

基于空间域与频域特征自适应融合和类间边界区域增强的三维海马分割

白贺^{1,2} 滕野^{1,2} 冯蕾^{1,2} 孟海伟^{1,2} 汤煜春^{1,2} 刘树伟^{1,2*}

(1. 山东大学基础医学院解剖学与神经生物学系, 断层解剖学与数字人研究院, 济南 250012;
2. 山东大学脑与类脑科学研究院, 济南 250012)

[摘要] **目的** 海马萎缩是诊断阿尔茨海默病等诸多精神疾病的临床重要标志, 因此准确分割海马是一个重要的科学问题。随着深度学习的发展, 人们提出了大量先进的自动分割方法。然而, 由于 MRI 中各种噪声的影响以及海马不同类别之间不清晰的边界, 三维海马分割仍然具有挑战性。因此本文旨在提出新的自动分割方法来更精确地分割海马头、体、尾。**方法** 为了克服这些挑战, 本文提出了两个策略。一种是空间域与频域特征自适应融合策略, 通过快速傅立叶变换和卷积自动选择合适的频率组合, 减少噪声对特征提取的影响。另一种是类间边界区域增强策略, 它允许网络通过加权每个类之间边界区域的损失函数来增强对边界区域的学习, 以达到精确定位边界和调节海马头、体、尾大小的目的。**结果** 在 50 例青少年大脑 MRI 数据集上进行的实验表明, 我们的方法实现了较先进的海马分割, 海马头、体、尾相较于现有的方法都取得了一定的提升。消融实验证明我们提出的两种策略有效, 我们还在 260 例 Task04_Hippocampus 数据集上验证了网络具有强大的泛化能力, 说明本文提出的方法可用于更多的海马分割场景。**结论** 我们提出的方法可以帮助临床医生更清楚地观测海马萎缩, 并完成更精确的病情诊断和追踪。

[关键词] 空间域与频域特征自适应融合; 类间边界区域增强; 医学图像分割; 海马分割; 人

[中图分类号] R445.2 **[文献标志码]** A **[DOI]** 10.16098/j.issn.0529-1356.2024.01.011

3D hippocampal segmentation based on spatial and frequency domain features adaptive fusion and inter-class boundary region enhancement

BAI He^{1,2}, TENG Ye^{1,2}, FENG Lei^{1,2}, MENG Hai-wei^{1,2}, TANG Yu-chun^{1,2}, LIU Shu-wei^{1,2*}

(1. Institute of Sectional Anatomy and Digital Human, Department of Anatomy and Neurobiology, School of Basic Medical Sciences, Shandong University, Ji'nan 250012, China; 2. Institute of Brain and Brain-Inspired Sciences, Shandong University, Ji'nan 250012, China)

[Abstract] **Objective** Hippocampal atrophy is a clinically important marker for the diagnosis of many psychiatric disorders such as Alzheimer's disease, so accurate segmentation of the hippocampus is an important scientific issue. With the development of deep learning, a large number of advanced automatic segmentation method have been proposed. However, 3D hippocampal segmentation is still challenging due to the effects of various noises in MRI and unclear boundaries between various classes of the hippocampus. Therefore, the aim of this paper is to propose new method to segment the hippocampal head, body, and tail more accurately. **Methods** To overcome these challenges, this paper proposed two strategies. One was the spatial and frequency domain features adaptive fusion strategy, which reduced the influence of noise on feature extraction by automatically selecting the appropriate frequency combination through fast Fourier transform and convolution. The other was an inter-class boundary region enhancement strategy, which allowed the network to focus on learning the boundary regions by weighting the loss function of the boundary regions between each class to achieve the goal of pinpointing the boundaries and regulating the size of the hippocampal head, body and tail. **Results**

[收稿日期] 2023-07-24 **[修回日期]** 2023-08-25

[基金项目] 国家自然科学基金(81371533, 31571237, 31771328); 山东省重点研发计划项目(2017CXGC1501, 2019JZZY020106)

[作者简介] 白贺(1999—), 男(汉族), 吉林省松原市人, 在读硕士研究生。

* 通讯作者(To whom correspondence should be addressed)

E-mail: liusw@sdu.edu.cn Tel: (0531) 88382171

Experiments performed on a 50-case teenager brain MRI dataset show that our method achieves state-of-the-art hippocampal segmentation. Hippocampal head, body and tail had been improved compared to the existing method. Ablation experiments demonstrated the effectiveness of our two proposed strategies, and we also validated that the network had a strong generalization ability on a 260-case Task04_Hippocampus dataset. It was shown that the method proposed in this paper could be used in more hippocampal segmentation scenarios. **Conclusion** The method proposed in this paper can help clinicians to observe hippocampal atrophy more clearly and accomplish more accurate diagnosis and follow-up of the condition.

[**Key words**] Spatial and frequency domain features adaptive fusion; Inter-class boundary region enhancement; Medical image segmentation; Hippocampal segmentation; Human

据统计,全球约有 5000 万老年人患有痴呆症,其中约 70% 为阿尔茨海默病 (Alzheimer’s disease, AD)^[1]。研究证明,AD^[2] 和抑郁症^[3] 等精神疾病均会产生不同程度的海马萎缩。此外,海马体积萎缩已成为公认的 AD 生物标志^[4],用于疾病追踪和支持临床诊断^[5]。因此,需要对海马进行精确的分割,以便于观察和测量其体积变化。一些研究人员,将海马分为头、体、尾 3 部分对其进行了深入地功能研究。头部和体部占据了大部分区域,负责记忆、遗忘、情绪和注意力,而较小的尾部负责空间和时间上的感知和定位^[6]。不同的功能代表不同的疾病表现,精确分割海马的头部、体部和尾部具有重要的临床研究价值和意义。根据分割协议[例如,欧洲 AD 协会 (European Alzheimer’s Disease Consortium, EADC) 和 AD 神经成像倡议中心 (Alzheimer’s Disease Neuroimaging Initiative, ADNI) 海马分割协议^[7],即 EADC-ADNI 分割协议],手动分割海马不仅需要大量的时间成本,还需要专业且经验丰富的临床影像医生。因此,使用先进的深度学习方法自动分割海马是当务之急。

为了满足临床需求并减轻医生的负担,我们首先根据 EADC-ADNI 分割协议,对大脑 MRI 图像的海马头部、体部和尾部进行更详细的分割标注,并提出了一种基于深度学习的网络框架来完成多类自动分割任务。海马 MRI 图像分割面临着一些固有的挑战:(1)海马与其周围组织之间的边界不清晰,以及海马内部结构的相似性导致海马的前景与背景之间,以及海马的头部、体部和尾部之间的差异不明显,使得网络难以准确识别类间边界,并且容易对具有相似灰度值的一些像素进行错误分类;(2)由于扫描仪器和扫描操作等因素,所获得的 MRI 图像往往包含大量噪声,导致图像模糊,这意味着图像中包含大量隐藏的低频信息。然而,目前的深度学习网络大多只在空间域上进行卷积运算来完成特征提取,对高频信息(如边缘、纹理等)具有良好的学习能力,但却难以捕获和挖掘低频信息。为了解决上述问题,我们以极其简洁的 3D U-Net^[8] 为骨干网络,并在网络中添加了深度监督机制和两种新颖的策略。

资料和方法

1. 数据集

1.1 青少年大脑 MRI 数据集:招募 90 名中国志愿者(50 名男性和 40 名女性),平均年龄 (17.30 ± 1.58) 岁。使用具有标准 8 通道头部线圈的 3.0T GE (General Electric, Milwaukee, USA) MRI 扫描仪。T1 加权结构 MRI 扫描参数如下:重复时间 (repetition time, TR) = 18 ms,回波时间 (echo time, TE) = 10 ms,体素大小 = 0.47 mm × 0.47 mm × 0.70 mm,视野 (view, FOV) = 24.0 cm × 24.0 cm,矩阵大小 = 512 × 512,角度 = 10°,切片厚度 = 1.4 mm,切片间隙 = 0.7 mm,激发次数 (excitation frequency, NEX) = 2,总扫描时间 = 12 min。所有图像均沿穿过前后连续体的水平线进行扫描。随机选择 50 例作为我们的数据集进行注释,并以 4:1 的比例将 50 个样本分为训练集和测试集,以使用 5 折交叉验证方法来证明我们方法的有效性。

1.2 Task04_Hippocampus 数据集: Task04 _Hippocampus 数据集由 195 张 MRI 图像组成,这些图像来自美国纳什维尔范德比尔特大学医学中心的 90 名健康成年人和 105 名患有非情绪性精神病的成年人,使用 3D T1 加权 MPRAGE 作为成像序列。相应的靶感兴趣区域 (region of interest, ROI) 是海马体的前部和后部,从我们的数据集来看,前部代表海马头部,后部代表海马体部和尾部的组合^[9]。此数据集可从主要挑战网站下载 (http://medicaldecathlon.com/)。由于人脑同时拥有左右海马,195 例 MRI 图像包含 390 例数据,其中 260 例具有明显的标注,另外 130 例无。我们将这 260 个样本的数据以相同的 4:1 的比例划分为训练集和测试集,并使用 5 折交叉验证来证明方法的有效性。

2. 方法

2.1 总体框架:我们所提出的网络的总体框架如图 1A 所示,它使用类似于 3D U-Net 的编码器-解码器架构作为骨干网络。首先,从输入的 3D 大脑 MRI 数据中提取 ROI 海马区域,通过裁剪和零填充将体积调整为固定大小 (H × W × D),以便于后续的

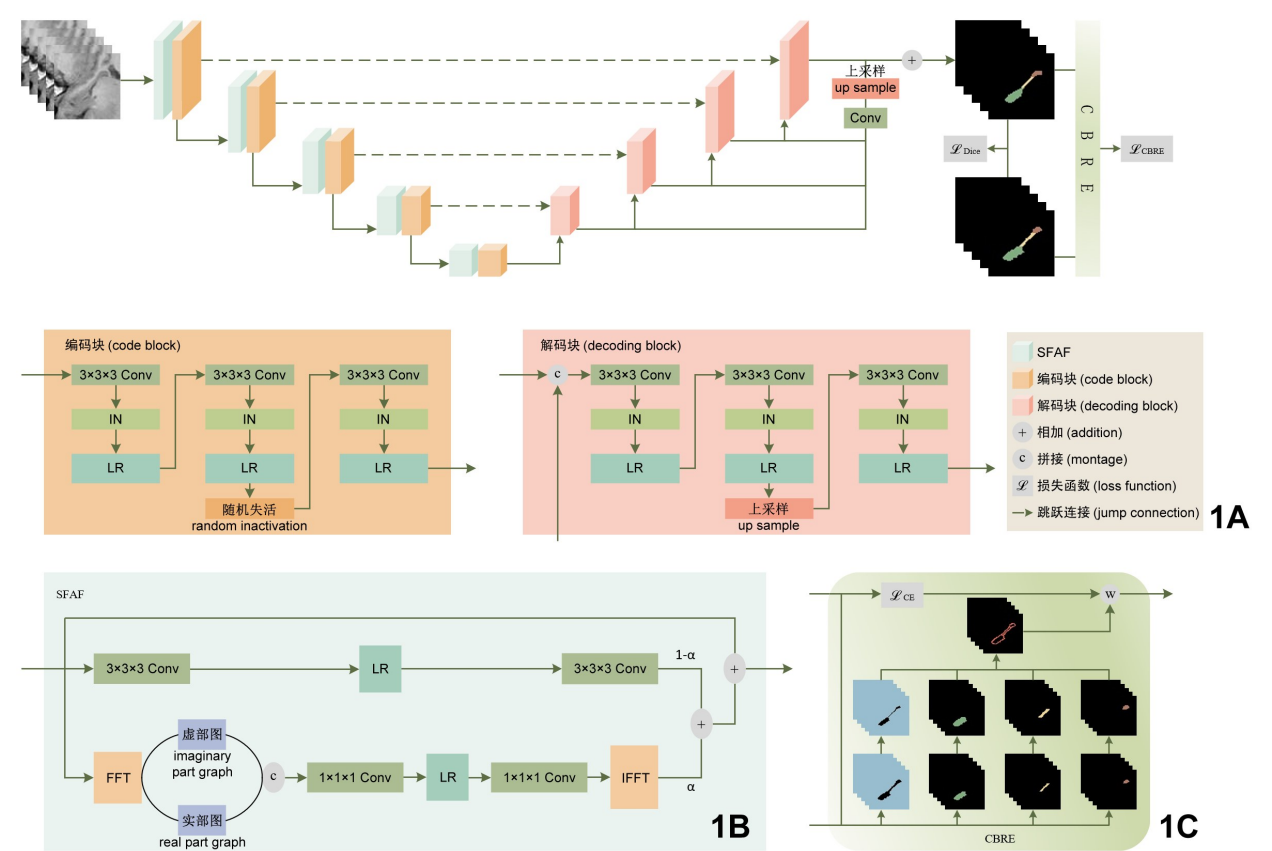


图1 所提出框架的概述,包含空间域与频域特征自适应融合模块和类间边界区域增强模块

A. 网络的主干为添加了深度监督机制的 3D U-Net; B. 空间域与频域特征自适应融合模块(SFAF),该模块由 3 个分支组成,顶部是原始特征保留分支,中间是空间域特征提取分支,底部是频域特征提取分支。通过可学习参数 α 自适应地融合来自中间分支和底部分支的特征,然后与顶部分支求和以获得最终输出特征; C. 类间边界区域增强策略(CBRE),底层表示从真实标签转换而来的每个类的两值标签,中间层表示每个类膨胀后的两值标签,顶层表示每两个膨胀后两值标签求交再合并后的边界区域图; Conv. 卷积; IN. 实体归一化; LR. 激活函数; FFT. 快速傅里叶变换; IFFT. 逆快速傅里叶变换

Fig.1 Overview of the proposed framework, which containing spatial and frequency domain features adaptive fusion module(SFAF) and inter-class boundary region enhancement strategy(CBRE)

A, The backbone of the network is 3D U-Net with added deep supervision mechanism; B, Spatial and frequency domain features adaptive fusion module, this module consists of three branches, the top is the input feature retention branch, the middle is the spatial domain feature extraction branch, and the bottom is the frequency domain feature extraction branch; The features from the middle and bottom branches are adaptively fused by the learnable parameter α , and then summed with the top branch to obtain the final output features; C, Inter-class boundary region enhancement module, the bottom layer represents the binary labels of each class transformed from ground truth, the middle layer represents the expanded binary labels of each class, and the top layer represents the intersection of every two expanded binary labels and then the union of them; CONV, Convolution; IN, Instance normal; LR, Leaky relu; FFT, Fast Fourier transform; IFFT, Inverse fast Fourier transform

网络训练。然后,将预处理后的 ROI 海马输入到由 5 个特征提取层堆叠组成的编码器中。每层都包含 1 个空间域与频域特征自适应融合模块和 1 个编码块,该编码块对特征图进行下采样以增加网络的感知野,同时提取高级语义信息,空间域与频域特征自适应融合模块通过快速傅里叶变换和卷积自动选择合适的频域特征图,并通过可学习的动态参数自适应地融合空间域与频域提取的特征。然后,提取的特征图将进入由 4 个特征恢复层组成的解码器,每层都包含 1 个解码块,解码块对特征图进行上采样以获得高分辨率表示并恢复特征图的大小。同时,解码器路径接收从跳跃连接传递而来的编码器特征图,在相应编码器的相应位置处获得更多上下文信

息。此外,我们还利用深度监督机制的思想,通过加权求和对解码器中第 2、3 和 4 层生成的多尺度特征图进行组合。在最后的损失监督部分,网络的预测分割结果和真实标签会进入到我们的类间边界区域增强策略,该策略通过增加损失函数中边界区域的权重来增强网络对边界区域的学习能力。在接下来的章节中,我们将详细描述所提出的空间域与频域特征自适应融合(spatial and frequency domain features adaptive fusion, SFAF)模块和类间边界区域增强模块(interclass boundary region enhancement strategy, CBRE)。

2.2 空间域与频域特征自适应融合模块:MRI 图像中总是含有大量的高频噪声,这些噪声掩盖了原始

图像中丰富的信息,导致我们的分割网络在提取特征时容易受到噪声的干扰,提取错误的信息并影响分割结果。因此,我们提出了一种空间域与频域特征自适应融合模块(图 1B)来解决这个问题,其中包含了自适应融合的 3 个分支:空间域特征提取分支,频域特征提取分支和原始特征保留分支,空间域与频域之间的转换是通过三维快速傅里叶变换和反快速傅里叶变换来实现的,我们将在下面详细介绍。

2.2.1 空间域特征提取:在空间域中提取特征的方法有很多,我们选择对特征图进行多次卷积运算以获得所需的效果。来自前一层的输入特征图 F_{in_i} 经过两个包含卷积和激活函数的卷积块,得到空间域特征图 F_{s_i} :

$$F_{s_i} = \hat{A}(F_{in_i} \otimes K_{s_i}) \otimes K_{s_i}, i = 1, 2, \dots, l \quad (1)$$

其中, \hat{A} 表示激活函数 Relu, \otimes 表示卷积运算, K 表示卷积核, l 表示特征提取的总层数。

2.2.2 频域特征提取:由于 MRI 图像中含有大量噪声,这些噪声掩盖了原始图像中丰富的信息,干扰了网络对特征的提取,因此我们希望通过在频域中自动选择合适的频率组合来消除 MRI 图像中噪声的影响,从而获得图像中真正有效的信息,并完成相应的特征提取。

首先,我们使用三维快速傅里叶变换 Φ 将特征图从空间域转换到频率域,得到实部图 $Freq_{real_i}$ 和虚部图 $Freq_{imag_i}$ 。实部图表示振幅图,其中包含图像的全局信息,如纹理、颜色等;虚部图表示相位图,其中包含图像的局部信息,如轮廓、形状等。

$$Freq_{real_i}, Freq_{imag_i} = \Phi(F_{in_i}) \quad (2)$$

众所周知,所有的波都可以用无限叠加的正弦波来表示,而正弦波又可以用频率、幅度和相位来表示。快速傅里叶变换的本质是将 1 个波形信号分解成无数个正弦波信号的叠加。因此,我们将特征图的实部和虚部拼接起来,通过卷积自动选择合适的波形组合 $Freq_{choi_i}$,以达到过滤无效特征的目的:

$$Freq_{choi_i} = \hat{A}[(Freq_{real_i} \oplus Freq_{imag_i}) \otimes K_{f_i}] \otimes K_{f_i} \quad (3)$$

其中 \oplus 表示拼接操作。然后将自动选取的频域图再次分为实部图和虚部图:

$$Freq_{choi_real_i}, Freq_{choi_imag_i} = \text{Chunk}(Freq_{choi_i}) \quad (4)$$

并重构为复数形式:

$$Freq_{comp_i} = \text{Complex}(Freq_{choi_real_i}, Freq_{choi_imag_i}) \quad (5)$$

然后,将复数形式的频域图通过反快速傅里叶反变换 Ψ 转换为空间域特征图 F_{f_i} ,用于后续的特征融合:

$$F_{f_i} = \Psi(Freq_{comp_i}) \quad (6)$$

2.2.3 空间域与频域特征的自适应融合:最后,我们通过一个初始值为 0.5,并在网络训练过程中不断调整的可学习参数 α ,自适应融合提取的空间域特征与频域特征,来提高网络的表征能力,同时利用残差连接打破网络的对称性。

$$F_{out_i} = \alpha \cdot F_{s_i} + (1 - \alpha) \cdot F_{f_i} + F_{in_i} \quad (7)$$

其中 F_{out_i} 为特征图经过空间域与频域自适应融合模块后的输出。

每 1 层的输入特征图由上 1 层的输出特征图转换而来:

$$F_{in_i} = \begin{cases} x, & i = 1, \\ (F_{out_{i-1}} \otimes K) \otimes K, & i = 2, \dots, l \end{cases} \quad (8)$$

其中 x 为原始输入数据。

在上述步骤之后,每层特征图的大小将从 $\mathbb{R}^{B \times C \times H \times W \times D}$ 变为 $\mathbb{R}^{B \times C \cdot 2 \times \frac{H}{2} \times \frac{W}{2} \times \frac{D}{2}}$,但第 1 层略有不同,因为它不执行下采样操作,大小将从原始输入数据 $x \in \mathbb{R}^{B \times 1 \times H \times W \times D}$ 变为 $\mathbb{R}^{B \times 8 \times H \times W \times D}$ 。

2.3 类间边界区域增强策略:对预分割组织边界区域的精确定位和预测仍然是医学图像分割中的一个难题,尤其是对于那些类别之间边界模糊的器官与组织,就如我们想要分割的海马。从整体上看,它与周围组织的边界并不清晰,从内部来看,海马头、体、尾 3 个部分在体素水平上没有清晰的边界,两侧的灰度值十分接近。这些导致了对类的边界区域划分的困难。为了解决这一问题,我们提出了一种类间边界区域增强策略(图 1C),该策略通过膨胀每个类的区域,然后将每 2 个类膨胀后的相交区域合并在一起获得类间边界区域,然后对相应位置的损失函数进行加权,使网络更加专注于边界区域信息的学习与推理。

2.3.1 类间边界区域的获取:首先,我们对输入标签数据 Y 的每个类分别进行编码,获得 C 个二进制标签,其中 C 表示包括背景在内的类的总数。该过程可以定义为:

$$Y_1, \dots, Y_c = \Omega(Y) \quad (9)$$

然后对所有 2 值标签进行膨胀操作:

$$Y_{d_i} = Y_j \odot K_d, i = 1, 2, \dots, C \quad (10)$$

其中 Y_d 为膨胀后的 2 值标签, K_d 是卷积操作的核。

对膨胀后的每 2 个不同 2 值标签求交集,得到 C 个类的 N 个边界区域图 BR :

$$\begin{cases} BR_1, BR_2, \dots, BR_n = \bigcap_{p! = q} (Y_{d_j}^p, Y_{d_j}^q), \\ [p, q \in (1, 2, \dots, C)] \\ N = [C \cdot (C - 1)] \end{cases} \quad (11)$$

其中 (\cdot) 表示求交运算,然后将 N 个边界区域

图合并,得到最终完整的类间边界区域图 CBR :

$$CBR = \cup \left(BR_1, BR_2, \cdots, BR_n \right)$$

(12)

其中 (\cdot) 表示求并运算。

2.3.2 基于损失函数加权的增强学习:我们使用提取的边界区域对体素级的多类交叉熵损失函数 \mathcal{L}_{vCE} 进行加权:

$$\mathcal{L}_{vCBRE} = \lambda \cdot CBR \cdot \mathcal{L}_{vCE} + \mathcal{L}_{vCE}$$

(13)

其中 λ 为权重因子, \mathcal{L}_{vCE} 被定义为:

$$\mathcal{L}_{vCE} = - \sum_{c=0}^C Y \cdot \log \hat{Y}$$

(14)

其中 Y 为真实标签, \hat{Y} 为多类分割预测。

然后我们对边界区域增强后的体素级交叉熵损失函数加和求平均:

$$\mathcal{L}_{CBRE} = \frac{1}{B \times H \times W \times D} \cdot \sum \mathcal{L}_{vCBRE}$$

(15)

整个网络最终的损失函数 \mathcal{L} 被定义为:

$$\mathcal{L} = \frac{1}{2} \cdot (\mathcal{L}_{Dice} + \mathcal{L}_{CBRE})$$

(16)

其中 \mathcal{L}_{Dice} 被定义为:

$$\mathcal{L}_{Dice} = 1 - \frac{2 \cdot |Y \cap \hat{Y}|}{|Y| \cup |\hat{Y}|}$$

(17)

其中 $|\cdot|$ 表示集合内元素的个数。

结 果

1. 对比实验

为了证明我们提出框架的优势,我们以 $Dice$ 相

表 1 在青少年大脑 MRI 数据集上海马分割结果的定量比较($\bar{x} \pm s$)

方法 methods	海马头 hippocampus head		海马体 hippocampus body		海马尾 hippocampus tail	
	DSC	HD95	DSC	HD95	DSC	HD95
3D U-Net	0.8555±0.0186 *	1.8481±0.3253 *	0.8265±0.0188 *	1.5034±0.2179 *	0.8349±0.0203 *	1.2142±0.1132
ResNet-VAE	0.8546±0.0170 *	1.6861±0.2667	0.8227±0.0320 *	1.5231±0.3401 *	0.8240±0.0250 *	1.9204±1.3742
Dense-Net	0.8541±0.0245 *	3.2027±3.4853	0.8273±0.0257 *	1.8521±1.1984	0.8343±0.0183 *	2.5614±2.6395
3D-UCaps	0.8546±0.0148 *	1.5840±0.1571 *	0.8274±0.0228 *	1.4052±0.3602	0.8021±0.0315 *	1.3237±0.1519 *
APAUNet	0.8658±0.0125 *	1.4780±0.1133	0.8373±0.0088 *	1.3323±0.1227 *	0.8425±0.0239	1.1811±0.2403
VT-UNet	0.8641±0.0125 *	1.4992±0.2159	0.8416±0.0157 *	1.2041±0.1185	0.8447±0.0171	1.1322±0.1251
UNETR	0.8604±0.0285	1.4536±0.6105	0.8388±0.0203 *	1.5934±0.7425	0.8384±0.0257 *	1.5119±0.7643
PANet	0.8676±0.0108 *	1.5656±0.2337	0.8444±0.0167 *	1.2478±0.1486	0.8451±0.0239 *	1.0939±0.0958
Ours	0.8868±0.0097	1.3370±0.1815	0.8639±0.0151	1.1422±0.0586	0.8582±0.0151	1.1669±0.0390

通过配对 t 检验证实我们的方法显著优于其他方法, * $P<0.05$
Our method significantly outperforms others, * $P<0.05$

表 2 消融实验分析我们提出的策略的贡献($\bar{x} \pm s$)

方法 methods	海马头 hippocampus head		海马体 hippocampus body		海马尾 hippocampus tail	
	DSC	HD95	DSC	HD95	DSC	HD95
Baseline	0.8555±0.0186	1.8481±0.3253	0.8265±0.0188	1.5034±0.2179	0.8349±0.0203	1.2142±0.1132
W/ SFAF	0.8731±0.0171	1.5370±0.2183	0.8413±0.0114	1.3641±0.2158	0.8461±0.0329	1.2761±0.2133
W/ CBER	0.8712±0.0145	1.7447±0.3591	0.8442±0.0091	1.3021±0.1850	0.8481±0.0110	1.1887±0.0369
Ours	0.8868±0.0097	1.3370±0.1815	0.8639±0.0151	1.1422±0.0586	0.8582±0.0151	1.1669±0.0390

数值是通过五折交叉验证获得的平均值(方差)
Values are the mean(stdev) obtained over five-fold cross-validation.

似系数($Dice$ similarity coeffieient, DSC)和 Hausdorff 95 距离(Hausdorff 95 distance, $HD95$)为指标,将其与一些主流的分割方法进行了比较:3D U-Net^[8]、ResNet-VAE^[10]、Dense-Net^[11]、3D-UCaps^[12]、APAUNet^[13]、VT-UNet^[14]、UNETR^[15]和 PANet^[16]。如表 1 所示,通过配对 t 检验证实,我们的方法(Ours)相比于其他具有竞争力的方法取得了明显的提升($P<0.01$)。图 2A 为我们的海马三维分割定性结果,展示了我们的方法在海马头、体、尾 3 类形态上的优势。

2. 消融实验

为了评估我们的分割框架中不同策略的有效性,我们进行了以下消融实验:(1) CNN 特征提取器(基线 Baseline);(2) Baseline+SFAF(W/SFAF);(3) Baseline+CBRE(W/CBER);(4) Baseline+CBRE+SFAF(Ours)。表 2 和图 2B 分别展示了实验的定量和定性结果。

3. 超参数选取

边界区域的膨胀运算核和损失函数权重可能是影响类间边界区域增强策略的重要超参数,我们分别使用 $K_d = 1, 2, 3, 4, 5$ 和 $\lambda = 0, 0.5, 1, 2, 3$ 进行了比较实验,探讨其最优设置。表 3、图 2C 和 3A 给出了膨胀核大小的实验结果,表 4、图 2D 和 3B 给出了类间边界区域损失函数加权系数大小的实验结果。

4. 鲁棒性分析

为了评估我们所提出方法的泛化能力,我们在

公开的 Task04_Hippocampus 数据集上进行了测试。实验结果见表 5 和图 2E。

5. 实验细节

我们使用一台配备 NVIDIA GeForce RTX 3090 GPU 的计算机,通过 Pytorch 框架对提出的网络进行端到端训练。使用 Adam 优化器更新权重,将学习率初始化为 0.01,将权重的衰减设置为 1e-4。学习率逐渐衰减公式如下:

$$lr = base_{lr} \cdot (1 - epoch_{cur}/epoch_{total}) \quad (18)$$

其中 lr 为每次迭代更新后的学习率, $base_{lr}$ 为初始学习率, $epoch_{cur}$ 为当前训练迭代次数, $epoch_{total}$ 为总训练迭代次数。该网络训练了 9000 次迭代,每次迭代训练的批量大小为 10,每 200 次迭代使用验证集进行测试,以保留更高的 *Dice* 精度。

表 3 在类间边界区域增强策略中不同 K_d 值下的分割结果($\bar{x} \pm s$)

K_d	海马头 hippocampus head		海马体 hippocampus body		海马尾 hippocampus tail	
	<i>DSC</i>		<i>DSC</i>		<i>DSC</i>	
	<i>HD95</i>	<i>HD95</i>	<i>HD95</i>	<i>HD95</i>	<i>HD95</i>	<i>HD95</i>
1	0.8731±0.0171	1.5370±0.2183	0.8413±0.0114	1.3641±0.2158	0.8461±0.0329	1.2761±0.2133
2	0.8868±0.0097	1.3370±0.1815	0.8639±0.0151	1.1422±0.0586	0.8582±0.0151	1.1669±0.0390
3	0.8767±0.0127	1.4527±0.1555	0.8620±0.0114	1.1409±0.0730	0.8522±0.0195	1.1276±0.1073
4	0.8738±0.0120	1.4240±0.2180	0.8549±0.0088	1.2181±0.0760	0.8536±0.0085	1.1176±0.0636
5	0.8730±0.0151	1.4691±0.2073	0.8506±0.0148	1.2170±0.0950	0.8516±0.0183	1.1204±0.0970

表 4 在类间边界区域增强策略中不同 λ 值下的分割结果($\bar{x} \pm s$)

λ	海马头 hippocampus head		海马体 hippocampus body		海马尾 hippocampus tail	
	<i>DSC</i>		<i>DSC</i>		<i>DSC</i>	
	<i>HD95</i>	<i>HD95</i>	<i>HD95</i>	<i>HD95</i>	<i>HD95</i>	<i>HD95</i>
0	0.8731±0.0171	1.5370±0.2183	0.8413±0.0114	1.3641±0.2158	0.8461±0.0329	1.2761±0.2133
0.5	0.8802±0.0117	1.3709±0.1441	0.8587±0.0070	1.1612±0.0618	0.8565±0.0209	1.1427±0.0946
1	0.8868±0.0097	1.3370±0.1815	0.8639±0.0151	1.1422±0.0586	0.8582±0.0151	1.1669±0.0390
2	0.8691±0.0160	1.5427±0.2538	0.8521±0.0107	1.2002±0.0753	0.8593±0.0096	1.0979±0.0387
3	0.8786±0.0143	1.4457±0.2723	0.8604±0.0101	1.2002±0.1459	0.8575±0.0153	1.2026±0.1430

表 5 在 Task04_Hippocampus 数据集上海马分割结果的定量比较($\bar{x} \pm s$)

方法 methods	海马前部 hippocampus anterior		海马后部 hippocampus posterior	
	<i>DSC</i>		<i>DSC</i>	
	<i>HD95</i>	<i>HD95</i>	<i>HD95</i>	<i>HD95</i>
3D U-Net	0.8733±0.0039 *	1.3707±0.0533 *	0.8645±0.0020 *	1.3515±0.0530
ResNet-VAE	0.8774±0.0064 *	1.6083±0.1517 *	0.8634±0.0024 *	1.5232±0.0967 *
Dense-Net	0.8738±0.0028 *	2.0864±0.3194 *	0.8598±0.0109	1.4912±0.1962
3D-UCaps	0.8809±0.0017 *	1.4198±0.1303	0.8635±0.0028 *	1.3565±0.0672
APAUNet	0.8744±0.0052 *	1.4958±0.0833 *	0.8606±0.0025 *	1.4698±0.0653 *
VT-UNet	0.8689±0.0056 *	1.4443±0.0665 *	0.8563±0.0039 *	1.4119±0.0442 *
UNETR	0.8720±0.0039 *	1.5318±0.0987 *	0.8597±0.0051 *	1.4843±0.1235 *
PANet	0.8818±0.0020 *	1.3384±0.0293 *	0.8673±0.0014 *	1.3270±0.0672
Ours	0.8917±0.0025	1.2873±0.0119	0.8734±0.0017	1.3029±0.0578

配对 *t* 检验证实我们的方法显著优于其他方法, * $P<0.05$
Our method significantly outperforms others, * $P<0.05$

讨 论

在青少年大脑 MRI 数据集上进行对比实验的定量结果发现,将本方法与 8 种不同的方法进行了比较,在海马头、体、尾分别为 0.8868±0.0097、0.8639±0.0151 和 0.8582±0.0151 的情况下,获得了较为理想的结果。与主流的 3D U-Net 医学图像分割模型相比,我们的方法对海马头部、体部和尾部的 DSC 分别提高了 3.12%、3.75% 和 2.33%。海马三维分割的定性结果展示了我们的方法在海马头、体、尾 3 类形态上的优势。在定性结果的呈现中可以看出,我们的方法不仅在各自海马头、体、尾的形态和大小上获得了与真实标签更接近的结果,而且在 3 类的边界区域和分界位置上也获得了最优的性能。这使得临床医生可以更准确地观察海马头、体、尾的形态,获得更精确的体积大小,以确定其萎缩情

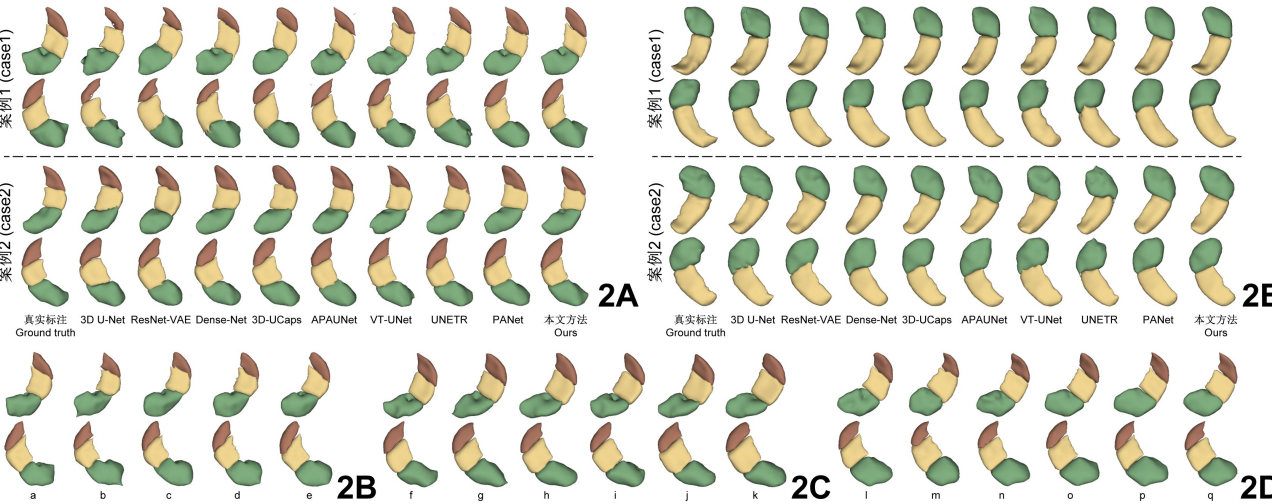


图 2 海马分割的定性结果,顶行和底行表示两种不同的查看视角(前和后)

A. 在我们的青少年大脑 MRI 数据集上海马分割结果的定性比较; B. 展示使用 SFAF 和 CBRE 影响的案例; C. 在类间边界区域增强策略中,不同 K_d 值的可视化结果; D. 在类间边界区域增强策略中,不同 λ 值的可视化结果; E. 在 Task04_Hippocampus 数据集上海马分割结果的定性比较; a. 真实标注; b. 基线; c. 基线+SFAF; d. 基线+CBRE; e. 基线+SFAF+CBRE; f. 真实标注; g. 膨胀核=1; h. 膨胀核=2; i. 膨胀核=3; j. 膨胀核=4; k. 膨胀核=5; l. 真实标注; m. 权重=0; n. 权重=0.5; o. 权重=1; p. 权重=2; q. 权重=3

Fig.2 Qualitative results of hippocampal segmentation, the top and bottom rows indicating two different viewing perspectives (front and back)

A, Qualitative comparison of hippocampal segmentation results achieved by tested methods on our Teenager Brain MR dataset; B, Examples showing the impact of using SFAF and CBRE; C, Visual results for different number of K_d in our inter-class boundary region enhancement strategy; D, Visual results for different number of λ in our inter-class boundary region enhancement strategy; E, Qualitative comparison of hippocampal segmentation results achieved by tested methods on Task04_Hippocampus dataset; a, Ground truth; b, Baseline; c, Baseline+SFAF; d, Baseline+CBRE; e, Baseline+SFAF+CBRE; f, Ground truth; g, $K_d=1$; h, $K_d=2$; I, $K_d=3$; j, $K_d=4$; k, $K_d=5$; l, Ground truth; m, $\lambda=0$; n, $\lambda=0.5$; o, $\lambda=1$; p, $\lambda=2$; q, $\lambda=3$

况,帮助他们更准确地诊断和追踪病情。

我们还进行了消融实验来研究这两个模块对分割效果的影响。基线为 3D U-Net 网络,类间边界区域增强策略(CBRE)通过提取类间边界区域并对该区域的损失函数重加权,帮助网络更清晰地识别多个类之间的边界,从而精确控制每个类的体积大小。通过图 2B 中(d)和(b)的对比,我们可以更直观地看到对类间边界定位的改善。该模块提高了 1.57%、1.77% 和 1.32% 的 *Dice* 精度,同时降低了 *HD95* 和平均表面距离 (average surface distance, *ASD*)。空间域与频域特征自适应融合模块(SFAF)通过快速傅里叶变换将特征图从空间域变换到频域,然后通过卷积在频域自动选择合适的频率组合,最后通过可学习参数 α 自适应融合空间域与频域的特征。该模块的加入有助于减弱网络对噪声的影响,从而捕获更丰富的图像信息和特征,与基线相比,对海马头、体、尾的 *Dice* 准确率分别提高了 1.76%、1.48% 和 1.12%。两个模块的结合也达到了互补的效果和更高的 *Dice* 精度。

我们还对网络中使用的超参数进行了实验,以验证不同值的选择对模型性能的影响。当膨胀核大小为 2 时,膨胀操作的性能达到最佳。这是因为海马本身非常小,现在它更是被分为 3 个部分:头、体、

尾。此时,选择太大的核进行膨胀可能会导致提取出太大的边界区域,从而无法实现精确定位类别分界位置的目标。表 4、图 2D 和图 3B 给出了类间边界区域损失函数加权系数大小的实验结果,结果表明,当 $\lambda=1$ 时,结果最优。

为了评估我们提出的方法的泛化能力,我们在 Task04_Hippocampus 数据集上进行了实验。如表 5 和图 2E 所示,我们的方法在海马前部和后部的准确率分别达到了 0.8917 ± 0.0025 和 0.8734 ± 0.0017 ,在 *Dice*、*HD95* 和 *ASD* 3 个指标上都优于其他 8 种具有竞争力的方法。与主流的 3D U-Net 方法相比,本方法的 *Dice* 精度分别提高了 1.84% 和 0.89%。这表明我们的方法具有很强的泛化性,可以应用于其他海马数据集进行精确分割。

目前有许多自动分割方法来处理各种模态的医学图像,包括 CT、MRI、X 射线、超声图像等。一些方法通过增强边界来提高分割精度。Ho 等^[17]通过移位、缩放和旋转的方式对数据进行增强后,使用边缘轮廓方法提高了脑肿瘤边缘的稳健性,Wilcoxon 符号秩检验表明,增强和非增强边缘轮廓的结果明显不同。Zhang 等^[18]通过引入 Sobel 边缘检测器,提出了一种用于无监督的边界增强策略,该检测器允许网络获得用于医学图像血管分割的额外边缘先

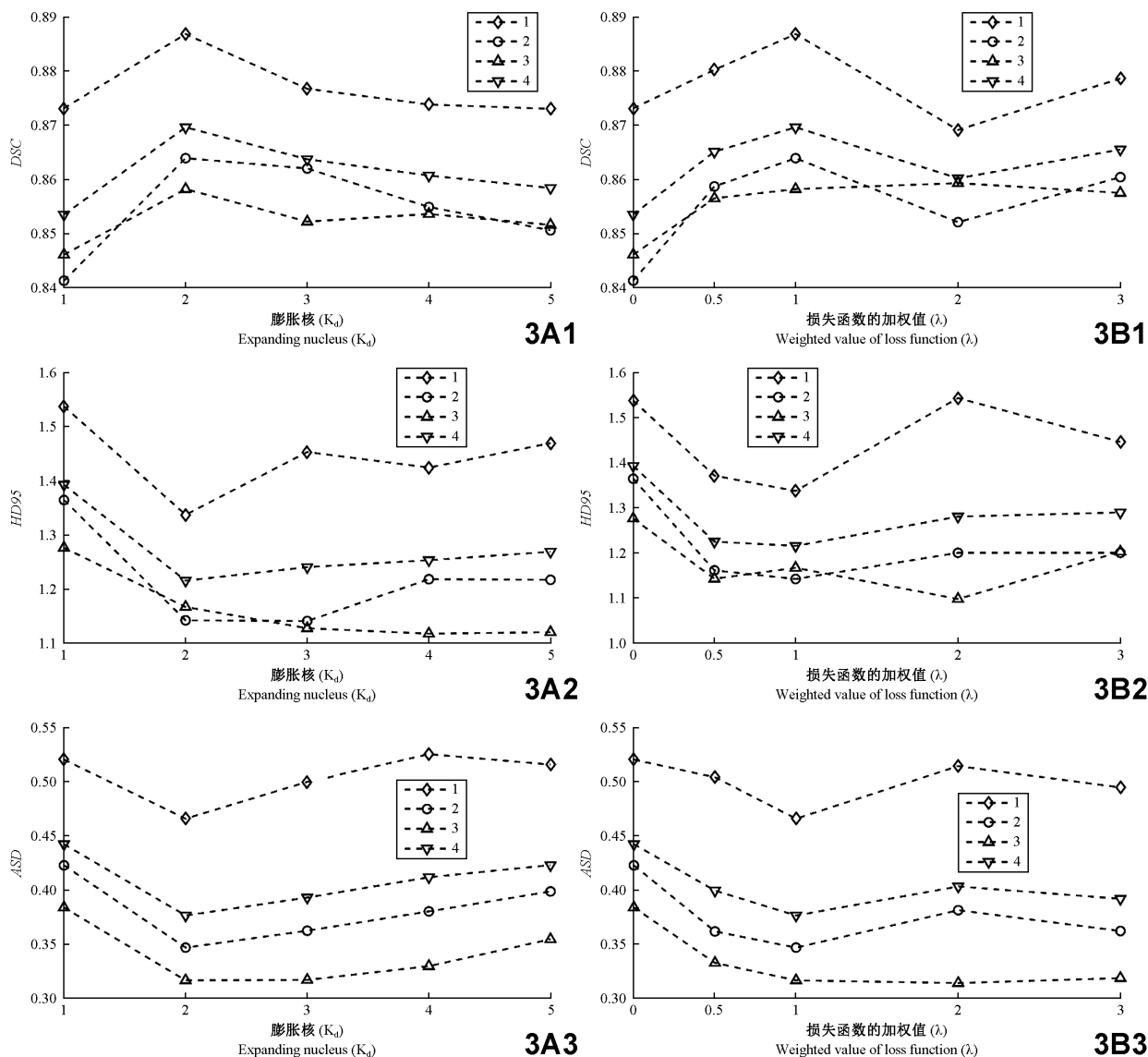


图 3 在类间边界区域增强策略中不同 K_d 和 λ 值的验证性能

A1. 不同 K_d 值下海马头体尾的 DSC 值和平均值; A2. 不同 K_d 值下海马头体尾的 $HD95$ 值和平均值; A3. 不同 K_d 值下海马头体尾的 ASD 值和平均值; B1. 不同 λ 值下海马头体尾的 DSC 值和平均值; B2. 不同 λ 值下海马头体尾的 $HD95$ 值和平均值; B3. 不同 λ 值下海马头体尾的 ASD 值和平均值; 其中, DSC 表示 *Dice* 相似性系数, 衡量真实标签与预测分割结果之间的重叠程度; $HD95$ 表示 Hausdorff 95 距离, 衡量两个目标边界之间的最大差异; ASD 表示平均表面距离, 用于计算从预测分割结果到真实标签的平均表面距离; 1. 海马头; 2. 海马体; 3. 海马尾; 4. 平均值

Fig. 3 Validation performances of the different number of K_d and λ in our inter-class boundary region enhancement strategy

A1, DSC and mean values of hippocampal head, body and tail with different K_d values; A2, $HD95$ and mean values of hippocampal head, body and tail with different K_d values; A3, ASD and mean values of hippocampal head, body and tail with different K_d values; B1, DSC and mean values of hippocampal head, body and tail with different λ values; B2, $HD95$ and mean values of hippocampal head, body and tail with different λ values; B3, ASD and mean values of hippocampal head, body and tail with different λ values; DSC denotes *Dice* similarity coefficient, which measures the degree of overlap between the real label and the predicted segmentation results; $HD95$ denotes Hausdorff 95 distance, which measures the maximum difference between the two target boundaries; and ASD denotes the average surface distance, which is used to compute the average surface distance from the predicted segmentation results to the real labels; 1, Hippocampal head; 2, Hippocampal body; 3, Hippocampal tail; 4, Average value

验。Wang 等^[19]引入边缘检测算子构建肺结节边缘增强数据集,提高了数据集的利用率,提供了更多目标病变边界的先验知识,解决了肺结节分割时病变边界和位置信息的丢失问题。还有一些方法考虑从去噪的角度去除噪声的干扰,使机器能够从清晰的

图像中获得丰富的本质特征。Nagarajan 等^[20]提出了一种新的基于块差的 MRI 图像去噪方法,通过计算正常 MRI 和噪声 MRI 强度之间的块差,将其与正常 MRI 图像的块强度进行比较,并更新去噪后的 MRI 图像的每个块的像素权重,以达到 MRI 图像去

噪声的目的。Ziyad 等^[21]通过遗传算法计算阈值,提出了一种基于自适应阈值的离散小波变换去噪技术,用于低剂量 CT 预处理。Huynh 等^[22]提出了一种映射和减少 MRI 图像噪声的方法,通过可证明的最优策略,利用复值多通道 MRI 数据的固有冗余来减少数据矩阵的奇异值,使其在降噪方面优于当时最先进的方法。

这些方法无论是传统的图像分割方法还是和人工智能相结合的自动分割方法,依旧被海马边界不清晰、MRI 图像中噪声较多等问题所困扰,无法达到更好的海马分割效果。因此,本研究提出了一种先进的大脑 MRI 图像海马三维分割网络。采用空间域与频域特征自适应融合模块,降低图像中噪声的影响,使网络能够提取更丰富、更有效的信息。同时,采用类间边界区域增强策略,对损失函数进行加权,提高网络的边界区域学习能力。在我们将海马分为头、体、尾 3 部分的青少年大脑 MRI 数据集上进行的实验表明,我们的方法实现了较先进的海马分割。消融实验证明了我们所提出的两种策略的有效性。此外,我们还在 Task04_Hippocampus 数据集上验证了网络的泛化能力。总之,我们所提出的海马分割方法的准确性和泛化能力可以帮助临床医生更精准地计算海马体积,观察海马形态^[23]并检测海马萎缩情况,完成更精确的疾病诊断与追踪。

参 考 文 献

[1] Tatulian SA. Challenges and hopes for Alzheimer’s disease[J]. Drug Discov Today, 2022, 27(4): 1027-1043.

[2] Schröder J, Pantel J. Neuroimaging of hippocampal atrophy in early recognition of Alzheimer’s disease-a critical appraisal after two decades of research[J]. Psychiatry Res Neuroimaging, 2016, 247: 71-78.

[3] Cole J, Costafreda SG, McGuffin P, et al. Hippocampal atrophy in first episode depression; a meta-analysis of magnetic resonance imaging studies[J]. J Affect Disord, 2011, 134(1-3): 483-487.

[4] Frisoni GB, Jack Jr CR, Bocchetta M, et al. The EADC-ADNI harmonized protocol for manual hippocampal segmentation on magnetic resonance; Evidence of validity[J]. Alzheimers Dement, 2015, 11(2): 111-125.

[5] Boccardi M, Bocchetta M, Apostolova LG, et al. Delphi definition of the EADC-ADNI harmonized protocol for hippocampal segmentation on magnetic resonance [J]. Alzheimers Dement, 2015, 11(2): 126-138.

[6] Lisman J, Buzsáki G, Eichenbaum H, et al. Viewpoints: how the hippocampus contributes to memory, navigation and cognition[J]. Nat Neurosci, 2017, 20(11): 1434-1447.

[7] Apostolova LG, Zarow C, Biado K, et al. Relationship between hippocampal atrophy and neuropathology markers: a 7T MRI validation study of the EADC-ADNI harmonized hippocampal segmentation protocol [J]. Alzheimers Dement, 2015, 11(2): 139-150.

[8] Çiçek Ö, Abdulkadir A, Lienkamp SS, et al. 3D U-Net: learning

dense volumetric segmentation from sparse annotation[C].Medical Image Computing and Computer-Assisted Intervention-MICCAI 2016; 19th International Conference, 2016: 424-432.

[9] Antonelli M, Reinke A, Bakas S, et al. The medical segmentation decathlon[J]. Nat Commun, 2022, 13(1): 4128.

[10] Myronenko A. 3D MRI brain tumor segmentation using autoencoder regularization[C]. Brainlesion 4th International Workshop, 2019: 311-320.

[11] Dolz J, Gopinath K, Yuan J, et al. HyperDense-Net: a hyperdensely connected CNN for multi-modal image segmentation[J]. IEEE Trans Med Imaging, 2018, 38(5): 1116-1126.

[12] Nguyen T, Hua BS, Le N. 3d-ucaps: 3d capsules unet for volumetric image segmentation[C]. Medical Image Computing and Computer Assisted Intervention-MICCAI 2021; 24th International Conference, 2021: 548-558.

[13] Jiang Y, Zhang Z, Qin S, et al. APAUNet: axis projection attention UNet for small target in 3D medical segmentation[C]. Proceedings of The Asian Conference on Computer Vision, 2022: 283-298.

[14] Peiris H, Hayat M, Chen Z, et al. A robust volumetric transformer for accurate 3D tumor segmentation[C]. International Conference on Medical Image Computing and Computer-Assisted Intervention, 2022: 162-172.

[15] Hatamizadeh A, Tang Y, Nath V, et al. Unetr: transformers for 3D medical image segmentation[C]. Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision, 2022: 574-584.

[16] Zhao X, Zhang P, Song F, et al. Prior attention network for multi-lesion segmentation in medical images [J]. IEEE Trans Med Imaging, 2022, 41(12): 3812-3823.

[17] Ho TW, Qi H, Lai F, et al. Brain tumor segmentation using U-net and edge contour enhancement[C]. Proceedings of the 2019 3rd International Conference on Digital Signal Processing, 2019: 75-79.

[18] Zhang M, Yu F, Zhao J, et al. BEFD: Boundary enhancement and feature denoising for vessel segmentation [C]. Medical Image Computing and Computer Assisted Intervention-MICCAI 2020; 23rd International Conference, 2020: 775-785.

[19] Wang S, Jiang A, Li X, et al. DPBET: a dual-path lung nodules segmentation model based on boundary enhancement and hybrid transformer[J]. Comput Biol Med, 2022, 151(Pt B): 106330.

[20] Nagarajan I, Lakshmi Priya GG. Removal of noise in MRI images using a block difference-based filtering approach[J]. Int J Imaging Syst Technol, 2020, 30(1): 203-215.

[21] Ziyad SR, Radha V, Vaiyapuri T. Noise removal in lung LDCT images by novel discrete wavelet-based denoising with adaptive thresholding technique[J]. IJEHMC, 2021, 12(5): 1-15.

[22] Huynh KM, Chang W T, Chung S H, et al. Noise mapping and removal in complex-valued multi-channel MRI via optimal shrinkage of singular values [C]. International Conference on Medical Image Computing and Computer-Assisted Intervention, 2021: 191-200.

[23] JI DF, MA ZhB. Double modality fusion between CT and MRI for human head based on surface anatomic characters [J]. Acta Anatomica Sinica, 2019, 50(5): 638-644. (in Chinese)

季达峰, 马忠宾. 基于表面解剖特征的人头部计算机断层与磁共振图像双模态融合[J]. 解剖学报, 2019, 50(5): 638-644.