

# 基于多尺度注意力的特征自适应 聚合脑肿瘤图像分割

许学添<sup>1</sup>, 李玲俐<sup>1</sup>, 蔡跃新<sup>2</sup>

(1. 广东司法警官职业学院 信息管理系, 广州 510520;  
2. 中山大学孙逸仙纪念医院 耳鼻喉科听力学与言语研究所, 广州 510120)

**摘要:** 针对脑部磁共振图像中脑卒中病灶的自动分割因分割目标边缘复杂、尺度变化多样而造成的识别精度不高的问题, 提出一种基于多尺度注意力的多尺度特征聚合方法, 该方法利用注意力机制调节中间特征不同通道的权重, 并自适应地选择不同尺度的特征进行融合, 在缺血性脑卒中的公开数据集 ATLAS 上进行的一系列实验, 选取 Dice 系数、豪斯多夫距离、重叠度、准确率和召回率作为评价指标, 结果表明所提出的模型在脑卒中病变的分割问题上取得了较好的分割效果; 另外, 该模型还在 Kaggle 公开的脑肿瘤数据集上完成对比实验, 证明该模型具有良好的可泛化性。

**关键词:** 图像分割; 注意力机制; 多尺度; 深度学习

## Brain Tumor Image Segmentation Using Feature Adaptive Aggregation Based on Multi-Scale Attention

XU Xuetian<sup>1</sup>, LI Lingli<sup>1</sup>, CAI Yuexin<sup>2</sup>

(1. Department of Information Administration, Guangdong Justice Police Vocational College, Guangzhou 510520, China;  
2. Institute of Hearing and Speech-Language Science, Department of Otolaryngology,  
Sun Yat-sen Memorial Hospital, Guangzhou 510120, China)

**Abstract:** Aimed at the low recognition accuracy for complex segmentation target edges and diverse scale changes in the automatic segmentation of stroke lesions in brain magnetic resonance images, a multi-scale feature aggregation method based on multi-scale attention is proposed. This method utilizes attention mechanism to adjust the weights of different channels for intermediate features and adaptively selects the features of different scales for fusion. A series of experiments are conducted on the public dataset ATLAS for ischemic stroke, and the Dice coefficient, Hausdorff distance, overlap, accuracy and recall are selected as evaluation indicators. The experimental results show that the proposed model achieves a better segmentation performance in stroke lesion segmentation. Additionally, the comparative experiments of the model are also conducted on the publicly available brain tumor dataset Kaggle, proving that it has a good generalization ability.

**Keywords:** image segmentation; attention mechanism; multi-scale; deep learning

### 0 引言

胶质瘤是一种严重威胁人类健康的恶性肿瘤, 其在所有脑和中枢神经系统肿瘤中的占比大约为 30%, 而在所有恶性脑肿瘤中的占比更是高达 80%。它在临床诊断和治疗中, 呈现出较大的复杂性和难度。为了更精确地观察和判断病灶部分, 医生通常会采用脑部磁共振成像 (MRI, magnetic resonance imaging)。MRI 能高分辨率地展示软组织的解剖结构, 通过多方位、多序列成像, 从各个角度对病灶

部分进行详细观察, 为临床诊断提供更全面、更丰富的影像信息。

然而, 在临床实践中, 医生在获取到患者的磁共振头颅图像后, 需要进行大量的图像处理工作, 包括对每一层切片进行精细的勾画。这一步骤不仅极度依赖医生的专业判断, 而且操作难度大、耗时长。长时间高强度的重复工作会使医生疲劳, 可能引发误判和误诊。因此, 研发能自动化进行精准分割的磁共振图像处理技术, 具有重大的现

收稿日期: 2023-06-16; 修回日期: 2023-07-18。

基金项目: 国家自然科学基金项目(82271165); 广东省普通高校特色创新项目(2020KTSCX273); 广东司法警官职业学院第五届院级课题(2023YB02)。

作者简介: 许学添(1984-), 男, 硕士, 副教授。

李玲俐(1977-), 女, 硕士, 教授。

蔡跃新(1984-), 男, 博士, 教授。

引用格式: 许学添, 李玲俐, 蔡跃新. 基于多尺度注意力的特征自适应聚合脑肿瘤图像分割[J]. 计算机测量与控制, 2023, 31(12): 224-230.

实意义和实用价值。

近期有很多基于卷积神经网络 (CNN) 的模型在脑卒中分割任务中表现出显著效果。Qi 等<sup>[1]</sup>设计了一种基于深度可分离卷积的 X-Net 来检测脑卒中病变, 并提出一个特征相似性模块来捕获远程依赖关系; Yang 等<sup>[2]</sup>设计了一个跨层融合的上下文推理网络 (CLCI-Net) 来处理形态差异大的脑卒中病灶, 并通过使用跨层融合和基于卷积的长短期记忆模块来解决特征相似性问题, Chen<sup>[3]</sup>由两个卷积神经网络 (CNN) 自动提取 MRI 中的脑卒中病变; Zhang<sup>[4]</sup>提出了一种使用 3D 全卷积和密集连接的卷积网络, 有效利用三维上、下文信息自动学习特征分割脑卒中病变。Ronneberger 等<sup>[5]</sup>提出了 U 型网络将底层的类别信息和高层的位置信息相结合可获得良好的分割结果。通过在编码阶段同时进行 2D 卷积和 3D 卷积, 一个维度融合的 U 型网络<sup>[6]</sup>被设计为一种既能整合医学图像 3D 信息、又能节省计算资源的新架构; Zhang 等<sup>[7]</sup>结合了多种输入, 把包括灰质、白质和侧脑室在内的结构信息作为先验知识辅助病灶分割, 在 2D 和 3D 不同维度的实验中均实现出色的分割性能; Dolz 等<sup>[8]</sup>在多模态融合的方式上尝试改进, 提出了基于密集多路径多模态的 U-Net 分割模型, 然而, 现有的一些方法虽然可以有效地分割出不同尺度的病灶, 但是这些方法没有充分利用中间特征或忽略了不同尺度特征对融合特征的影响, 导致其存在计算量较大以及特征表达不够准确的问题。因此, 如何获取具有丰富表达能力的特征, 以提高不同物体的分割精度, 是当前面临的一项重要挑战<sup>[9]</sup>。

基于以上的动机, 我们在本文中提出了一种基于注意力机制的多尺度特征聚合的脑卒中病变分割方法。通过使用注意力机制, 我们可以自适应地调整不同通道和不同尺度特征的权重, 以提高特征的表达能力, 实现更准确的图像分割。在脑卒中分割数据集 ATLAS<sup>[10]</sup>以及 Kaggle 上公开的脑肿瘤脑部磁共振数据集<sup>[11]</sup>上的实验结果证明了我们方法的优越性和有效性。

## 1 基于多尺度注意力的特征自适应聚合模型

### 1.1 模型总体结构

基于多尺度注意力的特征自适应聚合模型 (FAAMA, feature adaptive aggregation model based on multi-scale attention) 总体结构如图 1 所示, 其主要由特征提取模块  $N_{\text{feats}}$ 、支路模块  $N_{\text{td}}$ 、通道注意力模块  $N_{\text{chl}}$ 、多尺度特征注意力模块  $N_{\text{scale}}$  和解码器  $N_{\text{dec}}$  组成。该模型的主要思路是通过  $N_{\text{feats}}$  和  $N_{\text{td}}$  提取不同尺度中间特征, 再利用  $N_{\text{chl}}$  和  $N_{\text{scale}}$  增强并融合中间特征表达, 最后将融合特征输入到  $N_{\text{dec}}$  得到图像的分割结果。这个模型的理念是提取并利用不同尺度的中间特征, 然后增强并融合这些中间特征以提高表达能力, 最后把融合后的特征输入到解码器以得到图像的分割结果。

$N_{\text{feats}}$  用 ResNet-101 作为骨干网络, 主要用来逐层地提取输入 MRI 的特征。在原始网络中, 为了扩大感受野并获得更多的上下文信息, 我们采用多次降采样, 这导致最终输出特征的分辨率是原始图像的 1/32。然而, 这种操作会导致大量细节信息的丢失。因此, 我们改用了空洞卷积, 把 ResNet 中 stage3 和 stage4 模块的池化步长从 2 降低到 1, 并将普通的  $3 \times 3$  卷积替换为空洞率为 2 和 4 的空洞卷积, 这使得特征的分辨率保持不变。网络最后输出的特征分辨率是输入图像尺寸的 1/8。这是因为空洞卷积可以覆盖较大的特征范围, 所以随着网络深度的增加, 模型的感受野也会逐渐增大, 可以获取更多的上下文信息。

骨干网络由多个卷积阶段 (Stage) 组成, 每个阶段输出不同尺度的特征。 $N_{\text{chl}}$  模块以单个尺度特征  $f_i$  作为输入, 采用通道间的注意力机制调整  $f_i$  中每个通道的权重, 以增强特征中每个通道的表达能力。该模块将  $f_i$  划分为不同区域, 通过平均池化操作得到不同区域的平均值, 该值代表了所在区域的上下文信息。然后在通道维度采用全连接操作学习通道间的依赖关系, 计算不同通道的权重。最后利用该权重值对通道进行重新加权, 调整不同通道激活值的大

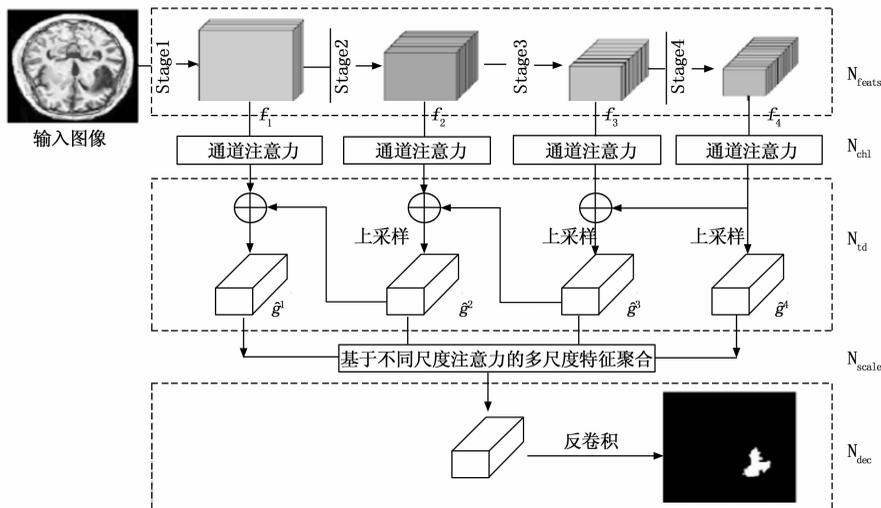


图 1 模型结构图

小, 得到新的特征  $\hat{f}_s$ 。

经过通道重加权后, 每个阶段的输出具有不同的上下文信息。低层的特征具有比较大的分辨率, 包含脑组织的轮廓等细节信息, 而高层特征的分辨率较低, 主要表示脑组织的语义信息。为了增强特征的表达能力,  $N_{id}$  模块采用自顶向下的方式将不同尺度的特征进行聚合。具体地, 对于经过通道注意力模块权重调整后的高层特征  $\hat{f}_s$ , 首先采用带有  $1 \times 1$  卷积的横向连接减小特征的通道数, 使其具有和上一阶段输出的特征  $\hat{f}_{s-1}$  相同的通道数量。然后将  $\hat{f}_s$  与  $\hat{f}_{s-1}$  相加, 对特征进行融合以获得同时具有高层语义信息和低层细节信息的中间层特征。经过多次操作,  $N_{id}$  不需要缩放输入图像就可以得到不同尺度的中间特征。

$N_{scale}$  的主要目的是对不同尺度的特征进行聚合。对于不同尺度特征中的每一个空间位置,  $N_{scale}$  首先计算每一个尺寸特征在该位置的权重, 其描述了该尺度特征在当前位置的重要程度。然后利用计算的权重对不同尺度的特征进行重加权, 并进一步聚合得到具有多尺度信息的最终特征。

最后,  $N_{dec}$  以最终聚合的特征作为输入得到脑部 MRI 的病灶分割结果。

### 1.2 通道重加权

$N_{feat}$  直接输出的中间特征并没有考虑通道之间的依赖关系。在 SENet<sup>[12]</sup> 中作者采用压缩和膨胀的方式计算池化后全局向量之间的依赖关系, 以增强特征的表达能力。这种方式可以获得图像全局的上下文信息, 满足分类任务的要求。但是在分割任务中, 脑部组织的病灶周围的信息对识别病灶具有重要作用。因此, 我们提出区域池化的策略, 将中间特征分成均匀的特征块, 然后对每个特征块进行池化操作获得代表该局部区域的特征向量, 其值为:

$$v_c = \frac{1}{w^r \cdot h^r} \sum_{i=1}^{w^r} \sum_{j=1}^{h^r} u_c(i, j) \quad (1)$$

其中:  $r$  为特征块所在区域,  $w^r$  和  $h^r$  分别为该区域的宽度和高度,  $v_c$  为池化后的特征向量在通道  $c$  的值,  $u_c(i, j)$  为区域中位置  $(i, j)$  的值。分块和池化的降采样操作如图 2 所示。

随后, 降采样后每一个位置的特征向量  $v$  输入到编码器—解码器网络, 学习向量中不同通道在特征表达中的重要程度。在编码器中, 本文采用  $1 \times 1$  的卷积减小向量的通道数, 对通道维度的特征进行压缩, 以挖掘通道之间的相互依赖关系。在解码器中, 同样采用  $1 \times 1$  的卷积对压缩后的通道进行扩展, 使扩展后的通道数与原通道数相同。其公式如下:

$$m_c = \sigma(W_2 \delta(W_1(v_c))) \quad (2)$$

其中:  $W_1$  和  $W_2$  分别为编码器和解码器中卷积层的可学习参数,  $\delta$  为 ReLU 函数。  $\sigma$  表示 Sigmoid 函数, 使通道  $c$  处的权重值位于  $[0, 1]$  区间。通过学习通道间的依赖关系, 对特征表达贡献较大的通道权重值增大, 而贡献较小的通道权重值则变小该权重对通道中的特征值进行重新调整, 避免了使用同一权重值造成的表达能力不足的问题。

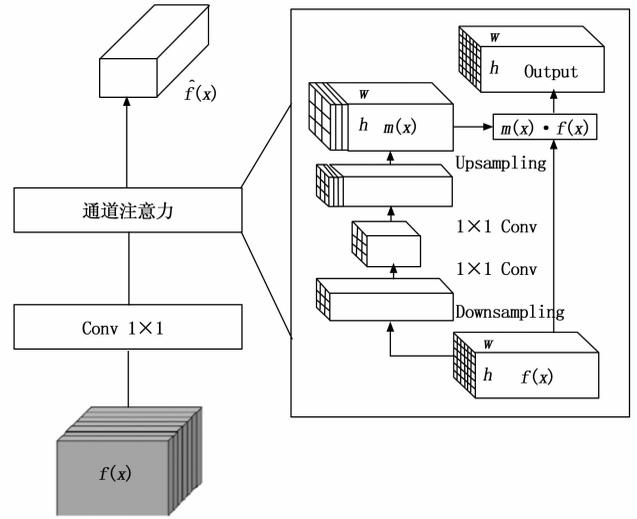


图 2 通道注意力操作示意图

经过重加权, 通道的值可以表示为:

$$\hat{f}_{i,j}^c(x) = m_c \cdot f_{i,j}(x) \quad (3)$$

这里  $\hat{f}_{i,j}^c$  为经过权重调整的通道  $c$  的值。这里没有使用残差的结构, 原因是残差连接会保留特征原有的分布, 减轻通道加权的影响。通道注意力机制如图 2 所示。

通道注意力机制利用区域池化获得像素所在区域的上下文信息, 并通过学习通道间的依赖关系对通道进行重加权, 增强通道中有效特征的表达能力, 减小和物体无关特征的影响。经过调整后的特征被送入到  $N_{id}$ , 以获得融合高层语义信息与低层细节信息的中间特征。

### 1.3 多尺度特征聚合

本小节采用尺度间的注意力机制对获得的不同尺度的中间层特征进行融合。我们将  $N_{id}$  自顶向下融合得到的中间层特征分别记作  $g^1, g^2, g^3, g^4$ , 这些特征包含了不同的上下文信息。对于较低层的中间特征如  $g^1$ , 其含有较多的组织轮廓等细节信息。而相对于较高层的特征如  $g^4$ , 其具有更大的感受野以及更多的脑组织语义信息。因此, 通过注意力机制自适应地选择不同尺度的特征, 可以获得更准确的脑组织信息表达。由于中间层特征具有不同的分辨率, 这里通过上采样操作使特征的分辨率相同, 如下式所示:

$$\hat{g}^s = F_{up}(g^s, g^1), \forall i, 1 \leq s \leq S \quad (4)$$

其中:  $S$  表示不同尺度的特征数量,  $F_{up}$  通过双线性插值的方式将中间特征  $g^s$  的分辨率上采样至与  $g^1$  相同的尺寸。通过以上操作, 低层、中层和高层的特征具有相同的分辨率。

接下来需要计算不同尺度特征在不同位置所对应的权重, 该权重表示特征所在尺度的重要程度。如图 3 所示, 我们首先将  $\hat{g}^1, \hat{g}^2, \dots, \hat{g}^S$  在通道维度聚合为一个向量, 得到包含有不同尺度信息的总特征。其次, 利用  $3 \times 3$  的卷积降低总特征维度, 融合并学习不同尺度特征之间的关系。最后, 采用  $1 \times 1$  的卷积输出通道数为  $S$  的特征  $e$ 。其过程可

以表示为:

$$e^s(i, j) = W_2 \cdot W_1 \cdot \hat{g}^s(i, j) \quad (5)$$

其中:  $W_1$  和  $W_2$  分别是  $3 \times 3$  和  $1 \times 1$  卷积层的可学习参数。

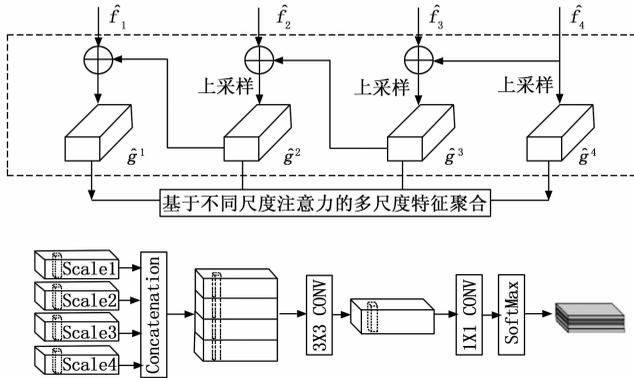


图 3 不同尺度注意力机制

本文采用  $\text{softmax}(\cdot)$  函数归一化不同尺度特征在位置  $(i, j)$  处的权重:

$$\alpha^s(i, j) = \frac{\exp[e^s(i, j)]}{\sum_{s=1}^S \exp[e^s(i, j)]} \quad (6)$$

其中:  $\alpha^s(i, j)$  表示尺度为  $s$  的特征在空间位置  $(i, j)$  的权重, 权重值越大, 表明该尺度特征所在位置对物体语义表达的重要性也越高。

在得到新的权重值  $\alpha^s(i, j)$  后, 其与不同尺度特征在对应位置相乘得到聚合特征, 即:

$$\bar{g}(i, j) = \sum_{s=1}^S \alpha^s(i, j) \cdot \hat{g}^s(i, j) \quad (7)$$

其中:  $\bar{g}(i, j)$  表示基于不同尺度注意力机制融合后的最终特征。该特征的分辨率与低层特征的分辨率相同, 为输入图像尺寸的  $1/4$ , 其送入到解码模块  $N_{\text{dec}}$  预测图像最终的分割结果。

#### 1.4 损失函数

在我们的基于注意力机制的多尺度特征聚合模型中, 选择了合适的损失函数是至关重要的。损失函数的设计会直接影响到模型学习的质量和效率, 它也是我们优化模型、调整模型参数的重要依据。在这个模型中, 我们采用了交叉熵损失作为网络的损失函数, 以实现端到端的训练, 自适应地学习不同通道和不同尺度的权重。交叉熵损失的选用是因为它适用于分类问题, 且具有非常好的性质。交叉熵损失可以衡量预测的概率分布与真实的概率分布之间的差距, 也就是说, 我们的模型会根据预测的概率分布与真实标签之间的交叉熵损失来进行学习和优化。此外, 交叉熵损失在分类问题中有一个明显的优势, 那就是它可以通过增大正确类别的预测概率, 同时减小错误类别的预测概率, 从而达到更好的分类效果。

然而, 我们的模型不仅仅关注于最终的分类结果, 也对中间的特征表示进行了关注。为了使每个中间特征都可以学到比较好的特征表示, 我们在模型中引入了附加的损

失对其进行监督训练。这种设计使得我们的模型在学习过程中, 不仅能关注到全局的最优, 也能关注到局部的优化, 从而获得更好的特征表示和学习效果。因此, 模型的总损失函数可以表示为如下形式:

$$Loss = -\frac{1}{HW} \sum_{j=1}^{HW} \sum_{t=1}^T \{y_j(t) \cdot \log[\hat{y}_j(t)] + \lambda \sum_{s=1}^S y_j(t) \cdot \log[\hat{y}_j^s(t)]\} \quad (8)$$

式中,  $H, W$  为预测图的高度和宽度,  $t$  为 MRI 中预测的目标类别, 类别总数是  $T$ , 由于本文仅关注从 MRI 中分割出病灶, 因此  $T=2$ , 表示附加损失的权重系数。

总的来说, 我们的损失函数设计能够充分考虑到不同尺度和通道的特性, 以及全局和局部的优化需求, 为模型提供了一个均衡且全面的学习目标, 从而有利于模型实现更好的训练和优化效果。

## 2 实验结果与分析

### 2.1 实验数据集

我们的模型在两个公开的数据集上进行了实验——卒中分割数据集 ATLAS 和 Kaggle 的脑肿瘤磁共振数据集。这两个数据集均由专业医疗影像所提供, 非常适合用于验证我们的模型的性能。

ATLAS 数据集是一组由专家收集的慢性卒中病例, 每个病例都包含 T1 加权磁共振扫描序列。这个数据集共有 229 个病例, 它们都被专业的医疗人员标记并手动分割, 确保数据的准确性。每个病例中都包含至少一个病变区域, 这对于我们的模型来说提供了极佳的训练材料。在这 229 个病例中, 每个三维序列都包含了 189 个切片, 每个切片的大小为  $233 \times 197$ 。因此, 总的来说, ATLAS 数据集由 43 281 个 2D 切片组成。为了对 ATLAS 数据集进行预处理, 我们采取了两步操作。首先, 我们把这 229 个案例进行打乱, 并随机拆分成 5 个互不相交的组别, 这样可以进行 5 折交叉验证, 提高模型的泛化性能和稳健性; 然后, 我们将所有图像裁剪为  $224 \times 192$  的大小, 使得它们能够适应我们的模型输入。

与 ATLAS 数据集相似, 我们还在 Kaggle 的脑肿瘤数据集上进行了测试。这个数据集包含了 138 张具有明显病灶的图像。尽管这些图像的疾病类型与 ATLAS 数据集的不同, 但是由于它们的成像方式和基本的脑部结构都相同, 因此可以作为我们模型泛化性能的测试材料。在处理这个数据集的时候, 我们首先将所有图像都缩放到  $224 \times 192$  的大小, 然后在这些图像中随机选择 109 张作为训练集, 余下的 29 张作为测试集。

### 2.2 实验设置

在实验设置阶段, 我们优选了随机梯度下降法 (SGD) 作为模型的训练算法。这种算法是由于其出色的优化性能和稳健的收敛性能而被广泛使用。我们的训练环境是两块性能强大的 2060Ti GPU, 这使我们能够有效地处理大规模的数据集, 同时也能在模型训练中保持较高的运行效

率。每次训练，我们会在每个 GPU 上处理 4 张图片，总共 8 张，这样的设置充分考虑了硬件资源的分配和模型训练的效率。

学习率策略是模型训练的重要组成部分，我们选取了“poly”策略。初始学习率设置为 0.01，相较于较大的学习率，这个设置可以使得模型训练过程更加稳定，避免过大的学习率导致的震荡问题。同时，动量和权值分别设置为 0.9 和 0.000 5，这是基于经验的设定，能够平衡训练速度和模型性能。

为了进一步提升模型的性能和泛化能力，我们对训练数据进行了一系列的数据增强操作，包括镜像旋转和尺寸缩放等。这种方式可以有效地扩展数据集，从而使模型有更多的学习机会。旋转的角度在  $[-10^\circ, 10^\circ]$  之间，尽可能地包含了可能出现的形态变化。尺寸缩放的倍数设置在  $[0.75, 1, 1.5, 2]$  之间，以便模型能够对不同尺寸的图片都有良好的处理能力。对于模型的输入，我们裁剪图像至  $513 \times 513$  的大小。这个尺寸的选择是基于我们模型设计的特性，以及对计算资源的考虑。这样的尺寸既可以包含足够的细节信息，又可以在计算复杂性和资源占用方面保持良好的平衡。在定义模型的损失函数时，我们对中间层特征损失设置了 0.4 的权重系数。这个设定有助于我们在训练过程中平衡各层特征的重要性，使得模型可以更好地学习和提取多尺度的特征，从而提高模型在复杂情况下的表现能力。

### 2.3 不同方法性能对比

对我们所设计的 FAAMA 模型与其他先进的图像分割算法进行比较的实验设置在 ATLAS 数据集上进行。我们选择了 8 种先进的流行方法：PSPNet<sup>[13]</sup>、Seg-Net<sup>[14]</sup>、DeepLabv3+<sup>[15]</sup>、U-Net<sup>[16]</sup>、ResUNet<sup>[17]</sup>、X-Net<sup>[1]</sup>、CLCI-Net<sup>[2]</sup>、MI-UNet<sup>[18]</sup> 进行对比。这些方法中，PSPNet、Seg-Net、DeepLabv3+、U-Net 和 ResUNet 是在各类图像分割任务上表现出色的算法，得到了广泛的应用和公认。而 X-Net、CLCI-Net 以及 MI-UNet 是专为解决卒中病变分割问题设计的模型，它们均在 ATLAS 数据集上进行过测试。我们使用这 8 种方法与我们的 FAAMA 模型进行对比，以凸显出 FAAMA 模型的有效性。

为了公正公平地比较各个方法的性能，我们采用了一系列标准化的评价指标，包括 Dice 系数、豪斯多夫距离、重叠度、准确率和召回率。通过这些定量的评价指标，我们对 FAAMA 模型的优越性进行了全面展示。具体的数据结果已在表 1 中进行了详细展示。

如表 1 所示，FAAMA 在 ATLAS 数据集上实现了整体最佳的脑卒中病变分割结果。FAAMA 所得到的 Dice 系数、重叠度、精度和召回率分别为 0.565 9, 0.433 0, 0.642 7 和 0.596 1，优于其他分割方法的结果，并且，经过计算，FAAMA 不仅在平均值上取得了最佳结果，还在 8 种方法中具有最小的标准差。FAAMA 所取得的 Dice 系数、重叠度、精度和召回率分别比基础框架 U-Net 高出 8.6%，

8.1%，4.0% 和 14.2%，豪斯多夫距离也小于 U-Net。CLCI-Net 和 X-Net 没有借助先验知识的帮助，但都实现了良好的分割效果，这可能是由于它们具有复杂的结构和众多参数。2D MI-UNet 结合了来自脑部其他组织信息的先验知识，虽然利用这些额外的信息有助于提高性能，但其在数据预处理步骤中使用 LDDMM 图像算法进行的脑部组织检测过程会引入额外的计算量并使模型的训练更加困难。我们的 FAAMA 模型取得了最优性能，表明采用不同通道与不同尺度注意力机制对多尺度特征进行聚合，可以有效提高特征的分割精度。

表 1 在 ATLAS 数据集上进行的模型量化对比结果

方法	Dice 系数	豪斯多夫距离	重叠度	准确率	召回率
PSPNet	0.358 3	65.68	0.250 8	0.377 3	0.423 8
SegNet	0.450 2	49.99	0.334 0	0.498 6	0.480 9
DeepLabv3+	0.497 5	48.54	0.375 8	0.559 1	0.530 3
UNet	0.521 3	48.01	0.400 3	0.618 9	0.520 5
ResUNet	0.534 8	47.96	0.416 1	0.572 7	0.563 3
CLCI-Net	0.543 9	46.92	0.416 7	0.622 6	0.570 0
X-Net	0.545 0	47.65	0.423 5	0.618 3	0.569 1
MI-UNet	0.550 3	46.44	0.425 3	0.622 9	0.571 4
FAAMA(ours)	0.565 9	44.28	0.433 0	0.642 7	0.596 1

我们还将我们的方法与当前的先进分割方法在 ATLAS 数据集上实现了可视化对比，图形化结果展示在图 4 中，其中图 (a) 列展示了原始输入图像，图 (b) 列标示了病灶真值区域，图 (c) ~ (f) 分别展示了 FAAMA、MI-UNet、X-Net、CLCI-Net 模型的分割结果。可以看出 FAAMA 捕捉到了更多病灶的细节信息，这在临床实际应用中是非常重要的。获取更多病灶的细节信息得益于 FAAMA 模型最终用于上采样的特征聚合了脑组织病灶不同上下文的信息，减轻了只使用空洞卷积造成的细节缺失问题，增强了病灶的特征表达。

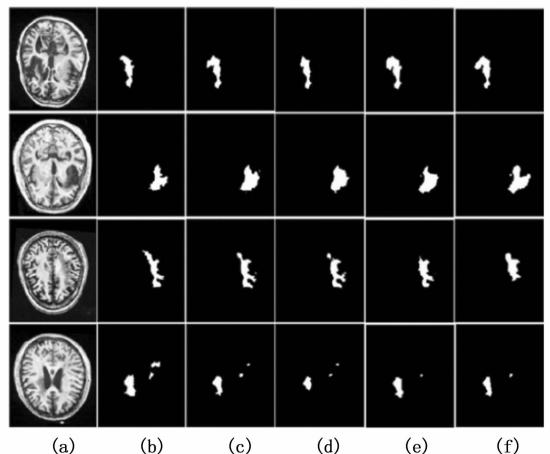


图 4 ATLAS 数据集上模型对比实验可视化结果

## 2.4 不同方法性能对比

我们在 ATLAS 数据集上对比了模型采用不同模块时的实验结果, 并对模块的性能进行分析, 实验结果见表 2。

表 2 模型消融实验结果

方法	Backbone	Top-down	C-attention	S-attention	Dice 系数	豪斯多夫距离
(a)	✓				0.521 3	48.30
(b)	✓	✓			0.534 1	47.35
(c)	✓	✓	✓		0.547 8	46.62
(d)	✓	✓		✓	0.543 7	46.35
(e)	✓	✓	✓	✓	0.565 9	44.28

表中“Top-Down”表示自顶向下的网络分支结构, “C-Attention”指通道间的注意力机制, “S-Attention”为不同尺度的注意力机制。实验 (a) 是以骨干网络为 ResNet101 的 U 型网络作为基准模型, 最终输出特征的分率为输入尺寸的 1/8。

实验 (b) 增加了自顶向下的特征聚合支路, 模型分割的 Dice 系数提高了 2.46%。这表明增大特征的分辨率并融合不同尺度的特征, 有利于恢复病灶的细节信息, 提高分割精度。

实验 (c) 和 (d) 分别增加通道间注意力机制以及不同尺度注意力机制模块, 使得模型分割精度相比实验 (b) 模型 Dice 系数提高了 2.57% 和 1.80%。结果表明调整不同通道的权重或自适应地聚合不同尺度的特征, 能够增强物体特征的表达力, 使模型的预测精度得到一定程度的提高。

实验 (e) 同时采用通道间注意力和不同尺度的注意力机制后, 模型的性能进一步增强, 相比单尺度基准模型 (c) 增加了近 3.30% Dice 系数。这表明不同的注意力机制之间不会相互干扰, 能够在不同的维度提高特征的表达力。

## 2.5 泛化实验

为了验证所提出的 FAAMA 的可拓展性, 我们在 Kaggle 脑肿瘤的磁共振数据集上进行了另一组对比实验, 实验结果展示在表 3 中。可以看出, FAAMA 与当前流行的先进模型相比仍然展现出优越的性能, 在全部 5 个指标中取得了最优的结果。

表 3 在 Kaggle 数据集上进行的模型量化对比结果

方法	Dice 系数	豪斯多夫距离	重叠度	准确率	召回率
PSPNet	0.525 4	59.67	0.360 2	0.595 6	0.564 9
SegNet	0.540 3	57.32	0.373 1	0.635 3	0.515 7
DeepLabv3+	0.572 1	57.12	0.404 0	0.641 4	0.521 5
UNet	0.590 2	53.44	0.421 8	0.670 4	0.535 6
ResUNet	0.598 3	54.75	0.428 9	0.661 0	0.558 6
CLCF-Net	0.653 2	56.13	0.490 3	0.744 7	0.579 6
X-Net	0.641 5	55.96	0.481 2	0.744 2	0.564 8
FAAMA(ours)	0.689 6	39.14	0.530 4	0.763 2	0.638 0

总的来说, 所设计的 FAAMA 模型不仅在 ATLAS 数

据集上展现出了优于其他先进方法的性能, 在另一数据集——Kaggle 脑肿瘤的磁共振数据集上, FAAMA 模型也证明了其优越的性能和良好的泛化能力。这些结果都为 FAAMA 模型在未来处理更多实际诊疗场景下的病灶分割问题提供了充分依据。

## 3 结束语

脑卒中在全球范围内都是一种常见的急性脑血管病变, 它的严重程度和影响使得精确快速的诊断和分割在临床上具有至关重要的意义。传统的医学图像分析往往需要医生手动进行, 这既消耗时间, 又受限于个体医生的经验和判断力。在这样的背景下, 我们提出了一种基于多尺度注意力机制的多尺度特征聚合模型, 以实现脑肿瘤病灶的自动化精准分割。

本研究中的 FAAMA 模型的关键在于其利用自顶向下融合的中间特征, 这种设计允许模型在不同深度层次获取并整合信息。进一步地, 我们采用不同通道和不同尺度的注意力机制来优化特征的选择和整合, 这使得我们的模型能够有效地学习通道间的依赖关系, 并自适应地选择和整合不同尺度的特征。这种设计让 FAAMA 模型在提取和了解图像特征时, 具有更高的灵活性和效率。此外, 我们的模型通过聚合包含物体高层语义信息和低层细节信息的特征, 显著提高了病灶分割的精度。高层次的语义信息为模型提供了对整个图像的全局理解, 包括图像的整体结构、目标的位置以及目标与背景的关系等, 而低层次的细节信息则使得模型能够更好地捕获和处理图像的微小变化, 从而提高模型的分割精度。为了验证我们模型的有效性和泛化能力, 我们在两个常用的公开数据集上进行了验证。在 ATLAS 数据集上, 我们的 FAAMA 模型在各项指标上均表现出色, 与多种先进的分割方法相比, 均显示出优越的性能。此外, 在 Kaggle 脑肿瘤磁共振数据集上, FAAMA 同样展现出了优秀的性能, 进一步证明了其在未来可以应用到更多实际诊疗场景下的病灶分割问题中。

## 参考文献:

- [1] QI K, YANG H, LI C, et al. X-net: Brain stroke lesion segmentation based on depthwise separable convolution and long-range dependencies [C] //International Conference on Medical Image Computing and Computer-assisted Intervention, 2019: 247 - 255.
- [2] YANG H, HUANG W, QI K, et al. CLCF-Net: Cross-level fusion and context inference networks for lesion segmentation of chronic stroke [C] //International Conference on Medical Image Computing and Computer-assisted Intervention, 2019: 266 - 274.
- [3] CHEN L, BENTLEY P, RUECKERT D. Fully automatic acute ischemic lesion segmentation in DWI using convolutional neural networks [J]. Neuroimage: Clinical, 2017, 15: 633 - 643.
- [4] ZHANG R, ZHAO L, LOU W, et al. Automatic segmentation

- of acute ischemic stroke from DWI using 3D fully convolutional DenseNets [J]. IEEE Transactions on Medical Imaging, 2018; 2149 - 2160.
- [5] 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, 2015; 234 - 241.
- [6] ZHOU Y, HUANG W, DONG P, et al. D-unet: a dimension-fusion u shape network for chronic stroke lesion segmentation [C] // IEEE/ACM Transactions on Computational Biology and Bioinformatics, 2019, 18 (3): 940 - 950.
- [7] ZHANG Y, WU J, LIU Y, et al. Mi-unet: multi-inputs unet incorporating brain parcellation for stroke lesion segmentation from t1-weighted magnetic resonance images [J]. IEEE Journal of Biomedical and Health Informatics, 2020, 25 (2): 526 - 535.
- [8] DOLZ J, AYED I B, DESROSIERS C. Dense multi-path U-Net for ischemic stroke lesion segmentation in multiple image modalities [C] // International MICCAI Brainlesion Workshop, 2018; 271 - 282.
- [9] WANG S, CHEN Z, YOU S, et al. Brain stroke lesion segmentation using consistent perceptio generative adversarial network [J]. Neural Computing and Applications, 2022; 1 - 13.
- [10] LIEW S L, ANLIN J M, BANKS N W, et al. A large, open source dataset of stroke anatomical brain images and manual lesion segmentations [J]. Scientific Data, 2018, 5 (1): 1 - 11.
- [11] CHAKRABARTY N. Brain MRI images for brain tumor detection [EB/OL]. (2023 - 07 - 21) [2021 - 06 - 16]. <https://www.kaggle.com/navoneel/brain-mri-images-for-brain-tumor-detection>.
- [12] HU J, LI S, GANG S. Squeeze-and-excitation networks [C] // IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2018; 7132 - 7141.
- [13] BADRINARAYANAN V, KENDALL A, CIPOLLA R. Segnet: A deep convolutional encoder-decoder architecture for image segmentation [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2017, 39 (12): 2481 - 2495.
- [14] ZHAO H, SHI J, QI X, et al. Pyramid scene parsing network [C]. IEEE Conference on Computer Vision and Pattern Recognition, 2017; 2881 - 2890.
- [15] CHEN L C, PAPANDREOU G, KOKKINOS I, et al. DeepLab: Semantic image segmentation with deep convolutional nets, atrous convolution, and fully connected CRFs [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2017, 40 (4): 834 - 848.
- [16] 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, 2015; 234 - 241.
- [17] ZHANG Z, LIU Q, WANG Y. Road extraction by deep residual u-net [J]. IEEE Geoscience and Remote Sensing Letters, 2018, 15 (5): 749 - 753.
- [18] ZHANG Y, WU J, LIU Y, et al. Mi-unet: multi-inputs unet incorporating brain parcellation for stroke lesion segmentation from t1-weighted magnetic resonance images [J]. IEEE Journal of Biomedical and Health Informatics, 2020, 25 (2): 526 - 535.
- (上接第 179 页)
- [4] 赵莉华, 张亚超, 金 阳, 等. 基于 LabVIEW 和 Matlab 虚拟实验室的实现 [J]. 实验室研究与探索, 2014, 33 (4): 62 - 64, 67.
- [5] 裴 锋, 杨万生. LabVIEW 与 Matlab 混合编程 [J]. 电子技术应用, 2004, 30 (3): 4 - 6.
- [6] 艾迪昊, 李 聪, 荣 涛, 等. 基于 Labview 和 Matlab 的改进神经网络煤与瓦斯突出预测系统 [J]. 中国煤炭, 2014, 40 (7): 108 - 112.
- [7] 钟 崑, 郑立军, 俞自涛, 等. 基于“数字孪生”的智慧供热技术路线 [J]. 华电技术, 2020, 42 (11): 1 - 5.
- [8] 陶 飞, 戚庆林, 王力举, 等. 数字孪生与信息物理系统——比较与联系 [J]. Engineering, 2019, 5 (4): 132 - 149.
- [9] 晓 娜, 孔祥伟. 一种基于 LabVIEW 与 MATLAB 的混合编程的试飞监测系统的设计 [J]. 中国科技信息, 2019, (16): 44 - 45.
- [10] FARAMARZI A, HEIDARINEIAD M, MIRJALILI S, et al. Marine predator algorithm: a nature-inspired metaheuristic [J]. Expert Systems with Applications, 2020, 152: 113377.
- [11] 凌 云, 张志涌. Matlab 面向对象和 C/C++ 编程 [M]. 北京: 北京航空航天大学出版社, 2018.
- [12] 宋 铭. LABVIEW 编程详解 [M]. 北京: 电子工业出版社, 2017.
- [13] 任正雄, 韩 华, 崔晓钰, 等. LabVIEW 与 Matlab 联合编程在制冷系统故障诊断中的应用 [J]. 能源研究与信息, 2020, 7 (4): 193 - 200.
- [14] 吴小红. LabVIEW 中调用 Matlab 进行信号处理研究 [J]. 微型机与应用, 2014, 33 (10): 66 - 68.
- [15] 曲丽荣, 胡 荣, 范寿康. LABVIEW、Matlab 及其混合编程技术 [M]. 北京: 北京航空航天大学出版社, 2018.
- [16] BOOKER L B, GOLDBERG D E, HOLLAND J H. Classifier systems and genetic algorithm [J]. Artificial Intelligence, 1989, 40 (1): 235 - 242.
- [17] 高洪涛, 陈晓宁. 基于 LabVIEW 与 MATLAB 平台的神经网络设计与应用研究 [J]. 自动化技术与应用, 2012 (2): 51 - 54.
- [18] 史 峰, 王小川, 郁 磊, 等. Matlab 神经网络 30 个案例分析 [M]. 北京: 北京航空航天大学出版社, 2010.
- [19] 张德丰. MATLAB 神经网络应用设计 [M]. 北京: 机械工业出版社, 2011.
- [20] 杨 瑞. 智能化、数字化、网络化供热系统的构建与实施 [J]. 华电技术, 2020, 42 (6): 83 - 86.