1. **引言**

近年来，脑肿瘤的发病率呈上升趋势。在脑肿瘤中，胶质瘤是最常见的原发性恶性肿瘤。低级别胶质瘤预后较好，而高级别胶质瘤亲水性和侵袭性较强，患者生存期较短[1]。MRI是脑肿瘤分析、监测和手术治疗的主要诊断工具。多模态MRI图像可以为分析胶质瘤的不同区域提供补充信息。

在2017年和2018年MICCAI BraTS比赛中，基于卷积神经网络的分割方法取得了最先进的效果。超过80%的团队使用深度学习方法来分割脑肿瘤和病理组织，这表明这种方法受到了很多人的青睐。Havaei [2]提出了局部路径和全局路径的双路径架构，可以同时利用局部特征和更多全局上下文特征;Isensee等[3]提出了3D U-Net，改进了流行的U-Net架构，使用骰子损失函数处理职业不平衡;Pereira等[4]提出了一种深度小核图像补丁分类网络，从训练集中提取相同数量的各种图像补丁。

为了利用多模态信息，现有方法大多采用早期融合策略，将多模态图像拼接作为网络输入。然而，这种策略很难探索不同模态之间的非线性关系。为了缓解这一问题，最近的研究采用了一种层融合策略[5,6,7]，其中由不同编码器提取的模态特定特征在网络的中间层融合，并共享相同的解码器。MAML[6]通过不同的模态特定FCN嵌入多模态图像，然后应用模态感知模块回归注意图，以融合模态特定特征。然而，这些多模态融合方法没有建立模态内部和模态之间的远程空间依赖关系，无法充分利用不同模态之间的互补信息。

为了利用不同模态之间的互补信息，NestedFormer被提出。它是一种新的嵌套模式感知转换器，用于有效和鲁棒的多模态脑肿瘤分割。然而嵌套模式感知转换器对于图像输入后预处理不足，高层次特征融合不充分。针对上述问题，本文提出轴向门控注意力模块和修改的Inception模块。

1. **NestedFormer语义分割网络架构**

NestedFormer整体网络架构见图1。NestedFormer设计了一个有效的全局Poolformer，从不同的磁共振成像模式中来提取体积空间特征，更强调全局依赖性。为了更好地提取互补特征并实现任意数量的模态融合，我们提出了一种新的嵌套模态感知特征聚合(NMaFA)模块。它明确地考虑了单模态空间相关性和跨模态相关性，并利用嵌套转换来建立模态内和模态间的远程依赖关系，从而获得更有效的特征表示。

NestedFormer网络架构对于脑肿瘤分割有一下两个缺点：1）在图片输入后仅做简单的卷积预处理，缺少更进一步的处理，导致模型的精度降低。2）对多模态特征融合时过分关注模态间的特征融合，而忽略了单模态自身融合不足的问题。

针对上述问题，本文提出两个模块：轴向门控注意力模块(AGA)和修改的Inception模块(MI)来解决上述问题。



1. **轴向门控注意力模块**

改进后的NestedFormer模型整体框架如图2所示：图像输入后会先经过MI(Modified Inception)的处理再传入GPE（Global PoolFormer Encoder）下采样编码，在编码完成后将多个模态的特征拼接在一起，先进行NMaFA第一轮特征融合后，再由AGA进行最终的特征融合。



轴向门控注意力模块的整体结构见图3。输入图像通过调整通道排序，分别从图像的高度、宽度和深度去计算轴向门控注意力。轴向门控注意力的结构见图4。



轴向门控注意力中的轴向注意是为克服计算相关性的复杂性，将自注意分解为两个自注意模块。第一个模块在特征地图高度轴上进行自关注，第二个模块在宽度轴上进行操作，第三个模块在深度轴上进行操作。有效地模拟了原有的自注意机制，计算效率大大提高。在视觉模型中，位置信息通常用于捕获物体的结构，使相关性对位置信息敏感。这个偏置项被称为相对位置编码。这些位置编码是初始随机取值，可以通过训练学习的。提出的轴向注意能够以良好的计算效率计算非局部上下文，能够将位置偏差编码到机制中，并能够在输入特征映射中编码远程交互。然而，轴向注意更容易在键、值和查询上学习位置偏差。但是对于医学图像分割的小规模数据集，位置偏差很难学习，因此在编码远程交互时并不是准确的。在学习到的相对位置编码不够精确的情况下，将它们添加到各自的键、值和查询中会导致模型整体性能下降。所以引入门控单元，动态地调整相对位置编码，使得位置偏差更容易学习到。



1. **修改的Inception模块**

Inception是由谷歌提出，Inception模块的核心思想是同时使用多个不同大小的卷积核以及池化操作来提取特征，以此实现对图像特征的多尺度抓取。具体来说，一个典型的Inception模块包含以下几种类型的结构:1×1卷积层、3×3卷积层、5×5卷积层和一个最大池化层组成。所有这些结构的输出会在通道方向上被拼接(concatenated)起来，形成一个非常“宽”的特征图。这样的设计允许网络在不显著增加计算成本的情况下，探索多种不同的特征交叉组合，从而增加模型的表达能力。

本文重新设计的Inception模块如图5所示。相对于原本的Inception舍弃了5×5卷积和最大池化层，并且连续地2层去加强Inception的不同尺度特征提取能力，同时舍弃的5×5卷积和最大池化层也使重新设计的Inception模块的计算量远小于原本模块。



1. **实验验证**

本实验使用的数据集是来自BraTS2020数据集。BraTS2020训练数据集包含369个对齐的四模态MRI数据(即T1, T1Gd, T2, T2- flair)，。每种模态尺寸都是155×240×240。本实验会将这些图像统一裁剪边缘黑色无关区域，裁剪后为128×128×128大小。将数据集随机分为训练(315)、验证(17)和测试(37)。

我们的实验在NVIDIA GTX 3090 GPU上运行，使用Python3.6和PyTorch1.7.0深度学习框架。参数通过Xavier[8]初始化。损失函数是soft dice loss和cross-entropy loss 的组合，我们采用了权值衰减为10−5的Adam优化器[9]优化损失函数。学习率设定为10−4，batch size 为 2，动量（momentum）为 0.99，训练500个训练周期。

首先是消融实验，下表 1 所示的为本文所提出方法在BraTS2020数据集的表现。其中MI表示修改后的Inception模块，AGA表示轴向门控注意力模块。

表1 消融实验

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| 方法 | WT Dice(%) | TC Dice(%) | ET Dice(%) | Ave |
| NestedFormer | 91.1 | 86.4 | 80.5 | 86.0 |
| NestedFormer +MI | 91.1 | 86.2 | 81.1 | 86.1 |
|  |  |  |  |  |
| NestedFormer + AGA+ MI | 91.2 | 86.5 | 81.4 | 86.4 |

从表1中可以看出，MI在原来模型的基础上提升了0.1%，AGA+MI在原来模型的基础上提升了0.4%，验证了所提出的模块对于NestedFormer是有性能提升的。

对比实验将本文所提出的模型与其他模型在BraTS2020数据集上的效果进行比较。对比实验结果见表2。通过对比实验可以发现，改进后的模型在WT Dice、TC Dice、Ave中都优于其他模型，这证明了本文所提出的模型可以得到更好的脑肿瘤分割效果。

表2 对比实验

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| 方法 | WT Dice(%) | TC Dice(%) | ET Dice(%) | Ave |
| Res-Gated-3D UNet | 87.9 | 85.4 | 84.9 | 86.0 |
| NestedFormer | 91.1 | 86.4 | 80.5 | 86.0 |
| MVKS-Net | 89.5 | 83.1 | 78.2 | 83.6 |
| SwinBTS | 89.1 | 80.3 | 77.4 | 82.3 |
| Akbar et al | 89.3 | 80.2 | 72.9 | 80.8 |
| Our | **91.2** | **86.5** | 81.4 | **86.4** |

1. **结论**

本文提出了一种基于改进NestedFormer的脑肿瘤分割模型。针对NestedFormer模型图像输入后预处理不足和高层次特征融合不充分的问题。所提出轴向门控注意力模块和修改的Inception模块。针对NestedFormer模型图像输入后预处理不足的问题，提出修改的Inception，添加在输入图像进入网络主干中间。同时针对在网络高层次时特征融合不充分的问题，提出了轴向门控注意力模块，通过轴向分开计算门控注意力，提高对单模态的特征融合同时保持多模态之间的融合。实验结果表明，使用了轴向门控注意力和修改的Inception模块后的脑肿瘤分割网络的平均精度提升至了86.4。本文算法虽然在精度上取得了不错的结果，但分割精度相较于结构复杂的网络并没有明显提升，下一步需要针对损失函数进行修改，以进一步提升模型精度。

1. **摘要**

从磁共振三维图像(MRI)中准确分割脑肿瘤对临床决策和手术计划至关重要。人工分割虽然能达到较高的准确率，但费时费力，且不能批量处理。近年来，以U-Net为代表的深度神经网络已成功应用于医学图像分割，并取得了较好的效果。然而，没有一种标准的分割方法可以对所有的图像都产生理想的结果，各种分割方法都是针对特定的区域和成像方式进行优化的，具有很强的针对性。本文提出了一种基于改进NestedFormer的脑肿瘤，用于脑肿瘤分割模型。该模型在BraTS 2020训练数据集上进行了评估，达到了最先进的性能，整个肿瘤、肿瘤核心和增强肿瘤核心的Dice得分分别为91.2%,86.5%,81.4%。

关键词：脑肿瘤，分割，轴向，BraTS2020数据库

**参考文献**

[1]Bray, F., Ferlay, J., Soerjomataram, I., Siegel, R., Torre, L., Jemal, A.: Global cancer statistics 2018: Globocan estimates of incidence and mortality worldwide for 36 cancers in 185 countries. CA: A Cancer Journal for Clinicians 68(6), 394-424 (2018)

[2] Havaei M, Davy A, Warde-Farley D, et al. Brain tumor segmentation with deep neural networks [J]. Medical Image Analysis, 2017, 35: 18-31.

[3] Isensee F, Kickingereder P, Wick W, Bendszus M, Maier Hein K H: Brain Tumor Segmentation and Radiomics Survival Prediction: Contribution to the BRATS 2017 Challenge. In: Crimi A, Bakas S, Kuijf H, Menze B, and Reyes M.(eds.) Brainlesion: Glioma, Multiple Sclerosis, Stroke and Traumatic Brain Injuries. pp. 287–297. Springer International Publishing, Cham (2018).

[4]Pereira S, Pinto A, Alves V, et al. Brain tumor segmentation using convolutional neural networks in MRI images[J]. IEEE Transactions on Medical Imaging, 2016, 35(5): 1240-1251.

[5]Dolz, J., Gopinath, K., Yuan, J., Lombaert, H., Desrosiers, C., Ayed, I.B.:Hyperdense-net: A hyper-densely connected cnn for multi-modal image segmentation. IEEE Transactions on Medical Imaging 38(5), 1116{1126 (2019)

[6]Zhang, Y., Yang, J., Tian, J., Shi, Z., Zhong, C., Zhang, Y., He, Z.: Modality-aware mutual learning for multi-modal medical image segmentation. In: International Conference on Medical Image Computing and Computer-Assisted Intervention.pp. 589{599. Springer (2021)

[7]Zhou, T., Canu, S., Vera, P., Ruan, S.: 3d medical multi-modal segmentation network guided by multi-source correlation constraint. In: 25th International Conference on Pattern Recognition. pp. 10243{10250. IEEE (2020)

[8]Glorot, X., Bengio, Y.: Understanding the difficulty of training deep feedforward neural networks. In: Proceedings of the thirteenth international conference on artificial intelligence and statistics. pp. 249{256. JMLR Workshop and Conference Proceedings (2010)

[9]Loshchilov, I., Hutter, F.: Decoupled weight decay regularization. arXiv preprint arXiv:1711.05101 (2017)