An improved 3D-UNet-based brain hippocampus segmentation model based on MR images

Qian Yang[1†], Chengfeng Wang[2†], Kaicheng Pan[3], Bing Xia[3*], Ruifei Xie[3*] and Jiankai Shi[4*]

Abstract

Objective Accurate delineation of the hippocampal region via magnetic resonance imaging (MRI) is crucial for the prevention and early diagnosis of neurosystemic diseases. Determining how to accurately and quickly delineate the hippocampus from MRI results has become a serious issue. In this study, a pixel-level semantic segmentation method using 3D-UNet is proposed to realize the automatic segmentation of the brain hippocampus from MRI results. Methods: Two hundred three-dimensional T1-weighted (3D-T1) nongadolinium contrast-enhanced magnetic resonance (MR) images were acquired at Hangzhou Cancer Hospital from June 2020 to December 2022. These samples were divided into two groups, containing 175 and 25 samples. In the first group, 145 cases were used to train the hippocampus segmentation model, and the remaining 30 cases were used to fine-tune the hyperparameters of the model. Images for twenty-five patients in the second group were used as the test set to evaluate the performance of the model. The training set of images was processed via rotation, scaling, grey value augmentation and transformation with a smooth dense deformation field for both image data and ground truth labels. A filling technique was introduced into the segmentation network to establish the hippocampus segmentation model. In addition, the performance of models established with the original network, such as VNet, SegResNet, UNetR and 3D-UNet, was compared with that of models constructed by combining the filling technique with the original segmentation network. Results: The results showed that the performance of the segmentation model improved after the filling technique was introduced. Specifically, when the filling technique was introduced into VNet, SegResNet, 3D-UNet and UNetR, the segmentation performance of the models trained with an input image size of $48\times48\times48$ improved. Among them, the 3D-UNet-based model with the filling technique achieved the best performance, with a *Dice* score (*Dice* score) of 0.7989 ± 0.0398 and a mean intersection over union (*mIoU*) of 0.6669 ± 0.0540, which were greater than those of the original 3D-UNet-based model. In addition, the oversegmentation ratio (*OSR*), average surface distance (*ASD*) and Hausdorff distance (*HD*) were 0.0666 ± 0.0351, 0.5733 ± 0.1018 and 5.1235 ± 1.4397,

†Qian Yang and Chengfeng Wang contributed equally to this work.

*Correspondence:
Bing Xia
bingxia_hzch@163.com
Ruifei Xie
ruifei007@stu.xjtu.edu.cn
Jiankai Shi
shijiankai@alu.hdu.edu.cn

Full list of author information is available at the end of the article

respectively, which were better than those of the other models. In addition, when the size of the input image was set to 48×48×48, 64×64×64 and 96×96×96, the model performance gradually improved, and the *Dice* scores of the proposed model reached 0.7989±0.0398, 0.8371±0.0254 and 0.8674±0.0257, respectively. In addition, the *mIoUs* reached 0.6669±0.0540, 0.7207±0.0370 and 0.7668±0.0392, respectively. Conclusion: The proposed hippocampus segmentation model constructed by introducing the filling technique into a segmentation network performed better than models built solely on the original network and can improve the efficiency of diagnostic analysis.

Keywords Brain hippocampus segmentation, MRI, Deep learning, 3D-UNet, Filling technique

Introduction

Brain metastases are becoming an increasingly common complication of systemic cancers [1, 2]. Approximately 20–40% of patients with extracranial tumours develop cerebral metastases during the course of the disease, often leading to a poor prognosis [3, 4]. Currently, radiation therapy, including whole-brain radiation therapy (WBRT) and stereotactic radiosurgery (SRS), is still the primary treatment for both the prevention and treatment of intracranial metastases [5]. Unfortunately, patients with tumours undergoing radiation therapy often suffer brain injuries caused by radiation. Clinical trials by WBRT RTOG 0212 [6] and RTOG 0214 [7] have demonstrated that there is a significant decline in neurocognitive or intellectual function after receiving WBRT. In addition, the greater the dose is, the more severe the decrease. A study by Eric L Chang et al. indicated that patients receiving a combination of SRS and WBRT have a significantly greater risk of cognitive impairment than those who are treated with SRS alone [8]. In recent years, increasing attention has been focused on neural stem cells, which are critical sites susceptible to radiation-induced injuries [9]. Neural stem cells are located mainly in the brain hippocampus; these cells strongly proliferate and migrate and can serve as reserves, participating in the repair of cerebral lesions [9]. The subgranular zone of the hippocampus is a crucial neural centre for learning and memory [10]. Many clinical studies have demonstrated that there is a significant association between the hippocampus and cognitive state [11]. When killing tumour cells, radiation therapy can also inadvertently affect surrounding healthy cells, leading to undesirable side effects [12]. Bilateral or unilateral radiation-induced injury to the hippocampus can influence the processes involved in learning and memory formation [13]. This can lead to cognitive functional disorders and can significantly influence patients' quality of life [14]. Therefore, accurate delineation of the hippocampus region is crucial for decreasing the radiation dose and minimizing radioactive damage during radiotherapy [15, 16]. However, it is challenging to accurately delineate the hippocampus region in MR images due to the low signal-to-noise ratio (SNR), the poor quality of MR images, and the small size of the hippocampus in MR images [17]. In addition, among all brain tissues, the hippocampus is the most susceptible to ageing as life spans increase. Recent studies have shown that a decrease in hippocampal function is primarily attributed to the death of brain neurons [18, 19]. Moreover, numerous studies have shown that abnormalities in the hippocampus are closely related to the pathogenesis of epilepsy, intellectual impairment, Alzheimer's disease and other neurological pathologies [20–23]. Therefore, it is important to accurately recognize the hippocampus region for the prevention and early diagnosis of neurosystemic diseases. As a result, developing faster and more precise hippocampus segmentation methods for MR images has become a critical challenge.

Pruessner et al. proposed a manual hippocampus segmentation method, but it has limitations such as low efficiency, low accuracy and a high degree of subjectivity [24]. Strck et al. proposed a semiautomatic hippocampus segmentation method, which improved the efficiency of segmentation compared to manual methods. However, the accuracy of this segmentation method is still limited and cannot meet the demands of clinical application [25]. To achieve both high efficiency and high accuracy for hippocampus segmentation, several automatic segmentation methods based on deformation models [26–28], mapping technology [29–32] and machine learning [33] have been proposed. These methods have made significant advancements in hippocampus segmentation but still face challenges in accurately distinguishing the hippocampus from other tissues with similar grey values in MR images.

With the rapid development and application of deep learning in the field of image processing, many segmentation networks have been developed [34–36]. Ciresan et al. applied convolutional neural networks (CNNs) to segment the neuron cell membrane and achieved ideal segmentation results [37]. CNNs have been used for segmentation tasks in brain tumours [38], retina [34, 39], interstitial and epithelial tissues [35], liver tumours [40], and lung parenchyma [41]. For the segmentation of the hippocampus, Chen et al. used U-Seg-Net to segment the hippocampus with 2D sections of MR images first, and then 3D segmentation results were obtained via reconstruction of 2D segmentation results [42]. Liu and Yan combined deep learning and the lattice Boltzmann model to segment the hippocampus on MR images [43].

Yang *et al. BMC Medical Imaging*   (2024) 24:166

Page 3 of 12

A U-Net architecture was proposed by Weng et al. and applied to image segmentation [44].

These region of interest (ROI) segmentation methods achieve high accuracy, but these models were trained with 2D slices of MR images, ignoring the correlations between slices in 3D images, which can yield discontinuous and unsmooth results during segmentation [45]. To overcome this defect, Cieck et al. proposed a 3D-UNet segmentation method to detect the hippocampus with MRI [46]. This method is directly based on voxel-scale information and requires many parameters to be trained. In addition, the distribution of the target region is often discontinuous, leading to discontinuous segmentation or hollow segmentation, as shown in Fig. 1 (b) and (d), respectively.

Consequently, in this study, we introduce a filling technique into a 3D-UNet model to segment the hippocampus from MR images collected from patients who

underwent 3D-T1 magnetic resonance imaging of the brain. The methodology of this study is illustrated in Fig. 2.

Data and methods
Datasets
Magnetic resonance images
In this study, we collected 200 three-dimensional T1-weighted (3D-T1) sequence MRI scans without gadolinium contrast enhancement. This image dataset was acquired from 200 patients who underwent 3D-T1 sequence MRI scanning from June 2020 to December 2022 at Hangzhou Cancer Hospital. These data were divided into three groups for model training, validation and testing, with images for 145, 30 and 25 patients, respectively. The images for the first group of 145 patients were used to train the hippocampus segmentation model. The images for the second group of 30 cases were used to
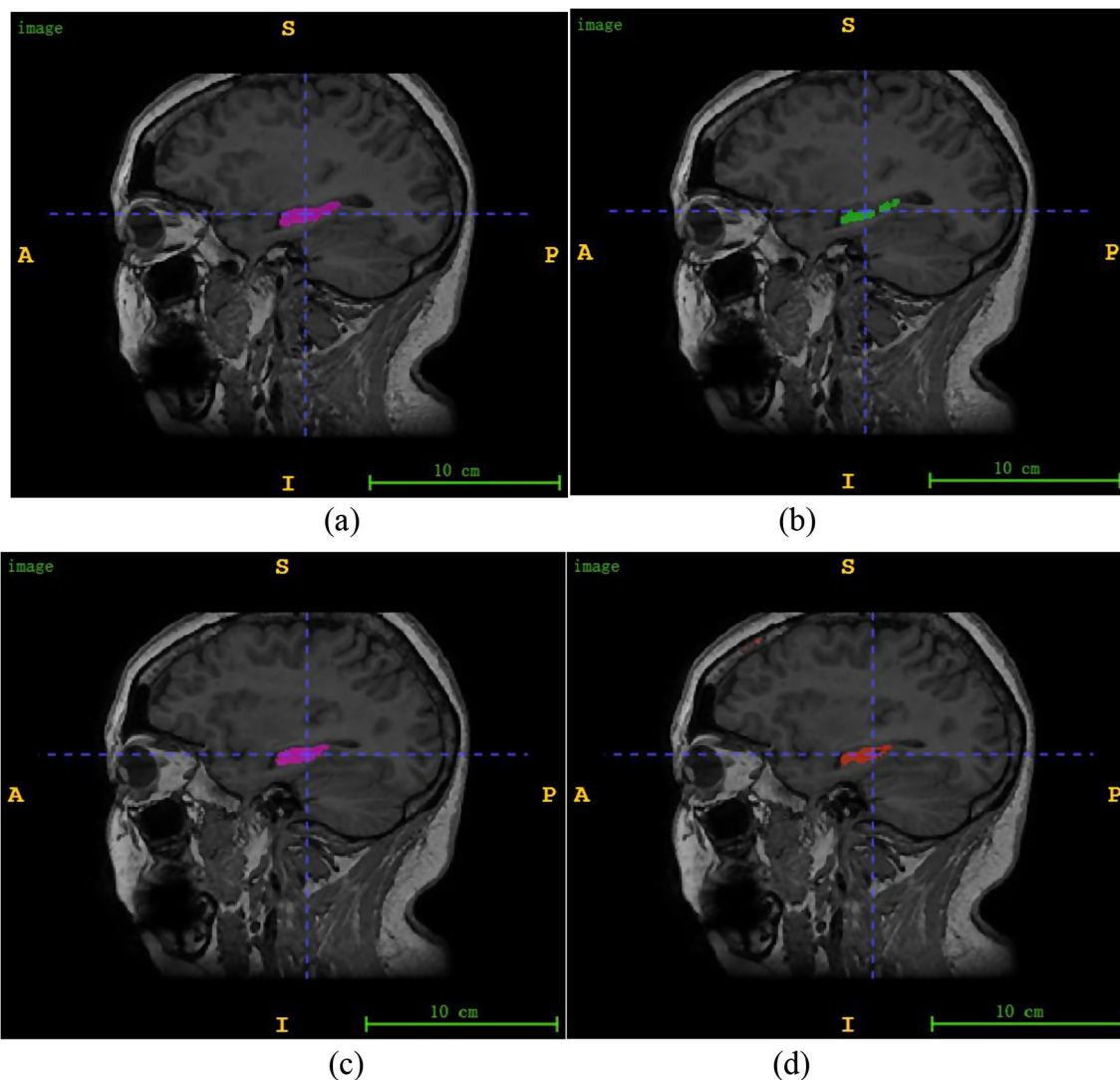

Fig. 1 **(a)** and **(c)** Ground truth for two cases; **(b)** and **(d)** Cases of 'discontinuous segmentation' and 'hollow segmentation', respectively
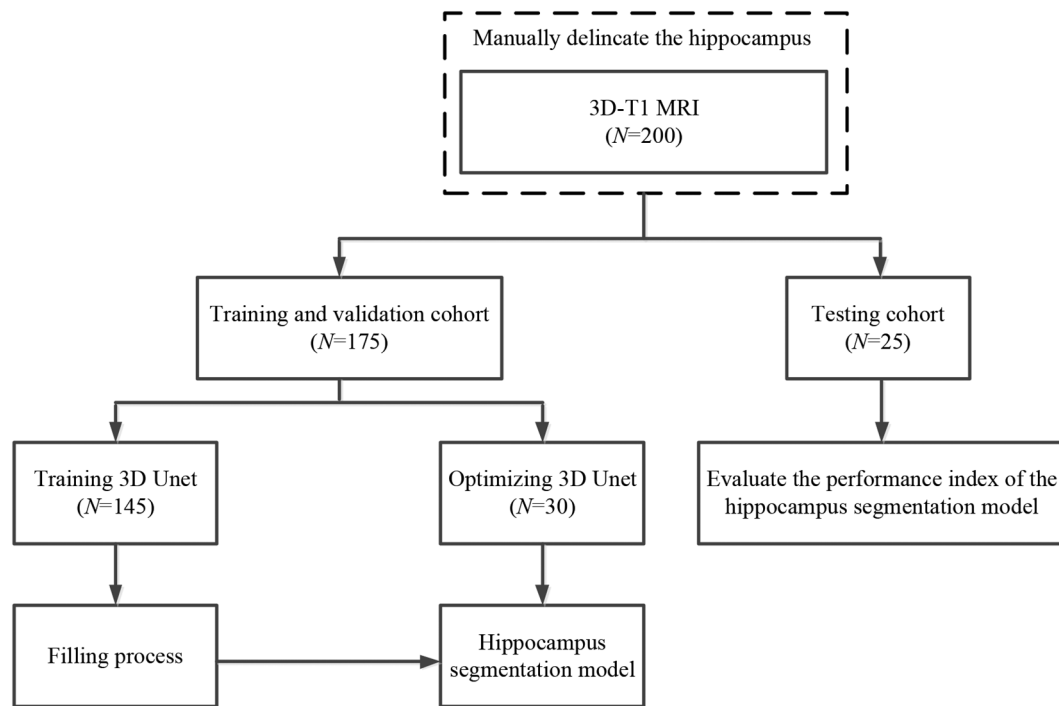
Fig. 2 The methodology of this study

fine-tune the hyperparameters of the model. The images for the third group of 25 cases were used as the test set to evaluate the performance of the segmentation model. All patients were adults over 18 years of age, and MRI confirmed that the hippocampus was not affected by any disease. The hippocampus was delineated manually from MRI scans by three deputy chief physicians following the RTOG 0933 hippocampal delineation guidelines [47, 48]. The 200 patients were numbered in sequence from 001 to 200. Images for patients who were numbered $001 \sim 070$, $071 \sim 140$ and $141 \sim 200$ were manually segmented by the first, second and third deputy chief physicians, respectively.

Automatic hippocampus segmentation model
A filling technique was introduced to 3D-UNet to establish an automatic hippocampus segmentation model. The details are as follows.

3D-UNet model
3D-UNet is a deep convolution neural network composed of an analysis path and a synthesis path, and each path has 4 resolution layers. The network structure of the 3D-UNet model is shown in Fig. 3.

The input of the network was a 128*128*128 voxel tile of an image with 3 channels. In the analysis path, each layer consists of two $3 \times 3 \times 3$ convolution layers that are activated by a rectified linear unit (ReLU); then, a $2 \times 2 \times 2$ max pooling operation with a step size of 2 is performed.

In the synthesis path, each layer is composed of an upconvolution operation with a kernel size of $2 \times 2 \times 2$ and a step size of 2; then, two $3 \times 3 \times 3$ convolution layers activated by a ReLU function are used. In the analysis path, layers with matching resolutions are connected via a shortcut, which provides the essential features for reconstruction. In the final layer of the synthesis path, the number of output channels is reduced to match the required number of output feature map channels using $1 \times 1 \times 1$ convolution. This architecture design enables highly efficient segmentation with relatively few annotated images by utilizing a weighted soft-max loss function. This approach has demonstrated excellent performance in various biomedical segmentation applications.

Filling technique for the segmentation of the hippocampus
In this model, a sliding window with a size of $96 \times 96 \times 96$ was used to identify the hippocampus from 3D-T1 sequence MR images. However, challenges were encountered due to the misrecognition of scattered voxels and the presence of continuous noise points in the image space. As a result, a maximum connected region algorithm was introduced to mitigate the influence of noise points on the recognition results. Additionally, a discontinuous distribution in the brain shell region often leads to the appearance of holes in the recognition outcomes. To address this issue, a filling technique was introduced to improve the performance of the segmentation model. Specifically, in each layer, for any two points P_1 and P_2,
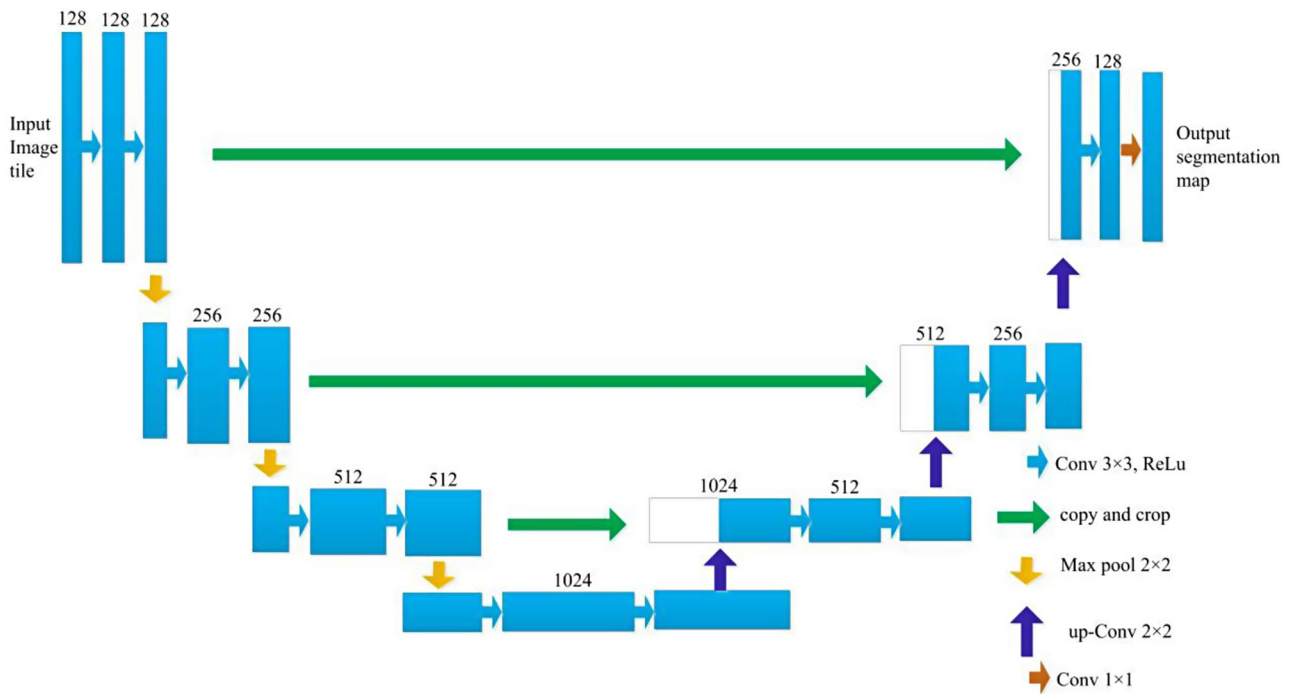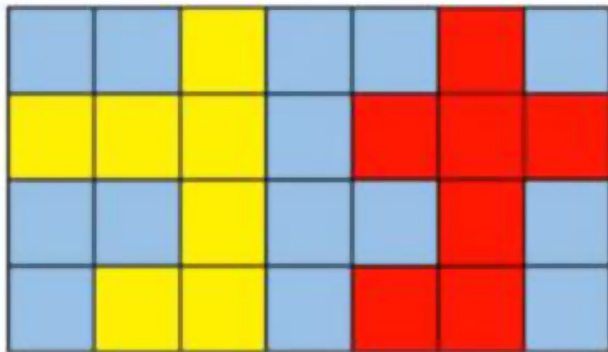
Fig. 3 Structure of 3D-UNet

Fig. 4 Illustration of the filling technique

with coordinates of (x_1, y_1) and (x_2, y_2), respectively, if the points are within a certain threshold range, i.e., the coordinates of points P_1 and P_2 satisfy formula (1), then a connection path is present.

$$(x_2 - x_1 + y_2 - y_1 + 2) \leq \theta \tag{1}$$

θ is the threshold value. An illustration of the filling technique is shown in Fig. 4. When θ is set to 1, the yellow region and the red region are two independent regions. When θ is set to 3, the yellow region and the red region are connected, forming one region, thus eliminating the disconnection between the two regions.

Training

Image enhancement was performed based on scaling, rotation and grey value augmentation in this study. In addition, a smooth dense deformation field method was used for both ground truth labels and image data. Specifically, random vectors were sampled with an interval of 32 voxels in each direction from a normal distribution. Then, *B*-spline interpolation was used. Due to the small proportion of ROIs in brain MR images and the imbalance among pixel categories between ROIs and background areas, network training with any loss function alone cannot achieve ideal results. As a result, a weighted cross-entropy loss function was used, in which the background weight was frequently decreased and the ROI weight was increased to overcome the imbalance between the pixel area of ROIs and the background region. The intensity of the input data was transformed to a range of [0, 500], which was found to provide the best contrast between the background and ROIs. Data augmentation was carried out in real time, producing a variety of different images for training iterations.

Evaluation indices for the segmentation model

To evaluate segmentation model performance, the *Dice* score [49–53], intersection over union (*IoU*) [54], over-segmentation ratio (*OSR*) [53], undersegmentation ratio (*USR*) [53], average surface distance (*ASD*) [49] and Hausdorff distance (*HD*) [49] were used in this paper, as shown in formulas (2) to (7), respectively.

The *Dice* score is one of the most commonly used metrics for assessing medical volume segmentation models [49–53]. The definition of the *Dice* score is shown in formula (2).

$$Dice\left(GT, Pred\right) = \frac{2 area_{GT} \cap area_{Pred}}{area_{GT} + area_{Pred}} \quad (2)$$

where $area_{GT}$ is the pixel area of the hippocampus in ground truth images, as delineated manually by a deputy chief physician. $area_{Pred}$ is the pixel area predicted with the segmentation model.

The intersection over union (*IoU*) is used to measure the accuracy of the segmentation model and quantify the degree of similarity between the annotated ground truth data and the region segmented with the model. The definition of the *IoU* is shown in formula (3).

$$IoU = \frac{TP}{TP + FP + FN} \quad (3)$$

In formula (3), *TP* is the pixel area of true positives, *FP* is the pixel area of false positives, and *FN* is the pixel area of false negatives.

The current medical image segmentation approaches have limitations in effectively solving the problems of oversegmentation and undersegmentation [50–53]. The metrics used to evaluate oversegmentation and undersegmentation focus on the proportions of incorrectly segmented and unsegmented pixels, which can reflect the performance of the segmentation model in detail [53]. The oversegmentation ratio (*OSR*) and undersegmentation ratio (*USR*) are defined in formulas (4) and (5), respectively.

$$OSR = \frac{FP}{R_s + T_s} \quad (4)$$

$$USR = \frac{FN}{R_s + T_s} \quad (5)$$

In formulas (4) and (5), *FP* is the pixel area of false positives, and *FN* is the pixel area of false negatives. R_s refers to the reference area of the ground truth ROI, which is delineated manually by a deputy chief physician, and T_s refers to the pixel area of the hippocampus estimated with the segmentation model.

Spatial distance-based metrics such as average surface distance and Hausdorff distance are widely applied to assess the performance of segmentation models. In this study, the average surface distance and Hausdorff distance were used to evaluate the performance of the segmentation model. The definitions of the average surface

Table 1 Experimental environment configuration

Name	Configuration
GPU	NVIDIA 2080 Ti
CPU	I9 13,900 k
MONAI	0.8.1.
Memory	128G
Operating system	Windows 11

Table 2 Experimental model parameters

Training parameters	Values
Input image size	48*48*48
Batch size	8
Epochs	600
Optimizer	Adam
Learning rate	0.001

distance and Hausdorff distance are shown in formulas (6) and (7), respectively.

$$ASD\left(A, B\right) = \frac{1}{S\left(A\right) + S\left(B\right)} \left(\sum_{s_A \in S(A)} d\left(s_A, S\left(B\right)\right) + \sum_{s_B \in S(B)} d\left(s_B, S\left(A\right)\right) \right) \quad (6)$$

In formula (6), *S(A)* and *S(B)* are the sets of surface voxels of *A* and *B*, respectively. $d\left(s_A, S(B)\right)$ indicates the shortest distance from an arbitrary voxel s_A to *S(B)*. $d\left(s_B, S(A)\right)$ indicates the shortest distance from an arbitrary voxel s_B to *S(A)*.

$$H\left(A, B\right) = \max\left(h\left(A, B\right), h\left(B, A\right)\right) \quad (7)$$

where *h(A,B)* and *h(B,A)* are the one-way Hausdorff distances between (*A, B*) and (*B, A*), respectively, as shown in Eqs. (8) and (9).

$$h\left(A, B\right) = \max_{a \in A}\left(\min_{b \in B} \|a - b\|\right\} \quad (8)$$

$$h\left(B, A\right) = \max_{b \in B}\left(\min_{a \in A} \|b - a\|\right\} \quad (9)$$

Results and analysis
Experimental environment
In this study, we carried out the experiments in a Windows system environment. Table 1 shows the information for the hardware environment. Table 2 shows the information regarding the model parameters.
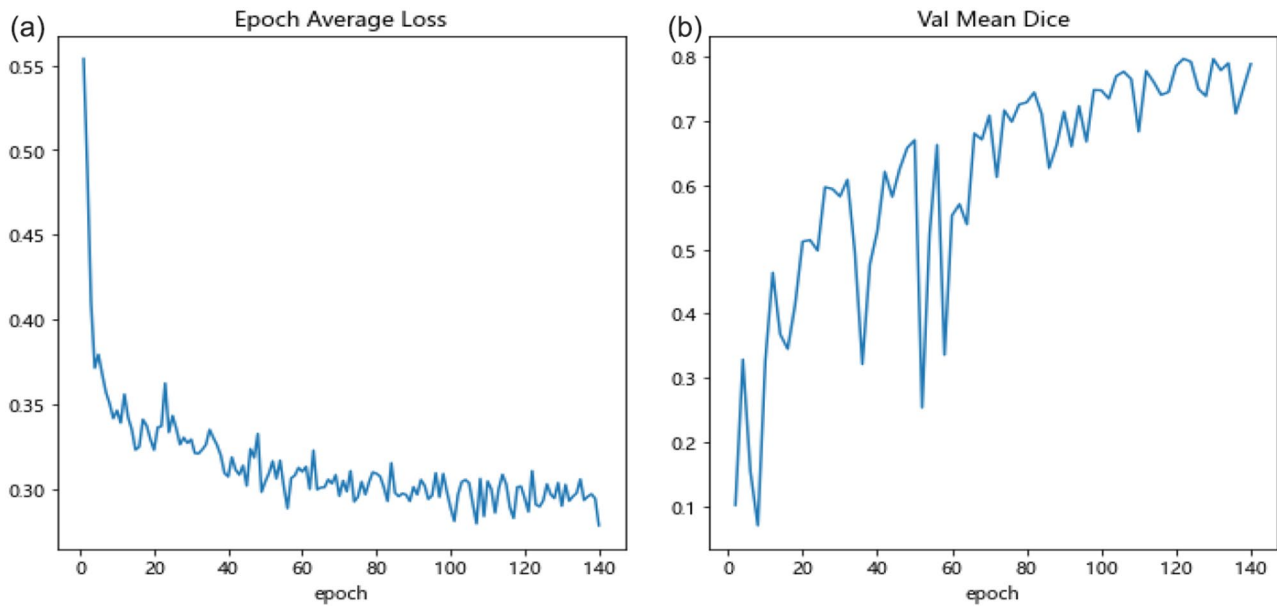
Fig. 5 Epoch average loss **(a)** and mean *Dice* score **(b)** of validation

Table 3 Comparison of segmentation performance for different models

Network model	Dice score	mIoU	OSR	USR	ASD	HD
VNet [55]	0.770	0.634	0.097	0.197	1.901	28.963
Filling technique + VNet	0.776	0.642	0.092	0.207	0.640	6.821
SegResNet [56]	0.761	0.618	0.076	0.248	3.112	75.963
Filling technique + SegResNet	0.771	0.630	0.068	0.259	0.600	7.806
UNetR	0.800	0.668	0.104	0.159	1.183	41.672
Filling technique + UNetR	0.802	0.670	0.102	0.160	0.654	5.157
3D-UNet [44]	0.5457	0.3794	0.3618	0.1281	63.3934	169.2131
Filling technique + 3D-UNet	0.7989	0.6669	0.0666	0.2218	0.5733	5.1235

Note Dice score is the *Dice* score similarity, *mIoU* is the mean intersection over union, *OSR* is the oversegmentation ratio, *USR* is the undersegmentation ratio, *ASD* is the average surface distance, and *HD* is the Hausdorff distance

Comparison of the performance of hippocampus segmentation models established with different deep learning networks

In this paper, an improved hippocampal segmentation model was proposed by introducing a filling technique to 3D-UNet. Figure 5 shows that the trend curves of the average loss and mean *Dice* coefficient of validation vary with the number of epochs. As illustrated in Fig. 5 (a) and (b), a continuous decrease in loss is observed around the 140th epoch. In addition, the loss reaches a minimum at this point. Moreover, the *Dice* coefficient gradually increases and reaches a peak around the same epoch.

To assess the effect of the filling technique, we compared the performance of models established using the original networks VNet, SegResNet, UNetR and 3D-UNet and models constructed by introducing the filling technique into different original networks in this paper. A comparison of the performance of the different models is shown in Table 3.

As shown in Table 3, the model established by introducing the filling technique into 3D-UNet performs better than the original 3D-UNet base model. The *Dice* coefficient is 0.7989, and the *mIoU* is 0.6669, which are higher than those of the original 3D-UNet-based segmentation model. The *P* value of the *Dice* score was $5.9605 \times 10^{-8} < 0.001$ according to the Wilcoxon signed-rank test, indicating a significant improvement after introducing the filling technique into the 3D-UNet-based model. In addition, the *OSR*, *ASD* and *HD* were 0.0666, 0.5733 and 5.1235, respectively, which are lower than those of the original 3D-UNet-based model. Furthermore, for other segmentation networks, such as VNet, SegResNet, and UNetR, we found that the performance of the models improved when the filling technique was added to the base models. Specifically, the *Dice* score and the *mIoU* improved compared to those of the original network-based models. Other metrics, such as *OSR*, *ASD* and *HD*, also displayed reductions. The findings demonstrate that the proposed method improved the

Table 4 Comparison of performance for models established with the 3D-UNet network and the proposed model trained with images of different sizes

size	Model	Dice score	mIOU	OSR	USR	ASD	HD
48	Original	0.5457±0.0828	0.3794±0.0774	0.3618±0.0984	0.1281±0.0580	64.3934±13.7736	169.2131±19.7572
	Proposed	0.7989±0.0398	0.6669±0.0540	0.0666±0.0351	0.2218±0.0849	0.5733±0.1018	5.1235±1.4397
64	Original	0.7271±0.0510	0.5736±0.0615	0.2023±0.0637	0.1116±0.0378	36.9662±7.9704	176.8759±13.2424
	Proposed	0.8371±0.0254	0.7207±0.0370	0.082±0.0372	0.1392±0.0447	0.5319±0.1151	4.3937±1.4050
96	Original	0.8674±0.0257	0.7668±0.0392	0.0621±0.0274	0.1241±0.0445	0.4125±0.0752	3.9032±1.3248
	Proposed	0.8674±0.0257	0.7668±0.0392	0.0621±0.0274	0.1241±0.0446	0.4124±0.0750	3.9665±1.3177

Fig. 6 Performance of the 3D-UNet-based model and the proposed model trained with images of different sizes

Table 5 Comparison of the performances of models established with the original segmentation network and with the fusion of the filling technique and the original segmentation network for images of size 96×96×96

Model	Dice score	mIOU	OSR	USR	ASD	HD
3D-UNet	0.8674±0.0257	0.7668±0.0392	0.0621±0.0274	0.1241±0.0445	0.4125±0.0752	3.9032±1.3248
Filling technique + 3D-UNet	0.8674±0.0257	0.7668±0.0392	0.0621±0.0274	0.1241±0.0446	0.4124±0.0750	3.9665±1.3177
VNet96	0.867±0.026	0.765±0.039	0.061±0.026	0.127±0.043	0.408±0.069	3.803±1.068
Filling technique + VNet96	0.867±0.026	0.765±0.039	0.061±0.026	0.127±0.043	0.408±0.069	3.866±1.065
UNetR	0.860±0.023	0.755±0.034	0.076±0.026	0.112±0.038	0.589±0.202	59.437±40.484
Filling technique + UNetR	0.861±0.023	0.756±0.035	0.075±0.026	0.113±0.038	0.459±0.069	3.834±0.709
SegResNet	0.865±0.025	0.763±0.038	0.069±0.030	0.116±0.040	0.427±0.085	7.297±16.789
Filling technique + SegResNet	0.865±0.025	0.763±0.038	0.069±0.030	0.116±0.040	0.423±0.080	3.935±0.870

performance of the hippocampus segmentation model and reduced the occurrence of false and missed detections. The model established with the proposed method can improve the precision of hippocampus segmentation based on MR images.

In addition, we further analyzed the performance of the models trained with images of different sizes. The results are shown in Table 4; Fig. 6.

Table 4; Fig. 6 show that as the size of the input image was gradually increased from 48×48×48 to 64×64×64 and 96×96×96, the performance of the segmentation model improved. When the size of training images was set to 48×48×48, 64×64×64 and 96×96×96, the *Dice* scores of the 3D-UNet-based models were 0.5457±0.0828, 0.7271±0.0510 and 0.8674±0.0257, respectively, displaying a gradually increasing trend. In addition, the mIoUs were 0.3794±0.0774, 0.5736±0.0615 and 0.7668±0.0392, respectively. Similarly, the *Dice* score and the mIoU of the proposed model gradually increased as the training image size was increased. Furthermore, when the image size was set to 48×48×48 or 64×64×64,

the performance of the proposed model was better than that of the original segmentation network-based model.

Table 5 shows that when the size of the input image was increased to 96×96×96, the performance of the segmentation model improved. For example, the *Dice* coefficient of 3D-UNet-based model trained with images of size 96×96×96 is 0.864±0.024, which is 7.5% greater than that of model trained with images of size 48×48×48 for the same network. In addition, the *mIoU* was improved by 10.91%. For other network-based segmentation models, we also found that the model performance improved when the input image size was increased. A comparison of the original network-based models and models that combined filling and the original network-based algorithms indicated that the performance levels were comparable. The improvement in performance achieved with large numbers of training images was not as obvious as the improvement in performance achieved with small numbers of training images. The possible reason is that the large input size mitigates discontinuous segmentation.
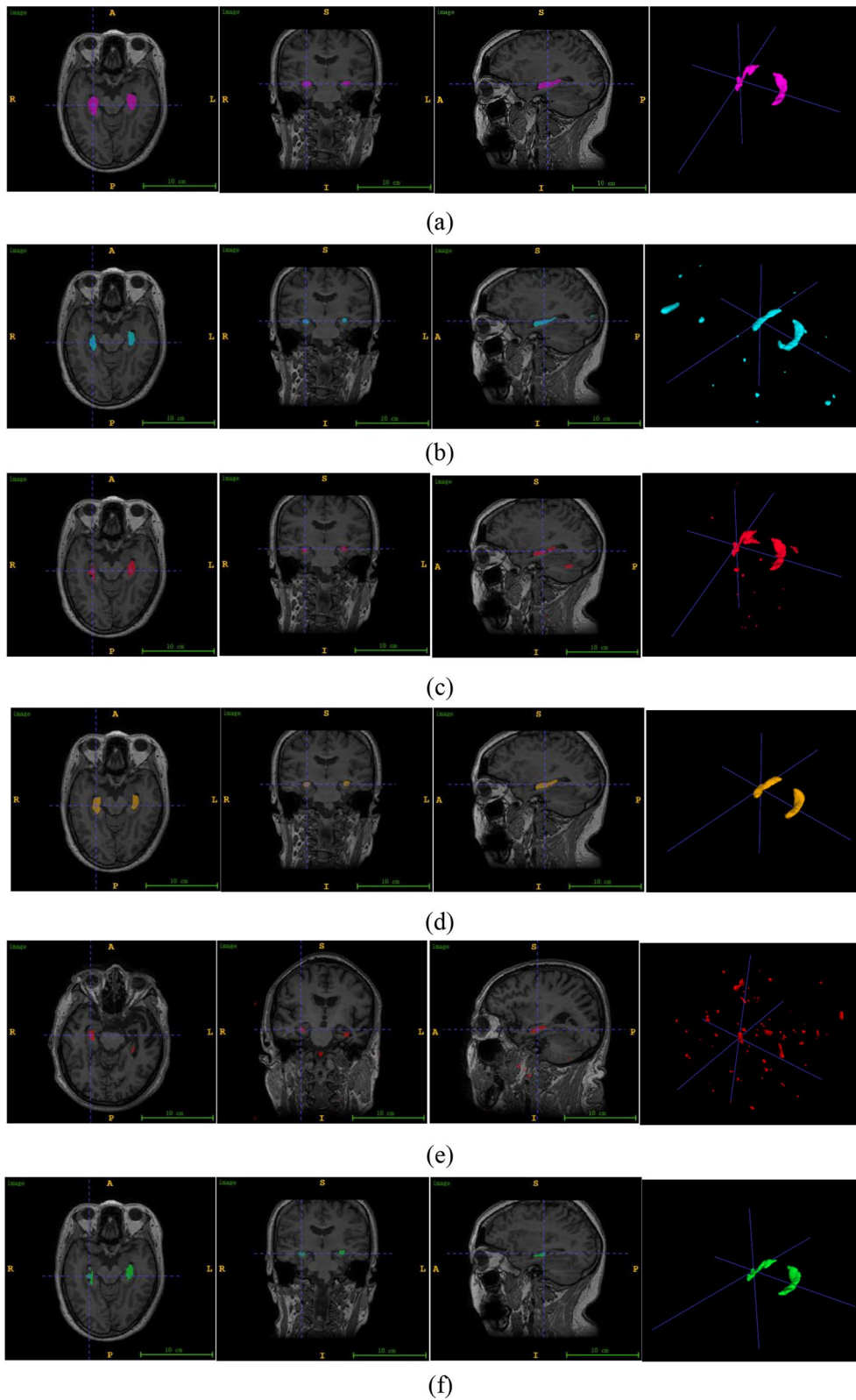
Fig. 7 **(a)** is the ground truth. **(b)**, **(c)**, **(d)** and **(e)** show the results of hippocampus segmentation with models established based on VNet, UNetR, SegResNet and 3D-UNet, respectively. **(f)** Results of hippocampus segmentation with the model established by introducing the filling technique into 3D-UNet

Yang *et al. BMC Medical Imaging*   (2024) 24:166

Page 10 of 12

Figure 7 (a) shows the ground truth for one MR image. Figure 7 (b) to (f) show the results of hippocampus segmentation with models established based on VNet, UNet, SegResNet, and 3D-UNet and by introducing the filling technique into 3D-UNet with an image size of 48×48×48.

The performance of different models was assessed, and incorrect segmentation results were obtained with the original network-based segmentation models. By introducing the filling technique into 3D-UNet to establish the segmentation model, the accuracy of segmentation was improved, and details were enhanced and effectively extracted.

Discussion

Brain metastasis is a common complication of systemic cancer. Brain radiotherapy is the primary method used to prevent and treat intracranial metastatic tumours. However, radiotherapy can cause damage to the hippocampus, leading to cognitive impairment, which severely affects patients' quality of life. Therefore, accurate segmentation of the hippocampus from MR images is essential for minimizing radiation damage. With the rapid development and application of deep learning in the field of image processing, convolutional neural network-based segmentation algorithms, such as UNet and 3D-UNet, have been developed for hippocampus segmentation based on MR images. One limitation of the 3D-UNet-based model is the discontinuous distribution of the target region in the segmentation results, leading to poor recognition outcomes. Thus, there is a need to improve the accuracy of hippocampus segmentation models. In this study, we propose an improved model by introducing a filling technique to 3D-UNet for hippocampus segmentation based on high-precision MR images. Four deep learning-based models built on the basis of VNet, SegResNet, UNetR and 3D-UNet were constructed to facilitate analysis and comparison with our improved segmentation model. This work has several unique characteristics.

First, a filling technique is introduced into 3D-UNet. The results show that the performance of the model established by introducing the filling technique into 3D-UNet is improved. Notably, there was an increase of 3.22% for the *Dice* score and 4.47% for the *mIoU*. The *OSR* and *ASD* of the improved model were also better than those of the original network-based model.

Second, in the field of hippocampal segmentation, some improved models have been reported. For example, Tang et al. designed a multichannel, landmark large-deformation diffeomorphic mapping method to segment the hippocampus, which yielded a *Dice* score of 0.76 [57]. Hänsch et al. proposed a CNN-based hippocampal segmentation method that achieved a median *Dice* score of 0.76 [58]. Somasundaram and Genish proposed an atlas-based approach to segment the hippocampus from MRI results, and it achieved a *Dice* score of 0.82 [59]. Lin et al. proposed a 3D multiscale multiattention UNet for hippocampal segmentation, and it achieved a Dice coefficient of 0.827 [60]. Although different test datasets were applied in these studies, the *Dice* score in this study was found to be competitive with those of the reported models.

Third, when the filling technique was used for other segmentation networks, such as VNet, SegResNet, and UNetR, the performance of the models was all improved. The results indicate that introducing the filling technique into segmentation networks can improve the performance of hippocampus segmentation models.

Finally, when the size of the input images was set to 96×96×96, the performance of the segmentation model improved. The *Dice* coefficient of the 3D-UNet-based model trained with images of size 96×96×96 was 0.864±0.024, which was 7.5% greater than that of the model trained with images of size 48×48×48. In addition, the *mIoU* improved by 10.91%. This may be attributed to the fact that the large input size helped avoid discontinuous segmentation issues.

Conclusion

For patients with tumours, brain radiotherapy is often a necessary treatment. If the hippocampus is not adequately protected during brain radiotherapy, the patient's cognitive functions may be adversely affected. Therefore, when radiotherapy is conducted, it is crucial to precisely delineate the hippocampus to avoid irradiation, which could reduce the impact on the patient's cognition. In this paper, a hippocampus segmentation algorithm based on a deep learning network and MR images was studied. The current segmentation networks are limited by low accuracy and lengthy processing times. To address these issues, we collected MR images from patients who had undergone brain 3D-T1 magnetic resonance scans. Then, we performed segmentation experiments with the segmentation network. A filling technique was introduced into the original segmentation network to establish a hippocampal segmentation model. MR images were used to validate the accuracy of our automated hippocampal segmentation model. The experimental results demonstrated that our method can effectively segment the hippocampus and improve the *Dice* score, making it highly valuable for hippocampal segmentation tasks. However, due to the limitations of the dataset, further studies based on additional brain MRI data are needed.

Author contributions
Conceptualization, Ruifei Xie; Formal Analys, Qian Yang; Investigation, Ruifei Xie; Data curation, Ruifei Xie, Kaicheng Pan and Jiankai Shi; Funding acquisition, Ruifei Xie and Bing Xia; Methodology, Ruifei Xie and Qian Yang; Writing—original draft Chengfeng Wang and Bing Xia; Writing—Review, Qian

Yang *et al. BMC Medical Imaging*   (2024) 24:166

Page 11 of 12

Data availability
The data presented in this study are available upon request from the corresponding author.

Declarations

Ethics approval and consent to participate
This study was performed in accordance with the principles of the Declaration of Helsinki. Approval was granted by the Medical Ethics Committee of Hangzhou Cancer Hospital. Informed consent was obtained from all individual participants included in the study.

Consent for publication
Not applicable.

Competing interests
The authors declare no competing interests.

Author details
[1]Information Technology Center, Taizhou University, 1139 Shifu Dadao, Taizhou City, Zhejiang Province, China
[2]College of Mathematics and Computer Science, Zhejiang A & F University, 666 Wusu Street, Hangzhou 311300, China
[3]Hangzhou Cancer hospital, 34 YanGuan Lane, Hangzhou 310002, China
[4]School of Computer Science, Hangzhou Dianzi University, Xiasha Higher Education Zone, Hangzhou, Zhejiang 310018, People's Republic of China

References

1. Taillibert S, Le Rhun É. Epidemiology of brain metastases. Cancer Radiother. 2015;19(1):3–9.
2. Pan K, Zhao L, Gu S, et al. Deep learning-based automatic delineation of the hippocampus by MRI: geometric and dosimetric evaluation[J]. Radiation Oncol (London England). 2021;16(1):12.
3. Tsao MN, Xu W, Wong RK, et al. Whole brain radiotherapy for the treatment of newly diagnosed multiple brain metastases[J]. Cochrane Database of Systematic Reviews; 2018.
4. Adult Central. Nervous System Tumors Treatment (PDQ®)–Health Professional Version.
5. Habets EJJ, Dirven L, Wiggenraad RGJ, et al. Neurocognitive functioning and health-related quality of life in patients treated with stereotactic radiotherapy for brain metastases: a prospective study. Neurooncology. 2016;18(3):435–44.
6. Aaron H, Wolfson KB, Ritsuko Komaki CAM, Primary Analysis of a Phase II Randomized Trial Radiation Therapy Oncology Group (RTOG). 0212: impact of different total doses and schedules of prophylactic cranial irradiation on chronic neurotoxicity and quality of life for patients with Limited-Disease Small-Cell Lung Cancer. Int J Radiat Oncol Biol Phys. 2011;81(1):77–84.
7. Gore E, Bae K, Wong SJ, et al. Phase III comparison of prophylactic cranial irradiation versus observation in patients with locally advanced non-small-cell lung cancer: primary analysis of Radiation Therapy Oncology Group study RTOG 0214. J Clin Oncol. 2011;29(3):272–8.
8. Chang EL, Wefel JS, Hess KR, et al. Neurocognition in patients with brain metastases treated with radiosurgery or radiosurgery plus whole-brain irradiation: a randomised controlled trial. Lancet Oncol. 2009;10(11):1037–44.
9. Madden DJ, Spaniol J, Costello MC, et al. Cerebral white matter integrity mediates adult age differences in cognitive performance. J Cogn Neurosci. 2009;21(2):289–302.
10. Barani IJ, Benedict SH, Lin PS. Neural stem cells: implications for the Conventional Radiotherapy of Central Nervous System malignancies. Int J Radiat Oncol Biol Phys. 2007;68(2):324–33.
11. Bálentová SHE, Kinclová I. Radiation-induced long-term alterations in hippocampus under experimental conditions. Klin Onkol. 2012;25(2):110–6.
12. Chapman JD, Reuvers AP, Borsa J. Chemical Radioprotection and Radiosensitization of mammalian cells growing in Vitro. Radiat Res. 1973;56(2):291–306.
13. Tofilon PJ, Fike JR. The Radioresponse of the Central Nervous System: a dynamic process. Radiat Res. 2000;153(4):357–70.
14. Gondi V, Mehta MP, Pugh SL, et al. Memory preservation with conformal avoidance of the Hippocampus during Whole-Brain Radiation Therapy for patients with brain metastases: primary endpoint results of RTOG 0933. Int J Radiat Oncol Biol Phys. 2013;87(5):1186.
15. Mukesh M, Benson R, Jena R, et al. Interobserver variation in clinical target volume and organs at risk segmentation in post-parotidectomy radiotherapy: can segmentation protocols help? Br J Radiol. 2012;85(1016):e530–6.
16. Walker GV, Awan MJ, Tao R, et al. Prospective randomized double-blind study of atlas-based organ-at-risk autosegmentation-assisted radiation planning in head and neck cancer. Radiother Oncol. 2014;112(3):321–5.
17. Gondi V, Rowley WTHA. MP. Mehta. Hippocampal contouring: a contouring atlas for RTOG 0933. 2011.
18. Yang W, Ang LC, Strong MJ. Tau protein aggregation in the frontal and entorhinal cortices as a function of aging[J]. Brain Res Dev Brain Res. 2005;156(2):127–38. https://doi.org/10.1016/j.devbrainres.2005.02.004.
19. Yanhong Z, Chuangchuang T, Dong W, et al. The biomarkers for identifying preclinical Alzheimer's Disease via Structural and functional magnetic resonance Imaging[J]. Front Aging Neurosci. 2016;8(92):1–4. https://doi.org/10.3389/fnagi.2016.00092.
20. Šimić,Goran, Kostović, Winblad I et al. Volume and number of neurons of the human hippocampal formation in normal aging and Alzheimer's disease.[J].Journal of Comparative Neurology, 2015. DOI:10.1002/(SICI)1096-9861(19970324)379:4<482::AID-CNE2>3.0.CO;2-Z.
21. Moodley KK, Chan D. The hippocampus in neurodegenerative disease.[J].Monographs in neural sciences, 2014, 34:95–108. https://doi.org/10.1159/000356430.
22. Chatzikonstantinou A. Front Neurol Neurosci. 2014;34:121. https://doi.org/10.1159/000356435. Epilepsy and the hippocampus.[J].
23. Yi-En Quek, Fung YL, Mike W-L. Cheung,et al.Agreement Between Automated and Manual MRI Volumetry in Alzheimer's Disease: A Systematic Review and Meta-Analysis[J]. J Magn Reson Imaging: JMRI. 2022;256. https://doi.org/10.1002/jmri.28037.
24. Pruessner JC LiL, Serles M. w, Volumetry of hippocampus and amygdala with high-resolution MRI and three-dimensional analysis software: minimizing the discrepancies between laboratories[J]cerebral cortex, 2000,10(4):433–42.
25. Niessen WJ, Pluim JP, w, Viergever MA, et al. Medical Image Computing and Computer assisted Intervention-MICCAI 2000[M]. Berlin Heidelberg: Springer; 2000. pp. 129–42.
26. Cootes TF, Taylor CJ, Cooper DH, et al. Active shape models-their training and application[J]. Comput Vis Image Underst. 1985;61(1):38–59.
27. Kumar R, Snakes. Active contour Models[J]. Internmational J Comput Vis. 1988;1(4):321–31.
28. Cootes TF, Edwards GJ, Taylor CJ. Active appearance models[J]. IEEE Trans PatternAnalysis Mach Intell. 2001;23(6):681–5.
29. Haller JW, Christensen GE, Joshi S et al. Digital atlas-based segmentation of the hippocampus[C].Proc. of International Symposium on Computer and Communication Systems for Image GuidedDiagnosis and Therapy. Avenue,NY:Elsevier Science, 1995.152-157.
30. Rao A, Sanchez-Ortiz GI, Chandrashekara R, et al. Construction of a cardiac motion atlas from MR using non-rigid registration[J]. Lect Notes Comput Sci. 2003;2674(1):141–50.
31. Heckemann RA, Hajnal JV, Aljabar P, et al. Automatic anatomical brain MRI segmentation combining label propagation and decision fusion[J]. NeuroImage. 2006;33(1):115–26.
32. Anagnostis A, Tagarakis AC, Kateris D, et al. Orchard Mapp Deep Learn Semantic Segmentation[J] Sens. 2021;21(11):3813.
33. Hao Y, Wang T, Zhang X, et al. Local label learning (LLL) for Subcortical structure segmentation:application to Hippocampus Segmentation[J]. Hum Brain Mapp. 2014;35(6):2674–97.
34. Dasgupta A, Singh S. A fully convolutional neural network based structured prediction approach towards the retinal vessel segmentation[C]. In2017 IEEE 14th International Symposium on Biomedical Imaging (ISBI 2017),2017:248–251.

35. Xu J, Luo X, Wang G, Gilmore H, Madabhushi A. A deep convolutional neural network for segmenting and classifying epithelialand stromal regions in histopathological images[J]. Neurocomputing, 2016,191214–223.

36. Krizhevsky A, Sutskever I, Hinton GE. ImageNet classification with deep convolutional neural networks. Adv Neural Inf Process Syst. 2012;25:1–9.

37. Ciresan D, Giusti A, Gambardella L, Schmidhuber J. Deep neural networks segment neuronal membranes in electron microscopy images. Adv Neural Inf Process Syst. 2012;25:2852–60.

38. Bjoern H, Menze A, Jakab S, Bauer J, Kalpathy-Cramer et al. The multimodal brain tumor image segmentation benchmark (brats)[J].IEEE transactions on medical imaging, 2015, 34(10): 1993–2024.

39. Julian Zilly, Joachim M, Buhmann D, Mahapatra. Glaucoma detection using entropy sampling and ensemble learning for automatic optic cup anddisc segmentation[J]. Comput Med Imaging Graph. 2017;55:28–41.

40. Meng L, Tian Y, Bu S. Liver tumor segmentation based on 3D convolutional neural network with dual scale. J Appl Clin Med Phys. 2020;21(1):144–57.

41. Xu M, Qi S, Yue Y, Teng Y, Xu L, Yao Y, Qian W. Segmentation of lung parenchyma in CT images using CNNtrained with the clustering algorithm generated dataset[J]. Biomedical Engi-neering Online. 2019;18(1):1–21.

42. Chen Y, Shi B, Wang Z et al. Hippocampus segmentation through multi-view ensemble ConvNets[C].Proc. of the 14th International Symposium on Biomedical Imaging.Piscataway, NJ: IEEE,2017.192-196.

43. Liu Y, Yan Z. A Combined Deep-Learning and Lattice Boltzmann Model for Segmentation of the Hippocampus in MRI[J]. Sensors. 2020;20(13):3628.

44. Olaf Ronneberger P, Fischer, Brox T. U-net: Convolutional networks for biomedical image segmentation. In International Conference on Medical image computing and computer-assisted intervention,234–241. Springer, 2015.

45. Zhang C, Hua Q, Chu Y. Pengwei Wang,Liver tumor segmentation using 2.5D UV-Net with multi-scale convolution,Computers in Biology and Medicine,2021,133:104424.

46. Cicek O, Abdulkadir A, Lienkamp SS et al. 3DUnet: learning dense volumetric segmentation fromsparse annotation[C]. Proc. of the 19th International Conference on Medical Image Computing and Computer-Assisted Intervention. Berlin, German: Springer. 2016:424–432.

47. Vinai Gondi SL, Pugh WA, Tome et al. Preservation of memory with conformal avoidance of the hippocampal neural stem-cell compartment during wholebrain Radiotherapy for Brain metastases (RTOG 0933): a phase II multi-institutional Trial[J]. J Clin Oncol 2014 32:34, 3810–6.

48. Gondi V, Tome WA, Rowley HA, Mehta MP. Hippocampalcontouring: a contouring atlas for RTOG 0933.www.rtog.org.

49. Taha AA, Hanbury A. Metrics for evaluating 3D medical image segmentation: analysis, selection, and tool [J]. BMC Med Imaging. 2015;15:29.

50. Fiandra C, Rosati S, Arcadipane F, et al. Active bone marrow segmentation based on computed tomography imaging in anal cancer patients: a machine-learning-based proof of concept [J]. Phys Med. 2023;113:102657.

51. Wu S, Bai X, Cai L, et al. Bone tumor examination based on FCNN-4s and CRF fine segmentation fusion algorithm. J Bone Oncol. 2023;42:100502.

52. Polizzi A, Quinzi V, Ronsivalle V, et al. Tooth automatic segmentation from CBCT images: a systematic review [J]. Clin Oral Invest. 2023;27:3363–78.

53. Femi D. Plant leaf infected spot segmentation using robust encoder-decoder cascaded deep learning model[J]. Manapakkam Anandan Mukunthan. Network: Computation in Neural Systems; 2023.

54. Vinayahalingam S, Kempers S, Schoep J, et al. Intra-oral scan segmentation using deep learning[J]. BMC Oral Health. 2023;23(1):643.

55. Milletari F, Navab N, Ahmadi SA. V-Net: fully convolutional neural networks for Volumetric Medical Image Segmentation[J]. arXiv e-prints; 2016.

56. Andriy Myronenko. 3D MRI brain tumor segmentation using autoencoder regularization[J].arXiv e-prints,2018.

57. Tang X, Mori S, Ratnanather T et al. Segmentation of hippocampus and amygdala using multi-channel landmark large deformation diffeomorphic metric mapping[C]//Bioengineering Conference (NEBEC), 2012 38th Annual Northeast.IEEE, 2012.https://doi.org/10.1109/NEBC.2012.6207140.

58. Hnsch A, Moltz JH, Geisler B, et al. Hippocampus segmentation in CT using deep learning: impact of MR versus CT-based training contours.[J]. J Med Imaging. 2020;6. https://doi.org/10.1117/1.JMI.7.6.064001.

59. Somasundaram K, Genish T, Kalaiselvi T. An atlas based approach to segment the hippocampus from MRI of human head scans for the diagnosis of Alzheimer's disease[J]. 2015.

60. Lin L et al. A 3D Multi-Scale Multi-Attention UNet for Automatic Hippocampal Segmentation, 2021 7th Annual International Conference on Network and Information Systems for Computers (ICNISC), Guiyang, China, 2021, pp. 89–93, https://doi.org/10.1109/ICNISC54316.2021.00025.

Publisher's Note