

融合大核门控及双注意力的骶髂关节分割网络^{*}

严武军 景莹 徐莹臣 张晓丽 王程

(太原师范学院计算机科学与技术学院 晋中 030600)

摘要: 强直性脊柱炎是一种慢性炎症性疾病,其早期诊断依赖于骶髂关节病变特征的准确识别。然而,由于骶髂关节解剖结构复杂、病灶呈现多尺度异质性,且易受CT部分容积效应及噪声干扰,传统分割方法的精度难以满足临床需求。为此,提出了一种基于多尺度注意力融合的网络模型(MAG-UNet)。该模型通过多尺度特征融合模块(MFF)强化局部-全局特征协同表征,结合双路径注意力机制(DA)的空间-通道自适应加权,并引入大核分组注意力门控(LGAG)以解决跨尺度特征耦合问题。在山西白求恩医院提供的数据集上进行的实验表明,MAG-UNet在骶髂关节CT分割中取得了显著的性能提升,Dice系数达到92.4%,IoU达到86.0%,较U-Net基线模型提升3.4%(IoU)。本文为强直性脊柱炎的早期诊断提供了可靠的技术支持,具有重要的临床应用价值与推广潜力。

关键词: 强直性脊柱炎;骶髂关节;医学影像分割;多尺度注意力融合;特征优化

中图分类号: TP391.41;TN911.73 **文献标识码:** A **国家标准学科分类代码:** 520.6040

Sacral-iliac joint segmentation network integrating large kernel gated mechanisms and dual attention

Yan Wujun Jing Ying Xu Yingchen Zhang Xiaoli Wang Cheng

(School of Computer Science and Technology, Taiyuan Normal University, Jinzhong 030600, China)

Abstract: Ankylosing spondylitis is a chronic inflammatory disease whose early diagnosis depends on the accurate identification of pathological features in the sacroiliac joint. However, due to the complex anatomical structure of the sacroiliac joint, the multiscale heterogeneity of lesions, as well as interference from partial volume effects and noise in CT imaging, the accuracy of traditional segmentation methods often fails to meet clinical demands. To address these challenges, this study proposes a Multiscale Attention-Guided U-Net (MAG-UNet). The model enhances local-global feature representation through a Multiscale Feature Fusion (MFF) module, integrates spatial-channel adaptive weighting via a Dual-path Attention (DA) mechanism, and introduces a Large-kernel Grouped Attention Gate (LGAG) to resolve cross-scale feature coupling issues. Experiments conducted on a dataset provided by Shanxi Bethune Hospital demonstrate that MAG-UNet achieves significant performance improvements in sacroiliac joint CT segmentation, with a Dice coefficient of 92.4% and an Intersection over Union (IoU) of 86.0%, surpassing the baseline U-Net model by 3.4% in IoU. This study provides a reliable technical solution for the early diagnosis of AS, offering substantial clinical value and broad potential for practical application.

Keywords: ankylosing spondylitis; sacroiliac joint; medical image segmentation; multiscale attention fusion; feature optimization

0 引言

骶髂关节炎(Sacroiliitis)是一种由免疫介导的慢性炎症性疾病,主要累及骶髂关节及脊柱。在众多相关疾病中,强直性脊柱炎(ankylosing spondylitis, AS)以其高发病率和典型性成为最具代表性的病症^[1]。流行病学统计显示,强直性

脊柱炎的患病率介于0.1%~0.5%,且男性患者比例较高^[2]。随着疾病的进展,患者常出现关节僵硬、疼痛及活动受限等症状,严重时甚至可能导致关节融合,显著影响生活质量^[3-4]。然而,由于早期症状隐匿且缺乏特异性,许多患者在确诊时往往已处于中晚期,这无疑增加了诊断和治疗的难度^[5-6]。医学影像技术的革新正驱动着临床诊断范式的深刻

变革。计算机断层扫描(computed tomography, CT)凭借其亚毫米级空间分辨率与多平面重建能力,在骨关节系统疾病诊断中展现出独特优势^[7]。特别是在强直性脊柱炎骶髂关节病变的诊断中,CT影像突破了传统X线平片的局限,实现了关节间隙、骨皮质和骨髓腔的三维可视化,精准量化了关节间隙狭窄和骨质侵蚀^[8]。然而,影像数据的精准挖掘依赖于高效的分割技术。传统图像分割技术(如阈值分割、区域生长和聚类分析)在解剖结构复杂、对比度低的骶髂关节影像中表现欠佳,制约了自动化诊断系统的临床应用效能。深度学习技术的快速发展为医学影像分割领域带来了革命性突破。2014年Long等^[9]提出的全卷积网络(fully convolutional network, FCN)首次实现了端到端的像素级预测,随后,Ronneberger团队^[10]设计的U-Net通过U形编解码架构和跳跃连接机制显著提高了医学影像分割精度。针对U-Net的局限性,研究者们提出了多种改进,如Çiçek等^[11]开发的3D U-Net,通过三维卷积核优化了医学体数据建模;Zhou等^[12]提出的U-Net++,优化了多尺度特征融合;Oktay等^[13]引入的Attention U-Net通过空间注意力机制提升了对病灶区域的关注;近期武丽团队^[14]进一步在U-Net框架中融入公平因子,通过半监督学习机制缓解了标注数据稀缺场景下的模型偏差问题。许超等^[15]改进Swin Transformer实现膝关节炎X光影像自动诊断;王霞霞等^[16]构建双阶段网络增强超声图像分割的抗噪能力。此外,结合目标检测与语义分割的模型(如Mask R-CNN^[17]和DeepLab系列^[18])在复杂病灶处理中表现优异。深度学习在骶髂关节疾病的自动诊断与图像分割中展现出巨大潜力。结合放射组学与3D CNN的方法显著提升了骶髂关节炎症病灶的识别与分级效果^[19-22]。

尽管医学影像分割技术取得进展,但在处理解剖结构

复杂、边界模糊的骶髂关节影像时,仍面临诸多挑战。传统的CNN和U-Net模型在多尺度特征建模、空间与通道信息交互以及局部与全局上下文平衡上存在不足,导致对微小骨侵蚀、低对比度炎性水肿等病灶分割不精确,难以满足临床高精度需求。为此,本文提出改进的3D U-Net,通过多层次特征优化提升分割性能。

具体而言,本研究的创新点如下:

1)针对骶髂关节CT图像中解剖结构与病灶尺度的显著差异,设计异构卷积核($1\times 1, 3\times 3, 5\times 5$)并行提取机制,增强了模型对微小病灶和复杂解剖边界的表征能力。

2)为解决传统卷积操作在特征交互中的局限性,提出了结合位置注意力(position attention module, PAM)和通道注意力(channel attention module, CAM)的双路加权策略,动态调节特征图的权重分布,精确提取细节特征,显著提高了复杂病灶分割任务的精度与鲁棒性。

3)针对局部细节与全局上下文信息融合不足的问题,设计了自适应门控机制,通过多级特征聚合,高效整合局部纹理与全局语义,提升了复杂结构分割的适应性及准确性。

1 网络构建

1.1 MAG-UNet 介绍

山西白求恩医院作为山西省大型三甲医院,其风湿免疫科在强直性脊柱炎(AS)及骶髂关节炎等相关疾病的诊疗方面积累了丰富的临床经验,骶髂关节CT影像分析在临床评估和诊断过程中发挥着重要作用。然而,传统手动分割方法存在三大核心问题:效率瓶颈、主观误差以及微小病灶漏诊,严重制约临床诊断效率与准确性。为解决这一问题,本文提出了一种创新的MAG-UNet模型,如图1所示,旨在提高CT影像的分割精度和效率。

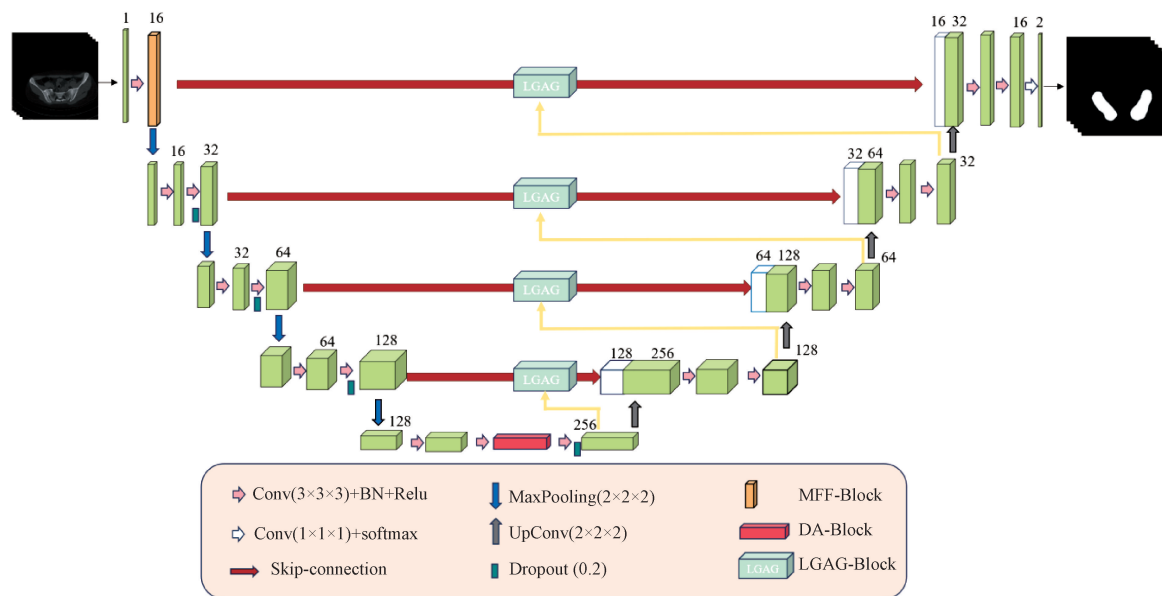


图1 MAG-UNet 网络模型

Fig. 1 MAG-UNet network model

该模型基于经典的 3D U-Net 架构,采用层次化的三维卷积编解码结构。输入为 $50 \times 256 \times 256$ 体素的三维特征图,编码路径通过四级下采样逐步提取特征并引入多尺度特征融合模块(multi-scale feature fusion, MFF)采用 $1 \times 1 \times 1$ 、 $3 \times 3 \times 3$ 和 $5 \times 5 \times 5$ 并行卷积核,提升对不同尺度病灶的检测能力。随后,通过四级编码器将空间维度压缩至 $3 \times 16 \times 16$ 并扩展通道维度至 256,以提取深层语义信息。解码路径通过四级反卷积恢复空间分辨率。模型创新性地集成了双路注意力机制(dual attention block, DA 模块)和大核分组注意力门控(large-kernel group attention gate, LGAG)模块,前者通过通道和空间的双重加权优化特征,后者则通过扩大感受野实现局部和全局特征的动态融合。此外,网络在每个卷积层后引入批归一化(batch normalization, BN)层,提升训练稳定性与泛化能力。

1.2 高效多尺度特征融合模块

在图像分析任务中,同时捕捉细节特征和整体结构对模型性能至关重要。传统方法通常依赖多个独立的卷积分支提取不同尺度的特征,但计算开销大且特征协调性较差。为此,本文设计了一个轻量化的多尺度特征融合模块(multi-scale feature fusion module, MFF),如图 2 所示,通过“共享基础层+多尺寸卷积”的协同设计,在保证精度的同时降低计算复杂度。具体而言,模块首先通过一个 $3 \times 3 \times 3$ 卷积核生成共享的基础特征图。与传统的多路径卷积方式不同,使用共享嵌入层能够有效减少计算冗余,从而提升了效率。在此基础上,构建 3 条并行的卷积路径,分别采用 $1 \times 1 \times 1$ 、 $3 \times 3 \times 3$ 和 $5 \times 5 \times 5$ 的卷积核来提取不同尺度的特征。小尺度卷积核专注于捕捉局部纹理细节,中尺度卷积核聚焦于提取区域结构信息,大尺度卷积核则负责捕捉全局空间关系,形成互补的多尺度特征。为控制模型复杂度, MFF 模块仅在浅层部署,避免深层堆叠导致的参数爆炸。同时,每条路径仅执行一次卷积,并通过最大池化压缩特征维度,降低计算负担并减少过拟合。最终,通过通道维度拼接不同尺度的特征,实现深度融合,增强模型表达能力。

MFF 模块通过共享嵌入机制和多尺度融合策略,有效解决了传统多路径卷积结构的计算冗余和参数膨胀问题,显著提升了特征提取效率和模型性能,提供了一种高效的多尺度特征建模方案。

1.3 DA-Block 模块

针对髋关节影像中解剖结构复杂、组织边界模糊的挑战,本文设计了双路注意力机制(dual attention block, DA-Block),如图 3 所示。该模块通过空间-通道注意力联合优化策略,在三维特征空间中同步增强空间定位精度与通道语义区分度,显著提升了细粒度解剖结构的分割效果。DA-Block 采用双分支架构设计:空间注意力分支(position attention, PAM)捕捉跨区域的空间关联性,通道注意力分支(channel attention, CAM)则聚焦于关键通道

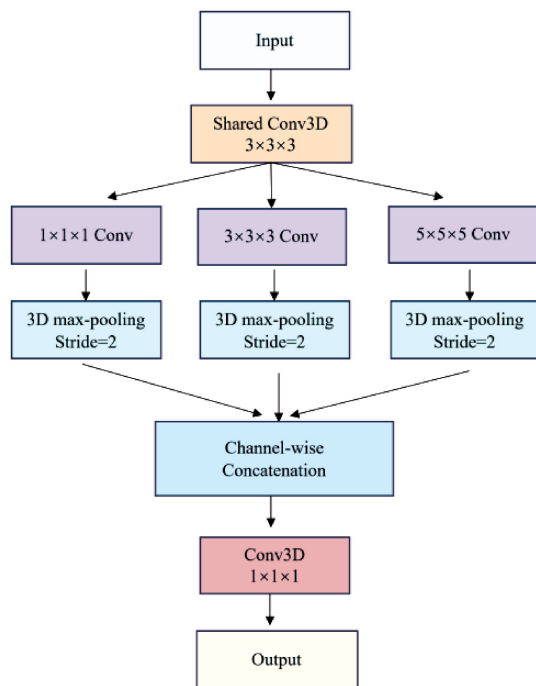


图 2 MFF 模块

Fig. 2 MFF module

的语义增强,两者的注意力信息通过加权融合机制进行动态调整,从而精准处理复杂结构,提升对细小病变(如骨质破坏、炎症灶等)的识别能力,抑制背景噪声干扰,并增强关键区域的区分度。

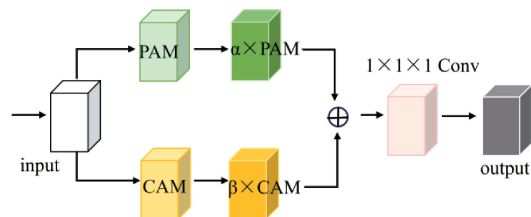


图 3 DA-Block 模块

Fig. 3 DA-Block module

其中,设计了两个可学习的标量权重参数 w_1 和 w_2 , 通过 softmax 归一化函数生成融合权重 α 和 β ,其数学表达为:

$$\alpha = \frac{e^{w_1}}{e^{w_1} + e^{w_2}} \quad (1)$$

$$\beta = \frac{e^{w_2}}{e^{w_1} + e^{w_2}} \quad (2)$$

在训练过程中, w_1 和 w_2 随网络其他参数共同优化,无需预设固定值。该机制赋予模型根据任务需求自主调节通道与空间注意力贡献度的能力,显著提升了特征融合的灵活性与泛化性能。

1) DA-PAM 模块

针对医学影像中目标区域空间分布不规则的特点,传

统卷积神经网络难以有效建模长程依赖关系。为此,本文设计了位置注意力模块(PAM),如图4所示。通过空间自适应注意力机制动态学习特征图中各位置的权重关系,显著提升关键区域的定位能力。

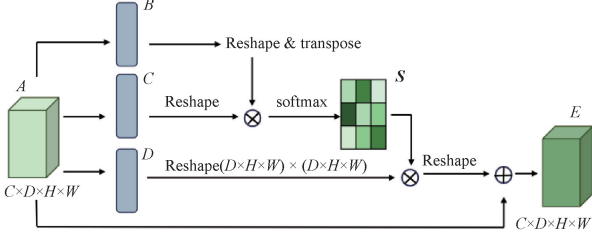


图4 PAM模块

Fig. 4 PAM module

给定输入特征图 $A \in R^{C \times D \times H \times W}$, PAM 的计算流程如下:

(1) 特征投影与重塑: 通过 $1 \times 1 \times 1$ 卷积生成 3 个特征 $B, C, D \in R^{C \times D \times H \times W}$, 并重塑为二维矩阵 $B', C', D' \in R^{C \times N}$ ($N = D \times H \times W$)。

(2) 空间注意力矩阵生成: 对 B' 进行转置并与 C' 进行矩阵乘法, 得到空间注意力矩阵 $S \in R^{N \times N}$, 通过 Softmax 函数归一化:

$$S_{ji} = \frac{\exp(B_i'^T \cdot C_j')}{\sum_{i=1}^N \exp(B_i'^T \cdot C_j')} \quad (3)$$

其中, S_{ji} 表示位置 j 对位置 i 的注意力权重, 反映了位置 j 对位置 i 的贡献程度。

(3) 加权特征图生成: 将空间注意力矩阵 S 与矩阵 D' 进行矩阵乘法后, 得到加权特征图。

$$E' = \alpha (D' S^T) \quad (4)$$

其中, α 为可学习的缩放参数, 用于控制注意力特征的贡献度。

(4) 输出特征图: 将输入特征图 A 与加权后的特征图逐元素相加, 得到最终输出特征图 E 。

$$E_j = \alpha \sum_{i=1}^N (S_{ji} \cdot D'_i) + A_j \quad (5)$$

DA-PAM 通过加权聚合空间上下文信息, 解决了传统卷积操作局部感受野的限制, 显著提升了模型对复杂空间特征的建模能力。该机制能够自适应地聚焦关键区域, 抑制背景噪声, 为后续分割任务提供鲁棒且判别性强的特征表示。

2) DA-CAM

医学图像中的关键信息常分布在特征图的不同通道, 因此本文设计了通道注意力模块(CAM), 如图5所示。该模块通过加权机制调整每个通道的贡献, 使模型关注临床重要特征, 从而增强关键特征的表达。

给定输入特征图 $A \in R^{C \times D \times H \times W}$, CAM 的计算流程如下:

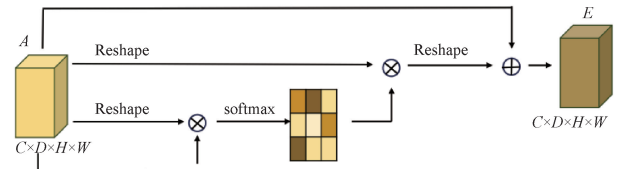


图5 CAM模块

Fig. 5 CAM module

(1) 特征图重塑: 将输入特征图 A 重塑为二维矩阵 $A' \in R^{C \times N}$ ($N = D \times H \times W$)。

(2) 通道注意力矩阵生成: 通过矩阵乘法计算通道注意力矩阵 $X \in R^{C \times C}$:

$$x_{ji} = \frac{\exp(A'_i \cdot A'_j)}{\sum_{i=1}^C \exp(A'_i \cdot A'_j)} \quad (6)$$

(3) 通道特征重构: 将通道注意力矩阵 X 与输入特征矩阵 A' 相乘, 实现通道特征的加权重构:

$$F = \beta (X \cdot A') \quad (7)$$

其中, β 为可学习的缩放因子, 初始值为 0, 通过训练过程自适应调整。

(4) 特征图恢复: 将重构后的特征矩阵 F 进行逆维度变换, 恢复为与输入特征图 A 相同的维度 $C \times D \times H \times W$, 确保空间结构的一致性。

(5) 输出特征图生成: 通过跳跃连接将重构后的特征图与输入特征图逐元素相加, 得到通道注意力增强后的输出特征图 E :

$$E_j = \beta \sum_{i=1}^C x_{ji} \cdot A'_i + A_j \quad (8)$$

与 DA-PAM 模块不同, DA-CAM 直接对原始通道特征进行加权融合, 无需额外的卷积层。这一设计保留了通道间的相关性, 显著降低了计算复杂度。

1.4 LGAG 模块

在医学图像分割任务中, 多尺度特征的融合对于提升模型性能至关重要。为此, 本文提出了大核分组注意力门控(large-kernel group attention gate, LGAG)机制, 如图6所示, 以优化特征图融合过程, 增强关键特征的表达并抑制无关信息。LGAG 通过学习动态注意力系数, 在网络不同阶段精细调控信息流, 提高分割精度。

与传统的 Attention UNet 采用 $1 \times 1 \times 1$ 卷积处理特征不同, LGAG 模块在 $q_{att}(\cdot)$ 函数中创新性地引入 $3 \times 3 \times 3$ 分组卷积(分别记为 $GC_g(\cdot)$ 和 $GC_x(\cdot)$), 用于分别处理门控信号 g 和输入特征图 x 。首先, 经过 $3 \times 3 \times 3$ 分组卷积后, 两者分别进行批归一化($BN(\cdot)$), 并通过元素级加法融合。随后, 融合特征经过 ReLU 激活函数($R(\cdot)$)进行非线性变换, 再通过 $1 \times 1 \times 1$ 卷积($C(\cdot)$)进一步提炼, 并经过批归一化生成单通道特征图。最终, 该特征图经过 Sigmoid 激活函数($\sigma(\cdot)$)计算出注意力系数, 并与输入特征图 x 进行逐元素乘法, 以生成经过注意力调控的特征图

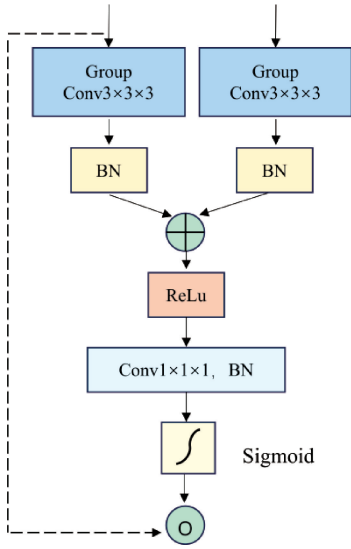


图 6 LGAG 模块
Fig. 6 LGAG module

些切片涵盖了不同病理阶段和解剖变异,为模型的训练和验证提供了大规模、多样化的临床样本。所有 CT 图像均采用标准化参数采集,单层图像分辨率为 $512 \text{ pixel} \times 512 \text{ pixel}$,体素间距范围为 $0.449 \text{ mm} \times 0.449 \text{ mm} \sim 0.801 \text{ mm} \times 0.801 \text{ mm}$,每例病例包含 150~200 层连续断层图像。为确保数据标注的准确性,所有 CT 图像均由山西白求恩医院经验丰富的影像科医生进行精细的手动分割,并通过多位影像科医生的验证,有效降低了个人主观因素对标注结果的干扰。在实验数据划分中,将 70 例骶髂关节 CT 数据按照 7:3 的比例随机划分为训练集(49 例)和验证集(21 例)。图 7 呈现了 CT 影像的横断面、冠状面和矢状面视图。

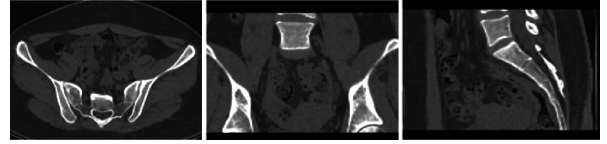


图 7 三方位 CT 图

Fig. 7 Three-dimensional CT image

$LGAG(g, x)$ 。LGAG 模块的整体操作可以用以下公式描述:

$$q_{att}(g, x) = R(BN(GC_g(g) + BN(GC_x(x)))) \quad (9)$$

$$LGAG_{(g, x)} = x \otimes \sigma(BN(C(q_{att}(g, x)))) \quad (10)$$

其中, $GC_g(g)$ 和 $GC_x(x)$ 代表对 g 和 x 进行 $3 \times 3 \times 3$ 分组卷积, $BN(\cdot)$ 表示批归一化, $R(\cdot)$ 表示 ReLU 激活函数, $C(\cdot)$ 为 $1 \times 1 \times 1$ 卷积操作, 而 $\sigma(\cdot)$ 为 Sigmoid 激活函数, 用于生成注意力系数。

相较于传统方法, LGAG 采用大核分组卷积, 在减少计算开销的同时, 能够捕获更丰富的空间上下文信息, 增强对关键结构的感知能力, 从而有效提升医学图像分割的准确性。

2 实验内容及结果分析

2.1 实验设置

1) 实验数据

本文共收集并筛选了 70 例来自山西白求恩医院的骶髂关节 CT 影像数据, 总计约 10 414 张高分辨率切片。这

2) 实验流程

实验流程如图 8 所示。首先, 针对 CT 序列图像数据量较大的特点(单例数据尺寸为 $N \times 512 \times 512$, $N \in [50, 200]$), 对原始数据进行多步预处理以提高训练效率并降低内存消耗: 将输入图像尺寸统一调整为 $N \times 256 \times 256$, 并通过方向规划确保图像空间方位的一致性; 对 CT 值进行线性归一化处理, 将其映射至 $[-200, 1\ 200]$ 区间以增强骶髂关节区域特征并抑制背景干扰; 此外, 通过筛选有效切片范围, 去除序列首尾不包含目标区域的切片, 确保数据集集中仅保留有效解剖结构。预处理后的数据通过基于 PyTorch 框架的仿射变换进行数据增强, 包括随机翻转、旋转、缩放和平移等操作, 以提升模型的泛化能力。增强后的数据输入改进后的分割网络进行训练, 通过超参数优化获得二值化的骶髂关节分割掩膜。最终, 将分割掩膜与原始 CT 数据进行矩阵乘法运算, 得到分割结果, 并使用 3D Slicer(v5. 6. 2) 进行三维重建与可视化, 实现分割结果的多角度观察与定量评估。

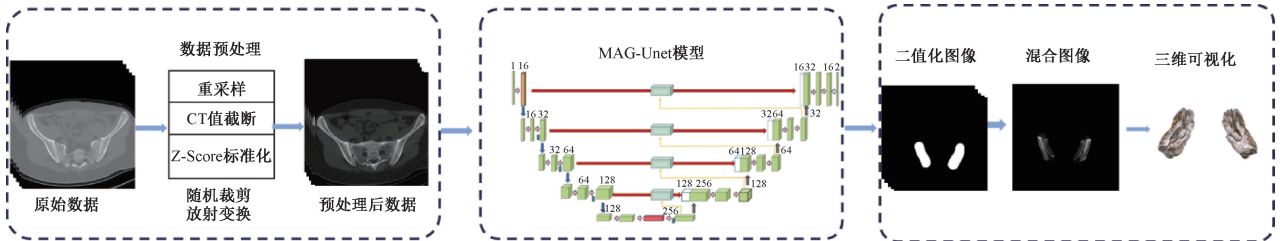


图 8 实验流程图

Fig. 8 Experimental flow chart

3) 实验环境

本文实验中模型训练所使用的硬件配置详细为: CPU

为 Intel Xeon Platinum 多核处理器(10 vCPU, 2.6 GHz), 配合 60 GB DDR4 内存, GPU 为与 NVIDIA GeForce RTX

4090 显卡(24 GB)。实验采用 Pytorch2.2.0 深度学习框架完成,结合 CUDA 11.8 和 cuDNN 8.9.4 加速库,采用 Adam 优化器,权重衰减为 1×10^{-4} 。batchsize 设置为 1。网络训练了 300 个 Epoch。初始学习率为 1×10^{-4} 。在训练过程中,采用的数据增强策略如下:以 0.5 的概率进行随机翻转、随机仿射变换、随机伽马校正等。

4) 评价指标

为了全面评估模型的分割性能,本文采用了 4 个常用指标: Dice 系数(Dice similarity coefficient, DSC)、交并比(intersection over union, IoU)、精确率(Precision)和召回率(Recall)。这些指标定义如下:

(1) Dice 系数衡量预测分割结果与真实标签的相似度,值越接近 1 表示分割效果越好,公式为:

$$Dice = \frac{2 \times |X \cap Y|}{|X| + |Y|} \quad (11)$$

其中, X 为预测结果, Y 为真实标签, $|X \cap Y|$ 为交集大小, $|X|$ 和 $|Y|$ 分别为预测结果和真实标签的像素数量。

(2) IOU 表示预测结果与真实标签的交集与并集的比值,值越接近 1 表示分割效果越好,公式为:

$$IOU = \frac{|X \cap Y|}{|X \cup Y|} \quad (12)$$

(3) 精确率衡量模型预测为正类的样本中,真实为正类的比例,公式为:

$$Precision = \frac{TP}{TP + FP} \quad (13)$$

其中, TP 表示正确预测为正类的像素数量, FP 表示错误预测为正类的像素数量。

(4) 召回率衡量真实为正类的样本中被正确预测为正类的比例,反映了模型对正类样本的覆盖能力,公式为:

$$Recall = \frac{TP}{TP + FN} \quad (14)$$

其中, FN 表示错误预测为负类的像素数量。

通过上述指标的综合评估,可以全面分析模型的分割性能。Dice 系数和 IoU 侧重于整体相似度,而 Precision 和 Recall 分别从预测准确性和覆盖能力两个角度评估模型的分割效果。

5) 损失函数

本文采用了医学图像分割中最常用的损失函数 Dice Loss。Dice Loss 在训练过程中更加关注对前景信息的挖掘,即更加关注骶髂关节区域部分,其定义如式(15)所示。但是 Dice Loss 在小目标的情况下容易出现不稳定的情况,极端情况下会导致梯度饱和现象。为了缓解这种问题,本研究还采用了交叉熵损失函数,其公式如式(16)所示。本文将二者结合起来,定义了骶髂关节分割任务的损失函数,如式(17)所示,从而指导模型优化与训练。

$$L_{Dice} = 1 - \frac{2 \times \sum_i (p_i \cdot g_i)}{\sum_i p_i + \sum_i g_i} \quad (15)$$

$$L_{CE} = - \sum_i [g_i \log(p_i) + (1 - g_i) \cdot \log(1 - p_i)] \quad (16)$$

其中, p_i 和 g_i 分别表示像素 i 的预测值和标签值, N 为像素点总个数。

$$L = L_{Dice} + L_{CE} \quad (17)$$

2.2 对比实验

为了验证 MAG-UNet 在骶髂关节分割任务中的性能,本文将其与 3D-Unet、3D V-Net^[23]、ResUNet^[24]、AttentionUnet、SegNet^[25]、UNTER^[26] 和 Deeplabv3 十七种主流分割模型进行了对比。实验结果如表 1 所示。MAG-UNet 在 Dice 系数(92.5%)、IoU(86.0%)和 Precision(93.3%)等关键指标上均取得了最优表现,明显优于其他模型。尽管在 Recall 指标上与 Attention UNet 接近, MAG-UNet 仍保持了更高的分割完整性。此外, MAG-UNet 采用轻量化设计,参数量仅为 23.98 M,远低于 TransUNet 和 ResUNet,在保证高精度的同时大幅降低了计算成本。特别地,在相近参数规模下, MAG-UNet 的 Dice 系数相比 3D U-Net 提升了 3.4%,展示了其在精度和效率上的优势。

表 1 不同模型的对比结果

模型	Dice	IoU	Pre	Recall	Params/M
3D U-Net	90.4	82.6	89.2	92.6	22.45
3D V-Net	89.5	81.0	88.5	91.0	54.14
ResUNet	91.0	83.6	91.1	91.3	55.92
Attention UNet	91.8	84.1	92.0	91.5	118.51
SegNet	88.6	79.5	87.5	89.0	76.80
UNTER	87.3	77.9	84.6	88.4	106.27
DeepLabv3+	90.7	82.9	91.0	90.6	52.85
Ours(MAG-UNet)	92.5	86.0	93.3	91.3	23.98

2.3 注意力对比试验

为验证本文提出的 DA-Block(融合空间-通道双路径注意力)机制的有效性,本研究设计了 6 组对比实验,实验结果如表 2 所示。评估不同注意力机制对三维图像分割性能的影响。实验结果基于常见的评估指标,包括 Dice 系数、IoU(交并比)和精确度(Precision)。实验的对比模型包括标准的 3D Unet、带有不同注意力模块的变种模型,以及本研究提出的 DA-Block 模块。

从表格中可以看出,应用不同注意力机制后,模型的性能普遍有所提升。特别是在引入 DA-Block 后,模型的 Dice 系数、IoU 和 Precision 分别达到了 91.5、84.4 和 90.9,相较于标准的 3D Unet(Dice 为 90.4, IoU 为 82.6, Precision 为 89.2),有了明显的提升。通过对比其他注意力模块,如 SE、CBAM 和 ECA, DA-Block 在 3 个指标上均展现了较大的优势,证明了其在融合空间和通道信息方面

表 2 不同注意力对比试验

Table 2 Comparison test of different attention

模型	Dice	IoU	Precision
3D Unet	90.4	82.6	89.2
+SE	90.7	83.0	89.5
+CBAM	91.0	83.2	89.8
+Non_Local	90.6	82.8	89.4
+ECA	88.4	79.6	87.6
+DA_Block	91.5	84.4	90.9

表 3 不同卷积核大小对比试验

Table 3 Comparison test of different convolution kernel sizes

Conv. kernels	[1]	[3]	[5]	[1,3]	[3,3]	[1,3,5]	[3,3,3]	[3,5,7]
IoU	79.6	82.6	80.2	82.5	81.6	83.2	80.8	81.0
Params	22.42	22.45	22.54	22.48	22.50	23.14	22.58	22.62

2.5 消融实验

为了验证 MAG-UNet 中各模块的贡献,本文进行了消融实验,逐步引入 DA-Block、MFF 和 LGAG 模块,并与基线模型(3D U-Net)进行对比。实验结果如表 4 所示,单独引入 DA-Block 模块后, IoU 从 82.6% 提升至 84.4%, Dice 系数从 90.4% 提升至 91.5%, 验证了空间-通道注意力机制的有效性。引入 MFF 模块后, IoU 提升至 83.2%, Dice 系数为

的有效性。
2.4 多尺度卷积核对比试验
为了探究多尺度卷积核对骶髂关节分割任务的影响,本研究测试了不同卷积核组合的性能,包括 $[1\times 1]$ 、 $[3\times 3]$ 、 $[5\times 5]$ 、 $[1\times 1, 3\times 3]$ 、 $[3\times 3, 3\times 3]$ 、 $[1\times 1, 3\times 3, 5\times 5]$ 和 $[3\times 3, 5\times 5, 7\times 7]$ 。实验结果如表 3 所示,多尺度组合的表现优于单一卷积核,且 $[1\times 1, 3\times 3, 5\times 5]$ 组合在 IoU (83.2%)和参数量(23.14 M)方面表现最佳。这表明多尺度卷积核能有效提高模型对不同尺度特征的感知能力,优化分割精度,并在性能和计算开销之间达到良好平衡。

90.8%, 证明了多尺度特征提取的作用。引入 LGAG 模块后, IoU 为 83.0%, Dice 系数为 90.6%, 表明局部-全局信息融合的优化效果。3 个模块联合使用时, IoU 达到 86.0%, Dice 系数为 92.4%, 参数量控制在 23.98 M, 充分展示了精度与计算效率的平衡。实验表明, DA-Block、MFF 和 LGAG 模块在 MAG-UNet 中协同作用, 显著提升了模型的分割性能。图 9 展示了消融实验可视化对比结果。

表 4 消融实验结果

Table 4 Ablation experiment results

基线	Da_Block	MFF	LGAG	IoU	Dice	Precision	Params/MB
✓				82.6	90.4	89.2	22.45
✓	✓			84.4	91.5	90.9	23.01
✓		✓		83.2	90.8	89.5	23.14
✓			✓	83.0	90.6	89.4	23.09
✓	✓	✓		84.8	91.7	91.2	23.45
✓	✓		✓	84.6	91.6	91.0	23.30
✓	✓	✓	✓	86.0	92.4	93.3	23.98

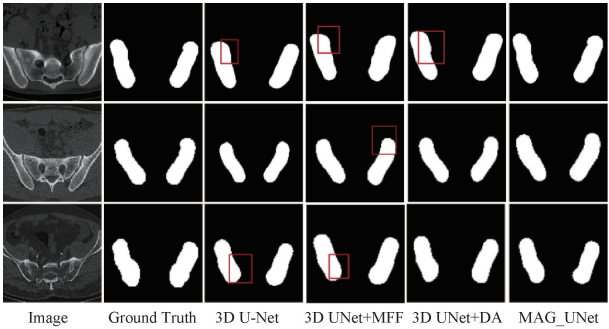


图 9 骶髂关节 CT 影像分割的消融实验结果对比

Fig. 9 Comparison of ablation experimental results of sacroiliac joint CT image segmentation

2.6 可视化结果分析

为了更直观地展示不同模型在骶髂关节分割中的差异,本研究对对比实验中选取的各个模型的分割结果进行了可视化分析。从测试集中随机选取了 4 例数据(a, b, c, d),具体可视化结果如图 10 所示。展示了 CT 图像、专家标注的 GT 和各模型的分割结果。结果显示, MAG-UNet 在分割精度上显著优于其他模型,尤其在复杂区域的分割中表现更强的鲁棒性。例如, MAG-UNet 减少了错分和漏分情况,能够更精准地捕获目标区域,尤其在数据 b、c、d 中明显优于其他模型。相比之下, UNETR 和 SegNet 等模型在目标边界刻画上仍存在较多错分和漏分,难以准确分割复杂结构。可视化分析验证了 MAG-UNet 在细节保留和

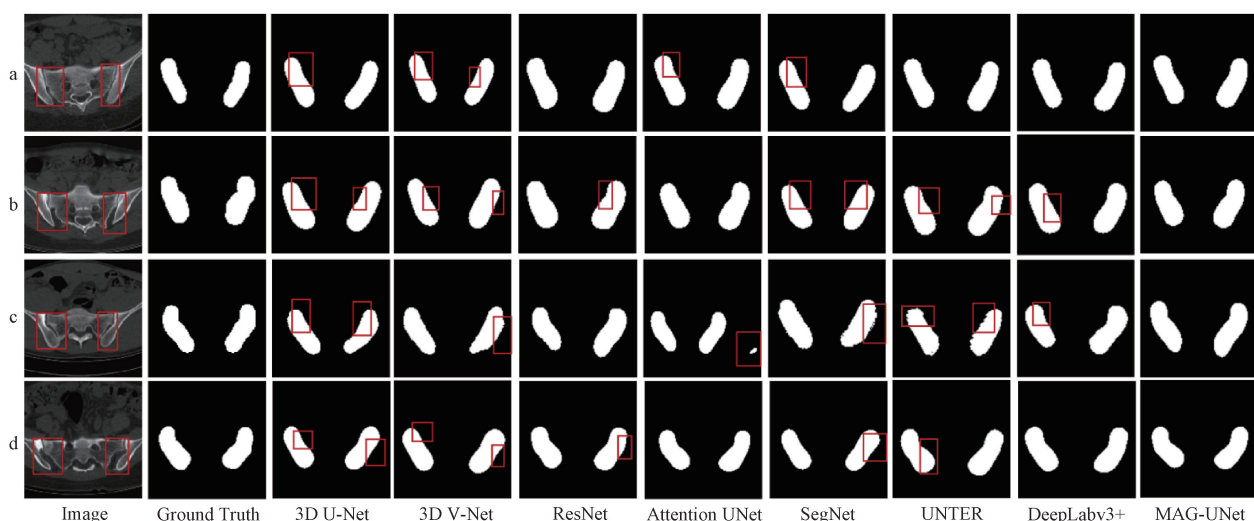


图10 分割结果可视化

Fig. 10 Segmentation result visualization

错误抑制方面的优势,展现了其在实际临床应用中的潜力。

3 结 论

本文与山西白求恩医院风湿免疫科合作,针对骶髂关节CT影像分割中自动化程度低、病灶边界模糊、早期微小病灶检出困难等临床挑战,提出了一种改进的MAG-UNet模型。该模型集成了多尺度特征融合(MFF)模块、全局注意力机制(DA)和局部引导自适应增强(LGAG)模块,通过多层次特征提取与增强策略,有效解决了复杂解剖结构和微小病灶的精准识别问题。实验结果表明,MAG-UNet在Dice系数和IoU等评价指标上优于主流分割方法,验证了其技术优势。本研究不仅为风湿免疫科和骨科的智能化辅助诊断提供了一个可靠工具,而且其创新的架构设计也为脊柱、关节等复杂解剖结构的医学影像分割任务提供了新的思路和改进方向。特别是在强直性脊柱炎早期诊断中,MAG-UNet展现出明显的临床应用价值,可为影像学分析提供更准确的解剖结构分割。未来的研究将重点围绕以下方向展开:引入自监督学习减少对大规模标注数据的依赖;优化推理效率以满足临床实时应用需求;结合多模态影像数据(如MRI)提升病变识别能力。这些改进将进一步提升模型的泛化性和临床实用性,推动智能辅助诊断系统在风湿免疫疾病诊疗中的深入应用。

参考文献

- [1] GARTENBERG A, NESSIM A, CHO W. Sacroiliac joint dysfunction: Pathophysiology, diagnosis, and treatment[J]. *European Spine Journal*, 2021, 30: 2936-2943.
- [2] 云雪, 郭华. 强直性脊柱炎诊断和治疗研究现状[J]. *临床个性化医学*, 2024, 3(4): 2643-2652.

YUN X, GUO H. Current status of diagnosis and treatment of ankylosing spondylitis [J]. *Journal of Clinical Personalized Medicine*, 2024, 3 (4): 2643-2652.

- [3] 严武军, 王家辉, 邱瑜茹. PAF-Net: 用于骶髂关节高效分割的并行注意力网络[J]. *计算机系统应用*, 2025, 34(1): 137-144.
- YAN W J, WANG J H, QIU Y R. PAF-Net: Parallel attention network for efficient sacroiliac joint segmentation[J]. *Computer Systems & Applications*, 2025, 34(1): 137-144.
- [4] 李府天. 基于骶髂关节CT影像的强直性脊柱炎自动分级研究[D]. 汕头: 汕头大学, 2022.
- LI F T. Study on automatic grading of ankylosing spondylitis based on sacroiliac joint CT images[D]. Shantou: Shantou University, 2022.
- [5] 董晚亭, 潘小文, 罗容智, 等. 基于MRI的伪CT图像对骶髂关节炎结构性病变的诊断效能[J]. *中国CT和MRI杂志*, 2022, 20(5): 175-178.
- DONG W T, PAN X W, LUO R ZH, et al. Diagnostic efficacy of MRI-based pseudo-CT images for structural lesions of sacroiliitis[J]. *Chinese Journal of CT and MRI*, 2022, 20(5): 175-178.
- [6] 刘倩倩, 马凯蒂. 强直性脊柱炎病人慢性疼痛潜在类别分析及其影响因素[J]. *护理研究*, 2025, 39(6): 925-930.
- LIU Q Q, MA K D. Latent class analysis and influencing factors of chronic pain in patients with ankylosing spondylitis[J]. *Nursing Research*, 2025, 39(6): 925-930.
- [7] ZHANG K, LUO G B, LI W J, et al. Automatic

- image segmentation and grading diagnosis of sacroiliitis associated with AS using a deep convolutional neural network on CT images [J]. Journal of Digital Imaging, 2023, 36(5): 2025-2034.
- [8] 谭晓倩, 宋玲玲, 李语, 等. MR 类似 CT 限制回波间隔的快速场回波成像技术在强直性脊柱炎骶髂关节骨质结构改变中的诊断价值[J]. 临床放射学杂志, 2025, 44(3): 536-540.
- TAN X Q, SONG L L, LI Y, et al. Diagnostic value of MR-like CT-limited echo interval rapid field echo imaging in the bone structure changes of sacroiliac joints in ankylosing spondylitis[J]. Journal of Clinical Radiology, 2025, 44(3): 536-540.
- [9] LONG J, SHELHAMER E, DARRELL T. Fully convolutional networks for semantic segmentation[J]. IEEE Transactions on Pattern analysis and Machine Intelligence, 2016, 39(4): 640-651.
- [10] RONNEBERGER O, FISCHER P, BROX T. U-net: Convolutional networks for biomedical image segmentation [C]. Medical Image Computing and Computer-Assisted Intervention-MICCAI 2015: 18th International Conference, Munich, Germany, October 5-9, 2015, proceedings, part III 18. Springer international publishing, 2015: 234-241.
- [11] ÇIÇEK Ö, ABDULKADIR A, LIENKAMP S S, et al. 3D U-Net: Learning dense volumetric segmentation from sparse annotation [C]. Medical Image Computing and Computer-Assisted Intervention-MICCAI 2016: 19th International Conference, Athens, Greece, October 17-21, 2016, Proceedings, Part II 19. Springer International Publishing, 2016: 424-432.
- [12] ZHOU Z, RAHMAN SIDDIQUEE M M, TAJBAKHS N, et al. Unet++: A nested u-net architecture for medical image segmentation[C]. Deep Learning in Medical Image Analysis and Multimodal Learning for Clinical Decision Support: 4th International Workshop, DLMIA 2018, and 8th International Workshop, ML-CDS 2018, Held in Conjunction with MICCAI 2018, Granada, Spain, September 20, 2018, Proceedings 4. Springer International Publishing, 2018: 3-11.
- [13] OKTAY O, SCHLEMPER J, FOLGOC L L, et al. Attention u-net: Learning where to look for the pancreas [J]. ArXiv preprint arXiv: 1804.03999, 2018.
- [14] 武丽, 丁琴, 葛彩成. 融合公平因子的半监督学习医学图像分割模型[J]. 电子测量技术, 2024, 47(23): 171-180.
- WU L, DING Q, GE C H. Semi-supervised learning medical image segmentation model integrating fairness factor [J]. Electronic Measurement Technology, 2024, 47(23): 171-180.
- [15] 许超, 王云健, 刘洋, 等. 基于改进 Swin Transformer 的膝骨关节炎 X 光影像自动诊断[J]. 电子测量技术, 2024, 47(19): 155-163.
- XU CH, WANG Y J, LIU Y, et al. Automatic diagnosis of knee osteoarthritis X-ray images based on improved Swin Transformer [J]. Electronic Measurement Technology, 2024, 47(19): 155-163.
- [16] 王霞霞, 张炫, 吴兴隆, 等. 双阶段噪声自适应超声图像分割网络[J]. 电子测量技术, 2024, 47(10): 175-183.
- WANG X X, ZHANG X, WU X L, et al. Two-stage noise adaptive ultrasound image segmentation network[J]. Electronic Measurement Technology, 2024, 47(10): 175-183.
- [17] HE K M, GKIOXARI G, DOLLÁR P, et al. Mask R-CNN [C]. IEEE International Conference on Computer Vision, 2017: 2961-2969.
- [18] CHEN L C, ZHU Y, PAPANDREOU G, et al. Encoder-decoder with atrous separable convolution for semantic image segmentation [C]. European Conference on Computer Vision (ECCV), 2018: 801-818.
- [19] LIU L, ZHANG H Y, ZHANG W F, et al. Grading diagnosis of sacroiliitis in ct scans based on radiomics and deep learning [C]. Chinese Conference on Biometric Recognition. Cham: Springer Nature Switzerland, 2022: 655-663.
- [20] ZHANG K, LIU CH R, PAN J L, et al. Use of MRI-based deep learning radiomics to diagnose sacroiliitis related to axial spondyloarthritis [J]. European Journal of Radiology, 2024, 172: 111347.
- [21] CHEN W, LIU B Q, PENG S T, et al. S3D-UNet: Separable 3D U-Net for brain tumor segmentation[C]. Brainlesion: Glioma, Multiple Sclerosis, Stroke and Traumatic Brain Injuries: 4th International Workshop, BrainLes 2018, Held in Conjunction with MICCAI 2018, Granada, Spain, September 16, 2018, Revised Selected Papers, Part II 4. Springer International Publishing, 2019: 358-368.
- [22] YANG Q, WANG CH F, PAN K CH, et al. An improved 3D-UNet-based brain hippocampus segmentation model based on MR images[J]. BMC Medical Imaging, 2024, 24(1): 166.
- [23] MILLETARI F, NAVAB N, AHMADI S A. V-net:

Fully convolutional neural networks for volumetric medical image segmentation [C]. 2016 Fourth International Conference on 3D Vision(3DV). IEEE, 2016: 565-571.

[24] DIAKOIANNIS F I, WALDNER F, CACCETTA P, et al. ResUNet-a: A deep learning framework for semantic segmentation of remotely sensed data [J]. ISPRS Journal of Photogrammetry and Remote Sensing, 2020, 162: 94-114.

[25] 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.

[26] HATAMIZADEH A, TANG Y, NATH V, et al. Unetr: Transformers for 3D medical image segmentation [C].

IEEE/CVF Winter Conference on Applications of Computer Vision, 2022: 574-584.

作者简介

严武军(通信作者), 硕士, 副教授, 硕士生导师, 主要研究方向为人工智能、深度学习、机器学习。

E-mail: yanwujun@tynu.edu.cn

景莹, 硕士研究生, 主要研究方向为深度学习、医学影像处理。

E-mail: jingying0423@163.com

徐莹臣, 硕士研究生, 主要研究方向为区块链、深度学习。

E-mail: 17615351517@163.com

张晓丽, 硕士研究生, 主要研究方向为深度学习, 计算机视觉。

E-mail: 2968196306@qq.com

王程, 硕士研究生, 主要研究方向为深度学习, 图像处理。

E-mail: chen9cool@163.com