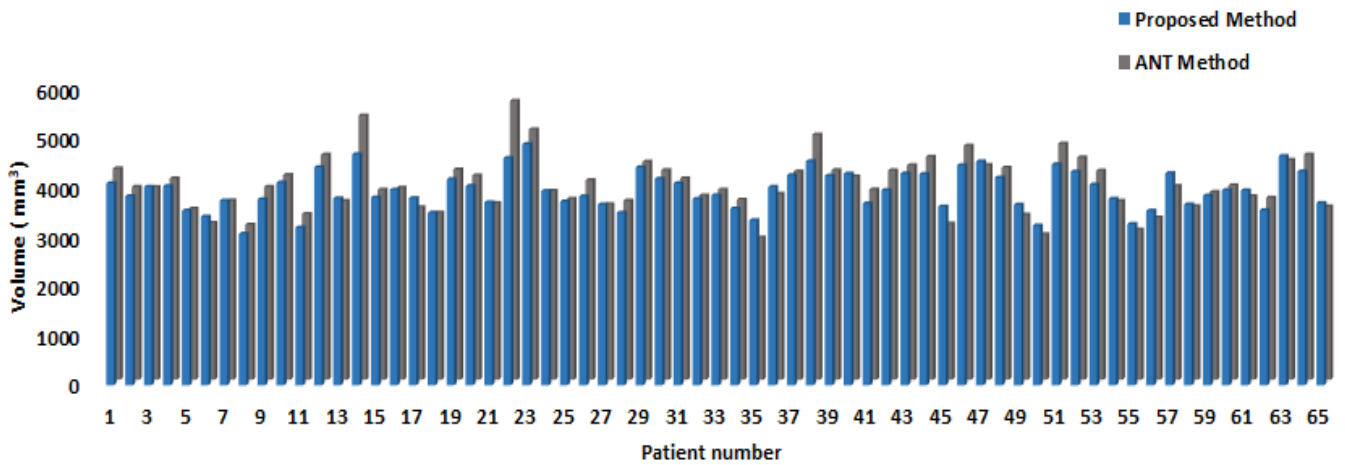

FIGURE 6. Measured right hippocampal volumes using our proposed approach and the ANT method on fifth fold.

number of voxels that form the hippocampus in any particular hemisphere.

$$A_{(X_\beta, Y_\beta, Z_\beta)} = \frac{1}{n} \left(\sum_{i=1}^{\beta*n} \frac{1}{3} \left(A_{X_i} + A_{Y_i} + A_{Z_i} \right) \right) \quad (4)$$

To calculate the discrete volume, the estimated number of voxels needs to be multiplied by the voxel spacing distance, in other words, the number of voxels will be multiplied by the voxel volume.

$$V_d = A_{(X_\beta, Y_\beta, Z_\beta)} * D_b SP \quad (5)$$

Here, V_d is the measured discrete volume of the target MRI scan. The whole procedure of our proposed approach of discrete volume measurement is shown in **Algorithm 1**. The measured discrete volume of the MRI scans of fifth fold using the proposed approach and the ANT method are reported using bar graphs in Fig. 5 and Fig. 6.

C. ERROR ESTIMATION FROM THE PREDICTED VOLUME

If the predicted discrete volume and discrete volume measured by the ANT method are $V_{predicted}$ and V_{Actual}, the prediction error of our proposed approach with the ANT method can be expressed as follows.

$$dE_{Prediction\ Error} = |V_{Actual} - V_{Predicted}| \quad (6)$$

Now, we can express the estimated errors in the discrete volume measurements as a percentage using the following expression.

$$dE_\% = \left(\frac{E_{Prediction\ Error}}{V_{Actual}} x100 \right) \% \quad (7)$$

The average errors in the measured volumes using the proposed approach and the ANT method for the left and right hippocampi are 4.3173 and 4.1562, respectively. In Tables 4 and 5, the predicted discrete volume with their corresponding errors are illustrated for 10 MRI scans (best case and worst case scenarios) along with their original discrete volumes.

TABLE 4. Estimated discrete volumes measured by our method and the ANT methods with the errors in percentage between the methods (left hippocampus).

MRI Scan	Volume - ANT method (mm^3)	Volume - Our method (mm^3)	Error%	Average Error% ±STD
Best 5 Cases				
MRI-37	3690.496	3684.828	0.154	
MRI-28	3010.56	3005.904	0.155	
MRI-54	3475.968	3468.881	0.204	
MRI-48	3324.416	3317.091	0.220	
MRI-2	3332.096	3343.411	0.340	
Worst 5 Cases				4.3173 ± 3.5436
MRI-1	4002.816	3618.242	9.608	
MRI-26	3534.336	3183.169	9.936	
MRI-62	3016.192	2675.152	11.307	
MRI-12	3950.592	3475.718	12.020	
MRI-38	4657.152	4032.773	13.407	

TABLE 5. Estimated discrete volumes measured by our method and the ANT method with the errors in percentage between the methods (right hippocampus).

MRI Scan	Volume - ANT method (mm^3)	Volume - Our method (mm^3)	Error%	Average Error% ± STD
Best 5 Cases				
MRI-39	4234.752	4231.811	0.069	
MRI-33	3841.024	3838.226	0.073	
MRI-29	4405.76	4401.285	0.102	
MRI-31	4068.352	4077.460	0.224	
MRI-60	3931.136	3941.947	0.275	
Worst 5 Cases				4.1562 ±3.5262
MRI-50	2939.904	3222.558	9.614	
MRI-14	5339.136	4664.470	12.636	
MRI-45	3158.016	3606.200	14.192	
MRI-35	2870.784	3330.632	16.018	
MRI-22	5635.584	4588.479	18.580	

D. STATISTICAL ANALYSIS

We performed a statistical analysis on the volumes measured by our proposed approach and the ANT method to determine whether our proposed approach has any statistical significance. To show the agreement in the volume measurements between our proposed method and those of the ANT method, Bland and Altman [54] mean-difference plots were generated using SPSS 16.0. The plots are shown in Fig. 7 and Fig. 8. The p-value measured in the test was less than .01 (p<.01), which indicates that it supported the null hypothesis; i.e., the methods that were utilized to measure the discrete volume of hippocampi were similar and statistically significant. We report the measured volume in Fig. 5 and Fig. 6.

The results were further analyzed to assess the suitability of the proposed approach as a substitute for atlas-based systems, such as the ANT algorithm. We calculated the squared Pearson correlation coefficients (R^2) comparing our method with the ANT method, where a large R^2 indicates that the method can be utilized as a proxy. The measured R^2 values were 0.834 (95% confidence intervals) and 0.848 (95% confidence intervals) for the left and right hippocampi, respectively. Two scatter plots of the discrete volumes of the left and right hippocampi measured by our method and the ANT algorithm with a linear line of best fit are shown in Fig. 9 and Fig. 10.
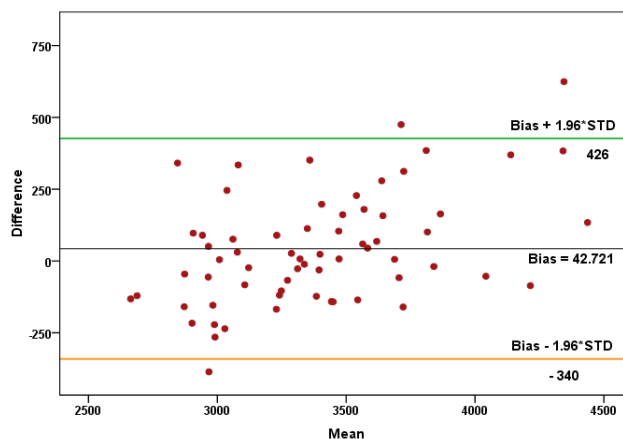

FIGURE 7. Left Hippocampus: fifth fold scatter plot (our approach minus the ANT method plotted against their mean) of the discrete volumes measured by our proposed approach and the ANT method with 95% limits of agreement (LOAs) confined by green and orange lines.

E. COMPARATIVE ANALYSIS WITH OTHER METHODS

We performed a comparison test of our method with the 2-D UNet [42] and SegNet [55]. The segmentation matrices such as IoU, F1 score, Dice coefficient are not applicable in our

Algorithm 1 Discrete Volume Measurement

1   Data: Input MRI scan, M_{XYZ},

2   Output: Estimated Volume, V_d,

3   Get the Hippocampus location using,

     $H_{(X,Y,Z)}$ = two-stage H-CNN(M_{XYZ}),

4   Extract 3-D patches from

     M_{XYZ} using $H_{(X,Y,Z)}$,

5   Separate the 2-D axial, coronal and sagittal

     slices from the 3-D patches,

6   Perform the normalization operation

     on each 2-D slice,

7   Resize the 2-D slices to 32x32,

8   Augment the resized 2-D slices

     by factor of n using a Θ-degree rotation,

9   Generate 3-channel 2-D patch sample

     $S_{\beta*n}$, where, $S_{\beta*n} = \{S_1 \cdots S_{\beta*n}\}$,

10   For i : 1 to $\beta * n$ do

11       $A_{X_{\beta*n}}, A_{Y_{\beta*n}}, A_{Z_{\beta*n}} \leftarrow S_{\beta*n}$

         to the trained CNN model;

12   $A_{(X_\beta, Y_\beta, Z_\beta)} = \frac{1}{n} \left(\sum_{j=1}^{\beta*n} \frac{1}{3} \left(A_{X_j} + A_{Y_j} + A_{Z_j} \right) \right)$

13   $V_d = A_{(X_\beta, Y_\beta, Z_\beta)} * D_b SP$

comparison study. Instead, average voxel/volume errors% are compared. The dice coefficents are also reported as references.

1) DATA PREPROCESSING FOR U-Net AND SegNet
We have trained the U-Net and SegNet using the same MRI scans that were used to train our proposed models. Similarly, the testing set was also kept the same. We have used the axial slices for training, validation and testing the U-Net and SegNet. Previously, using two-stage Hough-CNN, we have located the left and right hippocampi's locations of the MRI scans. Those voxels locations were used to extract 3-D patches of size $80 \times 80 \times 80$. We intentionally increased the size of patches to provide more global view of the brain regions. Then the axial slices (size: 80×80) were extracted
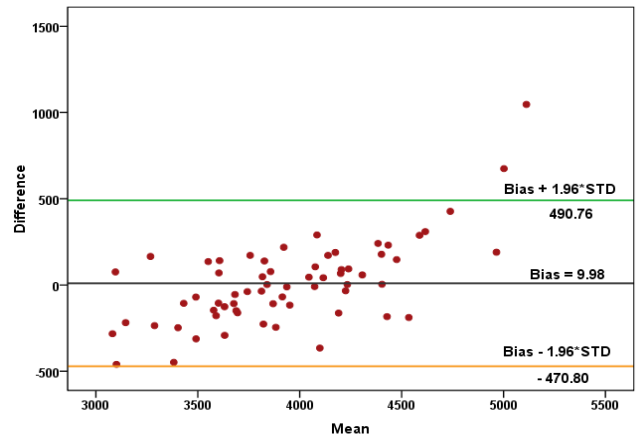

FIGURE 8. Right hippocampus: fifth fold scatter plot (our approach minus the ANT method plotted against their mean) of the discrete volumes measured by our proposed approach and the ANT method with 95% limits of agreement (LOAs) confined by green and orange lines.
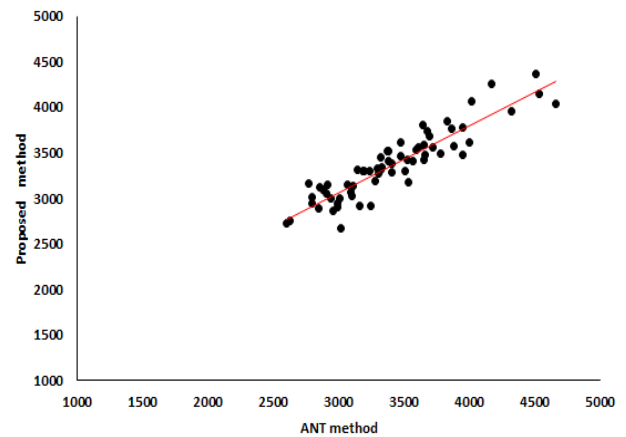

FIGURE 9. Left hippocampus: fifth fold scatter plot of the discrete volumes measured by our proposed approach and the ANT method with the linear line of best fit (not forced through the origin).

FIGURE 10. Right hippocampus: fifth fold scatter plot of the discrete volume measured by our proposed approach and the ANT method with the linear line of best fit (not forced through the origin).
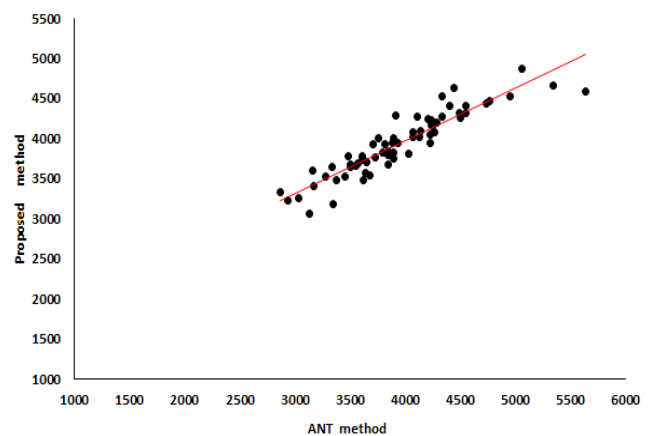
from the 3-D patches. The generated 2-D axial slices were normalized using a mean of zero and standard deviation of 1. The slice were reshaped into a size of $80 \times 80 \times 1$. The original

TABLE 6. U-Net and SegNet training information for left and right hippocampi on GARD Dataset.

Method		Dice loss and Dice Coefficient			
		Training Loss	Validation Loss	Testing Loss	Dice Coefficient
U-Net [42]	LH	0.1811	0.1929	0.1897	0.8114
	RH	0.1795	0.1998	0.1892	0.8130
SegNet [55]	LH	0.3109	0.3519	0.3498	0.8960
	RH	0.2778	0.3163	0.3113	0.9080

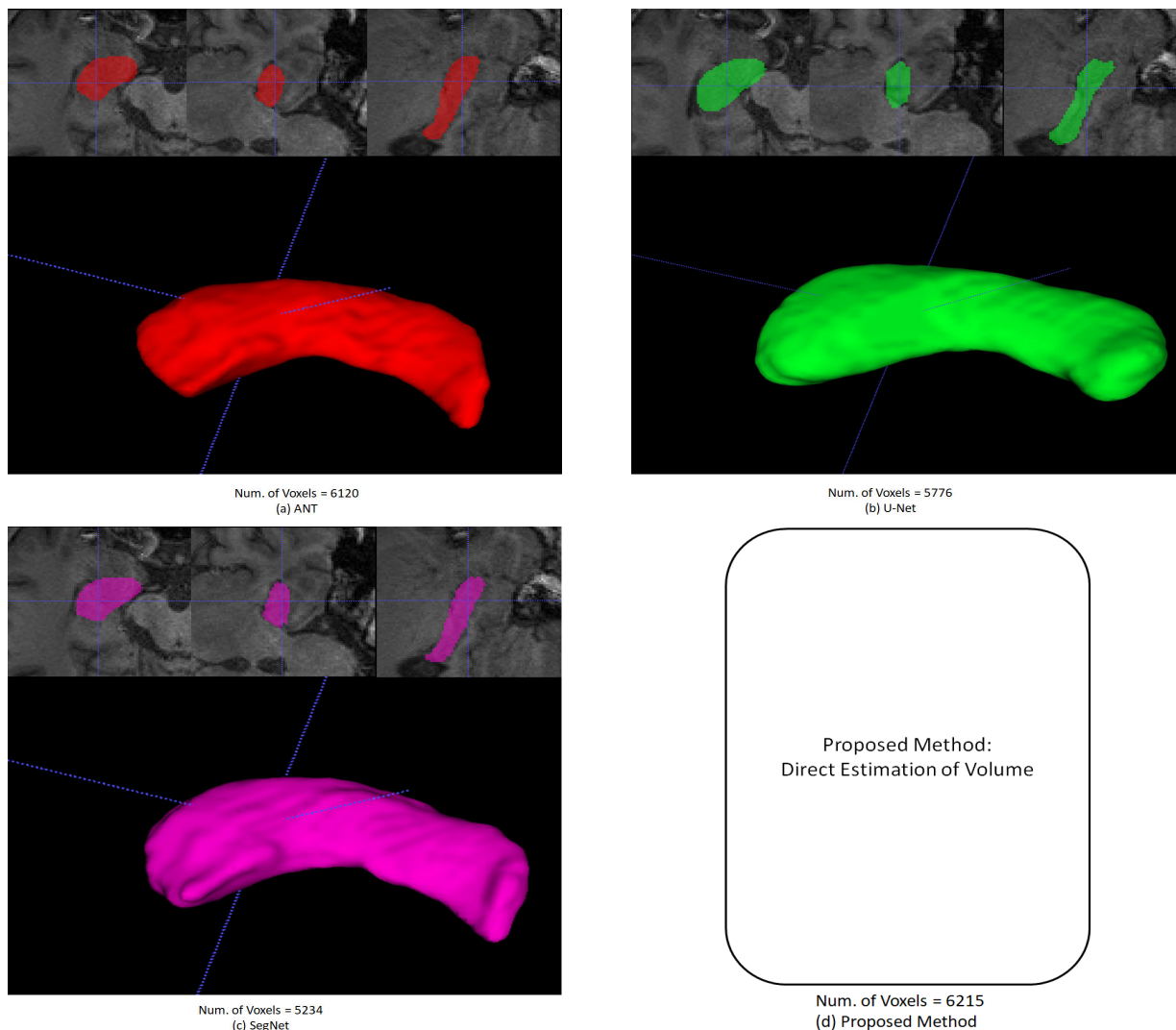*LH and RH stand for left and right hippocampi, respectively.

FIGURE 11. The visual interpretation of segmented volume with the voxel estimation are shown for the right hippocampus of a test sample for ANT method (a), U-Net (b), SegNet (c) and our proposed approach (d). As our method does not generate any mask instead it directly predicts the number of voxels attributed to the hippocampus, therefore, there is no graphical interpretation available for our method.

U-Net used 512×512 images [42] (ISBI 2012 dataset) which was much bigger than our 2-D patches of size 80×80. Similar to the T1-weighted MRI scans preparation, the segmented label MRI scans were pre-processed to generate the ground truths. We replaced HLV value (left hippocampus = 17 and Right hippocampus = 53) to 1 after generating the patches for both right and left hippocampi. The preprocessed data were used to train, validate and test the U-Net and SegNet. The

detail of U-Net and SegNet are illustrated in the following sections.

2) U-Net

We have implemented the U-Net [42] using Keras library (backend: Tensorflow). The original implementation was kept same with few modifications. The loss function has been modified in our implementation. We used dice loss with

TABLE 7. Comparative performance analysis: the average voxel/volume estimation errors%.

Method	Error%± STD for left hippocampus	Error%± STD for right hippocampus
U-Net [42]	13.086* ± 10.761	17.812* ± 14.748
SegNet [55]	10.7116 * ±9.0021	9.3530* ± 8.3483
Proposed Method(64x64 patch)	4.5968* ± 3.5470	4.8600 * ± 4.0470
Proposed Method(32x32 patch)	4.4027*±3.4963	4.5211*±3.7510

*The average errors are calculated on fifth fold.

TABLE 8. Inference time to estimate the voxels/volume.

Method	Inference Time
ANT [7]	15.7±2.0h
U-Net [42]	3.28 s
SegNet [55]	5.92 s
Proposed Method (64x64 patch)	6.57 s
Proposed Method (32x32 patch)	1.92 s

sigmoid activation in the last layer instead of cross entropy loss with softmax activation that was used in the original implementation using Caffe library. The network parameters optimization was performed using Adam optimizer with a learning rate of 1e-5 along with other default parameter settings instead of SGD optimizer.

We have trained the U-Net with a 20 epochs that had taken approximately 12 hours for each side of the hippocampus in the same machine that we have used to train the proposed models. The dice losses for training, validation and testing are shown in Table 6 with the dice coefficient of the predicted mask images for the test MRI scans (fifth fold) against the ground truth. The dice coefficient values were calculated after reconstruction of the 3-D volume (size:80 × 80 × 80) from axial slices of mask images predicted by the trained U-Net. The estimated average dice coefficients for left and right hippocampi are 0.8114±0.0471 and 0.8130±0.0359 (Dice±STD) for MRI scans from fifth fold, respectively. The average computational time to predict the mask image was 3.28 seconds for each MRI scan (each MRI scan (80×80×80, 3-D patch) consisted of axial slices of 80 sample patches).

3) SegNet
SegNet [55] (Encoder-Decoder Architecture) was implemented using Keras library (backend: Tensorflow). In the decoder side, we modified the last layer activation function from softmax to sigmoid and changed the loss function to dice loss. The official SegNet was trained on RGB images, therefore, the input of the encoder side is 3-channel data. In this case, we changed it to 1-channel encoder input. Rest of the architecture was kept same as the original SegNet except for the optimizer.

SegNet was trained with 40 epochs requiring more than 57 hours for each side of the hippocampus in the same machine. We replaced SGD optimizer by Adam and used

a learning rate of 1e-5 with other parameter settings as default. The detail dice losses of the SegNet training, validation and testing along with dice coefficients for the left and right hippocampi are shown in Table 6. After reconstructing 3-D masked volumes (size: 80 × 80 × 80) from the predicted axial slices of test MRI scans by the trained SegNet, the dice coefficients were estimated for the left and right hippocampi. The average dice coefficients are 0.8960±0.0356 and 0.9080±0.0230 for the left and right hippocampi, respectively. The average computational time for SegNet was 5.92 seconds for each reconstructed 3-D Volume.

IV. DISCUSSION
Hippocampal atrophy is a primary feature that contributes to the diagnosis of Alzheimer's disease. Tracing the hippocampus and automatically measuring its volume is a complicated and time-consuming task, even for expert neuro-radiologist [17]. Several methods were introduced to segment and measure the volume automatically, such as automatic segmentation using an atlas and/or a probabilistic atlas-based method [4], deformation-based morphometry [21], [56], and statistical parametric approach [20], [57], [58]. Very recently, machine learning and deep learning algorithms have been proposed to trace and estimate the volume from MRI scans. All these methods were developed to detect or segment and analyze ROIs in the MRI scans for neurodegenerative diseases.

In this research work, we analyzed the GARD cohort dataset of 326 MRI scans using a deep learning algorithm. The purpose of this study was to measure the discrete volumes of the MRI scans and compare them with the atlas-based system to determine whether we can use the proposed approach as a proxy method. We estimated the number of voxels attributed to the corresponding slices, which contribute to the left and right hippocampi on each slice. Then, the total number of voxels was multiplied by the voxel volume to measure the discrete volume of that particular MRI scan. Several research initiatives have been implemented to segment the ROIs based on voxels. In [28], based on probabilistic information, multiple ROIs, including the hippocampi, were segmented by assigning a specific label to the respective voxels for a particular ROI. We used a CNN model to determine the numbers of voxels of the hippocampi. After determining

the voxels, we calculated the volume multiplied by the voxel volume.

We performed a statistical analysis based on the measured volumes to determine whether the proposed approach is a viable substitute for the atlas-based system. We calculated the squared Pearson correlation coefficients of the reported discrete volumes. The values of R^2 were 0.834 and 0.848 (95% confidence intervals) for the left and right hippocampi, respectively. In [59], the authors' reported Pearson correlation coefficients of 0.83 / 0.82 for the manually measured volume and their method for the left hippocampus. Although it is a completely different case, as the datasets are not same, we can have an intuition of a proxy method from this statistical study. Two scatter plots with linear lines of best fit are shown in Fig. 9 and Fig. 10 based on our study. To assess the agreement in the values from our proposed CNN-based approach and the atlas-based approach, we used Bland and Altman [54] plots. The two measurements for the same MRI scan are expected to report the same result, i.e., the same volume, where the slop of regression line will be close to 1. We plotted the difference against the mean value which allows us to assess the bias and deviation. Bias values of 42.72 and 9.98 for the left and right hippocampi shown in Fig. 7 and Fig. 8 signify that, on average, the proposed method measures the volumes of 42.72 mm^3 and 9.98 mm^3 less than the compared ANT method [7] for the left and right hippocampi, respectively. However, we found that both methods were strongly correlated, and the measured p-value was less than 0.01.

At the end, the proposed method was compared with the state-of-the-art deep learning methods, such as U-Net and SegNet. The detail implementation and other changes made in the original architecture of U-Net and SegNet are explained in the previous sections. The U-Net performance was good in the context of dice coefficient (0.8114 ± 0.0471 and 0.8130 ± 0.0359 (Dice$\pm$STD) for the left and right hippocampi, respectively), however, the average voxel/volume errors% were comparatively very high (13.086 ± 10.761 and 17.812 ± 14.748 (Error%$\pm$STD) for the left and right hippocampi, respectively). Therefore, the proposed method outperformed the U-Net in the context of average voxel/volume estimation errors%. Moreover, SegNet performance was better than U-Net, however, our proposed method offers less average voxel/volume errors% than the SegNet. The average voxel/volume errors% for the left and right hippocampi of MRI scans from fifth fold along with the computational time are shown in Table 7 and Table 8. The proposed method's inference time is given for the final CNN model (Localization models (two-stage Hough-CNN) require 2 seconds to estimate the hippocampus position in an MRI scan.). ANT method's reported inference time is for the whole brain processing, whereas, the other methods only process the hippocampal region. The representative segmentation performance of U-Net, SegNet and ANT method with direct estimation of voxels on the same MRI scan by the proposed method are shown in Fig. 11.

A. LIMITATIONS
Although the proposed approach estimates the discrete segmentation volume, it cannot visualize the exact shapes of the hippocampi. The motivation of this research work is to measure the volume directly for on-site diagnosis and observe the volume changes in the baseline scans with other time point scans, such as 4-, 8-, and 12-Month visits.

V. CONCLUSION
In this research paper, we demonstrated a complete automatic neural network-based approach to measure discrete volumes from MRI scans. Our proposed method can accurately measure the left and right discrete hippocampal volumes from MRI scans and saves the time in a large margin by discarding the conventional method of measuring the volume by manually segmenting the region of interest. As volume atrophy is one of the primary biomarkers for Alzheimer's disease, the automatic estimation of discrete volumes from MRI scans can help in the diagnosis of Alzheimer's disease. The average errors of the predicted discrete volumes with the ANT method of the left and right hippocampi were 4.3173 $\pm$ 3.5436 (avg. error % $\pm$ STD) and 4.1562 $\pm$3.5262 (avg. error % $\pm$ STD), respectively. The statistical analysis of the measured volumes showed that the volumes measured by proposed approach were significantly correlated with those from the ANT method. In addition, the proposed approach outperforms the state-of-the-art deep learning methods, such as U-Net and SegNet in case of average voxel/volume estimation errors%. The proposed approach can be used as a proxy method for the measurement of the discrete volume from MRI scans.

REFERENCES

[1] J. L. Winterburn, J. C. Pruessner, S. Chavez, M. M. Schira, N. J. Lobaugh, A. N. Voineskos, and M. M. Chakravarty, "A novel *in vivo* atlas of human hippocampal subfields using high-resolution 3T magnetic resonance imaging," *NeuroImage*, vol. 74, pp. 254–265, Jul. 2013.

[2] K. Y. Choi, J. J. Lee, T. I. Gunasekaran, S. Kang, W. Lee, J. Jeong, H. J. Lim, X. Zhang, C. Zhu, and S.-Y. Won, "Apoe promoter polymorphism-219t/g is an effect modifier of the influence of apoe $\varepsilon4$ on alzheimer's disease risk in a multiracial sample," *J. Clin. Med.*, vol. 8, no. 8, p. 1236, 2019.

[3] J. Dolz, C. Desrosiers, and I. Ben Ayed, "3D fully convolutional networks for subcortical segmentation in MRI: A large-scale study," *NeuroImage*, vol. 170, pp. 456–470, Apr. 2018.

[4] J. E. Iglesias, K. Van Leemput, J. Augustinack, R. Insausti, B. Fischl, and M. Reuter, "Bayesian longitudinal segmentation of hippocampal substructures in brain MRI using subject-specific atlases," *NeuroImage*, vol. 141, pp. 542–555, Nov. 2016.

[5] E. R. Mulder, R. A. de Jong, D. L. Knol, R. A. van Schijndel, K. S. Cover, P. J. Visser, F. Barkhof, and H. Vrenken, "Hippocampal volume change measurement: Quantitative assessment of the reproducibility of expert manual outlining and the automated methods FreeSurfer and FIRST," *NeuroImage*, vol. 92, pp. 169–181, May 2014.

[6] I. B. Malone, K. K. Leung, S. Clegg, J. Barnes, J. L. Whitwell, J. Ashburner, N. C. Fox, and G. R. Ridgway, "Accurate automatic estimation of total intracranial volume: A nuisance variable with less nuisance," *NeuroImage*, vol. 104, pp. 366–372, Jan. 2015.

[7] N. J. Tustison, P. A. Cook, A. Klein, G. Song, S. R. Das, J. T. Duda, B. M. Kandel, N. van Strien, J. R. Stone, J. C. Gee, and B. B. Avants, "Large-scale evaluation of ANTs and FreeSurfer cortical thickness measurements," *NeuroImage*, vol. 99, pp. 166–179, Oct. 2014.
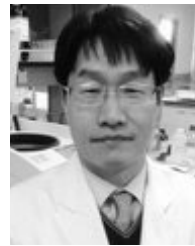
[8] F. Milletari, S.-A. Ahmadi, C. Kroll, A. Plate, V. Rozanski, J. Maiostre, J. Levin, O. Dietrich, B. Ertl-Wagner, K. Bötzel, and N. Navab, "Hough-CNN: Deep learning for segmentation of deep brain regions in MRI and ultrasound," *Comput. Vis. Image Understand.*, vol. 164, pp. 92–102, Nov. 2017.

[9] C. Wachinger, M. Reuter, and T. Klein, "DeepNAT: Deep convolutional neural network for segmenting neuroanatomy," *NeuroImage*, vol. 170, pp. 434–445, Apr. 2018.

[10] J. E. Iglesias, K. Van Leemput, P. Bhatt, C. Casillas, S. Dutt, N. Schuff, D. Truran-Sacrey, A. Boxer, and B. Fischl, "Bayesian segmentation of brainstem structures in MRI," *NeuroImage*, vol. 113, pp. 184–195, Jun. 2015.

[11] Y. Shao, J. Kim, Y. Gao, Q. Wang, W. Lin, and D. Shen, "Hippocampal segmentation from longitudinal infant brain MR images via classification-guided boundary regression," *IEEE Access*, vol. 7, pp. 33728–33740, 2019.

[12] K. Trojachanec, I. Kitanovski, I. Dimitrovski, and S. Loshkovska, "Longitudinal brain MRI retrieval for Alzheimer's disease using different temporal information," *IEEE Access*, vol. 6, pp. 9703–9712, 2018.

[13] T. Kaufmann, G. O. Nygaard, M. K. Beyer, P. Sowa, J. E. Nordvik, K. Kolskår, G. Richard, O. A. Andreassen, H. F. Harbo, and L. T. Westlye, "Cross-sectional and longitudinal MRI brain scans reveal accelerated brain aging in multiple sclerosis," *Frontiers Neurol.*, vol. 10, p. 450, Apr. 2019.

[14] O. T. Carmichael, H. A. Aizenstein, S. W. Davis, J. T. Becker, P. M. Thompson, C. C. Meltzer, and Y. Liu, "Atlas-based hippocampus segmentation in Alzheimer's disease and mild cognitive impairment," *NeuroImage*, vol. 27, no. 4, pp. 979–990, Oct. 2005.

[15] F. van der Lijn, T. den Heijer, M. M. B. Breteler, and W. J. Niessen, "Hippocampus segmentation in MR images using atlas registration, voxel classification, and graph cuts," *NeuroImage*, vol. 43, no. 4, pp. 708–720, Dec. 2008.

[16] A. R. Khan, N. Cherbuin, W. Wen, K. J. Anstey, P. Sachdev, and M. F. Beg, "Optimal weights for local multi-atlas fusion using supervised learning and dynamic information (SuperDyn): Validation on hippocampus segmentation," *NeuroImage*, vol. 56, no. 1, pp. 126–139, May 2011.

[17] N. Amoroso, T. Alzheimer's Disease Neuroimaging Initiative, M. L. Rocca, R. Bellotti, A. Fanizzi, A. Monaco, and S. Tangaro, "Alzheimer's disease diagnosis based on the hippocampal unified multi-atlas network (HUMAN) algorithm," *Biomed. Eng. OnLine*, vol. 17, no. 1, p. 6, Dec. 2018.

[18] Y. Liu, Y. Wei, and C. Wang, "Subcortical brain segmentation based on atlas registration and linearized kernel sparse representative classifier," *IEEE Access*, vol. 7, pp. 31547–31557, 2019.

[19] J. C. Rajapakse and J. Piyaratna, "Bayesian approach to segmentation of statistical parametric maps," *IEEE Trans. Biomed. Eng.*, vol. 48, no. 10, pp. 1186–1194, Oct. 2001.

[20] W. Penny, K. Friston, J. Ashburner, S. Kiebel, and T. E. Nichols, *Statistical Parametric Mapping: The Analysis of Functional Brain Images*. Jan. 2007.

[21] P. A. Yushkevich, B. B. Avants, S. R. Das, J. Pluta, M. Altinay, and C. Craige, "Bias in estimation of hippocampal atrophy using deformation-based morphometry arises from asymmetric global normalization: An illustration in ADNI 3 t MRI data," *NeuroImage*, vol. 50, no. 2, pp. 434–445, Apr. 2010.

[22] P. Coupé, J. V. Manjón, V. Fonov, J. Pruessner, M. Robles, and D. L. Collins, "Nonlocal patch-based label fusion for hippocampus segmentation," in *Medical Image Computing and Computer-Assisted Intervention*, T. Jiang, N. Navab, J. P. W. Pluim, and M. A. Viergever, Eds. Berlin, Germany: Springer, 2010, pp. 129–136.

[23] P. Coupé, J. V. Manjón, V. Fonov, J. Pruessner, M. Robles, and D. L. Collins, "Patch-based segmentation using expert priors: Application to hippocampus and ventricle segmentation," *NeuroImage*, vol. 54, no. 2, pp. 940–954, Jan. 2011.

[24] J. Kawahara, C. J. Brown, S. P. Miller, B. G. Booth, V. Chau, R. E. Grunau, J. G. Zwicker, and G. Hamarneh, "BrainNetCNN: Convolutional neural networks for brain networks; Towards predicting neurodevelopment," *NeuroImage*, vol. 146, pp. 1038–1049, Feb. 2017.

[25] F. Ren, C. Yang, Q. Qiu, N. Zeng, C. Cai, C. Hou, and Q. Zou, "Exploiting discriminative regions of brain slices based on 2D CNNs for Alzheimer's disease classification," *IEEE Access*, vol. 7, pp. 181423–181433, 2019.

[26] C. Feng, A. Elazab, P. Yang, T. Wang, F. Zhou, H. Hu, X. Xiao, and B. Lei, "Deep learning framework for Alzheimer's disease diagnosis via 3D-CNN and FSBi-LSTM," *IEEE Access*, vol. 7, pp. 63605–63618, 2019.

[27] S. Sarraf, D. D. Desouza, J. A. E. Anderson, and C. Saverino, "MCADNNet: Recognizing stages of cognitive impairment through efficient convolutional fMRI and MRI neural network topology models," *IEEE Access*, vol. 7, pp. 155584–155600, 2019.

[28] B. Fischl, "Whole brain segmentation: Automated labeling of neuroanatomical structures in the human brain," *Neuron*, vol. 33, no. 3, pp. 341–355, 2002.

[29] A. Basher, K. Y. Choi, J. J. Lee, B. Lee, B. C. Kim, K. H. Lee, and H. Y. Jung, "Hippocampus localization using a two-stage ensemble Hough convolutional neural network," *IEEE Access*, vol. 7, pp. 73436–73447, 2019.

[30] C. Watson, F. Andermann, P. Gloor, M. Jones-Gotman, T. Peters, A. Evans, A. Olivier, D. Melanson, and G. Leroux, "Anatomic basis of amygdaloid and hippocampal volume measurement by magnetic resonance imaging," *Neurology*, vol. 42, no. 9, p. 1743, Sep. 1992.

[31] C. R. Jack, "MRI-based hippocampal volume measurements in epilepsy," *Epilepsia*, vol. 35, no. s6, pp. 21–29, Dec. 1994.

[32] F. Cendes, F. Andermann, P. Gloor, A. Evans, M. Jones-Gotman, C. Watson, D. Melanson, A. Olivier, T. Peters, I. Lopes-Cendes, and G. Leroux, "MRI volumetric measurement of amygdala and hippocampus in temporal lobe epilepsy," *Neurology*, vol. 43, no. 4, p. 719, Apr. 1993.

[33] C. Omizzolo, D. K. Thompson, S. E. Scratch, R. Stargatt, K. J. Lee, J. Cheong, G. Roberts, L. W. Doyle, and P. J. Anderson, "Hippocampal volume and memory and learning outcomes at 7 years in children born very preterm," *J. Int. Neuropsychol. Soc.*, vol. 19, no. 10, pp. 1065–1075, Nov. 2013.

[34] C. R. Jack, R. C. Petersen, Y. C. Xu, P. C. O'Brien, G. E. Smith, R. J. Ivnik, B. F. Boeve, S. C. Waring, E. G. Tangalos, and E. Kokmen, "Prediction of AD with MRI-based hippocampal volume in mild cognitive impairment," *Neurology*, vol. 52, no. 7, p. 1397, Apr. 1999.

[35] J. E. Iglesias, J. C. Augustinack, K. Nguyen, C. M. Player, A. Player, M. Wright, N. Roy, M. P. Frosch, A. C. McKee, L. L. Wald, B. Fischl, and K. Van Leemput, "A computational atlas of the hippocampal formation using ex vivo, ultra-high resolution MRI: Application to adaptive segmentation of *in vivo* MRI," *NeuroImage*, vol. 115, pp. 117–137, Jul. 2015.

[36] B. Patenaude, S. M. Smith, D. N. Kennedy, and M. Jenkinson, "A Bayesian model of shape and appearance for subcortical brain segmentation," *NeuroImage*, vol. 56, no. 3, pp. 907–922, Jun. 2011.

[37] X. Hong, R. Lin, C. Yang, N. Zeng, C. Cai, J. Gou, and J. Yang, "Predicting alzheimer's disease using lstm," *IEEE Access*, vol. 7, pp. 80893–80901, 2019.

[38] L. Yue, X. Gong, J. Li, H. Ji, M. Li, and A. K. Nandi, "Hierarchical feature extraction for early Alzheimer's disease diagnosis," *IEEE Access*, vol. 7, pp. 93752–93760, 2019.

[39] S. Afzal, M. Maqsood, F. Nazir, U. Khan, F. Aadil, K. M. Awan, I. Mehmood, and O. Song, "A data augmentation-based framework to handle class imbalance problem for alzheimer's stage detection," *IEEE Access*, vol. 7, pp. 115528–115539, 2019.

[40] M. N. I. Qureshi, S. Ryu, J. Song, K. H. Lee, and B. Lee, "Evaluation of functional decline in Alzheimer's dementia using 3D deep learning and group ICA for RD-FMRI measurements," *Frontiers Aging Neurosci.*, vol. 11, p. 8, Feb. 2019.

[41] N. T. Duc, S. Ryu, M. N. I. Qureshi, M. Choi, K. H. Lee, and B. Lee, "3D-deep learning based automatic diagnosis of Alzheimer's disease with joint MMSE prediction using resting-state fMRI," *Neuroinformatics*, vol. 18, no. 1, pp. 71–86, Jan. 2020.

[42] O. Ronneberger, P. Fischer, and T. Brox, "U-net: Convolutional networks for biomedical image segmentation," in *Proc. 18th Int. Conf.*, Oct. 2015, vol. 9351, N. Navab, J. Hornegger, W. M. W. III, and A. F. Frangi, eds. Munich, Germany: Springer, 2015, pp. 234–241.

[43] J. Gall, A. Yao, N. Razavi, L. Van Gool, and V. Lempitsky, "Hough forests for object detection, tracking, and action recognition," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 33, no. 11, pp. 2188–2202, Nov. 2011.

[44] M. K. Abd-Ellah, A. I. Awad, A. A. M. Khalaf, and H. F. A. Hamed, "Two-phase multi-model automatic brain tumour diagnosis system from magnetic resonance images using convolutional neural networks," *EURASIP J. Image Video Process.*, vol. 2018, no. 1, p. 97, Dec. 2018.

[45] J. Islam and Y. Zhang, "Brain mri analysis for alzheimer's disease diagnosis using an ensemble system of deep convolutional neural networks," *Brain Informat.*, vol. 5, p. 2, May 2018.

[46] N. M. Khan, N. Abraham, and M. Hon, "Transfer learning with intelligent training data selection for prediction of Alzheimer's disease," *IEEE Access*, vol. 7, pp. 72726–72735, 2019.

[47] Y. Lecun, L. Bottou, Y. Bengio, and P. Haffner, "Gradient-based learning applied to document recognition," *Proc. IEEE*, vol. 86, no. 11, pp. 2278–2324, 1998.

[48] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "Imagenet classification with deep convolutional neural networks," in *Proc. Adv. Neural Inf. Process. Syst.*, P. L. Bartlett, F. C. N. Pereira, C. J. C. Burges, L. Bottou, and K. Q. Weinberger, eds., Lake Tahoe, NV, USA, 2012, pp. 1106–1114.

[49] Y. LeCun, Y. Bengio, and G. E. Hinton, "Deep learning," *Nature*, vol. 521, no. 7553, pp. 436–444, 2015.

[50] W. A. Al, H. Y. Jung, I. D. Yun, Y. Jang, H.-B. Park, and H.-J. Chang, "Automatic aortic valve landmark localization in coronary CT angiography using colonial walk," *PLoS ONE*, vol. 13, no. 7, 2018, Art. no. e0200317.

[51] S. Ioffe and C. Szegedy, "Batch normalization: Accelerating deep network training by reducing internal covariate shift," in *Proc. Int. Conf. Mach. Learn.*, vol. 37. Jul. 2015, pp. 448–456.

[52] V. Nair and G. E. Hinton, "Rectified linear units improve restricted Boltzmann machines," in *Proc. 27th Int. Conf. Mach. Learn. (ICML)*, J. Fürnkranz and T. Joachims, eds. Madison, WI, USA: Omnipress, 2010, pp. 807–814.

[53] D. P. Kingma and J. Ba, "Adam: A method for stochastic optimization," in *3rd Int. Conf. Learn. Represent. (ICLR)*, Y. Bengio and Y. LeCun, Eds., San Diego, CA, USA, May 2015, pp. 1–15.

[54] J. Martin Bland and D. Altman, "Statistical methods for assessing agreement between two methods of clinical measurement," *Lancet*, vol. 327, no. 8476, pp. 307–310, Feb. 1986.

[55] V. Badrinarayanan, A. Kendall, and R. Cipolla, "SegNet: A deep convolutional encoder-decoder architecture for image segmentation," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 39, no. 12, pp. 2481–2495, Dec. 2017.

[56] O. Colliot, O. Camara, and I. Bloch, "Integration of fuzzy spatial relations in deformable models—Application to brain MRI segmentation," *Pattern Recognit.*, vol. 39, no. 8, pp. 1401–1414, Aug. 2006.

[57] O. Muzik, D. C. Chugani, C. Juhász, C. Shen, and H. T. Chugani, "Statistical parametric mapping: Assessment of application in children," *NeuroImage*, vol. 12, no. 5, pp. 538–549, Nov. 2000.

[58] S. J. Kiebel, J. Ashburner, J.-B. Poline, and K. J. Friston, "MRI and PET coregistration—A cross validation of statistical parametric mapping and automated image registration," *NeuroImage*, vol. 5, no. 4, pp. 271–279, 1997.

[59] R. Maglietta, T. Alzheimers Disease Neuroimaging Initiative, N. Amoroso, M. Boccardi, S. Bruno, A. Chincarini, G. B. Frisoni, P. Inglese, A. Redolfi, S. Tangaro, A. Tateo, and R. Bellotti, "Automated hippocampal segmentation in 3D MRI using random undersampling with boosting algorithm," *Pattern Anal. Appl.*, vol. 19, no. 2, pp. 579–591, May 2016.

BYEONG C. KIM received the M.D., M.S., and Ph.D. degrees from the Department of Neurology, Chonnam National University Medical School, Gwangju, South Korea, in 1991, 1994, and 2000, respectively. He is currently a Professor with the Department of Neurology, Chonnam University Medical School, Gwangju. His current research interests include structural and molecular neuroimaging and CSF biomarkers in patients with neurodegenerative diseases.

KUN HO LEE received the B.S. degree from the Department of Genetic Engineering, Korea University, Seoul, South Korea, in 1989, and the M.S. and Ph.D. degrees from the Department of Molecular Biology, Seoul National University, Seoul, in 1994 and 1998, respectively. He is currently an Associate Professor with the Department of Biomedical Science, Chosun University, Gwangju, South Korea. He also works with the National Research Center for Dementia, Chosun University. His current research interests include brain image analysis and the development of prediction model for neurodegenerative diseases-based on magnetic resonance imaging and genetic variants.

ABOL BASHER received the B.S. degree in electrical and electronics engineering from the Mymensingh Engineering College, University of Dhaka, Bangladesh, in 2015, and the M.S. degree in computer engineering from Chosun University, Gwangju, South Korea, in 2020, where he is currently pursuing the Ph.D. degree with the Department of Computer Engineering. He is also working as a Research Assistant with the Computer Vision Laboratory, Chosun University. His research interests include medical image processing, computer vision, and deep learning.

HO YUB JUNG received the B.S. degree in electrical engineering from The University of Texas at Austin, in 2002, and the M.S. and Ph.D. degrees in electrical engineering and computer science from Seoul National University, in 2006 and 2012, respectively. He was a Senior Engineer with Samsung Electronics for a period of two years. Since 2017, he has been an Assistant Professor with the Department of Computer Engineering, Chosun University. His research interests include computer vision, machine learning, and medical imaging.

● ● ●