# A Modified U-Net Based Framework for Automated Segmentation of Hippocampus Region in Brain MRI

## NOSHEEN SOHAIL[1] AND SYED MUHAMMAD ANWAR[2]
[1]Department of Computer Engineering, University of Engineering and Technology, Taxila, Taxila 47050, Pakistan
[2]Department of Software Engineering, University of Engineering and Technology, Taxila, Taxila 47050, Pakistan

Corresponding author: Syed Muhammad Anwar (s.anwar@uettaxila.edu.pk)

**ABSTRACT** An accurate localization of the brain anatomical structure for correct and reliable diagnostic strategies is of great concern in many bio-medical applications. Towards this end, manual or semi-automated delineation methods used are found to be time consuming. Herein, to address this problem, we present an enhanced model for automated segmentation of two neighboring small structures of the brain in the Hippocampus region i.e., anterior and posterior. Our aim is to improve the segmentation performance, where the proposed architecture captures contextual information in encoding path and enables precise localization by utilizing the decoding path in a symmetric way. In particular, our proposed methodology enhances the original U-Net architecture with 3-dimensional (3D) data processing and employs spatial elastic deformation. Further, we evaluated the segmentation performance using recursive U-Net for comparison. The effectiveness of different optimization strategies are evaluated on a publicly available data comprising of 3D magnetic resonance imaging volumes from mono-modal hippocampus region. Our experimental results demonstrate the robustness of the proposed model by using patch-based augmentation technique for hippocampal segmentation.

**INDEX TERMS** Augmentation, deep learning, localization, magnetic resonance imaging, segmentation, U-net.

## I. INTRODUCTION

A small archi-cortical brain structure, that manages short-term anecdotal and critical memory while depositing it into the long term memory, is known as hippocampus. The hippocampus region is also responsible for vocal-based and musical emotions as it forms a part of the temporal limbic system. Hence, in plain interpretation of voices and musical emotions, the amygdala is particularly involved and enables the hippocampus to process even more complex information. Further, it contributes towards decoding heterogeneous emotions related to music, thereby creating an alliance between memory and contextual information [1]. The human hippocampus can be termed as a folded component of archi-cortex tissue, which is continuous with the neo-cortex [2]. The human brain consists of two hippocampi regions- shaped like seahorses- commonly termed as left- and right-hippocampi. They are also termed as anterior and posterior hippocampus sub-fields. From an imaging perspective, the region shows very little contrast on structural magnetic resonance imaging MRI) scans, due to the existence of nearby anatomical structures such as the thalamus, caudate nucleus, and amygdala. These structures have similar intensity levels in MRI scans to those for the hippocampus region [3].

In neuroimaging, segmentation of the hippocampus region plays a vital role in an early prognosis of certain brain related abnormalities. The gray matter tissue of the temporal cortex- hippocampus, is known to be primarily effected in the very initial stages of Alzheimer's disease (AD). This could further transform to cognitive decline with increasing age. In particular, for diagnosis of AD, the segmentation of hippocampus region is significant and is the most affected part of the human brain. To this end, a noticeable reduction in the hippocampal volume (HV) is a marker for AD diagnosis. There are other specific abnormalities that initially appear in the pre-clinical stages of AD in the hippocampus region such as tau pathology or $\beta$-amyloidand [4]. Such significant loss in

---

The associate editor coordinating the review of this manuscript and approving it for publication was Seifedine Kadry.

the volume of hippocampus region during the developmental phases of AD corresponds to mental cerebral and emotional decline [5]. During cognitive tests correlates were found binding short-term memory (STM) loss to AD [8]. According to a recent study, the accuracy of hippocampal volumetry via differential diagnostic techniques in normal subjects and MCI/AD patients is improved to a greater extent by normalizing the hippocampus volume with total intracranial volume (TIV) [9].

For the past few years, medical image analysis contributes to one of the popular and active research field in machine learning domain [10]. There is a fair chance that it will be the area where patients might comfortably interact with a completely functioning, practical artificial intelligence systems in the near future [11]. Artificial intelligence (AI) provides great opportunities to assist radiologist in diagnosing and streamlining complicated patterns within images, interpreting physiological characteristics into genetics as well as predicting clinical treatment results and developing strategies for prognostic planning [12]. The data obtained using structural brain MRI could contribute to various medical applications when analyzed using deep learning methods. Particularly, convolutional neural networks (CNNs) have shown remarkably robust results in the segmentation of brain tissues, tumors, and lesions [13]. With the help of a single CNN-based model being extensively trained, different tissues in MR brain images can be segmented. This could be extended to other regions such as the skeletal muscle in chest MRIs, and the blood circulatory arteries in cardiac computed tomography (CT) angiogram. Therefore, CNN based models have performed robustly in tissue classification, and visualisation of physical structures of the human body [14]. One of the peculiarities of CNN-based deep learning models is to provide an overall solution that requires nominal feature extraction with greater generalization and explainability [16]. These are also capable of performing object oriented classification by determining those features that characterize entire image objects unlike conventional classification techniques.

In many biomedical image segmentation tasks, localization is required to be included in the desired output as well i.e., assigning each voxel to a label class. One of the challenges faced is accessing labelled training images in a large quantity. In usual clinical practice, however, either delineation is still done manually by subjective analysis or via some underlying electromedical devices. However, there are many factors that contribute towards this gap between scientific research development and its practical applicability in the clinical routines. This is because of the complications in implementing these methods for practical use. While computer aided diagnosis (CAD) based models are dependant primarily on precise parameter calibration, that is essential in various medical applications. On the other hand, the facility of maintaining different atlases and its registration itself is usually not available to healthcare professionals. Hence, these models cannot be practically utilized to complex clinical applications. Moreover, earlier developed technique of multiple atlases

took considerable time in segmenting a single image. To this end, there is a recent focus in research to achieve minimal computation time while performing these segmentation [17]. An accurate performance evaluation of automated methods in application of hippocampus segmentation still poses another obstacle for its practical implementation. In most cases ample number of validation volumes are not used for comparing with a reference segmentation. This is due to the difficulty and time consuming process of obtaining manual segmentation for a large set of volumes. Hence, only limited scientific algorithms have so far been tested on a few bio-medical equipment as well for different acquisition scenarios [18].

### A. OUR CONTRIBUTIONS

Herein, we focus on hippocampi segmentation of brain MRIs, where our developed deep learning network has been derived from the famous U-Net architecture proposed by Ronneberger *et al.* (a standard benchmark algorithm in medical segmentation domain) [19]. Moreover, a comparison of different training strategies has also been performed to precisely locate and segment the actual regions of interest by classifying each voxel in an MR image as non-hippocampus (background or surrounding), left hippocampus, or right hippocampus. We utilized object-based approach to annotate each image pixels with class labels. Our framework includes patch-based segmentation along with dense data augmentation technique, so that during training phase the input images could easily be visualized on multiple scales [16]. Hence we resolve the issue of over-fitting and lower the error rate thereby enhancing the performance of our proposed model. Our main contributions are:

- We present a complete in house framework (built in pytorch) for extensive data augmentation including mirroring, cropping, and spatial elastic transformation to the input data and make it compatible to the proposed deep network model.
- We train our models from scratch with optimization strategy for robustness, using combination of weighted multi-class dice loss and cross entropy loss formulation. Further, we used overlap tile method for test data segmentation and evaluation using six different quality metrics for each hippocampus sub-region.
- A significant performance was achieved in terms of accuracy, precision, and recall in hippocampus segmentation.

### B. RELATED WORK

An automated segmentation of hippocampus region has gained a lot of significance in scientific and research community. In this regard, different approaches have been proposed and can be categorized in: atlas-based [20], machine learning-based [21], active-contour models [22], [23], and deep learning frameworks.Some of the popular deep learning algorithms include deep Boltzmann machines, CNNs, stacked auto-encoders, and deep neural networks. [24]. Atlas-based
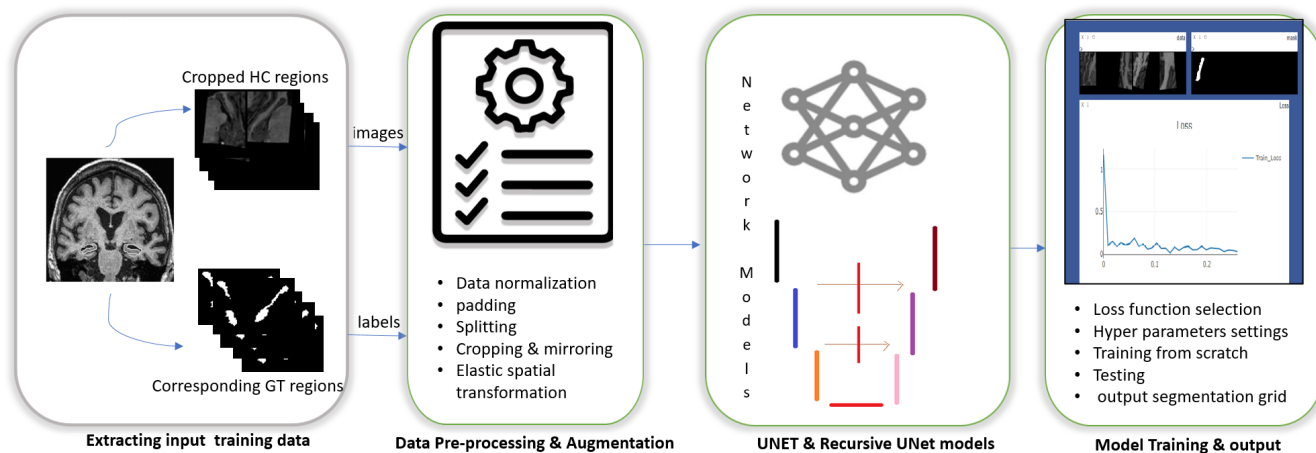
**FIGURE 1.** The overall sequence of our proposed methodology for hippocampus segmentation.

registration techniques have gained a significant level of popularity. First, various atlas-based images are registered to the unknown image, usually in a non-linear fashion. The manual segmentation masks are then transformed which further produces individual output segmentation against each of the used atlases. Once individual segmentation results are obtained, the final output is selected by concatenating a number of fusion algorithms [25], such as majority or average voting [26], use of global [27] or local weights [28], joint label fusion [29], and accuracy maps [30]. The active contour model (ACM) strategy evolves itself according to the intensities of the input images. The clear edge boundaries mainly affect the overall performance of the segmented object. To generate efficient segmentation results, ACMs mostly rely on prior information related to the shape of the structure for better evaluation of the contours [31]. In [23], the model used principal component analysis (PCA) for analysing shapes of both the hippocampus and the neighboring structures along with utilizing atlases being manually segmented. A fusion of multiple atlases based framework with ACM was presented and called 3D optimal local maps (OLMs) [32]. It was developed with enhanced multi-atlas concepts at voxel level, which locally controls the impact of each energy term of a hybrid ACM. In comparison to these conventional methods, different ML techniques for segmentation have been adopted [33], [34]. Usually in these methods, hand-crafted features were extracted from each training sample to make a training data. Further, classifiers were optimized towards generating the desired segmentation masks. In these methods, dictionaries and classifiers are learned simultaneously from a set of brain atlases, which can then be used for the reconstruction and segmentation of an unseen target image. The segmentation accuracy of such methods could be improved if more complicated classifiers are learned rather than linear classifiers [35]. Due to the rapidly increasing evolution in deep learning (DL) techniques for medical image analysis, convolutional neural networks are becoming a methodology

of choice [36] for tasks such as the segmentation of brain tissue [37], tumor [38] and lesions [39] anatomical structure segmentation for striatum [40], and for caudate and thalamus [41], [42]. Deep CNN-based sub-cortical nuclei of brain segmentation (including the hippocampus) has been also proposed by [43] and [44]. Additionally, CNNs have also shown remarkable performance for brain extraction [45] and full brain segmentation [46], [47].

It should be noted that for certain specific segmentation problems, various software platforms are available. For instance, for brain segmentation some examples include 'volBrain' by NITRC (Neuro Imaging Tools and Research Collaboratory) [48] and 3D Slicer [49]. However, it has not been possible so far to achieve the reported high levels of segmentation accuracy in actual clinical practice. Further, variations in the anatomical definitions as well as protocol specific issues could not be comprehensively addressed. In other words, we argue that such automatic methods may perform efficiently on user specific data, but may vary from that of the software algorithm developer on the basis of ground truth definitions. Moreover, recent automatic segmentation techniques incorporate expert prior knowledge of lesion appearance, anatomical shape, and other sophisticated high level features as model parameters. These methods are based on specific datasets that vary from the actual radiology imaging data acquired from various hospitals [50].

## II. PROPOSED METHODOLOGY
While in MR images, the exact localization of hippocampal structure is a critical and crucial task for correct diagnosis or treatment plannings. Moreover, human dependent or semi-automatic hippocampus region segmentation in 3D images is exhaustive and such methods have low reproducibility in clinical routines. Although, deep learning-based algorithms for automated segmentation have proven to be quite efficient and robust in the medical field [51], but there is still room for extensively evaluating its performance

on hippocampi segmentation tasks. Due to the very low intensity contrast of hippocampi structures with respect to the surrounding regions of the brain, the implementation of such frameworks is still a challenging task [18]. Towards this, we propose a deep learning-based method (shown in Figure 1), with details in the following subsections.

### A. DATA PRE-PROCESSING
The pre-processing steps included data normalization, padding, and splitting. A cropped image representing the hippocampal brain region was used. Towards this, a bounding box was created over the region of interest (ROI) training volumes and its corresponding labels respectively. The two vector matrices were merged to form the cropped images and their labels, which were then fed to the next step of data normalization. We independently normalized each volume modality of the cropped image region by subtracting the mean and divided with its standard deviation.

### B. DATA AUGMENTATION
In our proposed approach, we used random patches from the input images for training the deep learning models. The data were setup by extracting random patches from volumetric images and corresponding pixel label data. The patch-based technique was particularly useful in hardware systems where there is memory constraints while performing dense training from scratch with arbitrarily large volumetric images. In every pair of MR volumes and labels during training, we specified a patch size of $4 \times 4$ pixels, a mini-batch size of 64, and 16 patches per image from randomly positioned patches. Afterwards, data augmentation technique was applied, which comprised of transformations including mirroring and spatial transform. Hence, for each input, two transformed images were generated. All augmentation operations were applied on-the-fly with our own in-house built framework. We augmented the training and validation images by applying the transform operation to the random patches. In particular, we used random rotation and reflection for the input data to enable robust training. Further, we cropped the corresponding 2D random patches as per network size requirement.

### C. DEEP LEARNING MODELS
For segmentation we employed deep learning-based frameworks. First we used the proposed modified U-Net architecture as shown in Figure 2. We further used a recursive U-Net (Rec-UNet) model [52] to compare and validate the results. The training configuration included setting up the hyper parameters and optimizer settings as well as selection of the loss function. The Rec-UNet model comprised of two cascaded stages. In the first stage, a dense-UNet model was used to obtain the initial segmentation results. In the second stage, the resultant segmentation masks of the first stage have been used as prior knowledge by incorporating skip-connection blocks in the sub-module of the existing basic UNet model to obtain more accurate segmentation results. This implementation is done in a recursive way. It is therefore very easy to

configure the number of down-sampling steps. Also the type of normalization can be passed as a parameter as instance normalization. A dense Rec-UNet block consists of 'n' consecutive convolution layers, the cascaded block and transition layers for segmentation in MR images of the same resolution, each followed by a batch normalization (BN), rectified linear unit (ReLU), and dropout layers. The succeeding convolution layer takes the feature maps of all the previous layers as input. Our proposed modified UNet model follows the original encoder-decoder style UNet model [19], where layers such as convolution, up-convolution, and pooling (kernel size $2 \times 2$ and a stride of 1) are used. We also incorporated the skip connection for better outcomes. The filter sizes and layer configuration is shown in Figure 2.

### D. HIPPOCAMPUS SEGMENTATION
We performed test prediction on the segmented sequential batches and combined them to a hippocampal segmentation. The output against each testing MR images was compared with their corresponding ground truth images. For getting the final segmentation mask, the mean likelihood against each pixel has been computed from the values obtained at the output of softmax layer. The output predicted masks are stored separately against each network i.e., for U-Net and Recursive-UNet outputs.

## III. EXPERIMENTAL RESULTS AND DISCUSSION
### A. DATASET
We used a total of 263 3D mono modal MRI volumes of hippocampus head and body from the publicly available dataset of Medical Segmentation Decathlon competition MSD [53]. These 3D MRIs consist of axial IS (InterSpace) scans with a dimension of $34 \times 47 \times 40$ respectively, with an image spacing of 1$mm$ in each dimension. The structural data was acquired using an MPRAGE T1-weighted sequence with the following acquisition parameters: TI = 860ms TR = 8.0ms and TE = 3.7 ms, on a Philips Achieva scanner. The MSD challenge was organized by a number of teams and among several data contributors, the hippocampus data has been donated by Vanderbilt University for conceptual design and metrics committee. The data have been annotated and verified by human experts of their respective fields required for precise clinical use.

The segmentation result of our network after comparing with the ground truth labels has been calculated in terms of these evaluation metrics: mean dice score, accuracy, IOU index (Jaccard), positive predictive value (PPV), sensitivity, and specificity against foreground (non-hippocampus region), anterior hippocampus (left) and posterior hippocampus (right) and presented in Table 1. The performance of the resulting output segmentation mask has been evaluated by these quality metrics as described in [54].

### B. TRAINING OUR DL MODELS
The network training has been performed in a sliding window procedure that predicts each pixel class label to provide a local region (patch) around that pixel. The network gets
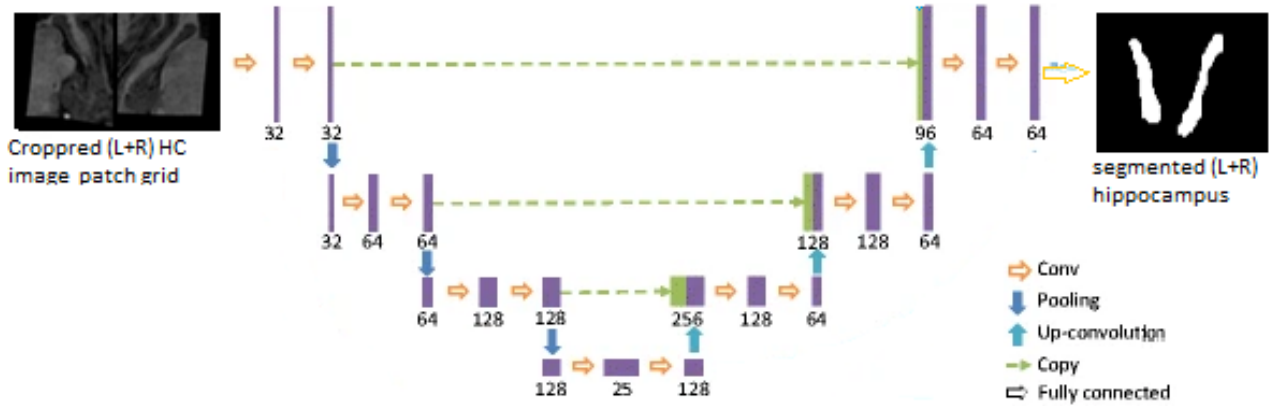
**FIGURE 2.** The overall sequence of our proposed methodology for hippocampus segmentation.

localized and the training data patches are much larger than the number of training images. The localization accuracy and the use of context has an inverse relationship. A higher number of patches need to process more max-pooling layers, thereby reducing the localization accuracy. On the other hand, for smaller patches, the network relaxes to process for little contextual data. However, right selection of both trade-offs i.e., localization and context leads to even better results when used at the same time.

A training from scratch with up to 20 epochs per network was performed on a single CPU (core i − 7, 6GB RAM). The test segmentation of the MRI volume ($34 \times 47 \times 40$) took about 50s for the 2D axial image. Table 2 presents a comparison of various most recent techniques of hippocampus subfield segmentation with our proposed scheme on the basis of Dice Similarity Coefficient and Jaccard Index, for which our model showed considerably better performance. The best results against each metric are highlighted in bold.

In particular, the nested dilation network (NDN) [15], residual blocks were nested with dilationsfor the segmentation tasks using CT, MRI, and endoscopic images. The data for the hippocampus segmentation task was taken from MSD as in our proposed method. Cao *et al.* [55] employed a 3D-Unet in multitask deep learning for joint hippocampus segmentation and clinical score regression.The authors evaluated their method on 407 subjects with MRI data from baseline Alzheimer's Disease Neuroimaging Initiative (ADNI). In [56], the authors proposed a combination of U-Seg-Net and Ensemble-Net framework of 110 healthy subjects from the ADNI. A multi-view ensemble approach that relies on neural networks to combine multiple decision maps for hippocampus segmentation was explored. In [25], segmentation masks were generated using an ensemble of three independent models, operating with orthogonal slices of the input volume, while erroneous labels were subsequently corrected by a combination of replace and refine networks. Experiments were performed on MICCAI dataset, achieving a mean

Dice value of 0.88 through transfer learning from the larger EADC-ADNI data. In comparison to these ensemble methods which are bound to be computationally expensive our proposed method shows considerably better results for both hippocampus regions.

### C. EVALUATION PARAMETERS

In this section we are presenting evaluation metrics for validation of each phase of the proposed system. All metrics can be used either by passing test and reference segmentations as parameters or by passing a *confusion matrix* object. The later is useful when many metrics need to be computed, because the relevant computations are only done once. All metrics assume binary segmentation inputs. Confusion matrix returns four integer values for true positives, false positives, true negatives and false negatives. For hippocampal region localization and detection we used greedy overlapping criteria of ground truth box and predicted box, known as intersection-over-union (IoU) or Jaccard index coefficient. The correctly predicted box is known as true positive, else false positive, and is computed as follows:

$$IoU(Jaccard) = 2 \times \frac{TP}{TP + FN + FP}. \quad (1)$$

For performance evaluation of segmentation phase, we considered the dice score (Dice), Jaccard coefficient (Jc), pixel level specificity (SP), pixel level sensitivity (SE), and pixel level accuracy (Ac) as the evaluation measures.

$$Dice = 2 \times \frac{2TP}{FN + 2TP + FP}, \quad (2)$$

$$SP = \frac{TP}{TP + FP}, \quad (3)$$

$$SE = \frac{TP}{TP + FN}, \quad (4)$$

$$Accuracy = \frac{TP + TN}{TP + FP + FN + TN}, \quad (5)$$

**TABLE 1.** Output segmentation results as performance metrics evaluated for the proposed models (Basic UNet & Recursive UNet).

| No. | Evaluation Metrics | Proposed UNet | | | Rec-UNet | | |
|---|---|---|---|---|---|---|---|
| | | Anterior | Posterior | Mean | Anterior | Posterior | Mean |
| 1. | Dice | 0.89 | 0.88 | 0.88 | 0.85 | 0.82 | 0.83 |
| 2. | Jaccard (IoU) | 0.80 | 0.77 | 0.78 | 0.74 | 0.69 | 0.83 |
| 3. | Sensitivity/Recall | 0.88 | 0.86 | 0.87 | 0.81 | 0.74 | 0.78 |
| 4. | Specificity/True Negative Rate | 0.99 | 0.99 | 0.99 | 0.99 | 0.99 | 0.99 |
| 5. | Precision/PPV | 0.90 | 0.88 | 0.89 | 0.90 | 0.92 | 0.908 |
| 6. | Accuracy | 0.997 | 0.998 | 0.997 | 0.996 | 0.996 | 0.996 |
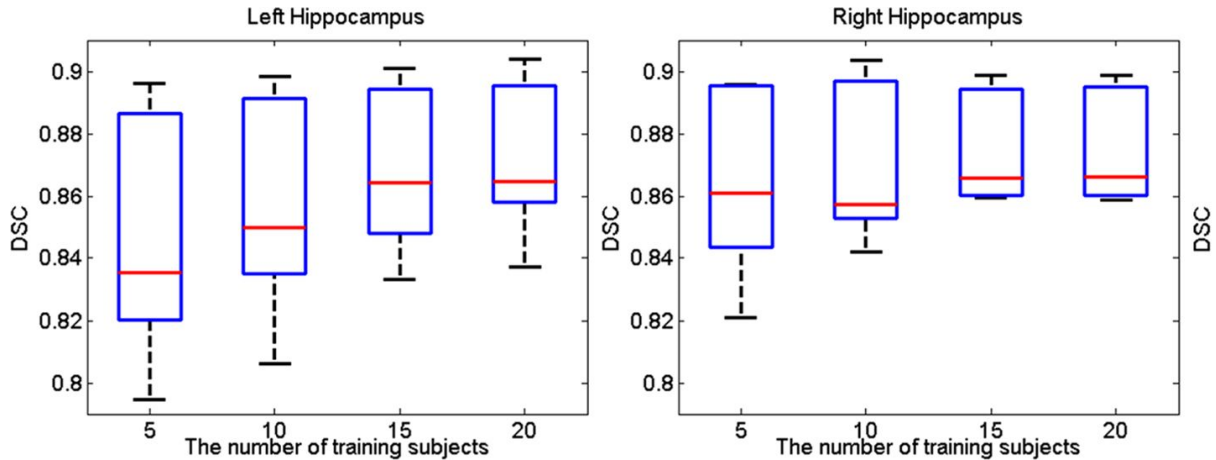


**FIGURE 3.** Dice plot - left & right hippocampus segmentation for randomly chosen training subjects.

**TABLE 2.** Comparative analysis of results between the proposed model and benchmark algorithms based on hippocampus segmentation.

| Methods | Dice | | Jaccard | |
|---|---|---|---|---|
| | Anterior | Posterior | Anterior | Posterior |
| NDN [15] | 0.87 | 0.88 | 0.78 | 0.79 |
| 3D-UNet [55] | 0.87 | 0.85 | 0.78 | 0.79 |
| U-Seg Ensemble Net [56] | 0.86 | 0.88 | 0.75 | 0.74 |
| Fast & Precise Net [25] | 0.88 | 0.88 | 0.79 | 0.79 |
| **Proposed (Ours)** | **0.89** | **0.88** | **0.80** | 0.77 |

where TP, TN, FP, and FN represents the number of true positive pixels, true negative pixels, false positive pixels, and false negative pixels, respectively.

We chose to implement both basic U-Net and Recursive U-Net network architectures and calculated statistical parameters across all cross folds validation. The network performed efficiently after being trained from 3D mono-modal axial slices, achieving a mean dice score of 0.885±0.01, IOU (Jaccard) of 0.796±0.02 and an accuracy of 0.996 ±0.01 as shown in Table 1. The training loss reduced to 0.0072 and validation loss minimizes to 0.012 at final iteration. We have implemented and trained the network from scratch on Pytorch 1.4.0 (python 3.6) platform.

Figure 3 represents the average of Dice Similarity Coefficient (DSC) results for the left and the right hippocampus regions for 5th, 10th, 15th and 20th training samples. The results show that mean Dice scores vary with both, the number of training subjects as well for left and right hippocampus regions respectively. It can be observed that

with the increase in number of training samples, the DSC scores tend to improve randomly, however, DSC variations in left (anterior) hippocampus region is slightly at higher end as compared to the right (posterior) hippocampus. The results indicates that our methods using two-side hippocampus segmentation strategy can achieve stable and accurate prediction. The final qualitative segmentation results for left and right hippocampus regions against randomly chosen input source volumes can be seen from Figure 4. The source at (a) has been chosen as grid of four (two pairs) both having left and right respectively, alongwith their corresponding ground truth labels in (b). Both implementation models i.e., basic UNet and Recursive UNet (Rec-UNet) results can be seen with argmax segmentation in respect of the input volumes. The results in terms of all performance metrics for the proposed modified UNet model slightly tends be at the higher side as compared to recursive implementation of our model. The noisy labels as artifacts are observed in many patients, which presented a significant problem during the evaluation of the segmentation. Because of such problematic slices present in the dataset, it was difficult for the proposed method to adequately handle such situations. Secondly, the segmentation of low contrast regions in the hippocampal volumetry has been more challenging due to close contact with the surrounding complex tissues of brain. However, the results for both network models showed considerably better performance when compared with benchmark algorithms in terms of accuracy, sensitivity, specificity, and dice score.
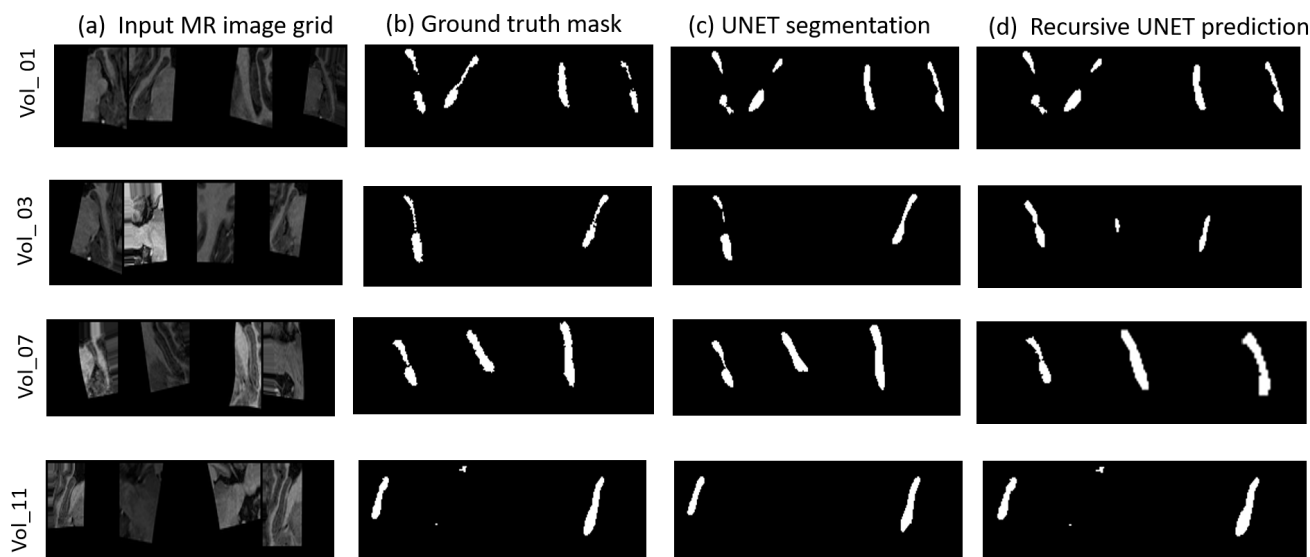
**FIGURE 4.** The qualitative results of MRI hippocampus segmentation (Left+Right). (a) indicates a random input images grid, (b) is the corresponding ground truth mask, (c) shows the output prediction on UNet and (d) indicates the corresponding Rec-UNet prediction with argmax segmentation of left and right hippocampus regions.

## D. DISCUSSION

Multiple atlas based methods perform efficiently to anatomical changes. However, the quality of image registration plays a critical role in its overall performance. Moreover, the effective time being consumed during segmentation has a direct relationship with registration i.e., it increases with the total number of registrations being performed. This segmentation strategy is based on image reconstruction, which is in contrast to the atlas-based labeling approaches that rely on comparing image similarities between atlases and target images. It may take several days to learn very good representative dictionaries and optimal discriminating classifiers offline. While deep neural networks provide superior performance as compared with conventional machine learning algorithms because of its ability to optimally use data representation learning in respect of various related tasks [57]. This technique of representation learning serves to be a significant feature in CNNs. Unlike traditional ML approaches, deep learning via CNNs resolves data computational problems by applying data representations strategies in very simpler ways [58]. Due to reduced hardware computational requirements and processing speed, the use of 2D CNNs still prevails effectively in medical research domain even applied on 3D brain image volumes as well [51]. Recent approaches [59], [60] also contribute towards 3D CNNs for brain image segmentation. In the current neuro-imaging studies where large sample sizes are required, CNN-based algorithms have proved its efficiency and robustness [61].

Herein, we have used a modified version of U-net architecture, hence deploying deep learning for medical image segmentation task. One of the biggest challenge in segmenting the hippocampus region includes the small anatomical structure and variation in the shape of the left and right hippocampus regions. In particular, our proposed method has a significant performance in all major evaluation parameters (Table 1). Further, the results are also significant when compared with other methods presented in literature (Table 2). Hence, we have shown that deep learning can be successfully used for challenging medical image segmentation tasks.

## IV. CONCLUSION

In this work, we develop a robust automated segmentation method for hippocampus sub-regions in an MRI based dataset, by using the benchmark algorithm of [19] for bio-medical imaging segmentation task. We trained the network from scratch using 3D mono-modal hippocampus dataset with the technique of data augmentation, spatial elastic deformation and a loss function chosen to be the union of cross-entropy and weighted multi-class dice-loss formulation. The dataset is sufficient for the evaluation of a deep learning based approach leading to promising results. We performed our model testing on the basis of six different evaluation metrics, after a dense training procedure we obtained significant results for all these metrics. In particular, the segmentation accuracy has been very high for both anterior and posterior regions. Various segmentation approaches have been proposed for brain region including tasks such as tissue classification or anatomical structure segmentation of hippocampus region. In comparison our proposed method is found to be effective in segmenting hippocampus region which is evident from our results. In addition, our proposed modified architecture can be utilized effectively for multiple bio-medical image analysis tasks.

## REFERENCES

[1] S. Fráhholz, W. Trost, and D. Grandjean, "The role of the medial temporal limbic system in processing emotions in voice and music," *Prog. Neurobiol.*, vol. 123, pp. 1–17, Dec. 2014.

[2] H. Duvernoy, *The Human Hippocampus: Functional Anatomy, Vascularization and Serial Sections With MRI*. Berlin, Germany: Springer, 2005.

[3] B. Fischl, D. Salat, E. Busa, M. Albert, and M. Dieterich, ''Whole brain segmentation: Automated labeling of neuroanatomical structures in the human brain,'' *Neuron*, vol. 33, no. 3, pp. 341–355, 2002.

[4] L. Serra, M. Bozzali, L. Fadda, M. S. De Simone, M. Bruschini, R. Perri, C. Caltagirone, and G. A. Carlesimo, ''The role of hippocampus in the retrieval of autobiographical memories in patients with amnestic mild cognitive impairment due to Alzheimer's disease,'' *J. Neuropsychol.*, vol. 14, no. 1, pp. 46–68, Mar. 2020.

[5] T. Joko, S. Washizuka, D. Sasayama, S. Inuzuka, T. Ogihara, T. Yasaki, T. Hagiwara, N. Sugiyama, T. Takahashi, T. Kaneko, T. Hanihara, and N. Amano, ''Patterns of hippocampal atrophy differ among Alzheimer's disease, amnestic mild cognitive impairment, and late-life depression,'' *Psychogeriatrics*, vol. 16, no. 6, pp. 355–361, Nov. 2016.

[6] I. O. Korolev, L. L. Symonds, and A. C. Bozoki, ''Predicting progression from mild cognitive impairment to Alzheimer's dementia using clinical, MRI, and plasma biomarkers via probabilistic pattern classification,'' *PLoS ONE*, vol. 11, no. 2, Feb. 2016, Art. no. e0138866.

[7] M. J. Lazarczyk, P. R. Hof, C. Bouras, and P. Giannakopoulos, ''Preclinical Alzheimer disease: Identification of cases at risk among cognitively intact older individuals,'' *BMC Med.*, vol. 10, no. 1, pp. 1–13, Dec. 2012.

[8] M. C. V. Hernández, R. Clark, S.-H. Wang, F. Guazzo, C. Calia, V. Pattan, J. Starr, S. D. Sala, and M. A. Parra, ''The striatum, the hippocampus, and short-term memory binding: Volumetric analysis of the subcortical grey matter's role in mild cognitive impairment,'' *NeuroImage, Clin.*, vol. 25, 2020, Art. no. 102158.

[9] S. Estévez-Santé and A. Jiménez-Huete, ''Comparative analysis of methods of volume adjustment in hippocampal volumetry for the diagnosis of Alzheimer disease,'' *J. Neuroradiol.*, vol. 47, no. 2, pp. 161–165, Mar. 2020.

[10] S. Yousaf, H. RaviPrakash, S. M. Anwar, N. Sohail, and U. Bagci, ''State-of-the-art in brain tumor segmentation and current challenges,'' in *Machine Learning in Clinical Neuroimaging and Radiogenomics in Neuro-Oncology*. Cham, Switzerland: Springer, 2020, pp. 189–198.

[11] S. Anwar, ''AIM and explainable methods in medical imaging and diagnostics,'' in *Artificial Intelligence in Medicine*. Cham, Switzerland: Springer, 2021, pp. 1–10.

[12] X. Chen, X. Zhang, H. Xie, X. Tao, F. L. Wang, N. Xie, and T. Hao, ''A bibliometric and visual analysis of artificial intelligence technologies-enhanced brain MRI research,'' *Multimedia Tools Appl.*, vol. 80, no. 11, pp. 17335–17363, 2020.

[13] 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.

[14] P. Moeskops, J. M. Wolterink, B. H. van der Velden, K. G. Gilhuijs, T. Leiner, M. A. Viergever, and I. Išgum, ''Deep learning for multi-task medical image segmentation in multiple modalities,'' in *Proc. Int. Conf. Med. Image Comput. Comput.-Assist. Intervent.* Cham, Switzerland: Springer, 2016, pp. 478–486.

[15] L. Wang, R. Chen, S. Wang, N. Zeng, X. Huang, and C. Liu, ''Nested dilation network (NDN) for multi-task medical image segmentation,'' *IEEE Access*, vol. 7, pp. 44676–44685, 2019.

[16] F. I. Diakogiannis, F. Waldner, P. Caccetta, and C. Wu, ''ResUNet—A: A deep learning framework for semantic segmentation of remotely sensed data,'' *ISPRS J. Photogramm. Remote Sens.*, vol. 162, pp. 94–114, Apr. 2020.

[17] M. Wang and P. Li, ''A review of deformation models in medical image registration,'' *J. Med. Biol. Eng.*, vol. 39, no. 1, pp. 1–17, Feb. 2019.

[18] V. Dill, A. R. Franco, and M. S. Pinho, ''Automated methods for hippocampus segmentation: The evolution and a review of the state of the art,'' *Neuroinformatics*, vol. 13, no. 2, pp. 133–150, Apr. 2015.

[19] O. Ronneberger, P. Fischer, and T. Brox, ''U-Net: Convolutional networks for biomedical image segmentation,'' in *Proc. Int. Conf. Med. Image Comput. Comput.-Assist. Intervent.* Cham, Switzerland: Springer, 2015, pp. 234–241.

[20] P. Aljabar, R. Heckemann, A. Hammers, and J. D. H. Rueckert, ''Multi-atlas based segmentation of brain images: Atlas selection its effect on accuracy,'' *NeuroImage*, vol. 46, pp. 726–738, Oct. 2009.

[21] S. Powell, V. Magnotta, H. Johnson, V. Jammalamadaka, and R. N. Pierson, ''Andreasen registration machine learning-based automated segmentation of subcortical cerebellar brain structures,'' *NeuroImage*, vol. 39, pp. 238–247, Apr. 2008.

[22] Y. Jiang, ''Active contour model for image segmentation based on salient fitting energy,'' *Int. J. Inf. Commun. Technol.*, vol. 19, pp. 219–230, Apr. 2021.

[23] J. Yang, L. H. Staib, and J. S. Duncan, ''Neighbor-constrained segmentation with level set based 3-D deformable models,'' *IEEE Trans. Med. Imag.*, vol. 23, no. 8, pp. 940–948, Aug. 2004.

[24] N. Yamanakkanavar, J. Choi, and B. Lee, ''MRI segmentation and classification of human brain using deep learning for diagnosis of Alzheimer's disease: A survey,'' *Sensors*, vol. 20, p. 3243, Jan. 2020.

[25] D. Ataloglou, A. Dimou, D. Zarpalas, and P. Daras, ''Fast and precise hippocampus segmentation through deep convolutional neural network ensembles and transfer learning,'' *Neuroinformatics*, vol. 17, no. 4, pp. 563–582, Oct. 2019.

[26] D. L. Collins and J. C. Pruessner, ''Towards accurate, automatic segmentation of the hippocampus and amygdala from MRI by augmenting ANIMAL with a template library and label fusion,'' *NeuroImage*, vol. 52, no. 4, pp. 1355–1366, Oct. 2010.

[27] T. R. Langerak, U. A. van der Heide, A. N. T. J. Kotte, M. A. Viergever, M. van Vulpen, and J. P. W. Pluim, ''Label fusion in atlas-based segmentation using a selective and iterative method for performance level estimation (SIMPLE),'' *IEEE Trans. Med. Imag.*, vol. 29, no. 12, pp. 2000–2008, Dec. 2010.

[28] P. Coupé, J. V. Manjón, V. Fonov, J. Pruessner, M. Robles, and D. L. Collins, ''Nonlocal patch-based label fusion for hippocampus segmentation,'' in *Proc. Int. Conf. Med. Image Comput. Comput.-Assist. Intervent.* Berlin, Germany: Springer, 2010, pp. 129–136.

[29] H. Wang and P. Yushkevich, ''Multi-atlas segmentation with joint label fusion and corrective learning—An open source implementation,'' *Frontiers Neuroinform.*, vol. 7, p. 27, Nov. 2013.

[30] M. Sdika, ''Combining atlas based segmentation and intensity classification with nearest neighbor transform and accuracy weighted vote,'' *Med. Image Anal.*, vol. 14, no. 2, pp. 219–226, Apr. 2010.

[31] D. Shen, S. Moffat, S. M. Resnick, and C. Davatzikos, ''Measuring size and shape of the hippocampus in MR images using a deformable shape model,'' *NeuroImage*, vol. 15, no. 2, pp. 422–434, Feb. 2002.

[32] D. Zarpalas, P. Gkontra, P. Daras, and N. Maglaveras, ''Accurate and fully automatic hippocampus segmentation using subject-specific 3D optimal local maps into a hybrid active contour model,'' *IEEE J. Transl. Eng. Health Med.*, vol. 2, 2014, Art. no. 1800116.

[33] J. H. Morra, Z. Tu, L. G. Apostolova, A. E. Green, A. W. Toga, and P. M. Thompson, ''Comparison of AdaBoost and support vector machines for detecting Alzheimer's disease through automated hippocampal segmentation,'' *IEEE Trans. Med. Imag.*, vol. 29, no. 1, pp. 30–43, Jan. 2010.

[34] T. Tong, R. Wolz, P. Coupé, J. V. Hajnal, and D. Rueckert, ''Segmentation of MR images via discriminative dictionary learning and sparse coding: Application to hippocampus labeling,'' *NeuroImage*, vol. 76, pp. 11–23, Aug. 2013.

[35] J. Mairal, F. Bach, J. Ponce, G. Sapiro, and A. Zisserman, ''Discriminative learned dictionaries for local image analysis,'' in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2008, pp. 1–8.

[36] S. M. Anwar, M. Majid, A. Qayyum, M. Awais, M. Alnowami, and M. K. Khan, ''Medical image analysis using convolutional neural networks: A review,'' *J. Med. Syst.*, vol. 42, no. 11, pp. 1–13, Nov. 2018.

[37] P. Moeskops, M. A. Viergever, A. M. Mendrik, L. S. de Vries, M. J. N. L. Benders, and I. Isgum, ''Automatic segmentation of MR brain images with a convolutional neural network,'' *IEEE Trans. Med. Imag.*, vol. 35, no. 5, pp. 1252–1261, May 2016.

[38] M. Havaei, A. Davy, D. Warde-Farley, A. Biard, A. Courville, Y. Bengio, C. Pal, P. M. Jodoin, and H. Larochelle, ''Brain tumor segmentation with deep neural networks,'' *Med. Image Anal.*, vol. 35, pp. 18–31, Apr. 2017.

[39] T. Brosch, L. Y. W. Tang, Y. Yoo, D. K. B. Li, A. Traboulsee, and R. Tam, ''Deep 3D convolutional encoder networks with shortcuts for multiscale feature integration applied to multiple sclerosis lesion segmentation,'' *IEEE Trans. Med. Imag.*, vol. 35, no. 5, pp. 1229–1239, May 2016.

[40] H. Choi and K. H. Jin, ''Fast and robust segmentation of the striatum using deep convolutional neural networks,'' *J. Neurosci. Methods*, vol. 274, pp. 146–153, Dec. 2016.

[41] M. Shakeri, S. Tsogkas, E. Ferrante, S. Lippe, S. Kadoury, N. Paragios, and I. Kokkinos, "Sub-cortical brain structure segmentation using F-CNN'S," in *Proc. IEEE 13th Int. Symp. Biomed. Imag. (ISBI)*, Apr. 2016, pp. 269–272.

[42] 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.

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

[44] K. Kushibar, S. Valverde, S. González-Villé, J. Bernal, M. Cabezas, A. Oliver, and X. Lladó, "Automated sub-cortical brain structure segmentation combining spatial and deep convolutional features," *Med. Image Anal.*, vol. 48, pp. 177–186, Aug. 2018.

[45] J. Kleesiek, G. Urban, A. Hubert, D. Schwarz, K. Maier-Hein, M. Bendszus, and A. Biller, "Deep MRI brain extraction: A 3D convolutional neural network for skull stripping," *NeuroImage*, vol. 129, pp. 460–469, Apr. 2016.

[46] A. de Brebisson and G. Montana, "Deep neural networks for anatomical brain segmentation," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. Workshops (CVPRW)*, Jun. 2015, pp. 20–28.

[47] R. Mehta, A. Majumdar, and J. Sivaswamy, "BrainSegNet: A convolutional neural network architecture for automated segmentation of human brain structures," *J. Med. Imag.*, vol. 4, no. 2, Apr. 2017, Art. no. 024003.

[48] *volBrain: Online MRI Segmentation*. Accessed: Oct. 20, 2021. [Online]. Available: https://www.nitrc.org/projects/vol

[49] *3D Slicer*. Accessed: Oct. 20, 2021. [Online]. Available: https://www.slicer.org/

[50] H. Wang, S. R. Das, J. W. Suh, M. Altinay, J. Pluta, C. Craige, B. Avants, and P. A. Yushkevich, "A learning-based wrapper method to correct systematic errors in automatic image segmentation: Consistently improved performance in hippocampus, cortex and brain segmentation," *NeuroImage*, vol. 55, no. 3, pp. 968–985, Apr. 2011.

[51] G. Litjens, T. Kooi, B. E. Bejnordi, A. A. A. Setio, F. Ciompi, M. Ghafoorian, J. A. W. M. van der Laak, B. van Ginneken, and C. I. Sánchez, "A survey on deep learning in medical image analysis," *Med. Image Anal.*, vol. 42, pp. 60–88, Dec. 2017.

[52] Ö. Çiçek, A. Abdulkadir, S. S. Lienkamp, T. Brox, and O. Ronneberger, "S3D U-Net: Learning dense volumetric segmentation from sparse annotation," in *Proc. Int. Conf. Med. Image Comput. Comput.-Assist. Intervent.* Cham, Switzerland: Springer, 2016, pp. 424–432.

[53] A. L. Simpson, "A large annotated medical image dataset for the development and evaluation of segmentation algorithms," 2019, *arXiv:1902.09063*.

[54] H.-H. Chang, A. H. Zhuang, D. J. Valentino, and W.-C. Chu, "Performance measure characterization for evaluating neuroimage segmentation algorithms," *NeuroImage*, vol. 47, no. 1, pp. 122–135, Aug. 2009.

[55] L. Cao, L. Li, J. Zheng, X. Fan, F. Yin, H. Shen, and J. Zhang, "Multi-task neural networks for joint hippocampus segmentation and clinical score regression," *Multimedia Tools Appl.*, vol. 77, no. 22, pp. 29669–29686, Nov. 2018.

[56] Y. Chen, B. Shi, Z. Wang, P. Zhang, C. D. Smith, and J. Liu, "Hippocampus segmentation through multi-view ensemble convnets," in *Proc. IEEE 14th Int. Symp. Biomed. Imag. (ISBI)*, Apr. 2017, pp. 192–196.

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

[58] N. Nogovitsyn, R. Souza, M. Müller, A. Srajer, S. Hassel, S. R. Arnott, A. Davis, and G. Hall, "Testing a deep convolutional neural network for automated hippocampus segmentation in a longitudinal sample of healthy participants," *NeuroImage*, vol. 197, pp. 589–597, Aug. 2019.

[59] L. Yuan, X. Wei, H. Shen, L.-L. Zeng, and D. Hu, "Multi-center brain imaging classification using a novel 3D CNN approach," *IEEE Access*, vol. 6, pp. 49925–49934, 2018.

[60] E. Hosseini-Asl, G. Gimel'farb, and A. El-Baz, "Alzheimer's disease diagnostics by a deeply supervised adaptable 3D convolutional network," 2016, *arXiv:1607.00556*.

[61] G. Wang, W. Li, M. A. Zuluaga, R. Pratt, P. A. Patel, M. Aertsen, T. Doel, A. L. David, J. Deprest, S. Ourselin, and T. Vercauteren, "Interactive medical image segmentation using deep learning with image-specific fine tuning," *IEEE Trans. Med. Imag.*, vol. 37, no. 7, pp. 1562–1573, Jul. 2018.
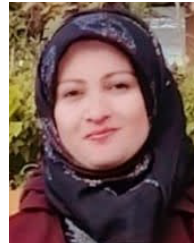
**NOSHEEN SOHAIL** received the M.Sc. degree in computer engineering from the University of Engineering and Technology, Taxila, Pakistan, in 2013, where she is currently pursuing the Ph.D. degree in computer engineering. Her research interests include deep learning and medical image segmentation.

**SYED MUHAMMAD ANWAR** received the B.Sc. degree (Hons.) in computer engineering from the University of Engineering and Technology (UET), Taxila, Pakistan, in 2005, and the M.Sc. degree (Hons.) in data communications and the Ph.D. degree in electronic and electrical engineering from The University of Sheffield, U.K., in 2007 and 2012, respectively. He is currently an Associate Professor with the Department of Software Engineering, UET Taxila. His research interests include medical imaging, data communication, and human–computer interaction.

• • •