Medical & Biological Engineering & Computing (2023) 61:2329–2339
https://doi.org/10.1007/s11517-023-02836-9

ORIGINAL ARTICLE

Deep convolutional neural network for hippocampus segmentation with boundary region refinement

Guanghua He[1] · Guying Zhang[1] · Lianlian Zhou[1] · Hancan Zhu[1]

Received: 3 December 2022 / Accepted: 5 April 2023 / Published online: 17 April 2023

Abstract

Accurately segmenting the hippocampus from magnetic resonance (MR) brain images is a crucial step in studying brain disorders. However, this task is challenging due to the low signal contrast of hippocampal images, the irregular shape, and small structural size of the hippocampi. In recent years, several deep convolutional networks have been proposed for hippocampus segmentation, which have achieved state-of-the-art performance. These methods typically use large image patches for training the network, as larger patches are beneficial for capturing long-range contextual information. However, this approach increases the computational burden and overlooks the significance of the boundary region. In this study, we propose a deep learning–based method for hippocampus segmentation with boundary region refinement. Our method involves two main steps. First, we propose a convolutional network that takes large image patches as input for initial segmentation. Then, we extract small image patches around the hippocampal boundary for training the second convolutional neural network, which refines the segmentation in the boundary regions. We validate our proposed method on a publicly available dataset and demonstrate that it significantly improves the performance of convolutional neural networks that use *single-size* image patches as input. In conclusion, our study proposes a novel method for hippocampus segmentation, which improves upon the current state-of-the-art methods. By incorporating a boundary refinement step, our approach achieves higher accuracy in hippocampus segmentation and may facilitate research on brain disorders.

Keywords Hippocampus segmentation · U-Net · Deep learning · Boundary refinement

1 Introduction

With the advancement of artificial intelligence technology and computer hardware, computer-aided diagnosis (CAD) has attracted increasing attention [1]. Computer-aided diagnosis (CAD) refers to the use of computer technology and algorithms to assist in the diagnosis of diseases or abnormalities in medical images, such as X-rays, MRI, CT scans, and ultrasound. CAD systems can help identify patterns and features in medical images that may be difficult for a human observer to detect or analyze and provide quantitative measurements or other diagnostic information that can aid in the diagnostic process. These systems can help improve the accuracy and efficiency of diagnosis, reduce inter-observer variability, and assist in early disease detection and treatment.

The hippocampus is a bilateral brain structure involved in many brain disorders, such as epilepsy, Alzheimer's disease (AD), and Parkinson's disease [2–4]. As a critical step in CAD, it is important to accurately and automatically segment the hippocampus from MR images to study these brain disorders. In the past decade, multi-atlas segmentation method has been one of the most popular medical image segmentation methods and has been widely used in the hippocampus segmentation [5–9]. The multi-atlas segmentation method uses a set of atlases (an atlas consists of an image and its segmentation label) to segment the target image and usually includes three steps, i.e., atlas selection [10, 11], image registration [12, 13], and label fusion [14–16]. In the atlas selection step, atlases that are most similar to the target image are selected. Then, each selected atlas image is registered to the target image independently, obtaining the deformation field from the atlas image to the

✉ Hancan Zhu
hancanzhu@yeah.net

1 School of Mathematics, Physics, and Information Science, Shaoxing University, 900 ChengNan Rd, Shaoxing 312000, Zhejiang, China

target image. By using the deformation field, each atlas label is warped to the target image. These warped atlas labels are fused to obtain the final segmentation of the target image in the label fusion step.

The multi-atlas segmentation method has achieved a considerable degree of success. However, it usually takes several hours to segment the hippocampi of a subject, as it requires a number of time-consuming image registrations. In recent years, deep learning–based methods have been developed rapidly and have been widely used in the medical image segmentation [17, 18]. Most deep learning–based segmentation methods use fully convolutional networks (FCNs) for dense prediction [19]. FCNs take an image or image patch of arbitrary size as input and produce a correspondingly sized segmentation label as output with efficient inference and learning. U-Net is one of the most commonly used FCNs in the medical image segmentation and consists of a contracting path and a symmetric expanding path [20]. The contracting path is used to exact context information, and the expanding path is used to obtain precise localization. The feature maps in the contracting path are copied and concatenated to the feature maps in the expanding path to provide the detailed image information that is lost during the successive pooling operators. Based on the structure of U-Net, Oktay et al. [21] added an attention mechanism to U-Net, and proposed Attention U-Net. Gao et al. [22] used a graph convolutional network in U-Net and developed graph U-Nets. To better capture long-distance image semantic information, the transformer structure was also applied in the U-Net [23].

Deep learning–based methods have been utilized for hippocampus segmentation [24–31]. By referring to Lao et al. [24], they constructed a 3D multi-task convolutional neural network (CNN) for joint automatic hippocampal segmentation and AD classification. Nogovitsyn et al. [25] evaluated a CNN-based segmentation algorithm for the hippocampus, demonstrating superior performance compared to FreeSurfer. In [26], Folle et al. proposed a new network for hippocampus head and body segmentation using dilated convolutions and deep supervision in 3D U-Net. Ataloglou et al. [27] introduced deep CNN ensembles and transfer learning for fast and accurate hippocampus segmentation. Xie et al. [28] proposed a patch-based canonical neural network for near real-time hippocampus segmentation. Cao et al. [29] proposed multi-task neural networks for joint hippocampus segmentation and clinical score regression. Kim et al. [30] proposed an unsupervised deep learning method for hippocampus segmentation. Finally, Zhu et al. [31] introduced a deep learning–based label correction method, which was applied in multi-atlas label fusion for hippocampus segmentation.

Most of the aforementioned deep learning–based hippocampus segmentation approaches are primarily focused on network architecture design. Recently, Isensee et al. [32, 33] demonstrated that a well-trained U-Net outperforms most existing deep learning models and offered several recommendations for training deep networks, such as employing larger image patches. According to the human visual system, greater emphasis should be placed on the boundary region when recognizing or segmenting an image [34, 35]. It could increase computational overhead and, more importantly, overlook the significance of the boundary region when employing *single-size* image patches for training. In this study, we present a novel two-stage deep convolutional neural network method for hippocampus segmentation. The first neural network is used for initial hippocampus segmentation, while the second neural network is utilized for hippocampal boundary refinement. The proposed method is evaluated on a publicly available dataset, and our results demonstrate its efficacy. Our primary contributions can be summarized as follows:

(1) We propose a new two-stage deep convolutional neural network method for hippocampus segmentation.(2) Multi-scale image patches are employed in our approach. Large image patches are used in the first stage to capture contextual information, while small image patches are employed for boundary refinement in the second stage.(3) Our experimental findings reveal that our method significantly improves the performance of convolutional neural networks that employ *single-size* image patches as input.

2 Methods

The framework of the proposed hippocampus segmentation method is shown in Fig. 1. It consists of two steps. In the first step, a neural network is proposed for initial segmentation. Then, the boundary region is refined with the second neural network. In the following subsections, we will introduce these two steps, and also the details of training and inference.

2.1 Initial hippocampus segmentation

U-Net [20, 36] is applied for the initial segmentation of hippocampi. U-Net consists of a contracting path and an expanding path, which are shown in the top subfigure of Fig. 2. The contracting path is built by alternating one convolution block and one $2 \times 2 \times 2$ max pooling operation with stride 2. The convolution block consists of two $3 \times 3 \times 3$ convolutions, each of which is followed by a batch normalization layer and a rectified linear unit (ReLU). Padded convolution layers with a padding size of 1 are used to ensure that the spatial dimension of the feature maps is preserved. In the contracting path, three max pooling operations are used to contract feature maps by a scale of $\frac{1}{8} \times \frac{1}{8} \times \frac{1}{8}$. The last max pooling operation is followed by a convolution block.

Correspondingly, the expanding path is built by alternating one $2 \times 2 \times 2$ transposed convolution with stride 2, and one convolution block that has the same structure as that in the contracting path. In total, three transposed convolution operations are used to recover the resolution of feature maps in the expanding path. The expanding path is followed by a $1 \times 1 \times 1$ convolution and a softmax layer, which outputs two feature maps for the probabilities of the hippocampus and background. The feature maps in the contracting path are concatenated to the corresponding feature maps in the expanding path to provide the detailed image information that is lost during the successive down-sampling steps.

In the initial segmentation, we randomly extract image patches with a size of $128 \times 128 \times 128$ in the whole brain image as the input of the network. The channel dimension of the first convolution block is set to 16 (denoted as the base channel dimension) and is multiplied by 2 after each pooling operation, and divided by 2 after each transposed convolution operation.

2.2 Boundary region refinement

Although large image patches are used for training the neural network in the initial hippocampus segmentation, there still exist some small isolated false positives outside the hippocampal region. Two post-processing steps are used to remove these artifacts automatically by searching the voxels of each automated segmentation to find the connected regions and selecting two regions with maximum volumes for the left and right hippocampus. Then, boundary regions are extracted for the hippocampi, by performing the dilation operation and erosion operation successively, with the same structure element, i.e., the $3 \times 3 \times 3$ tensor whose elements are all 1.

The segmentation of boundary regions is refined by a second U-Net, which is shown in the bottom subfigure of Fig. 2. To train the network, image patches with sizes of $16 \times 16 \times 16$ are randomly extracted from the obtained boundary regions. As small image patches are used for training, the contracting path of the network contains only two max pooling operations to contract feature maps by the scale of $\frac{1}{4} \times \frac{1}{4} \times \frac{1}{4}$, and the expanding path contains two transposed convolution operations to recover the resolution of feature maps. The base channel dimension is set to 32, which is multiplied by 2 after each pooling operation and divided by 2 after each transposed convolution operation.

2.3 Details of training and inference

The networks are trained by the Adam optimizer with a batch size of 4 in the network for the initial segmentation and 15 in the network for the boundary region refinement. This is implemented in Pytorch on a single NVIDIA GeForce RTX 2080 Ti GPU [37]. The learning rates are initially set to

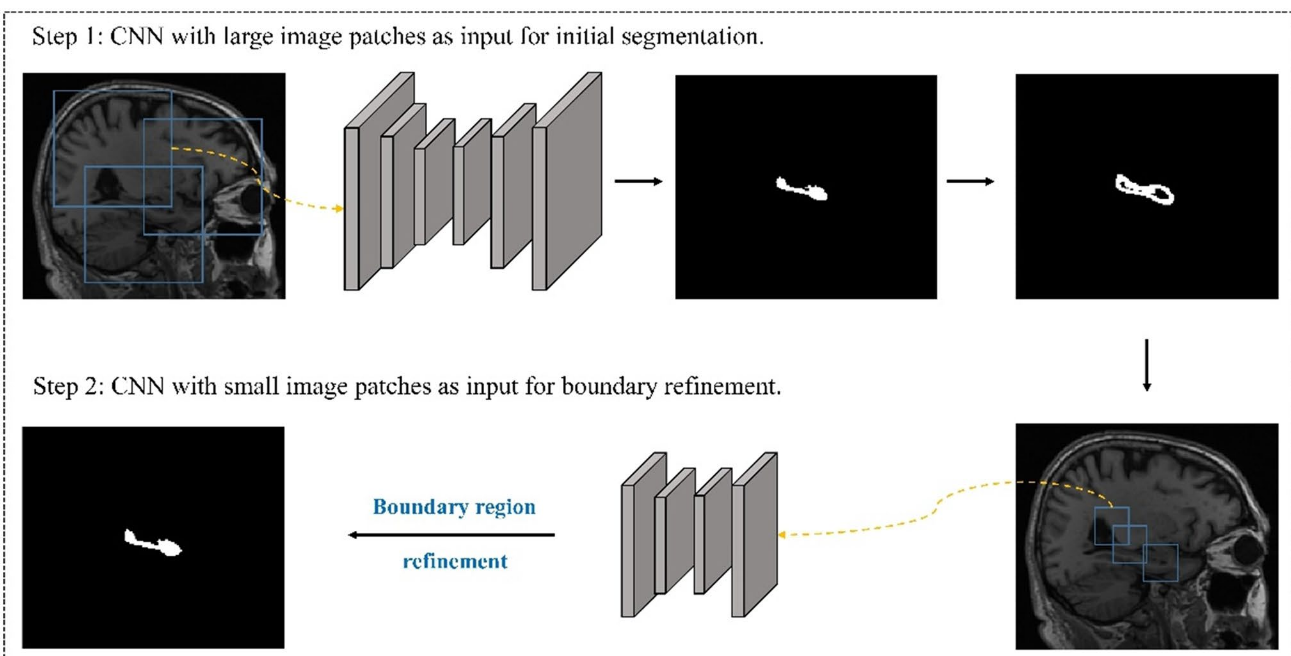

Fig. 1 The framework of the proposed method

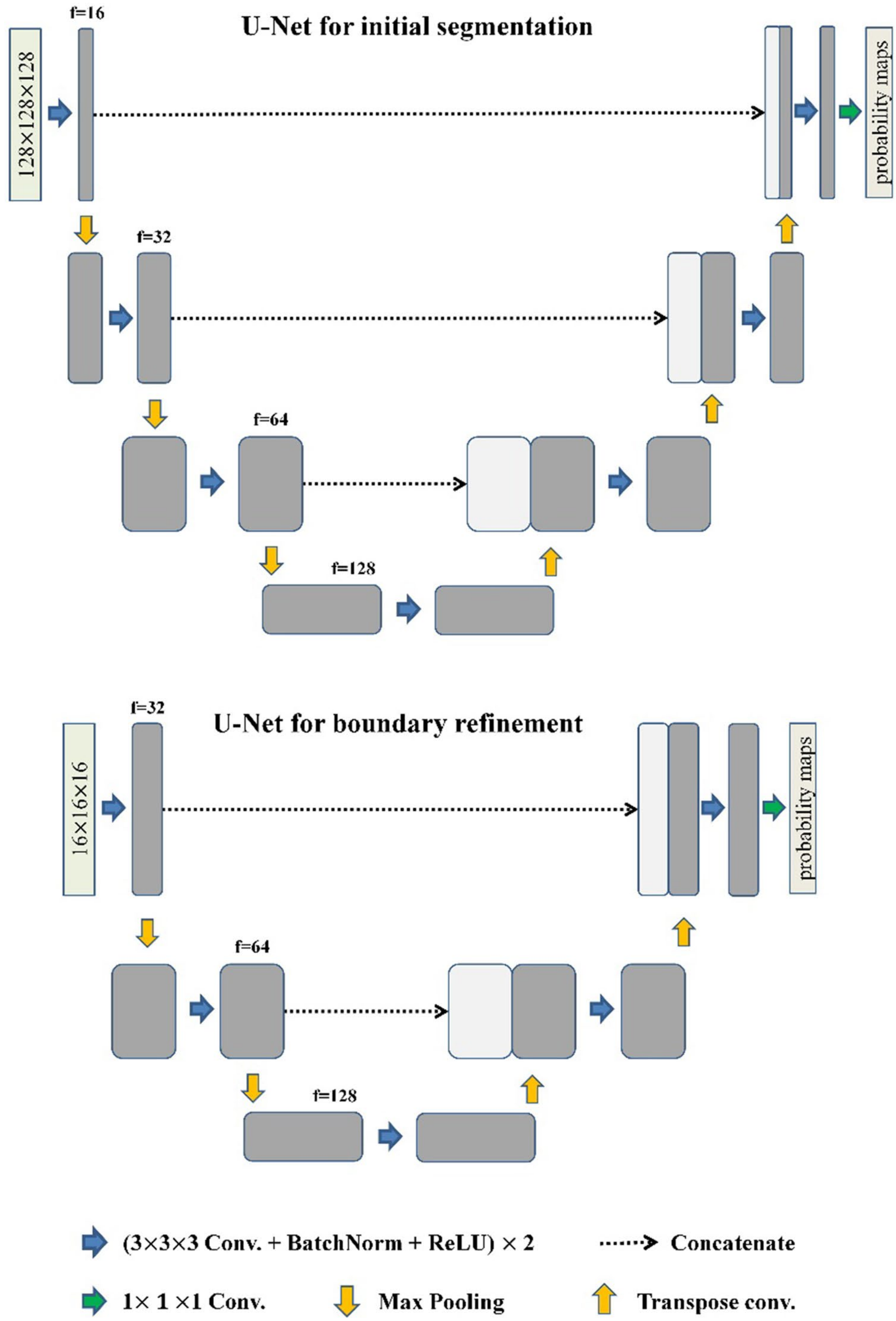

Fig. 2 The structure of network for initial segmentation (top) and the structure of network for boundary refinement (bottom)

0.001 and decay by each iteration with a power of 0.9 using a poly learning rate strategy. To increase the training data, several data augmentation techniques are used, including random cropping image patches; random mirror flipping across the axial, coronal, and sagittal planes by a probability of 0.5; and random intensity shift between [−0.1, 0.1] and scale between [0.9, 1.1]. The softmax Dice loss is employed to train the network [38], which is defined as,

$$\mathrm{L}\left(\mathrm{Y}, \widetilde{Y}\right) = -\frac{1}{N}\sum_{n=1}^{N}\frac{2 \bullet Y_n \bullet \widetilde{Y}_n}{Y_n + \widetilde{Y}_n}, \tag{1}$$

where Y_{n} and $\widetilde{Y}_{\mathrm{n}}$ are the ground truth and predicted probabilities of the n-th image patch, respectively, and N is the batch size. L2 Norm is applied for model regularization with a weight decay rate of 10^{-5}. The training process is

Table 1 Four selected metrics (mean ± std) of segmentation results using different methods with 30 training subjects. ↑ indicates that a larger value corresponds to a higher segmentation accuracy, and ↓ indicates that a smaller value corresponds to a higher segmentation accuracy. Best results in each row are typeset in bold

	UNet64	AttionUNet64	UNet128	AttionUNet128	Proposed
Dice (↑)	0.852 ± 0.031	0.836 ± 0.038	0.885 ± 0.023	0.885 ± 0.023	**0.893 ± 0.019**
Jaccard(↑)	0.744 ± 0.046	0.720 ± 0.054	0.794 ± 0.036	0.794 ± 0.037	**0.808 ± 0.032**
HD (↓)	19.619 ± 26.917	29.899 ± 31.454	11.560 ± 22.337	4.747 ± 8.358	**3.844 ± 1.639**
HD95 (↓)	1.728 ± 1.157	3.737 ± 10.946	1.101 ± 0.222	1.648 ± 5.579	**1.041 ± 0.133**

stopped after 1000 epochs in the initial segmentation and 4000 epochs in the boundary refinement.

In the testing stage, patches with a size of $128 \times 128 \times 128$ are extracted to feed into the trained models with a non-overlapped sliding window strategy for the initial segmentation. For the boundary refinement, patches with a size of $16 \times 16 \times 16$ are extracted to feed into the trained models with an overlapped sliding window strategy with a stride of $8 \times 8 \times 8$, and the average of the probability maps for the overlap regions are used to obtain the boundary refinement.

2.4 Evaluation metrics

We evaluate the segmentation results by four segmentation evaluation measures: Dice coefficient, Jaccard index, Hausdorff distance (HD), and Hausdorff 95 distance (HD95) [39]. By denoting A as the reference segmentation label, B as the automated segmentation label, and $V(S)$ as the volume of segmentation S, these evaluation measures are defined as:

$$\mathrm{Dice} = 2\frac{\mathrm{V(A \cap B)}}{\mathrm{V(A) + V(B)}}, \tag{2}$$

$$\mathrm{Jaccard} = \frac{\mathrm{V(A \cap B)}}{\mathrm{V(A \cup B)}}, \tag{3}$$

$$\begin{aligned}\mathrm{HD} &= \max(\mathrm{H(A,B)}, \mathrm{H(B,A)}),\ \text{where } \mathrm{H(A,B)} \\ &= max_{e \in \partial A}\left(min_{f \in \partial B} d(e,f)\right),\end{aligned} \tag{4}$$

$$\begin{aligned}&\mathrm{HD95}: \text{similar to HD, except that 5\% data points with the} \\ &\quad \text{largest distance are removed before calculation,}\end{aligned} \tag{5}$$

where ∂A is the boundary voxels of A, and $d(\bullet, \bullet)$ is the Euclidean distance of two points. In these metrics, the first two, namely, Dice and Jaccard, are used to evaluate the volumetric overlap between the automated segmentation and the reference segmentation. Higher values for these metrics indicate better segmentation results. The last two metrics, HD and HD95, are used to assess the agreement between segmentation boundaries. Lower values for these metrics indicate better segmentation results.

3 Experiments and results

3.1 Dataset and pre-processing

We downloaded a dataset from the Alzheimer's Disease Neuroimaging Initiative (ADNI) database (adni.loni.usc.edu/) as well as the corresponding hippocampus labels, which are provided by the EADC-ADNI harmonized segmentation protocol (www.hippocampal-protocol.net) [33]. The dataset consists of 135 T1 MRI scans that have been processed using a standard preprocessing protocol, including alignment along the line passing through the anterior and posterior commissures of the brain (AC-PC line), bias field correction, and spatial normalization to MNI152 template space using affine transformation.

Table 2 Four selected metrics (mean ± std) of segmentation results using different methods with 90 training subjects. ↑ indicates that a larger value corresponds to a higher segmentation accuracy, and ↓ indicates that a smaller value corresponds to a higher segmentation accuracy. Best results in each row are typeset in bold

	UNet64	AttionUNet64	UNet128	AttionUNet128	Proposed
Dice (↑)	0.873 ± 0.0285	0.872 ± 0.026	0.894 ± 0.020	0.896 ± 0.020	**0.900 ± 0.017**
Jaccard(↑)	0.776 ± 0.0437	0.774 ± 0.041	0.809 ± 0.032	0.812 ± 0.032	**0.819 ± 0.028**
HD (↓)	16.275 ± 29.2608	12.532 ± 17.91	6.986 ± 15.776	4.783 ± 10.202	**3.393 ± 1.549**
HD95 (↓)	1.215 ± 0.4787	1.247 ± 0.501	1.031 ± 0.111	1.029 ± 0.132	**1.000 ± 0.000**

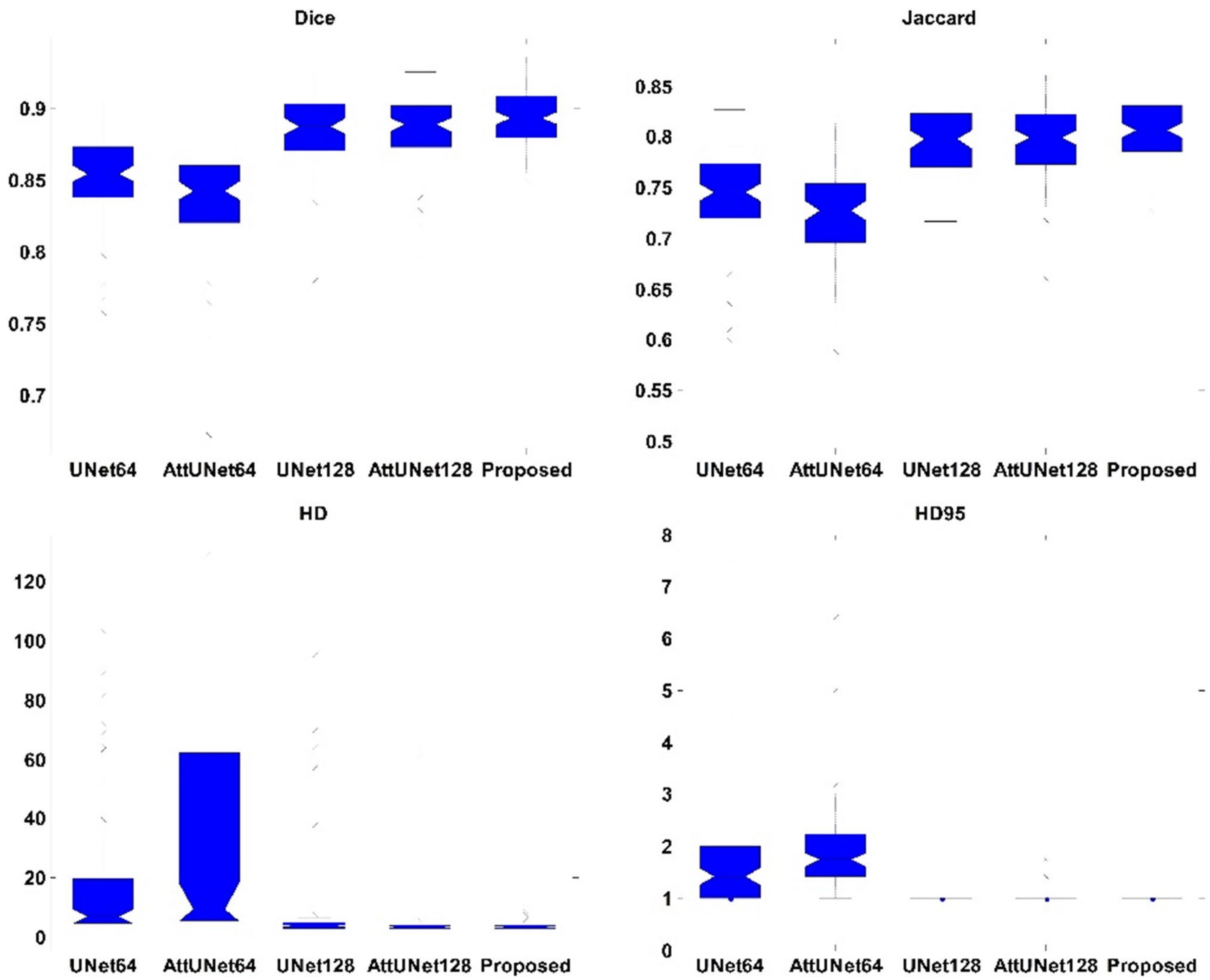

Fig. 3 Boxplots of four indexes for measuring hippocampus segmentation results by different methods with 30 training subjects

By carefully checking the hippocampus labels, we found 5 subjects whose hippocampus labels and images were not matched. Their subject IDs are 002_S_0938, 007_S_1304, 016_S_4121, 029_S_4279, and 136_S_0429. We used the remaining 130 subjects to validate the proposed method. This group was divided into training and testing sets in the experiment. To persuasively evaluate the proposed method, two different partitions were used, of which the first was randomly selecting 30 subjects as the training set and the remaining 100 subjects as the testing set, and the second was randomly selecting 90 subjects as the training set and the remaining 40 subjects as the testing set.

3.2 Experiment results

The proposed method was compared with U-Net [20, 36] and AttionUNet [21] using two different sizes of image patches, i.e., $64 \times 64 \times 64$ and $128 \times 128 \times 128$. For a fair comparison, the same settings were adopted in the learning and testing of these networks, including the same learning rate strategy, the same data augmentation techniques, and the same loss function. The batch size was set to 4, and the networks were trained until 4000 epochs.

Table 1 and Table 2 report the four metrics (mean ± std) of the segmentation results obtained by different methods including U-Net with $64 \times 64 \times 64$ image patches as input (UNet64), AttionUNet with $64 \times 64 \times 64$ image patches as input (AttionUNet64), U-Net with $128 \times 128 \times 128$ image patches as input (UNet128), AttionUNet with $128 \times 128 \times 128$ image patches as input (AttionUNet128), and the proposed method. Table 1 lists the results using 30 training subjects. It shows that the proposed method achieves the best segmentation results. For example, it outperforms UNet64, AttionUNet64, UNet128, and AttionUNet128 by 4.1%, 5.7%, 0.8%, and 0.8%, respectively, in terms of the Dice values.

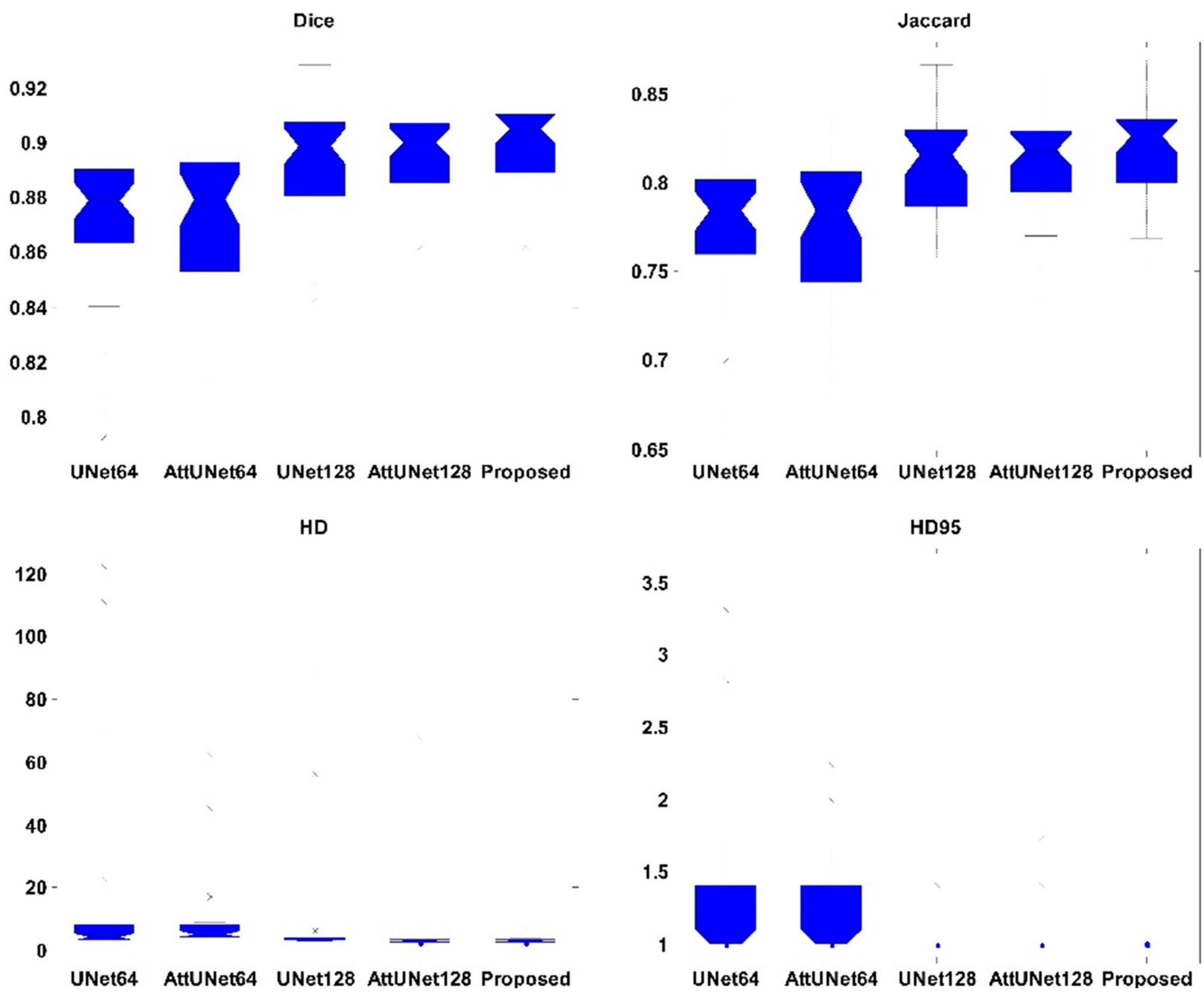

Fig. 4 Boxplots of four indexes for measuring hippocampus segmentation results by different methods with 90 training subjects

It also shows that U-Net/AttionUNet with $128 \times 128 \times 128$ image patches can obtain better segmentation results than that with $64 \times 64 \times 64$ image patches. Table 2 lists the results using 90 training subjects and supports the same conclusion.

Figures 3 and 4 show boxplots of the four metrics used to evaluate the segmentation performance: Dice, Jaccard, HD, and HD95. The boxplots display the performance of the proposed method and other methods under different sizes of training sets, i.e., 30 training subjects in Fig. 3 and 90 training subjects in Fig. 4. Based on the boxplots, the proposed method appears to outperform the other methods in terms of the four metrics. Specifically, the proposed method demonstrates higher values for the Dice and Jaccard metrics, indicating better volumetric overlap between the automated and reference segmentations. The proposed method also shows lower values for the HD and HD95 metrics, suggesting more accurate segmentation boundaries.

In Figs. 5 and 6, we list the hippocampus segmentation results of a randomly selected subject. The segmentation obtained by the proposed method is the most similar to the manual segmentation that is treated as the ground truth in the study. This suggests that the proposed method is likely to be accurate and reliable in identifying the boundaries of the hippocampus.

4 Discussion

Deep learning–based methods have been widely used in medical image segmentation, including hippocampus segmentation. The use of fully convolutional networks enables image segmentation to be implemented in an end-to-end fashion, and the segmental label is predicted with the same size as the input image. However, due to limited

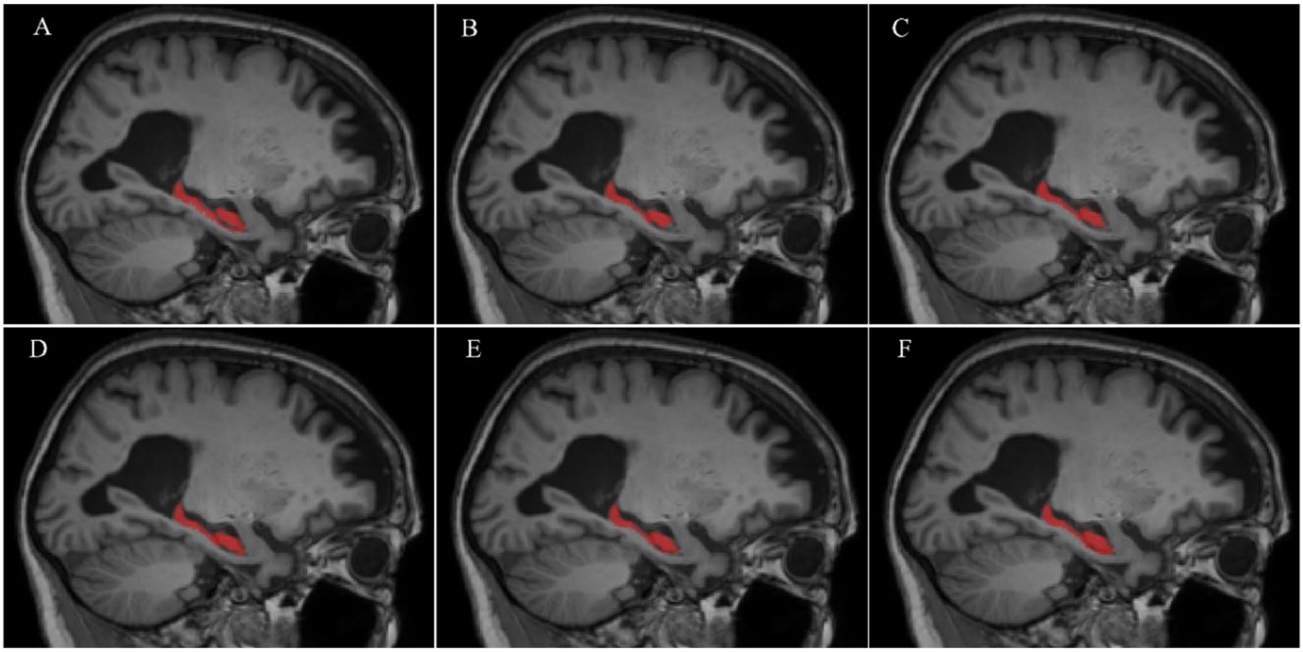

Fig. 5 Hippocampus segmentations of a randomly selected subject, obtained by manual segmentation and different deep learning methods using 30 training subjects. **A** Manual segmentation, **B** UNet64, **C** AttUNet 64, **D** UNet128, **E** AttUNet128, **F** proposed method

computation resources, image patches are often used to input deep networks in the field of 3D medical image segmentation, and the patch size is an important parameter. Previous studies have suggested that large image patches are beneficial for capturing context, while small image patches are conducive to learning local image information.

Following the idea of "look closer to segment better," this paper presents a two-stage deep learning method for

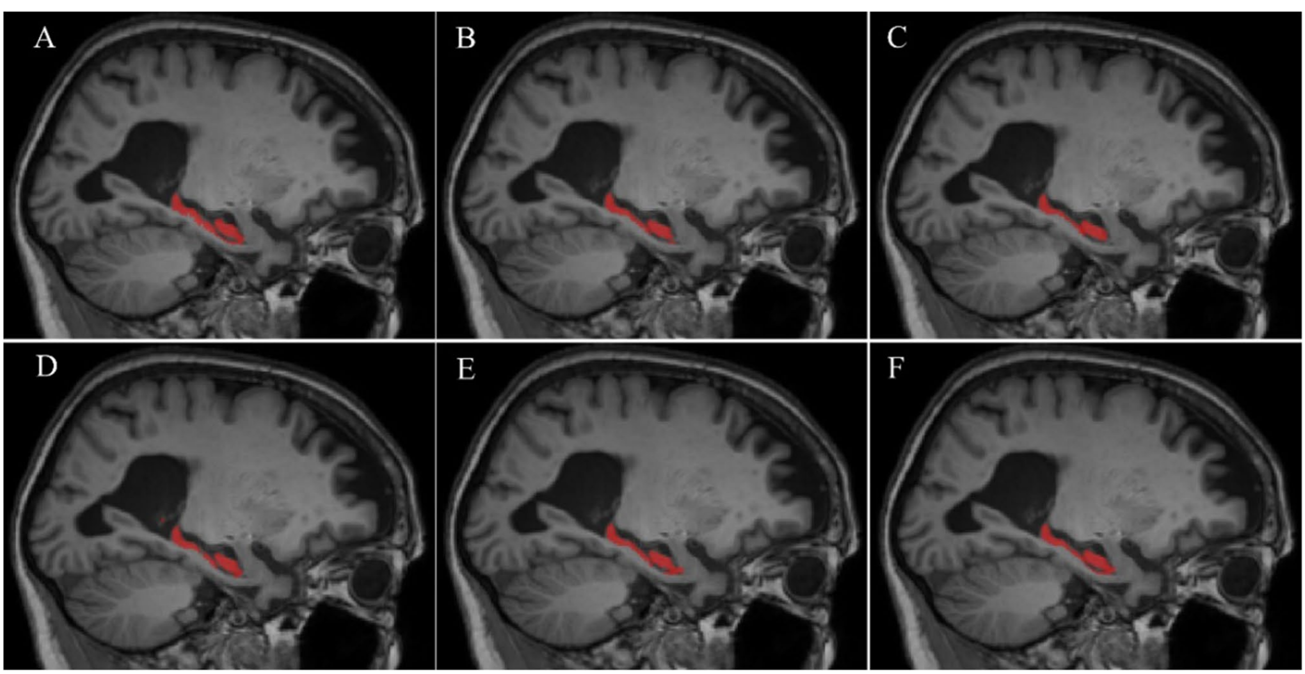

Fig. 6 Hippocampus segmentations of a randomly selected subject, obtained by manual segmentation and different deep learning methods using 90 training subjects. **A** Manual segmentation, **B** UNet64, **C** AttUNet 64, **D** UNet128, **E** AttUNet128, **F** proposed method

Table 3 Dice values of hippocampal segmentation results obtained by the proposed method using the base channels 8 and 16, respectively, in the first network

	With 30 training subjects	With 90 training subjects
Base channel = 8	0.893	0.900
Base channel = 16	0.893	0.900

hippocampus segmentation. The first step uses a U-Net with 3 pooling operators and large image patches to obtain initial hippocampal segmentation, while the second step refines the boundary region using another U-Net with 2 pooling operators and small image patches. The proposed method is validated using both a small training dataset containing 30 subjects and a large training dataset containing 90 subjects and is found to outperform U-Net and Attention U-Net with single-size image patches in both settings. Moreover, the proposed method has more obvious advantages than other methods in the small training dataset, illustrating its efficiency in exploiting a limited number of training subjects.

In the proposed method, the network of the first step can use light parameters to speed up the training and testing process, as it is only used to roughly locate the hippocampus. The performance of the proposed method is not compromised when using light parameters in the first step, as demonstrated in Table 3 by the Dice values of hippocampal segmentation results obtained with base channels 8 and 16 in the first network. The use of small image patches in the second step allows for fast training and testing, making the proposed method a promising approach for hippocampus segmentation.

5 Conclusion

The proposed two-stage deep learning method for hippocampus segmentation has exhibited remarkable performance compared to other methods that utilize single-size image patches. The method enhances UNet64, AttentionUNet64, UNet128, and AttentionUNet128 by 4.1%, 5.7%, 0.8%, and 0.8%, respectively, in terms of the Dice values, when validated on a publicly available dataset with 30 training subjects. The method is both efficient and effective, particularly when dealing with a restricted number of training subjects. Therefore, this approach holds potential for various applications in the diagnosis and treatment of brain disorders.

However, the proposed method also has some limitations. It was only evaluated on one publicly available dataset, and the generalizability of the method to other datasets remains to be validated. Future work could focus on exploring the generalizability of the proposed method to other datasets and evaluating its performance on larger datasets. It would also be interesting to investigate the potential of transfer learning or domain adaptation to improve the performance of the method when the target dataset is significantly different from the training dataset.

Acknowledgements The Canadian Institutes of Health Research is providing funds to support ADNI clinical sites in Canada. Private sector contributions are facilitated by the Foundation for the National Institutes of Health (www.fnih.org). The grantee organization is the Northern California Institute for Research and Education, and the study is coordinated by the Alzheimer's Therapeutic Research Institute at the University of Southern California. ADNI data are disseminated by the Laboratory for Neuro Imaging at the University of Southern California.

Funding This work was supported by the Scientific Research Project of Shaoxing University (20210038), Zhejiang Provincial Natural Science Foundation of China (LY19F020013), and National Natural Science Foundation of China (61602307, 61877039).

Data collection and sharing for this project was funded by the Alzheimer's Disease Neuroimaging Initiative (ADNI) (National Institutes of Health Grant U01 AG024904) and DOD ADNI (Department of Defense award number W81XWH-12–2-0012). ADNI is funded by the National Institute on Aging, the National Institute of Biomedical Imaging and Bioengineering, and through generous contributions from the following: AbbVie, Alzheimer's Association; Alzheimer's Drug Discovery Foundation; Araclon Biotech; BioClinica, Inc.; Biogen; Bristol-Myers Squibb Company; CereSpir, Inc.; Cogstate; Eisai Inc.; Elan Pharmaceuticals, Inc.; Eli Lilly and Company; EuroImmun; F. Hoffmann-La Roche Ltd and its affiliated company Genentech, Inc.; Fujirebio; GE Healthcare; IXICO Ltd.; Janssen Alzheimer Immunotherapy Research & Development, LLC.; Johnson & Johnson Pharmaceutical Research & Development LLC.; Lumosity; Lundbeck; Merck & Co., Inc.; Meso Scale Diagnostics, LLC.; NeuroRx Research; Neurotrack Technologies; Novartis Pharmaceuticals Corporation; Pfizer Inc.; Piramal Imaging; Servier; Takeda Pharmaceutical Company; and Transition Therapeutics.

Declarations

Conflict of interest The authors declare no competing interests.

References

1. Abdulla AA (2020) Efficient computer-aided diagnosis technique for leukaemia cancer detection. IET Image Proc 14:4435–4440
2. Akhondi-Asl A, Jafari-Khouzani K, Elisevich K, Soltanian-Zadeh H (2011) Hippocampal volumetry for lateralization of temporal lobe epilepsy: automated versus manual methods. Neuroimage 54:S218–S226
3. Wolz R, Schwarz AJ, Yu P, Cole PE, Rueckert D, Jack CR, Raunig D, Hill D, ASDN Initiative (2014) Robustness of automated hippocampal volumetry across magnetic resonance field strengths and repeat images. Alzheimer's & Dementia 10:430-438.e2
4. Pereira JB, Junqué C, Bartrés-Faz D, Ramírez-Ruiz B, Marti MJ, Tolosa E (2013) Regional vulnerability of hippocampal subfields and memory deficits in Parkinson's disease. Hippocampus 23:720–728
5. Zhu H, Wang S, Qu L, Shen D (2021) Hippocampus segmentation in MR images: multiatlas methods and deep learning methods. In Big Data in Psychiatry# x0026; Neurology, ed: Elsevier, 181–215
6. Wang Y, Ma G, Wu X, Zhou J (2018) Patch-based label fusion with structured discriminant embedding for hippocampus segmentation. Neuroinformatics 16:411–423

7. Platero C, Tobar M, Sanguino J, Velasco O (2013) A new label fusion method using graph cuts: application to hippocampus segmentation. XIII Mediterr Conf Med Biol Eng Comput 2014:174–177
8. Hao Y, Wang T, Zhang X, Duan Y, Yu C, Jiang T, Fan Y (2014) Local label learning (LLL) for subcortical structure segmentation: application to hippocampus segmentation. Hum Brain Mapp 35:2674–2697
9. Tong T, Wolz R, Coupé P, Hajnal JV, Rueckert D, AsDN Initiative (2013) Segmentation of MR images via discriminative dictionary learning and sparse coding: application to hippocampus labeling. NeuroImage 76:11–23
10. Dill V, Klein PC, Franco AR, Pinho MS (2018) Atlas selection for hippocampus segmentation: relevance evaluation of three meta-information parameters. Comput Biol Med 95:90–98
11. Langerak TR, Berendsen FF, Van der Heide UA, Kotte AN, Pluim JP (2013) Multiatlas-based segmentation with preregistration atlas selection. Med Phys 40:091701
12. Tang Z, Fan Y (2016) Groupwise image registration guided by a dynamic digraph of images. Neuroinformatics 14:131–145
13. Doshi J, Erus G, Ou Y, Resnick SM, Gur RC, Gur RE, Satterthwaite TD, Furth S, Davatzikos C, AsN Initiative (2016) MUSE: MUlti-atlas region Segmentation utilizing Ensembles of registration algorithms and parameters, and locally optimal atlas selection. NeuroImage 127:186–195
14. Wu G, Wang Q, Zhang D, Nie F, Huang H, Shen D (2014) A generative probability model of joint label fusion for multi-atlas based brain segmentation. Med Image Anal 18:881–890
15. Wang H, Suh JW, Das SR, Pluta JB, Craige C, Yushkevich PA (2013) Multi-atlas segmentation with joint label fusion. Pattern Anal Mach Intell, IEEE Trans 35:611–623
16. Zhu H, Cheng H, Yang X, Fan Y (2017) Metric learning for multi-atlas based segmentation of hippocampus. Neuroinformatics 15:41–50
17. Shen D, Wu G, Suk H-I (2017) Deep learning in medical image analysis. Annu Rev Biomed Eng 19:221–248
18. Litjens G, Kooi T, Bejnordi BE, Setio AAA, Ciompi F, Ghafoorian M, van der Laak JA, van Ginneken B, Sánchez CI (2017) A survey on deep learning in medical image analysis. Med Image Anal 42:60–88
19. Long J, Shelhamer E, Darrell T (2015) Fully convolutional networks for semantic segmentation. in Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition 3431–3440
20. Ronneberger O, Fischer P, Brox T (2015) U-net: convolutional networks for biomedical image segmentation. In Int Conf Med Image Comput Comput-Assist Interv 234–241
21. Oktay O, Schlemper J, Folgoc LL, Lee M, Heinrich M, Misawa K, Mori K, McDonagh S, Hammerla NY, Kainz B (2018) Attention u-net: learning where to look for the pancreas. arXiv preprint:1804.03999
22. Gao H, Ji S (2019) Graph u-nets. Int Conf Machine Learning 2083–2092
23. Hatamizadeh A, Tang Y, Nath V, Yang D, Myronenko A, Landman B, Roth HR, Xu D (2022) Unetr: Transformers for 3d medical image segmentation. Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision 574–584
24. Liu M, Li F, Yan H, Wang K, Ma Y, Shen L, Xu M, AsDN Initiative (2020) A multi-model deep convolutional neural network for automatic hippocampus segmentation and classification in Alzheimer's disease. NeuroImage 208:116459
25. Nogovitsyn N, Souza R, Muller M, Srajer A, Hassel S, Arnott SR, Davis AD, Hall GB, Harris JK, Zamyadi M (2019) Testing a deep convolutional neural network for automated hippocampus segmentation in a longitudinal sample of healthy participants. Neuroimage 197:589–597
26. Folle L, Vesal S, Ravikumar N, Maier A (2019) Dilated deeply supervised networks for hippocampus segmentation in MRI. In Bildverarbeitung für die Medizin 2019, ed: Springer, 68–73
27. Ataloglou D, Dimou A, Zarpalas D, Daras P (2019) Fast and precise hippocampus segmentation through deep convolutional neural network ensembles and transfer learning. Neuroinformatics 17:563–582
28. Xie Z, Gillies D (2018) Near real-time hippocampus segmentation using patch-based canonical neural network. arXiv preprint:1807.05482
29. Cao L, Li L, Zheng J, Fan X, Yin F, Shen H, Zhang J (2018) Multi-task neural networks for joint hippocampus segmentation and clinical score regression. Multimedia Tools Appl 77:29669–29686
30. Kim M, Wu G, Shen D (2013) Unsupervised deep learning for hippocampus segmentation in 7.0 Tesla MR images. In Mach Learn Med Imaging, ed: Springer 1–8
31. Zhu H, Adeli E, Shi F, Shen D (2020) FCN based label correction for multi-atlas guided organ segmentation. Neuroinformatics 18:319–331
32. Isensee F, Jaeger PF, Kohl SA, Petersen J, Maier-Hein KH (2021) nnU-Net: a self-configuring method for deep learning-based biomedical image segmentation. Nat Methods 18:203–211
33. Isensee F, Kickingereder P, Wick W, Bendszus M, Maier-Hein KH (2018) No new-net. In Int MICCAI Brainlesion Workshop 234–244.
34. Fu J, Zheng H, Mei T (2017) Look closer to see better: recurrent attention convolutional neural network for fine-grained image recognition. In Proc IEEE Conf Comput Vision Pattern Recognit 4438–4446
35. Tang C, Chen H, Li X, Li J, Zhang Z, Hu X (2021) Look closer to segment better: boundary patch refinement for instance segmentation. In Proc IEEE/CVF Conf Comput Vision Pattern Recognit 13926–13935
36. Çiçek Ö, Abdulkadir A, Lienkamp SS, Brox T, Ronneberger O (2016) 3D U-Net: learning dense volumetric segmentation from sparse annotation. In International Conference on Medical Image Computing and Computer-Assisted Intervention 424–432
37. Paszke A, Gross S, Massa F, Lerer A, Bradbury J, Chanan G, Killeen T, Lin Z, Gimelshein N, Antiga L (2019) Pytorch: an imperative style, high-performance deep learning library. Adv Neural Inf Process Syst 32
38. Milletari F, Navab N, Ahmadi S-A (2016) V-net: fully convolutional neural networks for volumetric medical image segmentation, in 3D Vision (3DV). Fourth Int Conf 2016:565–571
39. Jafari-Khouzani K, Elisevich KV, Patel S, Soltanian-Zadeh H (2011) Dataset of magnetic resonance images of nonepileptic subjects and temporal lobe epilepsy patients for validation of hippocampal segmentation techniques. Neuroinformatics 9:335–346

Guanghua He Master, is an associate professor at School of Mathematics Physics and Information, Shaoxing University. Her research focuses on statistical modeling, medical image segmentation.

Lianlian Zhou PhD, is an assistant professor at School of Mathematics Physics and Information, Shaoxing University. Her research focuses on statistical modeling.

Guying Zhang is a graduate student at School of Mathematics Physics and Information, Shaoxing University. Her research focuses on medical image segmentation.

Hancan Zhu PhD, is a professor at School of Mathematics Physics and Information, Shaoxing University. His major research directions include machine learning, medical image analysis.