

基于2.5D Attention U-Net的颅脑CT图像分割研究

郭靖宇¹, 张云雷^{1,2*}, 陈文轩¹

¹应急管理大学, 计算机与信息安全学院, 北京

²河北省物联网监控技术创新中心, 北京

收稿日期: 2026年8月1日; 录用日期: 2026年8月31日; 发布日期: 2026年9月8日

摘要

颅脑CT图像中目标区域的自动分割能够为病灶定位、体积测量和后续定量分析提供基础。由于CT图像存在灰度差异小、边界不清、目标面积占比低以及空标签切片较多等问题, 常规二维分割网络容易出现漏检或边界缺失。针对上述问题, 本文构建一种2.5D Attention U-Net分割模型。该模型以当前切片及其前后相邻切片作为三通道输入, 在保留二维卷积计算效率的同时引入层间空间信息; 在U-Net跳跃连接处加入通道与空间联合注意力模块, 对编码端特征进行自适应加权; 并采用GroupNorm提高小批量训练时的稳定性。训练阶段使用二元交叉熵与Tversky损失的组合, 同时在验证集上选择预测阈值, 以减轻前景与背景不平衡对模型选择的影响。为检验模型效果, 本文设置单切片ResNet34-U-Net作为对照, 并采用Dice、交并比、精确率、召回率及空切片误报率进行评价。本研究可为小目标、弱边界颅脑CT图像的自动分割提供一种计算量与层间信息利用相兼顾的实现方案。

关键词

颅脑CT, 图像分割, 2.5D, Attention U-Net, ResNet34, 深度学习

Research on Craniocerebral CT Image Segmentation Based on 2.5D Attention U-Net

Jingyu Guo¹, Yunlei Zhang^{1,2*}, Wenxuan Chen¹

¹School of Computer Science and Information Security, University of Emergency Management, Beijing

²Hebei Internet of Things Monitoring Technology Innovation Center, Beijing

Received: August 1, 2026; accepted: August 31, 2026; published: September 8, 2026

*通讯作者。

文章引用: 郭靖宇, 张云雷, 陈文轩. 基于 2.5D Attention U-Net 的颅脑 CT 图像分割研究[J]. 计算机科学与应用, 2026, 16(9): 14-24. DOI: 10.12677/csa.2026.169285

Abstract

Automatic segmentation of target regions in brain CT images provides a foundation for lesion localization, volume measurement, and subsequent quantitative analysis. However, brain CT images often exhibit small grayscale differences, unclear boundaries, a low target-area proportion, and a large number of slices with empty labels. These issues can lead conventional two-dimensional segmentation networks to miss targets or produce incomplete boundaries. To address these problems, this study developed a 2.5D Attention U-Net segmentation model. The model uses the current slice together with its preceding and following slices as a three-channel input, thereby introducing inter-slice spatial information while retaining the computational efficiency of two-dimensional convolution. A combined channel and spatial attention module is incorporated into the skip connections of U-Net to adaptively reweight encoder features. Group Normalization is employed to improve training stability under small-batch conditions. During training, a combination of binary cross-entropy loss and Tversky loss is used. The prediction threshold is selected on the validation set to reduce the influence of foreground-background imbalance on model selection. To evaluate the model, a single-slice ResNet34-U-Net is used as a comparison model. Dice, Intersection over Union, precision, recall, and the false-positive rate on empty slices are used as evaluation metrics. This study provides an implementation approach for automatic segmentation of small-target and weak-boundary regions in brain CT images, balancing the use of inter-slice information with computational efficiency.

Keywords

Brain CT, Image Segmentation, 2.5D, Attention U-Net, ResNet34, Deep Learning

Copyright © 2026 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

1. 引言

医学图像分割是计算机辅助诊断中的基础环节，其任务是按照像素或体素对图像中的组织、器官或异常区域进行划分。对颅脑 CT 而言，分割结果不仅用于显示目标位置，还可用于计算面积、体积、形状和灰度等信息，并为病情评估、随访比较及影像组学分析提供输入。与单纯分类相比，分割要求模型同时回答“是否存在目标”和“目标边界在哪里”，因此对图像质量、标注一致性和模型的空间表达能力提出了更高要求。

U-Net 由 Ronneberger 等提出，其编码器提取多尺度语义特征，解码器逐步恢复空间分辨率，并通过跳跃连接融合浅层细节与深层语义信息[1]。该结构对样本规模要求相对较低，已经成为医学图像分割中常用的基础网络。在此基础上，Attention U-Net 通过注意力机制抑制与目标无关的响应，使网络在解码过程中更加关注目标区域[2]。U-Net++采用嵌套跳跃连接减小编码端和解码端特征之间的语义差异[3]，nnU-Net 则从数据预处理、网络配置、训练和后处理等方面建立自适应流程，在多项公开任务中表现出较强的适用性[4]。这些研究说明，U 形结构仍是医学图像分割的重要基础，但具体任务中的效果不仅取决于网络名称，还与输入方式、归一化、损失函数和评价策略有关。

从输入形式看，二维网络计算量较低，适合直接处理轴位切片，但无法显式利用相邻层面的连续信息；三维网络能够处理完整体数据，却需要更大的显存和更多训练样本。2.5D 方法位于二者之间，通常将连续切片作为多个输入通道，仍使用二维卷积完成特征提取。该方式不会建立完整的三维卷积计算图，

但能够向模型提供一定的层间变化信息, 适合显存受限且切片厚度相对稳定的 CT 任务。

除模型结构外, 类别不平衡也是影响分割结果的重要因素。Ma 等多种医学图像分割损失进行了系统比较, 指出不同任务对损失函数的响应存在差异, 不能仅依赖单一经验配置[5]。在目标较小或空标签较多的情况下, 二元交叉熵有利于维持整体像素分类, Dice 或 Tversky 类损失则更关注预测区域与真实标签的重叠。合理组合不同损失, 并将非空切片指标与空切片误报共同纳入评价, 能够减少“总体指标较高但目标漏检严重”的误判。

基于以上分析, 本文以 2.5D 输入和联合注意力为主要改进方向, 构建颅脑 CT 图像分割模型。模型将相邻三张切片组合为输入, 在 U-Net 跳跃连接处设置通道和空间注意力模块, 并使用适合小批量训练的 GroupNorm。为避免只与结构较弱的模型比较, 本文选用 ResNet34 作为编码器构建单切片 ResNet34-U-Net 对照模型[6]。

2. 国内外研究现状

2.1. 医学图像分割方法的发展

早期医学图像分割常采用阈值、区域生长、边缘检测、聚类和活动轮廓等方法。这类方法原理清晰, 但往往依赖人工设定参数, 对噪声、伪影和灰度变化较为敏感。当目标边界不完整或不同病例差异较大时, 固定规则很难保持稳定。卷积神经网络能够从训练样本中学习多层次特征, 逐渐取代以人工特征为主的处理流程。

U-Net 提出后, 编码器-解码器和跳跃连接成为医学图像分割的常见设计。其优点是能够同时利用局部纹理和高层语义, 但普通跳跃连接会把目标信息与背景信息一并传递到解码端。Attention U-Net 在跳跃连接处引入注意力门, 根据解码端上下文对编码特征进行筛选, 从而提高目标区域的响应。U-Net++ 进一步通过多层嵌套连接融合不同尺度特征。

近年的研究开始从自动配置、全局建模和通用分割三个方向扩展。nnU-Net 通过数据指纹和规则化配置自动确定预处理、网络规模和训练方案, 表明规范的训练流程与网络结构同样重要。UNETR 使用 Transformer 编码三维图像块, 以增强长距离依赖建模能力, 并通过 U 形解码器恢复分割结果[7]。MedSAM 利用大规模医学图像-掩膜数据对通用分割模型进行适配, 推动医学分割从单任务专用模型向基础模型发展[8]。SegMamba 则尝试以状态空间模型处理三维长序列, 在全局信息建模和计算效率之间寻找新的平衡[9]。

上述方法拓展了医学图像分割的技术范围, 但模型越复杂, 对数据规模、硬件和训练稳定性的要求通常越高。对于样本量有限、目标较小的颅脑 CT 任务, 直接使用大规模 Transformer 或基础模型并不一定优于经过合理配置的卷积网络。因此, 当前研究一方面关注更强的全局表示能力, 另一方面也重视小批量训练、类别不平衡和跨设备泛化等实际问题。

2.2. 国外颅脑 CT 自动分割研究

国外针对颅脑 CT 的深度学习研究已由单纯检出逐步发展到定位、分割和定量分析。MacIntosh 等汇集多个来源的非增强 CT 数据, 对自动分割得到的脑出血影像特征进行分析, 说明三维分割结果可用于不同病因和不同中心数据的定量研究[10]。这类工作强调外部数据验证, 因为单中心训练结果较好并不代表模型能够适应扫描协议和病例构成的变化。

Lin 和 Yuh 采用半监督学习研究颅内出血检测与分割, 希望利用未标注数据提高模型在分布外数据上的表现[11]。该研究反映出医学影像标注成本较高, 完全依赖人工逐层勾画会限制样本规模。半监督、自监督和伪标签方法能够扩大训练数据, 但错误伪标签也可能被模型重复学习, 因此仍需可靠的置信度

控制和外部测试。

Hoang 等提出面向 CT 颅内出血分割的高效卷积方法, 并通过数据增强与卷积结构改进提升重叠指标[12]。与大规模三维网络相比, 高效二维或 2.5D 网络更容易部署, 也适合显存有限的实验环境。总体来看, 国外相关研究较重视多中心验证、自动定量和模型泛化, 评价指标也由单一 Dice 扩展到灵敏度、特异度、精确率、体积误差及运行时间等。

2.3. 国内颅脑 CT 自动分割研究

国内在颅脑 CT 数据建设、网络改进和临床评价方面开展了较多研究。Ma 等发布 PHE-SICH-CT-IDS 数据集, 为自发性脑出血相关区域的语义分割、目标检测和影像组学特征提取提供了统一数据基础[13]。公开数据能够降低重复标注成本, 也便于不同算法在相同病例上的比较。但使用公开数据时仍需关注患者级划分, 避免同一患者的相邻切片同时进入训练集和测试集。

Hu 等对非增强 CT 中深度学习辅助颅内出血检测和分割研究进行系统评价, 结果表明相关方法总体具有较好的检测与分割能力, 但不同研究在数据来源、目标定义和验证方式上差异较大[14]。

Li 等构建残差编码 U-Net, 分别处理脑内出血和脑室内出血, 并通过外部数据检验模型及分割质量评估方法[15]。该研究说明残差编码器能够增强特征提取, 但模型效果仍受到目标类型和外部数据差异影响。国内研究还广泛使用多尺度卷积、注意力、Transformer 和组合损失改进 U-Net, 其中注意力主要用于突出目标区域, 多尺度结构用于处理目标大小变化, 组合损失用于缓解类别不平衡。

总体上, 国内相关研究已经由模型结构改进逐步转向数据集建设、外部验证和临床定量应用。不过, 部分研究仍存在数据量较小、仅进行切片级随机划分、对空标签处理说明不足、不同模型训练条件不一致等问题。

2.4. 现有研究的主要问题

结合国内外研究, 颅脑 CT 分割仍有四个值得关注的问题。第一, 二维切片缺少层间信息, 三维网络又存在显存占用大、训练时间长的问題。第二, 小目标与大量背景造成类别不平衡, 模型容易输出空掩膜, 普通 Accuracy 对此不敏感。第三, 注意力、Transformer 等模块能够增强特征表达, 但参数增加并不必然带来稳定提升。第四, 不同研究采用不同阈值、数据划分和后处理方式, 结果之间难以直接比较。

针对这些问题, 本文选择 2.5D 而非完整三维输入, 以较低成本利用相邻切片信息; 在跳跃连接处进行通道和空间加权, 减少无关特征传递; 训练中使用组合损失, 并根据验证集选择阈值; 测试时除 Dice 和 IoU 外, 同时统计精确率、召回率和空切片误报率。该方案重点解决模型容易漏检、边界不完整和评价不充分的问题。

3. 模型与方法

3.1. 模型总体结构

本文模型以 U-Net 为基本框架, 由编码器、瓶颈层、解码器和分割输出层组成。输入为连续三张轴位 CT 切片, 分别对应中心切片的前一层、当前层和后一层。三张图像按通道方向组合后送入二维卷积网络, 网络最终仅预测中心切片的二值掩膜。该输入方式保留了二维网络的计算形式, 同时使中心切片的判断能够参考相邻层面的结构变化。

编码器包含 4 个尺度, 每个尺度使用两次 3×3 卷积、GroupNorm 和 ReLU 激活, 特征通道依次为 64、128、256 和 512, 尺度之间使用 2×2 最大池化。瓶颈层通道数为 1024。解码器使用转置卷积逐级上采样, 并与经过注意力加权的同尺度编码特征拼接, 再通过双卷积完成融合。最后使用 1×1 卷积将通

道数压缩为 1，并通过 Sigmoid 得到每个像素属于目标区域的概率。

3.2. 2.5D 三切片输入

设第 k 层为中心切片，其前后相邻切片分别取自第 $k-1$ 层和第 $k+1$ 层。三张切片沿通道方向拼接，模型输入定义如式(1)所示。对于序列首尾没有相邻层的情况，采用边界切片复制方式补齐。监督标签为中心切片对应的掩膜，因此该任务仍属于二维监督分割。

$$X_k = [I_{k-1}, I_k, I_{k+1}] \in \mathbb{R}^{3 \times H \times W} \quad (1)$$

单张 CT 切片中某些目标可能只表现为低对比度或不完整边缘，而相邻层可以提供目标出现、扩大、缩小或消失的趋势。三切片输入使第一层卷积能够联合学习这些变化。与三维卷积相比，2.5D 方法参数和显存增长较小，也不需要整套体数据裁剪为三维块。但其局限是只利用有限层面，不能完全表示长距离三维关系；若不同序列的层厚差异较大，相邻切片提供的空间尺度也会不一致。

3.3. GroupNorm 与编解码特征融合

医学 CT 原始分辨率较高，训练时受显存限制，批量大小通常较小。BatchNorm 依赖一个批次的均值和方差，当批量过小时统计量容易波动。本文在主模型卷积块中使用 GroupNorm，将通道划分为若干组后在组内归一化。

编码器通过逐级下采样扩大感受野，深层特征包含较强语义信息，但空间细节减少；解码器通过上采样恢复分辨率。跳跃连接把浅层边缘和纹理直接传递给解码端，有利于恢复边界。不过，浅层特征也包含颅骨、高亮伪影和扫描背景等信息，因此本文不直接拼接全部编码特征，而是在拼接前进行注意力加权。

3.4. 通道 - 空间联合注意力模块

本文在 4 个跳跃连接处设置联合注意力模块。模块先计算通道权重，再计算空间权重。对于输入特征，通道分支先进行全局平均池化，再通过两个 1×1 卷积、ReLU 和 Sigmoid 得到通道权重，如式(2)所示。通道加权特征按式(3)计算，用于增强与目标有关的特征通道。

$$M_c(F) = \sigma(W_2 \delta(W_1 \text{GAP}(F))) \quad (2)$$

$$F_c = F \odot M_c(F) \quad (3)$$

空间分支对通道加权特征分别计算通道平均值和最大值，将两张二维特征图拼接后通过 7×7 卷积和 Sigmoid 生成空间权重，如式(4)所示。最终注意力特征按式(5)计算，并与对应的上采样特征拼接。

$$M_s(F_c) = \sigma(f^{7 \times 7}([\text{AvgPool}_c(F_c); \text{MaxPool}_c(F_c)])) \quad (4)$$

$$F_a = F_c \odot M_s(F_c) \quad (5)$$

该模块与经典 Attention U-Net 的门控形式并不完全相同。经典方法通常同时使用编码特征和解码端门控信号计算注意力系数；本文实现直接从编码特征中提取通道和空间权重，结构更接近顺序式通道 - 空间注意力。

3.5. 损失函数与预测阈值

本文采用二元交叉熵与 Tversky 损失的加权和，总损失定义如式(6)所示。二元交叉熵约束每个像素的分类概率，Tversky 损失通过真阳性、假阳性和假阴性控制预测区域与真实标签的重叠，定义如式(7)所

示。当前设置中, Tversky 损失权重为 0.5, α 取 0.6, β 取 0.4, s 为平滑项。对于空标签切片, 仅计算二元交叉熵。

$$L = 0.5L_{\text{BCE}} + 0.5L_{\text{Tversky}} \quad (6)$$

$$T = \frac{\text{TP} + s}{\text{TP} + \alpha\text{FP} + \beta\text{FN} + s}, L_{\text{Tversky}} = 1 - T \quad (7)$$

分割概率需要通过阈值转换为二值掩膜。固定使用 0.5 虽然简单, 但不一定适合前景比例很低的任务。本文在验证集上对多个候选阈值进行评价, 根据非空切片 Dice 选择与最佳模型对应的阈值, 并在测试阶段保持该阈值不变。该做法可以避免使用测试集调参。

3.6. ResNet34-U-Net 对照模型

ResNet 通过残差连接学习输入与目标映射之间的残差, 其基本形式如式(8)所示, 使梯度能够沿恒等分支传播, 缓解深层网络训练困难。本文使用 ResNet34 作为对照模型的编码器。其输入为单张灰度 CT 切片, 首层卷积通道数由 3 改为 1, 不加载 ImageNet 预训练权重; 解码端采用普通 U-Net 式转置卷积和跳跃连接。

$$y = F(x; \{W_i\}) + x \quad (8)$$

设置 ResNet34-U-Net 的目的在于检验较深残差编码器在相同数据上的表现, 并与 2.5D 注意力模型进行结构对照。

4. 实验与结果分析

4.1. 数据集与预处理

本研究采用 PHE-SICH-CT-IDS 数据集集中的 SubdatasetC_PNG 子集。当前项目共使用 119 个病例/CT 序列, 包含 3478 张配对图像和标签, 其中非空标签切片 676 张, 空标签切片 2802 张。按照病例/序列进行划分, 训练集、验证集和测试集分别包含 83、18 和 18 个病例, 对应 2402、550 和 526 张切片。三部分数据中的非空标签切片分别为 467、105 和 104 张。预处理阶段将图像和标签统一调整为 512×512 像素。图像采用区域插值, 标签采用最近邻插值, 以避免改变标签类别。图像灰度值由 0~255 归一化到 0~1, 标签中灰度值大于 127 的像素被转换为前景, 其余像素转换为背景。训练阶段使用水平翻转、旋转、缩放、平移和强度扰动等成对数据增强, 验证集和测试集不进行数据增强。

4.2. 实验环境与参数

Table 1. Experimental environment and training parameters

表 1. 实验环境与训练参数

项目	2.5D Attention U-Net	ResNet34-U-Net (对照模型)
输入方式	相邻 3 张切片 (3 通道)	单张灰度 CT 切片 (1 通道)
训练轮数	100 轮	150 轮
批量大小	2	2
优化器	AdamW	Adam
初始学习率	5×10^{-5}	1×10^{-4}
权重衰减	5×10^{-4}	0

续表

损失函数	0.5 BCE + 0.5 Tversky $\alpha = 0.6, \beta = 0.4$	0.5 BCE + 0.5 Dice
最佳阈值	0.005 (验证集选择)	0.5 (固定)
参数量	31,213,309 (约 3121 万)	24,445,313 (约 2445 万)

见表 1, 本实验分别采用 2.5D Attention U-Net 和 ResNet34-U-Net 进行对比。2.5D Attention U-Net 以当前切片及其前后相邻切片作为三通道输入, 训练 100 轮, 批量大小为 2, 采用 AdamW 优化器, 初始学习率为 (5×10^{-5}) , 权重衰减系数为 (5×10^{-4}) , 损失函数由 0.5 倍 BCE 损失和 0.5 倍 Tversky 损失组成, 其中 $(\alpha=0.6)$ 、 $(\beta=0.4)$ 。模型预测阈值在验证集上进行选择, 最终确定为 0.005, 模型参数量为 31,213,309, 约 3121 万。ResNet34-U-Net 以单张灰度 CT 切片作为单通道输入, 训练 150 轮, 批量大小同样为 2, 采用 Adam 优化器, 初始学习率为 (1×10^{-4}) , 不设置权重衰减, 损失函数由 0.5 倍 BCE 损失和 0.5 倍 Dice 损失组成, 预测阈值固定为 0.5, 模型参数量为 24,445,313, 约 2445 万。

4.3. 评价指标

分割结果采用 Dice、IoU、Precision、Recall 和 Empty FPR 进行评价, 各指标定义如式(9)~(13)所示。其中, TP、FP 和 FN 分别表示真阳性、假阳性和假阴性像素数; 空切片假阳性数表示真实标签为空但模型产生非空预测的切片数。

$$\text{Dice} = \frac{2\text{TP}}{2\text{TP} + \text{FP} + \text{FN}} \quad (9)$$

$$\text{IoU} = \frac{\text{TP}}{\text{TP} + \text{FP} + \text{FN}} \quad (10)$$

$$\text{Precision} = \frac{\text{TP}}{\text{TP} + \text{FP}} \quad (11)$$

$$\text{Recall} = \frac{\text{TP}}{\text{TP} + \text{FN}} \quad (12)$$

$$\text{Empty FPR} = \frac{N_{\text{FP-empty}}}{N_{\text{empty}}} \quad (13)$$

4.4. 对比实验结果

Table 2. Quantitative comparison of segmentation results
表 2. 分割结果定量对比

模型	Dice	Nonempty Dice	IoU	Precision	Recall	Empty FPR
2.5D Attention U-Net	0.6184	0.5316	0.3806	0.4079	0.5542	0.3602
ResNet34-U-Net	0.6982	0.4352	0.2914	0.3362	0.5851	0.2370

如表 2 所示, Dice 表示全部测试切片的平均 Dice 系数; Nonempty Dice 仅统计真实标注不为空的切片, 用于反映模型对目标区域的实际分割能力; IoU 表示非空标注切片的平均交并比。Precision 和 Recall 根据像素级混淆矩阵计算, 分别反映模型预测结果的准确性和目标区域的检出能力。Empty FPR 表示真

实标注为空但模型预测出目标区域的切片比例, 该指标越低, 说明模型产生的空切片误报越少。

由表中结果可知, 2.5D Attention U-Net 与 ResNet34-U-Net 在不同评价指标上各有特点。ResNet34-U-Net 的整体 Dice 为 0.6982, 高于 2.5D Attention U-Net 的 0.6184, 说明其在包含大量背景区域的全部测试切片上具有较好的整体重叠表现。然而, 在真实标签非空切片上, 2.5D Attention U-Net 的 Nonempty Dice 达到 0.5316, 高于 ResNet34-U-Net 的 0.4352; 其 IoU 和 Precision 分别为 0.3806 和 0.4079, 也高于对照模型的 0.2914 和 0.3362。这表明 2.5D Attention U-Net 在实际存在目标区域的切片中能够获得更准确的前景定位, 并在一定程度上改善目标区域的边界匹配效果。从 Recall 指标来看, ResNet34-U-Net 为 0.5851, 略高于 2.5D Attention U-Net 的 0.5542, 说明对照模型在目标区域召回方面具有一定优势, 能够覆盖更多真实标签像素。与此同时, 2.5D Attention U-Net 的 Empty FPR 为 0.3602, 高于 ResNet34-U-Net 的 0.2370, 表明其在部分真实标签为空的切片中存在误检现象。综合来看, 2.5D Attention U-Net 在非空切片的区域重叠质量和预测精度方面表现较好, 但仍需要进一步降低空切片误报并提升目标区域召回率。

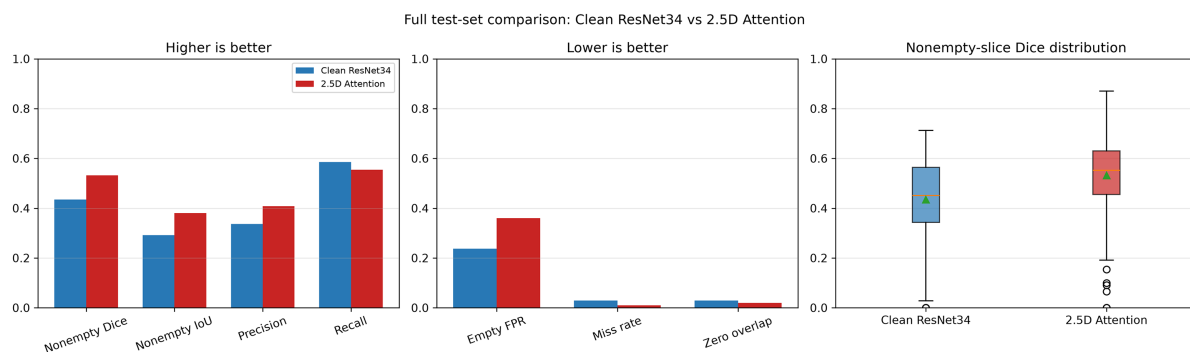


Figure 1. Global comparison of the two models

图 1. 双模型全局对比图

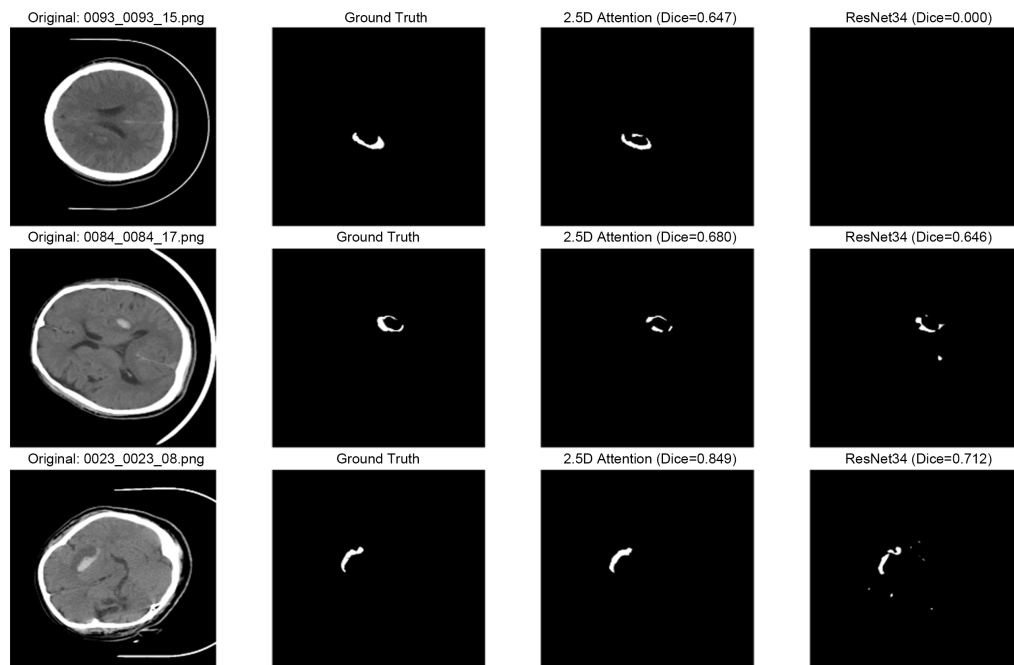


Figure 2. Comparison of the two models (1)

图 2. 双模型实验对比(1)

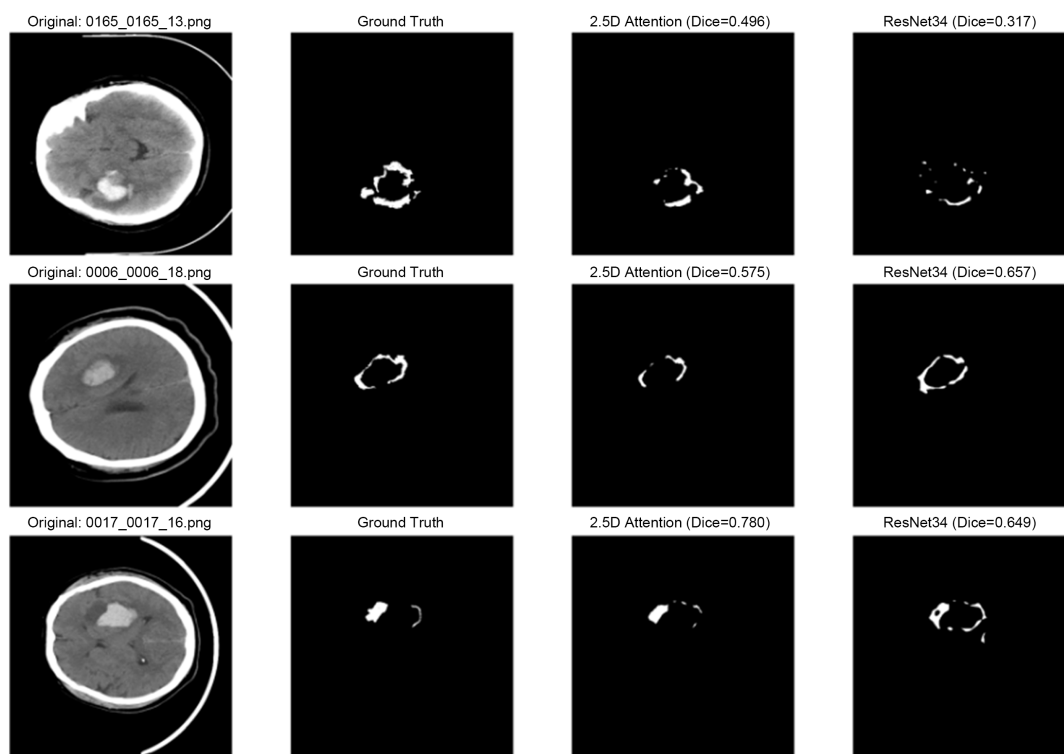


Figure 3. Comparison of the two models (2)

图 3. 双模型实验对比(2)

从预测结果图，图 1 可以看出，2.5D Attention U-Net 能够较好地识别多数非空切片中的目标区域，预测区域的位置与 Ground Truth 基本一致，部分样本的轮廓和面积也较为接近。但对于面积较小、边界不清晰或形态细长的目标，模型仍存在局部漏检和边界不完整现象。相比之下，ResNet34-U-Net 在部分样本中具有较高的召回能力，但预测区域相对粗糙，边界贴合程度和局部细节表现不足。

由图 2、图 3 可得出，2.5D Attention U-Net 对目标区域的整体轮廓具有较好的保持能力，在部分边界不清和形态不规则的切片中能够获得较完整的预测结果。相比之下，ResNet34-U-Net 在部分样本中出现区域缺失或边界断裂现象。

5. 结论

(1) 针对颅脑 CT 图像中目标区域小、边界模糊、灰度对比度低以及空标签切片较多等问题，本文构建了 2.5D Attention U-Net 模型。模型将当前切片与前后相邻切片组合为三通道输入，在保留二维卷积计算效率的同时，引入了一定的层间空间信息，有助于提升模型对目标连续性和局部结构的判断能力。

(2) 本文在 U-Net 跳跃连接中加入通道与空间联合注意力模块，并采用 GroupNorm 提高小批量训练过程的稳定性。同时结合二元交叉熵与 Tversky 损失，并在验证集上选择预测阈值，使模型能够更加关注非空标签切片中的前景区域。实验结果表明，该方法在非空切片 Dice、IoU、Precision 和 Recall 等指标上具有较好的表现，对目标区域的整体轮廓和局部边界能够进行较为准确地恢复。

(3) 通过与单切片 ResNet34-U-Net 进行对比，验证了相邻切片信息和注意力特征筛选对颅脑 CT 分割任务的积极作用。需要指出的是，当前模型对面积较小、边界不明显或形态差异较大的目标仍存在一定漏检和边界缺失现象。

(4) 未来将进一步扩大数据集规模，引入不同设备、不同扫描条件及多中心颅脑 CT 数据，以提高模

型的泛化能力。同时,针对当前模型在小目标、弱边界区域存在漏检和空切片误报的问题,将结合更有效的特征增强模块、三维上下文信息和后处理方法进行优化,并通过外部测试集验证模型的稳定性与实际应用价值。

(5) 需要指出的是,2.5D Attention U-Net 并非在所有评价指标上均优于对照模型。本实验中,该模型在非空切片 Dice、IoU 和 Precision 方面表现较好,说明其对实际目标区域具有较好的定位和区域匹配能力;但其 Recall 相对较低,且 Empty FPR 较高,表明模型在提高部分前景预测精度的同时,仍存在目标漏检和空标签切片误报问题。因此,该模型更适合作为颅脑 CT 图像分析中的辅助分割工具,而不宜直接作为独立诊断依据。从应用场景来看,在对目标区域定位精度和边界质量要求较高、且允许医生进行二次复核的辅助分析任务中,该模型具有一定应用价值。对于强调高召回率的临床初筛场景,可以通过适当降低预测阈值来增加目标检出数量,但这也可能进一步提高误报率,需要结合人工审核、连通域筛选或其他后处理方法共同使用。当前模型仍受到数据规模、数据来源、目标类型和 2.5D 有限层间信息等因素的限制,尚不能完全适应不同设备、不同扫描条件下的复杂病例。

基金项目

河北重大科技支撑计划项目(202622003010002)。

参考文献

- [1] Ronneberger, O., Fischer, P. and Brox, T. (2015) U-Net: Convolutional Networks for Biomedical Image Segmentation. In: Navab, N., Hornegger, J., Wells, W. and Frangi, A., Eds., *Medical Image Computing and Computer-Assisted Intervention—MICCAI 2015*, Springer, 234-241. https://doi.org/10.1007/978-3-319-24574-4_28
- [2] Oktay, O., Schlemper, J., Le Folgoc, L., et al. (2018) Attention U-Net: Learning Where to Look for the Pancreas. arXiv: 1804.03999.
- [3] Zhou, Z., Siddiquee, M.M.R., Tajbakhsh, N. and Liang, J. (2020) UNet++: Redesigning Skip Connections to Exploit Multiscale Features in Image Segmentation. *IEEE Transactions on Medical Imaging*, **39**, 1856-1867. <https://doi.org/10.1109/tmi.2019.2959609>
- [4] Isensee, F., Jaeger, P.F., Kohl, S.A.A., Petersen, J. and Maier-Hein, K.H. (2020) nnU-Net: A Self-Configuring Method for Deep Learning-Based Biomedical Image Segmentation. *Nature Methods*, **18**, 203-211. <https://doi.org/10.1038/s41592-020-01008-z>
- [5] Ma, J., Chen, J., Ng, M., Huang, R., Li, Y., Li, C., et al. (2021) Loss Odyssey in Medical Image Segmentation. *Medical Image Analysis*, **71**, Article ID: 102035. <https://doi.org/10.1016/j.media.2021.102035>
- [6] He, K., Zhang, X., Ren, S. and Sun, J. (2016) Deep Residual Learning for Image Recognition. 2016 *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Las Vegas, 27-30 June 2016, 770-778. <https://doi.org/10.1109/cvpr.2016.90>
- [7] Hatamizadeh, A., Tang, Y., Nath, V., Yang, D., Myronenko, A., Landman, B., et al. (2022) UNETR: Transformers for 3D Medical Image Segmentation. 2022 *IEEE/CVF Winter Conference on Applications of Computer Vision (WACV)*, Waikoloa, 3-8 January 2022, 1748-1758. <https://doi.org/10.1109/wacv51458.2022.00181>
- [8] Ma, J., He, Y., Li, F., Han, L., You, C. and Wang, B. (2024) Segment Anything in Medical Images. *Nature Communications*, **15**, Article No. 654. <https://doi.org/10.1038/s41467-024-44824-z>
- [9] Xing, Z., Ye, T., Yang, Y., Liu, G. and Zhu, L. (2024) SegMamba: Long-Range Sequential Modeling Mamba for 3D Medical Image Segmentation. In: Linguraru, M.G., et al., Eds., *Medical Image Computing and Computer Assisted Intervention—MICCAI 2024*, Springer, 578-588. https://doi.org/10.1007/978-3-031-72111-3_54
- [10] MacIntosh, B.J., Liu, Q., Schellhorn, T., Beyer, M.K., Groote, I.R., Morberg, P.C., et al. (2023) Radiological Features of Brain Hemorrhage through Automated Segmentation from Computed Tomography in Stroke and Traumatic Brain Injury. *Frontiers in Neurology*, **14**, Article 1244672. <https://doi.org/10.3389/fneur.2023.1244672>
- [11] Lin, E. and Yuh, E.L. (2024) Semi-Supervised Learning for Generalizable Intracranial Hemorrhage Detection and Segmentation. *Radiology: Artificial Intelligence*, **6**, e230077. <https://doi.org/10.1148/ryai.230077>
- [12] Hoang, Q.T., Pham, X.H., Trinh, X.T., Le, A.V., Bui, M.V. and Bui, T.T. (2024) An Efficient CNN-Based Method for Intracranial Hemorrhage Segmentation from Computerized Tomography Imaging. *Journal of Imaging*, **10**, Article 77. <https://doi.org/10.3390/jimaging10040077>

-
- [13] Ma, D., Li, C., Du, T., Qiao, L., Tang, D., Ma, Z., *et al.* (2024) PHE-SICH-CT-IDS: A Benchmark CT Image Dataset for Evaluation Semantic Segmentation, Object Detection and Radiomic Feature Extraction of Perihematoma Edema in Spontaneous Intracerebral Hemorrhage. *Computers in Biology and Medicine*, **173**, Article ID: 108342. <https://doi.org/10.1016/j.compbiomed.2024.108342>
 - [14] Hu, P., Yan, T., Xiao, B., Shu, H., Sheng, Y., Wu, Y., *et al.* (2024) Deep Learning-Assisted Detection and Segmentation of Intracranial Hemorrhage in NonContrast Computed Tomography Scans of Acute Stroke Patients: A Systematic Review and Meta-Analysis. *International Journal of Surgery*, **110**, 3839-3847. <https://doi.org/10.1097/js9.0000000000001266>
 - [15] Li, Y., Zhang, R., Li, Y., Zuo, X., Wang, Q., Zhang, S., *et al.* (2024) Deep Learning Models for Separate Segmentations of Intracerebral and Intraventricular Hemorrhage on Head CT and Segmentation Quality Assessment. *Medical Physics*, **51**, 8317-8333. <https://doi.org/10.1002/mp.17343>