

CS-SegNet: 新型的脑卒中病灶分割网络^{*}

刘煜莹¹ 李超生² 李恭新¹

(1. 江南大学轻工过程先进控制教育部重点实验室 无锡 214122; 2. 江南大学附属医院神经内科 无锡 214122)

摘要: 脑卒中是一种全球范围内导致高死亡率和残疾率的严重疾病,临幊上通过早期准确的影像诊断实现早期预防和及时治疗。然而,传统的影像诊断方法对医生的知识水平和经验存在很强的依赖性,容易漏检不明显的病变特征。为此,提出了一种新型图像分割模型CS-SegNet,旨在自动分割脑卒中CT图像中的病灶区域,从而辅助诊断。CS-SegNet基于UNet-Resnet50架构,在下采样阶段引入通道和空间注意力(CASA)模块,增强对关键区域上下文信息的提取能力;在上采样阶段,结合RDSConv模块,通过残差学习和稠密连接优化卷积操作,提高特征表达能力,改善复杂背景下的分割精度;跳跃连接部分引入多尺度通道注意力(MSCA)模块,整合低层次细节信息和高层次语义信息,提升分割结果的精度和一致性。实验结果显示,CS-SegNet在分割准确率、平均交并比和召回率上分别达到99.79%、91.52%和93.83%,比UNet-Resnet50基础网络的性能分别提升了0.14%、5.11%和4.05%,并在多个对比实验中表现最优,相比于现有主流模型,该模型在脑卒中病灶分割中具有有效性以及良好的分割精度与学习能力。

关键词: 深度学习;脑卒中病灶分割;特征融合;注意力机制;U-Net

中图分类号: TP391; TN911 文献标识码: A 国家标准学科分类代码: 520.60

CS-SegNet: Novel segmentation network for cerebral stroke lesions

Liu Yuying¹ Li Chaosheng² Li Gongxin¹

(1. Key Laboratory of Advanced Control in Light Industry Processes, Ministry of Education, Jiangnan University, Wuxi 214122, China; 2. The Affiliated Hospital of Jiangnan University, Department of Neurology, Wuxi 214122, China)

Abstract: Stroke is a serious disease that causes high mortality and disability rates worldwide. Early and accurate imaging diagnosis is used clinically to achieve early prevention and timely treatment. However, traditional imaging diagnostic methods have a strong dependence on the knowledge and experience of doctors, which can easily miss unclear lesion features. To this end, a novel image segmentation model CS-SegNet has been proposed, aimed at automatically segmenting lesion areas in stroke CT images to assist in diagnosis. CS-SegNet is based on the UNet-Resnet50 architecture and introduces Channel and Spatial Attention (CASA) modules in the down sampling stage to enhance its ability to extract contextual information from key regions; in the up sampling stage, combined with the RDSConv module, residual learning and dense connections are used to optimize convolution operations, improve feature expression ability, and enhance segmentation accuracy in complex backgrounds; the skip connection part introduces a multi-scale channel attention (MSCA) module, which integrates low-level detail information and high-level semantic information to improve the accuracy and consistency of segmentation results. The experimental results showed that CS-SegNet achieved segmentation accuracy, average intersection to union ratio, and recall rate of 99.79%, 91.52%, and 93.83%, respectively, which improved the performance of UNet Resnet50 basic network by 0.14%, 5.11%, and 4.05%, and performed the best in multiple comparative experiments. Compared with existing mainstream models, this model has effectiveness, good segmentation accuracy, and learning ability in stroke lesion segmentation.

Keywords: deep learning; segmentation of cerebral stroke lesions; feature fusion; attention mechanisms; U-Net

0 引言

脑卒中是一种急性脑血管疾病,其发病机理涉及血管、血液及血流动力学的异常,这些异常导致脑部血液供应受

阻,进而引发脑实质的缺血与水肿,最终造成脑组织的缺血性坏死,俗称“脑梗”或“脑中风”^[1],是全球导致死亡和残疾的首要因素之一^[2],据《中国脑卒中防治报告(2021)》的概要^[3]指出,我国40岁及以上人群中,脑卒中患者数量已高达

1 242 万,且发病呈年轻化趋势。《柳叶刀神经病学》杂志于 2021 年发布的全球疾病负担研究^[4]中揭示了当年全球卒中死亡人数惊人地达到了 730 万,预计从 2020 年~2050 年间,全球脑卒中死亡率将激增 46.83%,残疾率也将上升 30.74%^[5]。尤为值得注意的是,中国在脑卒中发病率和死亡率方面均位居全球前列。因此,对脑卒中实施精准诊断,对于挽救病患生命、改善预后具有至关重要的意义。

脑卒中病患的病理情况可以通过核磁共振(magnetic resonance imaging, MRI)、计算机断层扫描(computed tomography, CT)等医疗影像技术手段来反映。临床中,CT 扫描在脑组织病变检查中应用广泛,能清晰地显示出头颅不同横断面的组织结构及其解剖细节,对于脑出血及脑外伤的评估尤为敏感,并且在展示骨骼结构方面展现出显著优势。同时,CT 检查操作简便、耗时短、成本低廉、普及率高。CT 图像中出血区域呈现高密度,而正常的脑组织则显示为低密度。尽管医生手动分割病变区域被视为脑出血分割的金标准,但该过程繁琐、耗时且具有主观性,因此,如何实现对脑卒中患者信息的精确且客观的评估,是当前临床医学领域亟待解决的一项重大课题。

UNet 基于全卷积网络(fully convolutional networks, FCN)提出编码器-解码器结构的神经网络,编码器提取特征获取上下文信息,解码器精确定位,跳跃链接融合深浅层特征。它是目前医学影像分割里性能最优、应用最广的网络结构,脑卒中病灶分割的深度学习方法大多基于此展开研究与改进,比如引入注意力机制、稠密模块等,旨在提升分割准确性、运算效率与使用范围。2021 年 Zhang 等^[6]提出 ME-Net,嵌入空洞空间金字塔池化并设计了残差编码器,用于提取高维度特征。但是增加了模型复杂度,导致训练时间延长,且空洞卷积可能引发棋盘效应,影响分割精度。2022 年 Shin 等^[7]提出 e-UNets,设计全局特征注意力模块。但注意力机制计算开销大,对硬件要求高,小数据集上易过拟合。同年,Cao 等^[8]构建了 Swin-UNet,用 Swin Transformer 模块作为编码器的基本单元,在上采样过程中实现像素级分割预测和恢复特征图的分辨率。其不足在于 Transformer 计算量巨大,推理速度慢,不利于实时应用。Gu 等^[9]提出 STHarDNet,用 HarDNet 模块替代卷积模块并加入 Swin Transformer。但是新模块组合可能导致参数过多,模型收敛困难。METrans^[10]利用多尺度编码器提取多尺度特征和 Transformer 进行全局特征建模,但计算复杂度高,资源消耗大。Wu 等^[11]设计的 TSRL-Net 通过设计粗糙粒度残差学习模块修正残差连接,令解码器更关注于未预测区域;并提出一种新的目标感知损失,通过调节聚焦因子来扩展聚焦样本数据。但新模块设计可能使训练调参难度增加。Wu 等^[12]提出的双阶段病灶分割方法 W-Net,利用 CNN 和 Transformer 作为骨架网络,一阶段利用 CNN 的局部信息提取能力获得大致病变分割图,第二个阶段利用 Transformer 的强大全局上下文建模能力对

边界进行细致分割;同时设计边界变形模块和边界约束模块来应对模糊边界的苛刻条件。然而流程复杂,对硬件性能要求严苛。SAN-Net^[13]和 FAN-Net^[14]针对影像分布差异设计域适应分割算法,但在复杂分布场景下,算法适应性可能受限。2024 年 Ruan 团队^[15]结合了状态空间模型(如 Mamba)的优势以及 U 形架构设计了 VM-UNet 结构,特别适合于处理医疗图像中的长距离交互,通过引入视觉状态空间(visual state space, VSS)块来捕捉广泛的上下文信息,但非对称结构可能在特征平衡上存在缺陷。2025 年 Wu 等^[16]提出了 H-vmunet 结构,通过提出高阶 2D 选择性扫描(high-order 2D-selective-scan, H-SS2D)扩展了 SS2D 的适应性以及局部特征学习能力,但新模块设计可能在通用场景下泛化能力不足。

脑卒中病灶在不同患者间差异较大,一方面,基于 CNN 的深度学习方法因感受野小、缺乏长距离信息,难以应对。而注意力机制虽能强化有效特征,弥补卷积操作不足,提升全局特征提取能力,但对样本数据量要求高,不适用于小样本。另一方面,多尺度特征融合利用空洞空间金字塔池化嵌入 UNet,扩大感受野,可提升预测能力,不过空洞卷积对小病灶分割效果欠佳,易造成信息冗余。最后,UNet 无法建模长距离上下文信息,结合 Transformer 可弥补此缺点,优化网络模块如残差连接,能解决深层网络梯度问题,但需考虑设备和计算量^[17-19]。

基准 UNet 及其变体存在固有缺陷。为了解决上述问题,提高模型的分割效果、泛化性能以及边缘分割的准确度,本文基于 UNet 网络提出了一种创新的脑卒中病灶分割网络,主要包括以下工作:

1)为了增强 UNet 在处理脑卒中影像时对特征通道和空间信息的敏感度,在编码器引入通道与空间注意力(channel attention and spatial attention, CASA)模块。该模块通过结合通道和空间注意力机制,分别捕捉高效的通道间依赖关系和更广泛的空间注意力感受野,使模型能够全面地提取脑卒中病灶区域特征,提高泛化性能;

2)为了增强对病灶区域特征的捕捉能力,减少信息损失,提出残差-深度可分离卷积融合(residual and depth separable convolutions, RDSCConv)模块。将该模块引入到上采样阶段,不仅增强了网络的特征提取能力,还减少了病灶的欠分割和过分割现象,提升了泛化性能;

3)为了强化模型对脑卒中病灶特征和病灶边界的感知能力,在跳跃连接部分设计了多尺度通道注意力(multi-scale channel attention, MSCA)模块代替原有的 concat 操作。MSCA 具有多种感受野,可以进行多层次的特征提取和特征交互,使模型高效分割各种复杂的脑卒中病灶区域,准确预测目标边界;

4)为了模型更好地在脑卒中影像分割任务中发挥作用,该模型将 Focal 损失和 Dice 损失结合。Focal 损失能有效应对样本不均衡问题,让模型更关注小病灶;Dice 损失则

能精准衡量分割精度,优化边缘细节。二者结合,全方位提升模型性能,助力精准分割病灶。

1 CS-SegNet 网络结构设计

1.1 CS-SegNet 网络

本研究提出了一种新型的脑卒中病灶分割网络 CS-SegNet,该网络属于端到端的深度学习模型。CS-SegNet 的设计旨在处理脑卒中 CT 图像,并输出精确的病灶区域预测结果,具体网络结构如图 1 所示。

在网络架构上,CS-SegNet 融合了多种创新模块,包括 CASA 模块、RDSConv 模块以及 MSCA 模块,以增强其分割性能。从整体结构来看,CS-SegNet 借鉴了 UNet 的经典 U 型结构。其左半部分为编码器,负责提取图像特征;右半部分为解码器,用于逐步恢复图像的空间分辨率;而中间部分

为特征融合模块,旨在有效整合编码器和解码器的信息。

具体地,CS-SegNet 以 Resnet50 作为其核心 backbone,包含多层卷积层和残差连接结构,能够有效提取病灶区域特征并保持信息流动。在该网络的每个下采样阶段之后,加入 CASA 模块。该模块通过对特征图进行加权处理,能够显著提升模型对关键特征的关注度,从而增强模型对病灶区域的识别能力。在跳跃连接部分,引入 MSCA 模块以替代原有的特征图拼接操作,该模块利用注意力机制融合来自不同层次的特征图,实现底层细节特征与高层语义特征的完美结合,进而大幅提升模型的表现力。在上采样阶段,采用 RDSConv 模块将标准卷积分解为深度卷积和逐点卷积两部分,并融入了残差结构,这一设计不仅显著降低了模型的参数数量和计算复杂度,还能保持网络模型性能的同时,实现更高效的特征提取和图像重建。

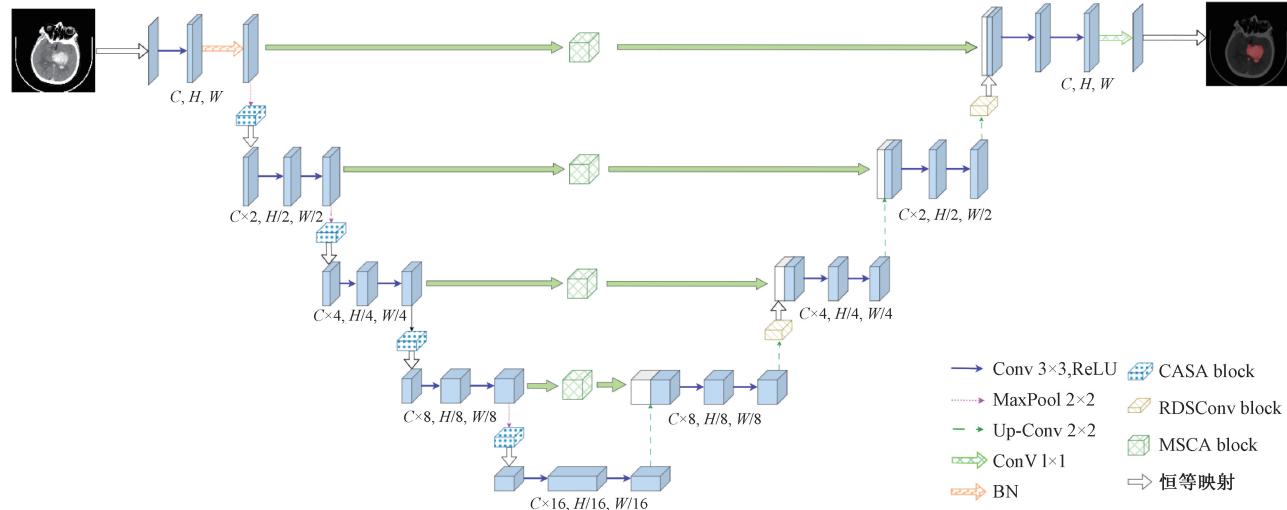


图 1 CS-SegNet 网络结构

Fig. 1 CS-SegNet network structure

1.2 CASA 模块

由于脑卒中病灶图像的复杂性,使得编码器提取的特征中含有影响分割精度的噪声因素,为了提高模型对于病灶特征的敏感度,本文在模型的编码器阶段提出了 CASA 模块,对该阶段提取的特征进行强化。其结构如图 2 所示,CASA 模块主要结合了通道注意力(channel attention, CA)和空间注意力(spatial attention, SA)增强模块,旨在通过这两个维度依次增强特征图的注意力效果。SA 用于强化局部特征编码器提取的特征;CA 用于强化全局特征编码器提取的特征。这一改进使得模型在编码阶段能够聚焦于病灶区域信息,具有更全面、更有效的特征提取能力,有助于执行后续分割任务。

在 CA 模块中,输入特征图先经全局最大池化与全局平均池化,生成两个 $1 \times 1 \times C$ 的特征图,分别捕捉全局最大值与平均值。接着,这两个特征图被送入两层的 MLP。第 1 层含 C/r 个神经元(r 为缩减比例),通过 ReLU 激活

函数增强非线性表达,降低计算复杂度并保留关键信息;第 2 层恢复至 C 个神经元,匹配输入通道数。随后,将 MLP 输出的两个特征图相加,融合池化结果,捕捉通道间复杂依赖。相加后的特征图经 sigmoid 映射,生成 CA 特征图 $H1$ 。 $H1$ 与原始输入特征图 H 逐元素相乘,分配通道权重,突出重要通道。最后,通过残差连接加法操作,引入初始输入特征,强化关键通道,防止因病灶的多变性而忽视重要通道。

在 SA 模块中,首先沿着通道方向对输入特征图进行全局最大池化和全局平均池化,从而得到两个 $H \times W \times 1$ 的特征图。接着将这两个特征图基于通道数进行拼接操作,以此补充 CA 机制所需的信息。然后通过一个 7×7 卷积操作对拼接后的特征图进行降维处理,将其通道数减少到 1 个,生成尺寸为 $H \times W \times 1$ 的特征图,旨在整合和提炼特征信息。随后使用 sigmoid 函数对降维后的特征图进行激活,生成 SA 特征图 $H2$ 。最后将 SA 特征图与模块的输

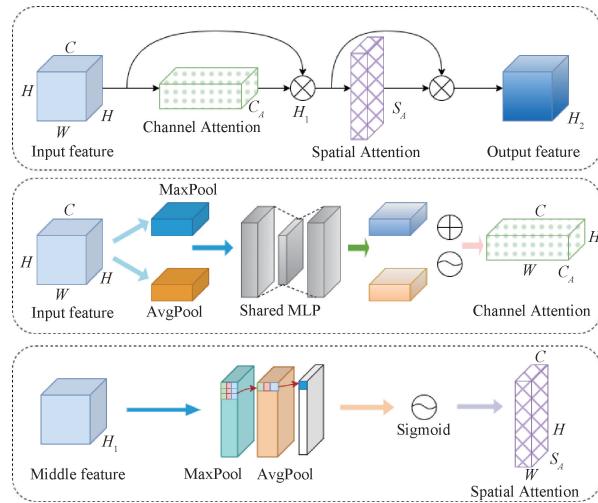


图 2 CASA 模块

Fig. 2 CASA module

入特征图进行逐元素乘法操作,生成最终的输出特征图,以实现注意力权重的分配。主要作用在于进一步增加主要目标区域重要性;由于脑卒中病灶具有多变性,使得部分复杂病灶区域在注意力学习过程中容易被忽略,原始特征的引入可以弥补该部分特征,供后续模块识别,减少信息损失。输入输出特性之间的关系可概括为:

$$H_1 = C_A(H) \otimes H \quad (1)$$

$$H_2 = S_A(H_1) \otimes H_1 \quad (2)$$

1.3 RDSConv 模块

RDSConv 模块由两个关键部分组成:逐深度卷积和逐点卷积。其中逐深度卷积专注于特征图的空间特征提取,它独立地对每个输入通道进行卷积操作,有效地减少了参数数量和计算复杂度。而逐点卷积负责整合来自不同通道的信息,即提取特征图的通道特征。通过逐点卷积将所有通道的信息合并起来,生成最终的输出特征图。这两部分在参数数量和计算需求上的显著减少,特别适合于卒中病灶的图像分割任务中,这种方法不仅降低了计算成本,还有提高了病灶分割的准确性和效率。同时,鉴于残差结构能够有效缓解深度网络性能退化的问题,并促进网络提取更为精细的语义特征。这一融合策略显著增强了网络的特征提取和泛化性能,对于缓解病灶分割中的欠分割和过分割问题具有积极作用。图 3 展示了 RDSConv 模块的架构,该模块的输入和输出之间的关系式表示如下:

$$x_{i+1} = F(x_i) + \lambda x_i \quad (3)$$

其中, $F(x)$ 表示 RDSConv 模块的映射关系, x_i 和 x_{i+1} 表示该单元的输入和输出。

1.4 MSCA 模块

不同于 AFF 特征融合模块,MSCA 模块不仅可以探索脑卒中病灶的边界信息,而且能够弥补多尺度特征信息的缺失,以它替代传统的 concat 操作,从而显著缩小编码

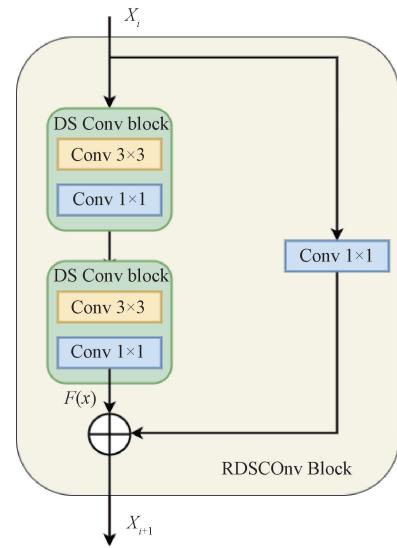


图 3 RDSConv 模块

Fig. 3 RDSConv module

器和解码器之间的语义差距,具体如图 4 所示。MSCA 模块的设计核心在于通过两个不同尺度的分支来提取通道注意力。其中一个分支利用全局平均池化技术,专注于全局特征的注意力提取;而另一个分支则直接采用 point-wise 卷积,以捕捉局部特征的通道注意力。这两个分支的结合,使得 MSCA 能够同时提取到多种不同尺度的脑卒中病灶特征信息以及多尺度特征在空间上的融合,提高对目标区域的探索效率;同时,感受野的多样性,能够实现对复杂区域进行多范围的探索,可以较容易地发现病灶区域和周围脑组织之间的差异,有利于在复杂的环境中准确预测病灶边界,引导模型准确识别目标。这种融合方式能够有效促进深层特征与浅层特征的融合,逐步丰富特征类别,增加模型所能感知的病灶样式。

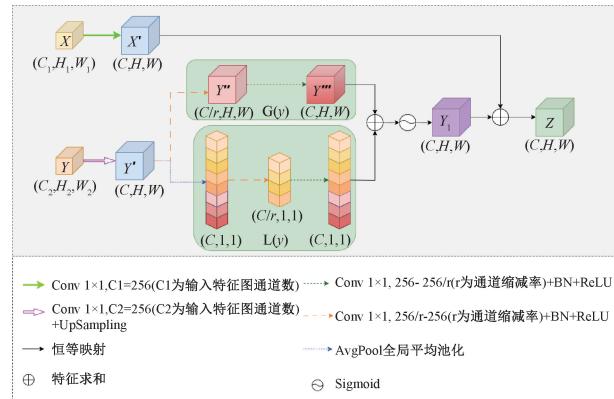


图 4 MSCA 模块

Fig. 4 MSCA module

在特征融合模块,给定两个特征 X, Y 进行特征融合(其中 Y 代表具有更大感受野的特征)。首先,对输入的特征 Y 应用多尺度通道注意力模块,并通过 sigmoid 激活函

数将输出值限制在 0~1 之间,再将处理后的的特征与输入的特征 X 做 concat 操作,以完成特征融合。MSCA 模块的输入和输出之间的关系用如下形式表示:

$$G(y) = \text{AvgPool}(B(\text{PWConv}_2(\delta(B(\text{PWConv}_1(Y')))))) \quad (4)$$

$$L(y) = B(\text{PWConv}_2(\delta(B(\text{PWConv}_1(Y'))))) \quad (5)$$

$$Y_1 = \sigma(L(y) \oplus g(y)) \quad (6)$$

$$Z = X' \oplus Y_1 \quad (7)$$

其中, $G(y)$ 表示全局特征的通道注意力, $L(y)$ 表示局部特征的通道注意力, PWConv_1 1×1 点卷积将输入的特征 X 通道数减少为原先的 $1/r$, B 表示 BatchNorm 层, δ 表示 ReLU 激活函数, 通过 PWConv_1 1×1 的卷积将通道数目恢复成与原输入通道数目相同, r 为通道缩放比。

1.5 损失函数

Focal 损失函数是一种交叉熵损失函数, 其核心在于通过加权操作来有效应对正负样本不平衡的问题。具体而言, 它将传统的 CE 损失函数的权重项进行拆分, 并引入一个可调参数, 用来灵活调节正负样本的权重。这一策略使得 Focal 损失函数可以减少容易分类样本的权重, 同时增加难以分类样本的权重, 以达到平衡训练的效果, 提高了模型对于分类错误样本的预测能力。

Dice 损失函数则是一种相似性度量函数, 常用于语义分割领域。它通过计算模型预测和真实标签之间的相似度来评估模型性能的优劣, 强调区分正负样本的差异性。Dice 损失函数的具体计算方法是, 将预测标签与真实标签的相交部分与其两者之和的比值作为模型预测的准确度指标。值越大, 表示预测结果与真实标签越相似。引入 Dice 损失函数不仅能有效应对类别不平衡的问题, 还能提高模型对类别界限模糊样本的预测精度。

本文将 Focal 损失函数和 Dice 损失函数相结合, 旨在全面解决脑卒中病灶分割任务中遇到的类不平衡和分类精度低的问题。在实际应用中, 将这两种损失加权求和作为脑卒中病灶分割模型的总损失函数。这样做既考虑了样本数量的分布差异, 又能更好的优化网络性能, 从而显著提高脑卒中病灶分割任务的精度。具体的公式表达如下:

$$\text{Focal loss}(P_t) = -\alpha(1 - p_t)^\gamma \log(p_t) \quad (8)$$

$$\text{Dice loss} = 1 - \frac{2 | X \cap Y | + \text{Smooth}}{| X | + | Y | + \text{Smooth}} \quad (9)$$

$$\text{Total loss} = \text{Focal loss} + \text{Dice loss} \quad (10)$$

其中, p_t 表示模型对于目标区域的预测概率; α 是平衡因子, 用于调整正负样本之间的权重; γ 是焦点因子, 用于调整难易样本之间的权重。 X 和 Y 分别表示预测结果和 Ground Truth, 分子代表两组交集的长度, 分母代表两组长度的总和, 为防止分母为 0 的情况出现, 引入 smooth, 值设为 1。

2 数据集与实验环境

图 5 展示了脑卒中 CT 影像分割流程的流程图, 包含数据处理、数据增强和 CS-SegNet 模型分割 3 个主要步骤。流程图中详细说明了从原始影像到分割结果的整个过程, 包括数据预处理、归一化以及模型的应用。最终, 该流程能够清晰地分割出脑卒中病灶区域。

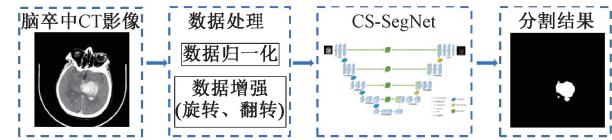


图 5 本文算法流程

Fig. 5 The algorithm process in this article

2.1 数据来源

本文所采用的影像数据源自江南大学附属医院神经内科回顾性分析的脑部 CT 影像, 该研究精心挑选了 50 名出血性脑卒中患者的 336 份脑部 CT 影像, 在确保排除假性脑卒中影像后, 最终整合成图像数据集。其中每例患者拍摄有 27~31 张脑部位 CT 图像, 图像分辨率为 512×512, 并将此数据集以 8:1:1 的比例划分为训练集、测试集和验证集, 以支持模型的有效训练与验证。

鉴于深度学习技术主要依赖于有监督学习方式, 本文的模型训练自然离不开精确的标签数据。在江南大学神经内科专家的专业指导下, 利用 labelme 软件对脑卒中影像中的高密度特征进行了细致标注, 生成了一套标签数据集。这套标签数据集不仅作为本文所提出算法的性能评估基准, 同时也为其他分割算法提供了不可或缺的参考标准。

2.2 数据预处理与增强

所有数据均为回顾性数据, 并在投入使用前执行了数据清洗, 旨在移除与患者个人信息相关的所有内容, 从而确保信息的安全性, 防止任何泄露风险。由于深度学习需要大量数据作为支持, 且医学影像数据具有稀缺性, 数据增强是模型训练前不可或缺的关键环节。数据增强旨在通过一系列操作扩大原始实验数据集, 这些操作基于几何变换原理, 能够生成语义不变且不失真的新图像。本研究主要采用了图像翻转和旋转两种数据增强技术。具体而言, 包括水平翻转、垂直翻转以及 90°、180° 和 270° 旋转在内的 5 种不同操作。数据增强并不会改变数据的形式, 而是调整模型训练时所使用的的输入数据形式, 不仅能够显著增加训练数据的数量, 还有效降低了模型过拟合的风险。这样将数据增强后的 269×5 份数据用于模型训练, 如图 6 所示。同时, 为了确保数据的一致性和准确性, 与原始图像对应的标签文件也同步进行了数据增强操作。

2.3 实验环境

本文实验环境搭建在 Windows 操作系统上, 配备了高性能的 NVIDIA GeForce RTX 4090 D(24 GB) 显卡, 搭配

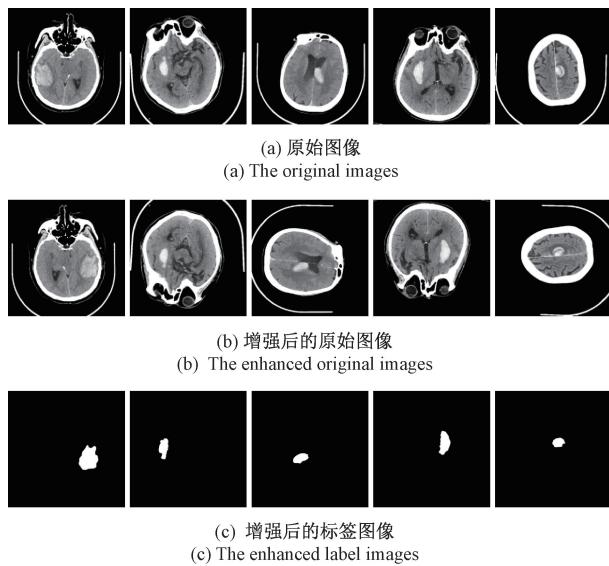


图 6 数据增强图例。从左往右依次为水平翻转、垂直翻转、 90° 旋转、 180° 旋转、 270° 旋转

Fig. 6 Data augmentation legend. From left to right, they represent horizontal flip, vertical flip, 90° rotation, 180° rotation, and 270° rotation

Intel[®] Xeon[®] Platinum 8481C 处理器(16 vCPU), 并配置 80 GB 内存和 30 GB 的系统盘, 旨在优化图像分割任务的训练和推理效率。实验的核心软件环境基于 Python 编程语言, 并采用了 Pytorch 框架, 后者为构建深度卷积神经网络模型提供了丰富的库支持。为了进一步提升训练和测试速度, 本研究还配置了 CUDA 11.0, 以充分利用 NVIDIA 显卡的加速能力。为确保实验结果的公平性和可比性, 所有参与实验的模型均采用了统一的损失函数, 即 Dice 损失函数与 Focal 损失函数的组合。表 1 详细列出了在训练和测试阶段所使用的各项超参数, 通过严格遵循这些实验条件和参数配置, 旨在获得准确、可重复的实验结果, 为图像分割任务的研究提供有力支持。

表 1 超参数设置

Table 1 Hyperparameter settings

参数名	具体数值
Batch-size	8
Training epoch	100
Training image size	512×512
Test image size	512×512
Learning rate	0.000 1
Optimizer	Adam

3 实验结果与讨论

3.1 评价指标

为了对算法进行精确评估, 本文采用平均交并比

(mIoU)、类别平均像素准确率 (mPA)、平均精准率 (mPrecision)、平均召回率 (mRecall) 和准确性 (Acc) 对模型输出进行评价。这些指标采用以下形式:

$$mIoU = \frac{TP}{FN + TP + FP} \quad (11)$$

$$mPA = \frac{TP + TN}{FN + TP + FP + TN} \quad (12)$$

$$mPrecision = \frac{TP}{TP + FP} \quad (13)$$

$$mRecall = \frac{TP}{TP + FN} \quad (14)$$

$$Acc = \frac{TP + TN}{TP + FP + TN + FN} \quad (15)$$

式中: 做出以下定义: 正确分割的病灶区域为真阳性 (true positive, TP), 错分为病灶的区域为假阳性 (false positive, FP), 正确分割的非病灶区域为真阴性 (true negative, TN), 错分为非病灶的区域为假阴性 (false negative, FN)。

3.2 损失函数对比试验

本研究通过实验对比了 Focal 损失函数、Dice 损失函数以及的 Focal 和 Dice 损失函数结合这 3 种情况的性能表现。具体训练-验证损失函数曲线如图 7 所示, 以 UNet-Resnet50 网络为训练模型为例, 其中 X 轴为训练的批次, Y 轴为精度值, 曲线类型包含训练损失曲线 (train loss)、验证损失曲线 (val loss)、训练损失平滑曲线 (smooth train loss) 和验证损失平滑曲线 (smooth val loss), 最终均能较好的拟合。

图 7(a) 代表 Focal 损失函数曲线, 当训练步长为 80 左右, 曲线趋于稳定, 其最低损失值稳定在 0.1 左右。图 7(b) 代表 Dice 损失函数曲线, 当训练步长为 85 左右, 曲线趋于稳定, 其最低损失值稳定在 0.1 左右。图 7(c) 代表 Focal+Dice 损失函数曲线, 当训练步长为 70 左右, 曲线趋于稳定, 其最低损失值稳定在 0.05 左右。综合对比可得, Focal 损失和 Dice 损失的结合在处理不平衡数据集和复杂分割任务时具有显著的优势。实验结果也充分证明了这种结合的有效性, 为脑卒中病灶分割领域的相关研究提供了新的思路和方法。

3.3 消融实验

1) 实验 1

在本次消融实验中, 深入对比了基于不同骨干网络的 UNet 模型在脑卒中 CT 图像分割任务中的性能。实验结果表明, 采用 Resnet50 作为骨干网络的 UNet (UNet-Resnet50) 模型显著优于采用 Vgg16 的 UNet (UNet-Vgg16) 模型。

具体实验结果如表 2 所示, UNet-Resnet50 的 mIoU 和 mPA 值均高于 UNet-Vgg16, 这表明 Resnet50 骨干网络在特征提取和图像表征方面具备更强的能力, 能够更精准地捕捉到脑卒中 CT 图像中的关键信息。此外, UNet-Resnet50 在 mPrecision 和 mRecall 值上的表现也进一步

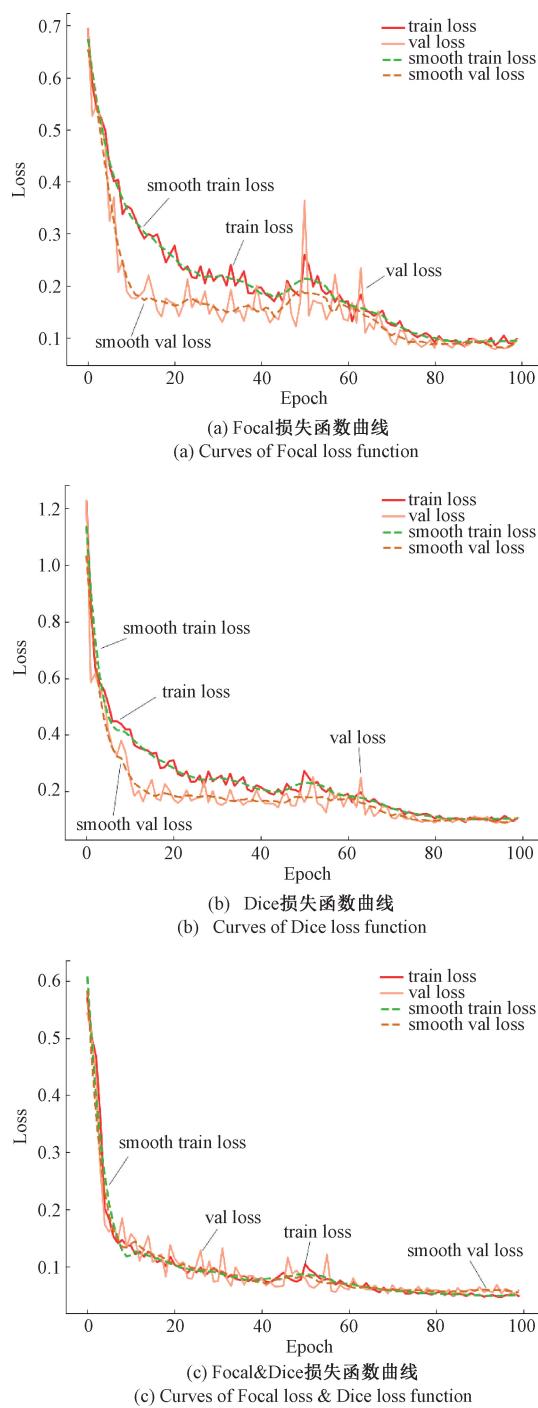


图 7 不同损失函数曲线

Fig. 7 Curves of different loss functions

验证了 Resnet50 骨干网络在提升模型整体性能方面的重要作用。这些性能参数的综合表现充分展现了 UNet-Resnet50 模型在脑卒中 CT 图像分割任务中的卓越性。

2) 实验 2

在本次消融实验中,深入探讨了 UNet-Resnet50 模型

表 2 不同 backbone 的 UNet 网络性能比较
Table 2 Comparison of UNet network performance of different backbones

Backbone	mIoU	mPA	mPrecision	mRecall	Acc	%
UNet-Vgg16	83.42	87.85	92.66	87.85	99.56	
UNet-Resnet50	86.41	89.78	94.86	89.78	99.65	

下采样阶段嵌入 CASA 模块对脑卒中 CT 图像分割性能的影响。实验设计以 Block i 代表 CASA 模块嵌入在第 i 层下采样阶段 ($i=1, 2, 3, 4$), Block1234 则表示第 1、2、3、4 层下采样均嵌入 CASA 模块。实验结果如表 3 所示, CASA 模块的嵌入对模型性能具有显著提升的作用。

未嵌入 CASA 模块的 UNet-Resnet50 的 mIoU 为 86.41%。随后,在下采样阶段的不同位置添加 CASA 模块后,发现模型性能普遍得到改善。这一改善不仅体现在单一嵌入位置上,也体现在不同累计嵌入数量上,充分证明了在下采样阶段引入 CASA 模块的有效性。最终,当在所有下采样阶段均嵌入 CASA 模块时(即 Block1234),模型的 mIoU 达到了 88.07%,高于其他所有嵌入情况,进一步证实了 CASA 模块对模型性能的显著提升作用。

因此,CASA 模块在网络下采样阶段的嵌入不仅提高了模型的 mIoU,还增强了整体分割性能。这一改进得益于更高效的特征提取和图像分割效果,为脑卒中 CT 图像的分割提供了更为准确和可靠的方法。

3) 实验 3

本次消融实验中,系统地评估了在实验 2 模型的上采样阶段嵌入 RDSCConv 模块对脑卒中 CT 图像分割性能的具体影响。实验设计以 Block j 代表 RDSCConv 模块嵌入在第 j 层上采样阶段 ($j=1, 2, 3, 4$), Block1234 则表示第 1、2、3、4 层上采样均嵌入 RDSCConv 模块。实验结果如表 4 所示,清晰地展示了 RDSCConv 模块的嵌入对模型分割性能的显著影响。

未嵌入 RDSCConv 模块的基础模型在 mIoU 上达到了 88.07%。当在上采样阶段添加一个 RDSCConv 模块时,例如在 Block1 位置,mIoU 略微下降至 86.85%,显示出单个模块在该位置的贡献有限。然而,在 Block4 位置嵌入 RDSCConv 模块时,mIoU 提高至 87.08%,显示出该位置能够有效增强特征表达。进一步评估多个 RDSCConv 模块的配置时,发现当使用两个 RDSCConv 模块时(如 Block12 和 Block24),mIoU 分别达到了 88.67% 和 88.81%,这表明在适当位置增加 RDSCConv 模块的数量能够显著提升模型性能。特别是在 Block24 位置,模型性能的提升最为显著,进一步证实了 RDSCConv 模块在深层上采样阶段的有效性。最终,在 Block1234 位置嵌入 4 个 RDSCConv 模块时,虽然 mIoU 未达到最高值(为 88.32%),但仍然能够显示出在上采样阶段嵌入 RDSCConv 模块的有效性。

表 3 CASA 模块在不同位置以及不同数量的性能对比

Table 3 Comparison of the performance of CASA modules at different locations and in different quantities

嵌入位置	嵌入数量	mIoU/%	mPA/%	mPrecision/%	mRecall/%	Acc/%
UNet-Resnet50	0	86.41	89.78	94.86	89.78	99.65
Block1	1	86.62	89.60	95.42	89.60	99.66
Block2	1	85.57	88.53	95.23	88.53	99.63
Block3	1	87.81	91.37	94.95	91.37	99.69
Block4	1	86.49	89.46	95.40	89.46	99.66
Block12	2	87.02	90.49	94.88	90.49	99.67
Block13	2	86.64	89.35	95.12	89.35	99.66
Block14	2	86.07	89.12	95.21	89.12	99.65
Block23	2	83.83	86.23	95.68	86.23	99.59
Block24	2	85.71	88.90	94.95	88.90	99.63
Block34	2	87.38	90.37	95.57	90.37	99.68
Block123	3	85.56	89.08	94.46	89.08	99.63
Block124	3	86.45	90.01	94.64	90.01	99.65
Block134	3	87.72	90.60	95.75	90.60	99.69
Block234	3	87.30	90.24	95.60	90.24	99.68
Block1234	4	88.07	91.53	95.78	91.53	99.69

表 4 RDSCConv 模块在不同位置以及不同数量的性能对比

Table 4 Comparison of the performance of RDSCConv modules at different locations and in different quantities

嵌入位置	嵌入数量	mIoU/%	mPA/%	mPrecision/%	mRecall/%	Acc/%
CASA1234+UNet-Resnet50	0	88.07	91.53	95.78	91.53	99.69
Block1	1	86.85	89.27	96.22	89.27	99.67
Block2	1	86.85	89.70	95.64	89.70	99.67
Block3	1	86.58	89.24	95.84	89.24	99.66
Block4	1	87.08	90.36	95.13	90.36	99.67
Block12	2	88.67	92.18	95.19	92.18	99.71
Block13	2	87.59	90.92	95.17	90.92	99.68
Block14	2	87.51	91.00	94.97	91.00	99.68
Block23	2	88.06	91.43	95.23	91.43	99.69
Block24	2	88.81	92.46	95.05	92.46	99.71
Block34	2	87.41	90.89	94.95	90.89	99.68
Block123	3	88.31	91.79	95.15	91.79	99.70
Block124	3	87.97	91.03	95.59	91.03	99.69
Block134	3	87.95	91.06	95.52	91.06	99.69
Block234	3	87.84	91.45	94.89	91.45	99.69
Block1234	4	88.32	91.72	95.24	91.72	99.70

综合看, RDSCConv 模块在网络上采样的第 2、4 阶段的嵌入性能提升效果最好。这表明在图像重建过程中, 使用 RDSCConv 模块有助于更好地恢复细节特征, 从而实现更高效的图像分割。

3.4 定量比较

在脑卒中 CT 图像分割的实验研究中, 为验证本文模型的性能, 针对一个特定的测试集, 将其与目前最优的一

些脑卒中分割 UNet 变体模型进行对比: PSPNet^[20]、DeepLab V3+、Attention-UNet、UNet++、HRNet、UNet++^[21]、ResUNet++^[22]、CA-UNet^[23]、Swin-UNet 和 BRNet^[24], 并用 UNet-Resnet50 和 UNet-Vgg16 作为基线进行对比。这些模型其中, UNet++ 和 UNet++ 是基于 UNet 改进的增强版本, Attention-UNet、ResUNet++ 以及 CA-UNet 模型是结合注意力机制改进的 UNet 模

型, PSPNet、Deeplab V3+ 以及 HRNet 都是较为常用的语义分割模型, SwinUNet 是基于 Transformer 与 UNet 结合的模型, BRNet 构建了边界生成和边界细化模块, 并用 UNet-Resnet50 作为基线进行对比。上述模型中, 除基准

模型外, 都是目前适合于脑卒中病灶分割任务的最优模型, 采用这些能与本文 CS-SegNet 有密切联系的模型做对比, 能够体现其先进性。所有模型设置相同的实验细节, 实验结果如表 5 所示。

表 5 不同网络性能比较

Table 5 Comparison of the performance of different networks

网络	mIoU	mPA	mPrecision	mRecall	Acc	%
UNet-Resnet50	86.41	89.78	94.86	89.78	99.65	
PSPNet	86.04	89.28	94.95	89.28	99.64	
Attention-UNet	80.70	87.75	88.46	87.75	99.55	
DeepLab V3+	84.76	87.31	95.63	87.31	99.62	
HRNet	88.95	91.91	95.89	91.91	99.72	
UNet++	86.37	89.94	95.98	89.94	99.65	
UNet+++	87.13	90.37	94.64	90.37	99.67	
ResUNet++	87.98	91.72	94.77	91.72	99.69	
CA-UNet	88.81	92.46	95.25	92.46	99.71	
Swin-UNet	91.16	93.78	96.43	93.78	99.78	
BRNet	90.91	93.04	95.99	93.04	99.75	
CS-SegNet	91.52	93.83	97.05	93.83	99.79	

通过细致的比较分析, 表 5 清晰地展示了本文所提出的 CS-SegNet 模型以及其他代表模型在脑卒中数据集上的病灶分割性能。从表中可以得知, 本文所提模型的 mIoU、mPA、mPrecision、mRecall 以及 Acc 分别为 91.52%、93.83%、97.05%、93.83% 以及 99.79%, 其中 4 项指标获得了最优值, 与基准模型 UNet-Resnet50 相比, mIoU、mPA、mPrecision、mRecall 以及 Acc 分别提升了 5.11%、4.05%、2.19%、4.05% 和 0.14%。与次优模型 BRNet 相比, mIoU、mPrecision、mRecall 以及 Acc 分别提升了 0.36%、0.62%、0.05% 和 0.01%。这些性能的提升, 充分展示了 CS-SegNet 模型在识别卒中区域时的高效性和可靠性。

综合来看, 本文所提出的 CS-SegNet 模型在脑卒中 CT 图像分割任务中展现出了卓越的性能, 不仅在各项关键指标上取得了显著优势, 还为临床诊断提供了有力的技术支持。这一研究成果有望为脑卒中患者的早期诊断和治疗带来积极的推动作用。

3.5 定性比较

本文针对多种深度学习模型在不同骨干网络作为模型进行训练损失函数的对比分析, 实验结果如图 8 所示。从实验结果可以看出, CS-SegNet 相比于其他网络收敛的较快, 大概在 60 个 epoch 之后, 模型整体损失趋于稳定, 且其最低损失值稳定在 0.01 左右。

为了直观展现各网络在脑卒中 CT 图像分割任务中的性能差异, 图 9 选取了不同脑 CT 图层的图像分割结果。从图中可以看出, 大多数分割网络都存在一些关于微小病灶的过分割和分割不足的问题, 这间接导致了脑卒中病灶

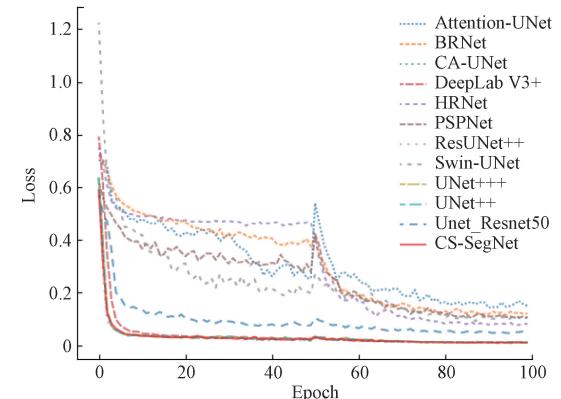


图 8 不同网络的损失函数曲线对比

Fig. 8 Comparison of loss function curves for different networks

分割精度的下降。具体地, 对于比较明显的病灶, 如第 1、6 行, 部分模型存在边界分割较为粗糙、边界不准确的问题; 对于多病灶情况, 如第 2、4 行, 多数模型存在漏检以及错检的问题, 将脑 CT 中的钙点错误的划分为病灶区域以及检测不出极小的病灶; 对于复杂病灶, 如第 3、5 行, 病灶区域紧靠着头骨区域, 甚至头骨区域有缺失, 多数模型不能有效感知与精准定位, 导致该部分区域被忽略。相比之下, 本文 CS-SegNet 网络的分割结果与标签图像高度一致, 能够较为精准的感知并定位脑卒中病灶区域, 抑制无关信息的干扰, 实现接近真实情况的分割效果, 有效避免了常见的错误分割、分割不足和过分割问题, 而且在处理微小病灶区域以及模糊边缘病灶区域时表现出了更高的准确性。这一结果表明, CS-SegNet 网络在处理复杂、多变

的脑卒中 CT 图像时具有更强的鲁棒性和适应性,为临床

诊断提供了更为可靠的技术支持。

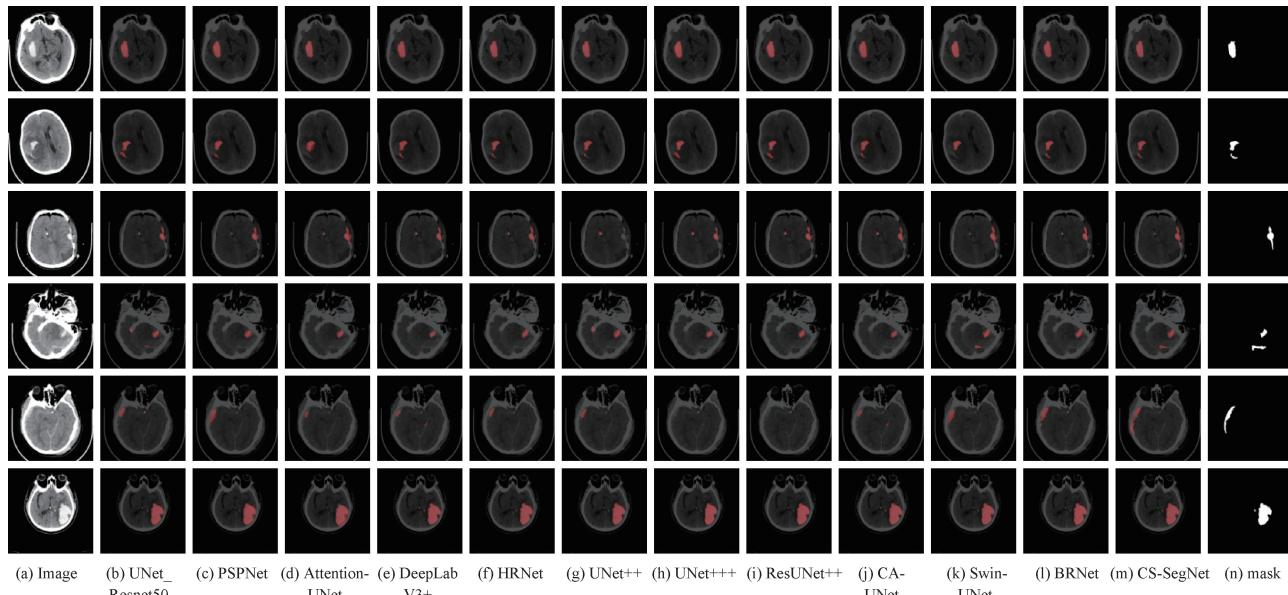


图 9 不同网络的分割结果对比图

Fig. 9 Segmentation results for different networks

4 结 论

为了进一步提升脑卒中图像分割的性能,本文创新性的提出了 CS-SegNet 网络模型,该模型融合了 CASA 模块、RDSConv 模块和 MSCA 模块,旨在全面提升模型的特征提取、图像重建及跨层特征融合能力。在下采样阶段,CASA 模块的加入显著增强了特征的表达能力,确保模型能够更精确地捕捉到图像中的关键信息,为后续分割任务奠定了坚实的基础。而在上采样阶段,RDSConv 模块则优化了图像重建过程,有效提升了细节恢复能力,使得分割结果更加精细、边缘更加清晰。此外,为了进一步加强跨层特征的融合,在跳跃连接部分引入了 MSCA 模块。这一创新设计使得模型能够更好地整合不同尺度的信息,从而进一步提升分割的准确性和鲁棒性。通过一系列消融实验充分验证了这些模块对模型性能的显著提升作用。尤其在平均交并比和精确度等关键指标上均取得了优异的成绩,在添加完 CASA 模块和 RDSConv 模块之后网络的整体分割准确率达到了 99.71%,平均交并比也达到了 88.81%,较基础模型增加了 0.06% 和 2.4%。最后添加 MSCA 模块与其他几种常用算法的比较实验也表明了 CS-SegNet 网络的有效性,其中整体的分割准确率、平均交并比以及召回率达到 99.79%、91.52% 和 93.83%,比 UNet-Resnet50 基础网络的性能分别提升了 0.14%、5.11% 和 4.05%,实验结果表明该模型的分割精度和分割性能较好,可以准确的分割出小区域病灶、复杂形态病灶等等,从而辅助医生进行脑卒中的精识别与定位,降低患者的风险,为脑卒中的临床诊疗提供重要的参考价值。

参考文献

- [1] 霍晓川,高峰. 急性缺血性卒中血管内治疗中国指南 2023[J]. 中国卒中杂志,2023,18(6): 684-711.
- [2] HUO X CH, GAO F. Chinese guidelines for endovascular treatment of acute ischemic stroke 2023[J]. Chinese Journal of Stroke, 2023, 18(6): 684-711.
- [3] 马丽媛,王增武,樊静,等.《中国心血管健康与疾病报告 2021》要点解读[J].中国全科医学,2022,25(27): 3331-3346.
- [4] MA L Y, WANG Z W, FAN J, et al. An essential introduction to the annual report on cardiovascular health and diseases in China (2021) [J]. Chinese General Practice, 2022, 25(27):3331-3346.
- [5] 王陇德.《中国脑卒中防治报告 2021》概要[J].中国脑血管病杂志,2023,20(11):783-793.
- [6] WANG L D. Brief report on stroke prevention and treatment in China, 2021 [J]. Chinese Journal of Cerebrovascular Diseases, 2023, 20(11): 783-793.
- [7] LEARY P J, LINDSTROM M, JOHNSON C O, et al. Global, regional, and national burden of pulmonary arterial hypertension, 1990-2021: A systematic analysis for the global burden of disease study 2021 [J]. The Lancet Respiratory Medicine, 2025, 13(1): 69-79.
- [8] 国家卫生健康委员会.中国卫生健康统计年鉴 2022[M].北京:中国协和医科大学出版社, 2022.
- [9] National Health Commission. China Health Statistical Yearbook 2022 [M]. Beijing: Peking Union Medical

- College Press, 2022.
- [6] ZHANG X, XU H, LIU Y, et al. A multiple encoders network for stroke lesion segmentation[C]. Chinese Conference on Pattern Recognition and Computer Vision. Cham: Springer, 2021: 524-535.
- [7] SHIN H, AGYEMAN R, RAFIQ M, et al. Automated segmentation of chronic stroke lesion using efficient U-Net architecture [J]. Biocybernetics and Biomedical Engineering, 2022, 42(1): 285-294.
- [8] CAO H, WANG Y Y, CHEN J, et al. Swin-Unet: Unet-like pure transformer for medical image segmentation[C]. European Conference on Computer Vision, Cham: Springer Nature Switzerland, 2022: 205-218.
- [9] GU Y, PIAO Z, YOO S J. STHarDNet: Swin transformer with HarDNet for MRI segmentation[J]. Applied Sciences, 2022, 12(1): 468.
- [10] WANG J, WANG SH Y, LIANG W. METrans: Multi-encoder transformer for ischemic stroke segmentation[J]. Electronics Letters, 2022, 58(9): 340-342.
- [11] WU Z, ZHANG X, LI F, et al. Multi-scale long-range interactive and regional attention network for stroke lesion segmentation [J]. Computers and Electrical Engineering, 2022, 103: 108345.
- [12] WU Z L, ZHANG X Y, LI F L, et al. W-Net: A boundary-enhanced segmentation network for stroke lesions[J]. Expert Systems with Applications, 2023, 230: 120637.
- [13] YU W Y, HUANG ZH ZH, ZHANG J P, et al. SAN-Net: Learning generalization to unseen sites for stroke lesion segmentation with self-adaptive normalization [J]. Computers in Biology and Medicine, 2023, 156: 106717.
- [14] YU W Y, LEI Y M, SHAN H M. Fan-net: Fourier-based adaptive normalization for cross-domain stroke lesion segmentation[C]. ICASSP 2023—2023 IEEE International Conference on Acoustics, Speech and Signal Processing. Rhodes Island:IEEE, 2023: 1-5.
- [15] RUAN J, LI J, XIANG S. VM-UNet: Vision mamba UNet for medical image segmentation [J]. Electrical Engineering and Systems Science, 2024, 14(8): 129447.
- [16] WU R, LIU Y, LIANG P, et al. H-VMUNet: High-order vision mamba unet for medical image segmentation [J]. Neurocomputing, 2025, 624: 129447.
- [17] 刘拥民, 张毅, 欧阳凌轩, 等. 基于改进 U-Net 的轻量级眼底病变分割算法设计[J]. 电子测量技术, 2024, 47(3):127-134.
- LIU Y M, ZHANG Y, OUYANG L X, et al. Design of lightweight fundus lesion segmentation algorithm based on improved U-Net[J]. Electronic Measurement Technology, 2024, 47(3): 127-134.
- [18] 武丽, 丁琴, 葛彩成. 融合公平因子的半监督学习医学图像分割模型[J]. 电子测量技术, 2024, 47(23): 171-180.
- WU L, DING Q, GE C CH. Semi-supervised learning medical image segmentation model fused with equity factors [J]. Electronic Measurement Technology, 2024, 47(23):171-180.
- [19] 张攀峰, 杨贺, 神显豪, 等. 融合局部和全局特征的息肉分割模型[J]. 电子测量技术, 2024, 47(16):100-109.
- ZHANG P F, YANG H, SHEN X H, et al. Polyp segmentation model based on fusion of local and global features [J]. Electronic Measurement Technology, 2024, 47(16):100-109.
- [20] WU D K, ZHAO J M, WANG Z F. AM-PSPNet: Pyramid scene parsing network based on attentional mechanism for image semantic segmentation [C]. International Conference of Pioneering Computer Scientists, Engineers and Educators. Singapore: Springer Nature Singapore, 2022: 425-434.
- [21] HUANG H M, LIN L F, TONG R F, et al. Unet 3+: A full-scale connected unet for medical image segmentation [C]. ICASSP 2020—2020 IEEE International Conference on Acoustics, Speech and Signal Processing. Barcelona:IEEE, 2020: 1055-1059.
- [22] FOIVOS I D, FRANCOIS W, PETER C, et al. ResUNet-a: A deep learning framework for semantic segmentation of remotely sensed data [J]. ISPRS Journal of Photogrammetry and Remote Sensing, 2020, 162: 94-114.
- [23] WANG K, HE B B, ZHU W P. CAUNet: Context-aware U-Net for speech enhancement in time domain[C]. 2021 IEEE International Symposium on Circuits and Systems. Daegu:IEEE, 2021: 1-5.
- [24] XU J, LIU X F, XU L M, et al. BRNet: A bilateral residual network for real-time semantic segmentation[C]. 2023 6th International Conference on Artificial Intelligence and Big Data. Chengdu:IEEE, 2023: 849-854.

作者简介

- 刘煜莹,硕士研究生,主要研究方向为医学图像处理、深度学习目标检测与图像分割。
E-mail: 6221913005@stu.jiangnan.edu.cn
- 李超生,副教授,副主任医师,博士,主要研究方向为中西医结合诊疗缺血性脑血管。
E-mail: wxsylcs@163.com
- 李恭新(通信作者),博士,副教授,硕士生导师,主要研究方向为软体机器人、微纳机器人以及生-机融合感知。
E-mail: gxli@jiangnan.edu.cn