Toward deep MRI segmentation for Alzheimer's disease detection

Hadeer A. Helaly[1] · Mahmoud Badawy[1,2] · Amira Y. Haikal[3]

Abstract

Alzheimer's disease (AD) is an irreversible, progressive, and ultimately fatal brain degenerative disorder, no effective cures for it till now. Despite that, the available treatments can delay its progress. So, early detection of AD plays a crucial role in preventing and controlling its progress. Hippocampus (HC) is among the first impacted brain regions by AD. Its shape and volume are measured using a structural magnetic resonance image (MRI) to help AD diagnosis. Therefore, brain hippocampus segmentation is the building block for AD detection. This study's main objective is to propose a deep learning Alzheimer's disease hippocampus segmentation framework (DL-AHS) for automatic left and right hippocampus segmentation to detect and identify AD. The proposed DL-AHS framework is based on the U-Net architecture and estimated on the baseline coronal T1-weighted structural MRI data obtained from Alzheimer's disease neuroimaging initiative (ADNI) and neuroimaging tools and resources collaboratory (NITRIC) datasets. The dataset is processed using the Medical Image Processing, Analysis, and Visualization (MIPAV) program. Besides, it is augmented using a deep convolutional generative adversarial network (DC-GAN). For left and right HC segmentation from other brain sub-regions, two architectures are proposed. The first utilizes simple hyperparameters tuning in the U-Net (SHPT-Net). The second employs a transfer learning technique in which the ResNet blocks are used in the U-Net (RESU-Net). The empirical results confirmed that the proposed framework achieves high performance, 94.34% accuracy, and 93.5% Dice similarity coefficient for SHPT-Net. Also, 97% accuracy and 94% Dice similarity coefficient are achieved for RESU-Net.

1 Introduction

Alzheimer's disease (AD) refers to a neurodegenerative disorder that chronically progressively attacks the brain tissues, including β-amyloid peptide (Aβ), neurofibrillary tangles, and neuronal degeneration [1, 2]. It also impacts both mental and memory functions [3, 4]. Until now, no effective medicines are available to cure AD, and the current drugs only hinder or delay its progress. As a result, the awareness of AD at its early stage is vital for controlling and preventing its development. For AD treatment and biomarkers diagnostic, several genes and pathways play critical roles [5]. The most significantly changed gene is PDHA1 which serves as a target gene in AD treatment [6].

Ventricles size, hippocampus shape, cortical thickness, and brain volume are the main AD-related variations of anatomical brain structures handled for early and accurate AD diagnosis [7]. Hippocampus is a small medial sub-cortical brain structure relevant to long and short-term

✉ Hadeer A. Helaly
hadeerhelaly@students.mans.edu.eg

Mahmoud Badawy
engbadawy@mans.edu.eg

Amira Y. Haikal
amirayh@gmail.com

1 Electrical Engineering Department, Faculty of Engineering, Damietta University, Damietta, Egypt

2 Department of Computer Science and Informatics, Taibah University, Medina, Saudi Arabia

3 The Head of Computers and Control Systems Engineering Department, Faculty of Engineering, Mansoura University, Mansoura, Egypt

memory [8–10]. Many pathologies such as the neurodegeneration associated with AD can affect its shape and volume [11], in which the hippocampus shape changes in time, as well as different rates of volume loss. Therefore, hippocampus shape and volume are vital factors for detecting Alzheimer's disease and distinguishing Alzheimer's patients from healthy persons [12, 13].

Image segmentation is a digital image clustering process that divides the image into several coherent sub-regions or segments according to the extracted features[14, 15]. It is the primary diagnosis, treatment, planning, monitoring, and pathology characterization in many medical applications [16, 17]. It helps a specialist derive different volumetric and morphological pointers. It supports the quantitative analysis and characterization of numerous neurological diseases and their development. Also, it supports experts' focus on specific regions in the medical image [18]. The accuracy of diagnosing AD can be improved by MRI segmentation of the brain structure because MRI is distinguished by high spatial resolution, high availability, and ability to contrast soft tissue [19, 20]. Besides, compared to positron emission tomography (PET) and computed tomography (CT) modalities, MRI is correlating with fewer health risks [21].

Machine learning and artificial algorithms have enabled radiologists to segment medical images, such as mammograms of breast cancer, brain tumors, brain lesions, and identifying lung nodules [18]. Deep learning-based medical imaging applications have exceeded the performance of traditional methods in complicated tasks [22]. It offers significant methodologies in segmentation [23], detection, and image pattern classification with great success [24–26]. Applying deep learning algorithms for medical image segmentation requires solving domain-specific difficulties associated with the quantity and quality of data and annotations [27].

To this end, the segmentation of natural images using deep convolutional neural networks (CNN) has achieved high performance [28]. This progress is mainly because of the paradigm shift from manual to automatic extraction of features provided by deep learning networks and the high significant computational power improvements [29]. The advancement of CNN has led to U-Net's creation with a modified U-shaped structure, which can segment biomedical images. U-Net joins the convolutional network with a de-convolutional network. The main benefit is that it performs the demanded tasks with fewer training images and gives more precise segmentation using a GPU. It is also proved that the U-Net architecture achieves high performance than several biomedical segmentation applications for medical imaging [30].

The main contribution in the proposed work is to solve and manipulate some of the MRI medical image segmentation problems. These issues revolve around the scarcity of medical images dataset, low anatomical structure contrast for T1, T2, and FLAIR modalities. Furthermore, manual segmentation for brain MRI is time-consuming and needs expert knowledge of brain anatomy. Besides, transfer learning, multi-task, or multi-module learning and the scalability of deep learning approaches are applied. Therefore, a clinical validation tool is proposed which manipulates the above-listed issues. Also, it gives support in analyzing the prognosis, the variation of HC region, and finding of AD.

In this study, the DL-AHS framework is proposed for automatic left and right hippocampus segmentation to detect and identify Alzheimer's disease. The proposed framework is based on the U-Net architecture and evaluated based on coronal T1-weighted structural MRI data obtained from ADNI and NITRIC datasets. The dataset is segmented and processed using the MIPAV program, and it is augmented using the DC-GAN. For medical MRI binary semantic-wise segmentation, two architectures are proposed. The first architecture is SHPT-Net which utilizes simple hyperparameters tuning. The second is RESU-Net that applies a transfer learning technique in which the ResNet block is used for feature extraction and contraction processes in the U-Net architecture. The empirical results confirmed that the proposed framework achieves high performance over other state-of-the-art techniques.

This paper is coordinated throughout the following way: in Sect. 2, the relevant works are reviewed. Section 3 outlines the major issues and the aims of this study. In Sect. 4, the techniques and materials are discussed. In Sect. 5, the experiments and the results are assessed. Section 6 summarizes the paper.

2 Related work

The hippocampus (HC) is among the first impacted regions in the brain by AD among all brain regions of interest (ROI). So, it is an imperative anatomical region in the AD etiology. The volume and shape attributes from the bilateral hippocampus are highly considered in MRI research for AD diagnosis. For subcortical structure segmentation, Dolz et al. [31] suggested 3D FCNNs where 3D and fully CNN for subcortical MRI brain structure segmentation was applied. That method was robust but required high computational power and memory due to the use of 3D architecture.

Allioui et al. [32] suggested a computer-aided diagnosis (CAD) framework. The proposed framework was the first 2.5D MRI system analysis based on U-Net architecture. The network was trained from scratch. It could segment images of the brain, detect mental disorders, and identify

AD. The system exploited the 3D MRI benefits. Besides, the complexity and computational costs are reduced. Jingwen et al. [33] proposed a 3D CNN based on V-Net for segmenting bilateral hippocampus from 3D brain MRI scans and diagnosing AD progression states. The V-Net convolutional blocks were replaced with a bottleneck architecture to compress the model. That network had higher robustness and accuracy. But it was applied to a small dataset.

Automatic hippocampus segmentation and AD classification were combined in a multi-task model by Manhua et al. [34]. That framework was based on CNN and evaluated on structural MRI data. The features of the 3D patches extracted based on the hippocampus segmentation results were learned using 3D densely connected convolutional networks (3D DenseNet). Then, it merged with the learned features from the multi-task CNN to detect AD. The framework gave promising performance but suffered from complexity. By several optimization techniques such as genetic algorithm (GA), lion optimization algorithm (LOA), artificial bee colony (ABC), BAT algorithm, and particle swarm optimization (PSO), the HC region is segmented by Chitradevi et al. [35]. After comparing these optimization techniques, LOA gave better performance than others due to its characteristics of escaping from local optima.

Carmo et al. [9] proposed a deep learning U-net method for hippocampus segmentation. It uses an extended 2D multi-orientation approach. Also, a public Alzheimer's disease hippocampus segmentation (HarP) dataset was used for developing and validating that method. The methodology gave a good performance but suffered from a low standard deviation between overall Dice and left/right Dice. Furthermore, a robust automatic atlas CNN-based hippocampus segmentation tool called DeepHarp for hippocampus delineation was developed by Sammaneh et al. [36]. The method was trained from scratch and evaluated by the ADNI harmonized hippocampal protocol (HarP). Table 1 describes the comparison among the relevant works listed above.

The proposed framework avoids the point of weaknesses listed before for Dolz et al. [31], Allioui et al. [32], Jingwen et al. [33]. It depends on the 2D architecture for medical images MRI and 2D layers in the proposed U-Net architectures. Therefore, the complexity represented in high computational power and high memory requirements is avoided. Moreover, Jingwen et al. [33] network is applied to a small dataset due to medical image scarcity. As a result, the proposed framework employs the DCGAN technique to increase the dataset size.

3 Problem statement and solution plan

Recently, biomedical MRI segmentation and classification using deep learning have been widely studied. However, it involves several issues and challenges such as (i) Medical images dataset scarcity, (ii) low anatomical structure contrast for T1, T2, and FLAIR modalities, (iii) manual segmentation for brain MRI which is time-consuming and need expert knowledge of brain anatomy, (iv) applying transfer learning, multi-task, or multi-module learning, and (v) the scalability of deep learning approaches [21]. Also, numerous architectures that can accommodate AD detection have been proposed in the literature, as seen in Sect. 2. However, most of them lack in manipulation the biomedical MRI segmentation and classification challenges listed above. So, according to other state-of-the-art techniques reviewed in Sect. 2, and to manipulate and solve some of the MRI segmentation challenges listed, the novelty of this study is organized as follows:

- As manual segmentation is time-consuming and needs proficient knowledge, a deep learning Alzheimer's disease automatic left and right hippocampus segmentation framework (DL-AHS) is proposed.
- Due to medical dataset scarcity, T1-weighted MRI dataset is collected from ADNI and NITRIC datasets and then segmented, denoised, and reconstructed.
- Due to low anatomical structure contrast for T1 modalities, the MIPAV program is used to segment, process, denoise, and prepare the dataset (bilateral filter and Inhomogeneity N3 correction algorithms are used).
- Due to the small dataset size, MRI dataset augmentation techniques are applied using the DC-GAN.
- Medical MRI binary semantic-wise segmentation is proposed using two new architectures based on U-Net architecture called SHPT-Net and RESU-Net.
- The SHPT-Net architecture employs simple hyperparameters tuning on U-Net architecture.
- The RESU-Net architecture uses the transfer learning concept in which the pre-trained model (ResNet) is used in the encoder and decoder layers of the U-Net architecture.
- The proposed architectures achieve high performance according to four performance metrics (Dice similarity coefficient, accuracy, sensitivity, and specificity)

4 Methods and materials

The Hippocampus (HC) segmentation is the primary factor for diagnosing AD [37]. Our goal is to propose a DL-AHS framework to segment the left and right hippocampus to

Table 1 A comparison among recent related work

	Dataset	Technique	Advantages	Disadvantages
Dolz et al. [31]	(ABIDE, ISBR)	3D FCNNs	1. The method is robust to various acquisition protocols, demographics, and clinical factors 2. The network has few parameters and, thus, is less prone to overfitting	1. High computational complexity 2. High memory requirements
Allioui et al. [32]	OASIS	U-Net architecture for each view of the 2.5D brain MRI after parsing them into transversal views	1. It takes advantage of the benefits of 3D architecture 2. It reduces complexity and computational costs	The network was trained from scratch and not benefits from transfer learning concepts
Jingwen et al. [33]	ADNI	3D convolutional neural network based on V-Net	1. That model had good robustness in the three-category classification task of pathological brain states 2. Accurately segment bilateral hippocampus	1. Applied to small dataset 2. Computational complexity when dealing with 3D images 3. Sample numbers of three-category with AD progression states are unbalanced
Manhua et al. [34]	ADNI	Multi-model deep CNNs for jointly learning hippocampus segmentation and disease classification evaluated on structural MRI data	1. The empirical results demonstrated that it achieved promising performance 2. The framework output the disease status and provided the hippocampus segmentation result	Computational complexity
Chitradevi et al. [35]	Hospital images	HC region is segmented by different optimization techniques, namely GA, ABC, BAT, PSO, and LOA	1. The system does not include highly complex computations and hardware implementations 2. The LOA method showed the best classification accuracy compared to all other methods	The system is not applied for MCI, which allows the doctor to examine AD in an early stage
Carmo et al. [9]	HarP	A deep learning U-net based CNN's method for hippocampus segmentation	Achieves state-of-the-art performance on the public HarP hippocampus segmentation benchmark	1. Low standard deviation between overall Dice and left/right Dice 2. That method was not ready to deal with hippocampus resection due to epilepsy treatment
Sammaneh et al. [36]	ADNI harmonized hippocampal protocol (HarP)	CNN-based hippocampus segmentation tool for MRI data called DeepHarp	This method gave high accuracy and robustness, which can aid atrophy measurements in a variety of pathologies	The CNN architecture was built from scratch and did not use the transfer learning concept

detect and identify Alzheimer's disease. There will be a comprehensive explanation of the proposed DL-AHS framework workflow, the preprocessing algorithms, and the techniques for medical image segmentation in the next subsections. U-Net and transfer learning will also be described.

4.1 Proposed framework

The proposed DL-AHS framework comprises five steps. The full pipeline of the framework, shown in Fig. 1, is as follows:

> *Step 1—The data acquisition step*: for medical image segmentation, needs the MRI scans and the ground truth masks for the input dataset. Due to the scarcity and curation of the medical dataset, a dataset is prepared by
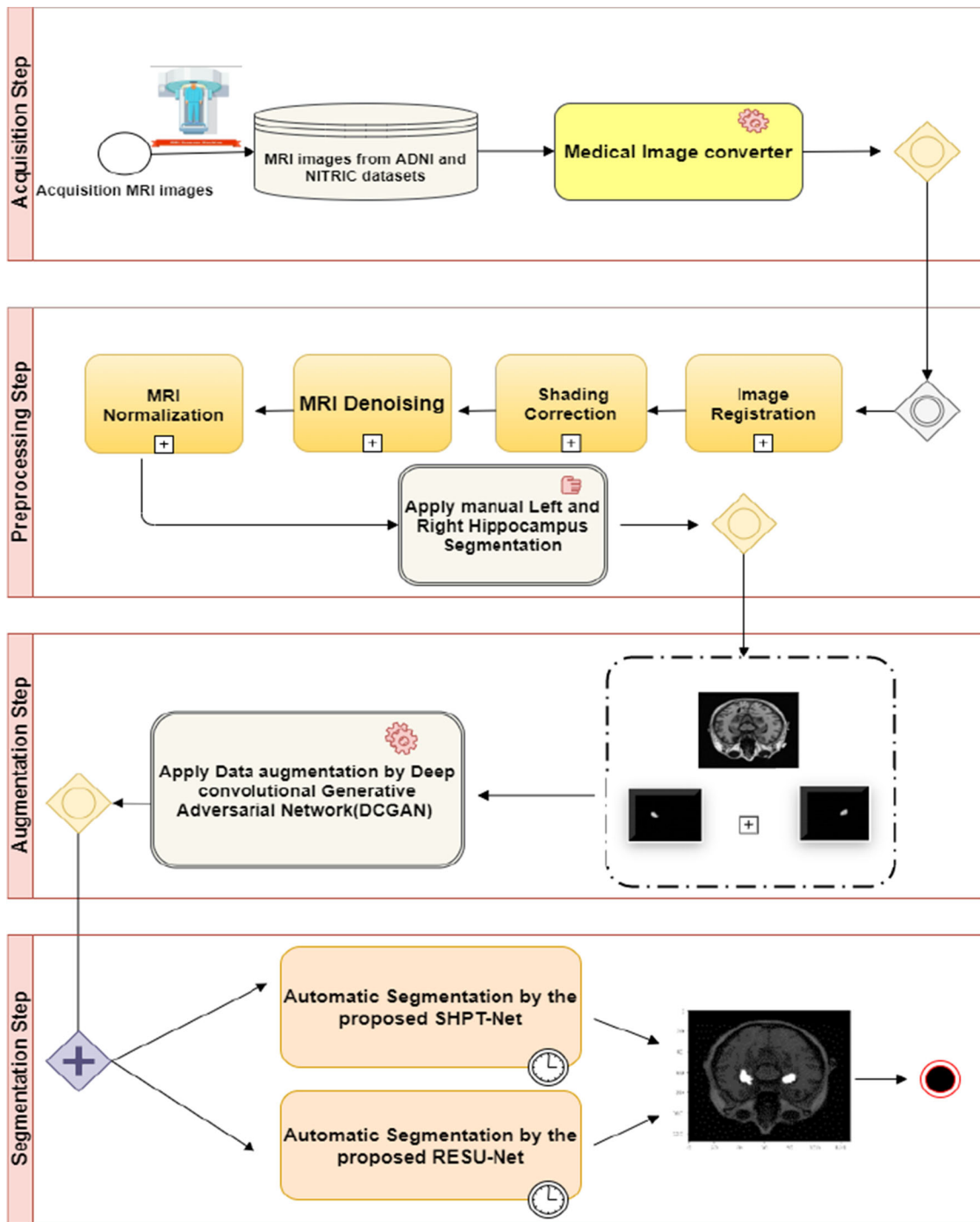
Fig. 1 The full pipeline of the proposed DL-AHS framework

collecting coronal, 2D, T1-weighted MRI slices in DICOM, hdr and IMG formats from the ADNI (https://www.loni.ucla.edu/ADNI) and NITRIC (https://www.nitrc.org/) datasets. Sixty-four patients with MRI historical scans representing various AD stages are collected.

Step 2—Preprocessing step: The MIPAV program is used for processing, segmenting the left and right hippocampus regions, denoising the dataset by bilateral filter algorithm, and shading correction of images by inhomogeneity N3 correction. Also, the dataset is normalized and registered in a suitable format. Thus, by the MIPAV program, manual segmentations have
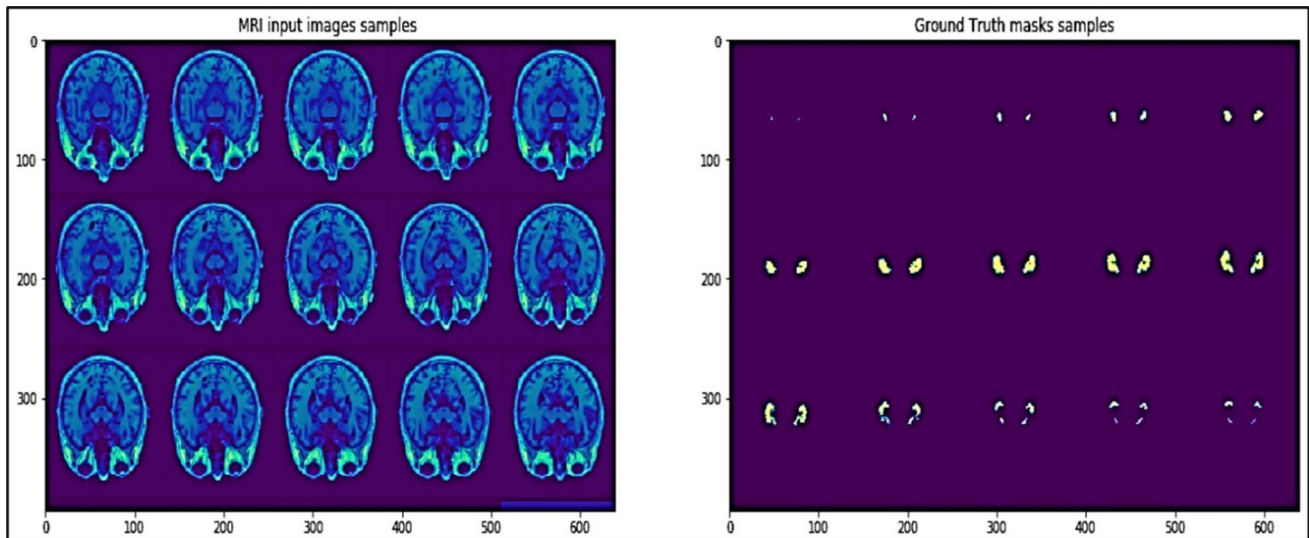
Fig. 2 Samples of MRI input data and the corresponding ground truth masks
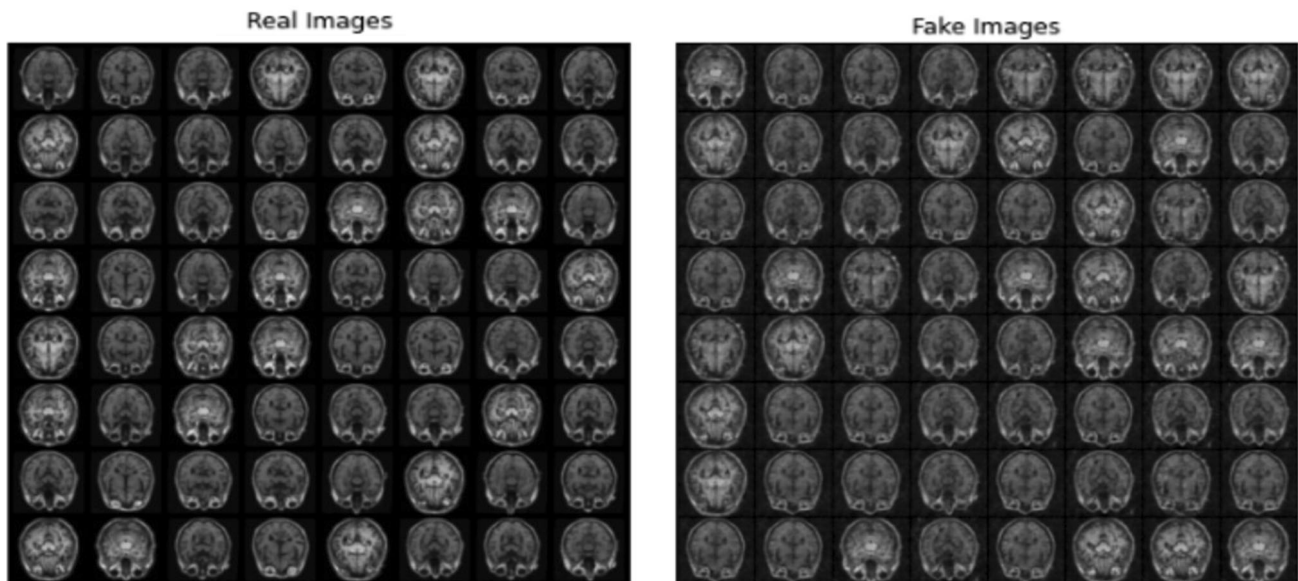

Fig. 3 The real and fake training samples after 5000 epochs of training the DCGAN

been used for preparing the ground truth in both segmentation model training and for final segmentation performance evaluation.

For each slice of the 64 patients slices collected, we segment the left and right HC in two slices using polygon volume of interest (VOI) and binary mask options. So, the dataset's size becomes 64 MRI slices for the original images, 64 for the left HC, 64 for the right HC. The samples of the original MRI images and their ground truth masks are shown in Fig. 2.

Step 3—Data augmentation step: due to the small dataset amount, data augmentation techniques are applied to maximize the dataset size and prevent the overfitting problem. The DCGAN is used to augment the data. The real and fake training samples we get after 5000 epochs of training the DCGAN are shown in Fig. 3. After applying DCGAN and other simple augmentation techniques, the dataset's size becomes 1500 MRI slices, 500 for the original slices, 500 for the left HC slices, and 500 for the right HC.

Step 4—Segmentation step: in this step, two proposed architectures are applied based on U-Net, which are called SHPT-Net and RESU-Net. The SHPT-Net utilizes simple hyperparameters tuning on U-Net architecture.
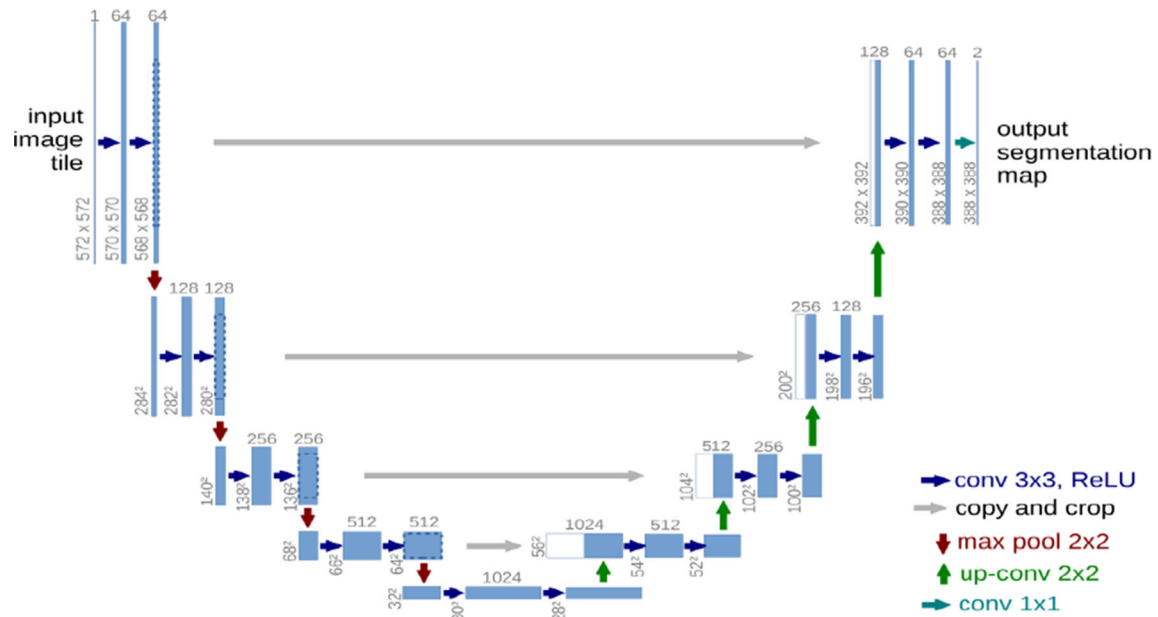
Fig. 4 The default U-Net architecture [30]

The default U-Net architecture [30] and the proposed SHPT-Net architecture are described in Figs. 4 and 5. The RESU-Net architecture uses the transfer learning technique where the ResNet pre-trained model is used. The ResNet blocks are used in encoder and decoder layers in the U-Net architecture.

Step 5—Evaluation step: in this step, we monitor and review the model's performance to ensure effective results and accurate performance by following the DL-AHS algorithm steps. The acronyms used in the algorithm are shown in Table 2.

4.2 Preprocessing algorithms

4.2.1 Shading correction by inhomogeneity N3 correction

Inhomogeneity N3 correction algorithm is a robust, accurate, and fully automatic algorithm used to correct shading artifacts often seen in MRI. It roughly calculates both distributions of true tissue intensities and multiplicative bias field. So, it is an iterative algorithm. Also, it makes no assumptions about the kind of anatomy present in a scan referred to as nonparametric intensity non-uniformity normalization (N3). For MRI segmentation, it is vital to apply shading correction for MRI images for highly and efficient results.

The determination of the model is a joint problem with tissue intensity model-based methods. In the N3 correction algorithm, the modeling assumption that the data histogram is blurred by non-uniformity. Therefore, it can be observed, measured, and deleted. This non-uniformity blurring distribution is referred to as the blurring kernel. The basis of the N3 approach is in Eq. 1.

$$v(x) = u(x) \times f(x) + n(x) \tag{1}$$

where n is white, Gaussian noise assumed to be independent of the true signal emitted by the tissue u. f is an unknown smoothly varying bias field, and the measured signal is v this at location x. The bias field f multiplies by signal u to corrupt it. Then, the noise n is added to get the output signal v.

DL-AHS Algorithm

Input: Structural T1, Coronal MR Images

Output: h_s, accuracy, dice simitlary coefficient, sensitivity, specificity

1: Preprocessing $(A_d + N_d) \rightarrow d_p$

2: Segmentation $d_p \rightarrow h_l$, h_r

3: Augmentation d_p , h_r , h_l

4: Random(d) /* partition the data into training d_{train} (80%) and validateion set val (10%) portions*/

5: d_{test}=d $- (d_{train}+ d_{val})$ used for models evaluation.

6: while (training (SHPT-Net or RESU-Net))

    If (network test performance != network training performance), then

      The network suffers from overfitting or underfitting or other problem

      Repeat the step from 3-6 after making modifications in the proposed models

    else:

      continue training until reach high-performance metrics.

7. Test the models on unknown Coronal MRI images and get h_s.

4.2.2 Image registration

Registration is used in image and data compilation and analysis. It converts various sets of data such as many images data from different times, sensors, depths, or viewpoints into one coordinate system. Therefore, to compare data acquired from several measurements such as our dataset collected from ADNI and NITRIC datasets in various formats, it is essential to do image registration.

4.2.3 Bilateral filter

The bilateral filter is a highly effective, nonlinear, edge-preserving, and noise-reducing smoothing filter for images. However, its operation is slower than other filters because the intensity of each pixel is substituted by the weighted average of intensity from the nearby pixels. The MRI medical image before and after applying the bilateral filter is shown in Fig. 6. The bilateral filter is specified as in Eqs. 2, 3.

$$I^{filtered}(x) = \frac{1}{w_p} \sum_{x_i \in \Omega} I(x_i) f_r(||I(x_i) - I(x)||) g_s(||x_i - x||),$$

(2)

And normalization term w_p is defined as

$$w_p = \sum_{x_i \in \Omega} f_r(||I(x_i) - I(x)||) g_s(||x_i - x||)$$

(3)

where $I^{filtered}$ is the filtered image; I is the original input image to be filtered; x are the coordinates of the current pixel to be filtered; Ω is the window centered in x, so $x_i \in \Omega$ is another pixel. f_r is the range kernel for smoothing differences in intensities (this function can be a Gaussian function); g_s is the spatial (or domain) kernel for smoothing differences in coordinates (this function can be a Gaussian function).

4.2.4 MRI normalization

Normalization is a process of converting the MRI data to a comprehensive anatomic template. It maps the data acquired from discrete subject space to a reference space containing a template and a source image. The main advantages of normalization are: it facilitates comparing brain MRI images and interprets them onto a common
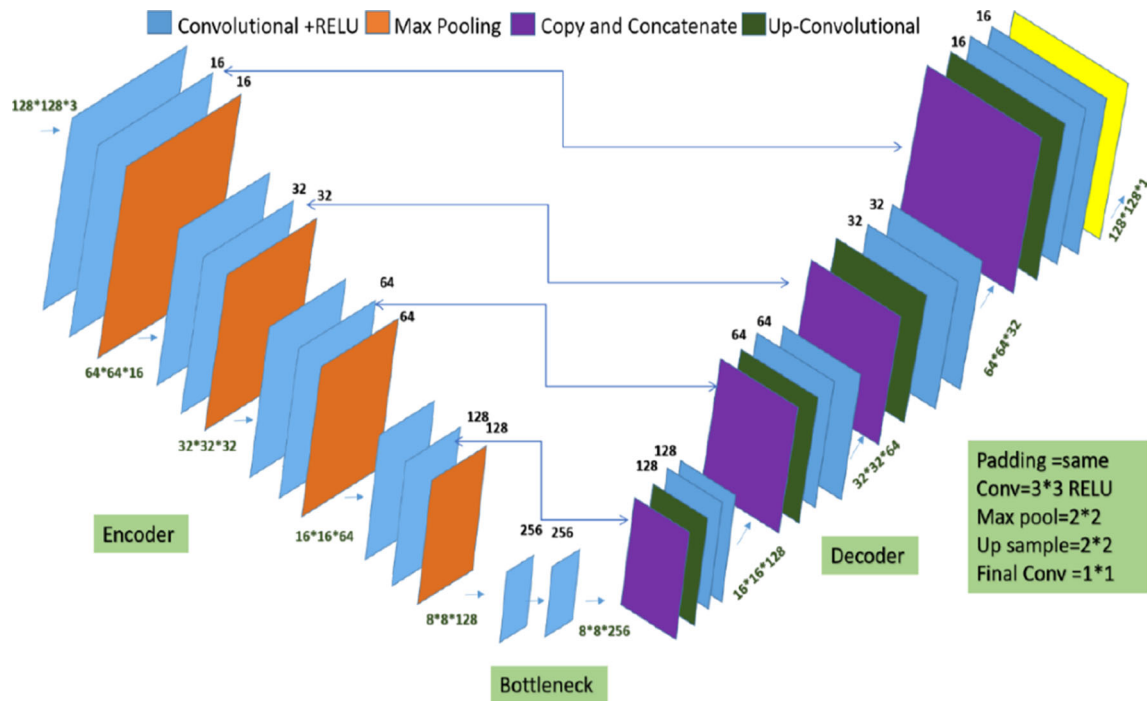
Fig. 5 The proposed SHPT-Net architecture

Table 2 Acronyms used in DL-AHS algorithm

Acronyms	Description
A_d	ADNI dataset
N_d	NITRIC dataset
d'_{train}	Dataset for training
d_{val}	Dataset for validation
d_{test}	Dataset for testing
d_p	Preprocessed MRI images
h_l	Left hippocampus mask
h_r	Right hippocampus mask
d	The dataset that includes of numbers of features and the ground truth ($h_r + h_l$)
h_s	Left and right hippocampus segmented image (output)

shape and size [38]. It removes some variations in the data, such as different subject pose or differences in image contrast, to simplify the detection of subtle differences. There are several types of normalization, such as intensity normalization (IN), spatial normalization (SN), Z-score normalization (ZN), and numerical normalization (NNM).

4.3 Medical image segmentation types and techniques

Image segmentation before deep learning is applied by several techniques such as (i) thresholding, (ii) clustering methods, (iii) histogram-based methods, (iv) edge detection, (v) region-growing methods, (vi) graph-based approaches, (vii) watershed transformations, and (vii) feature-based techniques. Deep learning techniques give efficient and high-performance results [15]. Additionally, deep learning approaches utilize several applications behind image segmentation, such as classification, object detection, genotype detection, speech recognition [39]. Stacked auto-encoders, deep neural networks, CNN, and deep Boltzmann machines are popular deep learning algorithms [17].

Using the CNN architecture, several segmentation methods include cascaded, multi-modality, single-modality, patch-wise, and semantic-wise. In our work, we depend on semantic-wise segmentation because of its benefits over other methods. Semantic-wise segmentation links each pixel of the input image with its class label; it is also called the dense prediction process because every pixel is
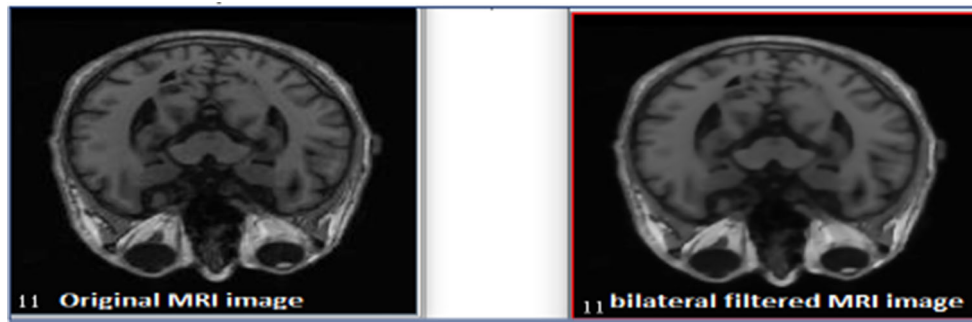
Fig. 6 Original MRI image and bilateral filter image

predicted from the whole input image [40]. Its main advantages are that it minimizes the loss function and generates segmentation maps for any image size, and its computational complexity is lower than other methods [21, 41].

4.4 Deep convolutional generative adversarial networks (DCGAN)

DCGAN is a deep-learning-based generative model. The primary usage of it is for data augmentation, in which new plausible artificial examples from the input data are created [42]. In image data, applying techniques such as flips, crops, zooms, or other simple transformations on the existing images of the training dataset refer to primitive augmentation techniques. Thus, we also apply DCGAN, a perfect choice for accurate and high-efficiency model results.

GAN consists of two parts; the first part is the generator that learns to generate new plausible artificial training examples, which become negative examples for the discriminator's second part. The discriminator learns to differentiate the faked generated data from the real one [32, 43, 44]. The workflow for generator and discriminator processes is shown in Fig. 7, and the general structure of the applied DCGAN is shown in Fig. 8. Figure 9 shows the generator and discriminator loss during training through 5000 training epochs.

4.5 U-Net

U-Net is a convolutional network with a modified U-shaped structure, used in biomedical image segmentation due to its excellent performance in the ImageNet. Its structure consists of a fully convolutional network with a de-convolutional network or an encoder for feature extraction or image contraction and a decoder for image expansion and segmentation. Its key advantage is that it performs the tasks demanded with fewer training images and gives high precise segmentation results using a GPU.

The proposed SHPT-Net architecture is applied based on the default U-Net architecture. It consists of three sections: (i) a down-sampling or encoding section, (ii) an up-sampling or decoding section, and (iii) the bottleneck section, as shown in Fig. 5. The encoding section has four contraction or encoding blocks. Each block has input with two convolutional layers, kernel size (3 × 3), Relu activation function, stride equal to 1, kernel initializer technique used is " he_normal." The padding is "same" followed by the dropout layer and max-pooling layer with the kernel (2 × 2), where the number of feature maps gets double at each block.

The up-sampling or decoding has four expansion blocks; each block has de-convolutional layers with kernel size (2 × 2), the stride is (2 × 2), and padding "same". This is followed by two convolutional layers as in the encoding path with Relu activation function, the stride is 1, kernel initializer technique is "he_normal" and padding is "same". Then, it is followed by the dropout layer. The number of output filters for the four contraction or encoding blocks' convolution layers is (16, 32, 64, 128). The bottleneck section consists of two convolutional layers with filter numbers equal to 256 with a dropout layer. The two convolutional layers reduce the number of feature maps from the concatenation of de-convolutional feature maps with feature maps from the encoding path. We use a convolutional layer with kernel 1 × 1 and a sigmoid activation function for our binary segmentation process for the output layer.

4.6 Transfer learning

Transfer learning is a pre-training and fine-tuning concept [45] that initializes networks with pre-trained parameters rather than randomly set parameters. Its advantages can speed up the learning process and increase the generalization capability. It is also very common in models based on CNNs [34]. As a result, the proposed RESU-Net architecture uses ResNet pre-trained model blocks for encoder and decoder layers.
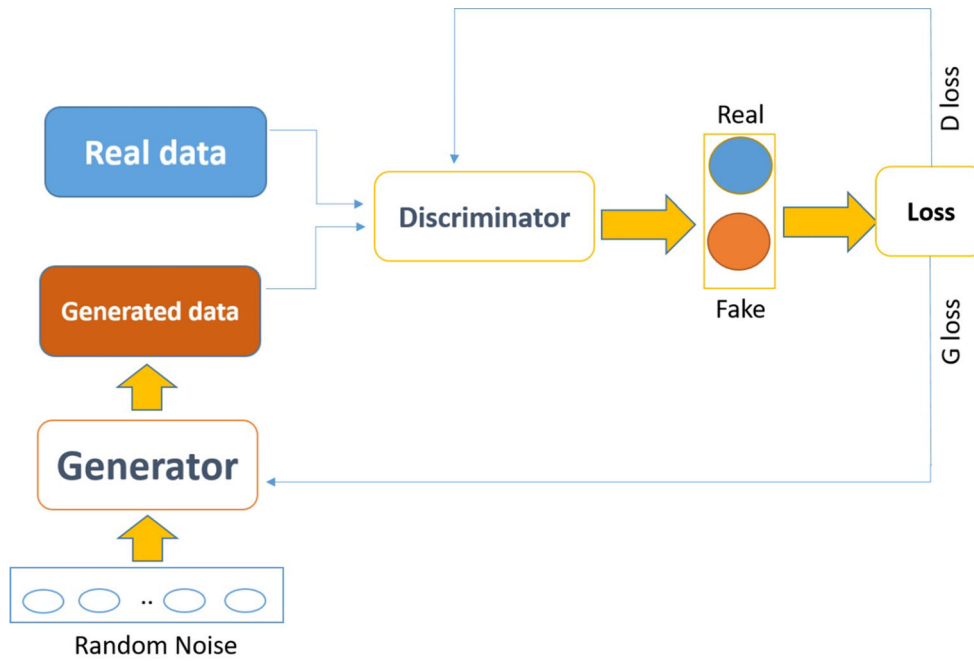
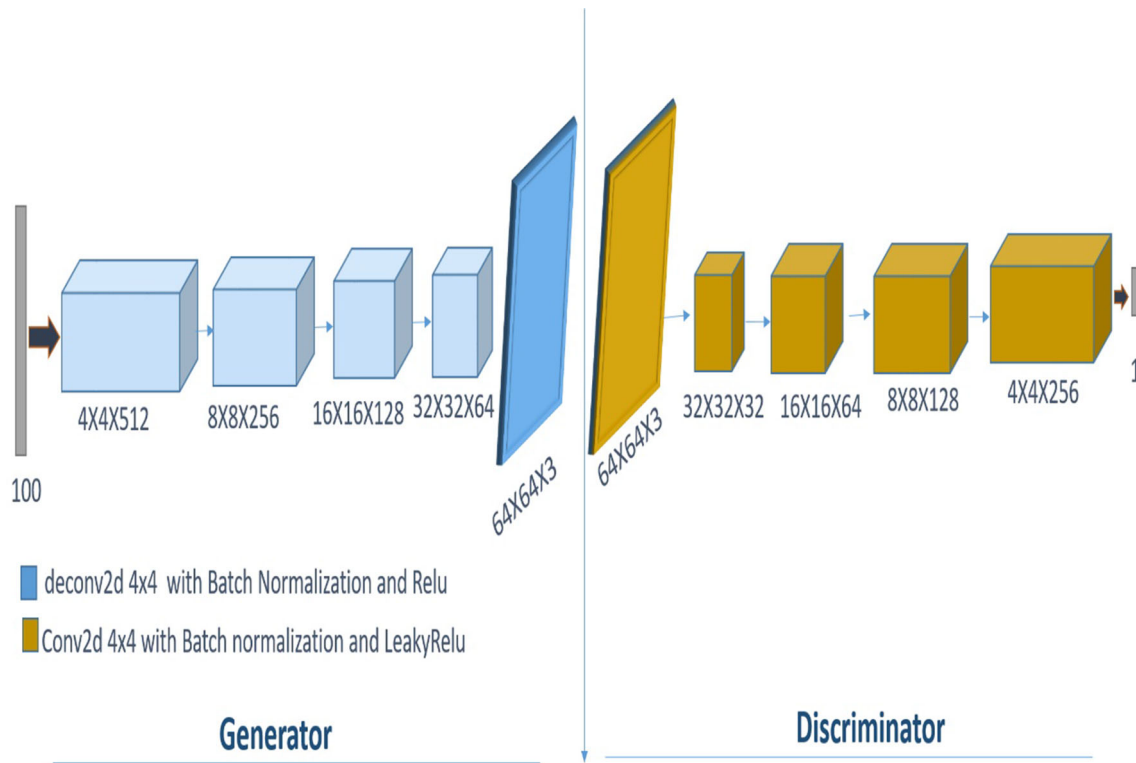**Fig. 7** The workflow for generator and discriminator processes

Fig. 8 The general structure of the applied DCGAN

4.6.1 ResNet pre-trained model

ResNet is a short name for a residual network which used as an image classification model. ResNet is based on CNN

architecture and consists of residual block series (Res-Block) identified with skip connections that separate the ResNet network from other CNNs. It is pre-trained on the large classification ImageNet dataset that includes over
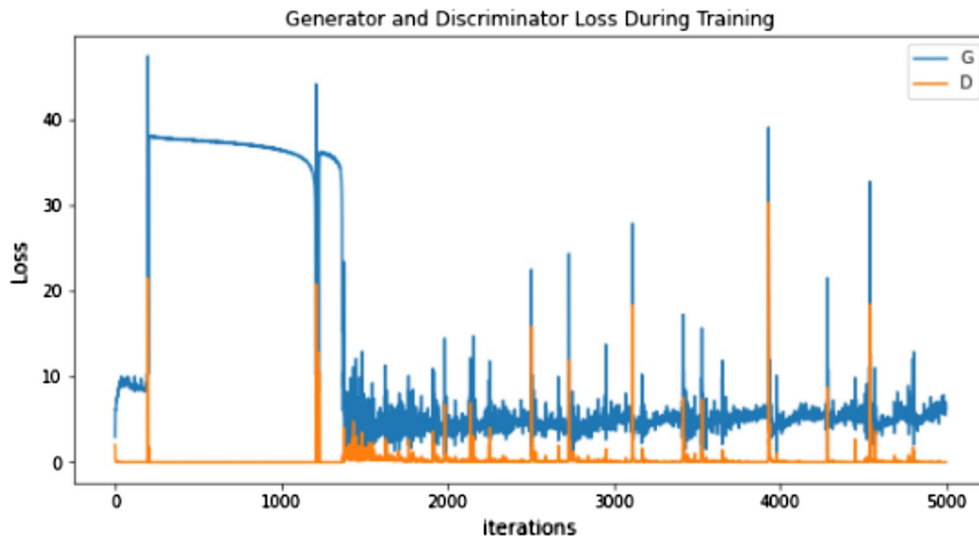
Fig. 9 The loss of the generator and discriminator during 5000 epochs of training

100,000 images belonging to 200 classes. Therefore, it is fine-tuned and used its blocks as encoder and decoder in the U-Net architecture for feature extraction and contraction processes to get the advantages of pre-trained weights.

Each ResBlock has two connections from its input called identity, cross or skip connections, in which the first connection from a series of convolutions and a batch normalization with linear functions. The other connection is skipping over that the series of convolutions and functions. The outputs of both connections are added together, as shown in Fig. 10.

Figure 11 describes the proposed RESU-Net architecture. The encoding section has an input block that contains a convolutional layer with kernel size (3 × 3), stride equal to 1, and padding is "same". Also, it contains Conv-block and convolutional layer with kernel size (1 × 1), stride 1, and "same" padding with Relu activation function followed by batch normalization layer. Four residual blocks follow the input block. Each residual block consists of two Conv-block with kernel size 3 × 3 and a convolutional layer with kernel 1 × 1, all with stride equal to 1 and padding "same" ended by batch normalization layer. The bottleneck section has two Conv-block. The decoder section has four blocks, and each block consists of an up-sampling concatenation layer with the kernel (2 × 2) and a residual block. The output block is a convolutional layer with kernel size (1 × 1) and a sigmoid activation function for the binary segmentation process. The Conv-block has only a convolutional layer and a batch normalization layer.

5 Experimental results and evaluation

5.1 Performance metrics

5.1.1 Dice similarity coefficient (DSC)

DSC is a statistical indicator that measures the similarity of two samples [46]. If the Dice coefficient equals 1, it signifies that the two samples we compare are exactly equal. Thus, the closer it to 1, is better. DSC is calculated by Eq. 4.

$$\text{DSC} = \frac{2TP}{2TP + FP + FN} \tag{4}$$

5.1.2 The accuracy (ACC)

ACC is the number of correct predictions to the total number of predictions. ACC is calculated by Eq. (5).

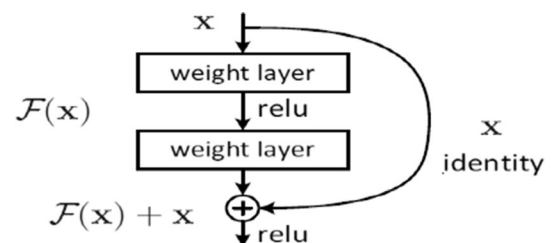$$\text{ACC} = \frac{TP + TN}{TP + TN + FP + FN} \tag{5}$$

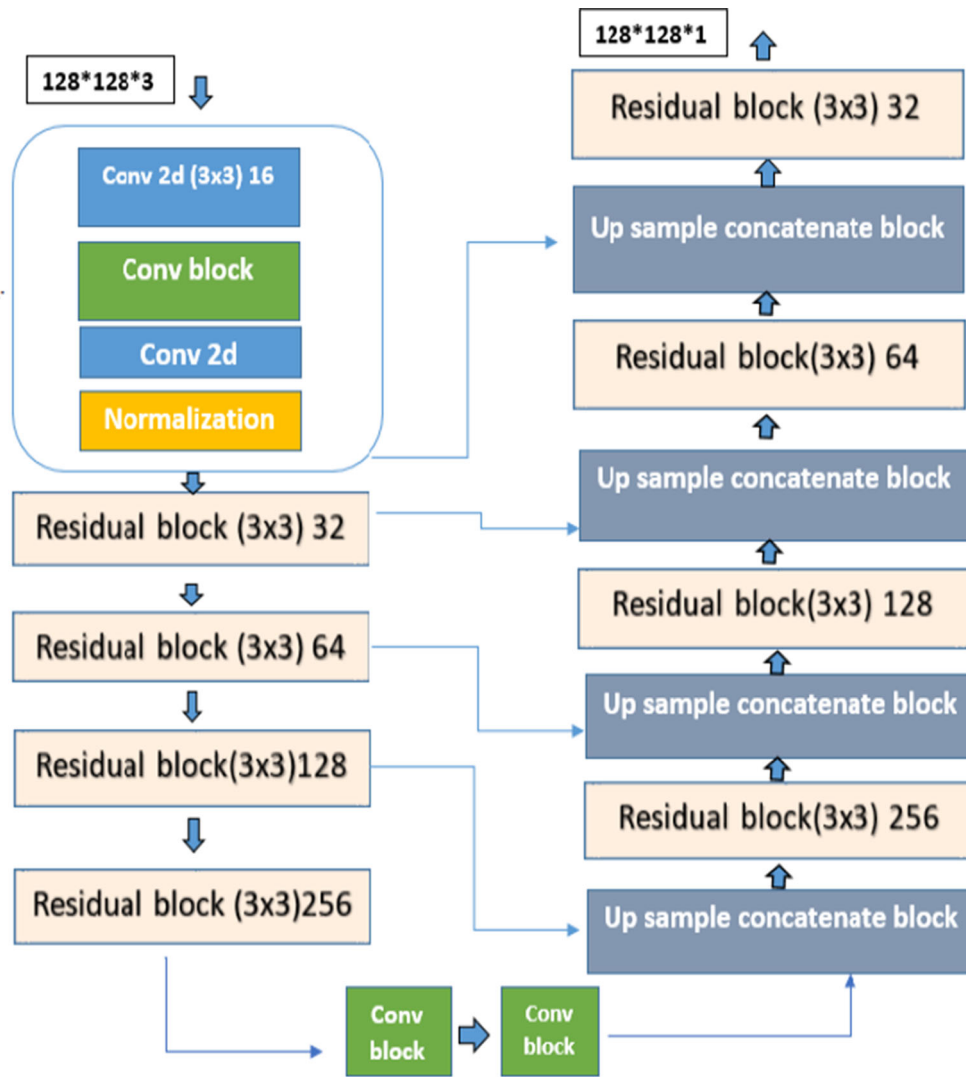**Fig. 10** The building block of the ResNet (ResBlock)

Fig. 11 The proposed RESU-Net architecture

5.1.3 True positive rate (sensitivity)

It is similar to the recall metric. It represents positive data-point proportion, which is correctly considered positive to all positive data points and calculated using Eq. (6).

$$\text{Sensitivity} = \frac{TP}{TP + FN} \tag{6}$$

5.1.4 False positive rate (1-specificity)

It is a negative data-point proportion incorrectly considered positive to all negative data points. It can be calculated using Eq. (7).

$$\text{False Positive Rate} = \frac{FP}{FP + TN}, \qquad \text{Specificity} = \frac{TN}{FP + TN} \tag{7}$$

TP, FP, FN, and TN indicate true positives, false positives, false negatives, and true negatives.

5.1.5 The receiver operating curve (ROC) and area under the curve (AUC)

It picks a good cut-off threshold for the model from plotting true positive rate (TPR) against false positive rate (FPR) for different threshold values in the range of [0, 1].

5.2 Experimental results

The proposed framework will be assessed by comparing its performance with the other state-of-the-art models. The evaluation is based on four performance metrics accuracy (ACC), Dice similarity coefficient (DSC), sensitivity, and specificity, as shown in Table 3.

Table 3 The comparison of our proposed architectures to other states of the art results

| Performance metrics	Accuracy	Sensitivity	Specificity	DSC (Dice similarity coefficient)
Shaken et al. [16]	N/A	N/A	N/A	89%
Hosseini et al. [7]	90.31%	91.18%	90.51%	N/A
Dong et al. [14]	89.75%	85.85%	87.34%	N/A
Kalavathi et al. [19]	55.83%	56.28%	51.11%	N/A
Dolz et al. [31]	N/A	N/A	N/A	92%
Alliouiet al. [32]	92.71%	94.43%	91.59%	N/A
Jingwenet al. [33]	N/A	N/A	N/A	0.9162 ± 0.023
Manhua et al. [34]	88.9%	N/A	N/A	87.0%
Chitradevi et al. [35]	95%	94%	93%	N/A
Carmo et al. [9]	N/A	N/A	N/A	0.9
Sammaneh et al. [36]	N/A	N/A	N/A	.88
Proposed SHPT-Net	94.34%	91.5%	92%	93.5%
Proposed RESU-Net	97%	95%	94.7%	94%

Fig. 12 A comparison between the proposed architectures (SHPT-Net and RESU-Net)
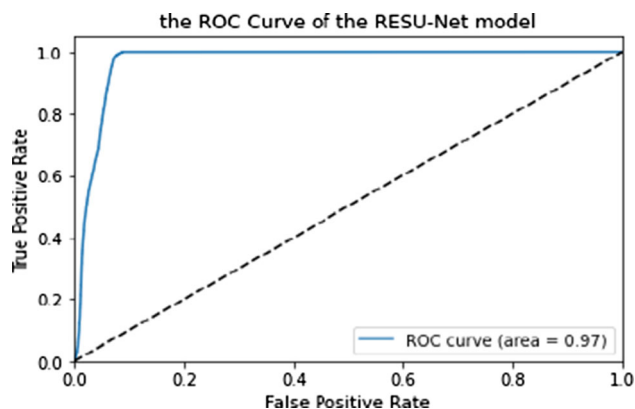
Fig. 13 The ROC curve for the RESU-Net model

Table 3 shows that the proposed RESU-Net architecture achieved the highest accuracy of 97%, better than other methods. LOA [35] obtained the second-highest accuracy of 95%. The authors used several optimization techniques such as GA, PSO, BAT, ABC but LOA gives high performance due to its characteristics of escaping from local optima. The proposed SHPT-Net achieved the third-highest accuracy of 94.34%. Quick fuzzy C means (FFCM) [19] gets the lowest accuracy of 55.83%, which depended on the contour-based brain segmentation method (CBSM) to remove the hug regions. Also, it is the lowest value in terms of sensitivity and specificity.

In terms of sensitivity, the proposed RESU-Net achieved the highest value of 95%. The 2.5 D MRI method [32] achieved the second-highest sensitivity of 94.43%, which utilizes 2.5D MRI depended on the deep learning U-Net architecture. The third value of sensitivity was achieved by Chitradevi et al. [35] of 94%. In terms of specificity, also the highest value for the proposed RESU-Net is 94.7%. The second value of specificity is achieved by LOA [35] of value 93%. The proposed SHPT-Net of 92% achieves the third-highest value of specificity.

In terms of DSC, the first and second-highest values for the two proposed methods are 94% for the RESU-Net and 93.5% for the SHPT-Net. The third value of the DSC is achieved by J. Dolz, C et al. [31] of 92%. The lowest value of DSC is 87.0% from the multi-task model [34], which combined hippocampal segmentation and AD classification in a model. According to the four performance metrics, the proposed RESU-Net architecture is the best. It depends on the deep learning U-Net architecture merged with applying the transfer learning concept.

Figure 12 presents a graphical description of the proposed architectures (SHPT-Net, RESU-Net) with each other. Also, when evaluating the RESU-Net by ROC curve metric, it gives a promising result, as shown in Fig. 13.

Figure 14 gives a graphical comparison among the proposed architectures (RESU-Net, SHPT-Net) and other states of the art results.

The proposed framework has been tested by various MRI images for real patients uploaded and stored in the

Fig. 14 The comparison of our proposed architectures to other states of the art results
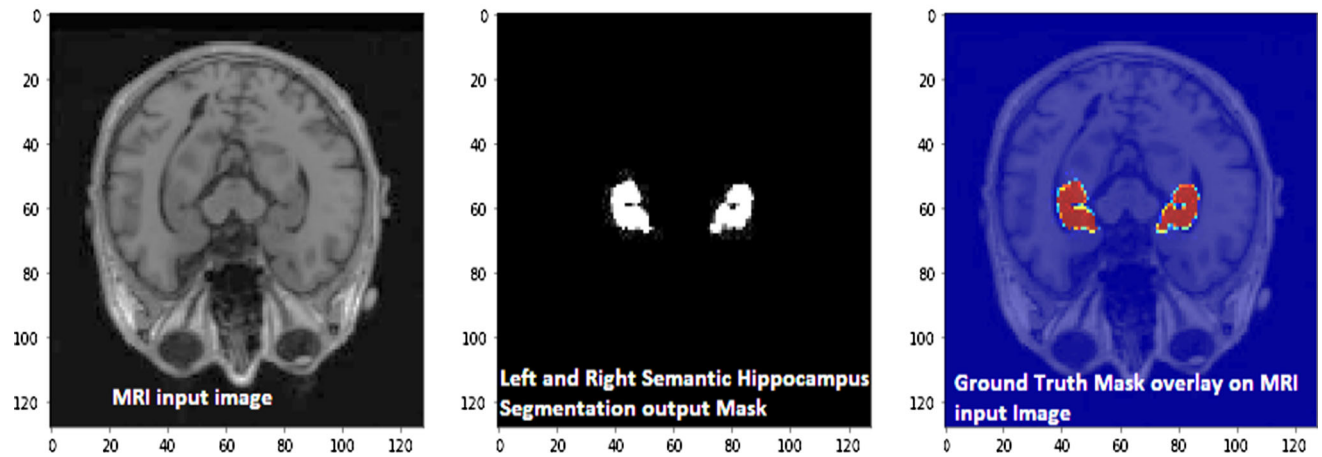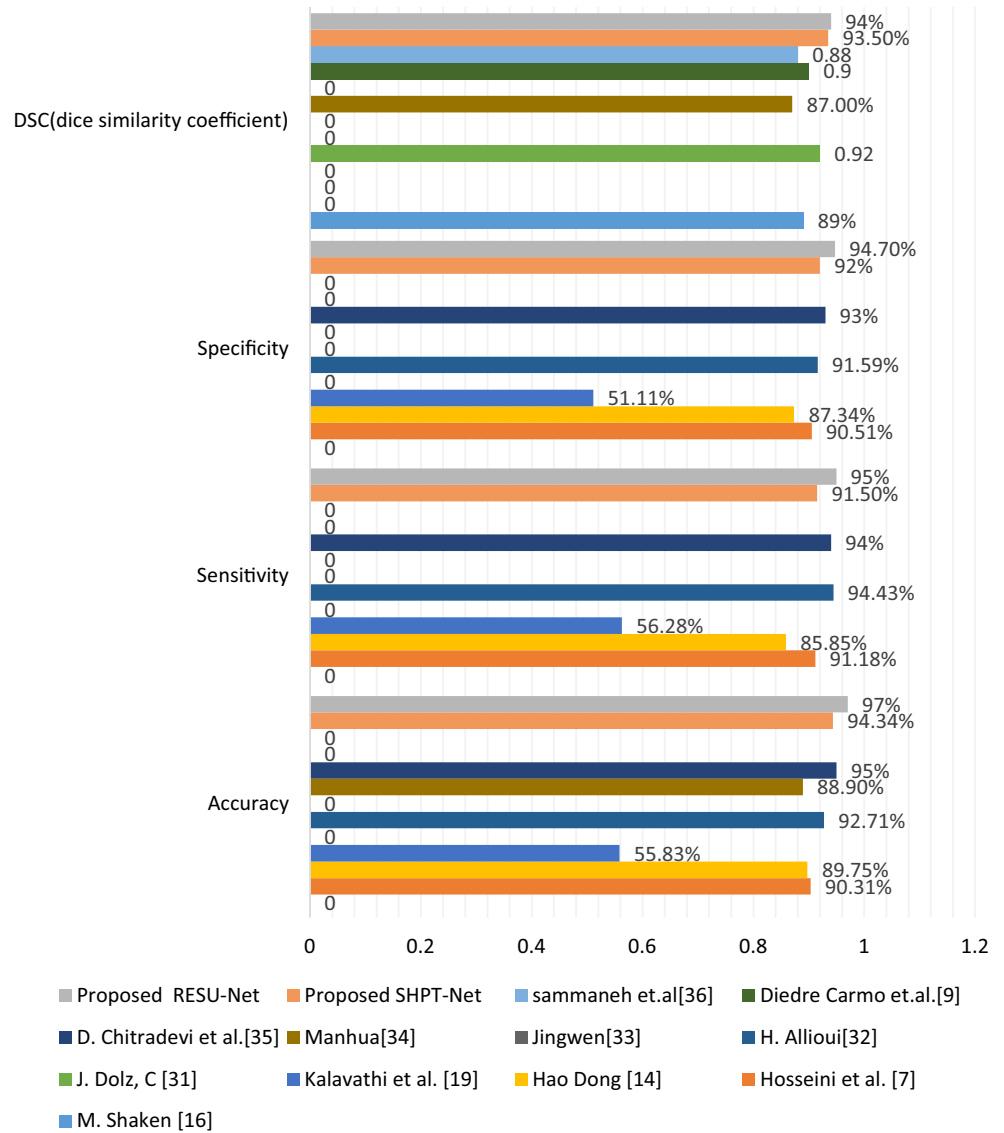

Fig. 15 Input image, the left and right semantic hippocampus segmentation output mask, and the ground truth mask overlay the input MRI image

ADNI dataset in the evaluation and testing step. The MRI input image, the left as well as the right hippocampus segmentation output image, and furthermore the left and right hippocampus segmentation that overlaid the input images are shown in Fig. 15.

6 Conclusion

The structural variations have been observed in the HC region, which assists in recognizing AD. Therefore, there is a necessity to analyze the microstructural variations, which is a very vital task. In this study, the DL-AHS framework for automatic left and right hippocampus segmentation and AD detection is proposed. The DL-AHS framework is based on the U-Net architecture and evaluated by coronal T1-weighted structural MRI data obtained from ADNI and NITRIC datasets. The dataset is processed using the MIPAV program and augmented by DCGAN. SHPT-Net and RESU-Net models are proposed. The empirical results prove that the proposed models are robust and accurate. They achieve high performance in accuracy, sensitivity, specificity, and Dice similarity coefficient over other state-of-the-art models. Also, the proposed framework is considered a clinical validation tool that analyzes the prognosis, the variation of the HC region, and the finding of AD. Then, it would be useful to identify and track the progress of mental impairment. The source code and the dataset are available at https://github.com/hadeerhelaly/hippocampus-segmentation (will be publicly opened after the acceptance of the manuscript).

7 Discussion

The current study uses the transfer learning concept that applies the ResNet blocks in the encoder and decoder layers of the U-net architecture. It is planned that applying other pre-trained models such as sufficient net and check the performance. Also, the dataset is augmented by a DCGAN. It is expected that using another type of GAN will give high performance. In addition, we aim to apply the same proposed models for multi-segmentation of MRI brain features of Alzheimer's disease.

Declarations

Conflict of interest The authors declare that they have no conflict of interest.

References

1. Li H, Habes M, Wolk DA, Fan Y (2019) A deep learning model for early prediction of Alzheimer's disease dementia based on hippocampal magnetic resonance imaging data. Alzheimer's Dement 15(8):1059–1070
2. Sevigny J et al (2016) The antibody aducanumab reduces Aβ plaques in Alzheimer's disease. Nature 537(7618):50–56
3. Wen J et al (2020) Convolutional neural networks for classification of Alzheimer's disease: Overview and reproducible evaluation. Med Image Anal 63:101694
4. Jain R, Aggarwal A, Kumar V (2021) Chapter 1—A review of deep learning-based disease detection in Alzheimer's patients. In: Jude HD (ed) Handbook of decision support systems for neurological disorders. Academic Press, pp 1–19
5. Ding J, Kong W, Mou X, Wang S (2019) Construction of a transcriptional regulatory network of Alzheimer's disease based on PANDA Algorithm. Interdiscip Sci Comput Life Sci 11(2):226–236
6. Yang F et al (2020) Identification of key regulatory genes and pathways in the prefrontal cortex of Alzheimer's disease. Interdiscip Sci Comput Life Sci 12(1):90–98
7. Hosseini-Asl E, Gimel'farb G, El-Baz A (2016) Alzheimer's disease diagnostics by a deeply supervised adaptable 3Dconvolutional network 502
8. Nadal L et al (2020) Differential annualized rates of hippocampal subfields atrophy in aging and future Alzheimer's clinical syndrome. Neurobiol Aging 90:75–83
9. Carmo D, Silva B, Yasuda C, Rittner L, Lotufo R (2021) Hippocampus segmentation on epilepsy and Alzheimer's disease studies with multiple convolutional neural networks. Heliyon 7(2):e06226
10. Andersen P, Morris R, Amaral D, Bliss T, O'Keefe J (2006) The hippocampus book. Oxford University Press
11. Petersen RC et al (2010) Alzheimer's disease neuroimaging initiative (ADNI). Neurology 74(3):201–209
12. Wang L et al (2003) Changes in hippocampal volume and shape across time distinguish dementia of the Alzheimer type from healthy aging. Neuroimage 20(2):667–682
13. Duraisamy B, Shanmugam JV, Annamalai J (2019) Alzheimer disease detection from structural MR images using FCM based weighted probabilistic neural network. Brain Imaging Behav 13(1):87–110
14. Dong H, Yang G, Liu F, Mo Y, Guo Y (2017) Automatic brain tumor detection and segmentation using U-Net based fully convolutional networks. In: Annual conference on medical image understanding and analysis, pp 506–517
15. Ghosh S, Das N, Das I, Maulik U (2019) Understanding deep learning techniques for image segmentation. ACM Comput Surv 52(4):1–58
16. Shaken M et al (2016) Sub-cortical brain structure segmentation using F-CNN'S. Proc—Int Symp Biomed Imaging 2016:269–272
17. Seo H et al (2020) Machine learning techniques for biomedical image segmentation: an overview of technical aspects and introduction to state-of-art applications. Med Phys 47(5):e148–e167
18. Singh SP, Wang L, Gupta S, Goli H, Padmanabhan P, Gulyás B et al (2020) 3D deep learning on medical images: a review. Sensors 20(18):5097
19. Kalavathi P, Christy AAA, Priya T (2017) Detection of Alzheimer disease in MR brain images using FFCM method. Comput Methods, Commun Tech Inf 140–144
20. Biju KS, Alfa SS, Lal K, Antony A, Akhil MK (2017) Alzheimer's detection based on segmentation of MRI image. Procedia Comput Sci 115:474–481

21. Yamanakkanavar N, Choi JY, Lee B (2020) MRI segmentation and classification of the human brain using deep learning for diagnosis of Alzheimer's disease: a survey. Sensors (Switzerland) 20(11):1–31

22. Shen D, Wu G, Suk H-I (2017) Deep learning in medical image analysis. Annu Rev Biomed Eng 19:221–248

23. Litjens G et al (2017) A survey on deep learning in medical image analysis. Med Image Anal 42:60–88

24. Li F, Tran L, Thung K-H, Ji S, Shen D, Li J (2015) A robust deep model for improved classification of AD/MCI patients. IEEE J Biomed Heal Inf 19(5):1610–1616

25. Zeng N, Zhang H, Song B, Liu W, Li Y, Dobaie AM (2018) Facial expression recognition via learning deep sparse autoencoders. Neurocomputing 273:643–649

26. Lin W et al (2018) Convolutional neural networks-based MRI image analysis for the Alzheimer's disease prediction from mild cognitive impairment. Front Neurosci 12:777

27. Chen Y (2019) Intelligent systems reference library 171 deep learning in healthcare

28. Noh H, Hong S, Han B (2015) Learning deconvolution network for semantic segmentation. In: Proceedings of the IEEE international conference on computer vision, pp 1520–1528

29. Lguensat R, Sun M, Fablet R, Tandeo P, Mason E, Chen G (2018) EddyNet: a deep neural network for pixel-wise classification of oceanic eddies. In: IGARSS 2018–2018 IEEE international geoscience and remote sensing symposium, pp 1764–1767

30. Ronneberger O, Fischer P, Brox T (2015) U-net: Convolutional networks for biomedical image segmentation. Lect. Notes Comput. Sci. (includingSubser. Lect. Notes Artif. Intell. Lect. Notes Bioinformat) 9351:234–241

31. Dolz J, Desrosiers C, Ben Ayed I (2018) 3D fully convolutional networks for subcortical segmentation in MRI: a large-scale study. Neuroimage 170:456–470

32. Allioui H, Sadgal M, Elfazziki A (2019) Deep MRI segmentation: A convolutional method applied to Alzheimer's disease detection. Int J Adv Comput Sci Appl 10(11):365–371

33. Sun J, Yan S, Song C, Han B (2020) Dual-functional neural network for bilateral hippocampi segmentation and diagnosis of Alzheimer's disease. Int J Comput Assist Radiol Surg 15(3):445–455

34. Liu M et al (2020) A multi-model deep convolutional neural network for automatic hippocampus segmentation and classification in Alzheimer's disease. Neuroimage 208:116459

35. Chitradevi D, Prabha S, Prabhu AD (2020) Diagnosis of Alzheimer disease in MR brain images using optimization techniques. Neural Comput Appl 7:223–237

36. Nobakht S, Schaeffer M, Forkert ND, Nestor S, Black SE, Barber P (2021) Combined atlas and convolutional neural network-based segmentation of the hippocampus from MRI according to the ADNI harmonized protocol. Sensors 21(7):2427

37. Chitradevi D, Prabha S (2020) Analysis of brain sub-regions using optimization techniques and deep learning method in Alzheimer disease. Appl Soft Comput J 86:105857

38. Noor MBT, Zenia NZ, Kaiser MS, Al Mamun S, Mahmud M (2020) Application of deep learning in detecting neurological disorders from magnetic resonance images: a survey on the detection of Alzheimer's disease, Parkinson's disease, and schizophrenia. Brain Inf 7(1):1–21

39. Jo T, Nho K, Saykin AJ (2019) Deep learning in Alzheimer's disease: diagnostic classification and prognostic prediction using Neuroimaging data. Front Aging Neurosci 11:220

40. Liu X, Deng Z, Yang Y (2019) Recent progress in semantic image segmentation. Artif Intell Rev 52(2):1089–1106

41. Russakovsky O et al (2015) Imagenet large scale visual recognition challenge. Int J Comput Vis 115(3):211–252

42. Goodfellow IJ, Pouget-Abadie J, Mirza M, Xu B, Warde-Farley D. Generative adversarial nets, pp 1–9

43. Frid-Adar M, Diamant I, Klang E, Amitai M, Goldberger J, Greenspan H (2018) GAN-based synthetic medical image augmentation for increased CNN performance in liver lesion classification. Neurocomputing 321:321–331

44. Venu SK, Ravula S (2021) Evaluation of deep convolutional generative adversarial networks for data augmentation of chest x-ray images. Futur Internet 13(1):1–13

45. Deepak S, Ameer PM (2019) Brain tumor classification using deep CNN features via transfer learning. Comput Biol Med 111:103345

46. Eelbode T et al (2020) Optimization for medical image segmentation: theory and practice when evaluating with dice score or Jaccard index. IEEE Trans Med Imaging 39(11):3679–3690