

基于移位窗口自注意力机制的 新生儿脑区域图像分割

张啸成^{1,2}, 王涛^{1,2}, 田昕³, 张永刚^{2,4}

- 吉林大学 软件学院, 长春 130012;
- 吉林大学 符号计算与知识工程教育部重点实验室, 长春 130012;
- 吉林大学第二医院, 长春 130062;
- 吉林大学 计算机科学与技术学院, 长春 130012)

摘要: 通过改进 Swin Transformer 编码解码网络, 结合跳跃连接和深度监督机制, 针对新生儿脑部核磁共振成像(MRI)图像分割面临的低信噪比、组织对比度差等问题, 提出一种新的基于移位窗口自注意力机制的新生儿脑区域图像分割方法. 该方法经过对 MRI 图像的预处理后, 能实现新生儿脑部图像多功能区域的精确分割, 并采用最大连通域算法进一步提高了分割精度. 在数据集 dHCP 上的实验结果表明, 该方法优于现有方法, 为早期发现和干预新生儿脑部损伤提供了潜在的可能性.

关键词: 脑部图像区域分割; Swin Transformer 编码解码网络; 新生儿 MRI; 自注意力; 移位窗口

中图分类号: TP391.4 **文献标志码:** A **文章编号:** 1671-5489(2024)05-1129-09

Image Region Segmentation of Neonatal Brain Based on Self-attention Mechanism of Shifted Windows

ZHANG Xiaocheng^{1,2}, WANG Tao^{1,2}, TIAN Xin³, ZHANG Yonggang^{2,4}

- College of Software, Jilin University, Changchun 130012, China;
- Key Laboratory of Symbol Computation and Knowledge Engineering of Ministry of Education, Jilin University, Changchun 130012, China;
- The Second Hospital of Jilin University, Changchun 130062, China;
- College of Computer Science and Technology, Jilin University, Changchun 130012, China)

Abstract: By improving the Swin Transformer coding and decoding network, combined with the skip-linking and depth supervision mechanisms, we proposed a new image region segmentation method of neonatal brain based on self-attention mechanism of shifted windows to address the issues of low signal-to-noise ratio and poor tissue contrast in segmentation of nuclear magnetic resonance imaging (MRI) images of the neonatal brain. The method could achieve accurate segmentation of multifunctional regions of the neonatal brain images after preprocessing the MRI images, and further improve the segmentation accuracy by using the maximum connected domain algorithm. The

收稿日期: 2023-09-25.

第一作者简介: 张啸成(1998—), 男, 汉族, 硕士研究生, 从事计算机视觉的研究, E-mail: xc Zhang21@outlook.com. **通信作者简介:** 张永刚(1975—), 男, 汉族, 博士, 教授, 从事约束求解与约束优化的研究, E-mail: zhangyg@jlu.edu.cn.

基金项目: 吉林省自然科学基金(批准号: 20200201447JC).

experimental results on the dHCP dataset show that the method is superior to existing methods, providing potential possibilities for early detection and intervention of neonatal brain injury.

Keywords: brain image region segmentation; Swin Transformer coding and decoding network; neonatal MRI; self-attention; shifted window

近年来,随着新生儿护理技术的不断发展,新生儿存活率显著提高,但临床研究显示新生儿仍有发育迟缓和脑损伤等问题^[1-2].研究新生儿大脑发育,旨在了解新生儿独特能力的获得方式以及早期脑损伤对新生儿发育的影响^[3-4].

目前,核磁共振成像(MRI)技术广泛应用于神经系统研究中,但手动标记脑部结构耗时且较难.神经网络在医学图像分割领域可取代手动勾画,加快精准器官图像区域划分研究^[5-6].不同器官图像自动化分割难度不同,新生儿脑部图像分割尤其挑战,需考虑低信噪比和组织对比度.

Swin Transformer 模块^[7]作为一种高效网络模块,将 Transformer^[8]结构用于计算机视觉,降低了网络的复杂度和内存消耗.其移位窗口模块与传统滑动窗口模块不同,减少了硬件需求,提高了运行速度,扩大了感受野,提升了模型精度.

本文沿用移位窗口方法,加入深度监督机制,并配合全尺度的跳跃连接,搭建出一个全新的网络模型.其中,全尺度跳跃连接技术是通过在网络各层次之间建立跳跃连接,将低层的特征图与高层的特征图进行连接,从而使模型可同时考虑不同层次的特征信息,有效提高了模型性能.深度监督机制通过在深度神经网络的不同层级间进行监督训练,将底层网络的输出与标签进行对比,使梯度可以在底层网络中进行有效传递,从而解决了梯度消失及过拟合等问题,并提高了网络的收敛速度.此外,本文提出了一个端到端的新生儿脑部图像多功能区域分割框架,主要包括医学影像预处理、网络模型训练、模型预测及预测结果后处理等模块,可替代传统的手工勾画流程,得到新生儿脑部图像各区域分割模型,以协助医生或研究人员做进一步的新生儿脑部发育研究.

1 本文方法

图像分割算法在医学图像分析上具有重要意义,目前已从传统的基于阈值、基于区域的分割算法,发展到基于神经网络的卷积算法及 Transformer 算法,医学图像分割精度不断提高,特征表达能力不断增强.但对应用广泛的卷积神经网络,其自身不可避免地存在感受野低的问题,无法捕获远距离的特征,同时由于需要 Pooling 层进行下采样,导致丢失了图像的位置及细节信息,而这些问题 Transformer 网络能解决.由于 Transformer 网络本身来自于自然语言任务,虽然已经开发出了适应视觉任务的 ViT^[9]版本,但由于它的全局自注意力机制导致其学习能力较弱,需要大量数据训练,计算复杂度较高,因此不适合小数据量的医学数据集.本文使用移位窗口自注意力模块,其是由 Swin Transformer 网络提出,区别于传统的多头自注意力(MSA)模块,其由两个模块构成,分别为基于窗口的多头注意力(W-MSA)模块和基于移位窗口的多头自注意力(SW-MSA)模块.基于此模块本文提出基于移位窗口自注意力模块的新生儿脑部图像多功能区域分割网络,该网络以移位窗口自注意力模块作为主要模块,搭建出类似 Swin-Unet^[10]的网络结构,同时结合深度监督技术以及全尺度的跳跃连接技术优化网络模型,提高网络精度,增强其特征提取能力.此外,本文还提出一个端到端的脑部图像区域分割框架,包括新生儿脑部 MIR 图像预处理、训练、后处理等措施,进一步完成分割操作,同时在公开数据集 dHCP 上评估了其对新生儿脑部图像区域分割的性能,得到了优于 UNet^[11], TransUnet^[12]和 Swin-Unet 的结果.

1.1 框架设计

端到端的新生儿脑部 MRI 影像分割框架主要分为数据预处理,网络模型训练和分割结果后处理 3 个阶段,如图 1 所示.第一阶段为图像预处理,由于采集到的图像因采集机器、环境、个人技术等因素会使采集磁场发生变化,导致 MRI 图像采集强度不均匀,从而导致成像可能会出现忽明忽暗的现象,这种变化称为偏置场.这样的图像可能会影响网络性能,所以需要 N4 偏置场矫正技术纠正磁

场. 为增强 MRI 影像中各部分区域的差异化, 提高各部分区域的辨识度, 可使用平衡对比度增强技术 (BCET)^[13] 增强影像的对比度, 从而达到提高图像分割精度的目的. 第二阶段为模型训练部分, 本文基于移位窗口自注意力模块设计一个新的 Transformer 网络, 网络主体结构与 Swin-Unet 类似, 同时结合深度监督机制和全尺度跳跃连接技术提高网络模型能力. 第三阶段为分割结果后处理, 由于脑部图像区域存在相似性, 可能会导致错误分割出部分不正确区域的问题, 因此为提高模型精度, 在得到分割结果后还会经过后处理阶段, 通过最大连通域算法^[14], 将存在误差的区域去除, 以提高预测结果精度.

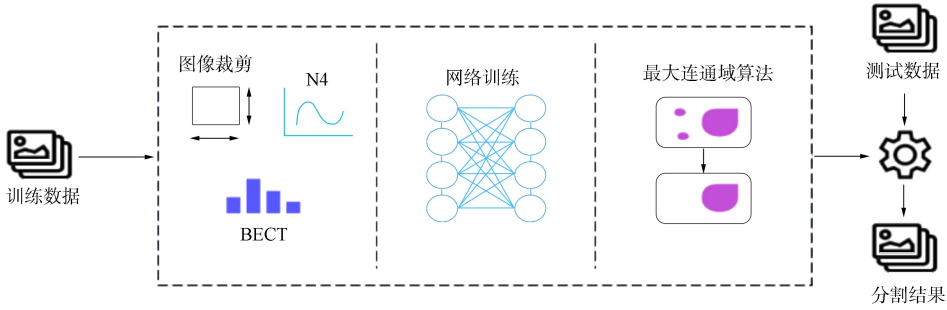


图 1 脑部 MRI 影像分割框架

Fig. 1 Segmentation framework of brain MRI images

1.2 图像预处理

1.2.1 图像裁剪

由于医学图像的特殊性, 需要针对其特性做一些预处理, 从而帮助模型更好地捕获图像特征, 本文预处理方法主要有图像裁剪、N4 偏置场校正^[15] 和平衡对比度增强.

图像裁剪是图像分割最常用的预处理措施, 其目的首先是为保证输入网络的图像符合网络要求, 其次是对医学图像这种感兴趣区域基本在相同位置的图像进行图像裁剪, 去除一些明显不感兴趣的区域, 从而增大感兴趣区域所占图像的百分数, 有利于网络捕获图像特征. 最后, 当训练数据匮乏时, 进行一些随机的裁剪可以增强数据, 有益于帮助网络收敛, 防止过拟合, 增强网络稳定性.

1.2.2 N4 偏置场矫正

N4 偏置场校正是提高 MRI 图像质量的重要方法, 主要用于避免环境或人为因素导致的 MRI 磁场分布不均匀现象, 如果不对这种现象做修正, 将会降低图像质量, 干扰网络特征捕获能力, 降低网络分割精度. N4 偏置场校正算法是专门针对该现象的图像校正技术, 是对 N3 校正技术的改进, 效果如图 2 所示.

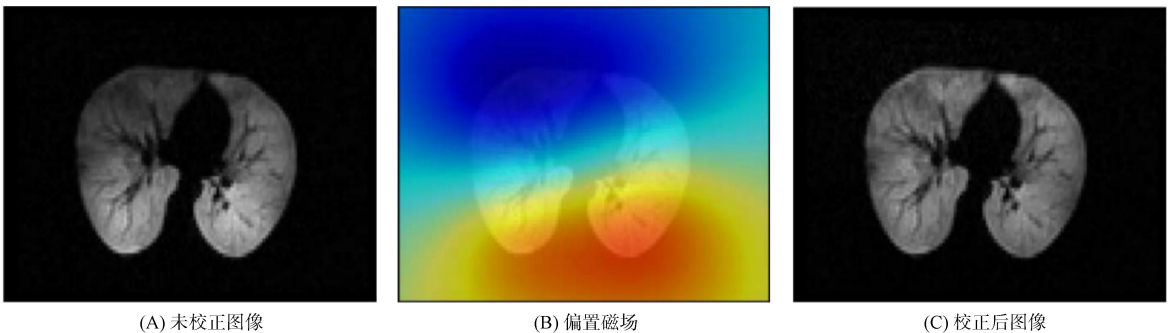


图 2 N4 偏置场校正效果

Fig. 2 Correction effect of N4 bias field

1.2.3 平衡对比度增强技术

平衡对比度增强技术是在不改变图像直方图分布的情况下提高图像的对比度, 从而使计算机更好地捕获到图像特征. 该技术有两种实现方式, 分别是基于二次函数的 BCETP 及基于三次函数的 BCETC, 本文主要使用二次函数对图像进行预处理, 公式如下:

$$y = a(x - b)^2 + c, \quad (1)$$

其中 y 为输出图像, x 为输入图像, 该二次函数通过 a, b, c 3 个系数控制. a 由 x 中的最大灰度值 h 、最小灰度值 l 及想要得到的 y 中的最大灰度值 H 、最小灰度值 L 决定, 公式如下:

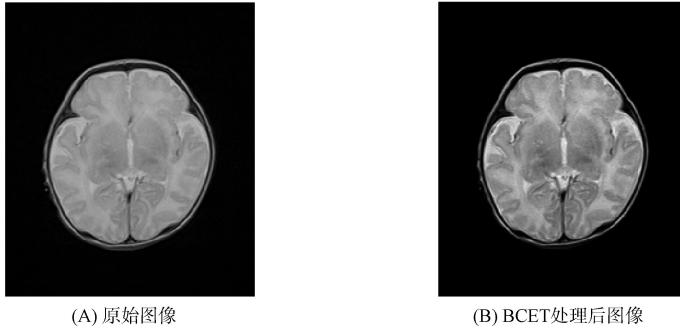
$$a = \frac{H - L}{(h - l)(h + l - 2b)}. \quad (2)$$

系数 b 和 c 为二次函数的转折点, 计算公式如下:

$$b = \frac{h^2(E - L) - s(H - L) + l^2(H - E)}{2[h(E - L) - e(H - L) + l(H - E)]}, \quad (3)$$

$$c = L - a(l - b)^2. \quad (4)$$

通过 BCET 技术对图像进行预处理, 提高图像对比度, 使图像获得更具表现力的细节, 效果如图 3 所示.



(A) 原始图像

(B) BCET处理后图像

图 3 平衡对比度增强技术效果

Fig. 3 Balance contrast enhancement technique effect

1.3 网络模型

1.3.1 模型结构

在网络结构部分, 本文针对新生儿脑部 MRI 影像特点设计一个全新的网络结构, 使用移位窗口自注意力模块作为主要模块, 搭建出类似 Swin-UNet 的网络结构. 通常增加神经网络的深度可在一定程度上提高网络的特征提取能力^[16], 所以为增强网络的特征学习能力, 本文将网络深度定为 4 层, 同时为防止由于网络深度过深而导致的梯度爆炸和梯度消失等问题, 本文还加入了全尺度的跳跃连接以及深度监督机制, 具体结构如图 4 所示, 其中: 紫色箭头即为深度监督机制, 从上采样的图像中选出 3 层, 通过 Patch Expanding 和 Linear Projection 层直接产生分割结果, 用于计算相应的损失; 黑色箭头为全尺度跳跃连接, 通过传入不同尺度的特征, 提高深层网络的感受野, 进一步优化网络.

网络仍采用经典的 U 型结构, 即编码器-解码器结构构成. 编码器部分的输入图像需先经过一个 Patch Partition 层, 该层主要将输入图像转换为 4×4 的 patches 分辨率图像, 每个 patch 视为一个 token, 其中 Token_feature 为 48, 经过本层处理即将图像分辨率编码为 $H/4 \times W/4 \times 48$, 再经过 Linear Embedding 将 Token_feature 转换为需要的维度 96, 至此得到需要输入网络的图像编码. 网络的主要特征编码模块由 4 层相同的移位窗口和 Patch Merging 模块构成, 移位窗口为网络的主要特征提取模块, Patch Merging 模块主要进行图像的下采样操作, 将图像最终编码为 $H/64 \times W/64 \times 16C$, 其中 C 为所需的通道数. 下采样后会经过一个瓶颈结构, 由 4 层移位窗口模块构成, 之后不再进一步下采样. 解码器部分与编码器部分基本相同, 使用 Patch Expanding 模块进行图像的上采样操作, 将图像逐步恢复为原来大小, 最后经过 Linear Projection 输出分割结果.

1.3.2 移位窗口模块

移位窗口模块由 Swin Transformer 提出, 主要在降低计算复杂度的同时提高窗口之间的相互通信能力, 从而解决了普通窗口自注意力模块限制了自注意力只能覆盖局部而忽略了全局信息, 导致精度降低的问题. 但移位窗口模块自注意力计算的过度不平滑, 所以本文在 Swin Transformer 的基础上改进图像特征捕获过程, 在 Swin Transformer 先经过 W-MSA 再经过 SW-MSA 特征捕获过程的基础

上增加两个移位窗口模块, 如图 5 所示, 其中: Layer 1 表示常规的窗口划分方式, 相当于移位窗口的初始状态, 类似于 ViT 中的划分, 将特征图划分大小相同的 4 个窗口, 在每个窗口内进行单独的自注意力计算; Layer 2 是将原始状态的窗口以步长为 2 向右滑动, 此时针对分割后的窗口分别计算即得到左右两个窗口间的通信信息; Layer 3 和 Layer 2 相似, 是将原始状态的窗口以步长为 2 向下滑动, 此时针对分割后的窗口分别计算即得到上下两个窗口间的通信信息; Layer 4 将所有窗口中的特征相融合, 窗口以步长为 2 向右下方移动, 分别计算各窗口. 从而使整幅图像的各位置信息都得到良好的通信.

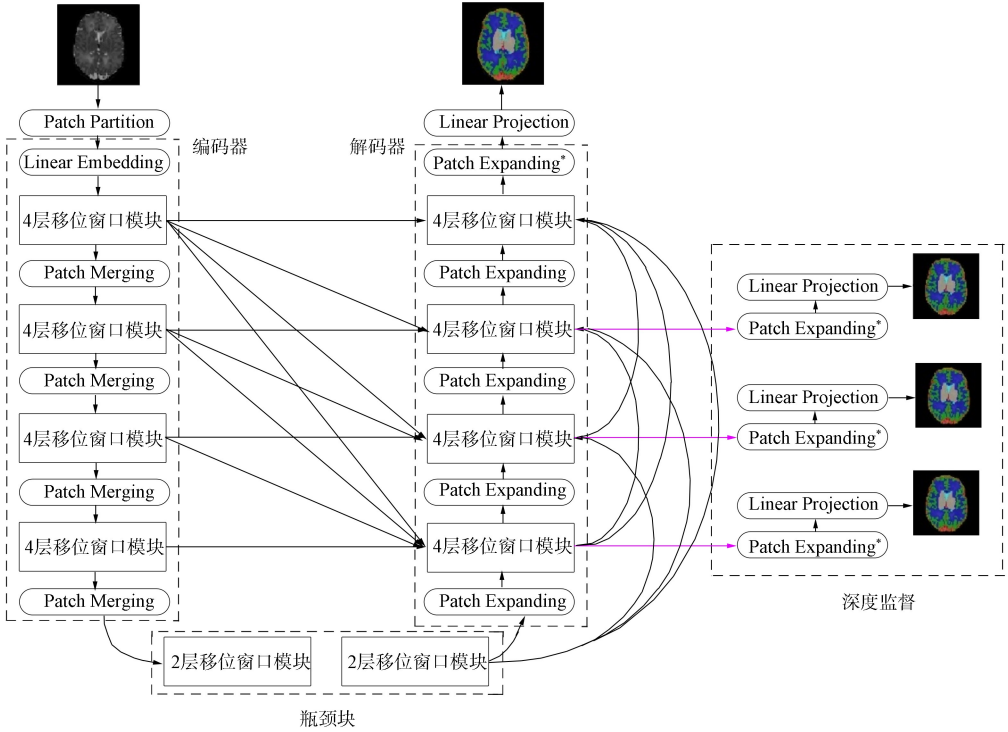


图 4 基于移位窗口模块的模型结构

Fig. 4 Model structure based on shifted window module

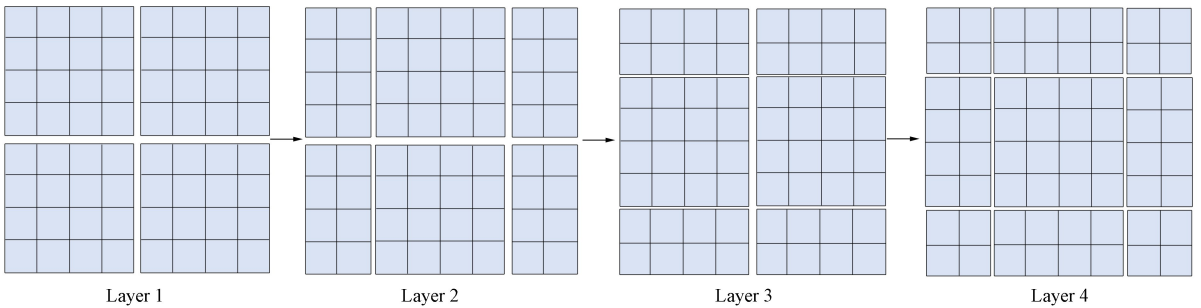


图 5 移位窗口模块 Layer

Fig. 5 Shifted window module Layer

1.3.3 深度监督模块

深度监督模块可提高网络中隐藏层学习过程的直接性和透明度^[17-18], 可在网络早期阶段提高特征学习的辨别力以及学习的稳健性, 同时面对网络中可能出现的梯度消失及梯度下降等问题有很好的抑制作用.

深度监督模块的核心思想是为隐藏层提供一个直接监督层, 用来引导网络的训练, 而不是仅监督最后的输出层, 这种监督同样通过反向传播达到优化网络的目的. 通常情况下, 使用更多的特征判别分类器将具有更高的性能, 因此, 可利用深层次隐藏层输出的特征图训练判别分类器, 以评估这些隐

藏层的质量,并不断优化,从而实现全局优化的目标,同时能有效提高神经网络的训练速度和准确性.对于小型训练数据和相对较浅的网络,针对分类准确性和学习特征问题,深度监督可以提供强大的纠正以及加速特征学习的能力,通过改善深层次有问题的收敛行为提升网络的性能.

2 实验设计

2.1 数据集

数据集 dHCP 提供了目前最新并且庞大多样的新生儿脑部区域 MRI 影像分割数据^[19],该数据集中的图像是从 24~45 周出生的婴儿脑部区域成像获得的.

2.2 评价指标

本文使用 Dice 系数和 95% Hausdorff 距离(95% Hausdorff distance, 95HD)^[20]两个指标从不同的角度评估模型的精度. Dice 系数是基于区域重合度进行评价,主要评估分割区域的相似程度,取值范围为 0~1,数值越大分割效果越好,表示其结果与目标区域相似度越高. 95HD 是基于两个点云之间距离进行评价,它将一个点云的每个点与另一个点云中的所有点计算距离,并选择距离最近那个点的距离值,然后取这些距离值中的最大值.在此基础上,选择 95%分位数获得一个较稳健的距离值.

Dice 系数是一种根据交集衡量集合相似度的度量函数,常被用于衡量图像分割算法中分割结果与真实标注之间的相似度,其取值范围为 0~1,计算公式为

$$s = 2|X \cap Y| / (|X| + |Y|), \quad (5)$$

其中: $|X \cap Y|$ 表示目标区域 X 和分割结果 Y 之间的交集; $|X|$ 和 $|Y|$ 分别表示其中非零元素个数,因为分母存在重复计算 X 和 Y 之间共同元素的行为,所以分子系数为 2. Dice 损失计算方式为 $1-s$,比较适合正负样本严重不平衡的场景,这种情形在医学图像中较常见,所以 Dice 损失在医学图像分割模型中经常使用,效果相比于其他的损失函数也相对较好.

95HD 是用于比较两组点集相似度的一种度量方式,其描述了两个点集之间距离的定义形式,即假设两个集合 $A = \{a_1, a_2, \dots, a_p\}$, $B = \{b_1, b_2, \dots, b_q\}$,则这两个点集之间的 Hausdorff 距离定义为

$$H(A, B) = \max\{h(A, B), h(B, A)\}, \quad (6)$$

$$h(A, B) = \max_{a \in A} \left\{ \min_{b \in B} \|a - b\| \right\}, \quad (7)$$

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

其中 $\|\cdot\|$ 是点集 A 和点集 B 之间的距离范式. 95HD 就是取两个点集之间最大距离的 95%的量化值,在模型中一般都使用 95%最大距离表示,这是为避免可能出现的少数极端异常值的影响,有利于提高系数的说服力.

2.3 环境及训练参数

为避免实验设备或环境导致的误差,本文对比的所有模型都在同一环境下完成.不同模型的训练参数也进行了统一:批处理大小为 24,优化器为 Adam,初始化学学习率为 1×10^{-3} ,输入图像大小为 $256 \times 256 \times 203$,最大迭代次数(epoch)为 50.

3 实验结果与分析

3.1 多功能区域分割结果

本文选取几种目前较流行的卷积神经网络以及 Transformer 结构网络在数据集 dHCP 上对 9 个不同的脑部区域图像分割结果进行定量比较,分别对分割结果的 Dice-Similarity coefficient (DSC)以及 95HD 系数做出比较,综合评估网络性能.实验结果列于表 1 和表 2.由表 1 可见,本文方法在多区域分割任务中的 DSC 和 95HD 评价指标都有提升,特别在 95HD 上提升更明显.由 DSC 系数可见,相比于卷积网络 Unet, Unet++ 和 Unet3P, TransUnet 和 Swin-Unet 这种 Transformer 结构网络在多区域分割上性能更优.本文网络相比于卷积神经网络中效果最好的 Unet++ 在 DSC 上提升 1.31 个百分点,在 HD95 上提升 2.86 mm,相比于 Transformer 网络中效果最好的 Swin-Unet 在 DSC 上提升

0.42 个百分点, 在 HD95 上提升 5.34 mm. 实验结果表明, 无论是在区域还是在边缘上, 本文提出的网络均能达到更好的效果.

表 1 多区域 DSC 和 95HD 结果

Table 1 Multi-region DSC and 95HD results

方法	评价指标		方法	评价指标	
	DSC/%	95HD/mm		DSC/%	95HD/mm
Unet	94.54	22.31	TransUnet	95.31	20.76
Unet++	95.22	18.06	Swin-Unet	96.11	20.54
Unet3P	94.63	33.88	本文	96.53	15.20

表 2 为对单个区域分别进行分析所得的结果. 表 2 中共划分 9 个区域, 分别为 Cerebrospinal fluid (CSF), Cortical Gray matter (cGM), White matter (WM), Background, Ventricle, Cerebellum, Deep Grey Matter (GM), Brainstem 和 Hippocampus. 由表 2 可见, 本文网络在大多数的区域中分割效果最好. 卷积神经网络相比于 Transformer 结构网络在一些小目标的分割上精度更准确, 而在一些大目标的分割效果不如 Transformer 结构网络, 这可能是其结构特性所致. 卷积神经网络获取全局信息能力较差, 感受野也层层递减, 导致其分割大目标时边界确定不好, 而 Transformer 结构网络更注重全局注意力信息的提取, 更容易确定大的边界信息, 如 Cortical Gray matter (cGM) 区域分割精度. 对于一些小目标, 由于 Transformer 结构网络的注意力机制过多地关注了全局信息, 而丢失了局部细节, 从而使一些小目标特征区域能力不如卷积神经网络, 如 Hippocampus 区域分割精度. 本文方法由于加入了全尺度跳跃连接机制以及深度监督机制, 其性能无论是在小目标还是大目标的分割精度上都较均衡, 在大目标的任务上精度仍优于卷积神经网络, 在小目标任务上其精度比 Swin-Unet 这类 Transformer 结构网络更优.

表 2 单区域 DSC 结果

Table 2 Single-region DSC results

方法	DSC								
	CSF	cGM	WM	Background	Ventricle	Cerebellum	GM	Brainstem	Hippocampus
Unet	94.58	95.87	96.50	90.46	93.77	97.98	97.46	97.12	91.63
Unet++	95.71	95.98	97.61	94.36	97.12	98.34	98.18	98.21	93.14
Unet3P	95.41	97.63	98.39	94.02	96.79	98.35	97.90	98.17	93.15
TransUnet	96.24	97.56	97.08	93.50	96.20	97.21	97.23	97.92	92.56
Swin-Unet	96.99	97.27	98.51	93.57	96.89	97.91	97.55	94.66	91.39
本文	96.56	97.88	97.34	94.91	96.55	98.50	98.34	96.21	92.78

3.2 消融实验

为确定本文方法的有效性, 本文进行一系列的消融实验从多方面验证各结构的有效性, 其中包括是否添加全尺度的跳跃连接、是否加入深度监督机制以及各种不同的移位窗口模块.

表 3 列出了针对是否添加全尺度跳跃连接模块以及深度监督机制所得的实验结果, 其中: SW 表示仅使用本文提出的移位窗口模块, 其他两者都未添加; SW + SC 表示仅添加全尺度跳跃连接; SW-DS 表示仅添加深度监督机制.

表 3 全尺度跳跃连接及深度监督机制消融实验结果

Table 3 Results of full scale skip-linking and depth supervision ablation experiments

方法	DSC								
	CSF	cGM	WM	Background	Ventricle	Cerebellum	GM	Brainstem	Hippocampus
SW	94.79	95.67	96.45	91.56	93.71	95.18	95.86	94.12	90.56
SW+SC	95.51	96.68	97.22	94.16	96.53	97.44	98.10	96.17	91.14
SW+DS	95.89	96.13	97.16	94.51	96.41	97.15	97.89	96.10	91.56
本文	96.56	97.88	97.67	94.91	97.24	98.50	98.34	96.66	92.78

由表 3 可见, 仅添加全尺度跳跃连接有效, 使模型可同时考虑不同层次的特征信息, 从而提高网

络的特征提取能力,优化模型效果.由仅添加深度监督机制的模型结果可见,深度监督机制单独使用的效果提升较小,主要解决深层次神经网络训练时可能出现的梯度消失及收敛速度过慢等问题,提高了网络效率.将全尺度跳跃连接和深度监督机制相结合可以使网络从全尺度的聚合特征图中学习层次信息,融合低层的特征信息和高层的特征信息,效果得到进一步提高.

表 4 列出了移位窗口模块的消融实验结果,本文针对 3 种不同的移位窗口模块对比做出分析,其中窗口表示仅使用图 4 所示的 $\text{Layer1} \times 4$ 进行实验, Swin-W 是 Swin Transformer 中提出的移位窗口,即图 4 中 $(\text{Layer1} + \text{Layer4}) \times 2$. 分别将这两种移位窗口模块与本文提出的移位窗口模块进行对比.由表 4 可见:使用窗口模块的网络精度较低,主要因为其丢失了部分全局注意力,将图片窗口化后,各窗口之间缺少联系; Swin-W 优于加入了 Layer4 方法,加强了各窗口之间的联系,弥补了窗口的缺点,效果较好.但 Swin-W 中各窗口之间的联系不太紧密,且窗口之间的注意力过度不平滑,所以本文加入了 Layer2 和 Layer3 使各窗口之间的注意力过渡更平滑,增强窗口之间的联系.实验结果表明该方法有效,可以提高网络精度.

表 4 移位窗口模块消融实验结果

Table 4 Results of shifted window module ablation experiments

%

方法	DSC								
	CSF	cGM	WM	Background	Ventricle	Cerebellum	GM	Brainstem	Hippocampus
窗口	91.23	93.16	94.55	90.64	92.37	93.33	93.86	92.16	89.37
Swin-W	95.68	96.22	96.95	95.05	97.24	97.86	97.89	96.45	92.22
本文	96.56	97.88	97.34	94.91	96.55	98.50	98.34	96.21	92.78

综上所述,本文提出了一个端到端的新生儿脑部图像多功能区域分割框架,该框架包含预处理、网络训练及后处理三部分.在预处理阶段分别使用了图像裁剪、 N_4 偏置场校正和平衡对比度增强技术(BCET)增强图像的可训练性,同时加速网络收敛提高后续步骤的可靠性.网络训练阶段使用 Transformer 结构的网络,以 Swin Transformer 模型为基础,创建一个全新的移位窗口模块,用 4 个层层递进的移位窗口平滑窗口之间的注意力过渡,增强各窗口之间的联系,同时加入多尺度跳跃连接和深度监督机制,帮助网络快速收敛,可在网络的每层都获取到全尺度的特征信息.后处理阶段使用最大连通域算法,帮助纠正一些分割错误的区域,进一步提高框架的分割精度.在数据集 dHCP 上实验评估了本文提出的框架,分析了新生儿脑部图像 9 个区域的分割结果在 DSC 和 95HD 两个评估指标上的精度,同时对比了 Unet, Unet++, Unet3P, TransUnet 和 Swin-Unet 5 个网络.实验结果表明,本文提出的网络得到了较好的分割结果.最后为证明网络中各结构的有效性,进行了消融实验评估对网络改进的效果,包括添加全尺度跳跃连接和深度监督机制的必要性以及不同的移位窗口模块的有效性.实验结果表明,添加全尺度跳跃连接和深度监督机制有效,同时本文提出的 4 层移位窗口模块优于原有的两层结构,实验结果更优.

参 考 文 献

- [1] ZHANG Y G, LIU J Y, YANG W H, et al. Image Super-resolution Based on Structure-Modulated Sparse Representation [J]. IEEE Transactions on Image Processing, 2015, 24(9): 2797-2810.
- [2] DUBOIS J, ALISON M, COUNSELL S J, et al. MRI of the Neonatal Brain: A Review of Methodological Challenges and Neuroscientific Advances [J]. Journal of Magnetic Resonance Imaging, 2021, 53(5): 1318-1343.
- [3] FRANKE K, LUDERS E, MAY A, et al. Brain Maturation: Predicting Individual Brain Age in Children and Adolescents Using Structural MRI [J]. Neuroimage, 2012, 63(3): 1305-1312.
- [4] COLE J H, LEECH R, SHARP D J, et al. Prediction of Brain Age Suggests Accelerated Atrophy after Traumatic Brain Injury [J]. Annals of Neurology, 2015, 77(4): 571-581.
- [5] ZHANG Y, CHENG J Z, XIANG L, et al. Dual-domain Cascaded Regression for Synthesizing 7T from 3T MRI [C]//International Conference on Medical Image Computing and Computer-Assisted Intervention. Berlin: Springer, 2018: 410-417.

- [6] SULTANA F, SUFIAN A, DUTTA P. Evolution of Image Segmentation Using Deep Convolutional Neural Network: A Survey [J]. Knowledge-Based Systems, 2020, 201: 106062-1-106062-25.
- [7] LIU Z, LIN Y T, CAO Y, et al. Swin Transformer: Hierarchical Vision Transformer Using Shifted Windows [C]//Proceedings of the IEEE/CVF International Conference on Computer Vision. Piscataway, NJ: IEEE, 2021: 10012-10022.
- [8] VASWANI A, SHAZEER N, PARMAR N, et al. Attention Is All You Need [C]//Proceedings of the 31st International Conference on Neural Information Processing Systems. New York: ACM, 2017: 6000-6010.
- [9] DOSOVITSKIY A, BEYER L, KOLESNIKOV A, et al. An Image Is Worth 16×16 Words: Transformers for Image Recognition at Scale [EB/OL]. (2020-10-22)[2023-03-10]. <https://arxiv.org/abs/2010.11929>.
- [10] CAO H, WANG Y Y, CHEN J, et al. Swin-Unet: Unet-Like Pure Transformer for Medical Image Segmentation [EB/OL]. (2021-03-12)[2023-04-01]. <https://arxiv.org/abs/2105.05537>.
- [11] RONNEBERGER O, FISCHER P, BROX T. U-net: Convolutional Networks for Biomedical Image Segmentation [C]//International Conference on Medical Image Computing and Computer-Assisted Intervention. Berlin: Springer, 2015: 234-241.
- [12] CHEN J N, LU Y Y, YU Q H, et al. Transunet: Transformers Make Strong Encoders for Medical Image Segmentation [EB/OL]. (2021-02-08)[2023-01-10]. <https://arxiv.org/abs/2102.04306>.
- [13] GUO L J. Balance Contrast Enhancement Technique and Its Application in Image Colour Composition [J]. Remote Sensing, 1991, 12(10): 2133-2151.
- [14] ZHENG X, LEI Q Y, YAO R, et al. Image Segmentation Based on Adaptive K-Means Algorithm [J]. EURASIP Journal on Image and Video Processing, 2018, 2018(1): 1-10.
- [15] TUSTISON N J, AVANTS B B, COOK P A, et al. N4ITK: Improved N3 Bias Correction [J]. IEEE Transactions on Medical Imaging, 2010, 29(6): 1310-1320.
- [16] SUN B Y, WEI X H, CUI J X, et al. Social Activity Matching with Graph Neural Network in Event-Based Social Networks [J]. International Journal of Machine Learning and Cybernetics, 2023, 14(6): 1989-2005.
- [17] LEE C Y, XIE S, GALLAGHER P, et al. Deeply-Supervised Nets [C]//Artificial Intelligence and Statistics. [S.l.]: PMLR, 2015: 562-570.
- [18] WANG L W, LEE C Y, TU Z W, et al. Training Deeper Convolutional Networks with Deep Supervision [EB/OL]. (2015-03-11)[2023-02-15]. <https://arxiv.org/abs/1505.02496>.
- [19] HUGHES E J, WINCHMAN T, PADORMO F, et al. A Dedicated Neonatal Brain Imaging System [J]. Magnetic Resonance in Medicine, 2017, 78(2): 794-804.
- [20] SETIAWAN A W. Image Segmentation Metrics in Skin Lesion: Accuracy, Sensitivity, Specificity, Dice Coefficient, Jaccard Index, and Matthews Correlation Coefficient [C]//2020 International Conference on Computer Engineering, Network, and Intelligent Multimedia (CENIM). Piscataway, NJ: IEEE, 2020: 97-102.

(责任编辑: 韩 啸)