

Axial-FNet: 基于模糊卷积结合门控轴向自注意力的皮肤癌图像分割模型

姜舒¹, 陈琨², 丁卫平¹, 周天奕¹, 朱越²

(1. 南通大学人工智能与计算机学院, 江苏 南通 226019;

2. 南通大学张謇学院, 江苏 南通 226019)

摘要: 皮肤癌图像分割是医学图像处理领域中的一项关键任务, 现常用的分割算法在进行诊断时, 无法很好地平衡局部细节信息和全局上下文信息的计算资源请求量。此外, 肿瘤边界模糊且难以正确识别分割的问题也亟待解决。针对上述问题, 提出了模糊卷积结合门控轴向自注意力的皮肤癌图像分割模型 Axial-FNet, 该模型由门控轴向自注意力分支和模糊卷积神经网络分支构成。在门控轴向自注意力分支的尾部, 设置了门控权重控制器, 以控制捕捉局部细节信息和全局上下文信息的比例和程度; 融合模糊学习模块至卷积神经网络 (CNN) 中形成模糊神经网络分支, 提取图像局部信息。该模型在降低计算量的同时提升了分割精确度。在 ISIC 2017 数据集和 ISIC 2018 数据集上的平均交并比 (MIoU)、F1 分数 (F1-score) 和准确率 (Accuracy) 分别达到了 74.23%、83.05% 和 92.89%, 80.91%、88.13% 和 93.10%。实验结果表明, Axial-FNet 较其他多个先进分割模型而言, 分割的精确度和可靠性更佳。

关键词: 图像分割; 模糊学习; 门控轴向自注意力; 门控机制; 卷积神经网络

中图分类号: TP391

文献标志码: A

doi: 10.11959/j.issn.2096-6652.202520

Axial-FNet: skin cancer image segmentation model based on fuzzy convolution combined with gated axial self-attention

JIANG Shu¹, CHEN Kun², DING Weiping¹, ZHOU Tianyi¹, ZHU Yue²

1. School of Artificial Intelligence and Computer Science, Nantong University, Nantong 226019, China

2. School of Zhangjian, Nantong University, Nantong 226019, China

Abstract: The task of skin cancer image segmentation is a key task in the field of medical image processing. The commonly used segmentation algorithms can't well balance the computational resource requirements of local information and global context information when performing diagnosis. In addition, the problem of fuzzy tumor boundaries and difficulty in correctly identifying segmentation is also an urgent problem to be solved. Aiming at the above problems, a skin cancer image segmentation model Axial-FNet based on fuzzy convolution combined with gated axial self-attention was proposed. The model was composed of a gated axial self-attention branch and a fuzzy convolutional neural network branch.

收稿日期: 2025-02-12; 修回日期: 2025-04-21

通信作者: 丁卫平, dwp9988@163.com

基金项目: 国家重点研发计划项目 (No.2024YFE0202700); 国家自然科学基金项目 (No.62406153); 国家级大学生创新创业训练计划项目 (No.202410304067Z); 江苏省自然科学基金项目 (No.20231337); 江苏省高等学校自然科学研究面上项目 (No.23KJB520031, No.24KJB520032)

Foundation Items: The National Key R&D Plan of China (No.2024YFE0202700), The National Natural Science Foundation of China (No. 62406153), The National College Students' Innovation and Entrepreneurship Training Program Project (No. 202410304067Z), The Natural Science Foundation of Jiangsu Province (No.BK20231337), The General Program of the Natural Science Research of Higher Education of Jiangsu Province (No.23KJB520031, No.24KJB520032)

At the end of the gated axial self-attention branch, a gated weight controller was set to control the proportion and degree of capturing local information and global context information. The fuzzy learning module was fused into the convolutional neural network (CNN) to form a fuzzy neural network branch to extract the local information of the image. The segmentation accuracy was improved by the model while reducing the amount of calculation. The performance of the Axial-FNet model was evaluated on the ISIC 2017 dataset, achieving scores of 74.23%, 83.05%, and 92.89% for MIoU, F1-score, and accuracy, respectively, as well as 80.91%, 88.13%, and 93.10% for the same metrics on the ISIC 2018 dataset. The experimental results show that Axial-FNet has better segmentation accuracy and reliability than other advanced segmentation models.

Key words: image segmentation, fuzzy learning, gated axial self-attention, gating mechanism, convolutional neural network

0 引言

皮肤癌作为一类影响皮肤细胞的恶性肿瘤，主要包括基底细胞癌、鳞状细胞癌和恶性黑色素瘤等多种亚型，是全球范围内较为常见的癌症类型之一^[1]。其早期诊断与精准治疗对患者的生存率和生活质量至关重要。在医学图像分析领域，皮肤癌图像分割任务旨在从皮肤影像中自动提取癌变区域。首先，准确的皮肤癌分割能够极大地辅助临床医生进行诊断和治疗规划。通过精确地识别和分离癌变组织，医生可以更好地了解病灶的大小、位置及其扩散情况，从而为制定手术切除范围、放疗计划和药物治疗方案提供重要依据。其次，皮肤癌的形态特征，如边缘的不规则性、颜色变化、肿瘤体积和深度等，与患者的预后密切相关^[2]。自动化的图像分割工具能够提供更为量化和客观的癌症数据，有助于医生在治疗前后进行疗效评估，预测患者的生存期和生活质量。传统的皮肤癌诊断方法通常依赖医生的肉眼观察和手动标注，这不仅耗时费力，而且由于人为因素的影响，结果可能具有一定的主观性和不一致性^[3]。随着深度学习和计算机视觉技术的快速发展，基于自动化算法的皮肤癌分割方法逐渐成为研究热点。这些方法不仅能显著提高分割的效率和准确性，还能在大规模数据集上实现更为一致和可靠的分割结果，从而减轻医生的工作负担，并推动皮肤癌诊断与治疗向更加个性化和精准的方向发展。

卷积神经网络 (convolutional neural network, CNN) 广泛应用在皮肤癌图像分割任务中，其重要性已毋庸置疑。自 Olaf Ronneberger 提出 U-Net 架构^[4]以来，CNN 便在肿瘤图像分割领域得到了广泛应用，并迅速成为该领域的主流方法。由于其卓

越的性能，随后出现了众多基于 U-Net 的改进型 CNN，如 U-Net++^[5]、Res-UNet^[6]、TernausNet^[7]等。这些网络凭借强大的特征提取能力和对图像空间结构的深刻理解，已成为该领域的标准工具。然而，CNN 的归纳偏置，特别是其局部性和权重共享的特性，也带来了一些局限性，尤其是在处理远程依赖关系时。

在癌症图像分割中，远程依赖关系对于准确分割至关重要，而由于卷积操作的局部性，CNN 往往难以直接捕捉到这种远程依赖关系。而自注意力机制^[8]能够有效捕捉远程依赖关系，提升模型在处理长序列数据时的效果。因此，许多研究者开始将自注意力机制引入 CNN 中，以实现局部重要特征捕获和全局感知相结合的效果。Chen 等^[9]提出的 TransUNet 通过 CNN 提取关键特征，并利用多头自注意力机制^[10]学习全局信息；Chen 等^[11]提出的 SCUNet++ 则在编码器和解码器中引入 Swin Transformer^[12]以充分捕捉全局特征，同时通过 CNN 补充学习局部特征。然而，由于医学图像样本量少，加上自注意力机制参数数量庞大，其性能难以达到理想水平。此外，自注意力机制对计算资源的需求也非常高。为了解决这些问题，本文提出了 Axial-FNet，将门控轴向自注意力^[13]和 CNN 融合互补。

在皮肤癌图像分割任务中，肿瘤本身存在的一些问题也会直接影响图像分割的准确性和效率。图像在采集过程中，受采集设备、患者运动、扫描条件等多种因素的影响，可能会产生伪影，伪影会干扰图像的视觉效果，增加图像分割的难度^[14]。此外，病灶与正常皮肤的边界模糊，也会导致图像中的结构和细节难以辨识^[15]。而模糊学习^[16]是一种处理不确定性信息的学习算法，它通过建立模糊数

学模型来描述分割任务中的模糊边界、不清晰细节等不确定性信息,进而实现更好的分割效果。一些研究者试图引入模糊学习来处理图像中的不确定性信息以求达到更好的分割效果。例如,Deng等^[17]构建了一个基于层次化深度神经网络的模糊深度学习系统,该系统能够将模糊学习和神经网络所学习到的知识进行融合;Ding等^[18]提出的FTransCNN使用模糊积分来抑制融合特征中的异质性和不确定性,并使用模糊注意力融合模块处理分割边界的不确定性。而本文设计了模糊学习模块并将其引入CNN中,该CNN最终成为模糊卷积神经网络分支,减少了皮肤癌图像中的不确定性信息,提升了皮肤癌图像分割任务中的分割效果。

基于此,本文提出了门控轴向自注意力机制与模糊卷积神经网络相结合的Axial-FNet分割模型。本文主要贡献如下。

(1) 提出了一种新的网络结构,由模糊卷积结合门控轴向自注意力构成的皮肤癌图像分割模型,解决了传统CNN与自注意力结合时计算量过大的问题,同时在捕捉局部细节信息和全局上下文信息方面表现优异。

(2) 针对皮肤癌图像分割任务中的病灶边界模糊和细节不清晰等问题,本文创新性地将在模糊学习模块融入CNN中,在一定程度上抑制了上述不确定性,提升了分割的准确性。

(3) 在多个皮肤癌图像分割数据集上进行了大量的实验,实验结果表明本文方法较其他多个先进模型而言,分割的精确度和可靠性更佳,具有实际应用潜力,可以为皮肤癌图像分割领域提供一种新方案。

1 相关工作

1.1 基于卷积神经网络(CNN)的医学图像分割模型

医学图像分割是医学图像处理领域中的一项重要任务,旨在精确分割出医学图像中的目标点。CNN在医学图像分割领域有广泛的应用,通过卷积操作从原始图像中逐层提取特征,形成层次化的特征表示。这种局部感知和权值共享的特性使CNN在分割任务上表现出色。Olaf Ronneberger提出的U-Net^[4]由于其独特的U形架构和跳跃连接机制,使得其在医学图像分割任务中提供比较精确且详细的分割结果。由于U-Net的U形架构在医学图

像分割领域的出色表现,U-Net自提出后便成为CNN在医学图像分割任务中的基准模型。许多基于U-Net改进的算法也被相继提出,例如,Zhou等^[5]提出了UNet++,Ibtehaz等^[19]提出了MultiResUNet,Iglovikov等^[7]提出了TernausNet,Yan等^[20]提出了After-Unet,Ruan等^[21]提出了EGE-UNet。这些改进算法在各自的医学图像分割任务中均展现出了更优异的性能。

1.2 结合自注意力机制的医学图像分割模型

自注意力(self-attention, SA)机制在医学图像分割中也发挥着重要作用,它允许模型在处理图像数据时聚焦于图像中的关键区域,并捕捉这些区域之间的复杂依赖关系,从而提高模型对图像整体结构的理解能力和分割精确度。Vision Transformer^[22]的提出,首次把以SA为绝对核心的Transformer架构运用在了图像处理领域并取得良好成效。随后,Zheng等^[23]提出了SETR,它是基于Vision Transformer用于实现图像分割的第一个代表模型,以纯Transformer架构为编码器代替CNN编码。接下来,越来越多的研究者将SA运用到了医学图像分割任务中,如SegFormer^[24]、Swin-UNet^[25]和TransResNet^[26]。尽管SA在全局上下文建模方面表现出色,但它在捕获图像的细粒度细节方面却存在一定的局限性,医学图像通常包含丰富的细节信息,这些细节对于精确分割至关重要。

CNN以其局部感受野和权值共享的特性,在捕获图像的局部细节和纹理特征方面表现出色,SA则能够捕获全局上下文信息,两者结合,模型能够同时利用局部细节信息和全局上下文信息,更准确地理解和处理图像数据。Bello等^[27]提出的AAConvNet通过在每个卷积层之后添加一个SA模块,根据输入特征的重要性自适应地调整卷积核的权重,显著提升了模型精确度。Chen等^[9]提出的TransUNet首先通过精心设计的CNN模块,从输入图像中提取丰富且具有层次性的特征表示,这些特征涵盖了图像的局部细节,随后利用多头自注意力对这些特征进行全局性的整合,以捕捉图像中的远程依赖关系和上下文信息,在医学图像分割任务中表现良好。TransAttUnet^[28]在Transformer的基础上引入了多级引导注意力和多尺度跳跃连接,有效缓解了卷积层堆叠和连续采样操作带来的精细细节丢失问题。

1.3 基于模糊学习的医学图像分割模型

在医学图像获取过程中,由于人体解剖结构的复杂性、器官形状的不规则性以及不同个体间的差异性等因素,图像本质上具有不确定性。此外,由于人类视觉感知的局限性和医学影像技术的特点,医学图像中往往包含一些模糊特征^[29]。与此同时,噪声也是一个普遍存在的问题,它可能来自成像设备的固有噪声、患者的运动等因素。基于模糊理论的分割方法可以很好地处理这种不确定性,还可以在在一定程度上抑制噪声的影响,提高图像分割的精确度和稳定性。现如今,基于模糊学习的医学图像分割模型在不断问世。由Nan等^[30]提出的模糊注意力神经网络创建了一个模糊注意力层,该模糊注意力层通过学习一种可调整的、基于特征图和高斯隶属函数的注意力映射,使得网络在做出决策时能够考虑输入数据中的不确定性。Ding等^[18]提出的FTransCNN模型采用了模糊积分这一数学概念,其允许模型在提取和融合多源信息时,能够更准确地识别并降低那些可能导致分割结果不准确的噪声或不一致性。此外作者还设计了模糊注意力融合模块,该模块利用模糊逻辑和注意力机制,能够动态地关注图像中最重要的区域,尤其是分割边界处,故而该模型能够更好地处理边界的不确定性。

上述3类方法在平衡局部细节信息和全局上下文信息的获取、尽可能减少模型计算资源请求量、处理病灶模糊边界、精确分割癌症病灶等方面不能很好地同时实现。因此,本文创新性地设计了Axial-FNet模型。

2 Axial-FNet模型

2.1 整体框架概述

Axial-FNet整体框架如图1所示,Axial-FNet是一个拥有并行双分支的模型,两个分支架构分别为门控轴向自注意力分支和模糊卷积神经网络分支,有不同的工作方式和工作目标,但在最后进行效果融合。

(1) 门控轴向自注意力分支架构

该分支首先对医学图像进行了下采样操作,该下采样部分包含了3个连续的卷积操作模块,每个卷积操作模块都按序由一次卷积操作、一次批归一化操作^[31]和一个ReLU激活函数所构成;然后,本文方法设计了两块连续的编码器和两块连续的解码器^[32](也称上采样模块),以门控轴向自注意力为核心内容,先做一次卷积操作,接着做一次反卷积操作,再使用一次ReLU激活函数作为一个解码器,编码器的详细内容见第2.2节;在编码器和解码器之间,经过编码器和解码器处理后的特征图中,相同尺寸的特征图进行一次跳跃连接,再输入下一个解码器中;随后本文方法做了一次批归一化操作;最后,本文方法引入了一个可学习的门控权重控制器用于确定捕捉全局上下文信息的程度,该门控权重控制器的详细内容见第2.3节。

该分支的目的是把图像中的全局上下文信息充分利用起来,并且在尽可能减少计算量的同时提升分割精确度。

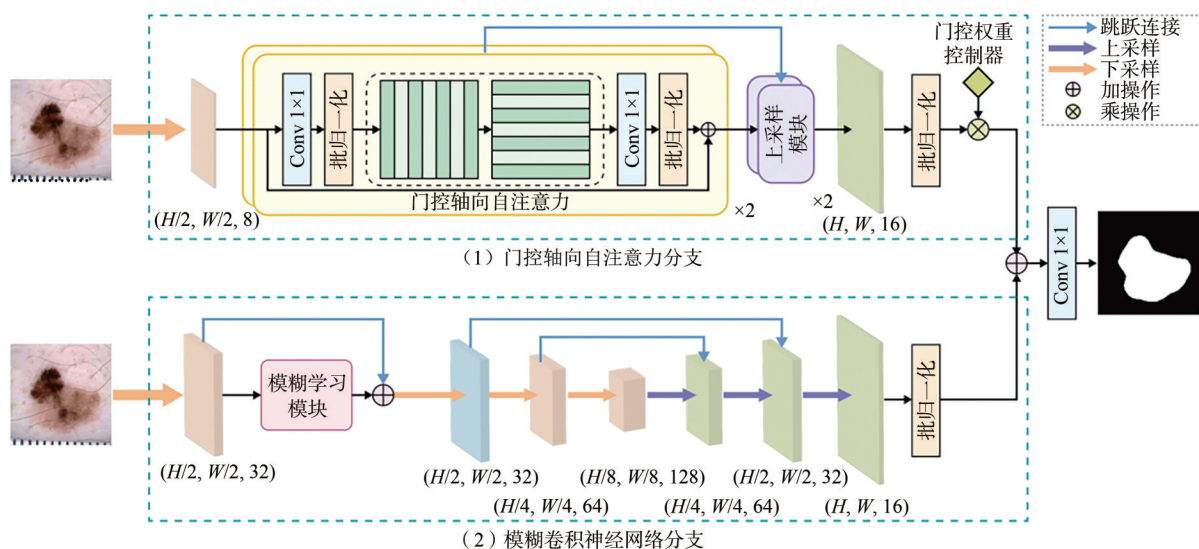


图1 Axial-FNet整体框架

(2) 模糊卷积神经网络分支架构

其下采样部分与门控轴向自注意力机制分支架构中的下采样模块并不相同, 是受 ResNet^[33] 的启发, 由本文创新性地自行设计, 在第2.4节中将详细阐述该下采样结构的内容。在下采样结构中的第一层网络架构之后引入了模糊学习模块进行处理, 模糊学习模块亦详见第2.4节; 其上采样部分共有3个上采样模块, 前两个上采样模块均按序由一次反卷积操作、一次卷积操作、一次批归一化操作 (BatchNorm) 以及一个 ReLU 激活函数^[34] 构成, 第三个上采样模块按序由一次反卷积操作、一次卷积操作和一次批归一化操作 (BatchNorm) 构成; 经过下采样部分处理过的特征图和上采样部分处理过的特征图之间, 相同尺寸的特征图进行一次跳跃连接, 然后输入下一个上采样模块中; 最后, 本文方法做了一次批归一化^[35] 操作。

该分支的目的是捕获图像中的重要局部细节, 充分学习图像中的局部特征信息, 并尽可能地解决医学图像本身存在的不确定性问题。

2.2 编码器

如图1所示, 在门控轴向自注意力分支中, 经过一次下采样操作后, 大小为 $x \in R^{\frac{H}{2} \times \frac{W}{2} \times C}$ 的特征图被输入本文方法所构建的第一块编码器中, 其中 H 、 W 、 C 分别表示特征图的高度、宽度、通道数, 这里将通道数 C 设置为8。

在第一块编码器中, 特征图先经过一次卷积核大小为 1×1 的卷积操作, 增加输入特征图的通道数, 以实现增加特征多样性和强化特征表达的作用, 然后再进行批归一化操作和一次 ReLU 激活函数操作。随后, 大小为 $y \in R^{\frac{H}{2} \times \frac{W}{2} \times C_1}$ 的特征图被输入门控轴向自注意力计算中, 其中 C_1 表示通道数。该自注意力在计算像素相似性的时候, 仅在高度轴和宽度轴这一行和一系列中进行操作, 二者计算方式和操作过程完全相同。为了便于理解门控轴向自注意力, 此处先以自注意力为引, 其计算方式如下:

$$z_{ij} = \sum_{h=1}^H \sum_{w=1}^W \text{softmax}(q_{ij}^T k_{hw}) v_{hw} \quad (1)$$

其中, q_{ij} 表示任意位置 $i \in \{1, \dots, H\}$ 且 $j \in \{1, \dots, W\}$ 处的查询, k_{hw} 和 v_{hw} 分别表示位置在 $h \in \{1, \dots, H\}$ 且 $w \in \{1, \dots, W\}$ 处的键和值, 再通过 $\text{softmax}()$ 函数

计算位置 (i, j) 处的信息和位置 (h, w) 处的相似度。

而门控轴向自注意力在高度轴上进行计算的方式如下:

$$z_{ij} = \sum_{h=1}^H \text{softmax} \left[\left(q_{ij}^T k_{ih} + G_Q q_{ij}^T r_{ih}^q + G_K k_{ih}^T r_{ih}^k \right) \cdot \left(G_{V1} v_{ih} + G_{V2} r_{ih}^v \right) \right] \quad (2)$$

式 (2) 依然遵循式 (1) 的计算方式, 仅将式 (1) 中计算每个位置上的信息和当前位置上的信息的相似度更改为计算当前位置高度轴上每个位置上的信息与当前位置上的信息的相似度。然后, 增加了在轴向上的位置偏差信息 r_{ih}^q 、 r_{ih}^k 和 r_{ih}^v , 增加了可学习的权重参数 G_Q 、 G_K 、 G_{V1} 和 G_{V2} 。

待门控轴向自注意力计算完毕后, 先进行一次 ReLU 激活函数操作, 再次进行卷积核大小为 1×1 的卷积操作和批归一化操作, 然后将当前的特征图信息与输入编码器前的特征图信息进行残差连接, 作为第一块编码器的最终输出, 同时也是第二块编码器的起始输入。

第二块编码器的计算方式和操作过程与第一块编码器类似, 仅改变了门控轴向自注意力的计算量, 即在高度轴和宽度轴上以步长为2进行计算, 最终得到大小为 $z_1 \in R^{\frac{H}{4} \times \frac{W}{4} \times C_2}$ 的特征图信息, 其中 C_2 表示通道数。由于第二块编码器处理完后输出的特征图和输入的特征图尺寸不一致, 需要对输入第二块编码器的特征图进行一次卷积操作和批归一化操作, 使二者尺寸一致, 然后再进行残差连接。

2.3 门控权重控制器

经过了两块编码器逐块编码运算之后, 便需要两块与编码器相对应的解码器进行上采样运算, 得到大小为 $z_3 \in R^{H \times W \times 16}$ 的特征图, 特征图的通道数固定为16。然后进行批归一化运算, 最终由本文方法所设计的门控权重控制器对所得到的整个特征图赋予权重。具体而言, 该门控权重控制器即一个一阶的随着模型训练可以不断学习的参数, 其动态控制学习局部细节信息和学习全局上下文信息的占比。该门控权重控制器工作原理的计算方式如下:

$$z_{\text{end}} = G_w \cdot R_o^{H \times W \times 16} \quad (3)$$

其中, z_{end} 代表门控权重控制器完成运算后的结果, G_w 为门控权重控制器, $R_o^{H \times W \times 16}$ 为完成批归一化运算后的特征图的参数信息。

2.4 下采样结构与模糊学习模块

受He等^[33]的ResNet的启发,本文设计了如图2所示的模糊卷积神经网络分支中的下采样结构,用于提取图像中的重要特征信息。

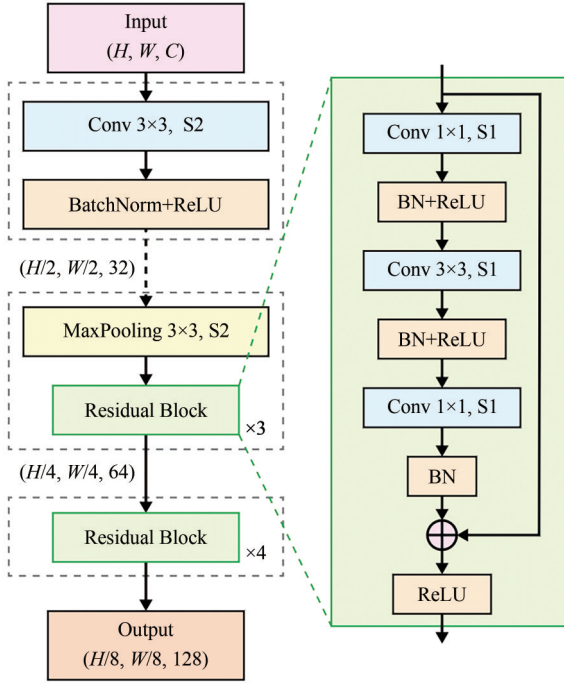


图2 模糊卷积神经网络分支中的下采样结构

该下采样结构包含3层网络架构,第一层网络架构按序由一个卷积层、一次批归一化操作(BN)和一个ReLU激活函数构成,其中卷积层的卷积核尺寸为 3×3 ,移动步长为2,本文将卷积操作对特征图尺寸的影响进行公式化,即如式(4)所示:

$$W_{\text{out}} = (W_{\text{in}} - k + 2P) / S + 1 \quad (4)$$

其中, W_{out} 是输出的特征图尺寸, W_{in} 是输入的医学图像尺寸, k 是卷积核尺寸, P 是零填充的大小, S 是卷积核移动步长的大小。

第二层网络架构中的3个残差模块采用串联方式进行工作,均采用图2中的Residual Block放大图所示的模块结构。每块残差模块的工作过程为:原始特征图首先经过卷积核尺寸为 1×1 、移动步长为1的卷积层、批归一化函数(BN)和ReLU激活函数;然后经过卷积核尺寸为 3×3 、移动步长为1的卷积层、批归一化函数和ReLU激活函数;最后经过卷积核尺寸为 1×1 、移动步长为1的卷积层和批归一化函数,得到中间过程特征图,其与原始特征图进行一次残差连接;再经过一个ReLU激

活函数,即可得到最终的输出特征图。

第三层网络架构由4个连续的残差模块构成,该残差模块的模块结构和第二层网络架构中的模块结构近乎相同,亦如图2中的Residual Block放大图所示,做出的改变仅仅为:对4个残差模块3次卷积操作过程中的特征图通道数的计算做了改变,同时设置第一个残差模块中卷积核为 3×3 的卷积操作的移动步长为2。由于第一个残差模块的3次卷积操作处理完后得到的中间过程图和输入特征图的尺寸不一致,需要对输入第一个残差模块的特征图进行一次卷积操作和批归一化操作,使二者尺寸一致,然后再进行残差连接。

针对皮肤癌图像中普遍存在的边界模糊和区域不确定问题,本文引入了一种如图3所示的创新性模糊学习模块。具体而言,该模糊学习模块包含模糊隶属度函数^[36]和模糊规则^[37],其核心目标在于建模特征图与像素语义类别之间复杂且模糊的关联规则。在模糊学习模块的输入端,本文针对特征图的每个通道独立应用 M 个模糊隶属度函数。特别的是,尽管模糊隶属度函数在同一通道内保持一致,但在不同通道间可根据特征性质灵活调整,以适应图像中多变的纹理和结构。对于每个通道内的每个特征点,每个模糊隶属度函数为其计算出模糊值,该模糊值描述其属于某个集合的程度。高斯隶属度函数广泛用于模糊集合、模糊逻辑等领域,该函数以高斯分布为原型,具有 C^2 级光滑连续性,其钟形曲线能精准描述元素对集合的渐变隶属关系,在0~1区间内呈现平滑过渡特征。通过指数运算量化像素值与高斯隶属度函数均值的偏离程度,经归一化处理后,输出值严格限定在0~1区间内,其中1表示完全隶属,0表示完全不隶属,能够很好地处理边界模糊和不确定性问题。故本文所有的模糊隶属度函数均采用高斯函数^[38],形式如下:

$$L_{x,y,m,c} = e^{-\left(\frac{P_{x,y,c} - \mu_{m,c}}{\sigma_{m,c}}\right)^2}, x=1, \dots, W, y=1, \dots, H, \\ m=1, \dots, M, c=1, \dots, C \quad (5)$$

其中, W 和 H 分别指特征图的宽度和高度, C 指特征图的通道数, M 为模糊隶属度函数的数量,是唯一需要预设的超参数, (x,y) 是通道 c 中特征点 $P_{x,y,c}$ 的位置坐标, $\sigma_{m,c}$ 和 $\mu_{m,c}$ 是第 m 个高斯隶属度函数的标准差和均值, $L_{x,y,m,c}$ 是通道 c 中位置在 (x,y) 处的特征点的第 m 个模糊输出的模糊值。特别的是,

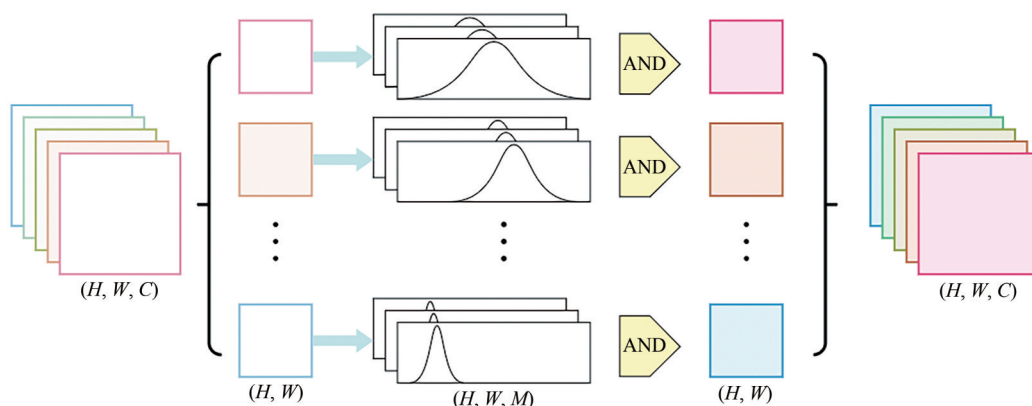


图3 模糊学习模块的结构

$\sigma_{m,c}$ 和 $\mu_{m,c}$ 是随着模型训练可以不断学习的数据信息。

最终，对于每个特征点的 M 个模糊输出，本文采用模糊逻辑“AND”进行融合，得到最终的模糊结果 $L_{x,y,c}$ ，过程如下：

$$L_{x,y,c} = \prod_{m=1}^M L_{x,y,m,c} \quad (6)$$

Axial-FNet的整体算法如算法1所示。

算法1: 基于模糊卷积结合门控轴向自注意力的皮肤癌图像分割模型 (Axial-FNet)

输入: 待分割图像 IMG, 训练批次大小 Batch-Size, 学习率 LR, 动量参数 Momentum, 权重衰减参数 WeightDecay, 迭代次数 Epoch

输出: 分割结果 y

for epoch in Epoch do

$y11 = \text{Downsample}(\text{IMG})$

 //在门控轴向自注意力分支中进行下采样

$y12 = \text{GatedAxialAttention}(y11)$

 //执行门控轴向自注意力

$y13 = \text{Upsample}(y12)$

 //执行上采样操作

$y14 = \text{BatchNorm}(y13)$

$y14 = \text{GatedWeight} * y14$

$y21 = \text{Downsample}(\text{IMG})$

$y22 = \text{FuzzyLearning}(y21)$

 //执行模糊学习模块操作

$y23 = y21 + y22$

$y24 = \text{Downsample}(y23)$

$y25 = \text{Downsample}(y24)$

$y26 = \text{Upsample}(y25)$

$y26 = y26 + y24$

$y27 = \text{Upsample}(y26)$

$y27 = y27 + y23$

$y28 = \text{Upsample}(y27)$

$y28 = \text{BatchNorm}(y28)$

$y = y28 + y14$

$y = \text{Conv}_1 \times 1(y)$

end

return y

3 实验和结果分析

第3.1节深入探讨了本文使用的数据集以及模型的具体参数配置。第3.2节详细介绍了对比实验的设置，包括评价指标的选择与对比模型的配置。第3.3节详细展示了对比实验的结果，经过分析，本文模型取得的效果更好。第3.4节通过消融实验和可视化分析，进一步验证了本文设计的门控权重控制器的有效性、模糊学习模块的有效性以及门控轴向自注意力机制的有效性。最后，第3.5节通过热力图揭示了模型的核心模块对图像特征的关注点，增强模型的可解释性。

3.1 实验设置

实验选用的数据集为公开的皮肤病变图像分割数据集，是由国际皮肤成像协作组织 (ISIC) 提供的，名称叫作 ISIC 2017^[39] 和 ISIC 2018^[40]，用于研究皮肤癌等疾病病灶的分割。具体而言，ISIC 2017 数据集包括 2 000 张图像用于训练，150 张图像用于验证，600 张图像用于测试。ISIC 2018 数据集包括 2 594 张图像用于训练，100 张图像用于验证，1 000 张图像用于测试。实验使用的深度学习框架为 PyTorch，设置输入模型的数据集图像的粒度块大小为 $128 \times 128 \times 3$ ，实验还对数据集使用随机翻

转技术,以增加数据的多样性。实验使用的损失函数为对数负似然损失函数,通过 Adam 优化器来优化模型,训练采用小批量训练方式,设置训练总次数 epoch 为 100,批次大小 batch_size 为 16,学习率 learning_rate 为 0.01,动量参数 momentum 为 0.9,权重衰减参数 weight_decay 为 0.00001。本文采用的实验平台为 NVIDIA A6000 GPU,开发工具为 PyCharm 2022.2.1 专业版,使用 Python 3.10 语言实现实验中相关算法。

3.2 对比实验

为了更全面地评估 Axial-FNet 的性能,本文将与其他多个先进模型进行比较,分别是 U-Net^[4]、Attention U-Net^[41]、Medical transformer^[13]、EGE-UNet^[21]和 UCTransNet^[42]。在实验中,本文通过使用平均交并比(MIoU)、F1 分数(F1-score)和准确率(Accuracy)这 3 种指标来衡量模型的表现。TP 表示真实为正类的像素被正确预测为正类的数量,FP 表示真实为负类的像素被错误预测为正类的数量,TN 表示表示真实为负类的像素被正确地预测为负类的数量,FN 表示真实为正类的像素被错误地预测为负类的数量。 A 表示预测的分割区域, B 表示真实的分割区域, N 表示类别的总数,即在分割任务中需要区分的不同目标的数量, IoU_i 表示第 i 个类别的交并比结果。

平均交并比是衡量图像分割模型整体性能的一个重要指标。它表示模型预测的分割区域与真实分割区域之间的重叠程度。交并比计算方式如下:

$$IoU = \frac{A \cap B}{A \cup B} \quad (7)$$

平均交并比的计算方式为:

$$MIoU = \frac{1}{N} \sum_{i=1}^N IoU_i \quad (8)$$

召回率表示在所有真实的正类像素中,模型正确预测为正类的比例,其计算方式如下:

$$recall = \frac{TP}{TP + FN} \quad (9)$$

精确度表示在所有被预测为正类的像素中,模型正确预测的比例,其计算方式如下:

$$precision = \frac{TP}{TP + FP} \quad (10)$$

准确率表示在所有预测结果中,正确的占总预测值的比例,其计算方式如下:

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

F1 分数是精确度和召回率的调和平均数,用于综合评估模型性能,其计算方式如下:

$$F1-score = \frac{2 \times precision \times recall}{precision + recall} \quad (12)$$

3.3 实验分析

表 1 和表 2 分别展示了 Axial-FNet 模型在 ISIC 2017 数据集和 ISIC 2018 数据集上与其他先进模型的定量对比结果。

表 1 ISIC 2017 数据集下不同模型对比实验的定量数据结果

模型	参数量	MIoU	F1-score	Accuracy
U-Net ^[4]	7 766 018	70.44%	79.97%	91.57%
Attention U-Net ^[41]	34 878 638	71.11%	80.62%	91.72%
Medical transformer ^[13]	1 560 116	73.37%	82.36%	92.13%
EGE-UNet ^[21]	53 383	72.20%	81.46%	92.12%
UCTransNet ^[42]	66 430 722	72.98%	82.01%	92.20%
Axial-FNet	2 353 974	74.23%	83.05%	92.89%

表 2 ISIC 2018 数据集下不同模型对比实验的定量数据结果

模型	参数量	MIoU	F1-score	Accuracy
U-Net ^[4]	7 766 018	77.90%	85.32%	91.82%
Attention U-Net ^[41]	34 878 638	78.19%	85.77%	91.69%
Medical transformer ^[13]	1 560 116	79.39%	86.54%	92.23%
EGE-UNet ^[21]	53 383	78.47%	85.62%	92.16%
UCTransNet ^[42]	66 430 722	80.38%	87.82%	92.99%
Axial-FNet	2 353 974	80.91%	88.13%	93.10%

对实验结果进行分析可以发现,相比于以往的一些先进模型,本文提出的 Axial-FNet 表现更加出色。在 ISIC 2017 数据集上, Axial-FNet 在 MIoU、F1-score 和 Accuracy 这 3 个指标上达到了 74.23%、83.05% 和 92.89%, 比 U-Net 高出 3.79%、3.08% 和 1.32%, 并且在计算资源需求上远小于 U-Net; 比 Attention U-Net 高出 3.12%、2.43% 和 1.17%, 同样在计算资源需求上远小于 Attention U-Net; 尽管在计算资源需求上略高于 Medical transformer, 但在 MIoU、F1-score 和 Accuracy 上高出 0.86%、0.69% 和 0.76%; 同样, 尽管在计算资源需求上高于 EGE-UNet, 但在 MIoU、F1-score 和 Accuracy 上高出 2.03%、1.59% 和 0.77%。比 UCTransNet 分别高出 1.25%、1.04% 和 0.69%, 同时在计算资源需求上远远小于 UCTransNet。同理, 在 ISIC 2018 数据集上, Axial-FNet 在 MIoU、F1-score 和 Accuracy 这

3个指标上达到了80.91%、88.13%和93.10%，比U-Net高出3.01%、2.81%和1.28%，并且在计算资源需求上远小于U-Net；比Attention U-Net高出2.72%、2.36%和1.41%，同样在计算资源需求上远小于Attention U-Net；尽管在计算资源需求上略高于Medical transformer，但在MIoU、F1-score和Accuracy上高出1.52%、1.59%和0.87%；同样，尽管在计算资源需求上高于EGE-UNet，但在MIoU、F1-score和Accuracy上高出2.44%、2.51%和0.94%。比UCTransNet分别高出0.53%、0.31%和0.11%，

同时在计算资源需求上远远小于UCTransNet。为了更直观地进行对比，图4和图5分别展示了不同模型在ISIC 2017数据集和ISIC 2018数据集上的部分图像分割结果的可视化示图。根据视图，可以很直接地发现本文模型在整体分割效果上优于其他模型，并且在分割病灶与正常皮肤的边界处时更加精准和清晰，充分展现了本文模型在处理边界模糊情况和整体分割情况上的优越性能。

3.4 消融实验

为了探讨本文提出的Axial-FNet中三大创新点

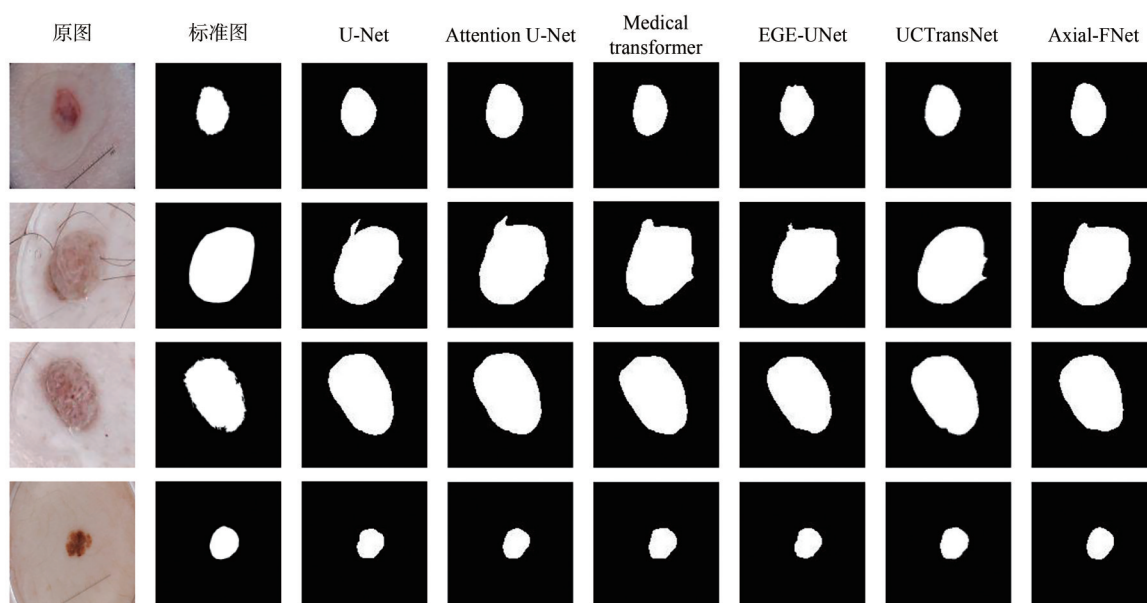


图4 不同模型在ISIC 2017数据集上的部分预测结果可视化

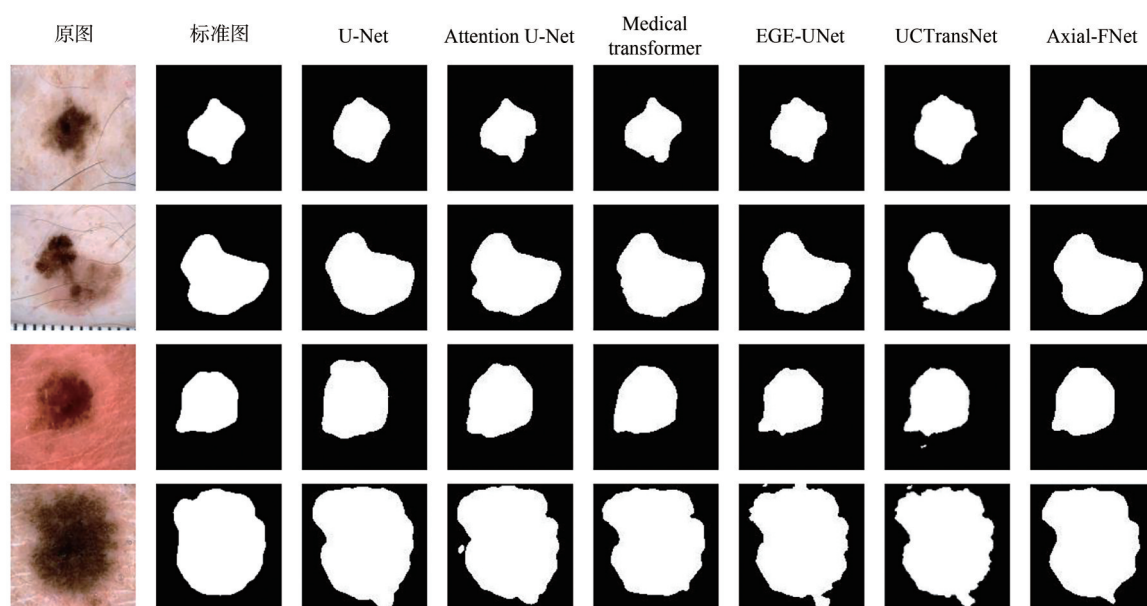


图5 不同模型在ISIC 2018数据集上的部分预测结果可视化

模块的有效性以及 Axial-FNet 中这些模块设置的最佳性, 本文同样在对比实验使用的 ISIC 2017 数据集和 ISIC 2018 数据集上设置消融实验, 实验参数细节和采用的评价指标均同对比实验。所需探讨的内容与门控权重控制器、模糊学习模块和门控轴向自注意力中的编码器和解码器数量的设置有关。为此, 本文设置了以下 6 组实验。

(1) 控制其他条件不变, 只保留和删除门控权重控制器, 删除时记为 N-Weight。

(2) 控制其他条件不变, 只保留和删除模糊学习模块, 删除时记为 N-Fuzzy。

(3) 控制其他条件不变, 同时将门控权重控制器和模糊学习模块删除与保留, 同时删除时记为 N-Weight and N-Fuzzy。

(4) 控制其他条件不变, 仅将门控轴向自注意力分支中的编码器和解码器的数量设置为 1 进行实验, 记为 1-Encoder。

(5) 控制其他条件不变, 仅将门控轴向自注意力分支中的编码器和解码器的数量设置为 3 进行实验, 记为 3-Encoder。

(6) 控制其他条件不变, 仅将门控轴向自注意力分支中的编码器和解码器的数量设置为 4 进行实验, 记为 4-Encoder。

通过表 3 和表 4 中不同数据集下消融实验的定量数据结果可见, 门控权重控制器与模糊学习模块任何一个发生缺失或者均发生缺失时, 分割效果都会下降, 尤其是当模糊学习模块发生缺失时, 分割

表3 ISIC 2017数据集下消融实验的定量数据结果

模型	MIoU	F1-score	Accuracy
N-Weight	73.67%	82.96%	92.27%
N-Fuzzy	72.52%	81.98%	91.98%
N-Weight and N-Fuzzy	72.43%	81.75%	91.93%
1-Encoder	71.69%	80.93%	92.23%
3-Encoder	70.45%	80.37%	91.46%
4-Encoder	66.79%	77.76%	90.02%
Axial-FNet	74.23%	83.05%	92.89%

表4 ISIC 2018数据集下消融实验的定量数据结果

模型	MIoU	F1-score	Accuracy
N-Weight	79.93%	87.43%	92.66%
N-Fuzzy	78.83%	86.74%	92.15%
N-Weight and N-Fuzzy	79.65%	86.37%	92.39%
1-Encoder	75.99%	84.37%	91.06%
3-Encoder	77.75%	85.91%	92.22%
4-Encoder	76.45%	84.18%	91.61%
Axial-FNet	80.91%	83.05%	93.10%

效果下降得较为明显。

通过图 6 和图 7 不同数据集下消融实验的部分结果可视化可见, 模糊学习模块发生缺失时, 病灶与正常皮肤边界的分割效果比门控权重控制器缺失时病灶与正常皮肤边界的分割效果要更差, 体现了模糊学习模块可以很好地处理边界模糊与细节不确定等问题。门控轴向自注意力中编码器的数量为 2 时分割效果最好, 此外, 当编码器数量增加时还会增加计算资源需求量。另外, 本文推测, 当

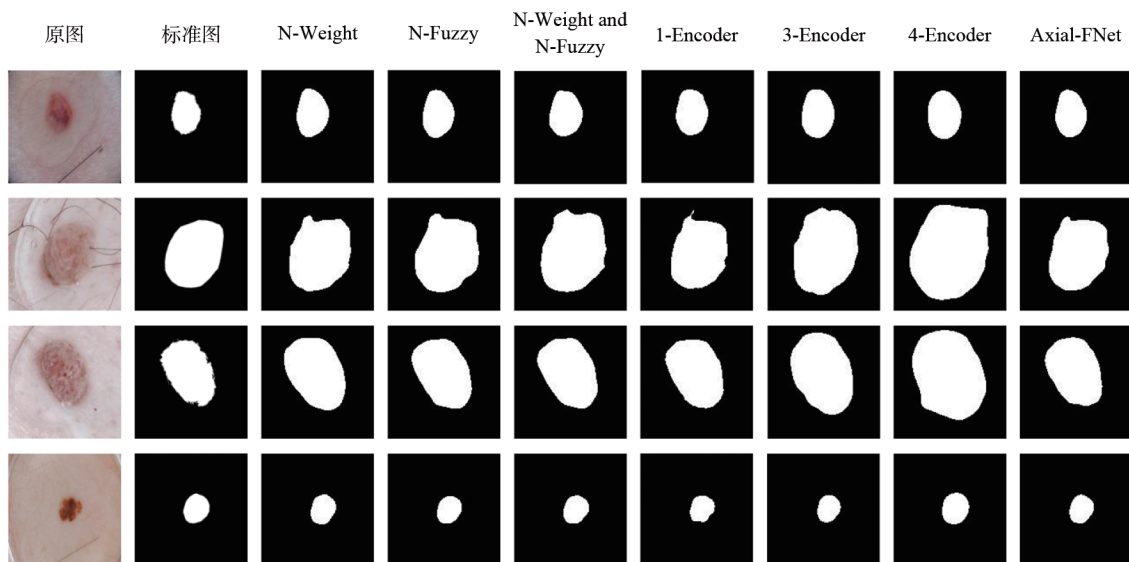


图6 ISIC 2017数据集下消融实验的部分结果可视化

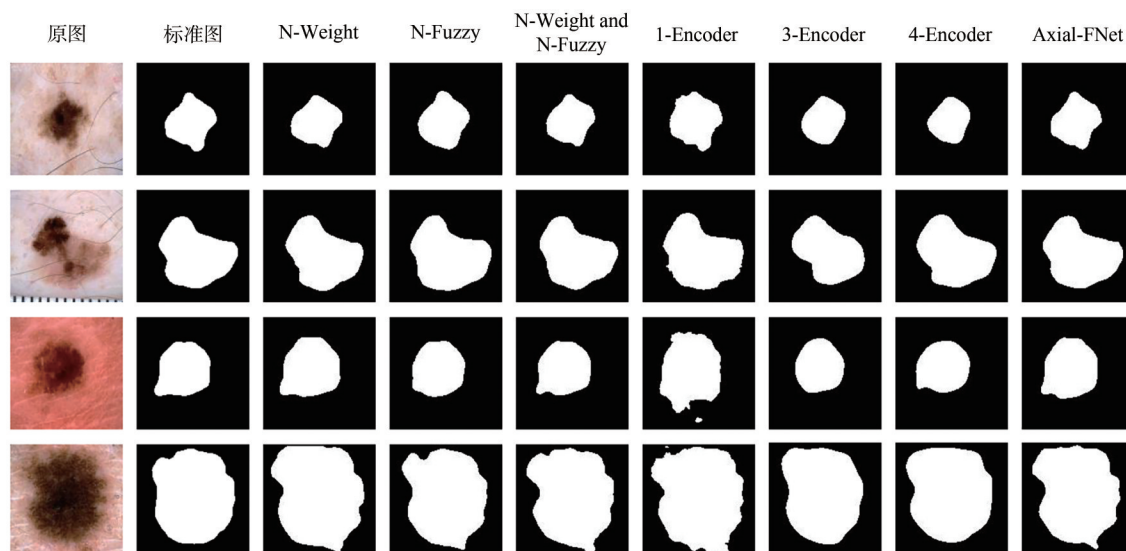


图7 ISIC 2018数据集下消融实验的部分结果可视化

编码器数量较多时,模型会过于注重整合特征图中的全局信息,当编码器数量较少时,模型对特征图中的全局整体信息的把握不够到位,均无法很好地处理特征图中的信息。在不同方法对比中,本文提出的Axial-FNet效果最佳,即本文提出的所有创新模块具有有效性,且创新模块的具体设置具有最佳性。

3.5 模型可解释性分析

为了分析模型中不同模块的协同工作效应以及解释模型的门控轴向自注意力分支和模糊卷积神经网络分支双分支并行架构的设计原理,本文将通过特征可视化方法来实现,即使用热力图来解释不同模块学习的图像信息。

不同模块对于特征学习的热力图可视化如图8所示。图8(a)随机选取了ISIC 2018数据集中测试集的两张图像作为分析对象,从上往下依次为原图、门控轴向自注意力学习到的特征热力图、模糊学习模块学习到的特征热力图;图8(b)随机选取了ISIC 2017数据集中测试集的两张图像作为分析对象,从上往下依次为原图、门控轴向自注意力学习到的特征热力图、模糊学习模块学习到的特征热力图。通过观察可以发现,门控轴向自注意力侧重于学习大面积病灶区域的特征,体现了门控轴向自注意力可以有效捕捉全局上下文信息的依赖关系,而模糊学习模块侧重于学习病灶的边界区域,能够有效区分病灶和正常皮肤,体现了模糊学习模块可以很好地处理边界模糊性和不确定性。因此,双分支关注点互补,有效提升了图像分割的精确度和可靠性。

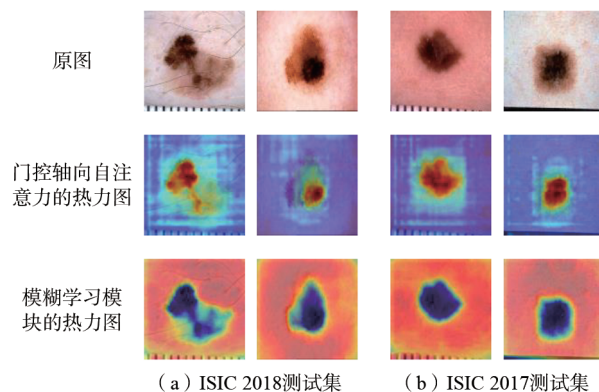


图8 不同模块对于特征学习的热力图可视化

4 结束语

本文提出的基于模糊卷积结合门控轴向自注意力的皮肤癌图像分割模型将门控轴向自注意力与CNN相结合,兼顾局部感知和全局上下文信息的提取,实现了皮肤癌图像分割效果的整体提升,同时解决了传统CNN与自注意力结合时计算量过大的问题。本文模型引入的模糊学习模块解决了图像中的病灶边界模糊性和噪声等问题,提升了分割的精确度和可靠性。

医学图像分割技术通过对患者的皮肤癌图像进行精确分割,使医生可以更准确地了解病变的位置、大小和形状等信息,从而制定更加个性化的治疗方案,提高疾病的治疗效果和患者的生活质量。而基于深度学习的医学图像分割模型不仅提高了分割精确度和效率,更推动了医学图像分割技术的不断

断发展和完善。随着今后人工智能技术的不断进步和应用场景的不断拓展,人工智能在医学领域的应用前景也将更加广阔。

参考文献:

- [1] AHMEDIN JEMAL DVM P. Cancer statistics, 2018[J]. CA: A Cancer Journal for Clinicians, 2018, 68(1): 7-30.
- [2] WU H S, CHEN S H, CHEN G L, et al. FAT-Net: feature adaptive transformers for automated skin lesion segmentation[J]. Medical Image Analysis, 2022, 76: 102327.
- [3] KHAN M A, SHARIF M I, RAZA M, et al. Skin lesion segmentation and classification: a unified framework of deep neural network features fusion and selection[J]. Expert Systems, 2022, 39(7): e12497.
- [4] RONNEBERGER O, FISCHER P, BROX T. U-Net: convolutional networks for biomedical image segmentation[C]//Proceedings of the Medical Image Computing and Computer-Assisted Intervention (MICCAI) 2015. Cham: Springer International Publishing, 2015: 234-241.
- [5] ZHOU Z W, RAHMAN SIDDIQUEE M M, TAJBAKHSH N, et al. UNet++: a nested U-Net architecture for medical image segmentation[C]//Proceedings of the Deep Learning in Medical Image Analysis and Multimodal Learning for Clinical Decision Support. Cham: Springer International Publishing, 2018: 3-11.
- [6] XIAO X, LIAN S, LUO Z M, et al. Weighted res-UNet for high-quality retina vessel segmentation[C]//Proceedings of the 2018 9th International Conference on Information Technology in Medicine and Education (ITME). Piscataway: IEEE Press, 2018: 327-331.
- [7] IGLOVNIKOV V, SHVETS A, BUSLAEV A V, et al. TernaUNet: U-Net with VGG11 encoder pre-trained on ImageNet for image segmentation[EB]. 2018.
- [8] VASWANI A, SHAZEER N, PARMAR N, et al. Attention is all you need[J]. Advances in Neural Information Processing Systems, 2017, 30.
- [9] CHEN J N, MEI J R, LI X H, et al. TransUNet: rethinking the U-Net architecture design for medical image segmentation through the lens of transformers[J]. Medical Image Analysis, 2024, 97: 103280.
- [10] CORDONNIER J B, LOUKAS A, JAGGI M. Multi-head attention: collaborate instead of concatenate[EB]. 2020.
- [11] CHEN Y F, ZOU B F, GUO Z X, et al. SCUNet: swin-UNet and CNN bottleneck hybrid architecture with multi-fusion dense skip connection for pulmonary embolism CT image segmentation[C]//Proceedings of the 2024 IEEE/CVF Winter Conference on Applications of Computer Vision (WACV). Piscataway: IEEE Press, 2024: 7744-7752.
- [12] LIU Z, LIN Y T, CAO Y, et al. Swin transformer: hierarchical vision transformer using shifted windows[C]//Proceedings of the 2021 IEEE/CVF International Conference on Computer Vision (ICCV). Piscataway: IEEE Press, 2021: 9992-10002.
- [13] VALANARASU J M J, OZA P, HACIHALILOGLU I, et al. Medical transformer: gated axial-attention for medical image segmentation[C]//Proceedings of the Medical Image Computing and Computer Assisted Intervention (MICCAI) 2021. Cham: Springer International Publishing, 2021: 36-46.
- [14] SABOUR S, VORA S, DUCKWORTH D, et al. RobustNeRF: ignoring distractors with robust losses[C]//Proceedings of the 2023 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Piscataway: IEEE Press, 2023: 20626-20636.
- [15] LEE H J, KIM J U, LEE S M, et al. Structure boundary preserving segmentation for medical image with ambiguous boundary[C]//Proceedings of the 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Piscataway: IEEE Press, 2020: 4816-4825.
- [16] GUAN C, WANG S L, LIEW A W. Lip image segmentation based on a fuzzy convolutional neural network[J]. IEEE Transactions on Fuzzy Systems, 2020, 28(7): 1242-1251.
- [17] DENG Y, REN Z Q, KONG Y Y, et al. A hierarchical fused fuzzy deep neural network for data classification[J]. IEEE Transactions on Fuzzy Systems, 2017, 25(4): 1006-1012.
- [18] DING W P, WANG H P, HUANG J S, et al. FTransCNN: fusing transformer and a CNN based on fuzzy logic for uncertain medical image segmentation[J]. Information Fusion, 2023, 99: 101880.
- [19] IBTEHAZ N, RAHMAN M S. MultiResUNet: rethinking the U-Net architecture for multimodal biomedical image segmentation[J]. Neural Networks, 2020, 121: 74-87.
- [20] YAN X Y, TANG H, SUN S L, et al. AFTer-UNet: axial fusion transformer UNet for medical image segmentation[C]//Proceedings of the 2022 IEEE/CVF Winter Conference on Applications of Computer Vision (WACV). Piscataway: IEEE Press, 2022: 3270-3280.
- [21] RUAN J C, XIE M Y, GAO J S, et al. EGE-UNet: an efficient group enhanced UNet for skin lesion segmentation[C]//Medical Image Computing and Computer Assisted Intervention (MICCAI) 2023. Cham: Springer Nature Switzerland, 2023: 481-490.
- [22] DOSOVITSKIY A. An image is worth 16x16 words: transformers for image recognition at scale[EB]. 2020.
- [23] ZHENG S X, LU J C, ZHAO H S, et al. Rethinking semantic segmentation from a sequence-to-sequence perspective with transformers[C]//Proceedings of the 2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Piscataway: IEEE Press, 2021: 6881-6890.
- [24] XIE E Z, WANG W H, YU Z D, et al. SegFormer: simple and efficient design for semantic segmentation with transformers[J]. Advances in Neural Information Processing Systems, 2021, 34: 12077-12090.
- [25] CAO H, WANG Y Y, CHEN J N, et al. Swin-unet: unet-like pure transformer for medical image segmentation[EB]. 2021.
- [26] LI C, CUI Y, LUO N, et al. Trans-ResNet: integrating transformers and CNNs for Alzheimer's disease classification[C]//Proceedings of the 2022 IEEE 19th International Symposium on Biomedical Imaging (ISBI). Piscataway: IEEE Press, 2022: 1-5.
- [27] BELLO I, ZOPH B, LE Q, et al. Attention augmented convolutional networks[C]//Proceedings of the 2019 IEEE/CVF International Conference on Computer Vision (ICCV). Piscataway: IEEE Press, 2019: 3286-3295.
- [28] CHEN B Z, LIU Y S, ZHANG Z, et al. TransAttUnet: multi-level attention-guided U-Net with transformer for medical image segmentation[J]. IEEE Transactions on Emerging Topics in Computational Intelligence, 2024, 8(1): 55-68.
- [29] JIANG Q, JIN X, CUI X H, et al. A lightweight multimodal medical image fusion method using similarity measure between intuitionistic fuzzy sets joint Laplacian pyramid[J]. IEEE Transactions on Emerging Topics in Computational Intelligence, 2023, 7(3): 631-647.
- [30] NAN Y, DEL SER J, TANG Z Y, et al. Fuzzy attention neural network to tackle discontinuity in airway segmentation[J]. IEEE Transactions on Neural Networks and Learning Systems, 2024, 35(6): 7391-7404.
- [31] 崔文成, 王可丽, 邵虹. 基于稠密块和注意力机制的肺部病理图像异常细胞分割[J]. 智能科学与技术学报, 2023, 5(4): 525-534.
- CUI W C, WANG K L, SHAO H. Abnormal cell segmentation for lung

pathological image based on denseblock and attention mechanism[J]. Chinese Journal of Intelligent Science and Technology, 2023, 5(4): 525-534.

- [32] HAN K, WANG Y H, CHEN H T, et al. A survey on vision transformer[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2023, 45(1): 87-110.
- [33] HE K M, ZHANG X Y, REN S Q, et al. Deep residual learning for image recognition[C]//Proceedings of the 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Piscataway: IEEE Press, 2016: 770-778.
- [34] 刘霞, 王迪. 深度ReLU神经网络的万有一致性[J]. 中国科学: 信息科学, 2024, 54(3): 638-652.
LIU X, WANG D. Universal consistency of deep ReLU neural networks[J]. Science China Information Sciences, 2024, 54(3): 638-652.
- [35] IOFFE S, SZEGEDY C, PARANHOS L, et al. Batch normalization: accelerating deep network training by reducing internal covariate shift[EB]. 2015.
- [36] DING W P, GENG S, WANG H P, et al. FDiff-Fusion: Denoising diffusion fusion network based on fuzzy learning for 3D medical image segmentation[J]. Information Fusion, 2024, 112: 102540.
- [37] 周天奕, 丁卫平, 黄嘉爽, 等. 模糊逻辑引导的多粒度深度神经网络[J]. 模式识别与人工智能, 2023, 36(9): 778-792.
ZHOU T Y, DING W P, HUANG J S, et al. Fuzzy logic guided deep neural network with multi-granularity[J]. Pattern Recognition and Artificial Intelligence, 2023, 36(9): 778-792.
- [38] 赵亮, 李梦威, 郑玉卿, 等. 二氧化氮浓度时空预测: 一种区间二型直觉模糊神经网络方法[J]. 智能科学与技术学报, 2024, 6(2): 253-261.
ZHAO L, LI M W, ZHENG Y Q, et al. Temporal and spatial prediction of nitrogen dioxide concentration: an interval type II intuitionistic fuzzy neural network method[J]. Chinese Journal of Intelligent Science and Technology, 2024, 6(2): 253-261.
- [39] CODELLA N C F, GUTMAN D, CELEBI M E, et al. Skin lesion analysis toward melanoma detection: a challenge at the 2017 International symposium on biomedical imaging (ISBI), hosted by the international skin imaging collaboration (ISIC)[C]//Proceedings of the 2018 IEEE 15th International Symposium on Biomedical Imaging (ISBI 2018). Piscataway: IEEE Press, 2018: 168-172.
- [40] CODELLA N, ROTEMBERG V, TSCHANDL P, et al. Skin lesion analysis toward melanoma detection 2018: a challenge hosted by the international skin imaging collaboration (ISIC)[EB]. 2019.
- [41] OKTAY O, SCHLEMPER J, LE FOLGOC L, et al. Attention U-Net: learning where to look for the pancreas[EB]. 2018.
- [42] WANG H, CAO P, WANG J, et al. Uctransnet: rethinking the skip connections in U-Net from a channel-wise perspective with transformer [C]//Proceedings of the AAAI Conference on Artificial Intelligence.

Palo Alto: AAAI Press, 2022, 36(3): 2441-2449.

[作者简介]



姜舒 (1993-), 女, 博士, 南通大学人工智能与计算机学院讲师, 主要研究方向为深度学习、大数据分析等。



陈琨 (2004-), 男, 南通大学张謇学院本科生, 主要研究方向为深度学习、模糊学习和图像处理等。



丁卫平 (1979-), 男, 博士, 南通大学人工智能与计算机学院教授、博士生导师, 主要研究方向为多模态机器学习、多粒度计算、演化计算和医学大数据分析等。



周天奕 (2000-), 男, 南通大学人工智能与计算机学院硕士生, 主要研究方向为模糊集、深度学习等。



朱越 (2004-), 男, 南通大学张謇学院本科生, 主要研究方向为深度学习、粗糙集等。